O'REILLY®

"Written by some of the top minds in SEO!"

—Danny Sullivan, Editor-In-Chief, SearchEngineLand.com
& Producer, SMX: Search Marketing Expo

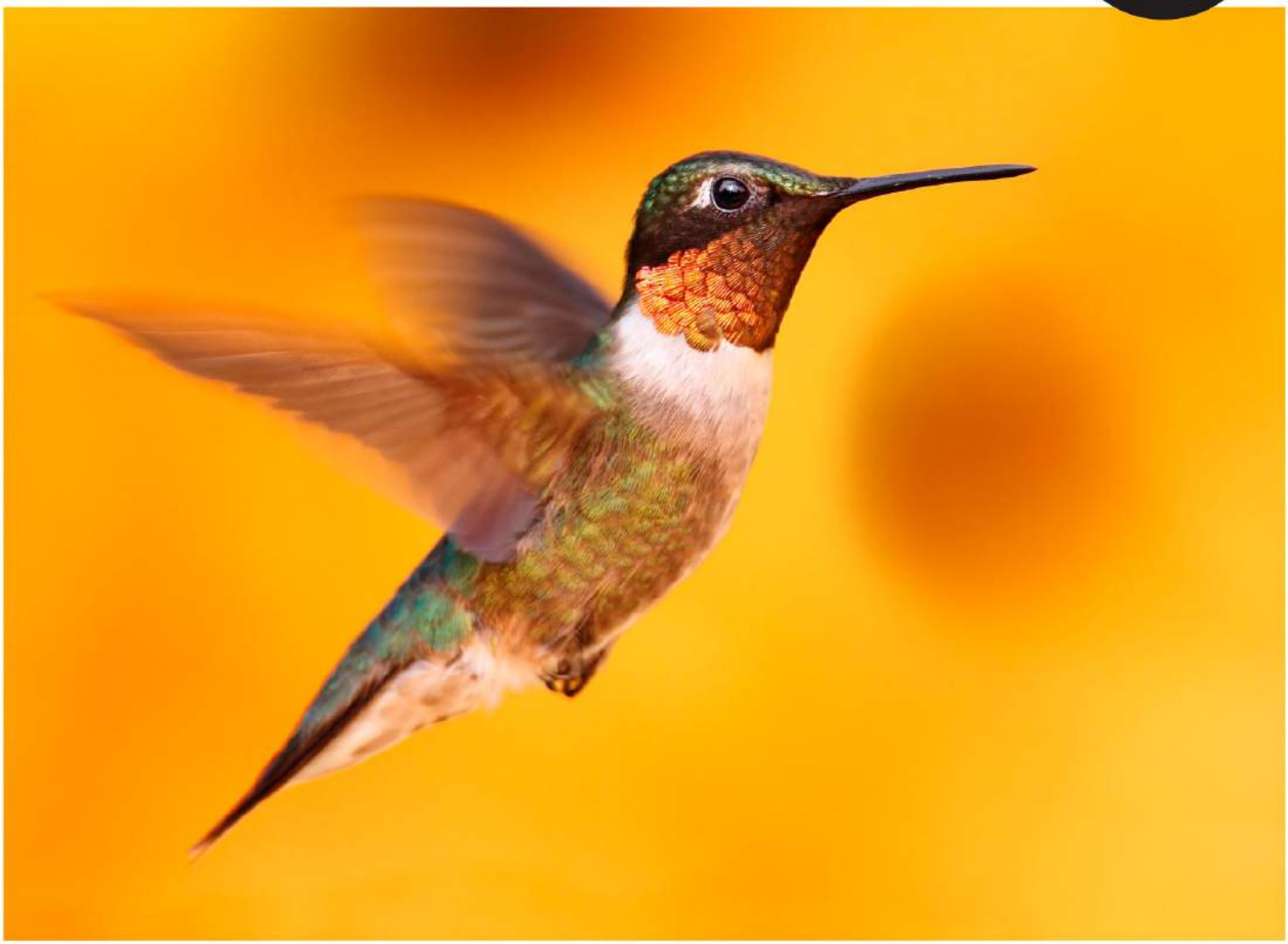
/ theory / in / practice

The Art of SEO

Mastering Search Engine Optimization

Eric Enge, Stephan Spencer,
and Jessie C. Stricchiola

3rd Edition



Dedication
Praise for The Art of SEO

How serious are you about maximizing and multiplying, monetizing your ability to harness the dynamic force and power of SEO? I say dynamic, because it is not something static. Like game theory, your strategy needs to be preeminent, preemptive, and ever-evolving. The Art of SEO is deep, serious (yet eminently, elegantly, and stunningly clear and profoundly actionable)! The authors examine, explain, explore, and expose the real truths, the real inner workings, the real ethical means of optimizing SEO in a fast-changing, ultra-competitive online environment. You will finally and meaningfully grasp how to gain the most prized outcome imaginable: the sustainable "gift" of your target market's fullest attention, presence, and trust! Every chapter is sincerely a complete short-course primer, masterfully distilled down to its most actionable, relevant, critical elements. If you've struggled trying to figure out who you can really trust to understand and meaningfully manage your SEO opportunties, read this book—then use it as a reality check against anyone you are now entrusting your most precious online relationship to.

Jay Abraham, best-selling author and Forbes Magazine "Top 5" Best Executive Coaches in the US

An essential guide to best practices and cutting-edge tactics that belongs on the desk of all search marketing professionals, especially in these days of nearly constant change, updates, and new approaches by the search engines.

Chris Sherman, Founding Editor, Search Engine Land, and VP Programming, Search Marketing Expo

The Art of War isn't about Chinese pottery, and The Art of SEO isn't a paint-by-numbers kit. This 900-page book is a comprehensive guide to search engine optimization strategies and tactics written by three SEO experts: Eric Enge, Stephan Spencer, and Jessie Stricchiola. The chapters in the third edition on Content Marketing as well as Panda, Penguin, and Penalties are must-reads for anyone interested in mastering search engine optimization.

Greg Jarboe, President of SEO-PR, and author of YouTube and Video Marketing

In a world where consumers spend more time on the Web than all other media channels combined, getting found is the a critical path to business success. The Art of SEO unlocks the core of what it means to get found and how to accelerate your growth through creating value for your consumers. This is a must have for any progressive marketing team.

Seth Besmertnik, CEO and cofounder, Conductor

Staying on top of your SEO is crucial to your marketing game—and for this, The Art of SEO is the best resource out there, by experts who know about the art and science of SEO.

Taki Moore, Coach Marketing Expert

The Art of SEO continues to be one of the best educational resources for having an SEO strategy that dominates. With SEO being a fast-paced, ever-evolving field, the third edition is a must to stay on top of the latest updates and changes.

Lewis Howes, two sport all-American, Olympic hopeful, entrepreneur, and host of The School of Greatness podcast

If I'm going to devote time to learn something, I want to learn from the best. The authors of The Art of SEO are simply the best at what they do... and they share their expertise in such a way that it's easy to read and understand. Whether you're a CEO or just climbing up the ladder, you should have a solid foundation of SEO and this book is where to start.

Steve Spangler, founder of Steve Spangler Science

The Art of SEO doesn't just arm you with facts and knowledge, it provides you with the strategies to dominate in the search engines. It continues to be the most comprehensive book out there for everyone, from absolute beginners to advanced SEOs looking to hone their skills.

Neil Patel, cofounder of Crazy Egg

This is the book to read on SEO, packed full of knowledge from beginner to expert to master.

Ramez Naam, former Relevance Group Program Manager, Bing, and author of Nexus

The Art of SEO represents a comprehensive and instructive guide to mastering search engine optimization for webmasters new and experienced. While the SEO industry continues to evolve, the primary teachings of this seminal book hold constant—SEO continues to be an art that can be learned and finessed when armed with the tools and tips recorded in these writings.

Kristopher B. Jones, founder of LSEO.com and best-selling author of Search Engine Optimization: Your Visual Blueprint to Effective Internet Marketing

Once again, The Art of SEO earns its place as the only SEO book on my bookshelf. It is the first resource I recommend to both new and experienced bloggers with good reason; it helped me to lay the foundation that resulted in more than 100 million organic hits to my site and has repeatedly achieved stellar results for those I've mentored. In its third edition, The Art of SEO expands to cover the latest on search engine logic, social implementation, and staying ahead of the ever-changing algorithms. A brilliant, concise, essential guide to search engine optimization that every online marketer should read and put into practice.

Jeni Larson Hott, 7 Figure blogger

If you ever want to sell anything, you can't overestimate the importance of understanding search behaviour. And there's no stronger team of A-players to write about search engine optimization. Everyone should read this book!

Chris Goward, founder and CEO of WiderFunnel, and author of You Should Test That!

After diving into the third edition of The Art of SEO, I felt like a gladiator armed with all the right knowledge and tools to take on the ever-changing world of SEO.

Ken Rutkowski, host of Business Rockstars radio show, founder of METal

The Art of SEO combines the expertise of the three leading search experts in the world, making it an invaluable resource.

Gokul Rajaram, Product Engineering Lead at Square, former Product Director, Ads at Facebook, former Product Director, AdSense at Google

The Art of SEO is an innovative book that can change your fortune and the future of your company forever, starting with the very first page. The book is full of valuable information that will save you countless hours—and perhaps make you millions of dollars—when promoting your business online. The concepts and ideas presented are easy to understand and follow, which is key for brands or companies that are busy focusing on their product or service, but need to keep well informed. The three authors behind this book, Stephan Spencer, Eric Enge and Jessie Stricchiola, bring together collectively decades of experience, and share some of their most innovative methods, research, and strategies to save you valuable time and money in accomplishing measurable results in your SEO. Being in its third edition, the authors of The Art of SEO are constantly following the latest changes, and providing the most up-to-date, comprehensive, tried-and-tested techniques to keep you ahead of the curve. As I've said in many of my talks, if you're not upgrading your skills, you're falling backwards. The Art of SEO gives you the latest information to stay competitive in the field, with all the knowledge readily available at your fingertips.

Brian Tracy, President of Brian Tracy International

When people ask me what SEO book they should read, The Art of SEO is the first and only one I mention. Plus, it's the only SEO book that's both technically sound and constantly updated. Highly recommended.

Cyrus Shepard, Director of Audience Development, Moz

Hands down the most valuable book on understanding SEO!

Jordan Koene, VP Professional Services, Searchmetrics

It doesn't matter if you have decades in the industry or just want to find out more about this thing people call SEO; The Art of SEO is an exceptional guide. Coming from some of the world's leading marketers, the updated theory, insights, and strategies are first class and a must have in your library of knowledge.

Sean Kainec, Head of SEO at The Home Depot

No industry has gone through as much change in the last five years as SEO, and no resource captures how we adapt to that change better than The Art of SEO. From planning your strategy and understanding all the technical considerations to demystifying mobile and local, The Art of SEO covers it and makes it very clear that SEO is not only very much alive, but quickly becoming the strongest traffic-driving channel. Anyone working in digital or anyone who owns a website needs to read this book.

Erin Everhart, Lead Manager, Digital Marketing, SEO, at The Home Depot

The moment you think you know everything about SEO, this book will put you back in your place.

Zac Johnson, Superaffiliate and Blogger at ZacJohnson.com

Hype-free, data-rich, and loaded with insight that's essential reading for anyone who needs a deep understanding of SEO.

Seth Godin, author of We Are All Weird

The Art of SEO is the perfect complement to the science of conversion optimization. This book is a must-read volume by three highly regarded industry veterans.

Bryan Eisenberg, New York Times best-selling author of Call to Action and Always Be Testing

Anyone who wants to know how SEO really works must read The Art of SEO. This is a true reference work.

John Chow, Superblogger, JohnChow.com

The third edition of The Art of SEO expands and enhances a book that was already the industry standard for SEO education and strategy. Anyone looking to optimize their website and get better rankings on the search engines should keep this book on their desk and refer to it daily. All of the advanced technical SEO strategies are covered in a straightforward method that is easy to understand and action-oriented. When you are finished reading this book, you will have a better grasp on how search engines work and how you can optimize your website with expert proficiency. If you want to drive more traffic to your website, engage your audience on a deeper level, generate more sales, and grow your business—this books lays the plan out for you.

Joseph Kerschbaum, author of Pay-Per-Click Search Engine Marketing: An Hour A Day

DO NOT BUY THIS BOOK. Please. I beg of you. If you compete with us or any of our clients, do not buy this book. It's become our go-to source for anything—and everything—we need to know about successful search engine optimization.

Amy Africa, CEO, Eight By Eight

I've known and followed the authors for years, each a bona fide expert in their own right. Their collective wisdom in one book is truly a gift to the industry. This has to be the #1 "must-read" SEO book ever written.

Eric Ward, a.k.a. Link Moses

In The Art of SEO, Eric Enge, Stephan Spencer, and Jessie Stricchiola have taken on the daunting task of compiling a comprehensive, step-by-step walk-through of what it takes to rank well on search. They go well beyond the usual tactical aspects, addressing fundamental challenges like understanding user intent, integrating an SEO culture within your organization, properly measuring success, and managing an SEO project. This is a deep, deep dive into the world of organic optimization, and you couldn't ask for better guides

than Enge, Spencer, and Stricchiola. Clear a place on your desk, because this is going to be your SEO bible.

Gord Hotchkiss, President, Enquiro Search Solutions

An amazingly well-researched, comprehensive, and authoritative guide to SEO from some of the most well-respected experts in the industry; highly recommended for anyone involved in online marketing.

Ben Jesson, cofounder and CEO of Conversion Rate Experts

There is an art (and science) to search engine optimization. It's not always easy, it's not always obvious, and the results depend a lot on what the major search engines are tinkering with under their own hoods. Thankfully, there is a book like The Art of SEO to shine a light, give you some clues, and help you get ahead of your competitors.

Mitch Joel, President of Twist Image, and author of Six Pixels of Separation

The world of SEO has had to adapt to many changes in the past few years, so the third edition of The Art of SEO is a must read. If you want to know the latest tried-and-true SEO techniques, then look no further.

Ross Dunn, CEO/founder, StepForth Web Marketing, Inc.

With over 80% of Internet sessions starting with a search, you should be looking for ways to develop traffic from search engines. The Art of SEO is a book I continually recommend to beginners and more experienced marketers. This book can shave years off the learning curve for anyone thinking of delving into the world of search marketing. The Art of SEO walks you through the most important steps in planning and executing a top-flight program. The authors of this book are trusted individuals whose repeated, proven success working with SEO and social media marks them as leaders in the field. Easy to understand and well written, this book walks you through everything you need to understand to be successful with your own SEO campaigns. Read now, prosper now and later.

Duane Forrester, author of New York Times best-seller How to Make Money

with Your Blog and Turn Clicks Into Customers, and Senior Product Manager, Bing, former SEMPO board member

The Art of SEO is really about the science of SEO. This detailed and practical guide to SEO mastery, from a panel of all-star practitioners, will give you the edge. Get it before your competitors do!

Tim Ash, CEO of SiteTuners.com, and author of the best-selling Landing Page Optimization: The Definitive Guide to Testing and Tuning for Conversions

SEO expertise is a core need for today's online businesses. Written by some of the top SEO practitioners out there, this book can teach you what you need to know for your online business.

Tony Hsieh, CEO of Zappos.com, Inc., author of New York Times best-seller Delivering Happiness

Rarely does a work so thoroughly deconstruct the art and science of SEO: what it is, how it works, who makes it happen, and why it is important to the modern firm.

Sara Holoubek, CEO, Luminary Labs

The Art of SEO, Third Edition, reads like an Ian Fleming novel; intriguing the reader with surprising insights and exciting new ideas…all while making SEO seem oh-so-sexy.

Sean Singleton, Digital Marketing Manager, American Apparel

As a coauthor of a book people refer to as the "Bible of Search Marketing," you might think that I wouldn't recommend other search books. Not so. But I recommend only excellent search books written by outstanding search experts. The Art of SEO easily clears that high standard and is a must-read for anyone serious about organic search success.

Mike Moran, coauthor of Search Engine Marketing, Inc., and author of Do It Wrong Quickly

Search is becoming increasingly important for marketers to understand. Businesses that don't include organic search as a core component to their marketing efforts risk being left behind as their competitors and their customers move forward. This book covers a lot of ground—it's full of stats, advice, and tools to equip marketers in their organic search efforts.

Vanessa Fox, author of Marketing in the Age of Google

Roll up your sleeves, buckle your seat belt, and take your foot off the brake. You are about to go on a journey from the very basics to the very high-end, enterprise level, and then into the future of the art of SEO. These three authors have been involved in Internet marketing from the very start and have hands-on experience. These are not pundits in search of an audience but practitioners who have actually done the work, know how it's done, and have the scars to prove it. This is a dynamite primer for the beginner and a valued resource for the expert. Clear, concise, and to the point, it may not make you laugh or make you cry, but it will make you smart and make you successful.

Jim Sterne, Producer of the eMetrics Marketing Optimization Summit and Chairmen of the Web Analytics Association

Regardless of whether you're a beginner or an expert search marketer, The Art of SEO delivers! From keyword research and search analytics to SEO tools and more!

Ken Jurina, President and CEO, Epiar

There are no better guides through the world of SEO—the combined experience of these authors is unparalleled. I can't recommend highly enough that you buy this book.

Will Critchlow, cofounder, Distilled

Simply put…The Art of SEO is a smart book on search engine optimization. Neatly laid out, comprehensive and clear…this edition explains the nuances of cutting-edge tactics for improving your SEO efforts. I refer to it constantly.

Allen Weiss, founder and CEO, MarketingProfs.com

There is only a small number of books I recommend to people who want to truly grok search, and this is one of them. The Art of SEO is practical, tactical, deep and wide. There is no other book more comprehensive in scope that also goes so deeply on the technical requirements of competing in today's search marketing war. Both beginners and experienced web developers can find wisdom and guidance in its pages, and will find themselves referring to it daily for examples and decision-making help.

Dana Todd, President of Interimpathic and cofounder of SEMPO (Search Engine Marketing Professional Organization)

In your hands is a definitive collection of SEO knowledge from three leading practitioners of the art. This book is required reading for my company, and we're also recommending it to our clients.

Adam Audette, SVP, Organic Search, Merkle, and lead SEO for Zappos.com

There are no better names in the search marketing industry to write a book on the art of SEO than these three authors. Each author has gems of knowledge to share individually, and all of them teaming up to create a single book is like discovering a treasure.

Barry Schwartz, News Editor, Search Engine Land, and Editor, Search Engine Roundtable

I have personally known and respected each author for many years, and this book is a superb collection of their collective wisdom for implementing SEO for your website. I trust the information presented in this book will help readers accomplish their traffic goals. You can never know too much about SEO in this ever-changing and competitive space. Read this book.

Bruce Clay, President, Bruce Clay, Inc.

Presenting the inner mechanics of search engine optimization is a daunting task, and this book has accomplished it with flair. The book reveals the closely guarded secrets of optimizing websites in a straightforward, easy-to-understand format. If you ever wanted to unravel the mysteries of the most

enigmatic discipline on the Internet, this is the book you want as your guide. This book is so comprehensive and well written, it just might put me out of a job.

Christine Churchill, President, KeyRelevance

Written by some of the top minds in SEO!

Danny Sullivan, Editor-in-Chief, SearchEngineLand.com, and Producer, SMX: Search Marketing Expo

The Art of SEO is a masterpiece in search engine optimization techniques. Whether you're technical or creative, whether a coder, a designer, a copywriter, or a PR professional, you need this book.

Andy Beal, coauthor of Radically Transparent, founder of Trackur and Marketing Pilgrim

As an online marketing guy, I'm aware of how much your search engine visibility can either skyrocket or impede your traffic and earning potential. I rely on traffic from Google to drive my business forward. The Art of SEO is hands down the most complete and user-friendly resource for real-world application of SEO, the kind that gets you paid.

Jeremy Schoemaker, founder of ShoeMoney Media and author of Nothing's Changed But My Change: The ShoeMoney Story

This must-have book by industry heavyweights is a milestone. The material is convincing and compelling. Most important of all, the ideas make powerful strategies for successfully marketing sites online.

Disa Johnson, CEO, SearchReturn

Stephan Spencer, Eric Enge and Jessie Stricchiola represent three of the most respected minds in SEO. This book is packed with both comprehensive "explainer" content, and also actionable strategies backed by data. It's an ideal foundational study for beginners, and a critical update for veteran

marketers who want to get current. If you're looking for just one book on SEO, this is it.

Jenise Henrikson, CEO of Search Engine Journal

In an ever-changing digital landscape, implementing a successful SEO program remains a crucial foundational priority for leading brands. It drives paid, owned, and earned visibility that is critical to capturing intent wherever it is expressed. The Art of SEO is a must-read for any brand or agency that wants to drive discoverability and results through SEO. This book takes deep, technical topics and breaks them down into something applicable, accessible and usable.

Michael Kahn, Global CEO, Performics

Integration of SEO into any strategic PR plan represents the evolution of our industry. Ultimately it's this combination of SEO and PR that realizes the greatest message pull-through. With its practical tips, The Art of SEO has been invaluable to our PR firm and to me as a PR professional, helping us form our content and social media strategy as well as acquire more valuable backlinks from top media outlets.

Heidi Krupp, CEO of Krupp Kommunications

After diving into the third edition of The Art of SEO, I felt like a gladiator armed with all the right knowledge and tools to take on the ever-changing world of SEO.

Ken Rutkowski, host of Business Rockstars

Written by in-the-trenches practitioners, The Art of SEO is a well-written step-by-step guide providing sensible and practical advice on how to implement a successful SEO program. The authors have created a readable and straightforward guide filled with concise and easily adopted strategies and tactics any online business can use. I now have a great resource to recommend when people ask, "Know any good books on SEO?"

Debra Mastaler, President, Alliance-Link and Marketing Tools 24/7

Fundamentals, data, science and hard work will get you only so far. It is the artistry of SEO that can make you truly great. The backgrounds and experience of the SEO artists—Stephan, Eric and Jessie—is all but guaranteed to help you reach that next level. Read this book and take special note of what is the real art of SEO.

Todd Friesen, Director of SEO, Salesforce

SEO can be a minefield if you're not armed with the knowledge and resources to safely navigate this ever-changing landscape. Think of The Art of SEO as your field survival manual. Simply put, don't leave home without it.

Jamie Salvatori, founder and CEO, Vat19.com

The definitive book on SEO just keeps getting better, as the new third edition of The Art of SEO is packed full of helpful new information.

Brett Tabke, founder and CEO of Pubcon, the premier optimization and new media conferences

The Art of SEO is the go-to book when it comes to SEO. It is what I recommend to everyone wanting to improve their understanding of SEO, novice or seasoned expert.

Brent Csutoras, social media strategist and founder of Kairay Media

Whether you are a neophyte or an advanced search engine marketer, this book will take your business to the next level and help you drive real revenue opportunities.

Joe Pulizzi, founder of Content Marketing Institute and author of Content Inc.

Since the science of search engine optimization changes daily, understanding those changes and executing from that understanding is critical to today's business. This map in book form can help you navigate the seas of change and take control of your ship. The essential SEO guide will move you into the captain's seat of online marketing.

Toni Sikes, CEO of CODAworx and founder of The Guild

The Art of SEO

Mastering Search Engine Optimization

Third Edition

Eric Enge, Stephan Spencer, and Jessie C. Stricchiola

The Art of SEO

by Eric Enge, Stephan Spencer, and Jessie C. Stricchiola Copyright © 2015 Eric Enge, Stephan Spencer, Jessie C. Stricchiola and Rand Fishkin. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*http://safaribooksonline.com*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

- Acquisitions Editor: Mary Treseler
- Editor: Angela Rufino
- Technical Editor: Rob Woods
- Production Editor: Matthew Hacker
- Copyeditor: Rachel Monaghan
- Proofreader: Jasmine Kwityn
- Indexer: Judith McConville
- Interior Designer: David Futato
- Cover Designer: Mark Paglietti
- Illustrator: Rebecca Demarest
- October 2009: First Edition
- March 2012: Second Edition
- August 2015: Third Edition

Revision History for the Third Edition

- 2015-08-11: First Release

See *http://oreilly.com/catalog/errata.csp?isbn=9781491948965* for release details.

The O'Reilly logo is a registered trademark of O'Reilly Media, Inc. *The Art of SEO*, the cover image, and related trade dress are trademarks of O'Reilly Media, Inc.

While the publisher and the authors have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the authors disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

978-1-49194896-5

[LSI]

Foreword

In 2004, when I did my first professional work in the field of search engine optimization, Google was estimated to receive ~200 million searches each day. In 2014, that number jumped to more than 6 billion, a 30x increase over 10 years. Since its inception, web search has been a powerful tool for people to find what they need, and as a result, it's also been a powerful channel for those seeking to attract attention, traffic, actions, and customers. But in 2004, discovering how search engines worked and how to drive that traffic was a daunting challenge.

Web forums, blogs, and a handful of industry news sites were the first to take on the problem. They built communities and published resources to help both industry insiders and those aspiring to learn SEO. That's where I first cut my teeth. But learning from those sources was hard work—and there were no alternatives. Schools didn't teach SEO (the great majority of them still don't!), online courses were extremely rare (and many were low quality or straight-up misleading), and there were almost no books on the subject.

When I founded the SEOmoz blog (which became the company known today as "Moz"), my mission was to educate and to learn by doing so. Nothing has made me a better student of SEO and of marketing than being forced to write about and teach it to others.

That's why it was a great honor to have been one of the contributing writers for the original two editions of this book. When we finished that first full draft, I remember skimming through it and thinking, "Man...I wish they'd had this when I was learning SEO." That was a good feeling.

Search engines are complex. The Web is massively complex. And the human beings doing the searching, sifting, clicking, and converting—they're the most complex of all.

But all that complexity shouldn't create an impenetrable wall for those seeking knowledge. In the past, when it has, the reputation of SEO itself has suffered. CMOs, marketing managers, and small-business owners have hired or contracted professionals to perform SEO and been frustrated by the process, the requirements, and the results, often because they themselves didn't know enough about the practice to make a good choice or to create the right expectations.

By pulling back the veil on SEO, those of us tasked with teaching are enabling the right expectations to be set, the right people to be hired, and the right results from SEO investments. Arguably, no one is more qualified or has done more than the team of writers, editors, and practitioners behind this new edition of *The Art of SEO*. In this book, you'll find not only the wisdom of its authors, but the work of thousands around the world who've contributed their insight directly (and are properly cited, of course) and indirectly (we thank you, too, unacknowledged teachers of us all) to transparency in a sometimes too-opaque field.

While I could not personally contribute to this edition as I would have liked, it was my honor and privilege to once again grace these pages (even if only in this paltry foreword) alongside such an excellent team. Welcome to *The Art of SEO*.

Rand Fishkin

July 2015

Preface

The book before you is designed to be a complete and thorough education on search engine optimization for digital marketers at all levels. This third edition has been completely revamped and updated from the second edition, taking into account the vast number of changes in the search engine industry, and the digital marketing field as a whole - from the rise of social media to the explosion of mobile. Nonetheless, as with the first and second editions, you can think of it as SEO 101, SEO 102, and SEO 500. Our goal has been to help simplify a very complex, layered topic and to make it easier for people to grasp, as well as to make it easier to focus on the most important aspects of SEO for individual businesses—while at the same time providing depth of knowledge and expertise that the most advanced search veteran can utilize. As a group, we have over 40 years' experience working on SEO, a discipline involving deep proficiency in all aspects of digital marketing—from website development, information architecture, and user experience (UX) to market research, content strategy, analytics, conversion optimization, and data-driven decision making. These technical skills, along with the ability to merge the analytical with the creative, compose the SEO professional's toolkit. We have seen how SEO works over a relatively long period of time, across thousands of different websites. Any one of us could have written this book individually (in fact, one of us tried to), but we discovered that by working together we were able to create something of much greater value for you, the savvy marketer who recognizes the value of the most highly converting digital channel.

Who Should Read This Book

A little-known fact in digital marketing is that many of the most successful growth hackers name SEO as their initial core expertise upon which they developed their growth strategies. Andy Johns (@ibringtraffic), a well-respected growth strategist, started in the SEO field. He was previously the SEO lead for

Facebook, Twitter, and Quora, among other roles, and is now the Director of Growth and Revenue at Wealthfront.com. People who are involved in digital marketing at any level—especially those tasked with developing online growth strategies—should consider this book invaluable. Web developers, development managers, product marketers, brand marketers, and key business personnel should all be exposed to the information in this book if online traction and growth are part of your business objectives. If you are currently focused on SEO as a growth channel for your company or your clients, or you are looking to dive deeply into this broad yet specialized area of digital marketing expertise, you would be well served to read this book from cover to cover. After reading the entire text, you will have been exposed to all aspects of the art of SEO, and will have laid the necessary groundwork for beginning to develop your SEO skills to become a more effective digital marketer and a more educated search marketing professional.

An experienced SEO veteran will find this volume invaluable as an extensive reference to support ongoing SEO engagements: both internally, within an in-house SEO group or SEO consultancy, and externally, with SEO clients. Finally, the book will serve as a refresher course for working SEO practitioners, from the novice to the pass:[professional].

Finally, this book is perfectly suited for the academic environment as a powerful resource for digital marketing, IT, and engineering disciplines. Georgetown University, UCLA, USC, UCSD, UC Davis, University of Wisconsin, the University of Mumbai, the City University of Hong Kong, and many other academic institutions have used *The Art of SEO* as either required or recommended reading in course syllabi.

Conventions Used in This Book

The following typographical conventions are used in this book:

Italic

Indicates new terms, URLs, email addresses, filenames, file extensions, and search terms.

`Constant width`

Used for program listings, as well as within paragraphs to refer to program elements such as variable or function names, databases, data types, environment variables, statements, and keywords.

`Constant width bold`

Shows commands or other text that should be typed literally by the user.

`Constant width italic`

Shows text that should be replaced with user-supplied values or by values determined by context.

TIP

This element signifies a tip or suggestion.

NOTE

This element signifies a general note.

WARNING

This element indicates a warning or caution.

Safari® Books Online

Safari Books Online is an on-demand digital library that delivers expert content in both book and video form from the world's leading authors in technology and business.

Technology professionals, software developers, web designers, and business and creative professionals use Safari Books Online as their primary resource for research, problem solving, learning, and certification training.

Safari Books Online offers a range of plans and pricing for enterprise,

government, education, and individuals.

Members have access to thousands of books, training videos, and prepublication manuscripts in one fully searchable database from publishers like O'Reilly Media, Prentice Hall Professional, Addison-Wesley Professional, Microsoft Press, Sams, Que, Peachpit Press, Focal Press, Cisco Press, John Wiley & Sons, Syngress, Morgan Kaufmann, IBM Redbooks, Packt, Adobe Press, FT Press, Apress, Manning, New Riders, McGraw-Hill, Jones & Bartlett, Course Technology, and hundreds more. For more information about Safari Books Online, please visit us online.

How to Contact Us

Please address comments and questions concerning this book to the publisher:

- O'Reilly Media, Inc.
- 1005 Gravenstein Highway North
- Sebastopol, CA 95472
- 800-998-9938 (in the United States or Canada)
- 707-829-0515 (international or local)
- 707-829-0104 (fax)

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at *http://bit.ly/art_of_seo_3e*.

You can also head over to *The Art of SEO* book website for FAQs and to post your own burning questions! You'll have access to exclusive offers, discounts, and coupon codes on various SEO tools and services. To get exclusive access to instructional videos related to the concepts in the chapter, simply send an email to *bonuses@artofseobook.com* and we'll send you some awesome links!

To comment or ask technical questions about this book, send email to *bookquestions@oreilly.com*.

For more information about our books, courses, conferences, and news, see our

website at *http://www.oreilly.com*.

Find us on Facebook: *http://facebook.com/oreilly*

Follow us on Twitter: *http://twitter.com/oreillymedia*

Watch us on YouTube: *http://www.youtube.com/oreillymedia*

Acknowledgments

- Alex Bennert—Google News SEO
- Seth Besmertnik—business of SEO
- John Biundo—local search and schema
- Matthias Blume—metrics
- Jessica Bowman—in-house
- Christoph Cemper—tools access
- Rand Fishkin—contributions to previous editions
- Michael Geneles—metrics, tools access, and review
- Chase Granberry—tools access
- Jon Henshaw—tools access
- Greg Jarboe—video/YouTube SEO
- Dixon Jones—tools access
- Adam Lella—metrics
- Dan Liebson—local search
- Michael Martin—SEO for mobile
- David Mihm—local search

- Mark Nunney—tools access
- Cyrus Shepard—tools access
- Yulia Shevardenkova—tools access
- Andrew Shotland—local search
- Danny Sullivan—for his role in launching this industry
- Marcus Tober—tools access
- Dana Todd—awesome work as a reviewer
- Mark Traphagen—authorship and social media
- David Warmuz—tools access
- Jeff Whelpley—Angular.js SEO

CHAPTER ONE

Search: Reflecting Consciousness and Connecting Commerce

Search has become integrated into the fabric of our society. With more than 197.9 billion searches performed worldwide each month as of April 2014, approximately 6.6 billion web searches are performed every day. This means that on average about 7,500 searches are performed every single second of every day.[1] In addition, users have grown to expect that the responses to their search queries will be returned in less than one second.

If you have ever come across the various "SEO is dead" claims that make the rounds periodically in the digital marketing world, rest assured: SEO is both alive and well—and has not, as many feared, been "killed" by social media marketing as a driver of web traffic. In fact, in November 2014, Twitter announced that by renewing its focus on SEO, it increased the number of logged-out visitors coming to the site tenfold—from 7.5 million to 75 million visitors per month.[2].

The high demand for search continues because people can now obtain in mere seconds information that 20 years ago would have required a trip to the library, the use of a card catalog and the Dewey Decimal System, and a physical search through halls of printed volumes—a process that could easily have consumed two hours or more. Through the new channel of search, people can conduct many of their shopping, banking, and social transactions online—something that has changed the way our global population lives and interacts.

This dramatic shift in behavior represents what investors like to label a *disruptive event*—an event that has changed something in a fundamental way. Search engines have been at the center of this disruptive event, and having a business's website content appear prominently in the search engines when people are looking for the service, product, or resource that business provides is critical to its survival. As is the case with most paths to success, obtaining such prime search result real estate is not a simple matter, but it is one that this book aims to deconstruct and demystify as we examine, explain, and explore the ever-changing art of search engine optimization (SEO).

The Mission of Search Engines

Web searchers are free to use any of the many available search engines to find what they are seeking, so the burden is on the search engines to develop a relevant, fast, and fresh search experience. For the most part, search engines accomplish this by being *perceived* as having the most relevant results and delivering them the fastest, as users will go to the search engine they think will get them the answers they want in the least amount of time.

As a result, search engines invest a tremendous amount of time, energy, and capital in improving their relevance. This includes performing extensive studies of user responses to their search results, comparing their results against those of other search engines, conducting eye-tracking and click-through rate studies (discussed later in this chapter), and constructing PR and marketing campaigns.

Search engines generate revenue primarily through paid advertising. The great majority of this revenue comes from a *pay-per-click* (or *cost-per-click*) model, in which the advertisers pay only for users who click on their ads. Because the search engines' success depends so greatly on the relevance of their search results, manipulations of search engine rankings that result in nonrelevant results (generally referred to as *spam*) are dealt with very seriously. Each major search engine employs a team of people who focus solely on finding and eliminating spam from their search results (generally referred to as "web spam" teams). This matters to SEO professionals because they need to be careful that the tactics they employ will not be considered spam by the search engines, as this would carry the risk of incurring penalties for the websites they work on.

The Market Share of Search Engines

Figure 1-1 shows the U.S. market share for search engines in June 2015, according to comScore. As you can see, Google is the dominant search engine on the Web in the United States.

comScore Explicit Core Search Share Report* (Desktop Only) June 2015 vs. May 2015 Total U.S. – Desktop Home & Work Locations Source: comScore qSearch			
Core Search Entity	Explicit Core Search Share (%)		
	May-15	Jun-15	Point Change
Total Explicit Core Search	*100.0%*	*100.0%*	*N/A*
Google Sites	64.1%	64.0%	-0.1
Microsoft Sites	20.3%	20.3%	0.0
Yahoo Sites	12.7%	12.7%	0.0
Ask Network	1.8%	1.7%	-0.1
AOL, Inc.	1.2%	1.2%	0.0

**"Explicit Core Search" excludes contextually driven searches that do not reflect specific user intent to interact with the search results.*

Figure 1-1. Search engine market share (June 2015)

In many European countries, the disparity is even greater. However, in some markets, Google is not dominant. In China, for instance, Baidu is the leading search engine. The fact remains, however, that in most world markets, a heavy focus on Google is a smart strategy for SEO.

The Human Goals of Searching

The basic goal of a human searcher is to obtain information relevant to a specific set of keywords and/or phrases entered into a search field, also known as a query. A searcher may formulate the query as a question, but the vast majority of searches are performed by users simply entering word combinations—leaving the search engines to do the work of determining "intent." One of the most important elements to building an SEO strategy for a website is developing a thorough understanding of the psychology of your target audience, and how they use words and concepts to obtain information about the services and/or products you provide. Once you understand how the average searcher—and, more specifically, your target market—utilizes query-based search engines, you can more effectively reach and keep those users.

Search engine usage has evolved over the years, but the primary principles of conducting a search remain largely unchanged. Most search engine use consists

of the following steps:

1. Users experience the need for information. They may be looking for information on a specific website, and they will search for that website (a navigational query); they might want to buy something (a transactional query); or they might want to learn something (an informational query). We will discuss this in more detail in the following section.

2. Users formulate that need using a string of words and phrases (comprising the query). Most people formulate their queries in one to three words, though as users are becoming more web savvy, they may use longer queries to generate more specific results more quickly. Table 1-1 gives a more detailed look at the percentages of searches per query length as of April 2014, the most recent comScore study on this aspect of search.

3. Users execute the query, check the results, and if they seek additional information, will try a refined query.

Table 1-1. Searches by query length (comScore, April 2014 data)

Words	Percent of searches
1	25.8%
2	22.8%
3	18.7%
4	13.2%
5+	19.5%

When this process results in the satisfactory completion of a task, a positive experience is created for the user, the search engine, and the site providing the information or result.

Who Searches and What Do They Search For?

comScore reported that the number of search queries performed worldwide on the Web was approximately 197.9 billion across all search engines in April 2014.

comScore data also shows over 1.7 billion people were using a search engine on a given day in that month. Search engine users in the United States were slightly more likely to be women than men (50.3% versus 49.7%). According to comScore, about two-thirds of U.S. Internet users had an income of $40,000 or more (Table 1-2).

Table 1-2. Internet users by household income (May 2015)

U.S. household income	Internet users
Less than $25,000	33,730 (13.2%)
$25,000–$39,999	32,941 (12.9%)
$40,000–$59,999	42,376 (16.5%)
$60,000–$74,999	27,861 (10.9%)
$75,000–$99,999	38,166 (14.9%)
$100,000 or more	81,015 (31.6%)

You can find additional data from studies, surveys, and white papers on Search Engine Land's Search Behavior page.

All of this research data leads us to some important conclusions about web search and marketing through search engines. For example:

- Search is a major source of valuable, targeted traffic.
- Google is the dominant search player in most world markets.
- Users tend to use shorter search phrases, but these are gradually getting longer, especially for non-navigational queries.
- Search covers all types of markets.

Search is undoubtedly still one of the best and most important ways to reach consumers and build a business, regardless of that business's size, reach, or focus.

Determining Searcher Intent: A Challenge for

Search Marketers and Search Engines

Good marketers are empathetic, and smart SEO practitioners—as well as the search engines—share the goal of providing searchers with results that are relevant to their queries. Therefore, a crucial element to building an online marketing strategy around SEO and organic (sometimes referred to as "natural") search visibility is understanding your audience and how they think about, discuss, and search for your service, product, and brand.

Search engine marketers need to be aware that search engines are *tools*—resources driven by intent toward a content destination. Using the search box is fundamentally different from entering a URL into the browser's address bar, clicking on a bookmark, or clicking on a link to go to a website. Searches are performed with *intent*—the user wants to find specific information, rather than just land on it by happenstance; search is also different from *browsing*, or clicking around links on a web page. Because of this specificity, search traffic is generally of higher value to marketers than these other types of web traffic.

What follows is an examination of the different types of search queries and their categories, characteristics, and processes.

Navigational Queries

Users perform navigational searches with the intent of going directly to a specific website. In some cases, the user may not know the exact URL, and the search engine serves as the "White Pages." Figure 1-2 shows an example of a navigational query.

Figure 1-2. Navigational query

Opportunities: Pull searcher away from destination; get ancillary or investigatory traffic.

Average traffic value: Very high when searches are for the publisher's own brand. These types of searches tend to lead to very high conversion rates. However, these searchers are already aware of the company brand, and some percentage of these queries may not represent new customers. For brands other than the one being searched on, the click-through rates will tend to be low, but this may represent an opportunity to take a customer away from a competitor.

Informational Queries

Informational searches involve an incredibly broad range of queries. Consider the many types of information people might look for: local weather, driving directions, a celebrity's recent interview, disease symptoms, self-help information, how to train for a specific type of career...the possibilities are as endless as the human capacity for thought. Informational searches are primarily non-transaction-oriented (although they can include researching information about a product or service); the information itself is the goal, and no interaction beyond clicking and reading is required for the searcher's query to be satisfied.

Figure 1-3 shows an example of an informational query.

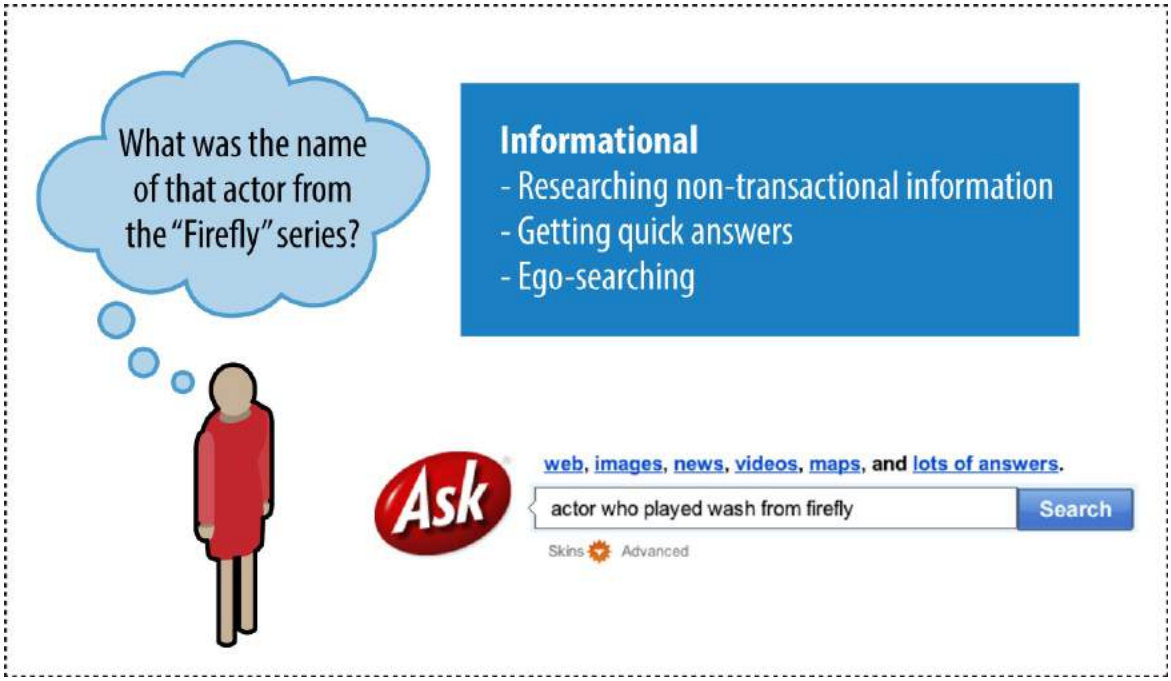


Figure 1-3. Informational query

Opportunities: Brand searchers with positive impressions of your site, information, company, and so on; attract inbound links; receive attention from journalists/researchers; potentially convert to sign up or purchase.

Average traffic value: The searcher may not be ready to buy anything just yet, or may not even have a long-term intent to buy anything, so the value tends to be "medium" at best. However, many of these searchers will later enter in a more targeted search, and this represents an opportunity to capture mindshare with those potential customers. For example, informational queries that are focused on researching commercial products or services can have high value.

Transactional Queries

Transactional searches don't necessarily have to involve a credit card or immediate financial transaction. Creating a Pinterest account, signing up for a free trial account at DomainTools, or finding the best local Japanese restaurant for dinner tonight are all transactional queries. Figure 1-4 shows an example of a transactional query.

Figure 1-4. Transactional query

Opportunities: Achieve transaction (financial or other).

Average traffic value: Very high.

Research from Pennsylvania State University and the Queensland University of Technology shows that more than 80% of searches are informational in nature, and only about 10% of searches are navigational or transactional.[3]

The researchers went further and developed an algorithm to automatically classify searches by query type. When they tested the algorithm, they found that it was able to correctly classify queries 74% of the time. The difficulty in classifying the remaining queries was vague user intent—that is, the queries could have multiple meanings.

Adaptive Search

The search engines also look at sequences of search queries to determine intent. This was confirmed in coauthor Eric Enge's interview with Jack Menzel, Product Management Director for Google Search. You can verify this by trying search sequences such as a search on *Rome* followed by a search on *hotels*.

Normally, a search on *hotels* would not include results for hotels in Rome, but

when the preceding query was for *Rome*, some results for hotels in Rome will be included. Keeping track of users' previous search queries and taking them into account when determining which results to return for a new query—known as *adaptive search*—is intended to help the search engines get a better sense of a user's intent. The search engines need to do this with care: excessive changes to the results they return based on recent query history are likely to lead to problems, so usually these types of changes are fairly limited in scope. Nonetheless, it is useful to be aware of the types of sequences of searches that users go through in their quest for information.

Searcher Intent

When you are building keyword research charts for clients or on your own sites, it can be incredibly valuable to determine the intent of each of your primary keywords. Table 1-3 shows some examples.

Table 1-3. Sample search queries and intent

Term	Queries	Intent	Monetary value
Beijing Airport	980	Nav	Low
Hotels in Xi'an	2,644	Info	Mid
7-Day China tour package	127	Trans	High
Sichuan jellyfish recipe	53	Info	Low

This type of analysis can help you determine where to place ads and where to concentrate content and links.

Hopefully, this data can help you to think carefully about how to serve different kinds of searchers based on their individual intents, and how to concentrate your efforts in the best possible areas.

Although informational queries are less likely to immediately convert into sales, this does not mean you should forgo pursuing rankings on these queries; getting your informative content in front of users seeking information can be incredibly valuable, and can turn users into potential customers. If you are able to build a relationship with users who find your site after an informational query, they may be more likely to come to you to make a related purchase at a later date, and they may also decide to share your information with others via their own website, or

through social media engagement—an indirect but potentially more valuable result than converting the single user into a paying customer.

One problem in search is that when most searchers formulate their search queries, their input is limited to just a handful of words (generally one to four). Because most people don't have a keen understanding of how search engines work, they often provide queries that are too general or that are presented in a way that does not provide the search engine (or the marketer) with what it needs to determine, with 100% accuracy 100% of the time, their specific intent. Additionally, search engine users may not have a specific intent for an individual search query beyond "curiosity" about a general trending topic, or subject matter. While this poses potential difficulty in delivering relevant results, it also poses great opportunity to capture the mind of someone who may not know what he is looking for, specifically, but who takes an interest in the subsequent variety of results the search engine (and search marketers) deliver in response.

These types of general queries are important to most businesses because they often get the brand and site on the searcher's radar, and this initiates the process of building trust with the user. Over time, the user will move on to more specific searches that are more transactional or navigational in nature.

If, for instance, companies buying pay-per-click (PPC) search ads bought only the high-converting navigational and transactional terms and left the informational ones to competitors, they would lose market share to those competitors. Over the course of several days, a searcher may start with *digital cameras*, hone in on *Olympus OMD*, and then ultimately buy from the store that showed up in her search for *digital cameras* and pointed her in the direction of the Olympus OMD model.

Given the general nature of how query sessions start, though, determining intent is quite difficult, and it can result in searches being performed where the user does not find what he wants—even after multiple tries. A July 2013 study conducted by ForeSee for the American Customer Satisfaction Index (ACSI) found that 77% of Google users and 76% of Bing users were satisfied with their experiences.

While 77% satisfaction is an amazing accomplishment given the complexity of building a search engine, this study still showed that more than 22% of users were not satisfied with their search results; additionally, these numbers represent

the lowest levels of customer satisfaction with search engines since 2003, and could reflect users' dissatisfaction with the number of ads that are infiltrating *search engine results pages* (SERPs). Greg Sterling of Search Engine Land provides a broad analysis of these issues (for the full discussion, see "Google Gets Lowest Score Ever in Customer Satisfaction Survey"). It will be interesting to see how these satisfaction numbers evolve over time. The important takeaway here is that in all instances, it is still difficult to determine exactly a searcher's intent; and when the searcher's intent is broad, there is ample opportunity to leverage your content assets with SEO.

As an SEO practitioner, you should be aware that some of the visitors that you succeed in attracting to your site may have arrived for the wrong reasons (i.e., they were really looking for something else), and these visitors are not likely to help your business goals. Part of your task as an SEO is to maintain a high level of relevance in the content placed on the pages you manage, to help minimize this level of waste—while still attempting to maximize SERP "mindshare" and gain brand exposure.

How People Search

Search engines invest significant resources into understanding how people use search, enabling them to produce better (i.e., faster, fresher, and more relevant) search engine results. For website publishers, the information regarding how people use search can be used to help improve the usability of a site as well as search engine compatibility.

User interactions with search engines can also be multistep processes, as indicated in the user search session documented by Microsoft and shown in Figure 1-5.

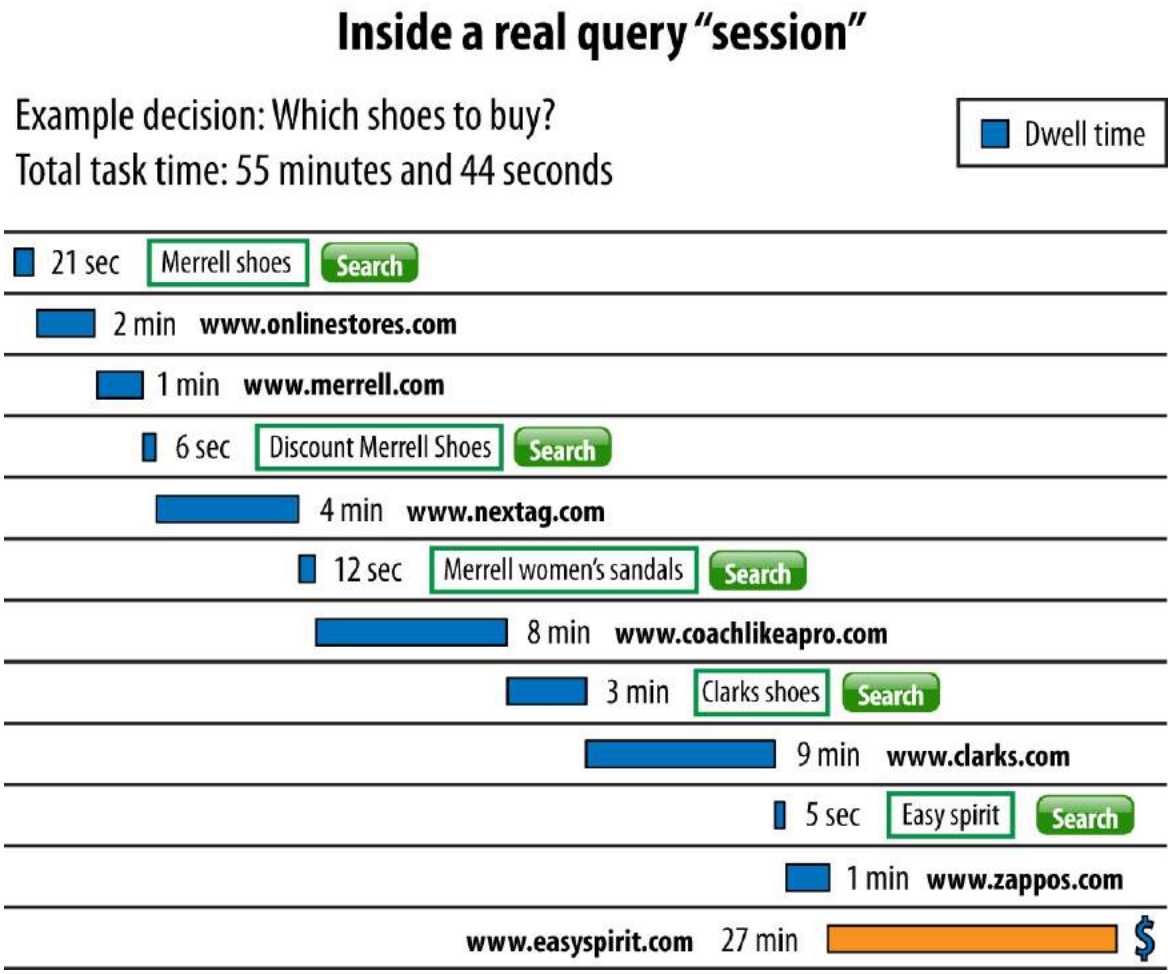


Figure 1-5. Merrell shoes user search session

In this sequence, the user performs five searches over a 55+ minute period before making a final selection. The user is clearly trying to solve a problem and works at it in a persistent fashion until the task is done.

Marin Software provided us with data on one consumer durable retailer (whose products represent high-cost, considered purchases) for whom 50% of the orders involved more than 10 clicks leading up to the conversion event.

For this particular retailer, when you look at the number of different ad groups that were clicked on in those 10 clicks, the clicks were mostly on the same keyword. In fact, for more than 75% of all conversions that came from multiple paid clicks, all the clicks were from the same ad group. Only 7% of conversions came from three different ad groups (and none from more than that).

Table 1-4 shows the average delay between the first click received by the site and the resulting purchase for this example retailer.

Table 1-4. Delay between first click and purchases

Delay between first click and purchases	Percentage of users
Same day	50%
2–7 days	9%
8–30 days	12%
31–90 days	26%
More than 90 days	3%

This behavior pattern indicates that people are thinking about their tasks in stages. As in our Merrell shoes example in Figure 1-5, people frequently begin with a general term and gradually get more specific as they get closer to their goal. They may also try different flavors of general terms. In Figure 1-5, it looks like the user did not find what she wanted when she searched on *Merrell shoes*, so she then tried *discount Merrell shoes*. You can then see her refine her search, until she finally settles on Easy Spirit as the type of shoe she wants.

This is just one example of a search sequence, and the variety is endless. Figure 1-6 shows another search session, once again provided courtesy of Microsoft.

In this search session, the user has a health concern. This particular user starts with a five-word search, which suggests that she may have some experience using search engines. At 3:01, her search on *headache pregnant 3rd trimester* leads her to Yahoo! Answers. After visiting this site, the user suddenly gets more specific in her search.

She begins to focus on gestational diabetes, perhaps because something she saw on Yahoo! Answers led her to believe she may have it. The session culminates in a search for *first signs of gestational diabetes*, which suggests that she has concluded that this is quite possibly the issue she is facing.

The session stops there. It may be that at this point the user feels she has learned what she can. Perhaps her next step is to go to her doctor with her concerns, prepared to ask a number of questions based on what she has learned.

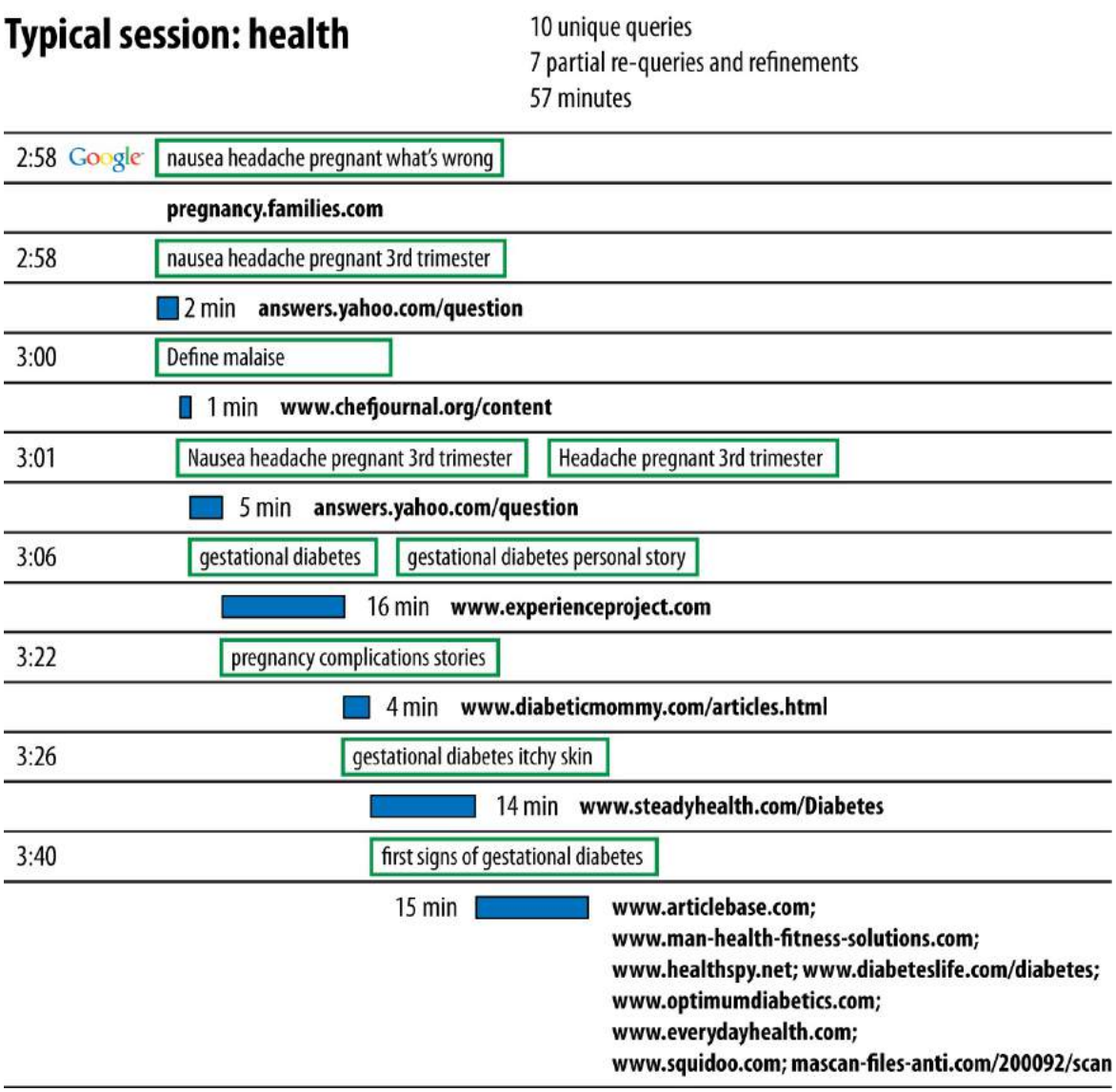


Figure 1-6. Health user search session

Our next search session example begins with a navigational search, where the user simply wants to locate the travel website Orbitz.com (see Figure 1-7). The user's stay there is quite short, and she progresses to a search on *Cancun all inclusive vacation packages*. Following that, she searches on a few specific resorts and finally settles on *cancun riviera maya hotels*, after which it appears she may have booked her hotel—the final site visited on that search is for Occidental Hotels & Resorts, and the direction of her searches changes after that.

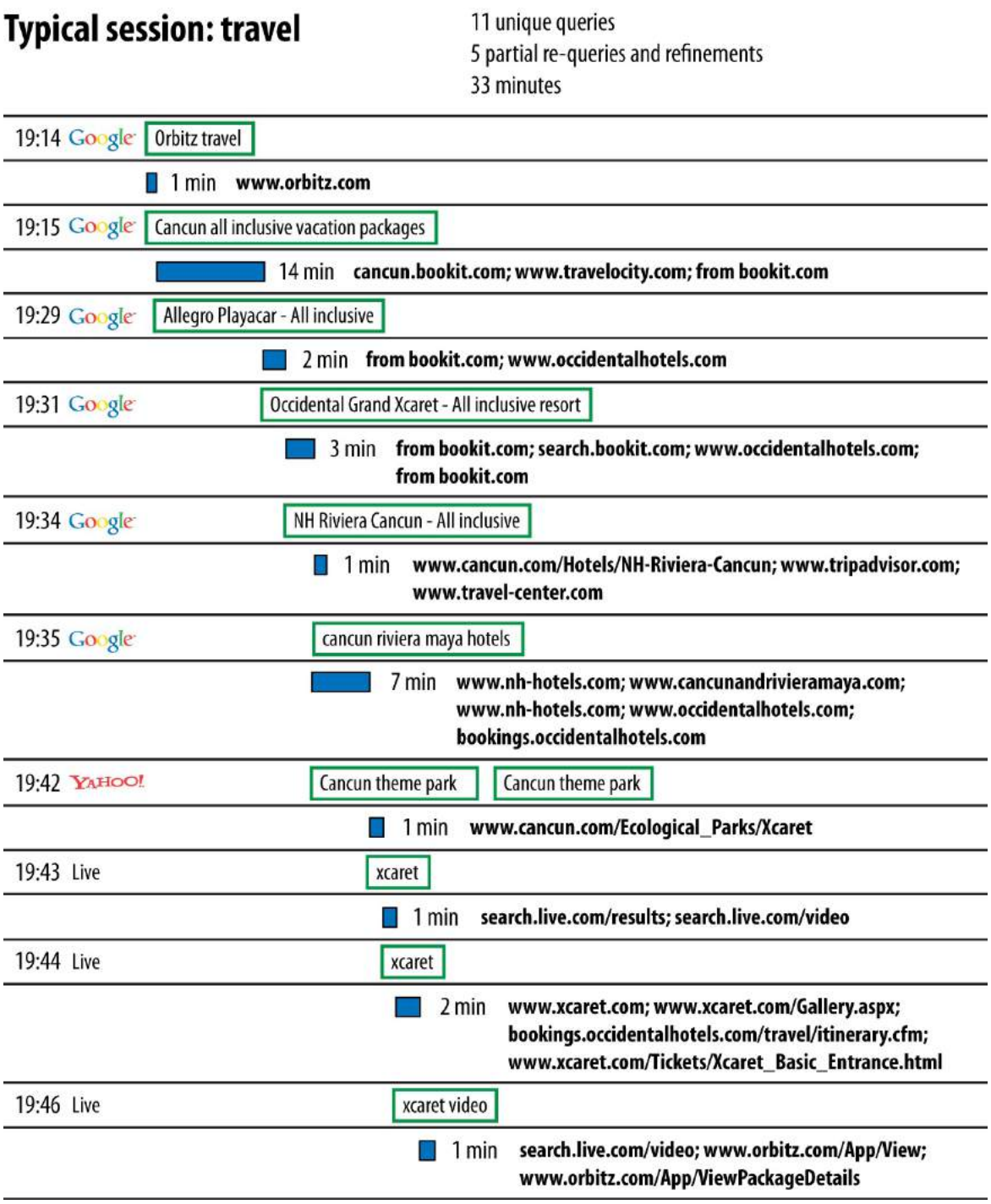


Figure 1-7. Travel user search session

At that point, the user begins to look for things to do while she is in Cancun. She conducts a search for *cancun theme park* and then begins to look for information on *xcaret*, a well-known eco park in the area.

Users traverse countless different scenarios when they are searching for something. These example search sessions represent traditional desktop interactions.

Recent data from mobile search shows different behavior for mobile searchers, who are more likely to be close to completing a transaction. Data from a May 2011 eMarketer study showed that 55% of people visited a business they found in the search results after searching for information on their smartphone devices. Search engines do a lot of modeling of these different types of scenarios to enable them to provide better results to users. The SEO practitioner can benefit from a basic understanding of searcher behavior as well. We will discuss this in more detail in Chapter 2.

How Search Engines Drive Ecommerce

People make use of search engines for a wide variety of purposes, with some of the most popular being to research, locate, and buy products. Ecommerce sales reported by the U.S. Census Bureau were a healthy $80.3 billion in the first quarter of 2015.[4] Forrester Research forecasts that U.S. ecommerce retail sales will reach $370 billion by 2017, outpacing sales growth at brick-and-mortar stores, as shown in Figure 1-8.

Figure 1-8. Forrester Research online retail forecast to 2017

It is important to note that search and offline behavior have a heavy degree of interaction, with search playing a growing role in driving offline sales. In April 2014, Google, in an effort to solve the online search/offline conversion attribution dilemma, announced the launch of a pilot program within its AdWords system called “In-Store Attribution Transaction Reporting,” teaming up with large data providers Axciom Corp. and DataLogix Holdings Inc. to combine cookie data with offline personal information databases. It will be interesting to see how this program performs, and to what extent it helps search marketers close the attribution gap between search and brick-and-mortar conversions.

The Mobile Shift

In March 2015, eMarketer published a study showing that mobile advertising spend continues to increase at the expense of desktop advertising, and projects that by 2019, mobile ad spending will rise to $65.87 billion, or 72.2% of total U.S. digital advertising spend.[5] Figure 1-9 shows the steep upward trend projection of mobile ad spend through 2019, reflecting consumers’ increasing use of mobile devices and tablets for search—and search marketers are taking notice.

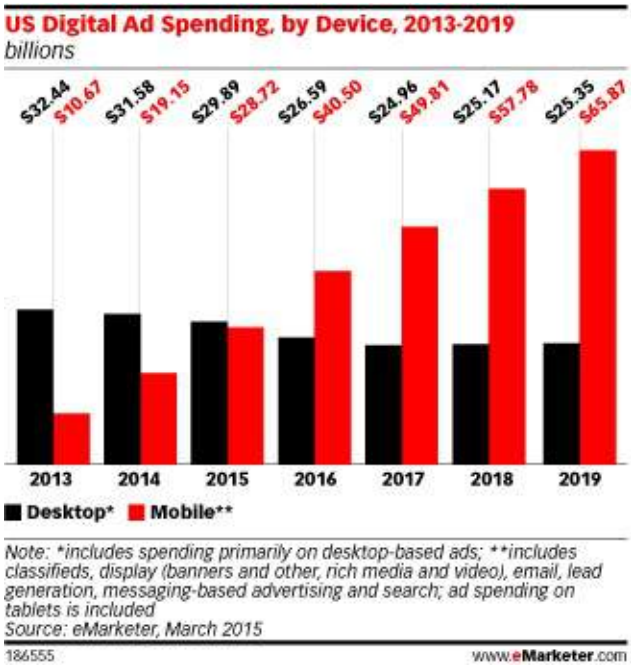


Figure 1-9. eMarketer projected mobile ad spending through 2019

The History of Eye Tracking: How Users Scan Results Pages

Way back in 2006, research firm Enquiro (now called Mediative) conducted heatmap testing with search engine users that produced fascinating results related to what users see and focus on when engaged in search activity.[6] Figure 1-10 depicts a heat map showing a test performed on Google; the graphic indicates that users spent the most amount of time focusing their eyes in the upper-left area, where shading is the darkest. This has historically been referred to in search marketing as the “Golden Triangle.”

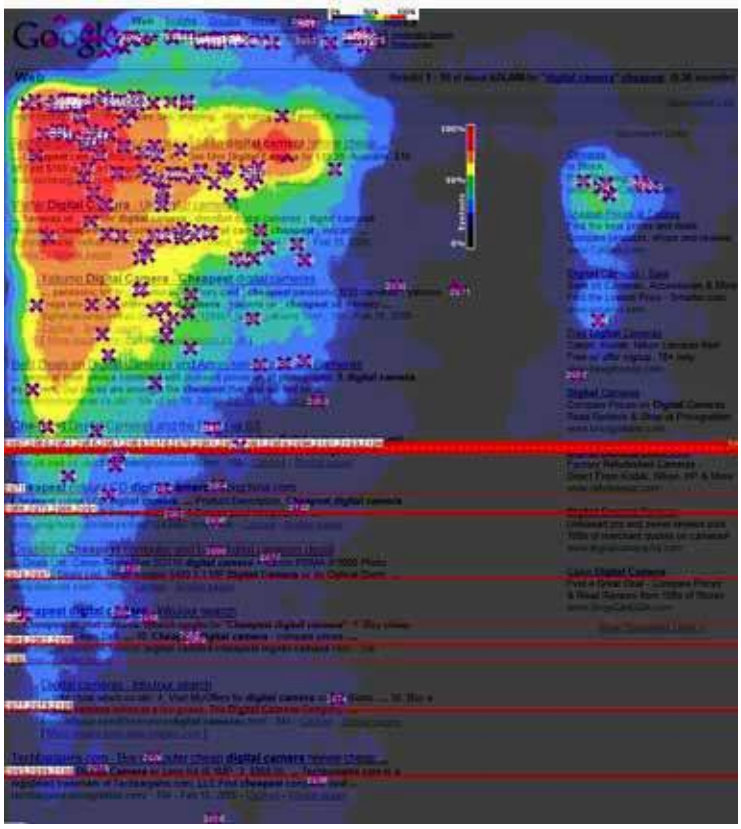


Figure 1-10. Enquiro eye-tracking results, 2006

This particular study perfectly illustrated how little attention has traditionally been paid to results lower on the page versus those higher up, and how users' eyes are drawn to bold keywords, titles, and descriptions in the organic results versus the paid search listings, which receive comparatively little attention. It also showed that different physical positioning of on-screen search results resulted in different user eye-tracking patterns. When viewing a standard Google results page, users tended to create this "F-shaped" pattern with their eye movements, focusing first and longest on the upper-left corner of the screen, then moving down vertically through the first two or three results, across the page to the first paid page result, down another few vertical results, and then across again to the second paid result. (This study was done only on left-to-right language search results—results for Chinese, Hebrew, and other non-left-to-right-reading languages would be different.)

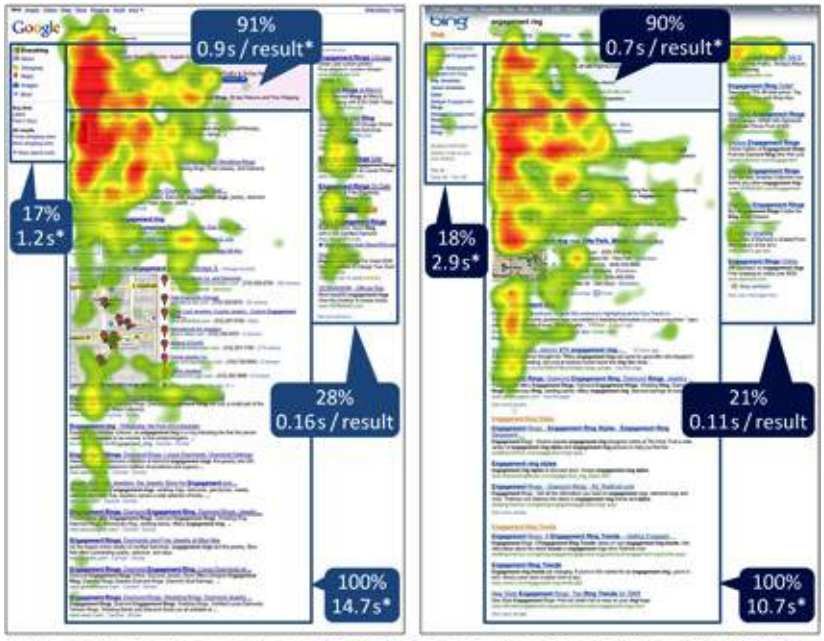
In May 2008, Google introduced *Universal Search* (discussed in more detail in Chapter 10), which was a move from simply showing the 10 most relevant web

pages (referred to as “10 blue links”) to showing other types of media, such as videos, images, news results, and so on, as part of the results of a basic web search. The other search engines followed suit within a few months, and the industry now refers to this general concept as *blended search*.

Blended search created more of a chunking effect of user eye attention, with the focus “chunks” hovering over various rich media objects such as images or video. Understandably, our eyes are drawn to the image first, then to the text beside it to see whether it corresponds to the image or video thumbnail. Based on an updated eye-tracking study published by Enquiro in September 2007, Figure 1-11 shows what the eye-tracking pattern on a blended search page looked like.

Figure 1-11. Enquiro eye-tracking results, blended search, 2007

A January 2011 study performed by User Centric[7] showed similar results, as shown in Figure 1-12.



Heatmaps showing the aggregate gaze time of all 24 participants on Google (left) and Bing (right) for one of the transactional tasks. The red color indicates areas that received the most total gaze time (4.5 seconds and above). Each callout includes the percentage of participants who looked at the area and the time (in seconds) they spent looking there. The numerical data are an average across all four tasks. Asterisks indicate values that were significantly different between Google and Bing at alpha = .1.

Figure 1-12. User Centric eye-tracking results, 2011

In 2010, Enquiro investigated the impact of Google Instant on search usage and attention, noting that for queries in its study:

- Percent of query typed decreased in 25% of the tasks, with no change in the others.
- Query length increased in 17% of the tasks, with no change in the others.
- Time to click decreased in 33% of the tasks and increased in 8% of the tasks.

In more recent studies, we can see how the continuously changing search results are shifting how users focus on a results page.

In October 2012, Mediative conducted a new eye-tracking study analyzing the areas of attention on Google Maps results pages for locally targeted searches, and as with a Google web search results page, users' eyes were most drawn to the topmost search result, as shown in Figure 1-13.

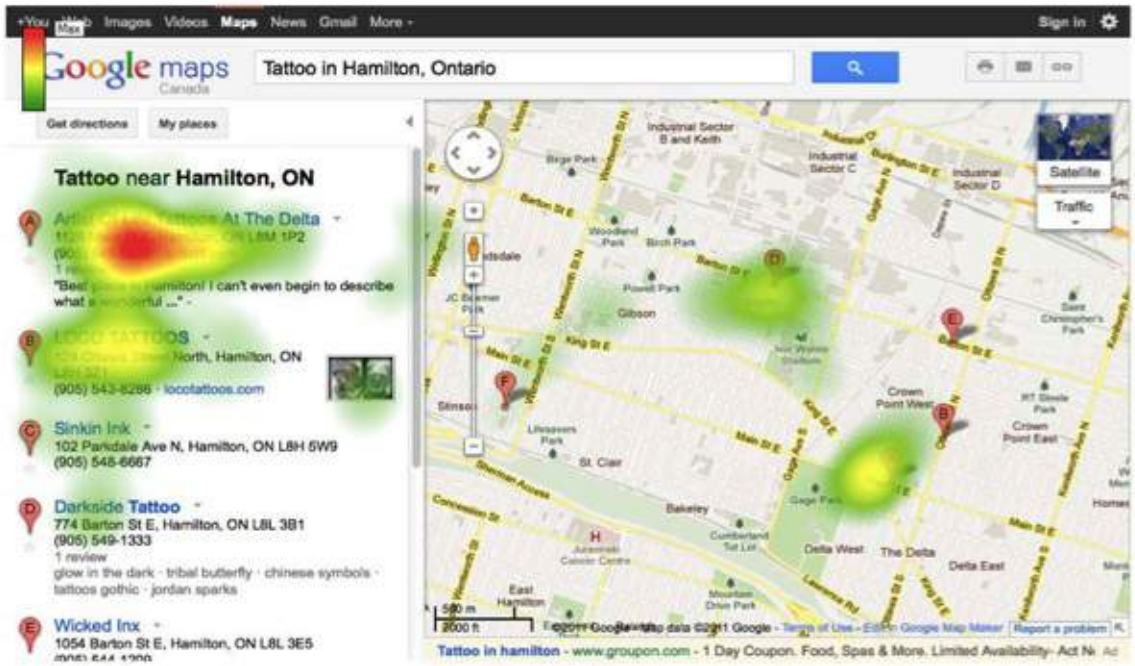


Figure 1-13. Mediative eye-tracking results, 2012

A more recent study by Mediative, issued in 2014, called "The Evolution of Google Search Results Pages and Their Effect on User Behavior" shows us a fascinating progression away from the "Golden Triangle" as the search results—as well as search platforms (mobile!)—evolve. Figure 1-14 is the heatmap from this study for an updated Google SERP showing how user attention is moving more vertically than horizontally. Some have theorized that this shift in attention is due to the increased adoption of mobile search by mobile device users, while others think that the ever-changing SERP makeup that pushes the "blue text link" results lower on the page (e.g., the increase in the number prominence and content of top-of-page paid search ads, as well as universal and local listings for various queries) is contributing to this change.

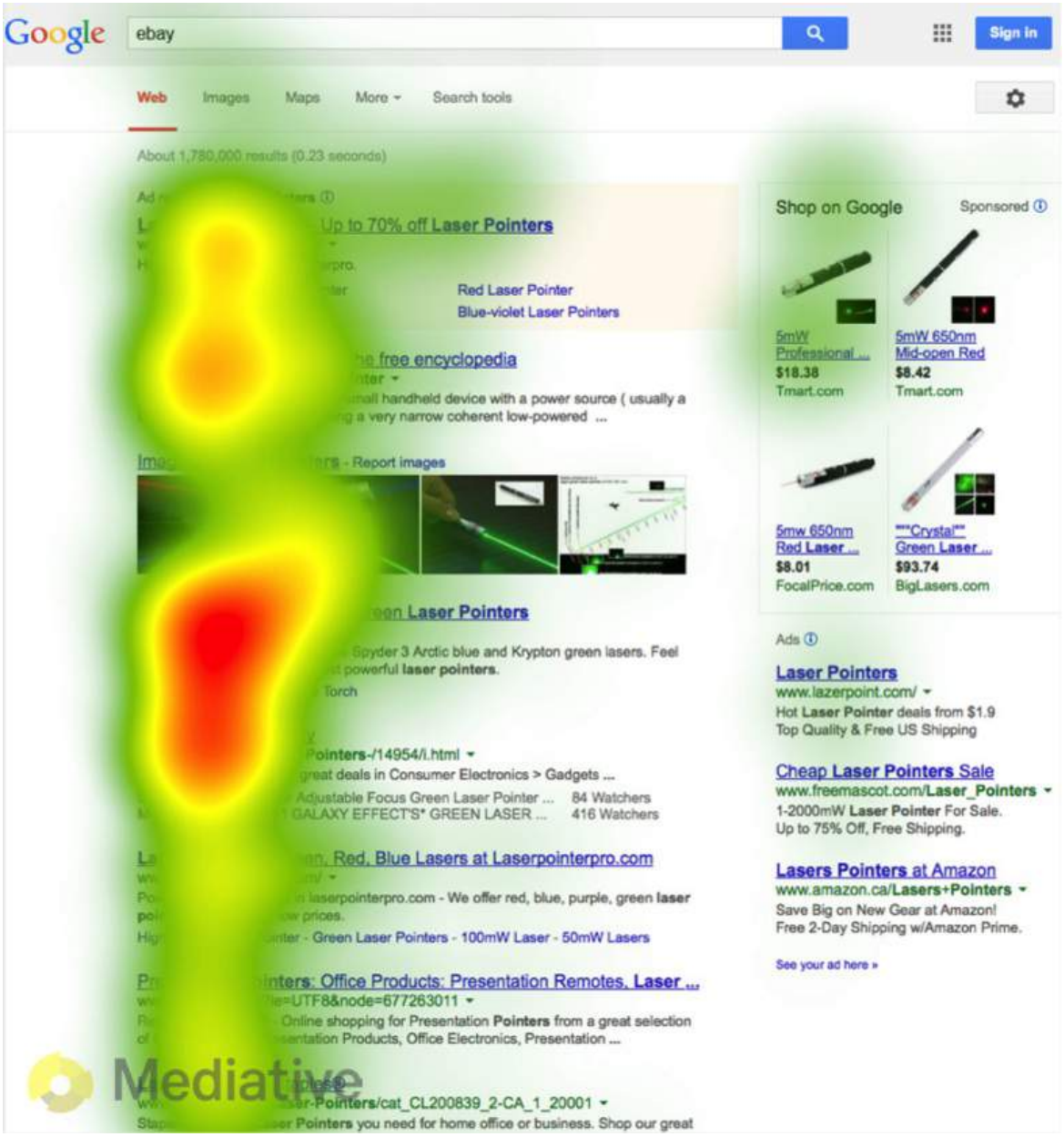


Figure 1-14. Google SERP eye-tracking results, 2014: users have moved from horizontal to vertical scanning

Another notable trend in user eye attention is that because the #1 organic site's actual position on the SERP is moving lower, sites that are positioned lower on the results page are seeing more click activity than in years past—making ranking in organic sites at these lower positions more valuable than in the past as well. Figure 1-15 highlights the increased attention being placed lower on the page.

Figure 1-15. Google SERP, eye-tracking results, 2014: lower SERPs get more attention

These types of studies are a vivid reminder of how important search engine results pages really are. And, as the eye-tracking research demonstrates, as

search engines—and therefore search results—continue to evolve, users' search and engagement patterns will follow suit. There will be more items on the page for searchers to focus on, more ways for searchers to remember and access the search listings, and more interactive, location-based delivery methods and results layouts—which will keep changing as other search environments and platforms, such as mobile, continue to evolve.

Figure 1-16 shows a "Mega SERP" published by Moz: a display of the current universe of potential results within a Google SERP.

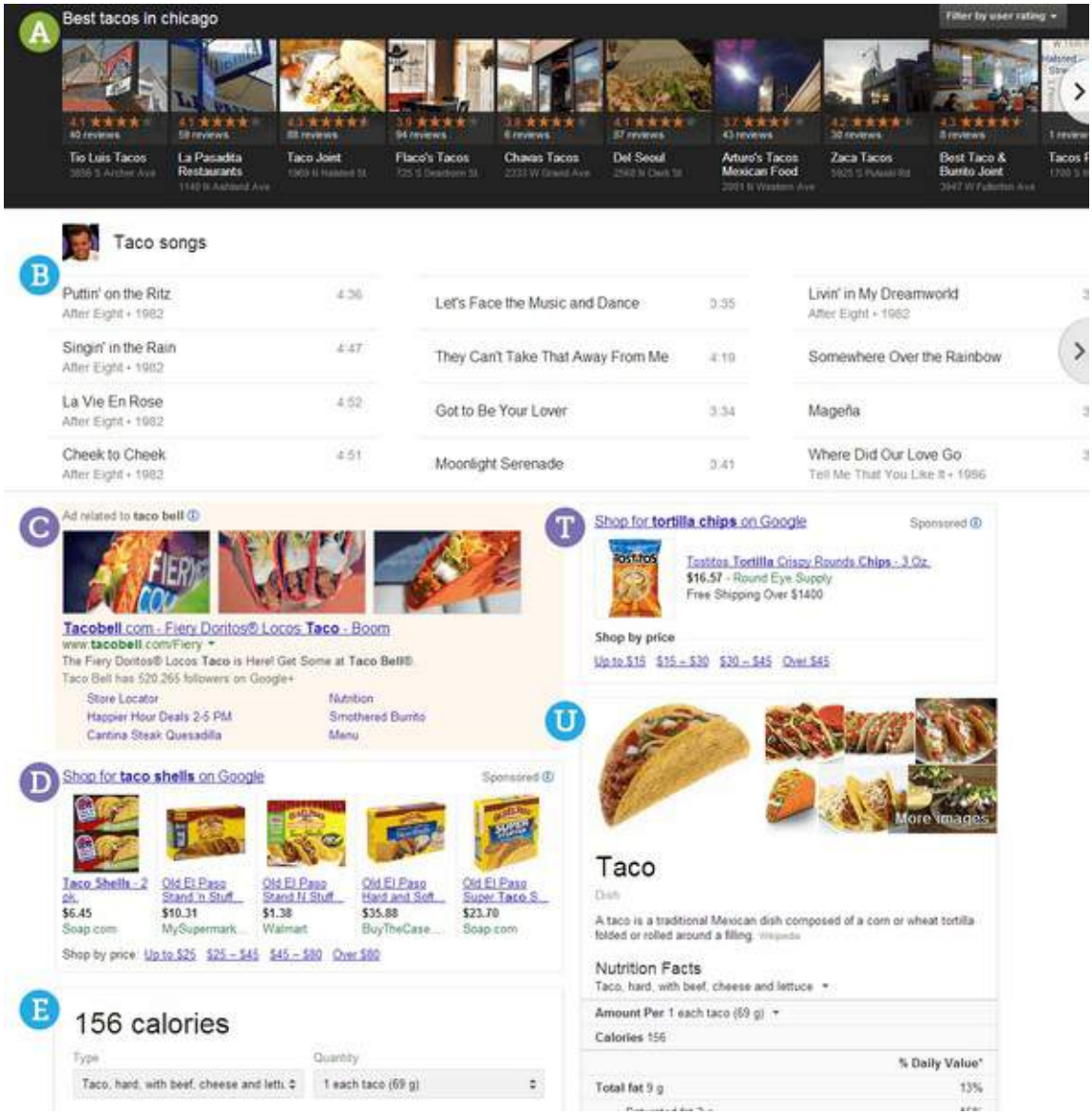


Figure 1-16. Moz mega-SERP

Click Tracking: How Users Click on Results (Organic Versus Paid)

By now, you should be convinced that you want your site content to be prominently displayed within SERPs. It never hurts to be #1 in the organic search results.

In contrast, data shows that you may *not* want to be #1 in the paid search results, because the resulting cost to gain the #1 position in a PPC campaign can reduce the total net margin on your campaign. A study released by AdGooroo in June 2008[8] found that:

> *Bidding for top positions usually makes financial sense only for high-budget, brand-name advertisers. Most other advertisers will find the optimal position for the majority of their keywords to lie between positions 5–7.*

Of course, many advertisers may seek the #1 position in paid search results, for a number of reasons. For example, if they have a really solid backend on their website and are able to make money when they are in the #1 position, they may well choose to pursue it. Nonetheless, the data from the survey suggests that there are many organizations for which being #1 in paid search does not make sense.

Even if your organic ranking is #1, you can still increase the ranking page's click rate by having a sponsored ad above it or in the righthand column. The AdGooroo survey showed that having a prominent paid ad on the same search results page makes your #1 organic ranking receive as much as 20% more clicks.

Distribution of Search Results and Traffic

To start breaking this down a bit, Figure 1-17 shows the screen real estate occupied by the two types of search results. This screenshot was taken prior to Google's January 2012 "Search, plus Your World" announcement, but it's the type of screen layout related to studies that will help you understand which portions of the search results receive the most clicks.

This example from Google shows how the paid results appear above and to the right of the organic search results. Note that Google often does not show paid results above the organic results, in which case the paid results show up only on

the right.

Your position in the results has a huge impact on the traffic you will receive. Studies on the impact of SERP position have shown widely varying results, but do agree that the advantage of higher positions is significant.

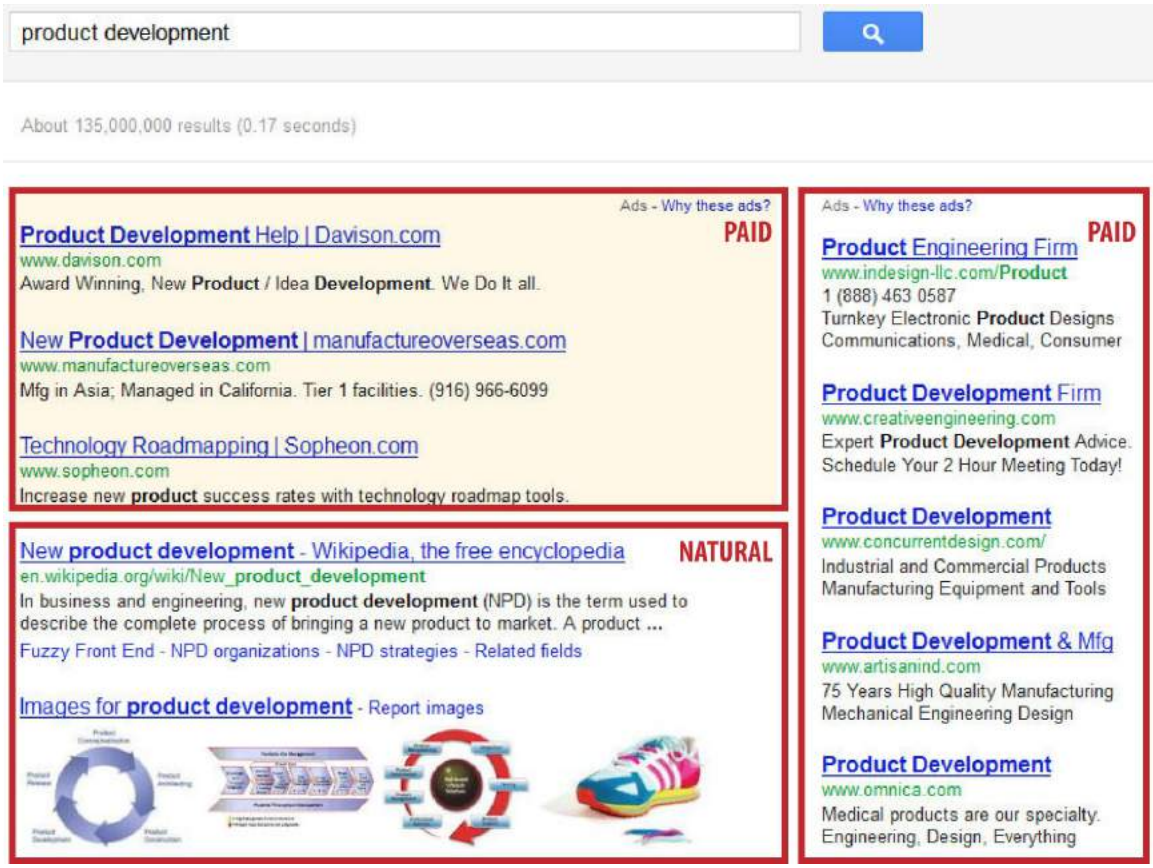


Figure 1-17. Paid and organic search results

In 2013, Catalyst Search Marketing published the most recent comprehensive study of Google CTR (click-through rate), which shows CTR data for both desktop and mobile searches, as well as for branded and unbranded searches.[9] In this study, Catalyst found that the top four Google SERP positions received 83% of first-page organic clicks, and that 48% of searches result in a first-page organic click. Table 1-5 shows the average CTR (percentage of clicks received) by Google SERP position.

Table 1-5. Google CTR by position

Average position	Average CTR
1.00	17.16
2.00	9.94
3.00	7.64
4.00	5.31
5.00	3.50
6.00	1.63
7.00	1.09
8.00	1.04
9.00	0.44
10.00	0.51

This data provides insight around platform CTR differentiators (desktop versus mobile), searcher intent CTR differentiators (branded versus unbranded), and query length CTR differentiators (one, two, three, and four-plus-word query data).

In Table 1-6, you can see that the more *refined* a user's search query is—that is, the more words used in the query—the greater the likelihood that the user will click on search results farther down the page.

Table 1-6. Google CTR by position and query length

Average position	One word	Two words	Three words	Four+ words
1.00	17.68	16.32	16.28	18.76
2.00	4.41	8.30	10.58	12.41
3.00	3.56	6.86	7.93	8.64
4.00	1.50	4.10	5.12	6.38
5.00	1.86	2.64	3.61	4.11
6.00	0.37	1.10	1.60	2.18
7.00	0.00	1.08	1.12	1.36
8.00	0.00	1.13	1.45	0.84
9.00	0.08	0.52	0.36	0.53
10.00	0.44	0.42	0.54	0.62

As the search engines, search engine results pages, and searcher behavior evolve,

the practice of search engine optimization must evolve in tandem.

Conclusion

Search has penetrated the very fabric of global society. The way people learn, work, share, play, shop, research, socialize, and interact has changed forever, and organizations, causes, brands, charities, individuals—almost all entities—need to have an Internet presence of some kind, and they need the search engines and search functionality within all platforms to generate exposure and facilitate engagement. As our society moves ever closer to a professional consumer ("prosumer") economy, the ways in which people create, publish, distribute, and ultimately find information and resources on the Web will continue to be of great importance. This book will investigate further just how search, and therefore search engine optimization, is at the center of the web ecosystem and thus is your key to success in the ever-evolving digital economy.

[1] comScore, "comScore Releases April 2014 U.S. Search Engine Rankings," May 16, 2014, http://www.comscore.com/Insights/Market-Rankings/comScore-Releases-April-2014-US-Search-Engine-Rankings.

[2] Danny Sullivan, "Twitter: Renewed Focus On SEO Generated 10 Times More Visitors," November 12, 2014, http://searchengineland.com/twitter-seo-more-visitors-208160.

[3] Bernard J. Jansen, Danielle L. Booth, and Amanda Spink, "Determining the Informational, Navigational, and Transactional Intent of Web Queries," *Information Processing and Management* 44 (2008) 1251–1266, *http://bit.ly/jansen_user_intent*.

[4] U.S. Census Bureau News, "Quarterly Retail ECommerce Sales, 1st Quarter 2015," May 15, 2015, *http://bit.ly/ecommerce_q1_2015*.

[5] eMarketer, "Mobile Will Account for 72% of US Digital Ad Spend by 2019: Consumer Usage and Better Ad Formats Drive Dollars to Mobile Apps," March 24, 2015, *http://bit.ly/mobile_ad_spend*.

[6] Mediative, published by Enquiro Search Solutions. "Eye Tracking Report 2: Google, MSN, and Yahoo! Compared!" November 2006, *http://bit.ly/eye_tracking_paper*.

[7] Aga Bojko, “Eye Tracking Bing Vs. Google: A Second Look,” GfK Insights Blog, January 27, 2011, *http://blog.gfk.com/2011/01/eye-tracking-bing-vs-google-a-second-look/*.

[8] AdGooroo, “Research Update: How Keyword Length and Ad Position Impact Clickthrough Rate and Cost-per-click on Google Adwords”. June 2008.

[9] Catalyst, “Google CTR Study,” *http://www.catalystsearchmarketing.com/pubs/google-ctr-study/*.

CHAPTER TWO

Search Engine Basics

In this chapter, we will begin to explore how search engines work. Building a strong foundation on this topic is essential to understanding the SEO practitioner's craft.

As we discussed in Chapter 1, people have become accustomed to receiving nearly instantaneous answers from search engines after they have submitted a search query. In Chapter 1 we also discussed the volume of queries (more than 7,500 per second). As early as 2008, Google knew about 1 trillion pages on the Web.[1] At SMX Advanced in Seattle in 2014, Google's Gary Illyes stated that Google now knows about 30,000 trillion pages on the Web. The scale of the Internet/Web (sometimes called the Interwebs) is growing fast!

Underlying the enormous problem of processing all these pages is the complex nature of the Web itself. Web pages include text, video, images, and more. It's easy for humans to understand these and to transition seamlessly between them, but software lacks the intelligence we take for granted. This limitation and others affect how search engines understand the web pages they come across. We'll discuss some of these limitations in this chapter.

Of course, this is an ever-changing landscape. The search engines continuously invest in improving their ability to process the content of web pages. For example, advances in image and video search have enabled search engines to inch closer to human-like understanding, a topic that will be explored more in the section "Vertical Search Engines".

Understanding Search Engine Results

In the search marketing field, the pages the engines return to fulfill a query are referred to as *search engine results pages* (SERPs). Each engine returns results in a slightly different format, and these may include *vertical results*—results that can be derived from different data sources or presented on the results page in a different format, which we'll illustrate shortly.

Understanding the Layout of Search Results Pages

Figure 2-1 shows the SERPs in Google for the query *stuffed animals*.

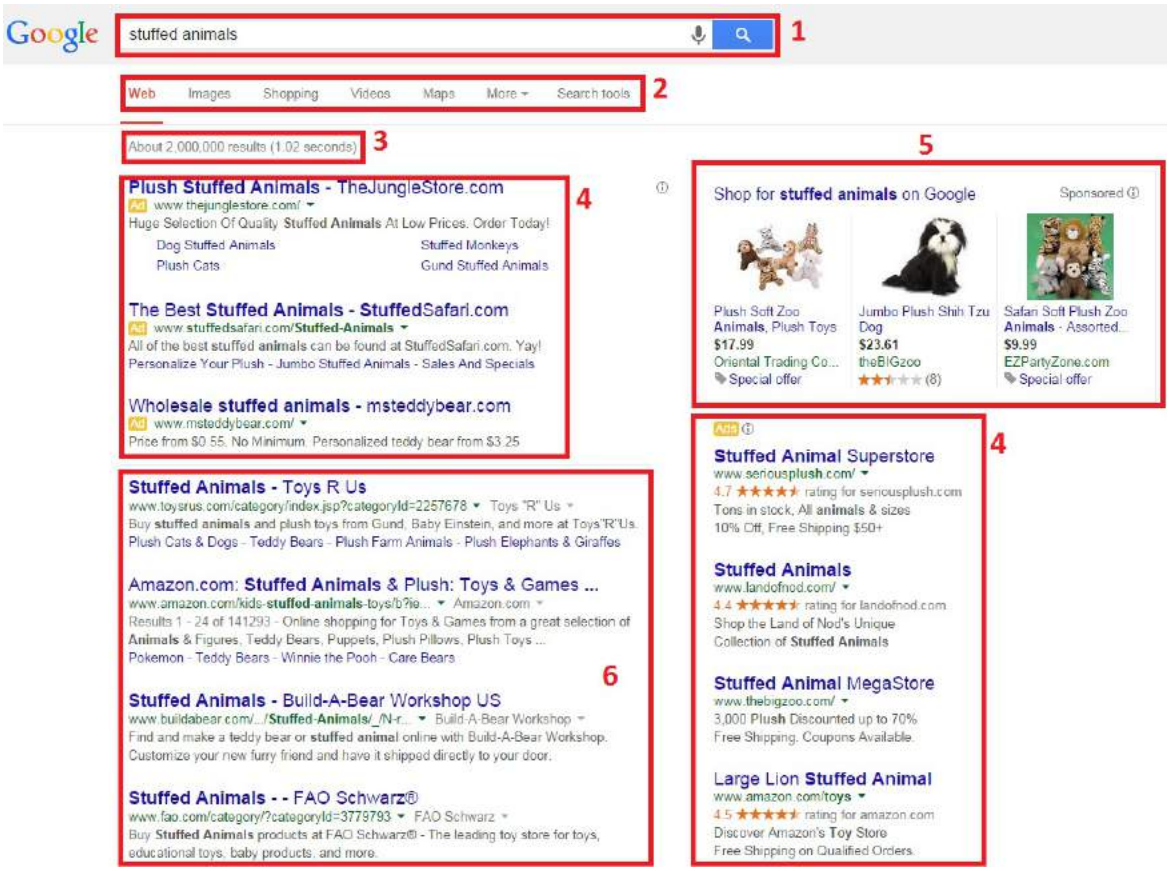


Figure 2-1. Layout of Google search results

The various sections outlined in the Google search results are as follows:

- Search query box (1)
- Vertical navigation (2)
- Results information (3)
- PPC advertising (4)
- Google product search results (5)
- Natural/organic/algorithmic results (6)

Even though Yahoo! no longer does its own crawl of the Web or provides its own search results information (it sources them from Bing), it does format the

output differently. Figure 2-2 shows Yahoo!'s results for the same query.

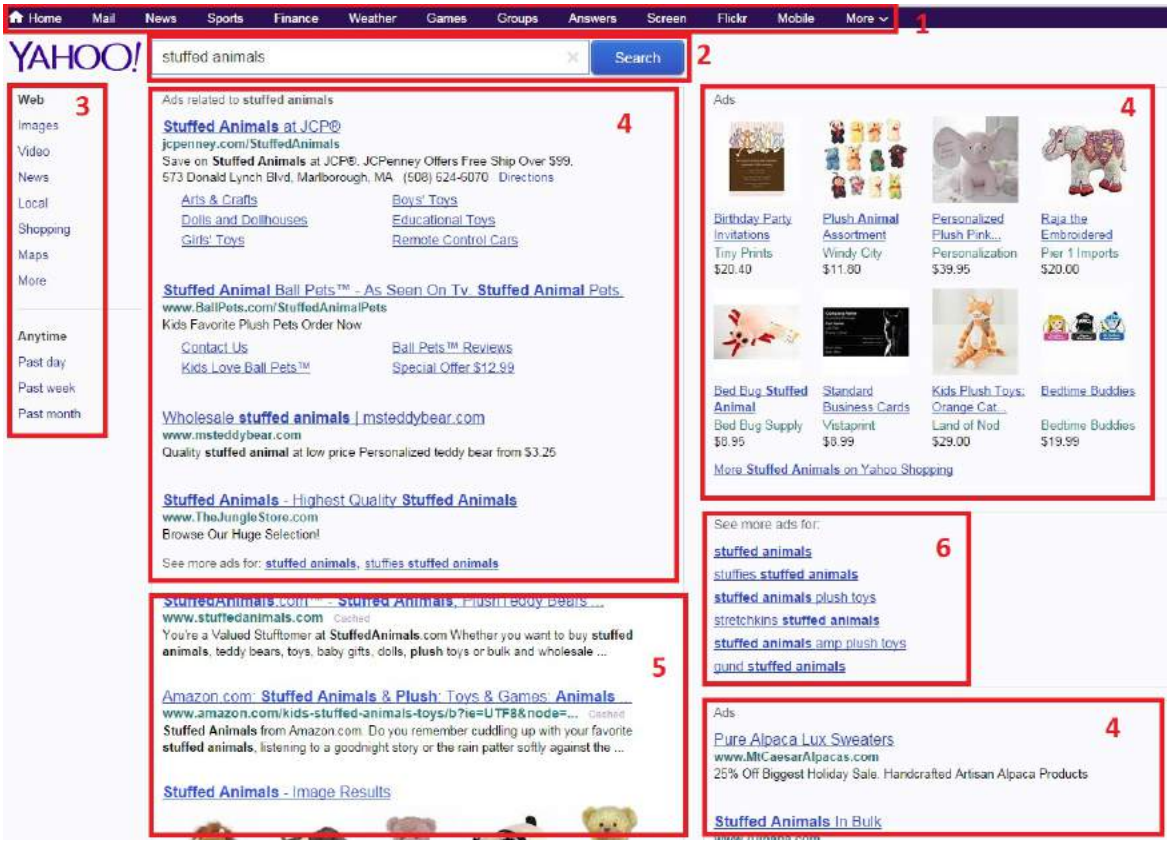


Figure 2-2. Layout of Yahoo! search results

The sections in the Yahoo! results are as follows:

- Vertical navigation (1)
- Search query box (2)
- Horizontal navigation (3)
- PPC advertising (4)
- Natural/organic/algorithmic results (5)
- Navigation to more advertising (6)

Figure 2-3 shows the layout of the results from Microsoft's Bing for *stuffed*

animals.

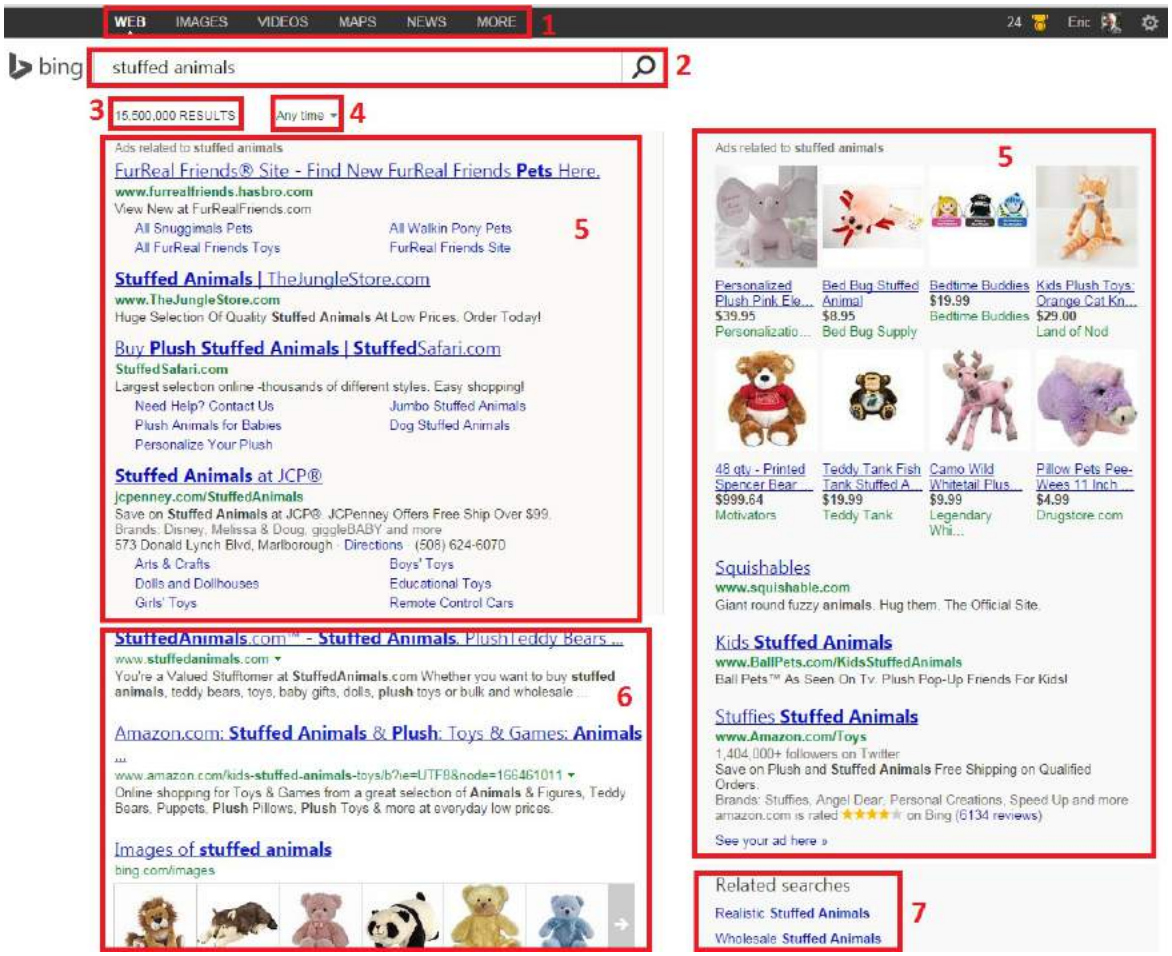


Figure 2-3. Layout of Bing search results

The sections in Bing's search results are as follows:

- Vertical navigation (1)
- Search query box (2)
- Results information (3)
- Time-based refinement options (4)
- PPC advertising (5)
- Natural/organic/algorithmic results (6)

- Query refinement options (7)

Each unique section represents a snippet of information provided by the engines. Here are the definitions of what each piece is meant to provide:

Vertical navigation

Each engine offers the option to search different verticals, such as images, news, video, or maps. Following these links will result in a query with a more limited index. In Figure 2-3, for example, you might be able to see news items about stuffed animals or videos featuring stuffed animals.

Horizontal navigation

All three engines used to have some form of horizontal navigation, but as of June 2015 only Yahoo! continues to include it.

Search query box

All of the engines show the query you've performed and allow you to edit or reenter a new query from the search results page. If you begin typing, you may notice that Google gives you a list of suggested searches below. This is the Google autocomplete suggestions feature, and it can be incredibly useful for targeting keywords. Next to the search query box, the engines also offer links to the advanced search page, the features of which we'll discuss later in the book. In addition, you will also see a microphone icon in the right of the search box that allows you to speak your query. In Google image search, this shows up as a camera icon that allows you to upload an image or get similar images back.

Results information

This section provides a small amount of meta-information about the results that you're viewing, including an estimate of the number of pages relevant to that particular query (these numbers can be, and frequently are, wildly inaccurate and should be used only as a rough comparative measure).

PPC (a.k.a. paid search) advertising

The text ads are purchased by companies that use either Google AdWords or Bing. The results are ordered by a variety of factors, including relevance (for which click-through rate, use of searched keywords in the ad, and relevance of the landing page are factors in Google) and bid amount (the ads require a maximum bid, which is then compared against other advertisers' bids).

Natural/organic/algorithmic results

These results are pulled from the search engines' primary indices of the Web and ranked in order of relevance and importance according to their complex algorithms. This area of the results is the primary focus of this section of the book.

Query refinement suggestions

Query refinements are offered by Google, Bing, and Yahoo!. The goal of these links is to let users search with a more specific and possibly more relevant query that will satisfy their intent.

In March 2009, Google enhanced the refinements by implementing Orion Technology, based on technology Google acquired in 2006. The goal of this enhancement is to provide a wider array of refinement choices. For example, a search on *principles of physics* may display refinements for the Big Bang, angular momentum, quantum physics, and special relativity.

Navigation to more advertising

Only Yahoo! shows this in the search results. Clicking on these links will bring you to additional paid search results related to the original query.

Be aware that the SERPs are always changing as the engines test new formats and layouts. Thus, the images in Figure 2-1 through Figure 2-3 may be accurate for only a few weeks or months until Google, Yahoo!, and Bing shift to new formats.

Understanding How Vertical Results Fit into the SERPs

These "standard" results, however, are certainly not all that the engines have to

offer. For many types of queries, search engines show *vertical* results, or *instant answers*, and include more than just links to other sites to help answer a user's questions. These types of results present many additional challenges and opportunities for the SEO practitioner.

Figure 2-4 shows an example of these types of results. The query in Figure 2-4 brings back a business listing showing an address and the option to get directions. This result attempts to provide the user with the answer he is seeking directly in the search results.

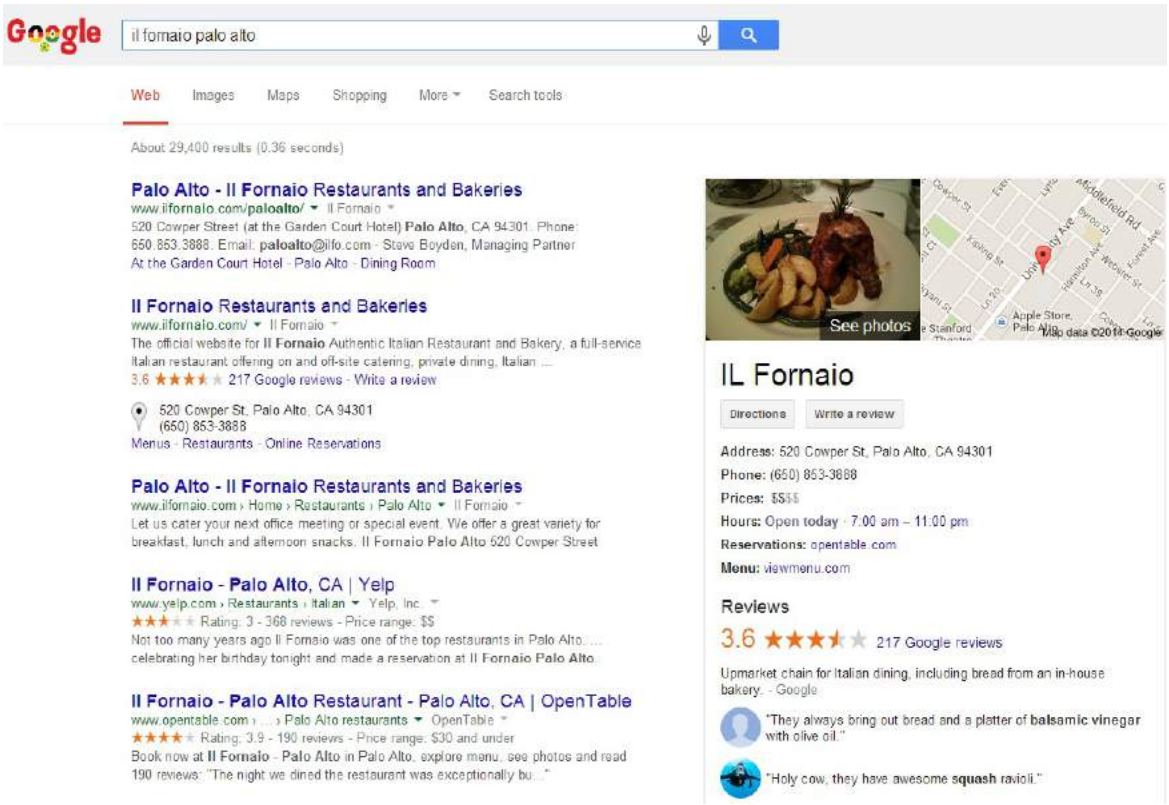


Figure 2-4. Local search result for a business

Figure 2-5 shows another example. The Google search in Figure 2-5 for weather plus a city name returns a direct answer. Once again, the user may not even need to click on a website if all she wanted to know was the temperature.

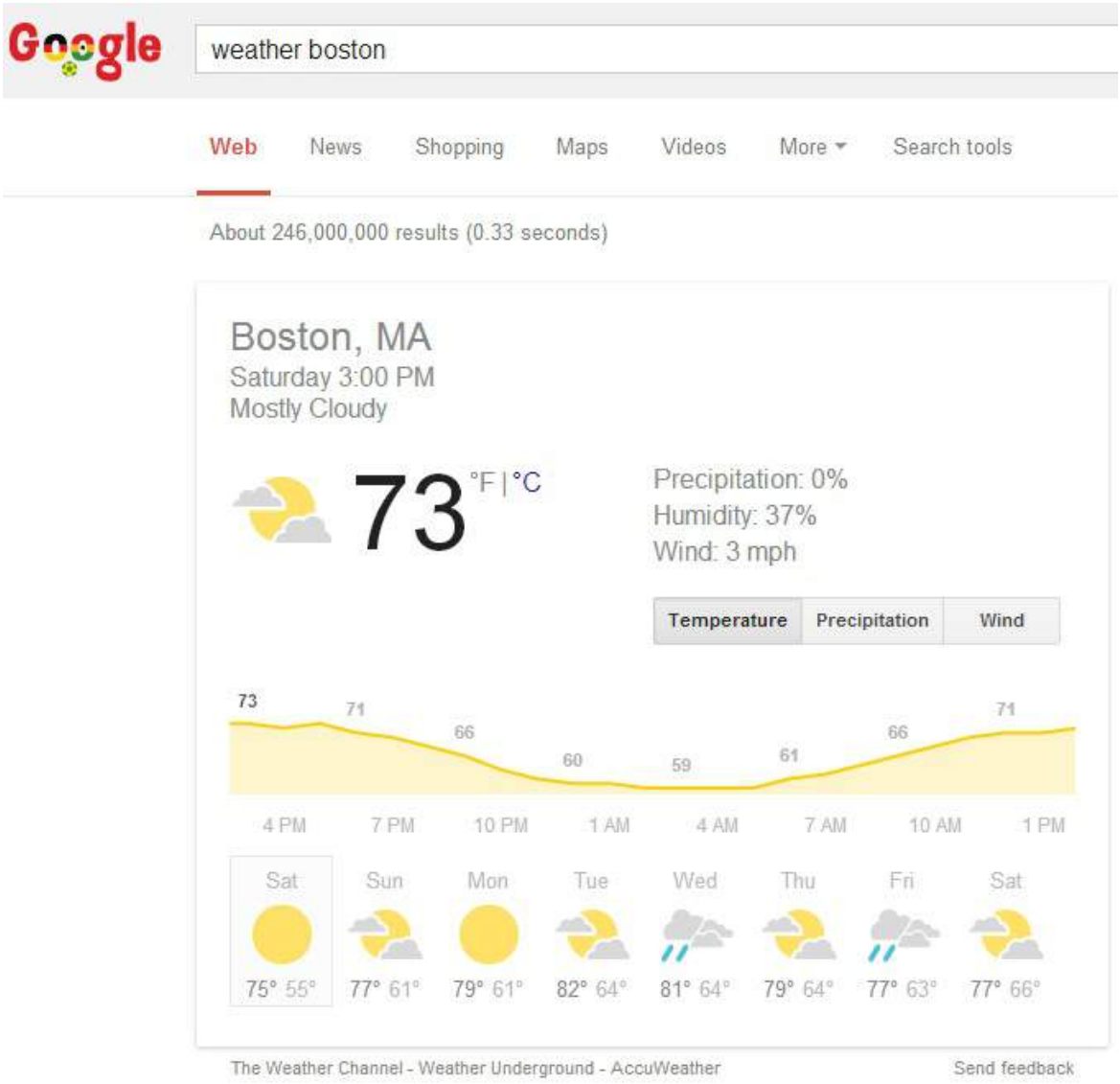


Figure 2-5. Weather search on Google

Figure 2-6 is an example of a search for a well-known painter. A Google search for the famous painter *Edward Hopper* returns image results of some of his most memorable works (shown in the lower-right of the screenshot). This example is a little different from the "instant answers" type of result shown in Figure 2-4 and Figure 2-5. If the user is interested in the first painting shown, he may well click on it to see the painting in a larger size or to get more information about it. For the SEO practitioner, getting placed in this vertical result could be a significant win.

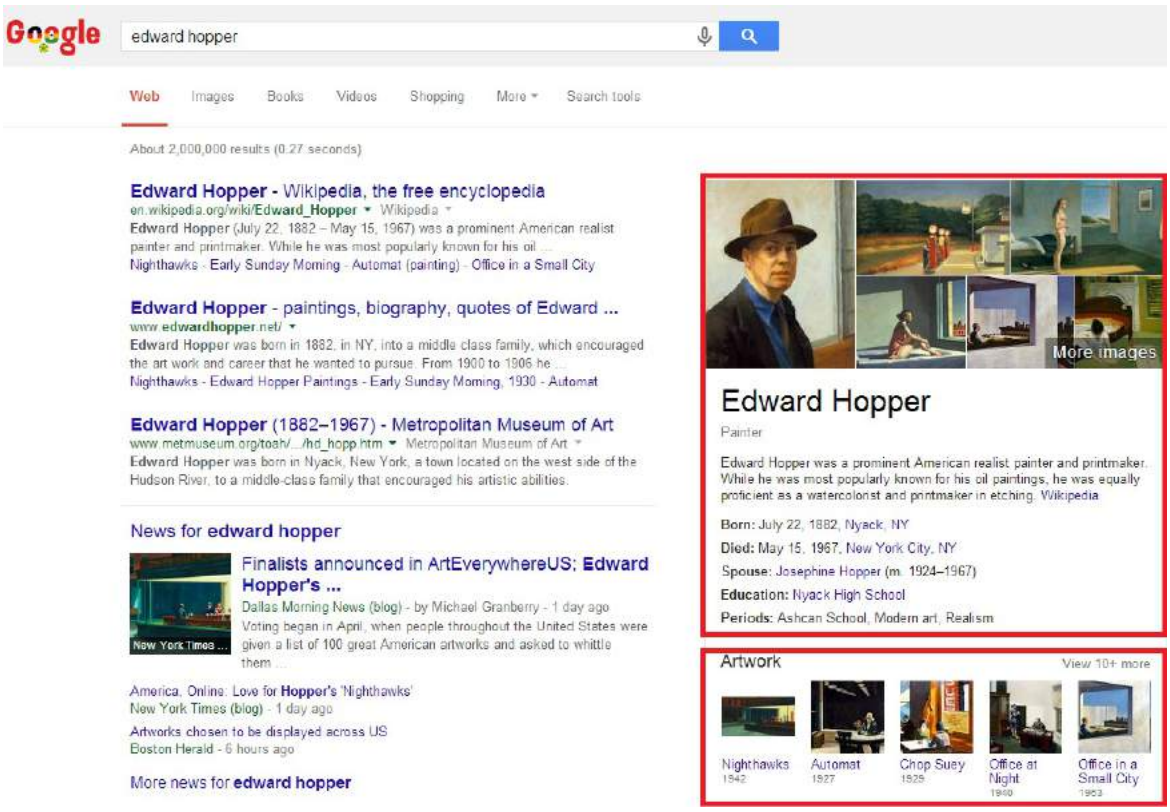


Figure 2-6. Google search on an artist's name

Figure 2-7 shows an example from Yahoo!. A query on Yahoo! for *chicago restaurants* brings back a list of popular dining establishments from Yahoo!'s local portal. High placement in these results has likely been a good thing for Lou Malnati's Pizzeria.

Figure 2-8 is an example of a celebrity search on Bing.

The results in Figure 2-8 include a series of images of the famous actor Charlie Chaplin. As a last example, Figure 2-9 is a look at the Bing search results for videos with Megan Fox.

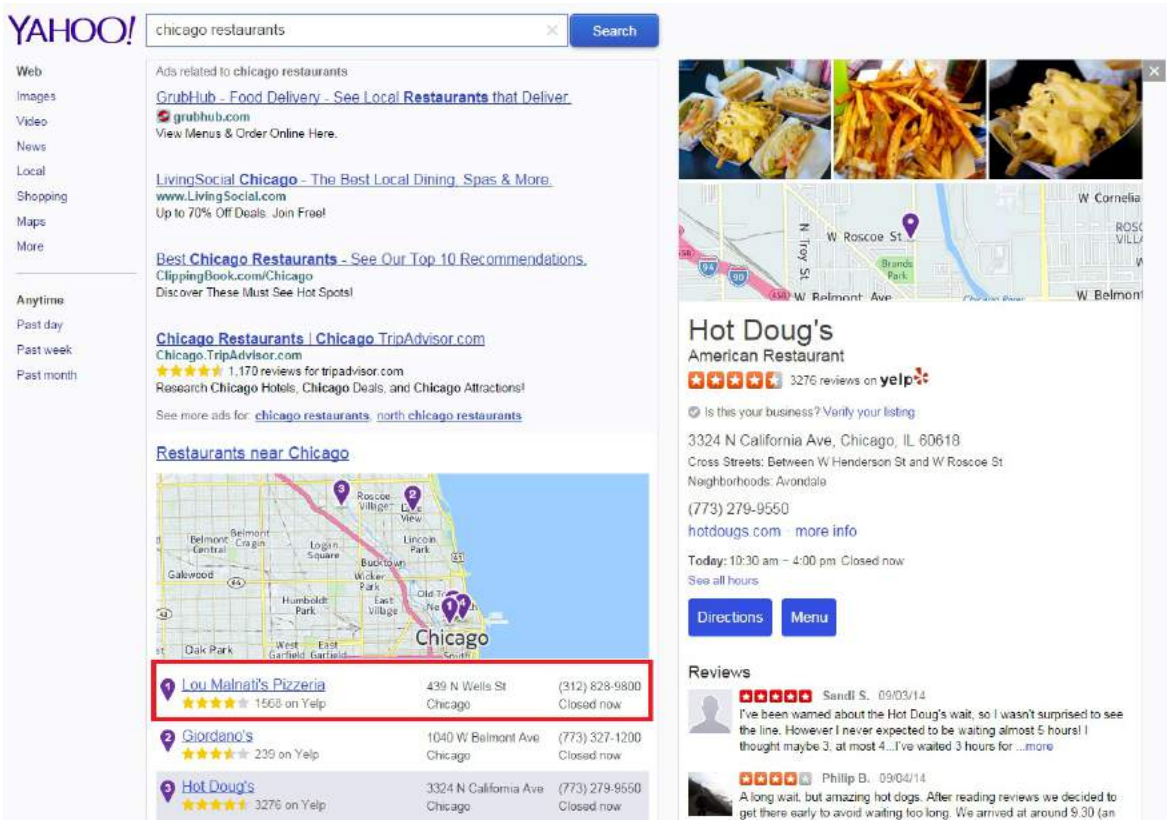


Figure 2-7. Yahoo! search for Chicago restaurants

At the top of the search results in Figure 2-9, you're provided with a series of popular videos. Click on a video in the results, and it begins playing right there in the search results.

As you can see, the vast variety of vertical integration into search results means that for many popular queries you can expect to receive significant amounts of information in the SERPs themselves. Engines are competing by providing more relevant results and more targeted responses to queries that they feel are best answered by vertical results, rather than web results.

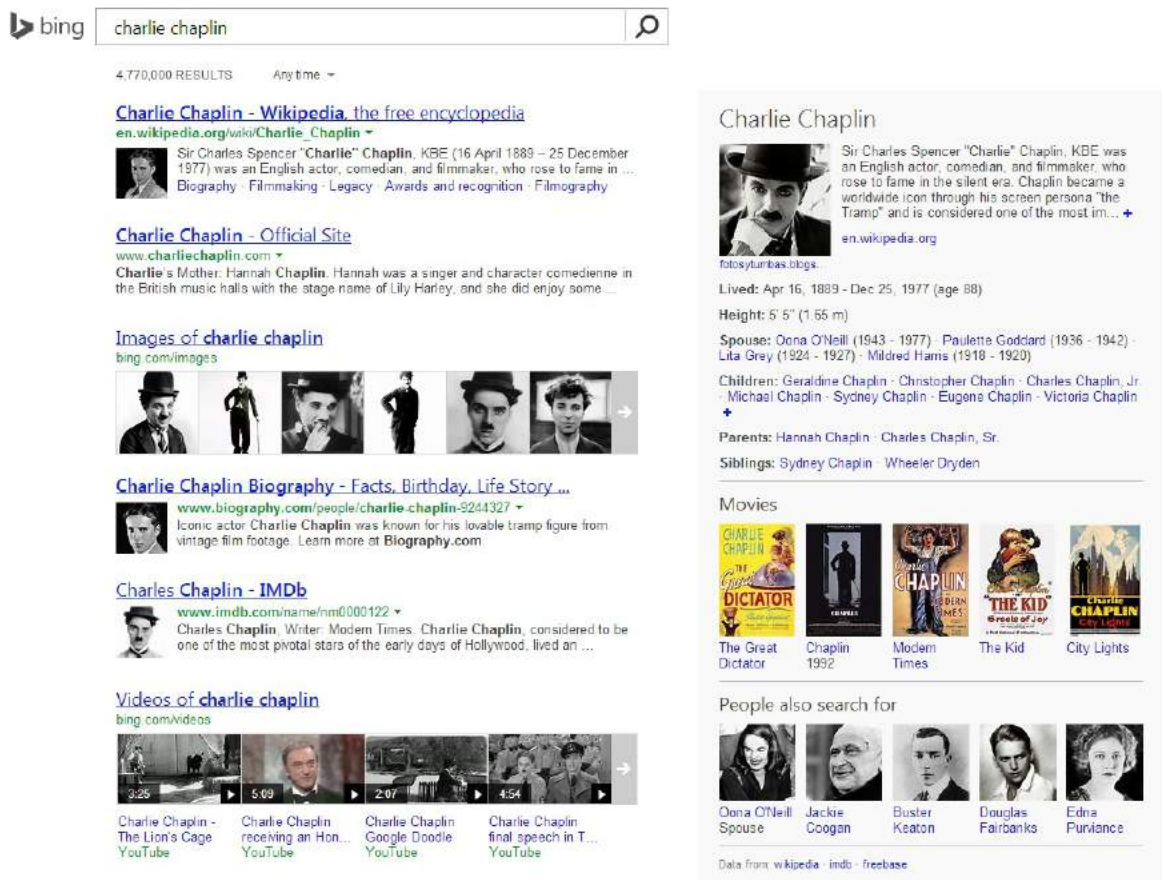


Figure 2-8. Bing result for Charlie Chaplin

As a direct consequence, site owners and web marketers must take into account how this incorporation of vertical search results may impact their rankings and traffic. For many of the searches shown in the previous figures, a high ranking—even in position #1 or #2 in the algorithmic/organic results—may not produce much traffic because of the presentation of the vertical results above them.

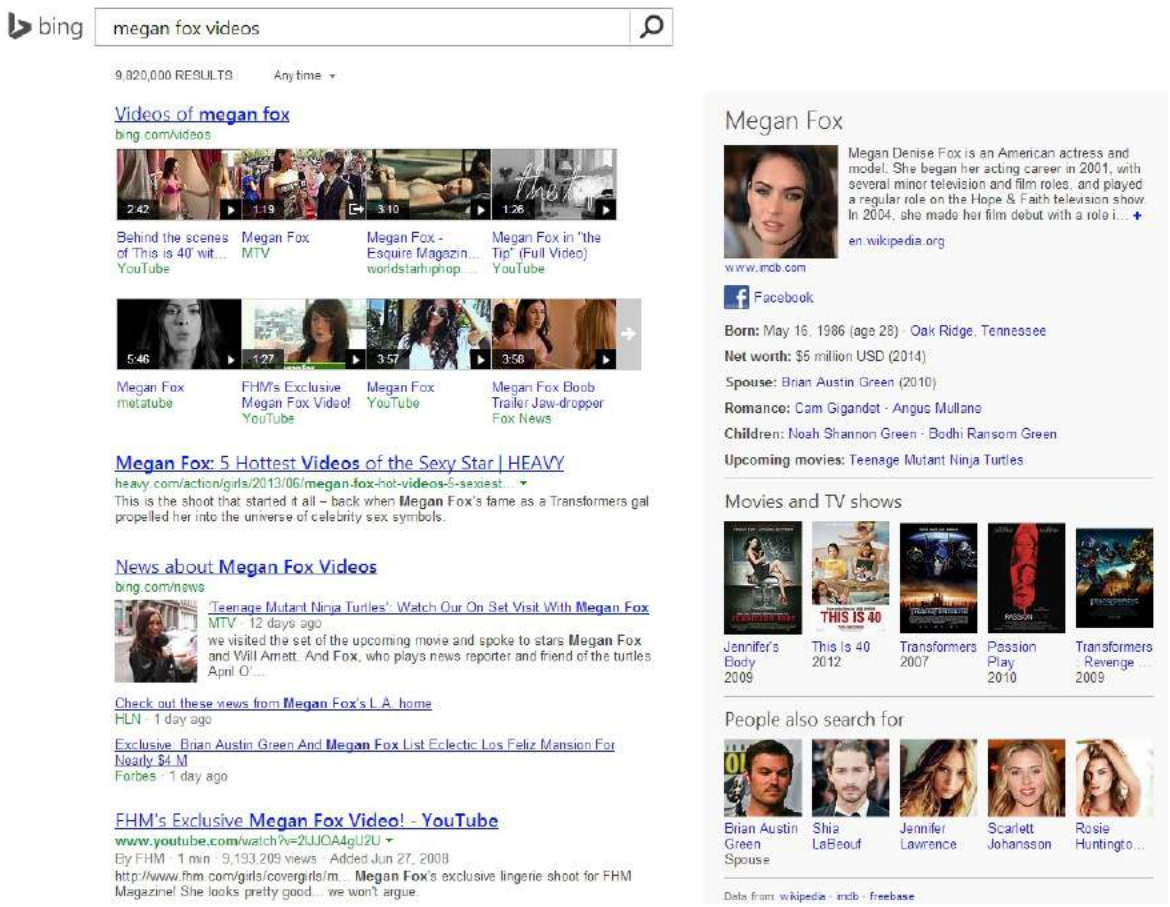


Figure 2-9. Bing result for Megan Fox videos

The vertical results also signify an opportunity, as listings are available in services from images to local search to news and products. We will cover how to get included in these results in Chapter 11.

Google's Knowledge Graph

The search engines are actively building structured databases of information that allow them to show answers to questions that are not simply links to web pages. In Figure 2-6, the information on the upper right is an example of this. Google provides direct answers in the result, including Edward Hopper's birth date, place of birth, the date and place of his death, his spouse, and more. In Figure 2-9, Bing provides similar information for Megan Fox.

Not only is additional information shown, but it is not just a data dump: it shows that the search engines are working to develop their own knowledge of the

relationships between people and things. In the case of Figure 2-6, we can see that Google understands that:

- Edward Hopper is the name of a person.
- People have dates and places of birth.
- People have dates and places of death.
- People might have spouses.

The search engines are actively mapping these types of relationships as part of their effort to offer more complete information directly in the search results themselves.

Algorithm-Based Ranking Systems: Crawling, Indexing, and Ranking

Understanding how crawling, indexing, and ranking works is useful to SEO practitioners, as it helps them determine what actions to take to meet their goals. This section primarily covers the way Google and Bing operate, and does not necessarily apply to other search engines that are popular in other countries, such as Yandex (Russia), Baidu (China), Seznam (Czech Republic), and Naver (Korea).

The search engines must execute many tasks very well to provide relevant search results. Put simplistically, you can think of these as:

- Crawling and indexing trillions of documents (pages and files) on the Web (note that they ignore pages that they consider to be “insignificant,” perhaps because the pages are perceived as adding no new value or are not referenced at all on the Web).
- Responding to user queries by providing lists of relevant pages.

In this section, we’ll walk through the basics of these functions from a nontechnical perspective. This section will start by discussing how search engines find and discover content.

Crawling and Indexing

To offer the best possible results, search engines must attempt to discover all the public pages on the World Wide Web and then present the ones that best match up with the user's search query. The first step in this process is *crawling* the Web. The search engines start with a seed set of sites that are known to be very high quality, and then visit the links on each page of those sites to discover other web pages.

The link structure of the Web serves to bind together all of the pages that were made public as a result of someone linking to them. Through links, search engines' automated robots, called *crawlers* or *spiders*, can reach the many trillions of interconnected documents.

In Figure 2-10, you can see the home page of USA.gov, the official U.S. government website. The links on the page are outlined in red. Crawling this page would start with loading the page, analyzing the content, and then seeing what other pages USA.gov links to.

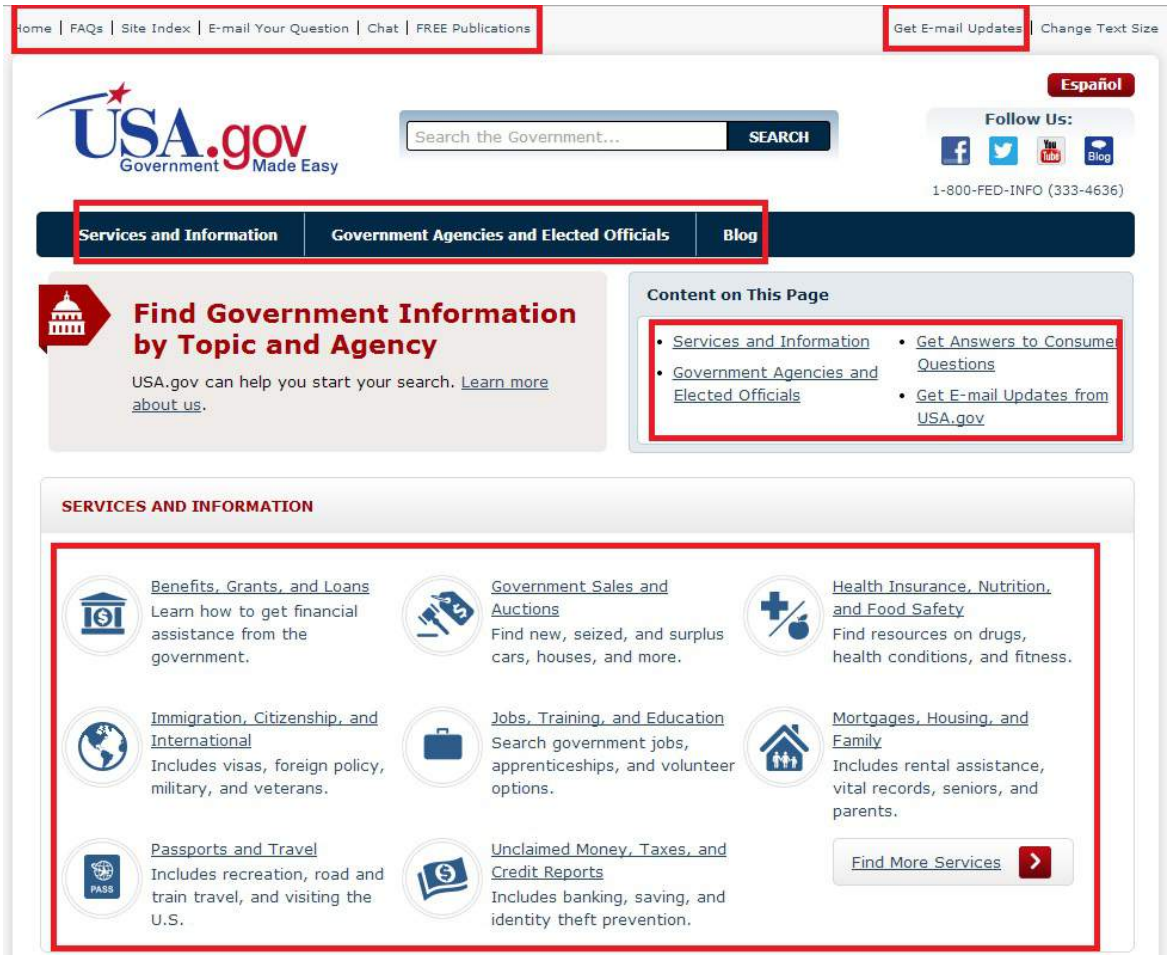


Figure 2-10. Crawling the U.S. government website

The search engine would then load those other pages and analyze that content as well. This process repeats over and over again until the crawling process is complete. This process is enormously complex, as the Web is a large and complex place.

NOTE

Search engines do not attempt to crawl the entire Web every day. In fact, they may become aware of pages that they choose not to crawl because those pages are not likely to be important enough to return in a search result. We will discuss the role of *importance* in "Retrieval and Ranking".

The first step in this process is to build an *index* of terms. This is a massive

database that catalogs all the significant terms on each page crawled by the search engine.

A lot of other data is also recorded, such as a map of all the pages that each page links to, the clickable text of those links (known as the *anchor text*), whether or not those links are considered ads, and more.

To accomplish the monumental task of holding data on hundreds of trillions of pages that can be accessed in a fraction of a second, the search engines have constructed massive data centers to deal with all this data.

One key concept in building a search engine is deciding where to begin a crawl of the Web. Although you could theoretically start from many different places on the Web, you would ideally begin your crawl with a trusted seed set of websites.

Starting with a known, trusted set of websites enables search engines to measure how much they trust the other websites that they find through the crawling process. We will discuss the role of trust in search algorithms in more detail in "How Links Historically Influenced Search Engine Rankings".

Retrieval and Ranking

For most searchers, the quest for an answer begins as shown in Figure 2-11.

Figure 2-11. Start of a user's search quest

The next step in this quest occurs when the search engine returns a list of relevant pages on the Web in the order it believes is most likely to satisfy the user. This process requires the search engines to scour their corpus of hundreds of billions of documents and do two things: first, return only the results that are related to the searcher's query; and second, rank the results in order of perceived importance (taking into account the trust and authority associated with the site). It is both relevance and importance that the process of SEO is meant to influence.

Relevance is the degree to which the content of the documents returned in a search matches the user's query intention and terms. The relevance of a document increases if the page contains terms relevant to the phrase queried by

the user, or if links to the page come from relevant pages and use relevant anchor text.

You can think of relevance as the first step to being “in the game.” If you are not relevant to a query, the search engine does not consider you for inclusion in the search results for that query. We will discuss how relevance is determined in more detail in “Determining Searcher Intent and Delivering Relevant, Fresh Content”.

Importance refers to the relative importance, measured via *citation* (the act of one work referencing another, as often occurs in academic and business documents), of a given document that matches the user’s query. The importance of a given document increases with every other document that references it. In today’s online environment, citations can come in the form of links to the document or references to it on social media sites. Determining how to weight these signals is known as *citation analysis*.

You can think of importance as a way to determine which page, from a group of equally relevant pages, shows up first in the search results, which is second, and so forth. The relative authority of the site, and the trust the search engine has in it, are significant parts of this determination. Of course, the equation is a bit more complex than this, and not all pages are equally relevant. Ultimately, it is the combination of relevance and importance that determines the ranking order.

So, when you see a search results page such as the one shown in Figure 2-12, you can surmise that the search engine (in this case, Bing) believes the Superhero Stamps page on eBay has the highest combined score for relevance and importance for the query *marvel superhero stamps*.

Figure 2-12. Sample search result for "marvel superhero stamps"

Importance and relevance aren't determined manually (those trillions of man-hours would require Earth's entire population as a workforce). Instead, the engines craft careful, mathematical equations—*algorithms*—to sort the wheat from the chaff and then rank the wheat in order of quality. These algorithms often comprise hundreds of components. In the search marketing field, they are often referred to as *ranking factors* or *algorithmic ranking criteria*.

We discuss ranking factors or signals (*signals* is the term Google prefers) in more detail in "Analyzing Ranking Factors".

Evaluating Content on a Web Page

Search engines place a lot of weight on the content of each web page. After all, it is this content that defines what a page is about, and the search engines do a detailed analysis of each web page they find during their crawl to help make that determination.

You can think of this as the search engine performing a detailed analysis of all the words and phrases that appear on a web page, and then building a map of that data for it to consider showing your page in the results when a user enters a related search query. This map, often referred to as a *semantic map*, seeks to

define the relationships between those concepts so that the search engine can better understand how to match the right web pages with user search queries.

If there is no semantic match of the content of a web page to the query, the page has a much lower possibility of showing up. Therefore, the words you put on the page, and the “theme” of that page, play a huge role in ranking.

Figure 2-13 shows how a search engine will break up a page when it looks at it, using a page on the *Forbes* website.

The navigational elements of a web page are likely similar across the many pages of a site. These navigational elements are not ignored, and they do play an important role, but they do not help a search engine determine what the unique content is on a page. To do that, the search engine focuses on the part of Figure 2-13 that is labeled “Unique Page Content.”

Determining the unique content on a page is an important part of what the search engine does. The search engine uses its understanding of unique content to determine the types of search queries for which the web page might be relevant. Because site navigation is generally not unique to a single web page, it does not help the search engine with that task.

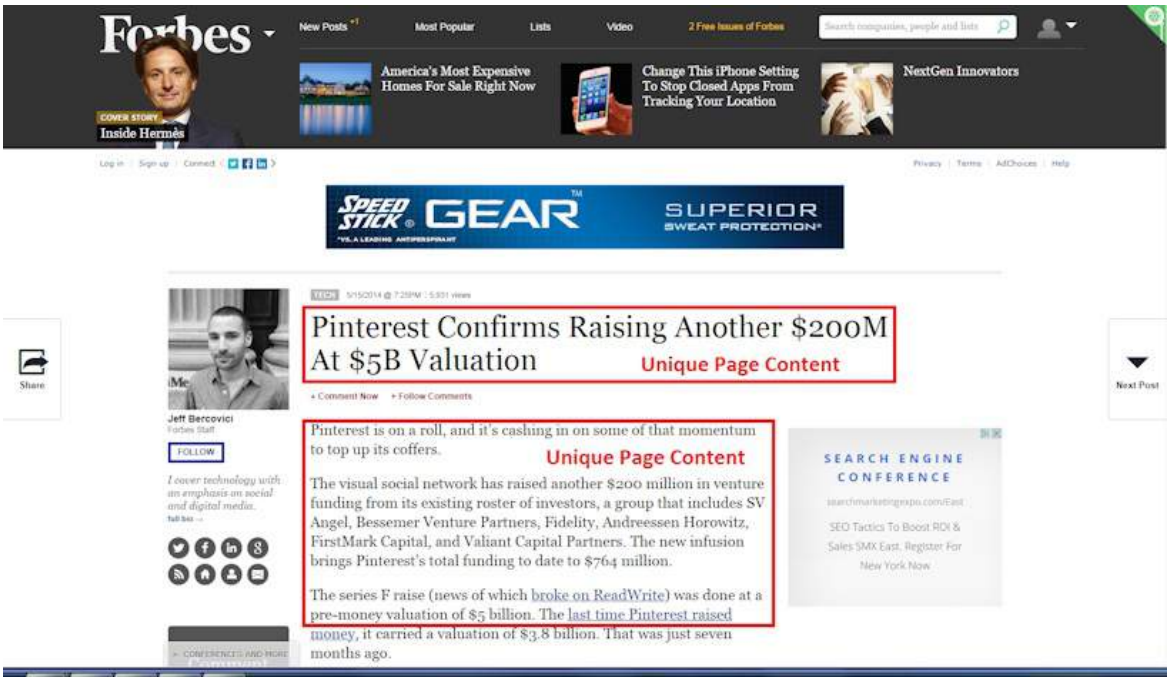


Figure 2-13. Breaking up a web page

This does not mean navigation links are not important—they most certainly are; however, they simply do not count when a search engine is trying to determine the unique content of a web page, as they are shared among many web pages.

One task the search engines face is judging the value of content. Although evaluating how the community responds to a piece of content using link analysis is part of the process, the search engines can also draw some conclusions based on what they see on the page.

For example, is the exact same content available on another website? Is the unique content the search engine can see two sentences long or 500 words long? Does the content repeat the same keywords excessively? These are a few examples of factors the search engine can evaluate when trying to determine the value of a piece of content.

Understanding What Content Search Engines Can "See" on a Web Page

Search engine crawlers and indexing programs are basically software programs. These programs are extraordinarily powerful. They crawl hundreds of trillions of web pages, analyze the content of all these pages, and analyze the way all these pages link to one another. Then they organize this into a series of databases that can respond to a user search query with a highly tuned set of results in a few tenths of a second.

This is an amazing accomplishment, but it has its limitations. Software is very mechanical, and it can understand only portions of most web pages. The search engine crawler analyzes the raw HTML form of a web page. If you want to see what this looks like, you can do so by using your browser to view the source.

Figure 2-14 shows how to do that in Chrome, and Figure 2-15 shows how to do that in Firefox. Typically you can access it most easily by right-clicking with your mouse on a web page to access a hidden menu.

Figure 2-14. Viewing source in Chrome: right-click on the web page to access the menu

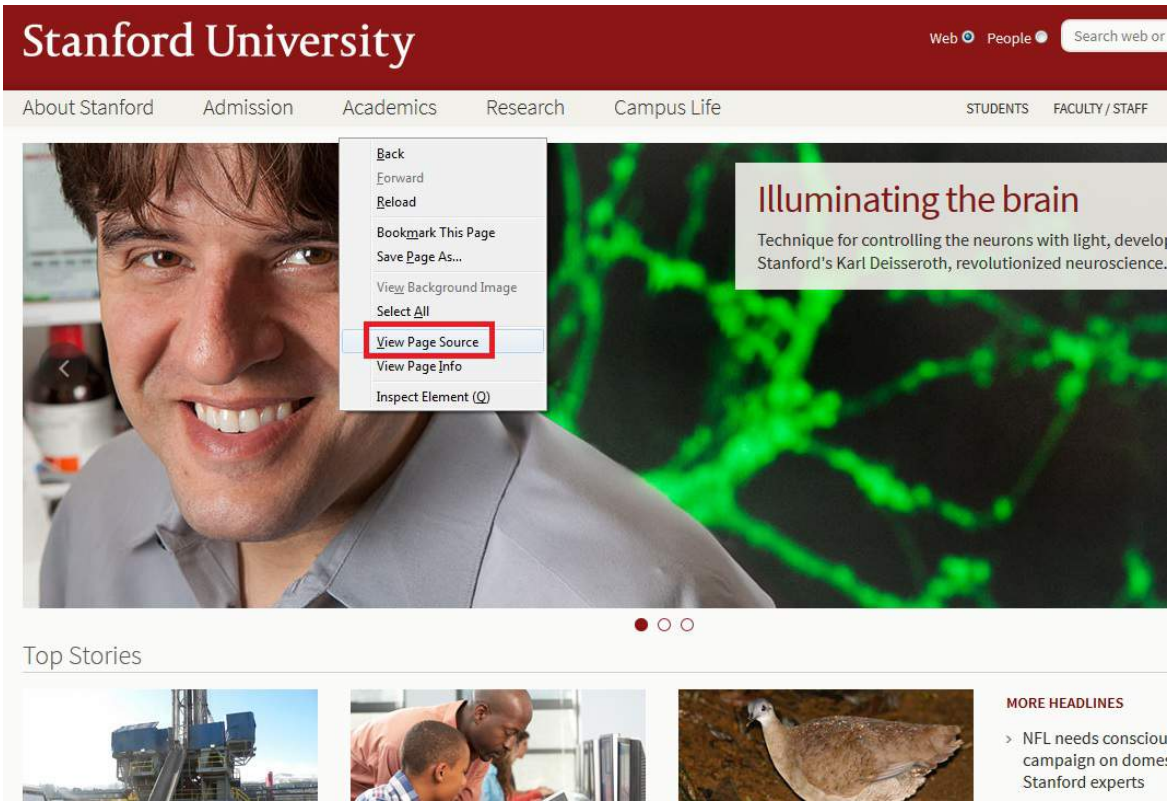


Figure 2-15. Viewing source in Firefox

There are also various in-browser web development tools (add-ons and extensions) that facilitate viewing source code in your browser of choice, as well as detecting web applications and JavaScript libraries. One of the most widely used code analysis tools is Web Developer, by Chris Pederick, available for Chrome, Firefox, and Opera. Once you view the source, you will be presented with the exact code for the web page that the web server sent to your browser. This is most of what the search engine crawler sees (the search engine also sees the HTTP headers for the page, which are status codes it receives from the web server where the page is hosted). In Some cases Google will execute JavaScript on the page as well. For more on how they do this, please refer to Chapter 6. When trying to analyze the user-visible content on a web page, search engines largely ignore code related to the navigation and display of the page, such as that shown in Figure 2-16, as it has nothing to do with the page's content.

```
<body class="home home-index  layout-default">
  <div id="root" class="no-overflow">
    <nav class="globalnav-container default">
<ul class="globalnav">
    <li class="active">
    <a href="http://moz.com" onclick="_gaq.push(['_trackEvent', 'multiproduct_nav', 'button', 'moz_com'">Moz.com</a>
  </li>
  <li class="">
    <a href="http://moz.com/pro" onclick="_gaq.push(['_trackEvent', 'multiproduct_nav', 'button', 'moz_pro'">Moz Pro</a>
  </li>
  <li class="">
    <a href="https://moz.com/local" onclick="_gaq.push(['_trackEvent', 'multiproduct_nav', 'button', 'moz_local'">Moz Local</a>
  </li>
</ul>
</nav><header class="background-blue-canvas masthead-container">
  <div class="container masthead">
    <div class="row">
      <div id="logo" class="span3 center">
        <a href="/">
          <svg width="120" height="35" alt="Moz">
<path fill="#fff" transform="scale(1)"
  d="M102.3,22.8L120,4.2H83.6c-0.7,0-1.3,0.6-1.3,1.310,3.8C78.9,3.7,72.3,0,64.6,0c-10,0-18.3,6.4-19.8,14.7
    V4.2h-6.3c-1.5,0-2.9,0.7-3.8,1.7L22.4,19.5L10.1,5.9c-1-1-2.3-1.7-3.8-1.7H0v27.2h7.4c0.7,0,1.3-0.6,1.3-1.3h0c0,0,0-0.1,0-0.1
    V17.3113.7,15.3113.7-15.310,12.8c0,0.7,0.6,1.3,1.3,1.3h7.4V20.3c1.5,8.3,9.8,14.7,19.8,14.7c11.1,0,20.1-7.9,20.1-17.5
    c0-1.6-0.3-3.2-0.8-4.8h16L82.4,31.4h1.8h34.4c0,0,0,0,0,0c0,0,0,0,0,0h0c0.7,0,1.3-0.6,1.3-1.3h0v-7.3H102.3z M64.6,26.2
    c-5.5,0-10-3.9-10-8.7c0-4.8,4.5-8.7,10-8.7c5.5,0,10,3.9,10,8.7C74.6,22.3,70.1,26.2,64.6,26.2z" />
</svg>
        </a>
      </div>
```

Figure 2-16. Sample web page source code

The search engine crawler is most interested in the HTML text on the page. Figure 2-17 is an example of HTML text for the Moz home page.

```
          <h2 class="h3 top0 bottom1"><a href="http://moz.com/blog/tips-and-tactics-for-amplifying-your-content-whiteboard-friday
class="slate">Tips and Tactics for Amplifying Your Content - Whiteboard Friday</a></h2>
          <p class="bottom1 small">
                <time datetime="2014-06-20 00:12:00">June 20th, 2014</time>

          - Posted by <a href="http://moz.com/community/users/17229">Ben Lloyd</a> to
                      <a href="http://moz.com/blog/category/content">Content</a> and
              <a href="http://moz.com/blog/category/whiteboard-friday">Whiteboard Friday</a>
                </p>
          <p>Content marketing should never be approached with a "set it and forget it" mentality. It needs to be structured and
shared in the right ways, and in today's Whiteboard Friday, the folks from Add3 are here to show you what that means.</p>
      <div class="pull-right">
      <a href="http://moz.com/blog/tips-and-tactics-for-amplifying-your-content-whiteboard-friday">Read Full Entry</a>
      |
      <span class="comment-count">
        <a href="http://moz.com/blog/tips-and-tactics-for-amplifying-your-content-whiteboard-friday#comments">
          <i class="icon icon-speech"></i>
          32 comments        </a>
```

Figure 2-17. Sample HTML text in the source code showing real content

Although Figure 2-17 still shows some HTML encoding, you can see the “regular” text clearly in the code. This is the unique content that the crawler is looking to find.

In addition, search engines read a few other elements. One of these is the page title. The page title is one of the most important factors in ranking a given web page. It is the text that shows in the browser’s title bar (above the browser menu and the address bar).

Figure 2-18 shows the code that the crawler sees, using Trip Advisor as an

example.

The first highlighted area in Figure 2-18 is for the `<title>` tag. The `<title>` tag is also often (but not always) used as the title of your listing in search engine results (see Figure 2-19).

```
<title>Reviews of Hotels, Flights and Vacation Rentals - TripAdvisor</title>
<meta http-equiv="pragma" content="no-cache"/>
<meta http-equiv="cache-control" content="no-cache,must-revalidate"/>
<meta http-equiv="expires" content="0"/>
<meta property="og:image" content="http://c1.tacdn.com/img2/postimg.jpg" height="150px"
width="150px"/>
<meta name="keywords" content="vacation, vacations, vacation packages, vacation package,
travel package, travel packages, travel, planning, hotel, hotels, motel, bed and
breakfast, inn, guidebook, review, reviews, popular, plan, airfare, cheap, discount, map,
maps, golf, ski, articles, attractions, advice, restaurants"/>
<meta name="description" content="TripAdvisor - Unbiased hotel reviews, photos and travel
advice for hotels and vacations - Compare prices with just one click."/>
<link rel="alternate" hreflang="en" href="http://www.tripadvisor.com/"/>
```

Figure 2-18. Meta tags in HTML source

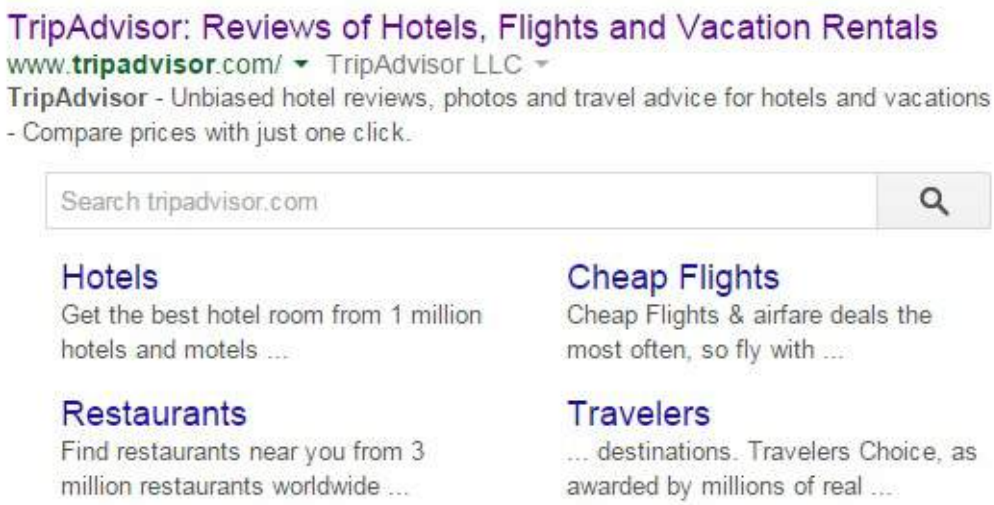


Figure 2-19. Search result showing the title tag

In addition to page titles, search engines previously used the meta `keywords` tag. This is a list of keywords that you wish to have associated with the page. *Spammers* (people who attempt to manipulate search engine results in violation of the search engine guidelines) ruined the SEO value of this tag many years ago, so its value is now negligible, as search engines don't use it anymore.

Spending time on meta keywords is not recommended because of the lack of SEO benefit.

The second highlighted area in Figure 2-18 shows an example of a meta `keywords` tag.

Search engines also read the meta `description` tag (the third highlighted area in the HTML source in Figure 2-18). However, the content of a meta `description` tag is not directly used by search engines in their ranking algorithms.[2]

Nonetheless, the `meta description` tag plays a key role, as search engines often use it as a part or all of the description for your page in search results. Therefore, a well-written meta description can have a significant influence on how many clicks you get on your search listing, and the click-through rate on your search listing can impact your ranking. As a result, time spent on meta descriptions is quite valuable. Figure 2-20 uses a search on *trip advisor* to show an example of the meta `description` tag being used as a description in the search results.

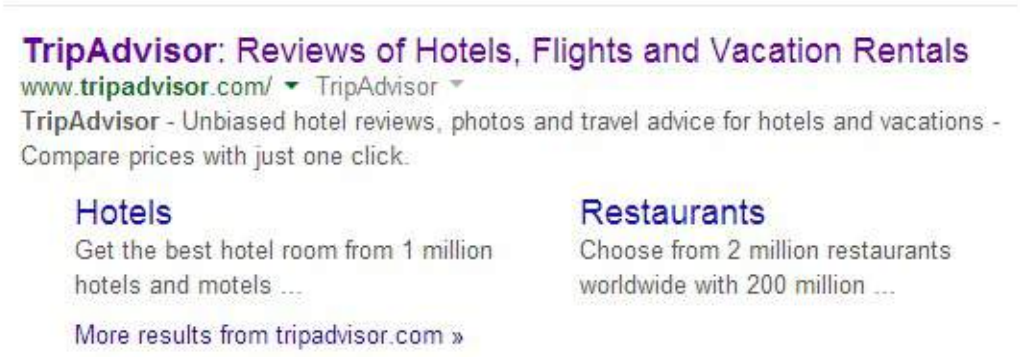


Figure 2-20. Meta description used in search results

NOTE

The user's keywords are typically shown in boldface when they appear in the search results (sometimes close synonyms are shown in boldface as well). As an example, in Figure 2-20, *TripAdvisor* is in boldface at the beginning of the description. This is called keywords in context (KWIC).

A fourth element that search engines read is the `alt` attribute for images. The

`alt` attribute was originally intended to allow something to be rendered for audiences who cannot view the images, primarily:

- Vision-impaired people who do not have the option of viewing the images.
- People who turn off images for faster surfing. This is generally an issue only for those who do not have a broadband connection.

Support for the vision-impaired remains a major reason for using the `alt` attribute. You can read more about this by visiting the W3C's Web Accessibility Initiative page.

Search engines also read the text contained in the `alt` attribute of an image tag (`<img>`). An image tag is an element that is used to tell a web page to display an image.

Another element that search engines read is the `<noscript>` tag. Historically, the ability of search engines to read JavaScript was quite limited, but this has been changing over time and Google says they execute more JavaScript today.[3] However, a small percentage of users do not allow JavaScript to run when they load a web page (our experience is that it is about 2%). For those users, nothing would be shown to them where the JavaScript is on the web page, unless the page contains a `<noscript>` tag.

Here is a very simple JavaScript example that demonstrates this:

```
<script type="text/javascript">
document.write("It is a Small World After All!")
</script>
<noscript>Your browser does not support JavaScript!</noscript>
```

The `<noscript>` portion of this is `Your browser does not support JavaScript!`. In this example, you could also choose to make the `<noscript>` tag contain the text `"It is a Small World After All!"`. The `<noscript>` tag should be used only to represent the content of the JavaScript.

What search engines cannot see

It is also worthwhile to review the types of content that search engines cannot "see" in the human sense.

For instance, although search engines are able to detect that you are displaying an image, they have little idea what the image is a picture of, except for whatever information you provide in the `alt` attribute, as discussed earlier. They can recognize only some very basic types of information within images, such as the presence of a face, or whether images have pornographic content by how much flesh tone they contain. A search engine cannot easily tell whether an image is a picture of Bart Simpson, a boat, a house, or a tornado. In addition, search engines typically don't recognize any text rendered in the image.

The reality is that the search engines have the technology to handle these types of tasks to some degree. For example, you can take a picture of the Taj Mahal and drag it into the search box in Google image search, and the search engine will recognize it. However, because of the processing power required for image recognition, search engines do not currently try to recognize all of the images they encounter across the Web.

Search engines are also experimenting with technology to use optical character recognition (OCR) to extract text from images, but it is not yet in general use within search. The main problem with applying OCR and image processing technology is that it's very computationally intensive, and not practical to apply at the scale of the Web.

In addition, conventional SEO wisdom has always held that the search engines cannot read Flash files, but this is a little overstated. Search engines have been extracting some information from Flash for years, as indicated by this Google announcement in 2008. However, the bottom line is that it's not easy for search engines to determine what is in Flash. One of the big issues is that even when search engines look inside Flash, they are still looking for textual content, but Flash is a pictorial medium and there is little incentive (other than the search engines) for a designer to implement text inside Flash. All the semantic clues that would be present in HTML text (such as heading tags, boldface text, etc.) are missing too, even when HTML is used in conjunction with Flash.

A third type of content that search engines cannot see is the pictorial aspects of anything contained in Flash, so this aspect of Flash behaves in the same way images do. For example, when text is converted into a vector-based outline (i.e., rendered graphically), the textual information that search engines can read is lost. Chapter 6 discusses methods for optimizing Flash.

Audio and video files are also not easy for search engines to read. As with images, the data is not easy to parse. There are a few exceptions where the search engines can extract some limited data, such as ID3 tags within MP3 files, or enhanced podcasts in AAC format with textual “show notes,” images, and chapter markers embedded. Ultimately, though, search engines cannot distinguish a video of a soccer game from a video of a forest fire.

Search engines also cannot read any content contained within a program. The search engine really needs to find text that is readable by human eyes looking at the source code of a web page, as outlined earlier. It does not help if you can see it when the browser loads a web page—it has to be visible and readable in the source code for that page.

One example of a technology that can present significant human-readable content that search engines cannot see is AJAX. AJAX is a JavaScript-based method for dynamically rendering content on a web page after retrieving the data from a database, without having to refresh the entire page. This is often used in tools where a visitor to a site can provide some input and the AJAX tool then retrieves and renders the correct content.

The problem arises because the content is retrieved by a script running on the client computer (the user’s machine) only after receiving some input from the user. This can result in many potentially different outputs. In addition, until that input is received, the content is not present in the HTML of the page, so the search engines cannot easily see it.

Similar problems arise with other forms of JavaScript that don’t render the content in the HTML until a user action is taken. New forms of JavaScript, such as AngularJS, make this even more challenging for search engines. For more information on SEO for single-page web applications, please see “Single-Page Applications”.[4]

As of HTML 5, a construct known as the embed tag (`<embed>`) was created to allow the incorporation of *plug-ins* into an HTML page. Plug-ins are programs located on the user’s computer, not on the web server of your website. The embed tag is often used to incorporate movies or audio files into a web page; it tells the plug-in where it should look to find the data file to use. Content included through plug-ins may or may not be invisible to search engines.

Frames and *iframes* are methods for incorporating the content from another web

page into your web page. Iframes are more commonly used than frames to incorporate content from another website. You can execute an iframe quite simply with code that looks like this:

```
<iframe src ="http://accounting.careerbuilder.com" width="100%" height="300">
  <p>Your browser does not support iframes.</p>
</iframe>
```

Frames are typically used to subdivide the content of a publisher's website, but they can be used to bring in content from other websites, as in *http://accounting.careerbuilder.com* on the *Chicago Tribune* website, shown in Figure 2-21.

Figure 2-21 is an example of something that works well to pull in content (provided you have permission to do so) from another site and place it on your own. However, the search engines recognize an iframe or a frame used to pull in another site's content for what it is, and therefore may ignore that content. In other words, they don't consider content pulled in from another site as part of the unique content of your web page.

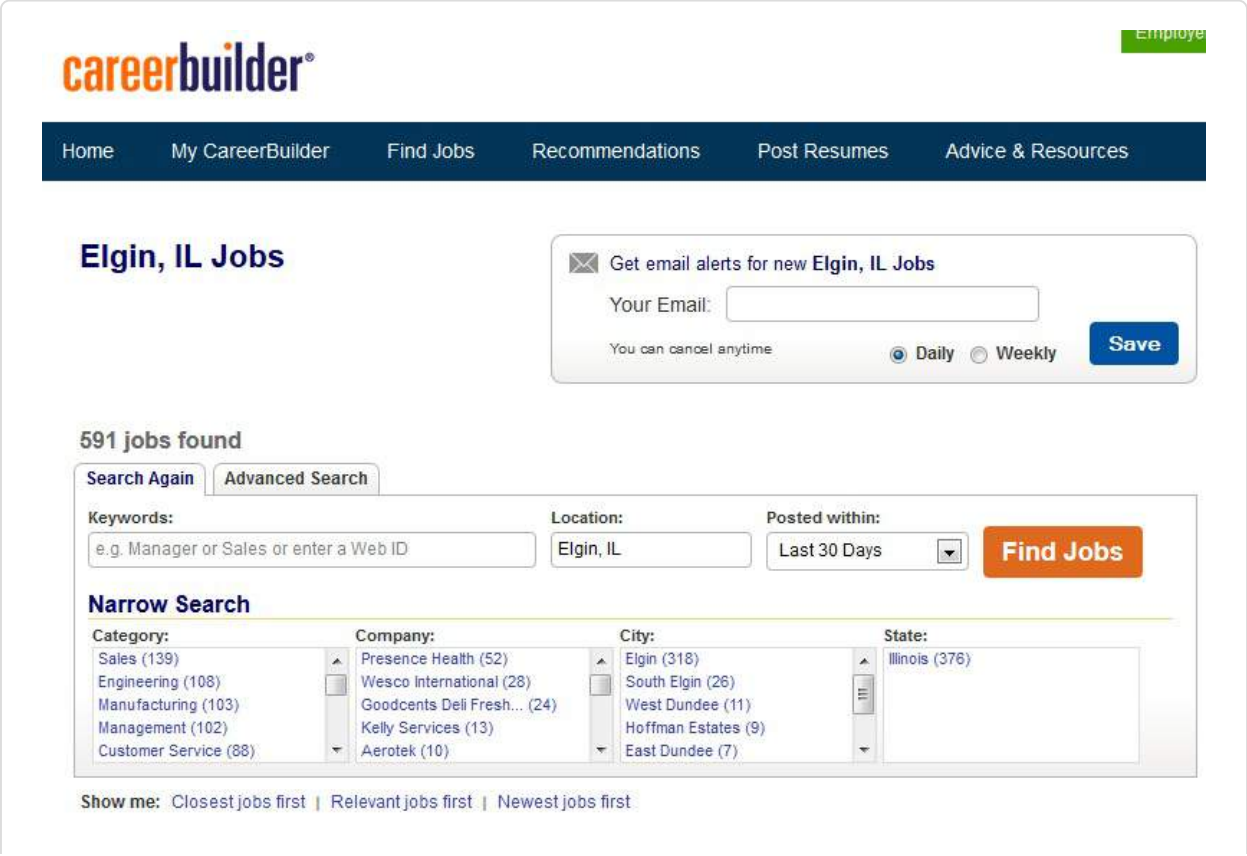


Figure 2-21. Framed page rendered in a browser

Determining Searcher Intent and Delivering Relevant, Fresh Content

Modern commercial search engines rely on the science of *information retrieval* (IR). This science has existed since the middle of the 20th century, when retrieval systems powered computers in libraries, research facilities, and government labs. Early in the development of search systems, IR scientists realized that two critical components comprised the majority of search functionality: relevance and importance (which we defined earlier in this chapter). To measure these factors, search engines perform document analysis (including semantic analysis of concepts across documents) and link (or citation) analysis.

Document Analysis and Semantic Connectivity

In *document analysis*, search engines look at whether they find the search terms in important areas of the document—the title, the metadata, the heading tags, and the body of the text. They also attempt to automatically measure the quality of the document based on document analysis, as well as many other factors.

Reliance on document analysis alone is not enough for today's search engines, so they also look at semantic connectivity. *Semantic connectivity* refers to words or phrases that are commonly associated with one another. For example, if you see the word *aloha*, you associate it with Hawaii, not Florida. Search engines actively build their own thesaurus and dictionary to help them determine how certain terms and topics are related. By simply scanning their massive databases of content on the Web, they can use fuzzy set theory and certain equations to connect terms and start to understand web pages and sites more like a human does.

The professional SEO practitioner does not necessarily need to use semantic connectivity measurement tools to optimize websites, but for those advanced practitioners who seek every advantage, semantic connectivity measurements can help in each of the following sectors:

- Measuring which keyword phrases to target
- Measuring which keyword phrases to include on a page about a certain topic
- Measuring the relationships of text on other high-ranking sites and pages
- Finding pages that provide “relevant” themed links

Although the source for this material is highly technical, SEO specialists need only know the principles to obtain valuable information. It is important to keep in mind that although the world of IR has hundreds of technical and often difficult-to-comprehend terms, these can be broken down and understood even by an SEO novice.

Common types of searches in the IR field include:

Proximity searches

A proximity search uses the order of the search phrase to find related documents. For example, when you search for *“sweet German mustard”* you are specifying only a precise proximity match. If the quotes are removed, the proximity of the search terms still matters to the search engine, but it will now show documents that don’t exactly match the order of the search phrase, such as *Sweet Mustard—German*.

Fuzzy logic

Fuzzy logic technically refers to logic that is not categorically true or false. A common example is whether a day is sunny (i.e., is 50% cloud cover a sunny day?). In search, fuzzy logic is often used for misspellings.

Boolean searches

These are searches that use Boolean terms such as AND, OR, and NOT. This type of logic is used to expand or restrict which documents are returned in a search.

Term weighting

Term weighting refers to the importance of a particular search term to the

query. The idea is to weight particular terms more heavily than others to produce superior search results. For example, the appearance of the word *the* in a query will receive very little weight in selecting the results because it appears in nearly all English language documents. There is nothing unique about it, and it does not help in document selection.

IR models (search engines) use fuzzy set theory (an offshoot of fuzzy logic created by Dr. Lotfi Zadeh in 1969) to discover the semantic connectivity between two words. Rather than using a thesaurus or dictionary to try to reason whether two words are related to each other, an IR system can use its massive database of content to puzzle out the relationships.

Although this process may sound complicated, the foundations are simple. Search engines need to rely on machine logic (true/false, yes/no, etc.). Machine logic has some advantages over humans, but it doesn't have a way of thinking like humans, and concepts that are intuitive to humans can be quite hard for a computer to understand. For example, oranges and bananas are both fruits, but oranges and bananas are not both round. To a human this is intuitive.

For a machine to understand this concept and pick up on others like it, semantic connectivity can be the key. The massive human knowledge on the Web can be captured in the system's index and analyzed to artificially create the relationships humans have made. Thus, a machine knows an orange is round and a banana is not by scanning thousands of occurrences of the words *banana* and *orange* in its index and noting that *round* and *banana* do not have great concurrence, while *orange* and *round* do.

This is how the use of fuzzy logic comes into play, and the use of fuzzy set theory helps the computer to understand how terms are related simply by measuring how often and in what context they are used together.

For example, a search engine would recognize that *trips* to the *zoo* often include *viewing wildlife* and *animals*, possibly as part of a *tour*.

To see this in action, conduct a search on Google for *zoo trips*. Note that the boldface words that are returned match the terms that are italicized in the preceding paragraph. Google is setting "related" terms in boldface and recognizing which terms frequently occur concurrently (together, on the same page, or in close proximity) in their indexes.

Search companies have been investing in these types of technologies for many years. In September 2013, Google quietly let the world know that it had rewritten its search engine and given it the name “Hummingbird”. This rewrite was in large part done to enable a whole new set of capabilities for recognizing the relationships between things.

For example, if you use Google’s voice search (click on the microphone icon at the right of the search box on Google.com) and ask it “Who is Tom Brady?” it will answer that question for you with a search result, but then use audio to tell you that he is an “American football quarterback for the New England Patriots of the National Football League.”

This shows that Google understands many aspects of Tom Brady. For example:

- He has an occupation: quarterback, playing American football (as distinct from the way the term *football* is used outside of the United States and Canada).
- He plays on a team: the New England Patriots.
- The New England Patriots belong to a league: the NFL.

This is far more sophisticated than search was in 2012. You can take this much further. For example, if you now use the voice search feature to ask “Who is his wife?” it will answer that question too (see Figure 2-22).

Figure 2-22. Gisele Bündchen is Tom Brady's wife

Notice too that in the second query we did not repeat Tom Brady's name, and that Google remembered the context of the conversation, in that "his" refers to Tom Brady. You can continue with a question such as "Does he have children?" and Google will answer that as well.

For SEO purposes, this usage opens our eyes to realizing how search engines recognize the connections between words, phrases, and ideas on the Web. As semantic connectivity becomes a bigger part of search engine algorithms, you can expect greater emphasis on the theme of pages, sites, and links. It will be important going into the future to realize search engines' ability to pick up on ideas and themes and recognize content, links, and pages that don't fit well into the scheme of a website.

Content Quality and User Engagement

Search engines also attempt to measure the quality and uniqueness of a website's content. One method they may use for doing this is evaluating the document

itself. For example, if a web page has lots of spelling and grammatical errors, that can be taken as a sign that little editorial effort was put into that page.[5]

They can also analyze the *reading level* of the document. One popular formula for doing this is the Flesch-Kincaid Grade Level Readability Formula, which considers factors like the average word length and the words per sentence to determine the level of education needed to be able to understand the sentence. Imagine a scenario where the product being sold on a page is children's toys and the search engine calculates a reading level of a college senior. This could be another indicator of a poor editorial effort.

The other method that search engines can use to evaluate the quality of a web page is measuring actual user interaction. For example, if a large number of users who visit the web page after clicking on a search result immediately return to the search engine and click on the next result, that would be a strong indicator of poor quality.

Engagement with a website began to publicly emerge as a ranking factor with the release of the Panda update by Google on February 23, 2011.[6] Google has access to a large number of data sources that it can use to measure how visitors interact with your website. Just because Google has access to this data, however, does not mean that it's definitely using the data as a ranking factor. That noted, some of those sources include:

Interaction with web search results

For example, if a user clicks through on a SERP listing and comes to your site, clicks the back button, and then clicks on another result in the same set of search results, that could be seen as a negative ranking signal. Or if the results below you in the SERPs are getting clicked on more than you are, that could be seen as a negative ranking signal for you and a positive ranking signal for them. Whether search engines use this signal or not, or how much weight they might put on it, is not known.

Google Analytics

It is hard to get a firm handle on just what percentage of websites run Google Analytics. A 2008 survey of websites by Immeria.net showed their share at 59%,[7] and the Metric Mail Blog checked the top 1 million sites in Alexa and

found that about 50% of those had Google Analytics.[8] Suffice it to say that Google is able to collect detailed data about what is taking place on a large percentage of the world's websites.

This provides Google with a rich array of data on that site, including:

Bounce rate

The percentage of visitors who visit only one page on your website.

Time on site

The time spent by the user on the site. Note that Google Analytics receives information only when each page is loaded, so if you view only one page it does not know how much time you spent on that page. More precisely, then, this metric tells you the average time between the loading of the first page and the loading of the last page, but does not take into account how long visitors spent on the last page loaded.

Page views per visitor

The average number of pages viewed per visitor on your site.

Google Toolbar

It is not known how many users out there use the Google Toolbar, but we believe that it numbers in the millions. For these users, Google can track their entire web surfing behavior. Unlike Google Analytics, the Google Toolbar can measure the time from when a user first arrives on a site to the time when she loads a page from a different website. It can also get measurements of bounce rate and page views per visitor.

Google +1 button

This enables users to vote for a page on the page itself. There is currently no evidence that Google uses this as a ranking factor, but in theory, it could. You can see a lot more about this in Chapter 8.

Chrome Personal Blocklist Extension

Google offers a Chrome add-on called the Personal Blocklist Extension. This enables users of the Chrome browser to indicate a search result they don't like. This was first used by Google as a part of its Panda algorithm, which attempts to measure the quality of a piece of content. You can read more about this algorithm in Chapter 9.

Goo.gl

Google has its own URL shortener. This tool allows Google to see what content is being shared, and which content is being clicked on, even in closed environments where Google web crawlers are not allowed to go.

What matters most is how your site compares to that of your competition. If your site has better engagement metrics, this is likely to be seen as an indication of quality and can potentially boost your rankings with respect to your competitors. Little has been made public about the way search engines use these types of signals, so the preceding comments are our speculation on what Google may be doing in this area. Social and user engagement ranking factors are discussed in more detail in Chapter 8.

Link Analysis

In link analysis, search engines measure who is linking to a site or page and what they are saying about that site/page. They also have a good grasp on who is affiliated with whom (through historical link data, the site's registration records, and other sources), who is worthy of being trusted based on the authority of sites linking to them, and contextual data about the site on which the page is hosted (who links to that site, what they say about the site, etc.).

Link analysis goes much deeper than counting the number of links a web page or website has, as all links are not created equal (one link can be worth 10 million times more than another one). Links from a highly authoritative page on a highly authoritative site will count more than other links of lesser authority. A search engine can determine a website or page to be authoritative by combining an analysis of the linking patterns and semantic analysis.

For example, perhaps you are interested in sites about dog grooming. Search engines can use semantic analysis to identify the collection of web pages that

focus on the topic of dog grooming. The search engines can then determine which of these pages about dog grooming have the most links from the set of websites relevant to the topic of dog grooming. These pages are most likely more authoritative on the topic than the others.

The actual analysis is a bit more complicated than that. For example, imagine that there are five pages about dog grooming with a lot of links from pages across the Web on the topic, as follows:

- Page A has 213 topically related links.
- Page B has 192 topically related links.
- Page C has 203 topically related links.
- Page D has 113 topically related links.
- Page E has 122 topically related links.

Further, it may be that Pages A, B, D, and E all link to one another, but none of them links to Page C. In fact, Page C appears to have the great majority of its relevant links from other pages that are topically relevant but have few links to them. In this scenario, Page C may not be considered authoritative because it is not linked to by the right sites.

The concept of grouping sites based on who links to them, and whom they link to, is referred to as grouping sites by *link neighborhood*. The neighborhood you are in says something about the subject matter of your site, and the number and quality of the links you get from sites in that neighborhood say something about how important your site is to that topic.

The degree to which search engines rely on evaluating link neighborhoods is not clear, and links from irrelevant pages can still help the rankings of the target pages. Nonetheless, the basic idea remains that a link from a relevant page or site should be more valuable than a link from a nonrelevant page or site.

Another factor in determining the value of a link is the way the link is implemented and where it is placed. For example, the text used in the link itself (i.e., the actual text that will go to your web page when the user clicks on it) is also a strong signal to the search engines.

This is referred to as *anchor text*, and if that text is keyword-rich (with keywords relevant to your targeted search terms), it can potentially do more for your rankings in the search engines than if the link is not keyword-rich. For example, anchor text of “Dog Grooming Salon” may bring more value to a dog grooming salon’s website than anchor text of “Click here.” However, take care. If you get 10,000 links using the anchor text “Dog Grooming Salon” and you have few other links to your site, this definitely does not look natural and could lead to a penalty.

The semantic analysis of a link’s value goes deeper than just the anchor text. For example, if you have that “Dog Grooming Salon” anchor text on a web page that is not really about dog grooming at all, the value of the link is lower than if the page is about dog grooming. Search engines also look at the content on the page immediately surrounding the link, as well as the overall context and authority of the website that is providing the link.

All of these factors are components of link analysis, which we will discuss in greater detail in Chapter 7.

Evaluating Social Media Signals

Sites such as Facebook, Twitter, and Google+ have created whole new ways for users to share content or indicate that they value it. This has led many to speculate that search engines could be using these signals as a ranking factor. Fueling that speculation, in August 2013, Moz released the data from its latest correlation study, and it showed a very strong correlation between +1s and ranking in Google.

Figure 2-23 shows the top 10 results in that data, and Google +1s had the second strongest correlation with rankings.

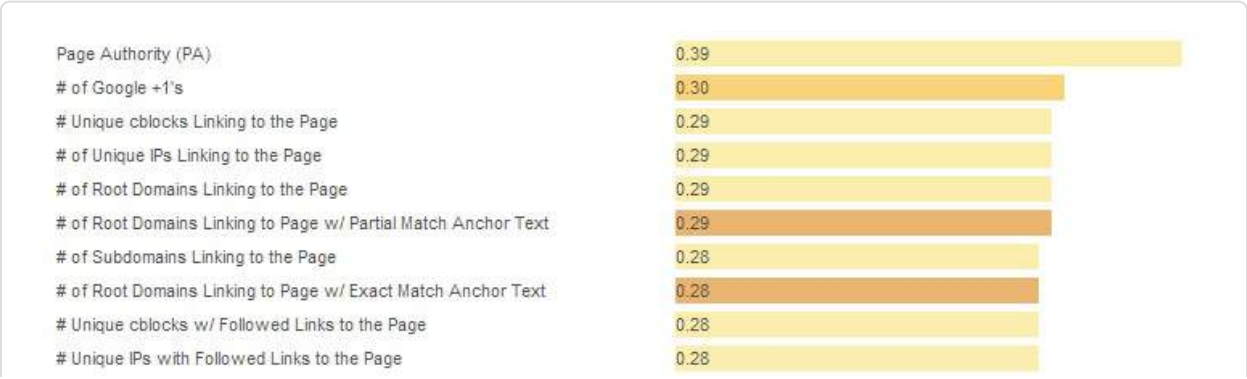


Figure 2-23. Top 10 results in Moz's 2013 correlation study

However, the fact that there is a correlation in no way means that +1s are used by Google as a ranking signal, or that they cause higher rankings. It can simply mean that good content that receives lots of links (which is known to be a signal that causes higher rankings) also happens to get many +1s.

In fact, Stone Temple Consulting did a different study targeted at measuring whether or not Google+ activity was used as a ranking factor by Google.[9] This study showed that there was no material evidence that Google+ shares or +1s had any material impact on rankings. The potential for social signals as a ranking factor is discussed in depth in Chapter 8.

Problem Words, Disambiguation, and Diversity

On the opposite side of the coin are words that present an ongoing challenge for the search engines. One of the greatest challenges comes in the form of disambiguation. For example, when someone types in *boxers*, does he mean the prize fighter, the breed of dog, or the type of underwear? Another example is *jaguar*, which is at once a jungle cat, a car, a football team, an operating system, and a guitar. Which does the user mean?

Search engines deal with these types of ambiguous queries all the time. The two examples offered here have inherent problems built into them, but the problem is much bigger than that. For example, if someone types in a query such as *cars*, does he:

- Want to read reviews?
- Want to go to a car show?
- Want to buy one?
- Want to read about new car technologies?

The query *cars* is so general that there is no real way to get to the bottom of the searcher's intent based on this query alone. One way that search engines deal with this is by looking at prior queries by the same searcher to find additional

clues to his intent. We discuss this a bit more in “Adaptive Search”.

Another solution they use is to offer diverse results. As an example, Figure 2-24 shows a generic search, this time for *GDP*.

Figure 2-24. Diverse results example

This brings up an important ranking concept. It is possible that a strict analysis of the relevance and link-driven importance scores in Figure 2-24 would not have resulted by itself in the Investopedia.com result being on the first page, *but* the need for diversity elevated the page’s ranking. This concept of altering the results in this manner is sometimes referred to as *query deserves diversity* (QDD).

A strict relevance-and importance-based ranking system might have shown a variety of additional government pages discussing the GDP of the United States. However, a large percentage of users will likely be satisfied by the government pages already shown, but for those users who are not, showing more of the same types of pages is not likely to raise their level of satisfaction with the results.

Introducing a bit of variety allows Google to also provide a satisfactory answer to those who are looking for something different from the government pages. Google’s testing has shown that this diversity-based approach has resulted in a higher level of satisfaction among its users.

For example, the testing data for the nondiversified results may have shown lower click-through rates in the SERPs, greater numbers of query refinements, and even a high percentage of related searches performed subsequently.

The idea to deliberately introduce diversity into the result algorithm makes sense and can enhance searcher satisfaction for queries such as:

- Company names (where searchers might want to get positive and negative press, as well as official company domains)
- Product searches (where ecommerce-style results might ordinarily fill up the SERPs, but Google tries to provide some reviews and noncommercial, relevant content)
- News and political searches (where it might be prudent to display “all sides” of an issue, rather than just the left-or right-wing blogs that did the best job of obtaining links)

Search engines also personalize results for users based on their search history or past patterns of behavior. For example, if a searcher has a history of searching on card games, and then does a search for *dominion*, the search engine may choose to push some of the results related to the *Dominion* card game higher in the results, instead of emphasizing the power company.

Where freshness matters

Much of the time, it makes sense for the search engines to deliver results from older sources that have stood the test of time. However, other times the response should be from newer sources of information.

For example, when there is breaking news, such as an earthquake, the search engines begin to receive queries within seconds, and the first articles begin to appear on the Web within 15 minutes.

In these types of scenarios, there is a need to discover and index new information in near real time. Google refers to this concept as *query deserves freshness* (QDF). According to the *New York Times*, QDF takes several factors into account,[10] such as:

- Search volume

- News coverage
- Blog coverage

QDF applies to up-to-the-minute news coverage, but also to other scenarios such as hot, new discount deals or new product releases that get strong search volume and media coverage. There has also been speculation that Google will apply QDF more to sites that have higher PageRank.[11]

Why These Algorithms Sometimes Fail

As we've outlined in this chapter, search engines do some amazing stuff. Nonetheless, there are times when the process does not work as well as you would like to think. Part of this is because users often type in search phrases that provide very little information about their intent (e.g., if they search on *car*, do they want to buy one, read reviews, learn how to drive one, learn how to design one, or something else?). Another reason is that some words have multiple meanings, such as the *jaguar* example we used previously in this section.

For more information on why search algorithms sometimes fail, you can read Hamlet Batista's Moz article, "7 Reasons Why Search Engines Don't Return Relevant Results 100% of the Time".

The Knowledge Graph

Traditional search results are derived by search engines crawling and analyzing web pages and then presenting that information in the search results. However, Google's mission "is to organize the world's information and make it universally accessible and useful". Google is actively pursuing initiatives to build databases of information that go far beyond traditional web-based search.

Note that earlier in this chapter we wrote about vertical search. Vertical search relates to breaking search into different categories, such as a search for images, videos, or local business information. The Knowledge Graph is more about providing richer answers directly in the search results, often answering the user's question directly without her having to click through to a website.

In May 2012, Google announced the Knowledge Graph. Initially, this was a set

of structured databases of information that allows Google to access information without deriving it from the Web. You can see an example of the type of data that Google might extract from its Knowledge Graph database in Figure 2-25.

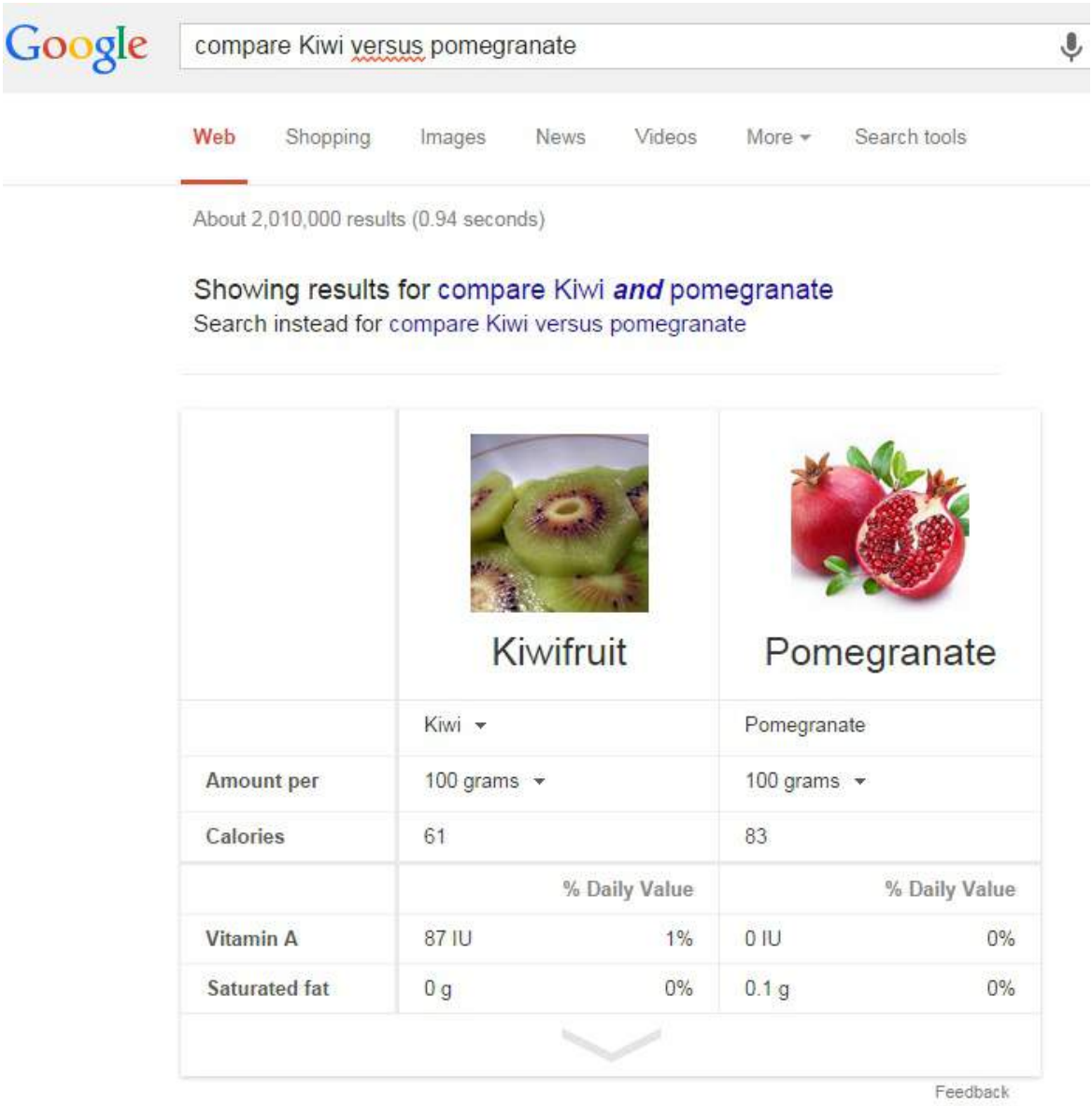


Figure 2-25. Kiwi versus pomegranate search result

Google initially built the Knowledge Graph using data from Freebase, Wikipedia, and the CIA Fact Book. This allowed Google to answer many questions, but really only satisfied a very small number of search queries. For that reason, Google is constantly working on expanding the information in the

Knowledge Graph.

In addition, Google is investing in ways to more reliably extract information from other sources, including websites, to present as direct answers in search. Google refers to these as “featured snippets.” Figure 2-26 shows the search result for buying a car.

In this result, Google provides a set of step-by-step instructions extracted from the CNN Money website. Note that two steps are omitted, so to get the complete procedure or additional details on each step, the user must click through to the CNN Money website.

In some cases, Google does provide the complete instructions in the search results, but most of the time it does not. A study performed by Stone Temple Consulting examimed 276 examples of step-by-step instructions, and found that 217 of these (79%) did not provide the complete instructions.

A related concept is *semantic search*, which overlaps the Knowledge Graph to some degree, but also takes into account many other factors to personalize results for the searcher. You can see a depiction of some of these factors in Figure 2-27.

Google

how to buy a car

Web Videos Books Shopping News More Search tools

About 29,600,000 results (0.84 seconds)

The Future of Car Buying - truecar.com
Ad www.truecar.com/
4.3 ★★★★★ rating for truecar.com
Never overpay and enjoy a hassle-free car buying experience
customer support · certified dealers · guaranteed savings · hassle-free pricing
Get upfront pricing before even visiting a dealership – Nationwide
Search Used Cars
See What Others Paid
2014 New Car Pricing
Select Make/Model

How To Buy A Car - Auto Buying Tips, Reviews, Pricing
Ad www.edmunds.com/
4.5 ★★★★★ rating for edmunds.com
Get Local Dealer Invoice Prices.
Ratings: Site content 10/10 - Site design 9/10 - Community 9/10
Edmunds.com has 1,788,746 followers on Google+
$15,000 to $25,000 - $55,000 to $85,000 - $45,000 to $55,000

CARFAX® Used Car Listings - CARFAX.com
Ad www.carfax.com/Cars_For_Sale 4.6 ★★★★★ advertiser rating
Find The Right Car With The Right History w/ CARFAX Used Car Listings

- Make sure you are getting the right vehicle. ...
- Assess the worth of your old car. ...
- Decide whether new or used is best for you. ...
- Consider whether leasing or buying makes more sense. ...
- Do your homework and set your target price. ...
- Shop for money before you shop for the car. ...
- Negotiating a lease. ...
- Negotiate a purchase.

2 more items

Buying a Car - Money Essentials, Lesson 17 - CNN Money
money.cnn.com/magazines/moneymag/money101/lesson17/ CNNMoney

Feedback

Figure 2-26. Step-by-step instructions for buying a car

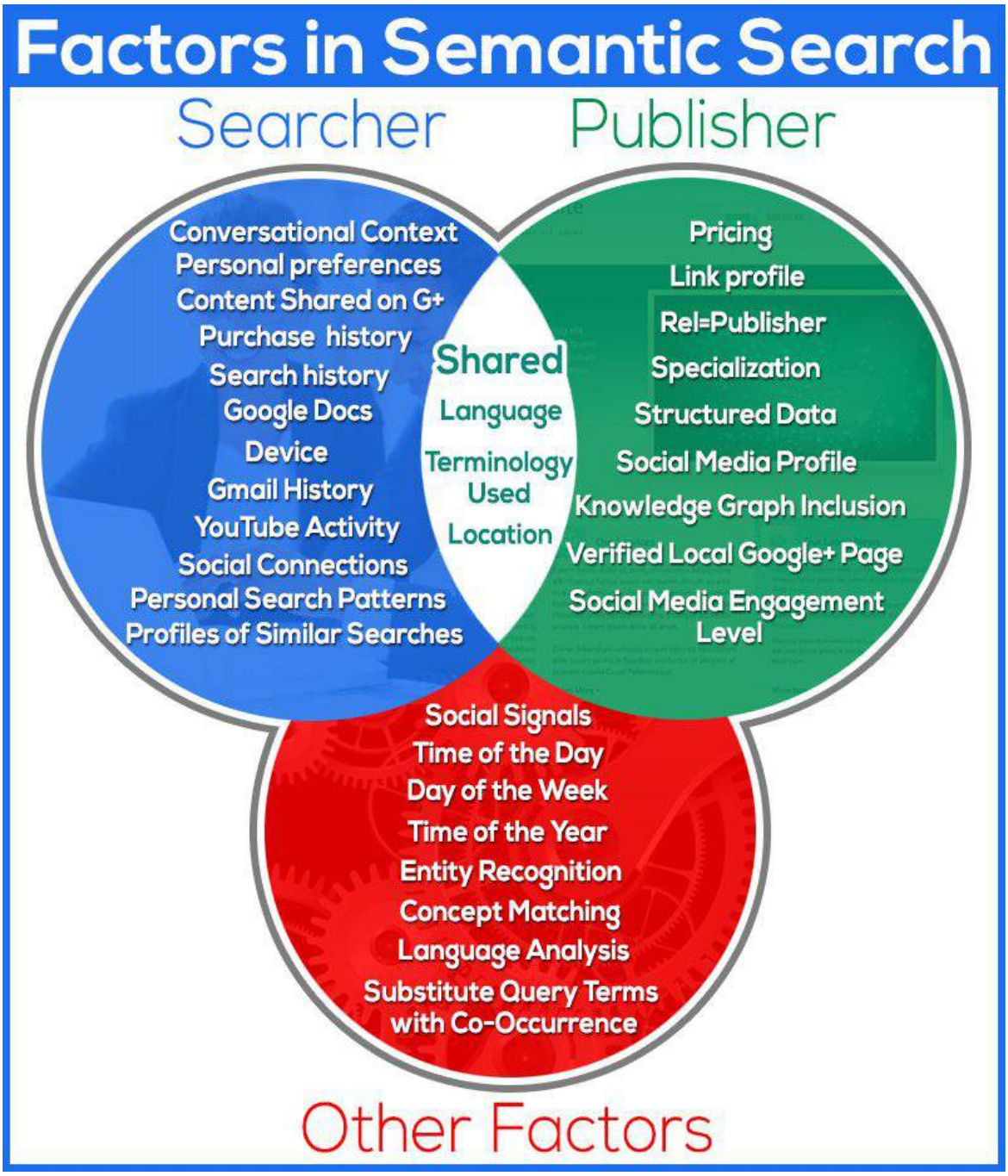


Figure 2-27. Factors involved in semantic search

The Knowledge Graph and semantic search are discussed in more detail in Chapter 6.

Analyzing Ranking Factors

Moz periodically conducts surveys of leading SEOs to determine what they think are the most important ranking factors.[12] Here is a high-level summary of the top nine results, in priority order (as suggested by the referenced study):

- Domain-level link authority features
- Page-level link metrics
- Page-level keywords and content
- Page-level keyword-agnostic features
- Domain-level brand metrics
- Usage and traffic/query data
- Page-level social metrics
- Domain-level keyword usage
- Domain-level keyword-agnostic features

Here is a brief look at each of these:

Domain-level link authority features

Domain-level link authority is based on a cumulative link analysis of all the links to the domain. This includes factors such as the number of different domains linking to the site, the trust/authority of those domains, the rate at which new inbound links are added, the relevance of the linking domains, and more.

Page-level link metrics

This refers to the links as related to the specific page, such as the number of links, the relevance of the links, and the trust and authority of the links received by the page.

Page-level keywords and content

This describes the use of the keyword term/phrase in particular parts of the HTML code on the page (`<title>` tag, `<h1>`, `alt` attributes, etc.).

Page-level features other than keywords

Factors included here are page elements such as the number of links on the page, number of internal links, number of followed links, number of "nofollow" links, and other similar factors.

Domain-level brand metrics

This factor includes search volume on the website's brand name, mentions, whether it has a presence in social media, and other brand-related metrics.

Page-level traffic/query data

Elements of this factor are click-through rate to the page in the search results, bounce rate of visitors to the page, and other similar measurements.

Page-level social metrics

Social metrics considered include mentions, links, shares, likes, and other social media site–based metrics. It should be emphasized that many SEO practitioners believe that this is a ranking factor even though studies have since shown otherwise, and representatives from Google clearly state that social signals are not part of their algorithm.

Domain-level keyword usage

This refers to how keywords are used in the root or subdomain name, and how impactful that might be on search engine rankings.

Domain-level keyword-agnostic features

Major elements of this factor in the survey include the number of hyphens in the domain name, number of characters in the domain name, and domain name length.

Negative Ranking Factors

It's also possible to have negative ranking factors. For example, if a site has a large number of low-quality inbound links that appear to be the result of artificial efforts by the publisher to influence search rankings, the site's rankings can be lowered. This is, in fact, exactly what Google's *Penguin* algorithm does. This algorithm is discussed more in Chapter 9. Some other potential negative ranking factors include:

Malware being hosted on the site

The search engines will act rapidly to penalize sites that contain viruses or Trojans.

Cloaking

Search engines want publishers to show the same content to the search engine as is shown to users.

Pages on the sites with links for sale

Google has a strong policy against paid links, and sites that sell links may be penalized.

Content that advertises paid links on the site

As an extension of the prior negative ranking factor, promoting the sale of paid links may be a negative ranking factor.

Page speed

Back in 2010, Google's Matt Cutts announced that Google was making page speed a ranking factor. In general, it is believed that this is a negative factor for pages that are exceptionally slow.

Other Ranking Factors

The ranking factors we've discussed so far are really just the basics. Search engines potentially factor in many more signals. Some of these include:

Rate of acquisition of links

If, over time, your site has acquired an average of 5 links per day, and then the links suddenly start to come in at a rate of 10 per day, that could be seen as a positive ranking signal. On the other hand, if the rate of new links drops to 2 per day, that could be a signal that your site has become less relevant.

However, it gets more complicated than that. If your site suddenly starts to get 300 new links per day, you have either become a lot more relevant or started to acquire links in a spammy way. The devil is in the details here, with one of the most important details being the origin of those new links. The concept of considering temporal factors in link analysis is documented in a U.S. patent held by Google, which you can look up by searching for patent number 20050071741.

User data

Personalization is one of the most talked-about frontiers in search. There are a few ways in which personalization can take place. For one, a search engine can perform a geolocation lookup to figure out where a user is approximately located. Based on this, the search engine can show results tailored to a user's current location. This is very helpful, for example, if the user is looking for a local restaurant.

Another way a search engine can get some data on a user is if he creates a profile with the search engine and voluntarily provides some information. A simple example would be a language preference. If the user indicates he prefers Portuguese, the search engine can tailor the results to that preference.

Search engines can also look at the search history for a given user. Basically, the search engine maintains a log of all the searches the user has performed when he is logged in. Based on this, it can see that he has been checking out luxury cars recently, and can use that knowledge to tweak the results he sees after he searches on *jaguar*. This is sometimes referred to as *adaptive search*.

To reduce the level of personalization, users can log out of their Google account. However, this does not disable *all* personalization, as Google may still tie some history to the person's computer. A user can disable all personalization by using Google's Chrome browser in *Incognito* mode. This

will allow her to see Google results that are not personalized based on search history. However, the results will still be personalized to her location.

A user can also depersonalize search results by performing her search query, and then appending *&pws=0* to the end of the search page URL and reloading the page. Note, this works only if she has turned off Google Instant (Google's feature of showing results instantly as the user types). Or, the user can choose the option "Disable customizations based on web history" under "webhistory" under the gear icon in the SERPs.

Using Advanced Search Techniques

One of the basic tools of the trade for an SEO practitioner is the search engines themselves. They provide a rich array of search operators that can be used to perform advanced research, diagnosis, and competitive analysis. The following are some of the more basic operators:

-keyword

Excludes the keyword from the search results. For example, *loans -student* shows results for all types of loans *except* student loans.

"key phrase"

Shows search results for the exact phrase—for example, *"seo company"*. You can also use *" "* to force the inclusion of a specific word. This is particularly useful for including *stopwords* (keywords that are normally stripped from a search query because they usually do not add value, such as the word *the*) in a query, or if your keyword is getting converted into multiple keywords through automatic stemming. For example, if you mean to search for the TV show *The Office*, you would want the word *The* to be part of the query. As another example, if you are looking for Patrick Powers, who was from Ireland, you would search for *"patrick powers" Ireland* to avoid irrelevant results for Patrick Powers.

keyword1 OR keyword2

Shows results for *at least one* of the keywords—for example, *google OR*

Yahoo!.

These are the basics, but for those who want more information, what follows is an outline of the more advanced search operators available from the search engines.

Advanced Google Search Operators

Google supports a number of advanced search operators that you can use to help diagnose SEO issues. Table 2-1 gives a brief overview of the queries, how you can use them for SEO purposes, and examples of usage.

Table 2-1. Google's advanced search operators

Operator	Short description	SEO application	Examples
site:	Domain-restricted search; narrows a search to one or more specific domains or directories.	Shows approximately how many URLs are indexed by Google.	For a website: *site:www.google.com* From a directory: *site:mit.edu/research/* Including all subdomains: *site:google.com* From a specific top-level domain (TLD): *site:org*
		▪ From a directory.	
		▪ Including all subdomains.	
		▪ From a specific top-level domain (TLD).	
inurl:/allinurl:	URL keyword restricted search; narrows the results to documents containing one	Find web pages having your keyword in a filepath.	*inurl:seo inurl:company* = *allinurl:seo company*

	or more search terms in the URLs.		
intitle:)/(allintitle:	Title keyword restricted search; restricts the results to documents containing one or more search terms in a page title.	Find web pages using your keyword in a page title.	*intitle:seo intitle:company = allintitle:seo company*
inanchor:/allinanchor:	Anchor text keyword restricted search; restricts the results to documents containing one or more search terms in the anchor text of backlinks pointing to a page.	Find pages having the most backlinks/the most powerful backlinks with the keyword in the anchor text.	*inanchor:seo inanchor:company = allinanchor:seo company*
intext:	Body text keyword restricted search; restricts the results to documents containing one or more search terms in the body text of a page.	Find pages containing the most relevant/most optimized body text.	*intext:seo*
ext:/filetype:	File type restricted search; narrows search results to the pages that end in a particular file extension.	A few possible extensions/file types: ▪ *.pdf* (Adobe Portable Document Format) ▪ *.html* or *.htm* (Hypertext Markup Language) ▪ *.xlsx* (Microsoft	*filetype:pdf ext:pdf*

Excel)

- *.pptx* (Microsoft PowerPoint)
- *.docx* (Microsoft Word)

related:	Similar URLs search; shows *related pages* by finding pages linking to the site and looking at what else they tend to link to (i.e., "co-citation"); usually 25–31 results are shown.	Evaluate how relevant the site's "neighbors" are.	Compare: *related:www.searchengineland.com* and *related:www.alchemistmedia.com*
info:	Information about a URL search; gives information about the given page.	Learn whether the page has been indexed by Google; provides links for further URL information; this search can also alert you to possible site issues (duplicate content or possible DNS problems).	*info:scienceofseo.com* will show you the page title and description, and invite you to view its related pages, incoming links, and page cached version.
cache:	What the page looked like when Google crawled it; shows Google's saved copy of the page.	Google's *text* version of the page works the same way as SEO browser.	*cache:www.stonetemple.com*

NOTE

When you use the *site:* operator, some indexed URLs might not be displayed (even if you use the "repeat the search with omitted results included" link to see the full list). The *site:* query is notoriously inaccurate. You can obtain a more accurate count of the pages of your site indexed by Google by appending *&start=990&filter=0* to the URL of a Google set for a search using the *site:* operator.

This tells Google to start with result 990, which is the last page Google will show you, as it limits the results to 1,000. This must take place in two steps. First, enter a basic *site:*

<yourdomain.com> search, and then get the results. Then go up to the address bar and append the *&start=990&filter=0* parameters to the end of the URL. Once you've done this, you can look at the total pages returned to get a more accurate count. Note that this works only if Google Instant is turned off.

To see more results, you can also use the following search patterns:

- *site:<yourdomain.com>/<subdirectory1> + site: <yourdomain.com>/<subdirectory2> + etc.* (the "deeper" you dig, the more/more accurate results you get)
- *site:<yourdomain.com> inurl:<keyword1> + site:<yourdomain.com> inurl: <keyword2> + etc.* (for subdirectory-specific keywords)
- *site:<yourdomain.com> intitle:<keyword1> + site:<yourdomain.com> intitle:<keyword2> + etc.* (for pages using the keywords in the page title)

To learn more about Google advanced search operators, check out Stephan Spencer's book *Google Power Search* (O'Reilly).

Combined Google queries

To get more information from Google advanced search, it helps to learn how to effectively combine search operators. Table 2-2 illustrates which search patterns you can apply to make the most of some important SEO research tasks.

Table 2-2. Combined Google search options

What for	Description	Format	Example
Competitive analysis	Find recent mentions of your competitor on other sites; use the `date range` option under the search tools in the SERPs; the following brand-specific search terms can be used: *<domainname.com>*, *<domain name>*, *<domainname>*, *<site owner name>*, and more.	*<domainname.com> - site: <domainname.com>* To select one day, pick "Search tools" → "Any time" → "Past 24 hours" .	*moz -site:moz.com during past 24 hours*
Keyword research	Evaluate the given keyword competition (sites that	*inanchor:<keyword> intitle: <keyword>*	*inanchor:seo intitle:seo*

	apply proper SEO to target the term).		
	Find more keyword phrases.	*key * phrase*	*free * tools*
SEO site auditing	Learn whether the site has canonicalization problems.	*site:<domain.com> -inurl:www*	*site:stephanspencer.com -inurl:www*
	Find the site's most powerful pages.	*site:<domain.com> -inurl:www*	*www site:alchemistmedia.com*
		inurl:<domainsite>: <domain.com>	*inurl:stonetemple site:stonetemple.com*
		<domain> site:<domain.com>	*alchemistmedia site:alchemistmedia.com*
	Find the site's most powerful page related to the keyword.	*site:<domain.com> <keyword>*	*site:scienceofseo.com seo*
		site:<domain.com> intitle: <keyword>	*site:stonetemple.com intitle:seo*
		site:<domain> inanchor: <keyword>	*site: moz.com inanchor:seo*
Link building	Find authority sites offering a backlink opportunity.	`site:`*`org bookmarks/links/"favorite sites"/`* `site:`*`gov bookmarks/links/"favorite sites"/`*] [`site:`*`edu bookmarks/links/"favorite sites"/`*	*site:org donors*
	Search for relevant forums and discussion boards to participate in discussions and probably link back to your site.	*inurl:forum OR inurl:forums <keyword>*	*inurl:forum OR inurl:forums seo*

Firefox plug-ins for quicker access to Google advanced search queries

You can use a number of plug-ins with Firefox to make accessing these advanced queries easier:

- Advanced Dork, for quick access to *intitle:*, *inurl:*, *site:*, and *ext:* operators for a highlighted word on a page, as shown in Figure 2-28

- SearchStatus, for quick access to a *site:* operator to explore a currently active domain, as shown in Figure 2-29

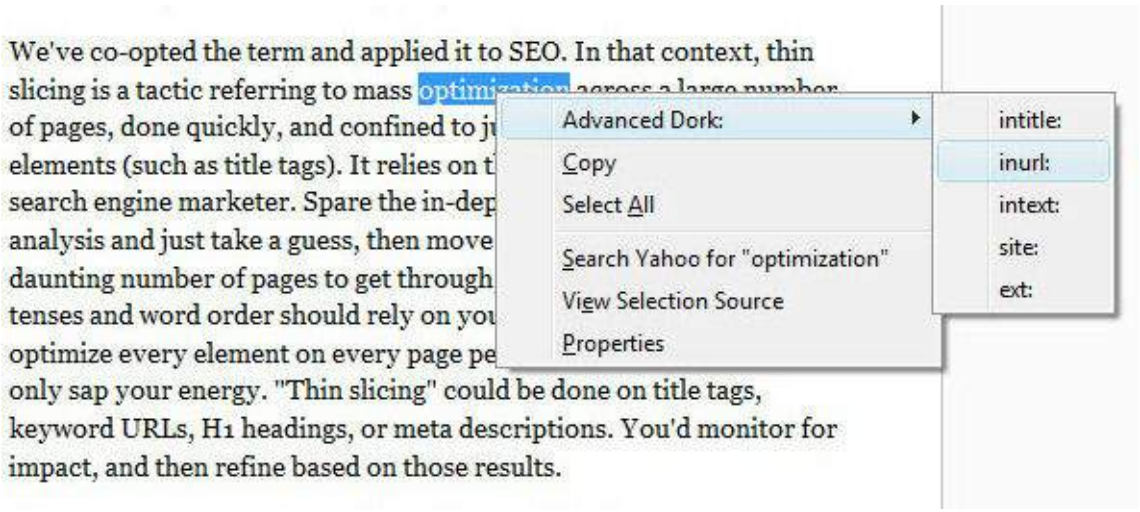


Figure 2-28. Advanced Dork plug-in for Firefox

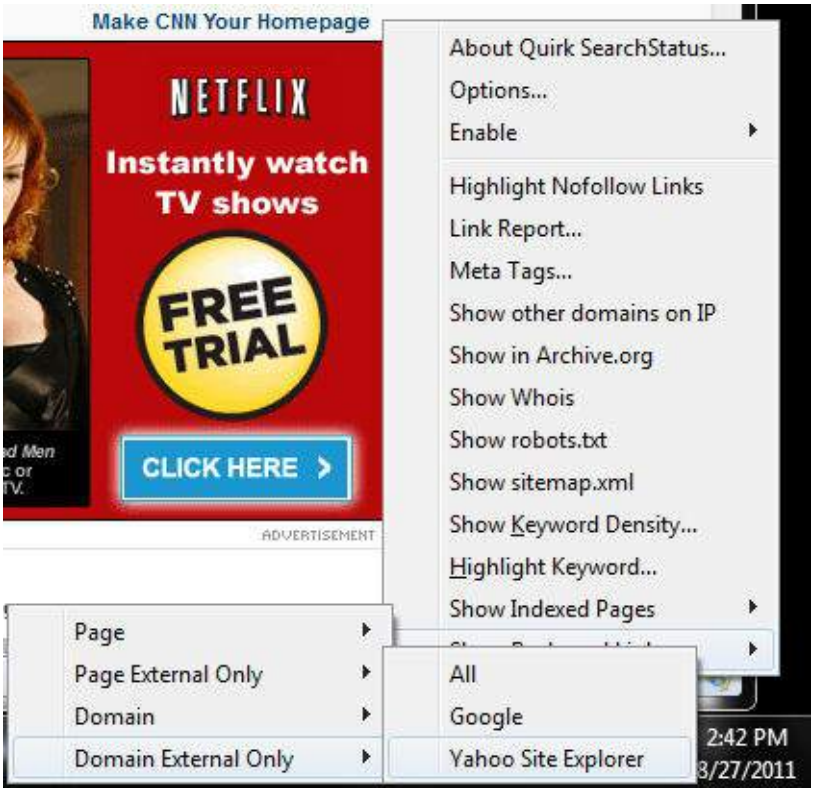


Figure 2-29. SearchStatus plug-in for Firefox

Bing Advanced Search Operators

Bing also offers several unique search operators worth looking into, as shown in Table 2-3.

Table 2-3. Bing advanced operators

Operator	Short description	SEO application	Examples
linkfromdomain:	Domain outbound links restricted search; finds all pages the given domain links out to.	Find most relevant sites your competitor links out to.	*linkfromdomain:moz.com seo*
contains:	File type links restricted search; `filetype:` is also supported; narrows search results to pages linking to a document of the specified file type.	Find pages linking to a specific document type containing relevant information.	*contains:wma seo*
ip:	IP address restricted search; shows sites sharing one IP address.	`ip:`*`xxx.xxx.xxx.xxx`*	*ip:207.182.138.245*
inbody:	Body text keyword restricted search; restricts the results to documents containing query word(s) in the body text of a page.	Find pages containing the most relevant/best optimized body text.	*inbody:seo* (equivalent to Google's *intext:*)
location:/*loc:*	Location-specific search; narrows search results to a specified location (multiple location options can be found under Bing's advanced search).	Find geospecific documents using your keyword.	*seo loc:AU*
feed:	Feed keyword restricted search; narrows search results to terms contained in RSS feeds.	Find relevant feeds.	*feed:seo*
hasfeed:	Feed keyword restricted search; narrows search results to pages linking to feeds that contain the specified keywords.	Find pages linking to relevant feeds.	*hasfeed:seo site:cnn.com*

More Advanced Search Operator Techniques

You can also use more advanced SEO techniques to extract more information.

Determining keyword difficulty

When you are building a web page, it can be useful to know how competitive the keyword is that you are going after, yet this information can be difficult to obtain. However, there are steps you can take to get some idea of how difficult it is to rank for a keyword. For example, the *intitle:* operator shows pages that are more focused on your search term than the pages returned without that operator (e.g., *intitle:"dress boots"*).

You can use different ratios to give you a sense of how competitive a keyword market is (higher results mean that it is more competitive). For example:

- *dress boots* (108,000,000) versus *"dress boots"* (2,020,000) versus *intitle:"dress boots"* (375,000)
- Ratio: 108,000/375 = 290:1
- Exact phrase ratio: 2,020/37 = 5.4:1

Another significant parameter you can look at is the *inanchor:* operator—for example, *inanchor:"dress boots"*. You can use this operator in the preceding equation instead of the *intitle:* operator.

Using number ranges

The number range operator can help restrict the results set to a set of model numbers, product numbers, price ranges, and so forth. For example: *site:stevespanglerscience.com "product/1700..1750"*

Unfortunately, the number range combined with *inurl:* is not supported, so the product number must be on the page. The number range operator is also great for copyright year searches (to find abandoned sites to acquire). Combine it with the *intext:* operator to improve the signal-to-noise ratio—for example, *intext:"copyright 1993..2011" -2014 blog*.

Using advanced doc type search

The *filetype:* operator is useful for looking for needles in haystacks. Here are a couple of examples:

- *confidential business plan -template filetype:doc*

- *forrester research grapevine filetype:pdf*

Determining listing age

You can label results with dates that give a quick sense of how old (and thus trusted) each listing is; for example, by appending the *&as_qdr=m199* parameter to the end of a Google SERP URL, you can restrict results to those within the past 199 months.

Uncovering subscriber-only or deleted content

You can sometimes get to subscriber-only or deleted content from the Cached link in the listing in the SERPs (found under the down arrow after the URL in the search listing) or by using the *cache:* operator. Don't want to leave a footprint? Add *&strip=1* to the end of the Google cached URL. Images on the page won't load.

If no Cached link is available, use Google Translate to take your English document and translate it from Spanish to English (this will reveal the content even though no Cached link is available):

- *http://translate.google.com/translate?prev=&hl=en&u=<URL-GOES-HERE>&sl=es&tl=en*

Identifying neighborhoods

The *related:* operator will look at the sites linking (the *linking sites*) to the specified site, and then see which other sites are commonly linked to by the linking sites. These are commonly referred to as *neighborhoods*, as there is clearly a strong relationship between sites that share similar link graphs.

Finding Creative Commons (CC) licensed content

Use the *as_rights* parameter in the URL to find Creative Commons licensed content. Here are some example scenarios to find CC-licensed material on the Web:

Permit commercial use

> *http://google.com/search?as_rights=(cc_publicdomain|cc_attribute|cc_sharealike|cc_nonderived).-(cc_noncommercial)&q=<KEYWORDS>*

Permit derivative works

> *http://google.com/search?as_rights=(cc_publicdomain|cc_attribute|cc_sharealike|cc_noncommercial).-(cc_nonderived)&q=<KEYWORDS>*

Permit commercial and derivative use

> *http://google.com/search?as_rights=(cc_publicdomain|cc_attribute|cc_sharealike).-(cc_noncommercial|cc_nonderived)&q=<KEYWORDS>*

Make sure you replace <*KEYWORDS*> with the keywords that will help you find content that is relevant to your site. The value of this to SEO is indirect. Creative Commons content can potentially be a good source of content for a website. An easier option if you don't need this same freedom in your Creative Commons searches is to use Google's Advanced Search page, where you can specify your Creative Commons license type.

Vertical Search Engines

Vertical search is a term sometimes used for specialty or niche search engines that focus on a limited data set. Examples of vertical search solutions provided by the major search engines are image, video, news, and blog searches. These may be standard offerings from these vendors, but they are distinct from the engines' general web search functions.

Vertical search engines sometimes come in the form of specialty websites, such as travel sites (such as TripAdvisor), and local business listing sites (such as YellowPages.com). Any site that focuses on vertically oriented niche markets could be considered a vertical search engine.

Vertical search results can provide significant opportunities for the SEO practitioner. High placement in these vertical search results can equate to high placement in the web search results, often above the traditional 10 blue links presented by the search engines.

Vertical Search from the Major Search Engines

The big three search engines offer a wide variety of vertical search products. Here is a partial list:

Google

> Google Maps, Google Images, Google Shopping, Google Blog Search, Google Video, Google News, Google Custom Search Engine, Google Book Search

Yahoo!

> Yahoo! News, Yahoo! Local, Yahoo! Images, Yahoo! Video, Yahoo! Shopping, Yahoo! Autos

Bing

> Bing Images, Bing Videos, Bing News, Bing Maps

Image search

All three search engines offer image search capability. Basically, image search engines limit the data that they crawl, search, and return in results to images. This means files that are in GIF, TIF, JPG, PNG, and other similar formats. Figure 2-30 shows the image search engine from Bing.

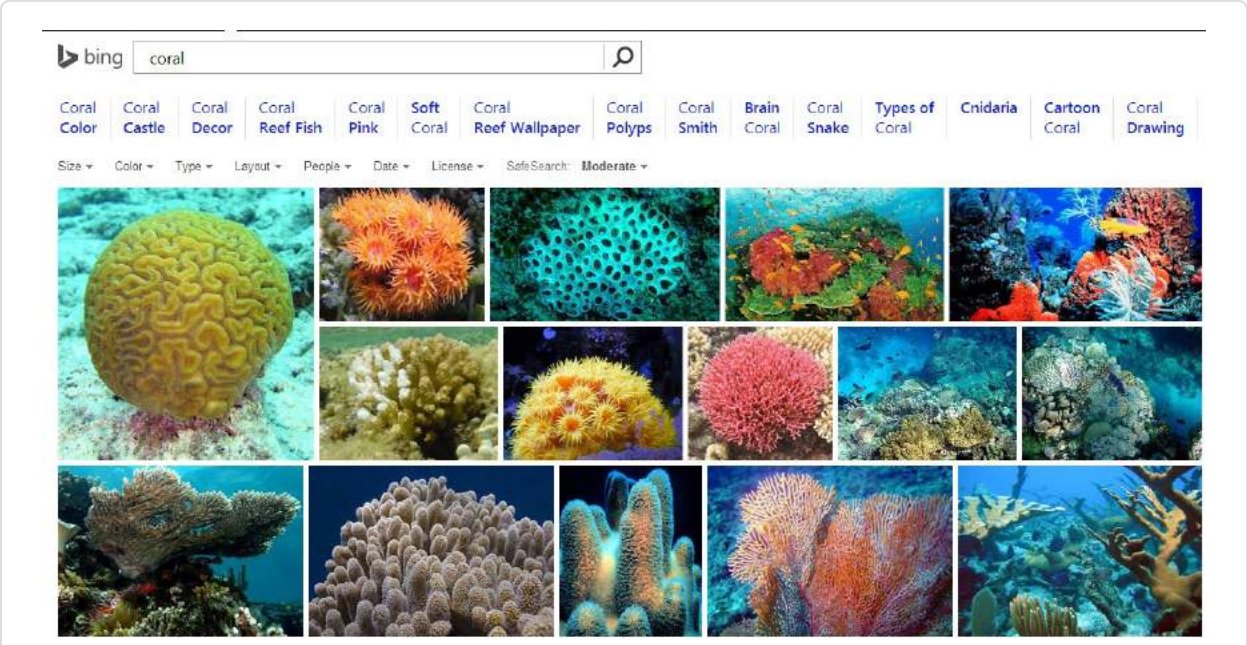


Figure 2-30. Image search results from Bing

Image search engines get a surprisingly large number of searches performed on them. Unfortunately, market data on these volumes is not often published, but according to comScore, more than 1 billion image searches are performed on Google Images per month. However, it is likely that at least that many image-related search queries occur within Google web search in the same timeframe. However, because an image is a binary file, it cannot be readily interpreted by a search engine crawler.

Search engines have had to historically rely on text surrounding the image, the `alt` attribute within the `<img>` tag, and the image filename. However, Google now offers a search by image feature that allows users to drag an image file into the Google Images search box and it will attempt to identify the subject matter of the image and show relevant results. Optimizing for image search is its own science, and we will discuss it in more detail in "Optimizing for Image Search".

Video search

As with image search, video search engines focus on searching specific types of files on the Web—in this case, video files such as MPEG, AVI, and others. Figure 2-31 shows a quick peek at video search results from YouTube.

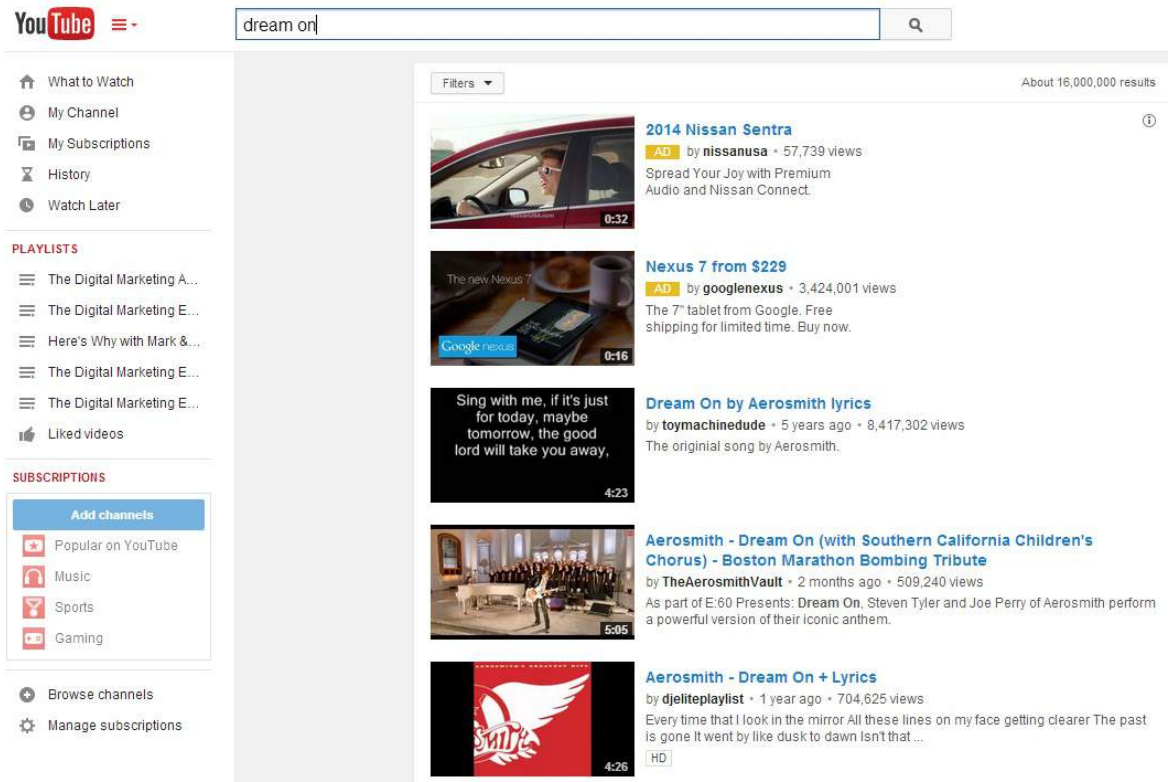


Figure 2-31. Video search results from YouTube

A very large number of searches are also performed in video search engines. YouTube is the dominant video search engine. Current data on total monthly searches is not readily available, but in June 2011, over 3.8 billion searches were performed on YouTube. This suggests that YouTube is the third largest search engine on the Web (Bing is larger when you consider the cumulative search volume of Bing and Yahoo!). As with image search, many video searches are also performed directly within Google web search.

You can gain significant traffic by optimizing for video search engines and participating in them. Once again, these are binary files and the search engine cannot easily tell what is inside them.

This means optimization is constrained to data in the header of the video and on the surrounding web page. We will discuss video search optimization in more detail in “Others: Mobile, Video/Multimedia Search” in “Optimizing for Video/Multimedia Search”.

However, each search engine is investing in technology to analyze images and

videos to extract as much information as possible. For example, the search engines are experimenting with OCR technology to look for text within images, and transcription and other advanced technologies are being used to analyze video content. In addition, flesh-tone analysis is being used to detect pornography or recognize facial features. The application of these technologies is in its infancy, and is likely to evolve rapidly over time.

News search

News search is also unique. News search results operate on a different time schedule; they have to be very, very timely. Few people want to read the baseball scores from a week ago when several other games have been played since then.

News search engines must be able to retrieve information in real time and provide near-instantaneous responses. Modern consumers tend to want their news information now. Figure 2-32 is a quick look at the results from a visit to Yahoo! News.

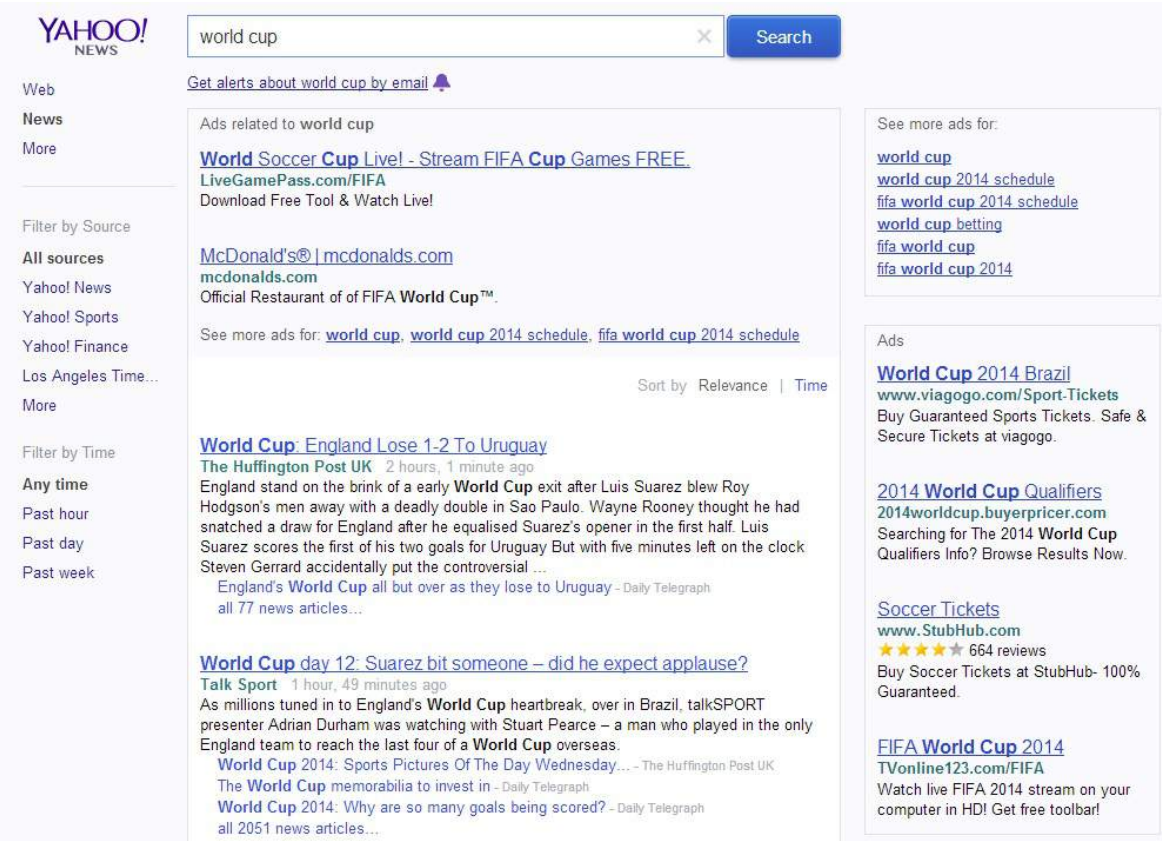


Figure 2-32. News search results from Yahoo!

As with the other major verticals, there is a lot of search volume here as well. To have a chance of receiving this volume, you will need to become a news source. This means timely, topical news stories generated on a regular basis. These and other requirements are discussed further in “Optimizing for News Search: Google News”.

Local search/maps

Next up in our hit parade of major search verticals is local search (a.k.a. map search). Local search results are now heavily integrated into the traditional web search results, so a presence in local search can have a large impact on organizations that have one or more brick-and-mortar locations. Local search engines search through databases of locally oriented information, such as the name, phone number, and location of local businesses around the world, or just provide a service, such as offering directions from one location to another. Figure 2-33 shows Google Maps local search results.

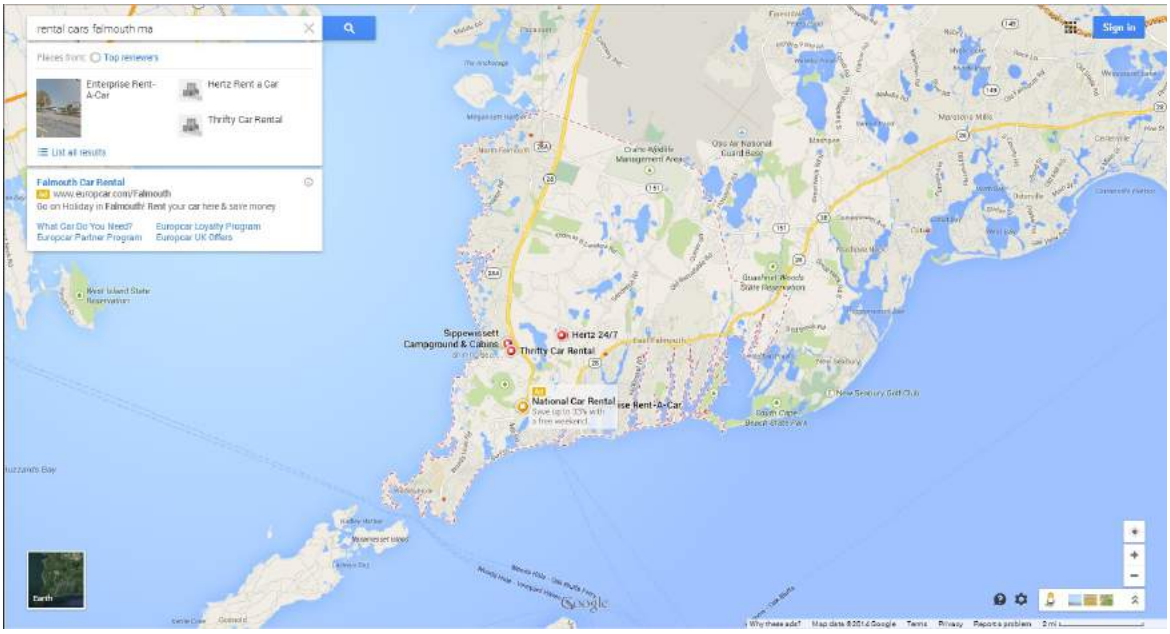


Figure 2-33. Local search results from Google Maps

The integration of local search results into regular web search results has dramatically increased the potential traffic that can be obtained through local search. We will cover local search optimization in detail in “Optimizing for Local Search”.

Blog search

Google has implemented a search engine focused just on blog search called Google Blog Search (misnamed because it is an RSS feed engine, not a blog engine). This search engine will respond to queries, but only searches blogs (more accurately, feeds) to determine the results. Figure 2-34 is an example search result for the search phrase *barack obama*.

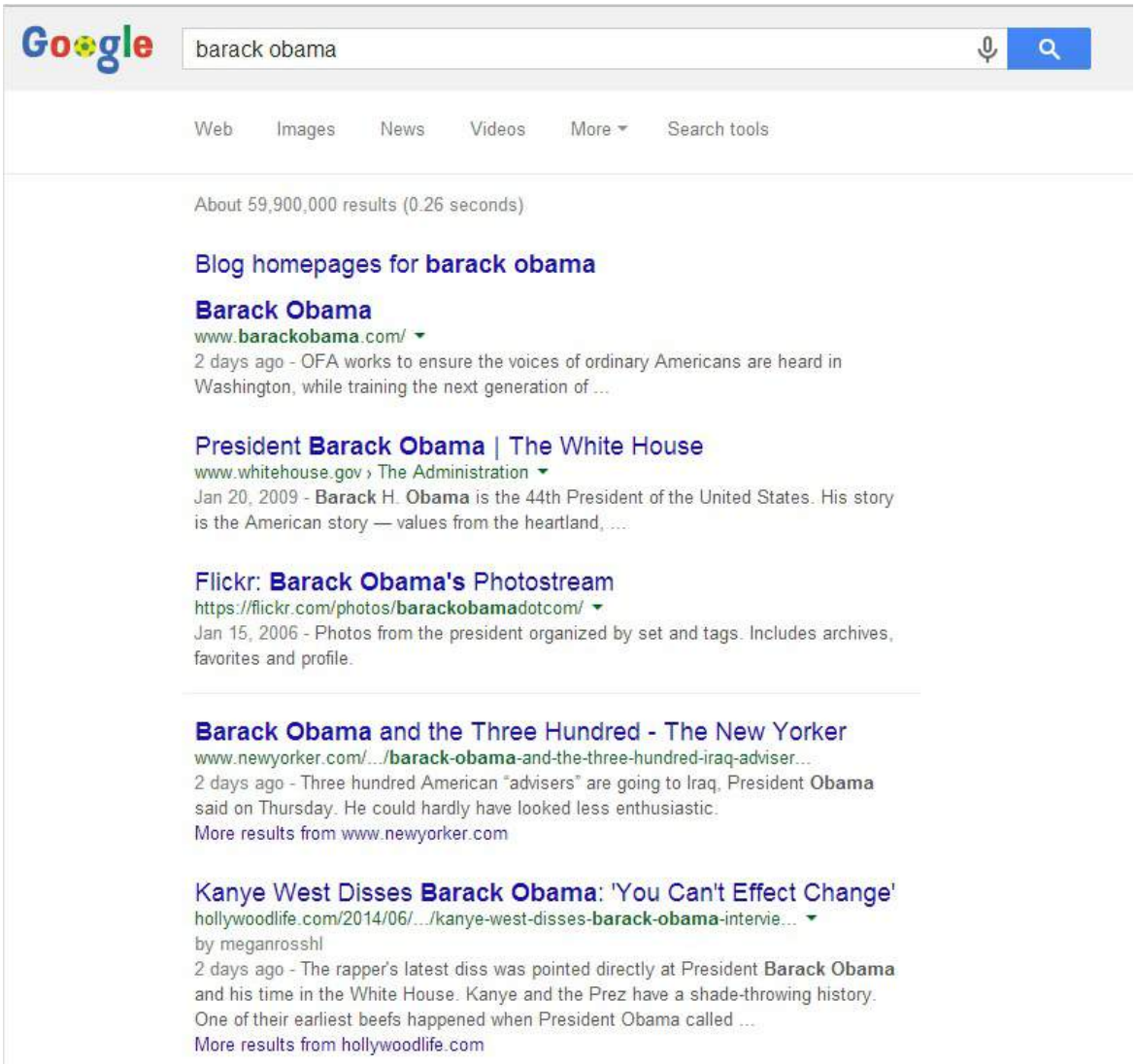


Figure 2-34. Results from Google Blog Search

We explore the subject of optimizing for Google Blog Search in "Optimizing for Blog Search".

Book search

The major search engines also offer a number of specialized offerings. One highly vertical search engine is Google Books search, which specifically searches only content found within books, as shown in Figure 2-35.

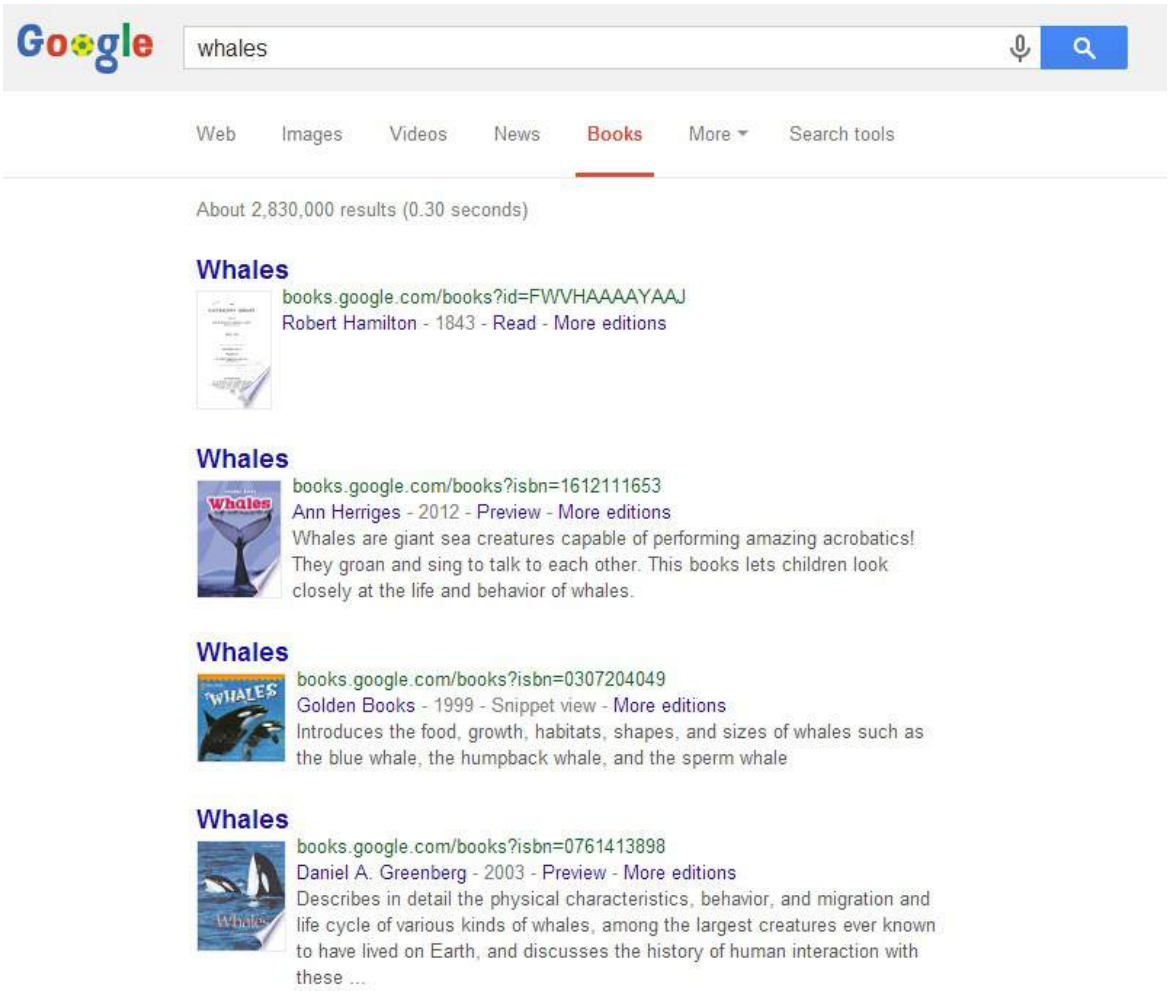


Figure 2-35. Google Books search

Product search

Bing also has some unique vertical search features. One of the more interesting ones is its product search solution. Instead of having a separate shopping search engine, Bing has integrated the product results into the main body of its search results, as shown on the right side of Figure 2-36.

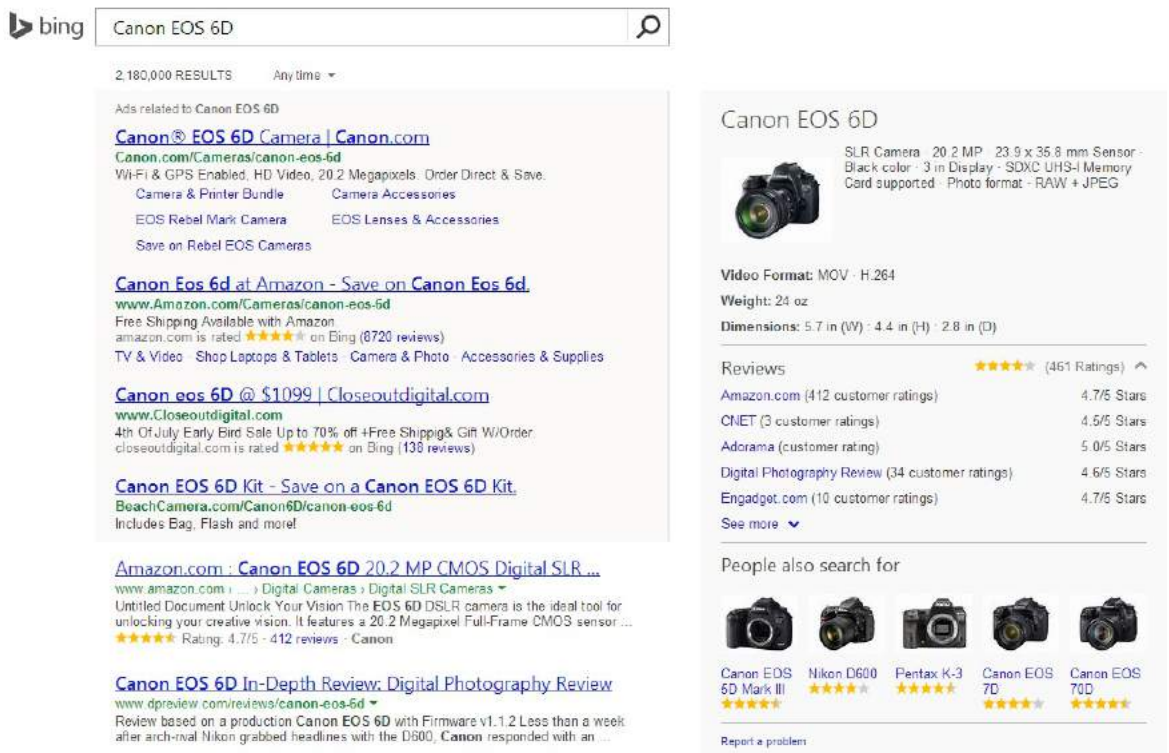


Figure 2-36. Bing product search

Universal Search/Blended Search

Google made a big splash in 2007 when it announced Universal Search, the notion of integrating images, videos, and results from other vertical search properties directly into the main web search results. Prior to this announcement, all the search engines showed their search results in separate vertical search engines. You have already seen an example of this in Figure 2-36, which shows Bing's way of integrating product search features directly into the main search results.

After Google's announcement, both Bing and Yahoo! quickly followed with their own implementations. Each type of result you see on a search results page offers different opportunities for obtaining traffic from search engines. People now refer to this general concept as *blended search* (because Universal Search is specifically associated with Google).

Figure 2-37 shows an example of blended search results from a Google search.

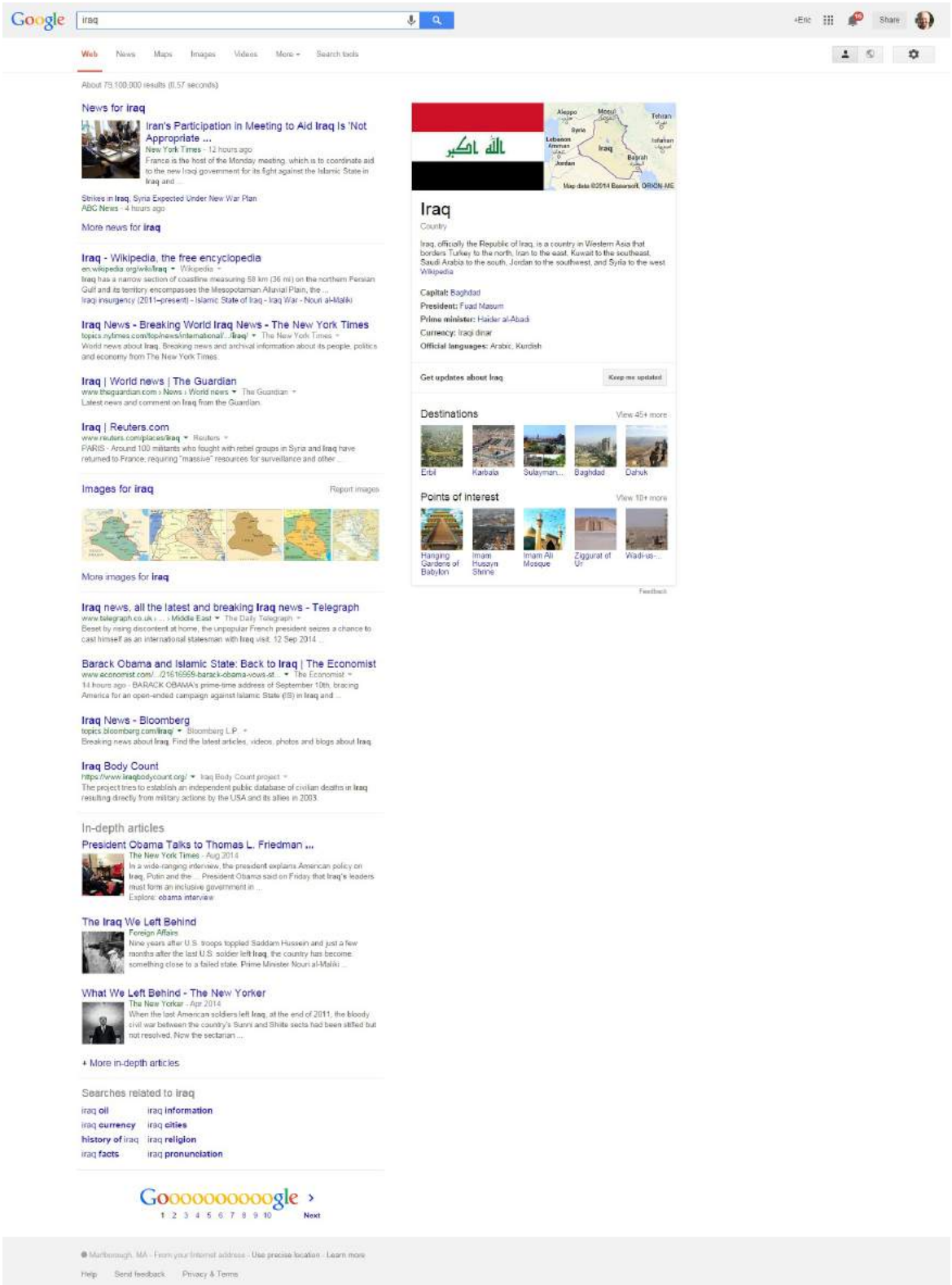


Figure 2-37. Google Universal Search results

More specialized vertical search engines

Vertical search can also come from third parties. Here are some examples:

- Comparison shopping engines (e.g., PriceGrabber, Shopzilla, and Nextag)
- Travel search engines (e.g., Expedia, Travelocity, and Kayak)
- Real estate search engines (e.g., Trulia and Zillow)
- Job search engines (e.g., Indeed, CareerBuilder, and SimplyHired)
- Music search engines (e.g., iTunes Music Store)
- B2B search engines (e.g., KnowledgeStorm and ThomasNet)

There is an enormous array of different vertical search offerings from the major search engines, and from other companies as well. We can expect that this explosion of different vertical search properties will continue.

Effective search functionality on the Web is riddled with complexity and challenging problems. Being able to constrain the data types (to a specific type of file, a specific area of interest, a specific geography, etc.) can significantly improve the quality of the results for users.

Country-Specific Search Engines

At this stage, search is truly global in its reach. Google is the dominant search engine in many countries, but not all of them. How you optimize your website depends heavily on the target market for that site, and the one or more search engines that are the most important in that market.

According to comScore, Google receives 54.3% of all searches performed worldwide as of April 2014. In many countries, that market share is 80% or more.

Here is some data on countries where other search engines are major players:

China

China Internet Watch reported in September 2014 that Baidu had about 70%

market share. This is significant because China has the largest Internet usage in the world, with 618 million users in 2010 according to China Internet Network Information Center.

Russia

According to figures reported by Yandex, the company's market share in Russia comprised about 62% of all searches in April 2014.[13]

South Korea

Naver was estimated to have about 70% market share in South Korea in March 2014.[14]

Czech Republic

In January 2014, the Startup Yard blog reported that Seznam had more than 60% market share in the Czech Republic.[15]

Optimizing for Specific Countries

One of the problems international businesses continuously need to address with search engines is identifying themselves as "local" in the eyes of the search engines. In other words, if a search engine user is located in France and wants to see where the wine shops are in Lyon, how does the search engine know which results to show?

Here are a few of the top factors that contribute to international ranking success:

- Owning the proper domain extension (e.g., *.com.au*, *.uk*, *.fr*, *.de*, *.nl*) for the country that your business is targeting
- Hosting your website in the country you are targeting (with a country-specific IP address)
- Registering with local search engines:

  — Google My Business

  — Yahoo! Small Business

— Bing Places

- Having other sites from the same country link to you
- Using the native language on the site (an absolute requirement for usability)
- Helping Google serve the correct language or regional URL in the search results by adding the `hreflang` attribute (*https://support.google.com/webmasters/answer/189077?hl=en*)
- Placing your relevant local address data on major pages of the site
- Setting your geographic target in Google Search Console (you can read more about this at *http://bit.ly/country_targeting*); note that Google does not really need you to do this if your site is on a country code top-level domain (ccTLD), such as *.de* or *.co.uk*, as the preferred regional target is assumed

All of these factors act as strong signals to the search engines regarding the country you are targeting, and will make them more likely to show your site as a relevant local result.

The complexity increases when you are targeting multiple countries. We will discuss this in more depth in "Best Practices for Multilanguage/Country Targeting".

Conclusion

Understanding how search engines work is an important component of SEO. The search engines are constantly tuning their algorithms. For that reason, the successful SEO professional is constantly studying search engine behavior and learning how search engines work.

[1] Google Official Blog, "We Knew the Web Was Big...", July 25, 2008, *http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html*.

[2] For more information, see Matt McGee, "21 Essential SEO Tips & Techniques," Search Engine Land, June 20, 2011, *http://searchengineland.com/21-essential-seo-tips-techniques-11580*.

[3] Google Webmaster Central Blog, “Understanding Web Pages Better,” May 23, 2014, *http://bit.ly/web_pages_understanding*.

[4] For even further discussion of this topic, see “How do search engines deal with AngularJS applications?” on StackOverflow.

[5] For more information, see Matt McGee, “Google: Low PageRank & Bad Spelling May Go Hand-In-Hand; Panda, Too?”, October 5, 2011, *http://bit.ly/pagerank_spelling*.

[6] See Danny Sullivan, “Google Forecloses On Content Farms With “Panda” Algorithm Update,” February 24, 2011, *http://bit.ly/panda_update*.

[7] Stéphane Hamel, “Web Analytics vendors market shares,” immeria - S.Hamel’s blog, January 4, 2008, *http://blog.immeria.net/2008/01/web-analytics-vendors-market-shares.html*.

[8] Metric Mail, “Google Analytics Market Share,” The Metric Mail Blog, August 4, 2010, *http://bit.ly/analytics_mkt_share*.

[9] Eric Enge, “Direct Measurement of Google Plus Impact on Search Rankings,” Stone Temple Consulting, September 17, 2013, *https://www.stonetemple.com/measuring-google-plus-impact-on-search-rankings/*.

[10] Saul Hansell, “Google Keeps Tweaking Its Search Engine,” New York Times, June 3, 2007, *http://www.nytimes.com/2007/06/03/business/yourmoney/03google.html?pagewanted=3&_r=0*.

[11] Sean Jackson, “The Surprising Effect of Freshness and Authority on Search Results By,” Copyblogger, February 21, 2013, *http://www.copyblogger.com/query-deserves-freshness/*.

[12] For more information, see “2013 Search Engine Ranking Factors,” *https://moz.com/search-ranking-factors*.

[13] Amy Gesenhues, “Yandex Reports 62% Share Of Russian Search Market With Q1 2014 Revenue Up 36%,” Search Engine Land, April 24, 2014, *http://bit.ly/russian_search_market*.

[14] kmc, “Should Korean Search Engine Naver Worry About Google?,” *http://www.korea-marketing.com/should-naver-worry-about-google/*.

[15] Lloyd Waldo, “Meet the Only Company in Europe that is Beating Google Seznam.cz,” Startup Yard Blog, January 3, 2014, *http://bit.ly/seznam_europe*.

CHAPTER THREE

SEO Planning: Customizing Your Strategy

SEO, once a highly specialized and often marginalized task relegated to the back rooms of website development teams, is now a mainstream digital marketing activity. Its dramatic rise can be attributed to various emerging trends, including:

- Search engines now drive a large volume of highly targeted traffic from people intent on accomplishing research-based, purchase-oriented, and content consumption goals. Businesses can earn significant revenues by leveraging the quality and relevance of this traffic for direct sales, customer acquisition, and content and branding/awareness campaigns.

- Visibility in search engines creates an implied endorsement effect, whereby searchers associate quality, relevance, and trustworthiness with sites that rank highly for their queries.

- Dramatic growth in the interaction between offline and online marketing necessitates investment by organizations of all kinds in a successful search strategy. Consumers are increasingly turning to the Internet for tasks ranging from researching purchases in verticals such as real estate, automobiles, and technology to finding local resources for health and wellness services. At this stage, businesses simply cannot afford to ignore their customers' needs as expressed through searches conducted on the major search engines.

Search engine optimization is a technical marketing function that relies upon both content creation and technological development, and success in organic search greatly depends upon various technical and analytical tools that provide access to valuable data necessary for ongoing SEO and website improvement, as well as overall user experience optimization. The intersection of and interdependence between website technology and online marketing has become clear, and essential—and SEO was at the forefront of this trend. New conference events such as MarTech (produced by Third Door Media, the company behind Search Engine Land and Search Marketing Expo) are beginning to provide platforms for learning and growth within the powerful intersection of marketing and technology.

As with all marketing functions, specific goal setting is required for success to

be measured and achieved. While SEO can be viewed as a *project* (and there are certainly situations in which SEO "projects" exist) the best investment, in our opinion, is to treat it as more of a *process*—one that is iterative and ongoing, and requires steady commitment from the stakeholders of an organization.

An ongoing investment is needed to create a site with the proper content, architecture, and overall foundation for SEO success. The results may not appear instantly, but a business that makes a patient and prudent commitment to SEO, invests in quality content development, constructs a crawlable site architecture, and earns legitimate linking relationships will be handsomely rewarded.

Strategic Goals SEO Practitioners Can Fulfill

Organic search, as a source of highly targeted traffic, is one of many viable online marketing channels. And while SEO is not a cure-all for businesses, it can fit into a company's overall business strategy in several critical ways.

Visibility (Branding)

Many consumers assume that top placement in search engines is like a stamp of approval on a brand. Surely a company could not rank highly in search results if it were not one of the best in its field, right?

If you are an experienced search engine user, you probably recognize that the preceding statement is not always true. However, the fact is that many consumers, and even sophisticated web searchers, interpret high search rankings as an implicit brand endorsement.

Therefore, for critical brand terms, an SEO strategy should be geared toward increasing organic search exposure for branded keywords.

In addition to targeting brand terms, you will want to rank well for nonbranded search terms that are specific to your core business (a highly competitive, but highly rewarding, branding effort). When searchers see you ranking highly on these types of search terms, they will associate your brand with the product or service and assume you are one of the best places to purchase from.

The list of situations where the brand can also limit the SEO strategy is quite long, and the opposite can happen too, where the nature of the brand makes a

particular SEO strategy pretty compelling. Ultimately, your goal is to dovetail SEO efforts with branding as seamlessly as possible.

Website Traffic

Long gone are the days of a “build it and they will come” paradigm online. Today’s search environment is highly competitive, and you need great SEO to capture targeted, high-quality traffic to your site.

While a business that engages with many of its customers through offline channels can drive traffic by telling those customers to visit its website, SEO fills the different, more critical role of bringing new prospects to your website from an audience of people who might not otherwise have been interested in, or perhaps even aware of, your business at all.

Experienced SEO professionals understand that users search for products, services, and information using an extraordinarily wide variety of search queries and query types. Developing an SEO strategy involves performing extensive keyword research (which we will discuss in Chapter 5) to determine which search queries people actually use. For example, when interested in purchasing a fuel-efficient car, a searcher might type in *hybrid efficient cars*. In this instance, the search marketing manager for a manufacturer of electric cars might be interested in search exposure for this traffic, even though technically, *electric cars* was not in the original search query.

Some users may not even know that a company specializing in a specific type of product exists until they perform that search. Or, if they have at one time learned about such a manufacturer, they might not remember enough about it to seek out the manufacturer’s website directly.

Capturing that traffic could provide the manufacturer with incremental sales of its electric vehicles that it probably would not have gotten otherwise. Knowing these factors, the SEO process involves developing a site architecture strategy (see Chapter 6) and a content development/editorial strategy (Chapter 7) to help the site’s pages achieve competitive search engine exposure for a broad range of potentially relevant terms.

High Return on Investment

Increasing online visibility and driving targeted traffic are the first steps to success with SEO. The next step is measuring the performance of that traffic to determine whether the overall website and business objectives are being achieved. This is an ongoing, iterative process that will always accompany all marketing efforts, and SEO is no exception. For most organizations, goals are defined as generating sales, leads, or advertising revenue. For others, goals may include promoting a particular message; consuming, sharing, or perhaps creating specific types of content; or gaining simple signups for a newsletter. An important component of SEO is to deliver not just traffic, but *targeted* traffic that has the potential to convert into these actions. Whether you are selling products and services, advertising for branding value, or trying to promote specific content to the world, a well-designed SEO strategy can result in a very high return on investment (ROI) when contrasted with other methods of marketing.

SEO generally brings a higher ROI when compared to TV, print, and radio advertising campaigns; and while traditional media is not in danger of being replaced by SEO, organic search can provide some high-margin returns that complement and enhance the use of offline media—especially when done properly over the long haul (see Figure 3-1).

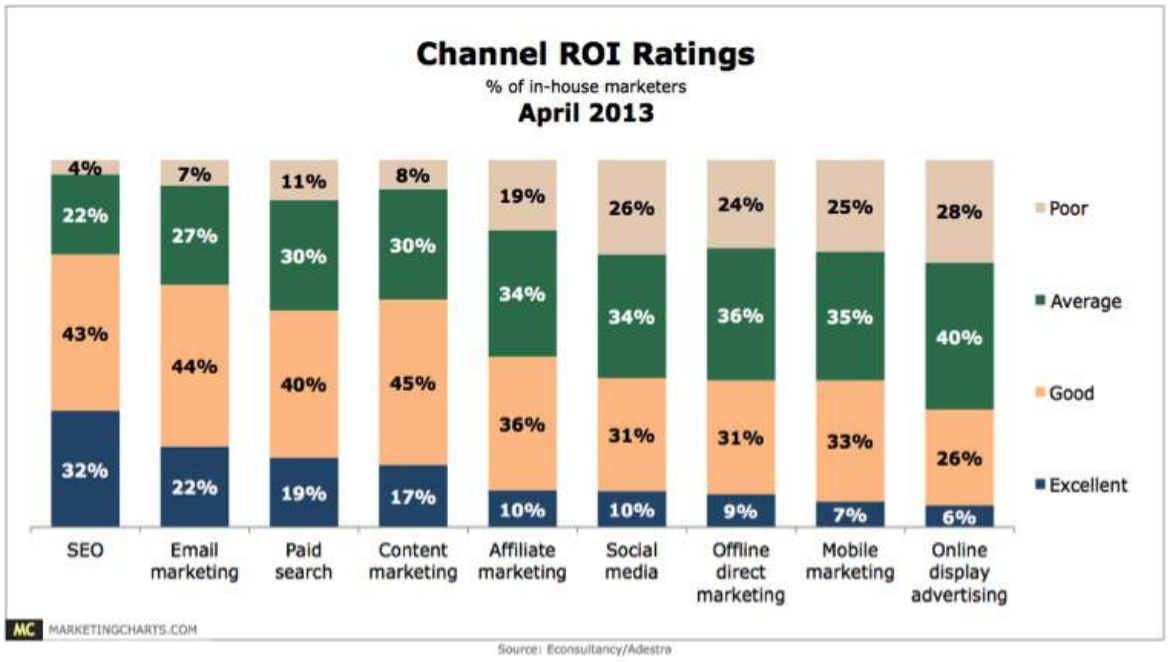


Figure 3-1. SEO, a high-ROI activity

In addition, a growing number of businesses operate purely online. Some examples of these are LinkedIn, Zappos, Amazon, and eBay.

Every SEO Strategy Should Be Customized

There is no such thing as a cookie-cutter SEO strategy (and if someone is trying to sell you one, be sure to dig deeper and get second opinions!). There is a difference between best practices, which are universal, and a business-and website-specific SEO strategy. The ever-changing, dynamic nature of the search marketing industry requires constant diligence, and SEO professionals must maintain a research process for analyzing how the search landscape is changing.

You must take the following factors into account when developing your SEO strategy:

- What the organization is trying to promote (service, product, content)
- Who the target market is (can be as simple as “women” or as detailed as personas)
- Brand (includes copy and messaging)
- Website structure (includes site architecture, navigational elements, and file/URL naming conventions)
- Current site content assets (includes images, videos, PDF files, white papers, case studies, articles)
- Ease with which the content and site structure can be modified (involves the CMS and web development teams)
- Editorial resources and calendar for content development (what content is developed, by whom, and on what timeline)
- Competitive landscape

Learning about the specific market a business is in is an obvious first step toward developing industry familiarity, trends, and vernacular—but it often makes sense

for two businesses offering the same products in the market to use different SEO strategies.

For example, if Competitor 1 published its website four years ago, and Competitor 2 is just launching its site now, Competitor 2 may gain the most traction by focusing on specific vertical areas where Competitor 1's offering or SEO implementation is weak.

It is important to not underestimate the importance of a well-thought-out SEO strategy. Skipping over this process or not treating it seriously can hamper your prospects for growth.

Understanding Search Engine Traffic and Visitor Intent

As we discussed in "The Mission of Search Engines", searchers enter many different types of queries. These can generally be classified into three major categories:

Navigational query

This is a query with the intent to arrive at a specific website or page (e.g., the person types in your company domain name, *www.companyname.com*, or simply types in the word *facebook*).

Informational query

This is a search performed to receive an answer to a broad or direct question, or to research and explore information around a specific topic with no specific source in mind (e.g., *yoga poses*).

Transactional query

A person who types in *digital camera* may be looking to buy one now, but it is equally possible that she is researching digital cameras to learn about how they are different from film cameras. This is an example of an initial transactional query, which can evolve in stages. For example, here are some other types of transactional queries that occur at a later stage in the buying cycle:

- The user types in *best online digital camera store*. Although there is no information in the query about which one she wants to buy, the intent is clearer that the searcher is seeking a store, not simply information about types of digital cameras.

- The searcher types in *olympus OMD lowest price*. The chances are very high that this user is looking to buy that particular camera.

Part of an SEO strategy is to understand how the various types of searches relate to the content and architecture of your website.

Developing an SEO Plan Prior to Site Development

It is widely understood in the SEO industry that SEO should be built in, as early as possible, to the entire site development strategy—from choosing a content management system (CMS) and planning site architecture to creating, optimizing, and publishing site content. As you will see in Chapter 6, SEO is deeply intertwined with these areas.

SEO plans have many moving parts, and SEO-related decisions can and usually do have a significant impact on other departments, such as web development, content development and editorial, other marketing groups (direct, offline, etc.), and sales. Integrating these moving parts and aligning them with each other is essential to developing an SEO-friendly website and establishing a strong foundation for enduring organic visibility.

Business Factors That Impact Your SEO Strategy

Here are some examples of business issues that can impact SEO:

Revenue and business models

The effective SEO strategy takes into account the purpose of the site—whether it is to sell products, sell advertising, obtain leads, or gain membership signups. We will discuss this more in the later sections of this chapter.

Target customers

Who are you trying to reach? This could be an age group, a gender group, or as specific as people looking to buy a house within a specific neighborhood of San Francisco, California.

Competitors

The competitive landscape is another big factor in your overall strategy. Competition may be strongly entrenched in one portion of the market online, and it may make sense to focus on a different segment. Or you may be the current leader in your market and want to protect this position while continuing to build your customer base.

Branding goals

There may be search terms for which it is critical that you have top search exposure, for branding reasons.

Content development

An important part of SEO and general online success is the creation and optimization of high-quality content for your users. For most businesses, your capacity to create quality content on an ongoing basis can significantly improve your SEO efforts, both from a content-availability perspective (more content in the search engines) and from a user engagement and link development perspective (great content breeds great links and social sharing, both of which can positively influence SEO).

How people search for products like yours

Understanding what customers do when they are searching for products or services like yours is one of the most basic functions of SEO, which we will discuss in detail in Chapter 5. This involves mapping the actual search queries your target customers use when they go to a search engine.

Understanding Your Audience and Finding Your

Niche

Your target audience should drive the initial site design, and should inform your SEO and content development strategies. As you will see in this section, there are many related considerations here, including your competition and the particular strengths or weaknesses of your own company.

Mapping Your Products and Services

Successful SEO requires a thorough understanding of your business. What products, services, and types of information and resources do you have to offer your site visitors and potential customers?

As we outlined in the preceding section, a critical component of SEO is understanding who is searching for what you are trying to promote, and this requires thoroughly understanding all aspects of your offering. You will need to understand both the broad and specific market categories that your products fall into, as each of these categories might relate to sections of your website that you will likely need to create. By having content areas of the site for those categories, you create the opportunity to obtain organic search traffic related to those categories.

You also should consider business development and the company's expansion strategy at the outset of the SEO planning process. Consider Amazon, which began as a bookseller but has evolved into a general-purpose etailer; sites that go through these types of changes may need to be substantially restructured, and such restructurings can be a source of major SEO headaches. Anticipating and planning for these types of changes in advance provides the opportunity to identify and execute appropriate architectural approaches to developing and optimizing the site.

Understanding That Content Is King

It has been debated since the late 1990s: whether content is king in SEO. Not only is content king for SEO purposes, but as the main driver of engagement on the Web, it is also one of the most important elements of any online effort. Development of high-quality, engaging content for your users increases the available real estate for organic search queries; provides your users reason to

enjoy your site and compels them to become customers; and invites promotion and exposure through users sharing your content online, on social media platforms and via direct links from other websites. Determining your available content assets is the first step toward leveraging them within the context of SEO. You may have a deep library of "how to" content, great testimonial or product demonstration videos, a unique photo gallery, or an awesome tool that people are interested in using. All of these content types can be invaluable in building a world-class website that does well in both search engines and the overall web ecosystem.

The content you have available to you will also affect your keyword research and your site architecture, as your content is the major source of information that search engines use to determine what your site is about. As we discussed in "Algorithm-Based Ranking Systems: Crawling, Indexing, and Ranking", you need relevant content to even be "in the game" in search (i.e., if someone searches for *hybrid efficient cars* and you don't have any content related to hybrid efficient cars, chances are good that you won't rank for that search query).

As we will discuss in Chapter 7, on-site content also affects content marketing and link development. Content marketing can be very similar to PR in that the success of your efforts is integrally related to what you are promoting (i.e., what type of content, and of what quality, are you hoping others will link to?).

Consider Site A, a site that has built a substantial, well-researched and well-written set of articles on a specific topic. However, 20 other sites out there have equally substantial sets of articles on the same topic, and many of these other sites have been in the major search engine indexes for much longer than Site A.

Site A has a content quality concern—namely, why would someone link to its articles over the articles from any of the other 20 websites? There is nothing new there. Chances are that Site A will succeed in getting some links to its articles; however, it will likely never be able to establish itself as a leader if it has nothing new or unique to offer.

To establish itself as a leader, Site A must create fresh, unique, engaging content that differentiates the site from its competitors. Perhaps it can offer a solution to a problem that no one else has been able to solve before, or perhaps it focuses on a specific vertical niche and establishes itself as a leader in that niche—for

example, by being the first to release a high-quality video series on the topic it covers.

One of the most important decisions Site A's leadership needs to make is where and how they are going to establish themselves as an authoritative resource in their industry. If they plan to make their website a major player in capturing market-related search engine traffic, this is not an optional step.

When looking at content plans and hopefully establishing an editorial calendar, it is critical to consider not only what you already have, but also what you could develop. This, of course, relates to budget. A publisher with no budget to spend on content development has few choices that she can make in her SEO strategy, whereas another publisher who has a team of in-house content developers has a myriad of options.

Segmenting Your Site's Audience

Let's not forget the audience! It is important for the SEO practitioner to understand the target audience. For example, Site A may be a website that sells sneakers. As a result, the site's developers go out and implement a brilliant campaign to rank for the "sneaker" terms they consider relevant. Being young and energetic, they focus on the way their peers search for sneakers—but what if the target audience for the sneakers Site A sells are age 50 or older? This segmentation goes to the very core of deciding what pages to create on your site, and what content to put on them.

The target audience for Site A (the over-50 crowd) may use different search terms than the younger generation to search for sneakers, which means Site A may well be bringing in search traffic from people who are not interested in its products, and not bringing in traffic from those who might be! For example, Skechers, a California-based shoe company, made the business decision to pivot from being a youth brand to targeting an older demographic with their shoes—so an SEO strategy for selling Skechers sneakers at this stage, then, would be very different from an SEO strategy for the more youth-oriented Nike sneakers.

Similar things can happen with gender, as women and men may not search for their shoes the same way. Figure 3-2 lists the top shoe-related search terms from Wordtracker for searches related to *womens shoes* and *mens shoes*.

Keyword	Volume
womens shoes	33,419
womens shoes online	4,465
womens boots shoes	2,912
womens high heel shoes	2,887
cheap womens shoes	2,740
womens boot shoes	2,680
salomon womens shoes sale	2,093
womens fashion shoes	2,084
clarks womens shoes	1,697
ebay womens shoes	1,679

Keyword	Volume
mens shoes	14,641
mens tennis shoes	2,035
mens dress shoes	1,440
mens shoes stores	1,387
buy mens shoes online	1,351
mens fashion shoes	1,217
mens clothing shoes	1,176
mens walking shoes	1,111
mens shoes online	1,049
mens sports shoes	1,007

Figure 3-2. Keyword search term variation examples by gender

As you can see in the figure, search terms used can vary significantly by gender.

Another major element to consider might be location. Searchers in Austin, Texas, may want a different version of your product than searchers in Chicago, Illinois. For that matter, because they want different products and are in different

regions, they may be looking for something specific to their region and may use different vernacular—all factors that influence a user's choice of search terms. Understanding these variables and identifying appropriate traffic targets requires extensive keyword research—yet another critical aspect of the SEO process.

Understanding Context: Market Competitiveness

A critical factor to understand before you launch your SEO efforts is the nature of the market in which you are competing. This tells you how competitive the environment is in general, and augmented with additional research, this information can tell you how competitive the SEO environment is specifically. In some markets, natural search is intensively competitive.

Market competitiveness does not mean you should give up on competing, especially if it is already the focus of your business; however, you might choose to focus your SEO efforts on less competitive terms that can still bring you many qualified leads.

To get a very quick read on keyword competitiveness, use the Google AdWords Keyword Planner to see what your cost per click (CPC) could be if you bid on your target phrase in a pay-per-click (PPC) campaign. Higher CPC prices in Google AdWords often mean that the terms are more competitive in organic search as well.

Your SEO strategy can also be influenced by your competitors' strategies, so understanding what they are doing is a critical part of the process for both SEO and business intelligence objectives. There are several scenarios you might encounter:

- The competitor discovers a unique, highly converting set of keywords.
- The competitor discovers a targeted, high-value link from an authoritative source.
- The competitor saturates a market segment, justifying your focus elsewhere.
- Weaknesses appear in the competitor's strategy, which provide opportunities for exploitation.

Understanding the strengths and weaknesses of your competition from an SEO perspective is a significant part of formulating your own SEO strategy. A final note on competitor analysis: be cautious when targeting competitor link strategies, as some of your competitors may be temporarily benefiting from link tactics that will get them penalized by the search engines (or they may already have been!).

SEO tools such as SEMRush and SearchMetrics can provide insight into your competitors' SEO performance.

SEO for Raw Traffic

Optimizing for search engines and creating thematically targeted content helps a site rank for key search queries, which typically leads to direct traffic, social sharing, and referral traffic from links as more and more people find, use, and enjoy the content you've produced. Thousands of sites on the Web leverage this traffic to serve advertising, directly monetizing the traffic sent from the engines. From banner ads to contextual services such as Google's AdSense, to affiliate and social media marketing, Internet advertising spending has become a massive industry. In October 2014, the Internet Advertising Bureau measured Internet advertising revenues for Q3 2014 at $12.4 billion, a 17% increase over Q4 2013.[1]

Here are some factors to think about when considering SEO for raw traffic:

When to employ SEO for raw traffic

Use it when you can monetize traffic without actions or financial transactions taking place on your site (usually through advertising).

Keyword targeting

Keyword targeting in this scenario can be very broad. The goal here isn't typically to select specific keywords, but rather to create high-quality content that naturally targets interesting, searched-for terms. Instead of singular optimization on specific terms, the focus is on accessibility and best practices throughout the site to earn traffic through both high-volume and long-tail queries (for more on what the *long tail* is, see Chapter 5). Concentrate efforts on great content, and use keyword-based optimization as a subsequent

application to confirm the titles, headlines, filenames, metadata, and other elements of the content you create.

Page and content creation/optimization

A shallow, highly crawlable link structure is critical to getting all of your content indexed—follow good information architecture practices (see "Creating an Optimal Information Architecture") and use intelligent, detailed category and subcategory structures to get the most benefit out of your work. You'll also need to employ good on-page optimization in `<title>` tags, headlines, internal links, and so on, and make your articles easy to share and optimized for viral spreading (see "Root Domains, Subdomains, and Microsites" and "Optimization of Domain Names/URLs" for more on this topic).

SEO for Ecommerce Sales

One of the most direct monetization strategies for SEO is driving relevant traffic to an ecommerce shop to boost sales. Search traffic is among the highest quality traffic on the Web, primarily because a search user has expressed a specific interest through his query, and when this matches a service, product, or brand a website carries, conversion rates are often extremely high. Here are some factors to think about when considering SEO for ecommerce sales:

When to employ SEO for ecommerce sales

Use it when you have products/services that are directly for sale on your website.

Keyword targeting

PPC advertising is an excellent way to test the efficacy and potential ROI of keyword targets. Find those that have reasonable traffic and convert well, and pursue them further. You'll often find that the more specific the query is—brand-inclusive, product-inclusive, and so on—the more likely visitors are to make the purchase. The best use of this tactic is for generic terms that you will find harder to win on than brand/company named terms, so you can decide if they are worth the effort.

Quality content creation and optimization

Your site will need to provide interesting, unique, and accessible content for both users and search engines in order to begin gaining traction in organic search. The links and social sharing that come from such content are highly influential in increasing overall site traffic, as well as improving organic search performance. Creating link-worthy, deeply engaging content should be the primary focus of any website seeking search and referral traffic, and should be incorporated into both the content development and SEO strategies. Manual link building is always an option, but scalable strategies that leverage a community or customers can be equally, or even more, valuable. Remember: content that keeps users on the page, instead of prompting them to click away rapidly, is a signal of quality and also serves to increase the value of traffic to the page (improving page monetization through advertising).

SEO for Mindshare and Branding

An equally powerful application of SEO is to use it for branding purposes. Bloggers, social media platforms, community websites, content producers, news outlets, and dozens of other web publishing outlets have found tremendous value in appearing atop the SERPs and using the resulting exposure to bolster their brand recognition and authority.

The process is fairly simple, much like the goal in traditional advertising of ad repetition to enter the buyer's consideration set. Online marketers have observed that being at the top of the search rankings around a particular subject has a positive impact on traffic, consideration, and perceived authority. A 2012 study by Conductor, Inc., highlights the significant branding value of a website appearing in organic search results.[2]

Here are some factors to think about when considering SEO for mindshare and branding:

When to employ SEO for mindshare/branding

Use it when branding or communicating a message is your goal.

Keyword targeting

A keyword focus is less critical here—you'll likely have a few broad terms that receive the high traffic you want, but the long tail may be far more achievable and the better target. Focus on keywords that are going to bring you visitors who are likely to be interested in and remember your brand.

Page and content creation/optimization

Make your site content easily crawlable by search engines, use intelligent linking structures to make the site intuitively navigable for users, and implement SEO best practices.

SEO for Lead Generation and Direct Marketing

Although online lead generation is less direct than an ecommerce transaction, it is arguably just as valuable and important for building customers, revenue, and long-term value. Millions of search queries have commercial intents that can't be (or currently aren't) fulfilled directly online. These can include searches for services such as legal consulting, contract construction, commercial loan requests, alternative energy providers, or virtually any service or product people source via the Internet.

Here are some factors to think about when considering SEO for lead generation and direct marketing:

When to employ SEO for lead generation and direct marketing

Use it for non-ecommerce products/services/goals that you want users to accomplish on your site, or for which you are hoping to attract inquiries and/or direct contact over the Web.

Keyword targeting

As with ecommerce, choose phrases that have measurable traffic and have previously converted well in PPC campaigns. In a "considered sale" or enterprise-scale business decision for users, include long-tail keywords that might pertain to customer research activities or pain points, to educate and persuade via content. Avoid arcane jargon that isn't typically understood

outside your company.

Page and content creation/optimization

Although you might think it would be easier to rank high in the SERPs for lead-generation programs than for ecommerce, it is often equally challenging. You'll need a solid combination of content development and on-site optimization to be competitive in the more challenging arenas.

SEO for Reputation Management

Your name—whether personal or corporate—is your identity, so establishing and maintaining the reputation associated with that identity is generally of great interest.

Imagine that you search for your brand name in a search engine, and a prominent result is a web page that is highly critical of your organization.

SEO for reputation management involves, in part, the process for neutralizing negative mentions of your name in the SERPs. In this type of SEO project, you would strive to occupy additional spots in the top 10 results to push the critical listing lower, and hopefully off the first page.

SEO enables this process through both content creation and promotion via link development, as well as through optimization of content on third-party platforms, such as Pinterest, Facebook, and LinkedIn. Although reputation management is among the most challenging of SEO tasks (primarily because you are optimizing many results for a query rather than one), demand for these types of services is rising as more and more companies become aware of the issue.

Here are some factors to think about when considering SEO for reputation management:

When to employ SEO for reputation management

Use it if you're trying to either protect your brand from negative results appearing on page one or push down already existing negative content.

Keyword targeting

Chances are, this is very easy—the keyword you are targeting is a person's name, your brand name, or some common variant (and you already know what it is). You might want to use keyword research tools just to see whether there are popular variants you're missing.

Page and content creation/optimization

Unlike the other SEO tactics, reputation management involves optimizing pages on many different domains to demote negative listings. This involves using social media profiles and other third-party platform pages, public relations, press releases, and links from networks of sites you might own or control, along with classic optimization of internal links and on-page elements. It is certainly among the most challenging of SEO practices, especially in Google, where the use of the *query deserves diversity* (QDD) algorithm can mean you have to work much harder because of how it favors diverse content.

On the topic of negative search results in Google, in May 2014 the Luxembourg-based European Union Court of Justice ruled that people can ask Google to remove search results containing information about them, with the new, EU-specific "right to be forgotten."[3] While some view this as a victory for privacy laws, others view it as censorship. Time will tell how this plays out internationally and on a case-by-case basis for individuals.

SEO for Ideological Influence

For those seeking to sway public (or private) opinion about a particular topic, SEO can be a powerful tool. By promoting ideas and content within the search results for queries likely to be made by those seeking information about a specific topic, you can influence the perception of even very large groups. Politicians and political groups and individuals are the most likely employers of this tactic, but it can certainly be applied to any subject, from the theological to the technical or civic.

Some factors to consider when using SEO for ideological influence:

When to employ

Use this tactic when you need to change minds or influence decisions/thinking around a subject—for example, a group of theoretical physicists attempting to get more of their peers to consider the possibility of alternative universes as a dark matter source.

Keyword targeting

Generally in these types of campaigns, you probably know the primary keywords you're chasing and can use keyword research query expansion to find others. Consider inclusion of both "pro" and "con" phrases for maximum visibility in a debate. For example, *Obamacare* typically is considered a conservative or critical label, while *Affordable Care Act* is the official label and is more commonly used by supporters.

Page and content creation/optimization

This is classic SEO, but with a twist. Because you're engaging in ideological warfare in the SERPs, chances are you've got allies you can rally to the cause. Leverage your combined links and content to espouse your philosophical preferences.

Advanced Methods for Planning and Evaluation

There are many methodologies for business planning. One of the better-known ones is the SWOT (strengths, weaknesses, opportunities, threats) analysis. There are also methodologies for ensuring that the plan objectives are the right type, such as the SMART (specific, measurable, achievable, realistic, timelined) plan. We will take a look at both of these in the context of SEO.

SWOT Analysis

Sometimes you need to get back to the basics and carry out a simple evaluation of where you are in the marketplace, as well as where you would like to be; for this, a SWOT analysis is a great starting point. It creates a grid from which to work and is very simple to execute.

As you can see from the SWOT chart in Figure 3-3, strengths and weaknesses

usually stem from internal (on-site, business operational, business resource) sources, whereas opportunities and threats are from external sources.

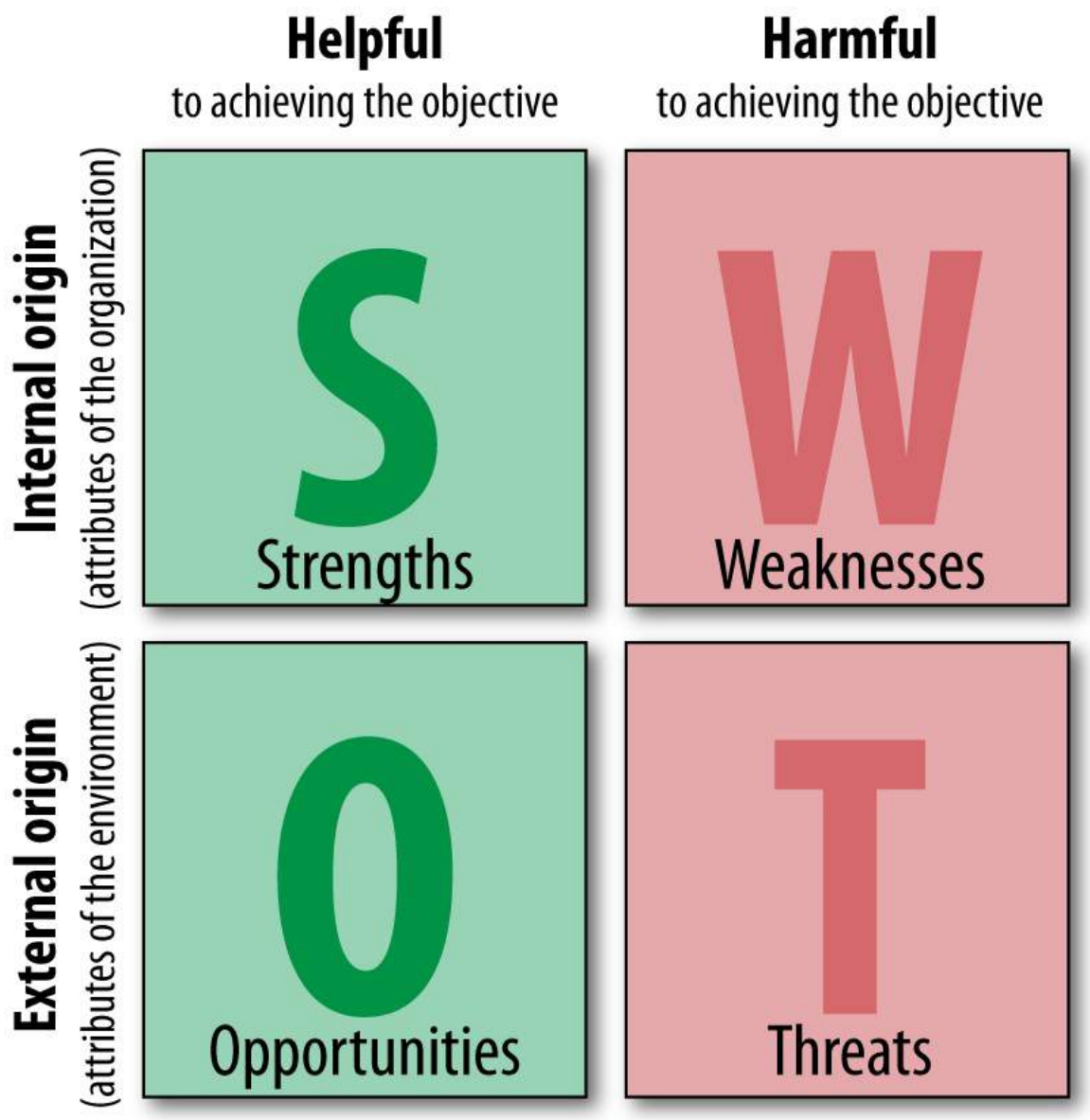


Figure 3-3. Example SWOT chart

Where does SEO fit in here? To explore this, we'll use an example. Take Business X. It has a website that was built on WordPress, makes use of category tagging, adds at least one page of content every two days, and has excellent knowledge of its industry. Its domain name isn't ideal —*Businessnameandkeyword.com*—but it is decent.

Business X does not get much traffic from search engines, but its rival, Business Y, does because Business Y has had its website up for a long period of time and received some great links along the way. Business Y doesn't have any SEO plan and relies on its main page to bring in all of its search traffic. This is because Business Y has a keyword-rich domain name and people have used those keywords in their links to Business Y's website (giving it keyword-rich anchor text), and because of its longevity on the Web.

There aren't a lot of target search queries; in fact, there are fewer than 50,000 searches per month for the core set of keywords. Business X's site ranks on the second page of Google results, whereas Business Y is ranked #3, with Wikipedia and About.com taking up the top two positions.

Neither of the businesses is spending money on PPC (paid search) traffic, and the niche doesn't have much room for other entrants (there may be 10–15 competitors). Both sites have similar link authority in terms of strengths and numbers. The businesses deal in impulse purchases; the products evoke strong emotions.

Figure 3-4 shows what the SWOT for Business X might look like.

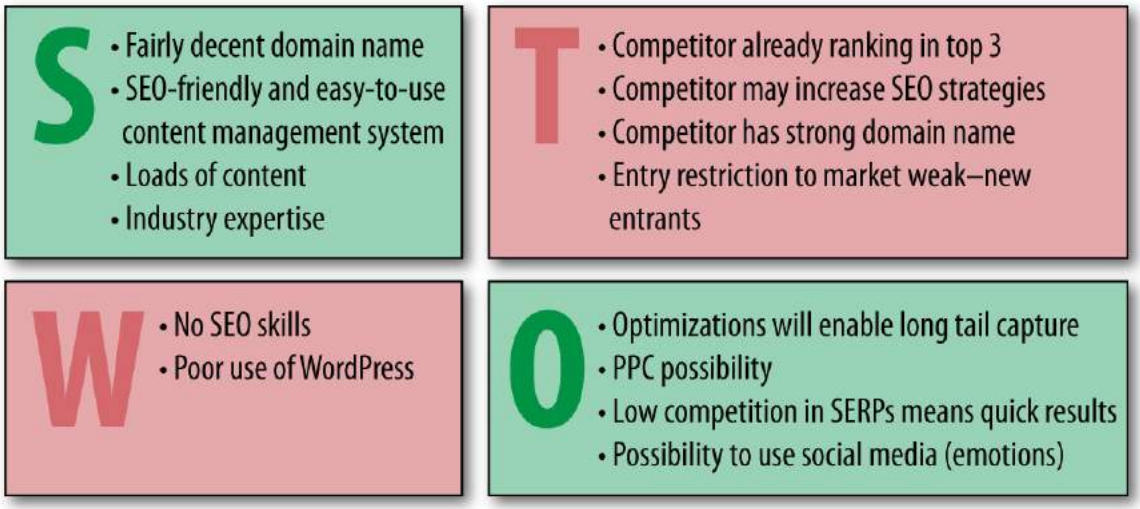


Figure 3-4. Sample SWOT chart data for Business X

The preceding analysis suggests where Business X can get some quick wins for its site, as well as where the priorities are. It also forms a great starting point for a long-term strategy and tactical maneuvers. This example is simplistic, but it

illustrates how instructive a fleshed-out SWOT can be. It does require you to have analyzed your site, your main competitor(s), your target keywords, and the SERPs.

SWOT Guidelines

Identifying *strengths* is typically one of the easier objectives:

- What sources of traffic are working well (converting on established site goals) for your site/business?
- Which projects/properties/partnerships are driving positive momentum toward traffic/revenue goals?
- Which of your content sections/types produces high traffic, conversions, and ROI?
- What changes have you made historically that produced significant value?

Determining the *weaknesses* can be tougher (and requires emotional separation from the data):

- What content is currently driving low levels of search/visitor traffic?
- Which changes that were intended to produce positive results have shown little or no value?
- Which traffic sources are underperforming or underdelivering?
- What projects/properties/partnerships are being leveraged poorly, or not at all?

Parsing *opportunities* requires a combination of strength and weakness analysis. You want to find areas that are doing well but have room to expand, as well as those that have yet to be explored:

- What brainstormed but undeveloped or untested projects/ideas can have a significant, positive impact?
- What traffic sources currently sending good-quality traffic could be expanded

to provide more value?

- What areas of weakness have direct paths to recovery?
- Which website changes have had positive results? Can these be applied more rigorously or to other areas for increased benefit?
- What new markets or new content areas are potentially viable or valuable for expansion?
- What sources of new content or new links have yet to be tapped?
- What third-party platforms (social media, content curators, etc.) can be utilized to expand reach and increase engagement?

Determining *threats* can be the most challenging of the tasks. You'll need to combine creative thinking with an honest assessment of your weaknesses and your competitors' strengths, and consider the possibilities of macro-events that could shape your website/company's future:

- In your areas of weakness, which players in your market (or other, similar markets) are strong? How have they accomplished this?
- What shifts in human behavior, web usage, or market conditions could dramatically impact your business/site? (For example, consider the "what if people stopped searching and instead navigated the Web in different ways?" perspective. It is a bit "pie in the sky," but we have already seen Craigslist make classifieds obsolete, and have witnessed Facebook start to take advertising market share from the search engines.)
- Which competitors have had the most success in your arena? How have they accomplished this? Where do they intersect with your business/customers?
- Are there any strategies implemented by startups in similar businesses that have had massive success in a particular arena that could be dangerous to your business if they were replicated in your market?

Conducting SWOT analysis from a web marketing and SEO perspective is

certainly one of the most valuable first steps you can take as an organization poised to expend resources. If you haven't taken the time to analyze the landscape from these bird's-eye-view perspectives, you might end up like a great runner who's simply gone off the course—sure, you'll finish fast, but where will it take you?

SMART Objectives

Every company is unique, so naturally its challenges are unique. Even a second SEO initiative within the same company will not be the same as the first. Your first SEO efforts will have changed things, creating new benchmarks, new expectations, and different objectives. Thus, each SEO effort is a new endeavor.

One way to start a new project is to set SMART objectives. Let's look at how to go about doing that in the world of SEO.

Specific objectives are important. It is easy to get caught up in the details of the plan and lose sight of the broader site objectives. You may think you want to rank #1 for this phrase or that, but in reality what you want is more granular than that: more leads, more page views, more customers. Perhaps you don't even need more customers from organic search, but you want higher sales volumes, so in fact having the same number of orders but with a higher average order value would meet your objectives better.

Measurable objectives are essential if you are to manage the performance in meeting them—you can't manage what you can't measure. SEO practitioners have to help their clients or organizations come to grips with analytics, and not just the analytics software, but the actual processes of how to gather the data, how to sort it, and most importantly, how to use it to make informed decisions.

Achievable objectives are ones that can be accomplished with the available resources. You could decide to put a man on Mars next year, for example, but it is just too big an undertaking to be feasible. You can be ambitious, but it is important to pick goals that can be met. You cannot possibly sell to more people than exist in your market. There are limits to markets, and at a certain point the only growth can come from opening new markets, or developing new products for the existing market.

Aside from basic business achievability, there are also limits to what can rank at

#1 for a given search query. The search engines want the #1 result to be the one that offers the most value for users, and unless you are close to having the website that offers the most value to users, it may be unreasonable to expect to get to that position, or to maintain it if you succeed in getting there.

Realistic objectives are about context and resources. It may be perfectly achievable to meet a certain objective, but only with greater resources than may be presently available. Even a top ranking on the most competitive terms around is achievable for a relevant product, but it is a realistic goal only if the resources required for such an effort are available.

Timelined objectives are the final part of the SMART methodology. If there is no timeline, no project can ever fail, because it can't run out of time. SEO generally tends to take longer to implement and gather momentum than a paid search advertising campaign. It is important that you set milestones and deadlines so that you can manage expectations and course-correct.

"We want to rank at #1 for loans" is not a SMART objective. It doesn't identify the specific reason why the company thinks a #1 ranking will help it. It doesn't have a timeline, so there is no way to fail. It doesn't state an engine on which to be #1, so there's a guaranteed argument if the intention is to rank well on both Google and Bing, but the result is only high rankings on Bing.

"We want to increase approved loan applications generated by organic search by 30% over six months" is a far better objective. There is a deadline, and the company can certainly gauge progress toward the specific objective. The company can look at its current market share and the resources committed to see whether this is an achievable and realistic goal.

Conclusion

To bring this all together successfully, your objectives, strategies, and tactics need to be aligned, and they need to take into account your market, your business, and the competition. Don't spread yourself too thin. Remember to ask yourself the tough questions, such as:

- Does your company need direct sales, traffic, branding, or some combination of these?

- Are there specific influencers you're trying to reach with a message?
- Is the organization/brand subject to potentially negative material that needs to be controlled/mitigated?
- Do you have products/services you sell, either directly over the Web or through leads established online?
- Do you have the resources to develop new, unique, and interesting content?

Getting the answers won't be easy, but it will be worth the effort!

[1] Interactive Advertising Bureau, "Q3 2014 Internet Advertising Revenues Hit $12.4 Billion, Making it the Highest Quarter on Record," December 18, 2014, *http://bit.ly/q3_2014_ad_revenues*.

[2] Nathan Safran, "Search Reimagined: The Branding Value of Page One," Conductor Blog, July 24, 2012, *http://bit.ly/search_reimagined*.

[3] Reuters, "Europe's top court: people have right to be forgotten on Internet," May 13, 2014, *http://bit.ly/right_to_be_forgotten*.

CHAPTER FOUR

SEO Implementation: First Stages

SEO efforts require forethought and planning to obtain the best results, and a business's SEO strategy is ideally incorporated into the planning stages of a website development or redevelopment project, as well as into ongoing web development efforts. Website architecture (including the selection of a content management system, or CMS), the overall marketing plan (including branding objectives), content development efforts, and more are affected by SEO.

In this chapter, we will discuss several aspects of how SEO efforts for both desktop and mobile generally begin, including:

- Putting together an SEO strategy
- Performing a technical SEO audit of site versions
- Setting a baseline for measuring results and progress

The Importance of Planning

As discussed in Chapter 3, your SEO strategy should be incorporated into the site planning process long before your site goes live, and your process should be well outlined before you make even the most basic technology choices, such as deciding on your hosting platform and content management system. However, this is not always possible—and in fact, more often than not, SEO efforts in earnest often begin well after a site has been launched and in use for some time.

In all scenarios, there are major components to any SEO strategy that need to be addressed long before you craft your first HTML `<title>` tag.

Identifying the Site Development Process and Players

Before you start the SEO process, it is important to identify your target audience, your message, and how your message is relevant. There are no web design tools or programming languages that tell you these things. Your company must have an identified purpose, mission, and vision—and an accompanying voice and tone

for communicating these to the outside world. Understanding these elements first helps drive successful marketing and branding efforts, whether the channel is email marketing, display, or organic search.

Your SEO team needs to be cross-functional and multidisciplinary, consisting of the team manager, the technical team, the creative team, the data and analytics team (if you have one), and the major stakeholders from marketing and website development. In a smaller organization or a startup environment, you and your team members may have to wear more than one of these hats (we never said SEO was easy!).

The SEO team leader wants to know who the website's target audience is. What does the marketing team know about them? How did we find them? What metrics will we use to track them? All of this is key information that should have an impact on various aspects of the project's technical implementation. This ties in to the PR messaging presented to the media to entice them into writing and talking about it. What message do they want to deliver? You have to mirror that message in your content. If they say you're relevant to organic cotton clothes, but your project plan says you're relevant to yoga attire, the whole project is in trouble. When you're creating visibility, the people who build up your brand have to see a clear, concise focus in what you do. If you provide them with anything less, their interest in your business will be brief, and they'll find someone else to talk about.

The technical, content development, and creative teams are responsible for jointly implementing the majority of an SEO strategy. Ongoing feedback is essential because the success of your strategy is determined solely by whether you're meeting your goals. A successful SEO team understands all of these interactions and is comfortable relying on each team member to do his part, and establishing good communication among team members is essential.

And even if you are a team of one, you still need to understand all of these steps, as addressing all aspects of the marketing problem (as it relates to SEO) is a requirement for success.

Development Platform and Information Architecture

Whether you're working with an established website or not, you should plan to research the desired site architecture (from an SEO perspective) as a core element of your SEO strategy. This task can be divided into two major components: technology decisions and structural decisions.

Technology Decisions

SEO is a technical process, and as such, it impacts major technology choices. For example, a CMS can facilitate (or possibly undermine) your SEO strategy: some platforms do not allow you to write customized titles and meta descriptions that vary from one web page to the next, while some create hundreds (or thousands) of pages of duplicate content (not good for SEO!). For a deep dive into the technical issues you'll need to familiarize yourself with in order to make the right technology decisions for your SEO needs, turn to Chapter 6. The technology choices you make at the outset of developing your site and publishing your content can make or break your SEO efforts—and it is best to make the right choices in the beginning to save yourself headaches down the road.

As we outlined previously in this chapter, your technology choices can have a major impact on your SEO results. The following is an outline of the most important issues to address at the outset:

Dynamic URLs

Dynamic URLs are URLs for dynamic web pages (which have content generated "on the fly" by user requests). These URLs are generated in real time as the result of specific queries to a site's database—for example, a search for *leather bag* on Etsy results in the dynamic search result URL *https://www.etsy.com/search?q=leather%20bag.* However, Etsy also has a static URL for a static page showing leather bags at *https://www.etsy.com/market/leather_bag*.

Although Google has stated for some time that dynamic URLs are not a problem for the search engine to crawl, it is wise to make sure your dynamic URLs are not "running wild" by checking that your CMS does not render your pages on URLs with too many convoluted parameters. In addition, be sure to make proper use of `rel="canonical"`, as outlined by Google

(*http://bit.ly/canonical_urls*).

Finally, while dynamic URLs are crawlable, don't overlook the value of static URLs for the purpose of controlling your URL structure for brevity, descriptiveness, user-friendliness, and ease of sharing.

Session IDs or user IDs in the URL

It used to be very common for a CMS to track individual users surfing a site by adding a tracking code to the end of the URL. Although this worked well for this purpose, it was not good for search engines, because they saw each URL as a different page rather than a variant of the same page. Make sure your CMS does not ever serve up session IDs. If you are not able to do this, make sure you use `rel="canonical"` on your URLs (what this is, and how to use it, is explained in Chapter 6).

Superfluous parameters in the URL

Related to the preceding two items is the notion of extra characters being present in the URL. This may bother search engines, and it interferes with the user experience for your site.

Links or content based in Flash

Search engines often cannot see links and content implemented with Flash technology. Have a plan to expose your links and content in simple HTML text, and be aware of Flash's limitations.

Content behind forms (including pull-down lists)

Making content accessible only after the user has completed a form (such as a login) or made a selection from an improperly implemented pull-down list is a great way to hide content from the search engines. Do not use these techniques unless you want to hide your content!

Temporary (302) redirects

This is also a common problem in web server platforms and content

management systems. The 302 redirect blocks a search engine from recognizing that you have permanently moved the content, and it can be very problematic for SEO, as 302 redirects block the passing of PageRank. Make sure the default redirect your systems use is a 301, or learn how to configure it so that it becomes the default.

All of these are examples of basic technology choices that can adversely affect your chances for a successful SEO project. Do not be fooled into thinking that SEO issues are understood, let alone addressed, by all CMS vendors out there—unbelievably, some are still very far behind the SEO curve. It is also important to consider whether a "custom" CMS is truly needed when many CMS vendors are creating SEO-friendly systems—often with much more flexibility for customization, a broader development base, and customizable, SEO-specific "modules" that can quickly and easily add SEO functionality. There are also advantages to selecting a widely used CMS, including portability in the event that you choose to hire different developers at some point.

Also, do not assume that all web developers understand the SEO implications of what they develop. Learning about SEO is not a requirement to get a software engineering degree or become a web developer (in fact, there are still very few college courses that adequately cover SEO). It is up to you, the SEO expert, to educate the other team members on this issue as early as possible in the development process.

Structural Decisions

One of the most basic decisions to make about a website concerns internal linking and navigational structures, which are generally mapped out in a site architecture document. What pages are linked to from the home page? What pages are used as top-level categories that then lead site visitors to other related pages? Do pages that are relevant to each other link to each other? There are many, many aspects to determining a linking structure for a site, and it is a major usability issue because visitors make use of the links to surf around your website. For search engines, the navigation structure helps their crawlers determine what pages you consider the most important on your site, and it helps them establish the relevance of the pages on your site to specific topics.

This section outlines a number of key factors that you need to consider before launching into developing or modifying a website. The first step will be to obtain a current site information architecture (IA) document for reference, or to build one out for a new site. From here, you can begin to understand how your content types, topics, and products will be organized.

Target keywords

Keyword research is a critical component of SEO. What search terms do people use when searching for products or services similar to yours? How do those terms match up with your site hierarchy? Ultimately, the logical structure of your pages should match up with the way users think about products and services like yours. Figure 4-1 shows how this is done on Amazon.com.

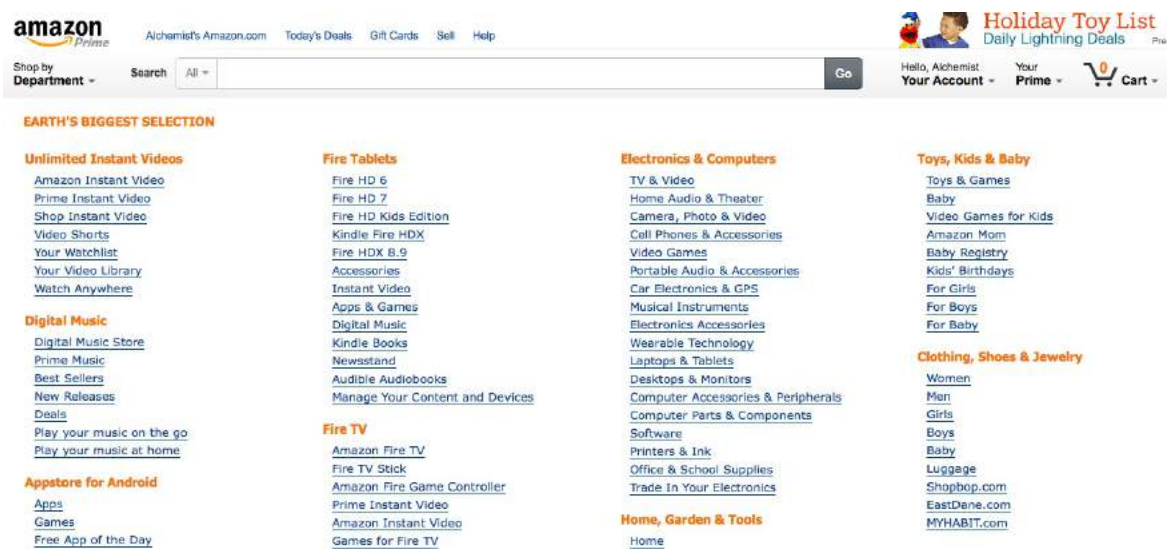


Figure 4-1. A well-thought-out site hierarchy

Cross-link relevant content

Linking between articles that cover related material can be very powerful. It helps the search engine ascertain with greater confidence how relevant a web page is to a particular topic. This can be extremely difficult to do well if you have a massive ecommerce site, but Amazon handles it nicely, as shown in Figure 4-2.

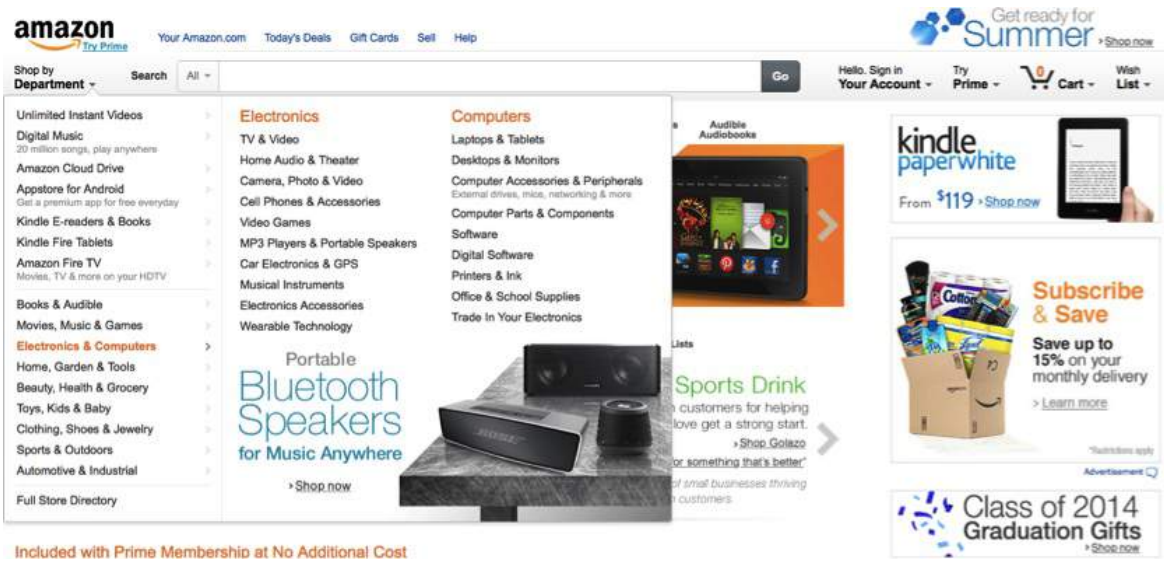


Figure 4-2. Product cross-linking on Amazon

The “Frequently Bought Together” and “Customers Who Bought This Item Also Bought” sections are brilliant ways to group products into categories that establish the relevance of the page to certain topic areas, as well as to create links between relevant pages.

In the Amazon system, all of this is rendered on the page dynamically, so it requires little day-to-day effort on Amazon’s part. The “Customers Who Bought...” data is part of Amazon’s internal databases, and the “Tags Customers Associate...” data is provided directly by the users themselves.

Of course, your site may be quite different, but the lesson is the same: you want to plan on having a site architecture that will allow you to cross-link related items.

Use anchor text, intuitively

Anchor text has generally been one of the golden opportunities of internal linking, and exact-match keyword anchor text was generally the protocol for internal linking for many years. However, in these days of aggressive anchor text abuse (and crackdown by the search engines), while keyword-infused anchor text in internal links is still often the most intuitive and user-friendly, we generally advocate for a more broad-minded approach to crafting internal anchor text. Use descriptive text in your internal links and avoid using irrelevant text

such as “More” or “Click here.” Try to be as specific and contextually relevant as possible and include phrases when appropriate within your link text. For example, as a crystal vendor, you might use “some of our finest quartz specimens” as anchor text for an internal link, versus “quality quartz crystals.” Make sure that the technical, creative, and editorial teams understand this approach, as it will impact how content is created, published, and linked to within your site.

Use breadcrumb navigation

Breadcrumb navigation is a way to show the user where he is in the navigation hierarchy. Figure 4-3 shows an example from PetSmart.

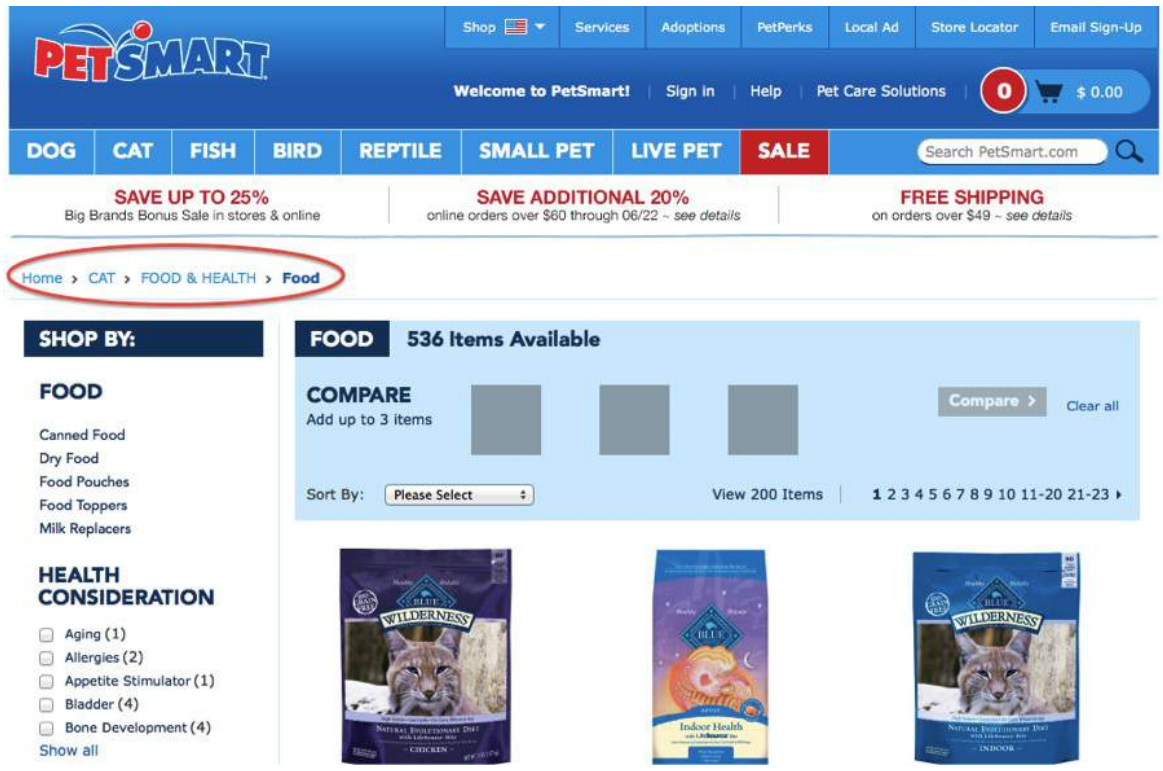


Figure 4-3. Breadcrumb bar on PetSmart’s website

This page is currently four levels down from the home page. Also, note how the anchor text in the breadcrumb is keyword-rich, as is the menu navigation on the left. This is helpful to both users and search engines.

Minimize link depth

Search engines (and people) look to the site architecture for clues as to what pages are most important. A key measurement is how many clicks from the home page it takes a person, and a search engine crawler, to reach a page. A page that is only one click from the home page is clearly important, while a page that is five or six clicks away is not nearly as influential. In fact, the search engine spider may never even find such a page, depending in part on the site's link authority.

Standard SEO advice is to keep the site architecture as flat as possible, to minimize clicks from the home page to important content. The bottom line is that you need to plan out a site structure that is as flat as you can reasonably make it without compromising your user experience.

In this and the preceding sections, we outlined common structural decisions that you need to incorporate into your SEO strategy prior to implementation. There are other considerations, such as how to make your efforts scale across a very large site (thousands of pages or more). In such a situation, you cannot feasibly review every page one by one.

Mobile Sites and Mobile Apps

If you are building a website, you need to build a mobile version if you want to take full advantage of organic search through SEO—and depending on your business, you may benefit from developing a mobile app as well. The main consideration regarding your site's mobile version is whether to host it on the same or separate URLs as your desktop version—and, if you're utilizing the same URLs, whether to choose responsive design or dynamic serving (a.k.a. adaptive design), discussed in detail in Chapter 10.

Single-Page Applications

Single-page applications (SPAs) are web applications that use AJAX and HTML5 to load a single HTML page in a web browser, and then dynamically update that page's content as the user interacts with the app. The majority of the work in loading page content, or *rendering*, is done on the client side (as opposed to the server side), which makes for a fast and fluid user experience and minimized page loads, often while the page URL remains the same. Commonly used frameworks for SPA development include Angular.js, Backbone.js, and

Ember.js, which are used by many popular applications including Virgin America, Twitter, and Square, respectively.

One of the main issues to address when you're building a site with one of these frameworks is URL *crawlability*—namely, ensuring that the search engines can crawl your site's URLs to access your site content. It is important that you have a publishing system that allows you to customize URLs to remove the # or #! (hashbang) from the URL, and to create user-friendly, bookmarkable, back-clickable URLs. There are various methods that developers can use to implement search-and user-friendly URLs, with the two most recent being `window.location.hash` and HTML5's `history.pushState`—both of which have advantages and disadvantages depending on your site and user objectives. An informative discussion of these two options can be found at StackOverflow.

Angular.js: Making it SEO-friendly

Detailing SEO-friendly development for all SPA frameworks is beyond the scope of this volume; however, the information below provides guidelines and workarounds specific to Angular.js thanks to SEO/Angular.js developer Jeff Whelpley, who assisted coauthor Jessie Stricchiola in addressing this issue for Alchemist Media client Events.com.

Server rendering versus client-Only rendering

If the website owner doesn't care about achieving top ranking, and the only goal is getting indexed—and, the only goal is to get indexed by Google, not any other search engine—then it is OK to do client-only rendering with Angular. Google has gotten very good recently at indexing client rendered HTML. If you go down this route, you will need to do the following:

1. Enable push state in Angular so you get pretty URLs without the hash.

2. Implement UI Router or the new Component Router in Angular so you can map URLs to pages.

3. Follow all normal SEO best practices for page titles, URLs, content, *etc.* Nothing changes here.

4. Optimize the heck out of the initial page load—a major mistake many make is thinking initial page load time for client-rendered apps doesn't

matter, but it does!

While the above approach will work for Google indexing, you will perform better in organic search with server rendering. The reasons for this are:

- Google is good at client rendering, but not perfect.
- Other search engines are really not good at it.
- Things like Facebook or Twitter link previews will not work.
- It is much easier to make server rendering fast than it is to make the initial load for client rendering fast—and that makes a big difference.

The options for server rendering are as follows:

- Implement in PHP or another language. This will work but requires that you duplicate all pages, which is generally not feasible unless you have a small/simple site.
- Use Prerender.io or a similar service. This works but can get expensive for larger sites, can be tricky to set up, and you need to be OK with long page cache times (i.e., server pages are one day old).
- Build a custom solution off the Jangular library. Although this works, it does require a lot of heavy lifting.
- Wait for Angular 2.0 (possibly to be released end of 2015)

As of July 2015, Jeff is working with the Angular core team to bake server rendering into Angular 2.0. It won't be ready until later in the year, but it will be by far the best option out of all these. For developers dealing with this issue prior to the 2.0 roll-out who don't have the resources for a comprehensive fix, it is always possible to implement the client routing solution in the interim and align the Angular app so it can easily be upgraded to 2.0 once it is ready.

NOTE

A special thanks to Jeff Whelpley for his contributions to the Angular.js portion of this chapter.

Auditing an Existing Site to Identify SEO Problems

Auditing an existing site is one of the most important tasks that SEO professionals encounter. SEO is still a relatively new field, and many of the limitations of search engine crawlers are nonintuitive. In addition, many web developers, unfortunately, are still not well versed in SEO. Even more unfortunately, some stubbornly refuse to learn, or worse still, have learned the wrong things about SEO! This includes those who have developed CMS platforms, so there is a lot of opportunity to find problems when you're conducting a site audit. While you may have to deal with some headaches in this department (trust us, we still deal with this on a regular basis), your evangelism for SEO, and hopefully support from key stakeholders, will set the stage for an effective SEO strategy.

Elements of an Audit

Your website needs to be a strong foundation for the rest of your SEO efforts to succeed. An SEO site audit is often the first step in executing an SEO strategy. Both your desktop and mobile site versions need to be audited for SEO effectiveness. The following sections identify what you should look for when performing a site audit.

Page load time

Is the page load time excessive? Too long a load time may slow down crawling and indexing of the site, and can virtually eliminate your site from competitiveness in mobile search.

Mobile-friendliness

Your site should have a fast, mobile-friendly version that is served to mobile devices.

Usability

Although this may not be seen as a direct SEO issue, it is a very good place to

start. Usability affects many factors, including conversion rate as well as the propensity of people to link to a site.

Accessibility/spiderability

Make sure the site is friendly to search engine spiders (discussed in more detail in "Making Your Site Accessible to Search Engines" and "Creating an Optimal Information Architecture").

Search engine health checks

Here are some quick health checks:

- Perform a *site:<yourdomain.com>* search in the search engines to check how many of your pages appear to be in the index. Compare this to the number of unique pages you believe you have on your site. Also, check indexation numbers in your Google Search Console and Bing Webmaster Tools accounts.

- Check the Google cache to make sure the cached versions of your pages look the same as the live versions.

- Check to ensure major search engine Search Console and Webmaster Tools accounts have been verified for the domain (and any subdomains, for mobile or other content areas). Google and Bing currently offer site owner validation to "peek" under the hood of how the engines view your site.

- Test a search on your brand terms to make sure you are ranking for them (if not, you may be suffering from a penalty; be sure to check your associated Search Console/Webmaster Tools accounts to see if there are any identifiable penalties, or any other helpful information).

Keyword health checks

Are the right keywords being targeted? Does the site architecture logically flow from the way users search on related keywords? Does more than one page target the same exact keyword (a.k.a. *keyword cannibalization*)? We will discuss these items in "Keyword Targeting".

Duplicate content checks

The first thing you should do is make sure the non-*www* versions of your pages (i.e., *http://yourdomain.com*) 301-redirect to the *www* versions (i.e., *http://www.yourdomain.com*), or vice versa (this is often called the *canonical redirect*). While you are at it, check that you don't have *https:* pages that are duplicates of your *http:* pages. You should check the rest of the content on the site as well.

The easiest way to do this is to take unique text sections from each of the major content pages on the site and search on them in Google. Make sure you enclose the string inside double quotes (e.g., *"a phrase from your website that you are using to check for duplicate content"*) so that Google will search for that exact string. If you see more than one link showing in the results, look closely at the URLs and pages to determine why it is happening.

If your site is monstrously large and this is too big a task, make sure you check the most important pages, and have a process for reviewing new content before it goes live on the site.

You can also use search operators such as *inurl:* and *intitle:* (refer back to Table 2-1 for a refresher) to check for duplicate content. For example, if you have URLs for pages that have distinctive components to them (e.g., *1968-mustang-blue* or *1097495*), you can search for these with the *inurl:* operator and see whether they return more than one page.

Another duplicate content task to perform is to make sure each piece of content is accessible at only one URL. This probably trips up big commercial sites more than any other issue. The problem is that the same content is accessible in multiple ways and on multiple URLs, forcing the search engines (and visitors) to choose which is the canonical version, which to link to, and which to disregard. No one wins when sites fight themselves; if you have to deliver the content in different ways, rely on cookies so that you don't confuse search engine spiders.

URL checks

Make sure you have clean, short, descriptive URLs. *Descriptive* means keyword-rich but not keyword-stuffed (e.g., *site.com/outerwear/mens/hats* is keyword-rich; *site.com/outerwear/mens/hat-hats-hats-for-men* is keyword-stuffed!). You don't want parameters appended (have a minimal number if you must have any), and you want them to be simple and easy for users (and search engine spiders) to

understand.

HTML <title> tag review

Make sure the `<title>` tag on each page of the site is unique and descriptive. If you want to include your company brand name in the title, consider putting it at the end of the `<title>` tag, not at the beginning, as placing keywords at the front of a page title (generally referred to as *prominence*) brings ranking benefits. Also check to ensure the `<title>` tag is fewer than 70 characters long, or 512 pixels wide.

Content review

Do the main pages of the site have enough text content to engage and satisfy a site visitor? Do these pages all make use of header tags? A subtler variation of this is making sure the number of pages with little content on the site is not too high compared to the total number of pages on the site.

Meta tag review

Check for a meta `robots` tag on the pages of the site. If you find one, you may have already spotted trouble. An unintentional `noindex` or `nofollow` value (we define these in "Content Delivery and Search Spider Control") could adversely affect your SEO efforts.

Also make sure every page has a unique meta description. If for some reason that is not possible, consider removing the meta description altogether. Although the meta description tags are generally not a direct factor in ranking, they may well be used in duplicate content calculations, and the search engines frequently use them as the description for your web page in the SERPs; therefore, they can affect click-though rate.

Sitemaps file and robots.txt file verification

Use the Google Search Console "Robots.txt fetch" to check your *robots.txt* file. Also verify that your `Sitemap` file is correctly identifying all of your site pages.

URL redirect checks

Check all redirects to make sure the right redirect is in place, and it is pointing to the correct destination URL. This also includes checking that the canonical

redirect is properly implemented. Use a server header checker such as Redirect Check or RedirectChecker.org, or when using Firefox, install the browser extension Redirect Check Client.

Not all URL redirect types are created equal, and some can create problems for SEO. Be sure to research and understand proper use of URL redirects before implementing them, and minimize the number of redirects needed over time by updating your internal linking and navigation so that you have as few redirects as possible.

Internal linking checks

Look for pages that have excessive links. As discussed earlier, make sure the site makes intelligent use of anchor text in its internal links. This is a user-friendly opportunity to inform users and search engines what the various pages of your site are about. Don't abuse it, though. For example, if you have a link to your home page in your global navigation (which you should), call it "Home" instead of picking your juiciest keyword. The search engines can view that particular practice as spammy, and it does not engender a good user experience.

NOTE

A brief aside about hoarding PageRank: many people have taken this to an extreme and built sites where they refused to link out to other quality websites, because they feared losing visitors and link authority. Ignore this idea! You should link out to quality websites. It is good for users, and it is likely to bring you ranking benefits (through building trust and relevance based on what sites you link to). Just think of your human users and deliver what they are likely to want. It is remarkable how far this will take you.

Avoidance of unnecessary subdomains

The engines may not apply the entirety of a domain's trust and link authority weight to subdomains. This is largely due to the fact that a subdomain could be under the control of a different party, and therefore in the search engine's eyes it needs to be separately evaluated. In the great majority of cases, content that gets placed within its own subdomain can easily go in a subfolder, such as *site.com/content*, as opposed to the subdomain *content.site.com*.

Geolocation

If the domain is targeting a specific country, make sure the guidelines for country geotargeting outlined in "Best Practices for Multilanguage/Country Targeting" are being followed. If your concern is primarily about ranking for *san francisco chiropractor* because you own a chiropractic office in San Francisco, California, make sure your address is on every page of your site. You should also claim and ensure the validity of your Google Places listings to ensure data consistency; this is discussed in detail in Chapter 10.

External linking

Check the inbound links to the site by performing a backlink analysis. Use a backlinking tool such as LinkResearchTools, Open Site Explorer, Majestic SEO, or Ahrefs Site Explorer & Backlink Checker to collect data about your links. Look for bad patterns in the anchor text, such as 87% of the links having the critical keyword for the site in them. Unless the critical keyword happens to also be the name of the company, this is a sure sign of trouble. This type of distribution is quite likely the result of link purchasing or other manipulative behavior, and will (if it hasn't already) likely earn you a manual Google penalty or trigger Google's Penguin algorithm to lower your rankings.

On the flip side, make sure the site's critical topics and keywords are showing up sometimes. A lack of the topically related anchor text is not entirely good, either. You need to find a balance, and err on the side of caution, intuitiveness, and usability.

Also check that there are links to pages other than the home page. These are often called *deep links*, and they will help drive the ranking of key sections of your site. You should look at the links themselves, by visiting the linking pages, and see whether the links appear to be paid for. They may be overtly labeled as sponsored, or their placement may be such that they are clearly not natural endorsements. Too many of these are another sure sign of trouble in the backlink profile.

Lastly, check how the backlink profile for the site compares to the backlink profiles of its major competitors. Make sure that there are enough external links to your site, and that there are enough high-quality links in the mix.

Image alt attributes

Do all the images have relevant, keyword-rich `alt` attribute text and filenames? Search engines can't easily tell what is inside an image, and the best way to provide them with some clues is with the `alt` attribute and the filename of the image. These can also reinforce the overall context of the page itself.

Code quality

Although W3C validation is not something search engines require, checking the code itself is a good idea (you can check it with the W3C validator. Poor coding can have some undesirable impacts. You can use a tool such as SEO Browser to see how the search engines see the page.

The Importance of Keyword Reviews

Another critical component of an architecture audit is a keyword review. Basically, this involves the following steps.

Step 1: Keyword research

It is vital to examine your topic and keyword strategy as early as possible in any SEO effort. You can read about this in more detail in Chapter 5.

Step 2: Site architecture

Coming up with architecture for a website can be very tricky. At this stage, you need to look at your keyword research and the existing site (to make as few changes as possible). You can think of this in terms of your site map.

You need a hierarchy that leads site visitors to your *high-value pages* (i.e., the pages where conversions are most likely to occur). Obviously, a good site hierarchy allows the parents of your "money pages" to rank for relevant keywords, which are likely to be shorter tail.

Most products have an obvious hierarchy they fit into, but for products with descriptions, categories, and concepts that can have multiple hierarchies, deciding on a site's information architecture can become very tricky. Some of the trickiest hierarchies, in our opinion, can occur when there is a location involved. In London alone there are London boroughs, metropolitan boroughs, tube stations, and postcodes. London even has a city ("The City of London") within it.

In an ideal world, you will end up with a single hierarchy that is natural to your users and gives the closest mapping to your keywords. But whenever there are multiple ways in which people search for the same product, establishing a hierarchy becomes challenging.

Step 3: Keyword mapping

Once you have a list of keywords and a good sense of the overall architecture, start mapping the major relevant keywords to URLs. When you do this, it is very easy to spot pages that you were considering creating that aren't targeting a keyword (perhaps you might skip creating these) and, more importantly, keywords that don't have a page.

If this stage is causing you problems, revisit step 2. Your site architecture should lead naturally to a mapping that is easy to use and includes your keywords.

Step 4: Site review

Once you are armed with your keyword mapping, the rest of the site review will flow more easily. For example, when you are looking at crafting your `<title>` tags and headings, you can refer back to your keyword mapping and see not only whether there is appropriate use of tags (such as the heading tag), but also whether it includes the appropriate keyword targets.

Keyword Cannibalization

Keyword cannibalization typically starts when a website's information architecture calls for the targeting of a single term or phrase on multiple pages of the site. This is often done unintentionally, but it can result in several or even dozens of pages that have the same keyword target in the title and header tags. Figure 4-4 shows the problem.

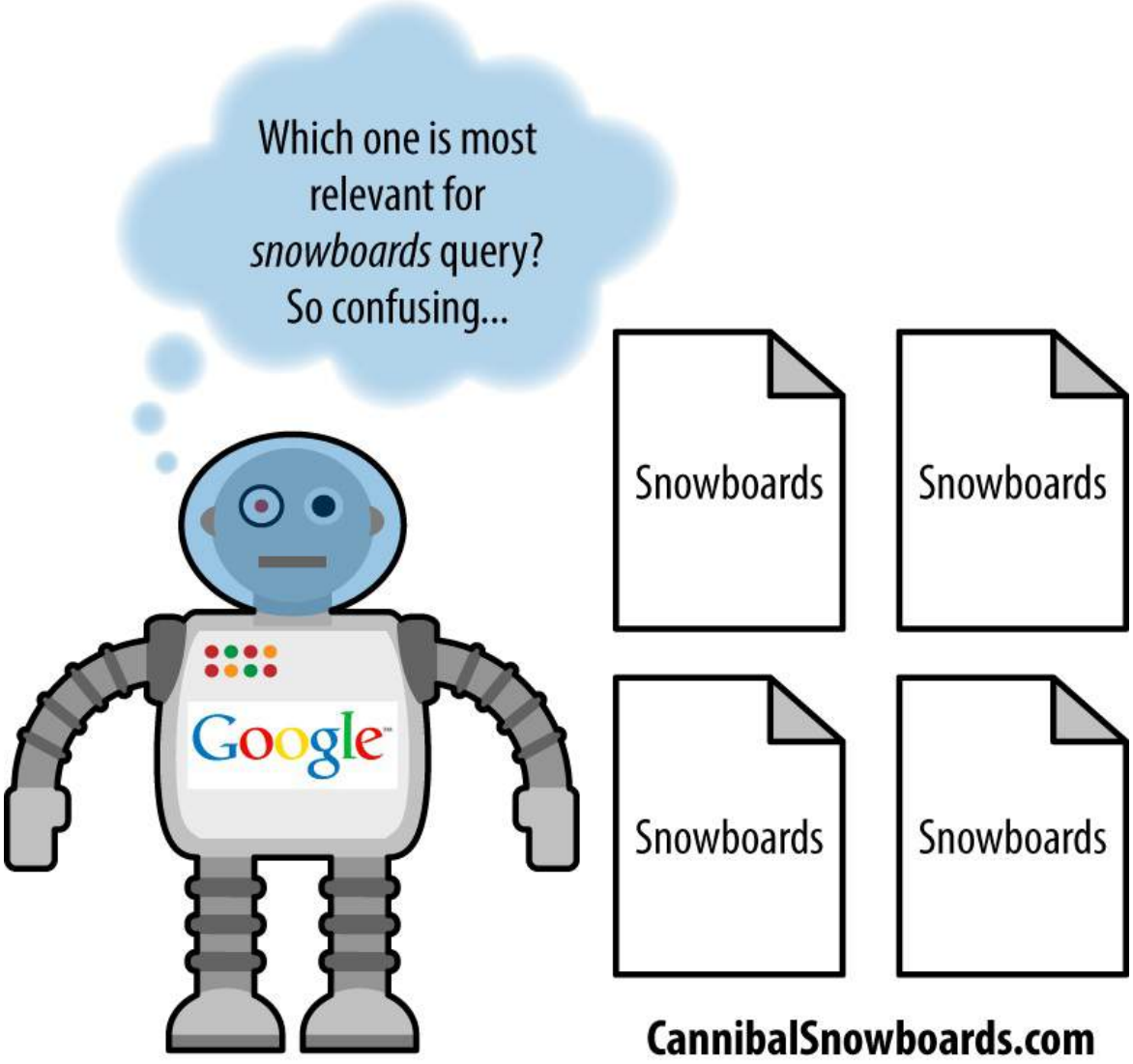


Figure 4-4. Example of keyword cannibalization

Search engines will spider the pages on your site and see 4 (or 40) different pages, all seemingly relevant to one particular keyword (in the example in Figure 4-4 the keyword is *snowboards*). For clarity's sake, Google doesn't interpret this as meaning that your site as a whole is more relevant to snowboards or should rank higher than the competition. Instead, it forces Google to choose among the many versions of the page and pick the one it feels best fits the query. When this happens, you lose out on a number of rank-boosting features:

Internal anchor text

Because you're pointing to so many different pages with the same subject, you can't concentrate the value and weight of internal, thematically relevant anchor text on one target.

External links

If four sites link to one of your pages on snowboards, three sites link to another of your snowboard pages, and six sites link to yet another snowboard page, you've split up your external link value among three topically similar pages, rather than consolidating it into one.

Content quality

After three or four pages about the same primary topic, the value of your content is going to suffer. You want the best possible single page to attract links and referrals, not a dozen bland, repetitive pages.

Conversion rate

If one page is converting better than the others, it is a waste to have multiple lower-converting versions targeting the same traffic. If you want to do conversion tracking, use a multiple-delivery testing system (either A/B or multivariate), such as Optimizely.

So, what's the solution? Take a look at Figure 4-5.

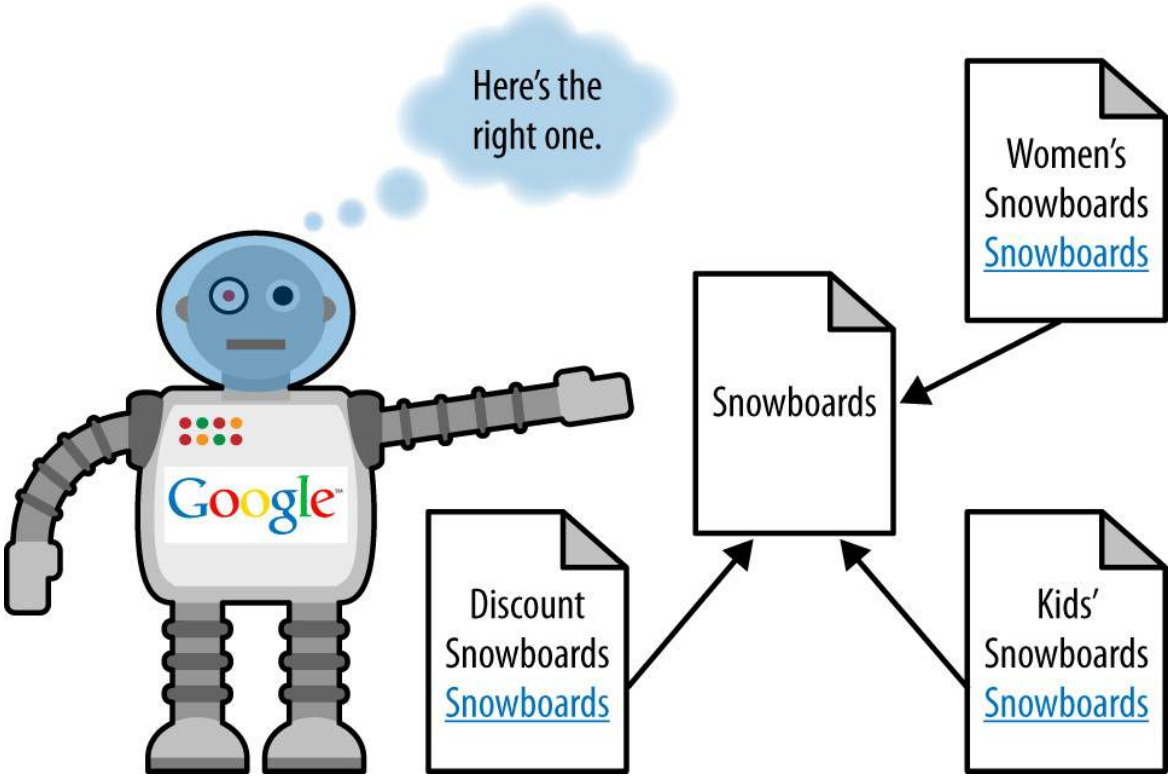


Figure 4-5. Solution to keyword cannibalization

The difference in this example is that instead of every page targeting the single term *snowboards*, the pages are focused on unique, valuable variations and all of them link back to an original, canonical source for the single term. Google can now easily identify the most relevant page for each of these queries. This isn't just valuable to the search engines; it also represents a far better user experience and overall information architecture.

What should you do if you already have a case of keyword cannibalization? Employ 301s liberally to eliminate pages competing with each other within SERPs, or figure out how to differentiate them. Start by identifying all the pages in the architecture with this issue and determine the best page to point them to, and then use a 301 from each of the problem pages to the page you wish to retain. This ensures not only that visitors arrive at the right page, but also that the link equity and relevance built up over time are directing the engines to the most relevant and highest-ranking-potential page for the query. If 301s aren't practical or feasible, use `rel="canonical"`.

Example: Fixing an Internal Linking Problem

Enterprise sites range from 10,000 to 10 million pages in size. For many of these types of sites, an inaccurate distribution of internal link juice is a significant problem. Figure 4-6 shows how this can happen.

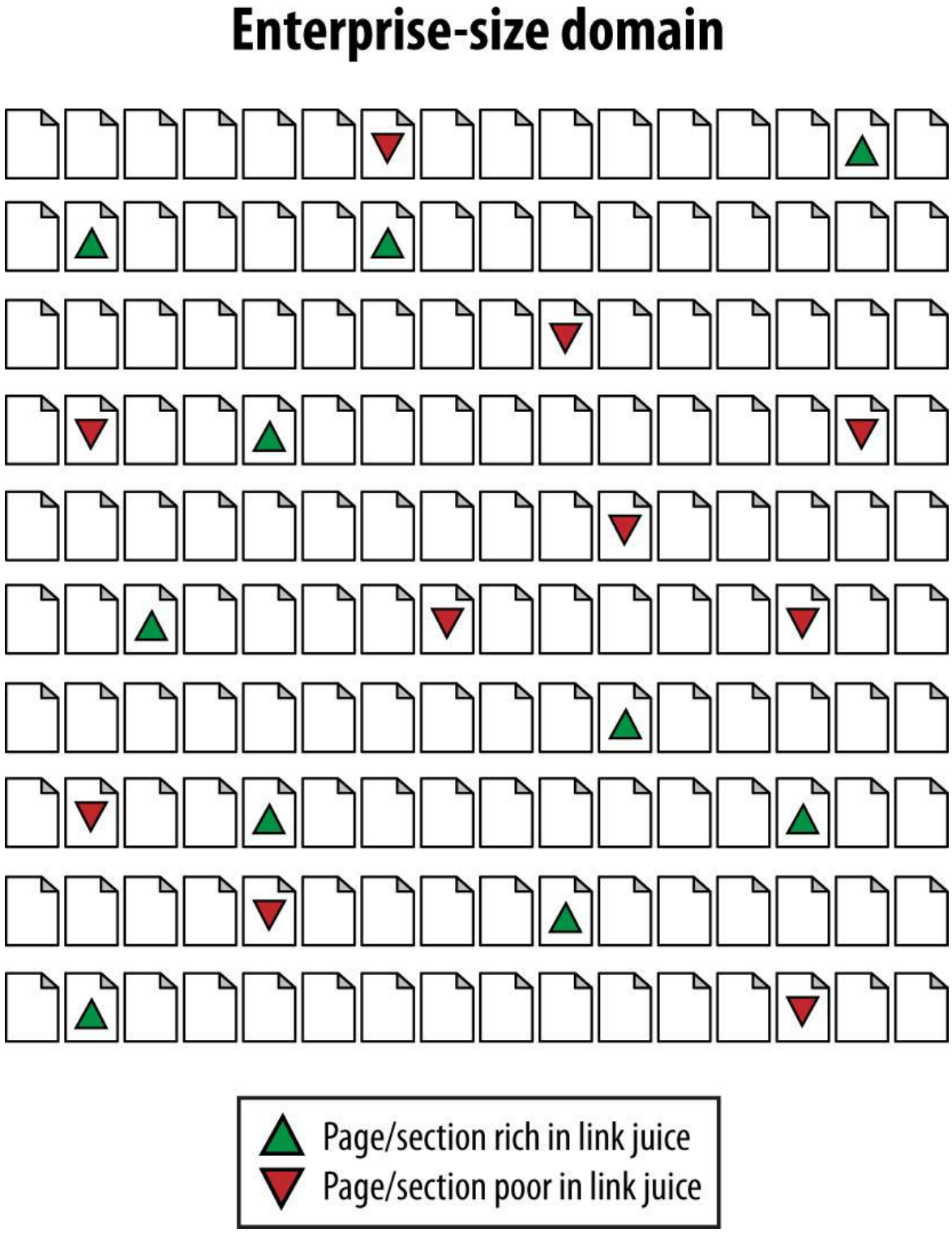


Figure 4-6. Link authority distribution on a very large site

Figure 4-6 is an illustration of the link authority distribution issue. Imagine that

each tiny page represents 5,000–100,000 pages in an enterprise site. Some areas, such as blogs, articles, tools, popular news stories, and so on, might be receiving more than their fair share of internal link attention. Other areas—often business-centric and sales-centric content—tend to fall by the wayside. How do you fix this problem? Take a look at Figure 4-7.

The solution is simple, at least in principle: have the link-rich pages spread the wealth to their link-bereft brethren. As easy as this may sound, it can be incredibly complex to execute. Inside the architecture of a site with several hundred thousand or a million pages, it can be nearly impossible to identify link-rich and link-poor pages, never mind adding code that helps to distribute link authority equitably.

The answer, sadly, is labor-intensive from a programming standpoint. Enterprise site owners need to develop systems to track inbound links and/or rankings and build bridges (or, to be more consistent with Figure 4-7, spouts) that funnel authority between the link-rich and link-poor.

An alternative is simply to build a very flat site architecture that relies on relevance or semantic analysis. This strategy is more in line with the search engines' guidelines (though slightly less perfect) and is certainly far less labor-intensive.

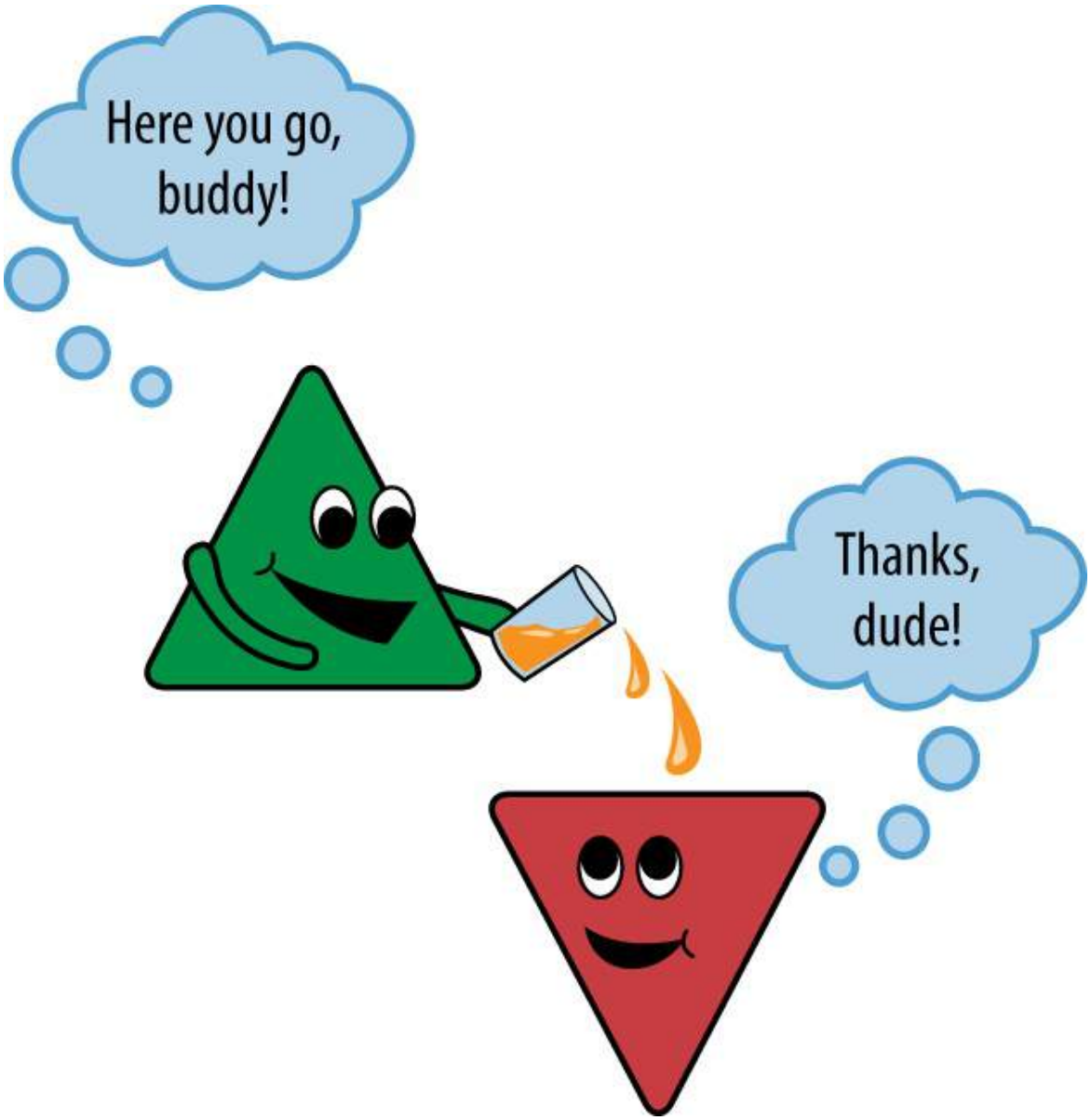


Figure 4-7. Using cross-links to push link authority where you want it

Server and Hosting Issues

Thankfully, only a handful of server or web hosting dilemmas affect the practice of search engine optimization. However, when overlooked, they can spiral into massive problems, and so are worthy of review. The following are some server and hosting issues that can negatively impact search engine rankings:

Server timeouts

If a search engine makes a page request that isn't served within the bot's time limit (or that produces a server timeout response), your pages may not make it into the index at all, and will almost certainly rank very poorly (as no indexable text content has been found).

Slow response times

Although this is not as damaging as server timeouts, it still presents a potential issue. Not only will crawlers be less likely to wait for your pages to load, but surfers and potential linkers may choose to visit and link to other resources because accessing your site is problematic. Again, net neutrality concerns are relevant here.

Shared IP addresses

Basic concerns include speed, the potential for having spammy or untrusted neighbors sharing your IP address, and potential concerns about receiving the full benefit of links to your IP address (discussed in more detail in "Why Using a Static IP Address Is Beneficial... Google Engineer Explains").

Blocked IP addresses

As search engines crawl the Web, they frequently find entire blocks of IP addresses filled with nothing but egregious web spam. Rather than blocking each individual site, engines do occasionally take the added measure of blocking an IP address or even an IP range. If you're concerned, search for your IP address at Bing using the *ip:address* query.

Bot detection and handling

Some system administrators will go a bit overboard with protection and restrict access to files to any single visitor making more than a certain number of requests in a given time frame. This can be disastrous for search engine traffic, as it will constantly limit the spiders' crawling ability.

Bandwidth and transfer limitations

Many servers have set limitations on the amount of traffic that can run

through to the site. This can be potentially disastrous when content on your site becomes very popular and your host shuts off access. Not only are potential linkers prevented from seeing (and thus linking to) your work, but search engines are also cut off from spidering.

Server geography

While the search engines of old did use the location of the web server when determining where a site's content is, Google makes it clear that in today's search environment, actual server location is, for the most part, irrelevant. According to Google, if a site is using a ccTLD or gTLD (country code top-level domain or generic top-level domain, respectively) in conjunction with Search Console to set geolocation information for the site, then the location of the server itself becomes immaterial. There is one caveat: content hosted closer to end users tends to be delivered more quickly, and speed of content delivery *is* taken into account by Google, impacting mobile search significantly.

Identifying Current Server Statistics Software and Gaining Access

In Chapter 11, we will discuss in detail the methods for tracking results and measuring success, and we will also delve into how to set a baseline of measurements for your SEO projects. But before we do that, and before you can accomplish these tasks, you need to have the right tracking and measurement systems in place.

Web Analytics

Analytics software can provide you with a rich array of valuable data about what is taking place on your site. It can answer questions such as:

- How many unique visitors did you receive yesterday?
- Is traffic trending up or down?
- What site content is attracting the most visitors from organic search?

- What are the best-converting pages on the site?

We strongly recommend that if your site does not currently have any measurement systems in place, you remedy that immediately. High-quality, free analytics tools are available, such as the powerful and robust Google Analytics, as well as the open source platform Piwik. Of course, higher-end analytics solutions are also available, which we will discuss in greater detail in Chapter 11.

Log file Tracking

Log files contain a detailed click-by-click history of all requests to your web server. Make sure you have access to the logfiles and some method for analyzing them. If you use a third-party hosting company for your site, chances are it provides some sort of free logfile analyzer, such as AWStats, Webalizer, or something similar (the aforementioned Piwik also provides a log analytics tool). Identify and obtain access to whatever tool is in use as soon as you can.

One essential function that these tools provide that JavaScript-based web analytics software cannot is recording search engine spider activity on your site. Although spidering will typically vary greatly from day to day, you can still see longer-term trends of search engine crawling patterns, and whether crawling activity is trending up (good) or down (bad). Although this web crawler data is very valuable, do not rely on the free solutions provided by hosting companies for all of your analytics data, as there is a lot of value in what traditional analytics tools can capture.

NOTE

Some web analytics software packages read logfiles as well, and therefore can report on crawling activity. We will discuss these in more detail in Chapter 11.

Google Search Console and Bing Webmaster Tools

As mentioned earlier, other valuable sources of data include Google Search Console and Bing Webmaster Tools. We cover these extensively in “Using Search Engine–Supplied SEO Tools”.

From a planning perspective, you will want to get these tools in place as soon as possible. Both tools provide valuable insight into how the search engines see your site, including external link data, internal link data, crawl errors, high-volume search terms, and much, much more. In addition, these tools provide important functionality for interacting directly with the search engine via sitemap submissions, URL and sitelink removals, disavowing bad links, setting parameter handling preferences, tagging structured data, and more—enabling you to have more control over how the search engines view, crawl, index, and output your page content within their index, and within search results.

Determining Top Competitors

Understanding the competition should be a key component of planning your SEO strategy. The first step is to understand who your competitors in the search results really are. It can often be small players who give you a run for your money.

Identifying Spam

Affiliates that cheat tend to come and go out of the top search results, as only sites that implement ethical tactics are likely to maintain their positions over time.

How do you know whether a top-ranking site is playing by the rules? Look for dubious links to the site using a backlink analysis tool such as Majestic SEO or Open Site Explorer (discussed earlier in this chapter). Because the number of links is one factor search engines use to determine search position, less ethical websites will obtain links from a multitude of irrelevant and low-quality sites.

This sort of sleuthing can reveal some surprises. For instance, here are examples of two devious link schemes:

- GiftCertificates.com's short-lived nemesis was FindGiftCards.com, which came out of nowhere to command the top two spots in Google for the all-important search term *gift certificates*, thus relegating GiftCertificates.com to the third position. How did FindGiftCards.com do it? It operated a sister site, 123counters.com, with a free hit counter that propagated "link spam" across thousands of sites, all linking back to FindGiftCards.com and other sites in its

network.

Sadly for FindGiftCards.com, coauthor Stephan Spencer discussed the company's tactics in an article he wrote for *Multichannel Merchant*, and Google became aware of the scam.[1] The end result? Down to two pages in the Google index, as shown in Figure 4-8.

- CraigPadoa.com was a thorn in the side of SharperImage.com, outranking the latter for its most popular product, the Ionic Breeze, by frameset trickery and guestbook spamming (in other words, defacing vulnerable websites with fake guestbook entries that contained spammy links back to its own site). As soon as The Sharper Image realized what was happening, it jumped on the wayward affiliate. It also restricted such practices in its affiliate agreement and stepped up its monitoring for these spam practices.

Figure 4-8. Site with only two pages in the index

Seeking the Best

Look for competitors whose efforts you would like to emulate (or "embrace and extend," as Bill Gates would put it)—usually a website that consistently dominates the upper half of the first page of search results in the search engines for a range of important keywords that are popular and relevant to your target audience.

Note that your "mentor" competitors shouldn't just be good performers, they should also demonstrate that they know what they're doing when it comes to SEO. To assess competitors' competence at SEO, you need to answer the following questions:

- Are their websites fully indexed by Google and Bing? In other words, are all

their web pages, including product pages, making it into the search engines' databases? You can go to each search engine and type in *site: <theirdomain.com>* to find out. A competitor with only a small percentage of its site indexed in Google probably has a site that is unfriendly to search spiders.

- Do their product and category pages have keyword-rich page titles (`<title>` tags) that are unique to each page? You can easily review an entire site's page titles within Google or Bing by searching for *site: <www.yourcompetitor.com>*.

  Incidentally, this type of search can sometimes yield confidential information. A lot of webmasters do not realize that Google has discovered and indexed commercially sensitive content buried deep in their sites. For example, a Google search for *business plan confidential filetype:doc* can yield a lot of real business plans among the sample templates.

- Do their product and category pages have reasonably high PageRank scores?

- Is anchor text across the site, particularly in the navigation, descriptive but not keyword-stuffed?

- Are the websites getting penalized? You can overdo SEO. Too much keyword repetition or too many suspiciously well-optimized text links can yield a penalty for overoptimization. Sites can also be penalized for extensive amounts of duplicate content, for lack of high-quality content (referred to as "thin" content), and for participating in guest blog/article writing schemes. You can learn more about how to identify search engine penalties in Chapter 9.

Uncovering Their Secrets

Let's assume your investigation has led you to identify several competitors who are gaining excellent search placement using legitimate, intelligent tactics. Now it is time to identify their strategy and tactics:

What keywords are they targeting?

You can determine this by looking at the page titles (up in the bar above the

address bar at the top of your web browser, which also appears in the search results listings) of each competitor's home page and product category pages. You can also use various online tools to see what keywords they may be targeting with PPC advertising; while it's not always an indication that they are investing in SEO, you can still get a solid grasp on their overall keyword strategy.

Who's linking to their home page, or to their top-selling product pages and category pages?

A link popularity checker can be quite helpful in analyzing this.

If it is a database-driven site, what technology workarounds are they using to get search engine spiders such as Googlebot to cope with the site being dynamic?

Nearly all the technology workarounds are tied to the ecommerce platforms the competitors are running. You can check to see whether they are using the same server software as you by using the "What's that site running?" tool. Figure 4-9 shows a screenshot of a segment of the results for HSN.com.

Server-Side

Includes all the main technologies that Netcraft detects as running on the server such as PHP.

Technology	Description	Popular sites using this technology
Atlas	A set of ASP.NET extensions for implementing Ajax functionality	www.ighome.com , www.idiz.com , www.geocaching.com
Using ASP.NET	ASP.NET is running on the server	www.wordreference.com , www.corriere.it , www.newegg.com

Client-Side

Includes all the main technologies that run on the browser (such as JavaScript and Adobe Flash).

Technology	Description	Popular sites using this technology
Image Map	No description	www.ebay.it , www.amazon.com , www.leboncoin.fr
JavaScript	Open source programming language commonly implemented as part of a web browser	www.imdb.com , mail.google.com

Client-Side Scripting Frameworks

Frameworks or libraries allow for easier development of applications by providing an Application Program Interface (API) or a methodology to follow whilst developing.

Technology	Description	Popular sites using this technology
Modernizr	No description	www.samsung.com , www.wunderground.com , www.gutefrage.net

Mobile Technologies

Mobile technology is the technology used for hand held mobile devices.

Technology	Description	Popular sites using this technology
Click to call	Markup language syntax intended for devices that can place calls (e.g. phones, VoIP, etc.)	www.eprice.it , www.istockphoto.com , www.nk.ca

Character Encoding

A character encoding system consists of a code that pairs each character from a given repertoire with something else such as a bit pattern, sequence of natural numbers, octets, or electrical pulses in order to facilitate the transmission of data (generally numbers or text) through telecommunication networks or for data storage.

Technology	Description	Popular sites using this technology
UTF8	UCS Transformation Format 8 bit	www.googleadservices.com

HTTP Compression

HTTP compression is a capability that can be built into web servers and web clients to make better use of available bandwidth, and provide greater transmission speeds between both.

Technology	Description	Popular sites using this technology
Gzip Content Encoding	Gzip HTTP Compression protocol	www.iffattiequotidiani.it , www.programme-tv.net , www.zerohedge.com

Figure 4-9. Sample Netcraft output

While you are at it, look at cached (archived) versions of your competitors' pages by clicking on the Cached link next to their search results in Google to see whether they're doing anything too aggressive, such as *cloaking*, where

they serve up a different version of the page to search engine spiders than to human visitors. The cached page will show you what the search engine actually saw, and you can see how it differs from the page you see.

What effect will their future SEO initiatives have on their site traffic?

Assess the success of their SEO not just by the lift in rankings. Periodically record key SEO metrics over time—the number of pages indexed, the PageRank score, the number of links—and watch the resulting effect on their site traffic. If you utilize one of the many web-based SEO tool platforms (read more about these in Chapter 11), you can set your competitors' sites up as additional sites or campaigns to track.

You do not need access to competitors' analytics data or server logs to get an idea of how much traffic they are getting. Simply go to Compete, Quantcast, Search Metrics, or SEMRush and search on the competitor's domain. If you have the budget for higher-end competitive intelligence tools, you can use comScore or Experian's Hitwise.

The data these tools can provide is limited in its accuracy, and can often be unavailable if the site in question receives too little traffic, but it's still very useful in giving you a general assessment of where your competitors are. The tools are most useful for making relative comparisons between sites in the same market space. To get an even better idea of where you stand, use their capabilities to compare the traffic of multiple sites, including yours, to learn how your traffic compares to theirs.

How does the current state of their sites' SEO compare with those of years past?

You can reach back into history and access previous versions of your competitors' home pages and view the HTML source to see which optimization tactics they were employing back then. The Internet Archive's Wayback Machine provides an amazingly extensive archive of web pages.

Assessing Historical Progress

Measuring the results of SEO changes can be challenging, partly because there are so many moving parts and partly because months can elapse between when

changes are made to a site and when results are seen in organic search exposure and traffic. This difficulty only increases the importance of measuring progress and being accountable for results. This section will explore methods for measuring the results from your SEO efforts.

Timeline of Site Changes

Keeping a log of changes to your site is absolutely recommended. If you're not keeping a timeline (which could be as simple as an online spreadsheet or as complex as a professional project management visual flowchart), you will have a harder time executing your SEO plan and managing the overall SEO process. Sure, without one you can still gauge the immediate effects of content additions/revisions, link acquisitions, and development changes, but visibility into how technical modifications to the website might have altered the course of search traffic, whether positively or negatively, is obscured.

One of the most useful ways to document this information is by utilizing annotations in Google Analytics, whereby you can document specific events along your site timeline for future reference and for ongoing visual association with your analytics data. Figure 4-10 shows how annotations appear along your timeline graph in Google Analytics reports; note the text balloons along the date timeline (each text balloon can be clicked on individually, to display the notes added at that time).

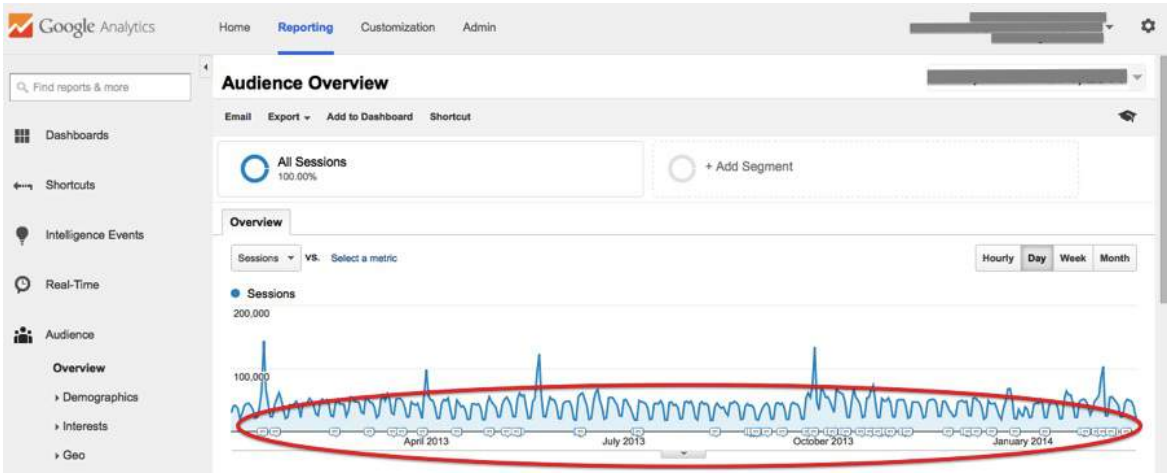


Figure 4-10. Annotations in Google Analytics

If you don’t record and track site changes—both those intended to influence SEO, and those for which SEO wasn’t even a consideration—you will be eventually optimizing blind and could miss powerful signals that could help dictate your strategy going forward. Ideally, you should track more than just site changes, as external factors that can have a big impact on your SEO results can be just as important; these include confirmed search engine algorithm updates, competitor news events (e.g., product or company launches), breaking news, and trending topics in social media. Factors within your own business can have an impact as well, such as major marketing or PR events, IPOs, or the release of earnings statements.

There are many scenarios in which you will want to try to establish cause and effect, such as:

If search traffic spikes or plummets

Sudden changes in organic traffic are obviously notable events. If traffic plummets, you will be facing lots of questions about why, and having a log of site changes will put you in a better position to assess whether any changes you recommended could have been the cause. Of course, if traffic spikes, you will want to be able to see whether an SEO-related change was responsible as well.

When gradual traffic changes begin

Changes do not always come as sudden spikes or drop-offs. If you see the traffic beginning a gradual climb (or descent), you will want to be able to assess the likely reasons.

To track and report SEO progress

Accountability is a key component of SEO. Budget managers will want to know what return they are getting on their SEO investment. This will inevitably fall into two buckets: itemizing specific items worked on, and analyzing benefits to the business. Keeping an ongoing change log makes tracking and reporting SEO progress much easier to accomplish.

Creation of new, noteworthy content onsite and/or tied to social media campaigns

Your social media efforts can directly impact your site's performance, both directly (traffic from third-party social media platforms) and indirectly (via links, social sharing, search query trending, and other factors). All of this can be tied back via annotations in Google Analytics for meaningful data analysis.

Types of Site Changes That Can Affect SEO

Your log should track all changes to the website, not just those that were made with SEO in mind. Organizations make many changes that they do not think will affect SEO, but that have a big impact on it. Here are some examples:

- Adding content areas/features/options to the site (this could be anything from a new blog to a new categorization system)
- Changing the domain of the site (this can have a significant impact, and you should document when the switchover was made)
- Modifying URL structures (changes to URLs on your site will likely impact your rankings, so record any and all changes)
- Implementing a new CMS (this is a big one, with a very big impact—if you must change your CMS, make sure you do a thorough analysis of the SEO shortcomings of the new CMS versus the old one, and make sure you track the timing and the impact)
- Establishing new partnerships that either send links or require them (meaning your site is earning new links or linking out to new places)
- Acquiring new links to pages on the site other than the home page (referred to as *deep links*)
- Making changes to navigation/menu systems (moving links around on pages, creating new link systems, etc.)
- Implementing redirects either to or from the site
- Implementing SSL/HTTPS

- Implementing new/updated sitemaps, canonical tags, schema markup, and so on

When you track these items, you can create an accurate storyline to help correlate causes with effects. If, for example, you've observed a spike in traffic from Bing that started four to five days after you switched your menu links from the page footer to the header, it is likely that there is a correlation; further analysis would determine whether, and to what extent, there is causation.

Without such documentation, it could be months before you noticed the impact—and there would be no way to trace it back to the responsible modification. Your design team might later choose to switch back to footer links, your traffic might fall, and no record would exist to help you understand why. Without the lessons of history, you cannot leverage the positive influencing factors.

Previous SEO Work

When you are brought on to handle the SEO for a particular website, one of the first things you need to find out is which SEO activities have previously been attempted. There may be valuable data there, such as a log of changes that you can match up with analytics data to gauge impact.

If no such log exists, you can always check the Internet Archive's Wayback Machine to see whether it has historical logs for your website. This offers snapshots of what the site looked like at various points in time.

Even if a log was not kept, spend some time building a timeline of when any changes that affect SEO (as discussed in the previous section) took place. In particular, see whether you can get copies of the exact recommendations the prior SEO consultant made, as this will help you with the timeline and the specifics of the changes made.

You should also pay particular attention to understanding the types of link-building activities that took place. Were shady practices used that carry a lot of risk? Was there a particular link-building tactic that worked quite well? Going through the history of the link-building efforts can yield tons of information that you can use to determine your next steps. Google undoubtedly keeps a "rap sheet" on us webmasters/SEO practitioners.

Benchmarking Current Indexing Status

The search engines have an enormous task: indexing the world's online content (well, more or less). The reality is that they try hard to discover all of it, but they do not choose to include all of it in their indexes. There can be a variety of reasons for this, such as the page being inaccessible to the spider, being penalized, or not having enough link juice to merit inclusion.

When you launch a new site or add new sections to an existing site, or if you are dealing with a very large site, not every page will necessarily make it into the index. You will want to actively track the indexing level of your site in order to develop and maintain a site with the highest accessibility and crawl efficiency. If your site is not fully indexed, it could be a sign of a problem (not enough links, poor site structure, etc.).

Getting basic indexation data from search engines is pretty easy. The major search engines support the same basic syntax for that: *site:<yourdomain.com>*. Figure 4-11 shows a sample of the output from Bing.

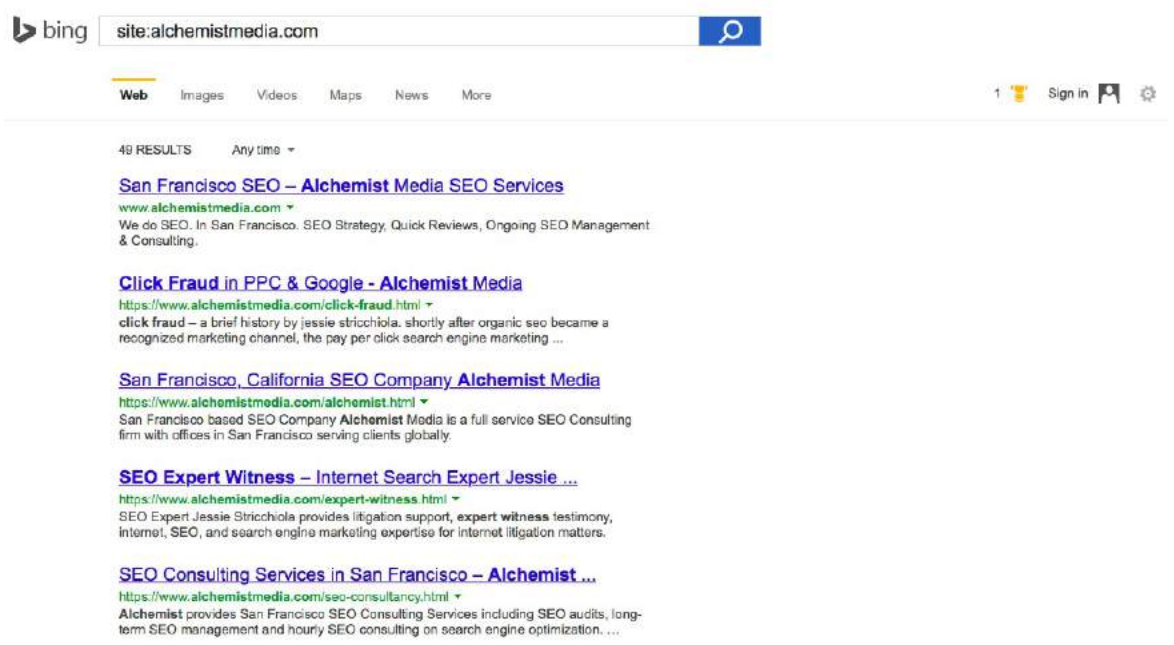


Figure 4-11. Indexing data from Bing

Keeping a log of the level of indexation over time can help you understand how things are progressing, and this information can be tracked in a spreadsheet.

Related to indexation is the crawl rate of the site. Google and Bing provide this data in their respective toolsets. Figure 4-12 shows a screenshot representative of the crawl rate charts that are available in Google Search Console.

Short-term spikes are not a cause for concern, nor are periodic drops in levels of crawling. What is important is the general trend. Bing Webmaster Tools provides similar data to webmasters, and for other search engines their crawl-related data can be revealed using logfile analyzers mentioned previously in this chapter—and then a similar timeline can be created and monitored.

Figure 4-12. Crawl data from Google Search Console

Benchmarking Organic Rankings

People really love to check their search rankings. Many companies want to use this as a measurement of SEO progress over time, but it is a bit problematic, for a variety of reasons. Here is a summary of the major problems with rank checking:

- Google results are not consistent:

  — Different geographies (even in different cities within the United States) often give different results.

  — Results are personalized for logged-in users based on their search histories, and personalized for non-logged-in users as well based on their navigational behaviors within the SERPs, browser history, location data, and more.

  — No rank checker can monitor and report all of these inconsistencies (at least, not without scraping Google hundreds of times from all over the world with every possible setting).

- Obsessing over rankings (rather than traffic) can result in poor strategic decisions:

  — When site owners obsess over rankings for particular keywords, the time and energy they expend on those few keywords and phrases often produces far less value than would have been produced if they had spent those resources on the site as a whole.

  — Long-tail traffic very often accounts for 70%–80% of the demand curve, and it is much easier to rank in the long tail and get valuable traffic from there than it is to concentrate on the few rankings at the top of the demand curve.

So indulge your desire to check rankings by going to the search engine and typing in a few queries or use a tool like AuthorityLabs and automate it, but be sure to also keep an ongoing focus on your visitor and conversion statistics.

Benchmarking Current Traffic Sources and

Volume

The most fundamental objective of any SEO project should be to drive the bottom line. For a business, this means delivering more revenue with favorable ROI. As a precursor to determining the level of ROI impact, the SEO pro must focus on increasing the volume of relevant traffic to the site. This is a more important objective than anything related to rankings or number of links obtained. More relevant traffic should mean more revenue for the business (or more conversions, for those whose websites are not specifically selling something).

Today's web analytics tools make gathering such data incredibly easy. Google Analytics offers a robust set of tools sufficient for many smaller to medium sites, while larger sites will probably need to use the paid version of Google Analytics (Google Analytics Premium) in lieu of or in conjunction with other paid solutions such as Adobe's Marketing Cloud, IBM Experience One, or Webtrends.

The number of data points you can examine in analytics is nearly endless. It is fair to say that there is too much data, and one of the key things that an SEO expert needs to learn is which data is worth looking at and which is not. Generally speaking, the creation of custom dashboards in Google Analytics can greatly reduce the steps needed to see top-level data points for SEO; in Chapter 10, we will discuss in greater detail how to leverage the power of analytics.

Leveraging Business Assets for SEO

Chances are, your company/organization has numerous digital media assets beyond the website that can be put to good use to improve the quality and quantity of traffic you receive through search engine optimization efforts. We discuss some of these assets in the subsections that follow.

Other Domains You Own/Control

If you have multiple domains, the major items to think about are:

- Can you 301-redirect some of those domains back to your main domain or to

a subfolder on the site for additional benefit? Be sure to check the domain health in Search Console and Webmaster Tools before performing any 301 redirects, as you want to ensure the domain in question has no penalties before closely associating it with your main domain(s).

- If you're maintaining those domains as separate sites, are you linking between them intelligently?

If any of those avenues produce valuable strategies, pursue them—remember that it is often far easier to optimize what you're already doing than to develop entirely new strategies, content, and processes.

Relationships On and Off the Web

Business relationships and partnerships can be leveraged in similar ways, particularly on the link development front. If you have business partners that you supply or otherwise work with—or from whom you receive service—chances are good that you can safely create legitimate, user-friendly linking relationships between your sites. Although reciprocal linking carries a bit of a bad reputation, there is nothing wrong with building a "partners," "clients," "suppliers," or "recommended" list on your site, or with requesting that your organizational brethren do likewise for you. Just do this in moderation and make sure you link only to highly relevant, trusted sites.

Content or Data You've Never Put Online

Chances are that you have content that you have never published on your website. This content can be immensely valuable to your SEO efforts. However, many companies are not savvy about how to publish that content in a search engine–friendly manner. Those hundreds of lengthy articles you published when you were shipping a print publication via the mail could be a great fit for your website archives. You should take all of your email newsletters and make them accessible on your site. If you have unique data sets or written material, you should apply it to relevant pages on your site (or consider building out if nothing yet exists).

Customers Who Have Had a Positive Experience

Customers are a terrific resource for earning links, but did you also know they can write? Customers and website visitors can contribute all kinds of content. Seriously, if you have user-generated content (UGC) options available to you and you see value in the content your users produce, by all means reach out to customers, visitors, and email list subscribers for content opportunities.

Followers, Friends, and Fans

If you run a respected business that operates offline or work in entertainment, hard goods, or any consumer services, there are likely people out there who've used your products or services and would love to share their experiences. Do you make video games? Hopefully you've established a presence on social media and can leverage third-party platforms (Facebook, Pinterest, Instagram, Twitter, Vine, and others) to market your content for link acquisition, content curation, positive testimonials, and brand buzz.

Conclusion

The first steps of SEO are often the most challenging ones. It's tempting to rush into the effort just to get things moving, but spending some focused time properly planning your SEO strategy before implementation will pay big dividends in the long run. Establish a strong foundation, and you will have a firm launching pad for SEO success.

[1] Stephan Spencer, "Casing the Competition," Multichannel Merchant, April 1, 2004, *http://multichannelmerchant.com/catalogage/ar/marketing_casing_competition/*.

CHAPTER FIVE

Keyword Research

The Theory Behind Keyword Research

Keyword research is one of the most important, valuable, and high-return activities in the search engine marketing field. Through the detective work of dissecting your market's keyword demand, you learn not only which terms and phrases to target with SEO, but also more about your customers as a whole.

With keyword research, you can predict shifts in demand, respond to changing market conditions, and produce the products, services, and content that web searchers are already actively seeking. In the history of marketing, there has never been such a low barrier to entry in understanding the motivations of consumers in virtually every niche.

Every search phrase that's typed into an engine is recorded in one way or another, and keyword research tools such as the ones we discuss in this chapter allow you to retrieve this information. However, those tools cannot show you (directly) how valuable or important it might be to rank for and receive traffic from those searches. This chapter seeks to expose the details of this process and the tools that can best assist.

Thinking Strategically

Keyword research tools provide valuable insight into the thinking of your potential customers. When users go to search engines and type out their search query, they may use language that is entirely different from what you expect. Even if your product or service provides a solution they can use, they may start with their problem. Someone with diabetes might simply type *diabetes* in the search box. Then his next search might be *diabetes medication* or *relief for diabetes symptoms*.

As we laid out in Chapter 1, searchers often go through a progression where they try certain searches, check out some sites, refine their searches, and repeat this process until they finally find what they want. Taking the time to understand typical search sequences is one aspect that impacts your keyword strategy.

Other elements that influence search behavior include searcher demographics

(male/female, age, income, etc.), geographical location, and time of year. Seasonal products such as Valentine's Day cards, for example, go through sharp peaks in volume during the time preceding February 14, and then decline rapidly once the season is past.

The keyword research tools in this chapter will provide you with methods to investigate all of these factors. Take the time to go beyond the surface and use the tools to learn how your customers think, get your thinking in alignment with theirs, and then build your website strategy (and perhaps even your product strategy) around this powerful information.

Understanding the Long Tail of the Keyword Demand Curve

It is wonderful to deal with keywords that have 5,000 searches per day, or even 500 searches per day, but in reality these "popular" search terms may actually comprise less than 30% of the overall searches performed on the Web. The remaining 70% lie in what's commonly called the "long tail" of search,[1] illustrated in Figure 5-1. The tail contains hundreds of millions of unique searches that might be conducted a few times in any given day, or even only once ever, but when assessed in aggregate, they comprise the majority of the world's demand for information through search engines.

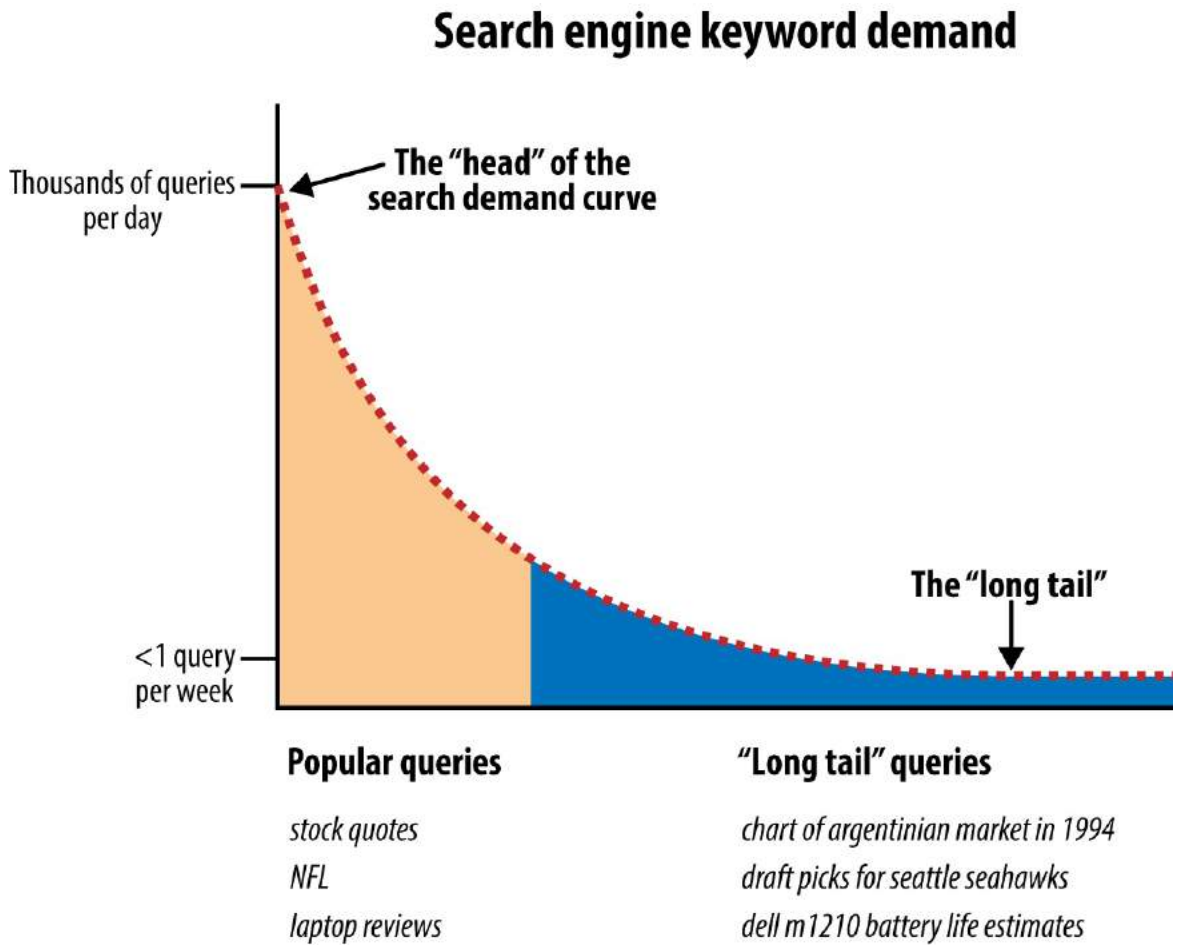


Figure 5-1. Long tail of search

Understanding the Impact of Google Hummingbird

In September 2013, to coincide with its 15th birthday, Google announced a major change to its search algorithm, called Hummingbird. Since then, although Hummingbird has been much discussed, Google has not released any more official documentation on this update. Hummingbird is a major change to the way Google interprets searches and in how searches relate to one another.

Hummingbird is intended to get at the heart of what users want, not just the exact keywords they search for. In large part, this is related to Google getting more prepared for mobile search. In mobile searches, users are less apt to type traditionally formatted queries, and in many cases, actually do voice searches. When users use voice search, their queries may also be much more conversational in format, such as *please find me the closest gas station*. In

addition, with Hummingbird, Google will use many other factors to determine the user's intent, such as considering previous related searches by that user.

The Hummingbird algorithm attempts to determine the true meaning behind what a user is searching for, rather than simply returning results for the exact query she uses. Indeed, in many cases, Google may simply relate the terms and consider them synonyms for the purposes of returning search results.

Let's look at this hypothetical, yet fully functional, example of a series of queries that a mobile user might ask Google:

- *where is the empire state building?*
- *who built it?*
- *pictures*
- *how tall is it?*
- *restaurants*
- *Italian restaurants*
- *show me the second one*
- *directions*

No longer is Google's response to the query *where is the empire state building?* simply the web page that is most closely optimized to that search query; Google uses its knowledge of entities—which include notable buildings and monuments and their attributes and locations, nearby restaurants, average review ratings of those restaurants (partially from semantic markup), and so forth—to return meaningful responses to these queries, including actual Italian restaurant recommendations near the Empire State Building, rather than merely a match on the keywords searched.

Note in the preceding example the presence of one-word queries that Google expands (interprets) into longer-tail queries. This changes the game in terms of chasing after longtail keywords. Now like never before, we must be like detectives who anticipate the searcher's intent and next move before he even

arrives on our website and be ready to meet his needs based on that intent, without the keywords to guide us.

In terms of Hummingbird's impact on keyword research, it's also important to consider co-occurrence analysis. While there is some debate in the search engine optimization marketing discipline about the impact of various types of co-occurrence, there is strong evidence that the co-occurrence of keywords in on-page content can affect search engine rankings. When it comes to keyword research, co-occurrence refers to an analysis of what words and phrases appear most commonly on a page, how often they appear, and how they relate to each other.

Coauthor Eric Enge gives a summary of a study by Searchmetrics in his post "How Content Quality Analysis Works with SEO". He shows how on a sample page terms such as *women*, *shoes*, *boots*, and *shoe* co-occur on an ecommerce page with words like *view*, *sale*, *cart*, *shipping*, and other ecommerce-related terms. This sends a signal to the search engine that this should be a page where users can shop for and buy women's footwear. The phrase *buy women's shoes* may never actually appear on the page and yet due to the co-occurrence of the product type terms and ecommerce terms, Google should understand that the page's "theme" or "purpose" is to offer a place where users can buy women's shoes, and should potentially rank this page for that term (depending on the myriad other quality signals surrounding the page). If an ecommerce page lacks those words and phrases that help to signal that the page's intent is to sell products, it might lower the chances of that page ranking for purchase-oriented search phrases.

There is a related type of keyword co-occurrence that may affect rankings similarly: co-occurrence across various search engine queries during a session. In this type of co-occurrence, the engines would look at how similar groups of keywords are searched across a users' search session. To continue the previous example, if a large number of queries for *women's shoes* co-occurred during users' search sessions with a query for *buy women's shoes online*, Google should know that it is likely that users searching for *women's shoes* are looking for where to buy women's shoes online. In this instance, sites or pages that are well optimized for the term *women's shoes* might also rank well for *buy women's shoes online* without that exact phrase having to appear in the page copy or page metadata. In this case, Google would extrapolate that the user's actual intent is to

find a place to buy women's shoes online, even though the query may simply have been *women's shoes*.

Ultimately, in both cases, the goal is for the search engines to understand the user's intent and to return a set of results that satisfy that intent, regardless of whether the exact phrase searched for explicitly appears on the pages returned as results for a given query.

Hummingbird is neither about shortening or lengthening queries—it's a complete rewrite of the search platform. You can read about this in detail at "A Hummingbird Isn't Just for Christmas, It's for Life". You can also watch "Hummingbird, Google Now, and the Future of Search", an hour-long video hangout with Danny Sullivan and Eric Enge.

Hummingbird was a platform rewrite, much like Caffeine, which rewrote the crawler and indexing parts of Google. Hummingbird rewrote the actual search engine piece (the part that receives user search queries and then pulls results out of the index).

At the SMX West search marketing conference in the Spring of 2014, Marcus Tober of Searchmetrics presented some results of the company's study of Google Hummingbird. Among its findings were that the context of the search (not just the exact query) and previous searches both matter, that Hummingbird attempts to understand conversational search, and that Google is trying to understand entities (people, places, things, etc.) and how they relate to one another. Searchmetrics found that Google is now delivering search results for many related queries as if they were the same search query.

What this means from an SEO perspective is that the exact query a user may be searching for is less important than the intent behind it. While keyword research is still crucial, creating pages highly optimized to a specific keyword is less important than creating extremely high-quality, unique content that answers the need or question behind the keyword query.

When doing your keyword research, be flexible and open to using tools that show you queries related to those at which you're specifically looking. Also watch for informational queries, where it is apparent that the user is attempting to solve a problem. Much of your keyword research should be centered on the true meaning behind a query and why a user searches on those specific terms.

To get your content to rank for the queries you uncover through your keyword research, focus on creating the best site and the best content while still incorporating keywords into your content where it is natural to do so.

Understanding Keyword "Not Provided" and Co-Occurrence Analysis

Two major changes in recent years that affect the way we approach keyword research are Google's (and now Bing's) move to obscuring referring keywords (*keyword "not provided"*) and a fundamental change to the way Google approaches looking at the entire body of keywords on a page, or *co-occurrence analysis*, as introduced in the preceding subsection.

With "not provided," Google has blocked the vast majority of the data that lets you see which keywords lead to organic search visits to your site. In June 2015, Bing announced it would follow suit in the summer of 2015 and also block keyword data from passing in the referring URL.[2] While there are some workaround solutions to try to retrieve or re-create this data, such as Searchmetrics, other changes in the way Google views individual keywords make this data somewhat less vital than it has been in the past. As noted, going forward it is going to become more important to not just think about how to replace lost keyword data, but to understand your site's relationship to keywords and to better understand your users' needs and intent. Rather than simply thinking about optimizing a given page to a specific keyword, think about how that page can more effectively answer your users' needs and provide a breadth of useful information to satisfy that need.

As just mentioned in the section "Understanding the Impact of Google Hummingbird", Google's intention is to now be able to understand the meaning behind a query rather than just the specific words used in that query. The company is also actively working toward understanding the meaning and theme of the content on your pages rather than just the specific keywords used within that content. Rather than simply ranking a given page for how well it is optimized for a specific query, Google will attempt to rank the content that most closely answers the true intent behind the user's query.

Related to this notion of intent is the aforementioned idea of co-occurrence analysis, which refers to an analysis of what words appear most commonly on a

page and under what circumstances they appear together. Rather than simply looking at keywords in isolation on a page, Google appears to have begun looking at what words commonly appear together under varying circumstances. We explore the notion of co-occurrence in depth in "Integrating Keyword Research, Co-Occurrence Analysis, and Knowledge of User Intent". For now it is useful to note that Google is not looking just specifically at the repetition of a single keyword phrase but rather at what kinds of words appear together to signal the page's true intent and theme and how well the content answers the true intent behind the user's query.

As you continue through this chapter, consider that the keyword research that you will perform is no longer simply about finding the exact best keyword phrases to specifically optimize pages on your site for, but rather to also help uncover what your potential users actually need to get from your site. Rather than highly optimizing each page on your site for one or two specific keyword phrases, think about which phrases might belong together on a page and identify the page's theme and intent. Consider each page more holistically and how you might improve the page to better answer your users' needs. Rather than trying to overoptimize for the exact number or combination of phrases, write naturally and for your users, using a wide variety of related terms and synonyms that you uncover during the keyword research process.

Traditional Approaches: Domain Expertise and Site Content Analysis

One of the smartest things you can do when initially conducting keyword research is to brainstorm original ideas with the participants in the business before getting keyword tools involved. This can be surprisingly effective for coming up with numerous critical keywords, and it can help you understand if your organization thinks about your offerings using different language than your customers—in which case, you may want to adapt!

Start by generating a list of terms and phrases that are relevant to your (or your client's) industry and what your site or business offers. The brainstorming phase should ideally result in a list of several dozen to several hundred or more keyword searches that will bring relevant, qualified visitors to your site. It can be

a great idea to get sales, customer services, or whichever branch works most directly with clients to participate in the brainstorm, as they may have input into keywords or phrases the customer uses or expresses interest in that aren't currently targeted.

One easy way to begin this process is to gather your team in a conference room and then follow these steps:

1. Produce a list of key one-to three-word phrases that describe your products/services.

2. Spend some time coming up with synonyms that your potential customers might use for those products and services. Use a thesaurus to help you with this process.

3. Create a taxonomy of all the areas of focus in your industry. It can be helpful to imagine creating a directory for all the people, projects, ideas, and companies connected to your site. You can also look at sites that are leaders in the industry and study their site hierarchy as a way to start your thinking about a taxonomy.

4. Broaden your list by thinking of higher-level terms and topics of which your products or services are a subset.

5. Review your existing site and extract what appear to be key phrases from your site.

6. Review industry association and/or media sites to see what phrases they use to discuss your topic area.

7. List all of your various brand terms.

8. List all of your products. If your site has a massive number of products, consider stepping back a level (or two) and listing the categories and subcategories.

9. Have your team imagine that they are potential customers, and ask them what they would type into a search engine if they were looking for

something similar to your product or service.

10. Supplement this by asking some people outside your business what they would search for, preferably people who are not directly associated with the company. Consider also the value of performing actual market research with a test group of consumers in your demographic, and ask them the same question.

11. Use various tools (such as Google Search Console) to see what terms people are already using to come to your site, or what terms they are using within your site search tool if you have one.

Gathering this kind of intelligence is what a traditional marketer might have done prior to initiating a marketing campaign before the Web existed. And of course, if any of this data is available to you from other departments of the company, be sure to incorporate it into your research process.

Including Competitive Analysis

Your competitors face the same problem, and unless you are very lucky, they are also probably resourceful and creative. You can likely count on their having invested in learning how their customers think and the best ways to appeal to them. So, add these steps to the process:

1. Review your competitors' websites and see what keywords and phrases they use for their products and services that compete with yours. Look also for unique variations and synonyms they incorporate into their language, and see if these indicate shifting trends in vernacular in your industry.

2. Record what nonbrand terms they use for their business.

3. Read any articles they have written that are published on sites other than their own.

4. Observe what the media may have had to say about them.

Add these ideas into the mix, and you will have a wonderfully robust set of keywords to use as a starting point.

You may ask why you should go through all this trouble. Don't the keyword tools take care of all this for you? There are two reasons why the extra effort is critical:

- Your internal team has a rich array of knowledge that the keyword tools do not: they know where to start. Keyword tools require the initial input of information, and the quality of the data they provide is only as good as the quality of the "seeds" you give them.

- The upfront brainstorming helps your organization's stakeholders better understand the market and the opportunities. It also helps them "buy into" the eventual prioritization of keywords.

Once you have completed these steps, you will have in hand a rich set of terms of interest. The next step is to expand those terms of interest using keyword research tools.

Integrating Keyword Research, Co-Occurrence Analysis, and Knowledge of User Intent

What changes like co-occurrence and Hummingbird mean for keyword research and strategy is that they are not as simple as they once were. No longer can you simply research what terms get the most traffic and build pages optimized to each one you would like to target. You still need to do keyword research to understand the universe of keywords your potential users might be looking for, and the relative scale of the opportunity in ranking well on each of those terms. You must also be able to uncover the intent behind your users' searches and be able to build quality pages and content that meet those needs.

Not only must your pages be well written and researched, and meet the actual needs of your potential users, but each page must present a coherent theme so that the engines can truly understand what need it is attempting to satisfy, and thus what body of related keywords it should be considered to rank for. Searchmetrics, which is mentioned elsewhere in this chapter, has tools in its Content Optimization section that can help you understand the terms and phrases that are highly relevant and related to your optimal keyword targets.

It is also vital that you understand what your customers' needs are and whether

they are able to meet those needs with the information they find on your pages. There are lots of ways to understand more about what your users want and whether they are getting it from your site—from leveraging services like Usertesting.com to simply spending time talking to your customers about how and why they use your site. It is becoming more and more important to ensure that your content is meeting these needs as simply as possible, that it includes keywords relevant to those needs, and that it is integrated into well-designed, clearly themed pages.

Keyword Research Options

A wide variety of options are available for performing keyword research, including tools provided by the search engines, tools developed by third parties, and tools for complex keyword analysis of terms culled during research. In this section, we will review each of these, but first let's provide some perspective on how to use these tools.

When you are using the various keyword research tools to brainstorm keywords, it is important to keep in mind that they are all based on relatively limited data. You should be aware that each tool will provide different search count numbers than the other tools. Rather than focusing on the exact search count of various terms, you should think of each tool as a good way to get a general comparison of two search terms, or of groups of terms.

For example, if you compare two terms and see that one term is more popular than the other because it returns a higher search count, you at least know that Term A may be more popular and searched for more often than Term B, but you can treat the search count values as only (rough) estimates.

If you are new to keyword research, consider starting with the Google AdWords Keyword Planner (which is free with an AdWords account) and at least one of the paid tools, such as KeywordDiscovery.com. This will give you a rich, varied data set with which to begin your keyword research. Over time you can experiment with the other tools and adjust your process as you find tools that you prefer for one task or another.

In addition, as you look at these keywords, keep in mind the semantic aspects of search, and the probable intent of the person performing the search. You don't

want to simply slap popular search terms on various pages of your site. You are using the keyword research information to provide you with insight into how your potential customers think, but you need to take into consideration many other inputs as well.

Keyword Research Data from Search Engines

Search engines provide a number of tools that can help you with keyword research. Many of these are not designed specifically for that purpose, but they can be used to obtain interesting keyword research information if they are used in the right manner. The data in these tools shows the number of pages that are related to a search phrase, not the number of searches on the phrase. This is still a useful indicator of the importance of a keyword phrase, as more web pages tend to get built for more popular topics.

Related terms

Several of the engines offer “related” terms, including Google, Yahoo!, Bing, Ask, and Yippy (which shows related terms in clusters, as shown in Figure 5-2). This data can be invaluable if you’re looking to find related terms that may not have come up through competitive analysis or brainstorming.

Figure 5-2. Yippy related terms clusters

Common usage and phrase combinations

Using a search with the * (asterisk) character can give you a good idea of what terms/phrases commonly precede or follow a given term/phrase. For example, using * *ringtones* can show you phrases that are commonly associated with the term *ringtones*, as shown in Figure 5-3.

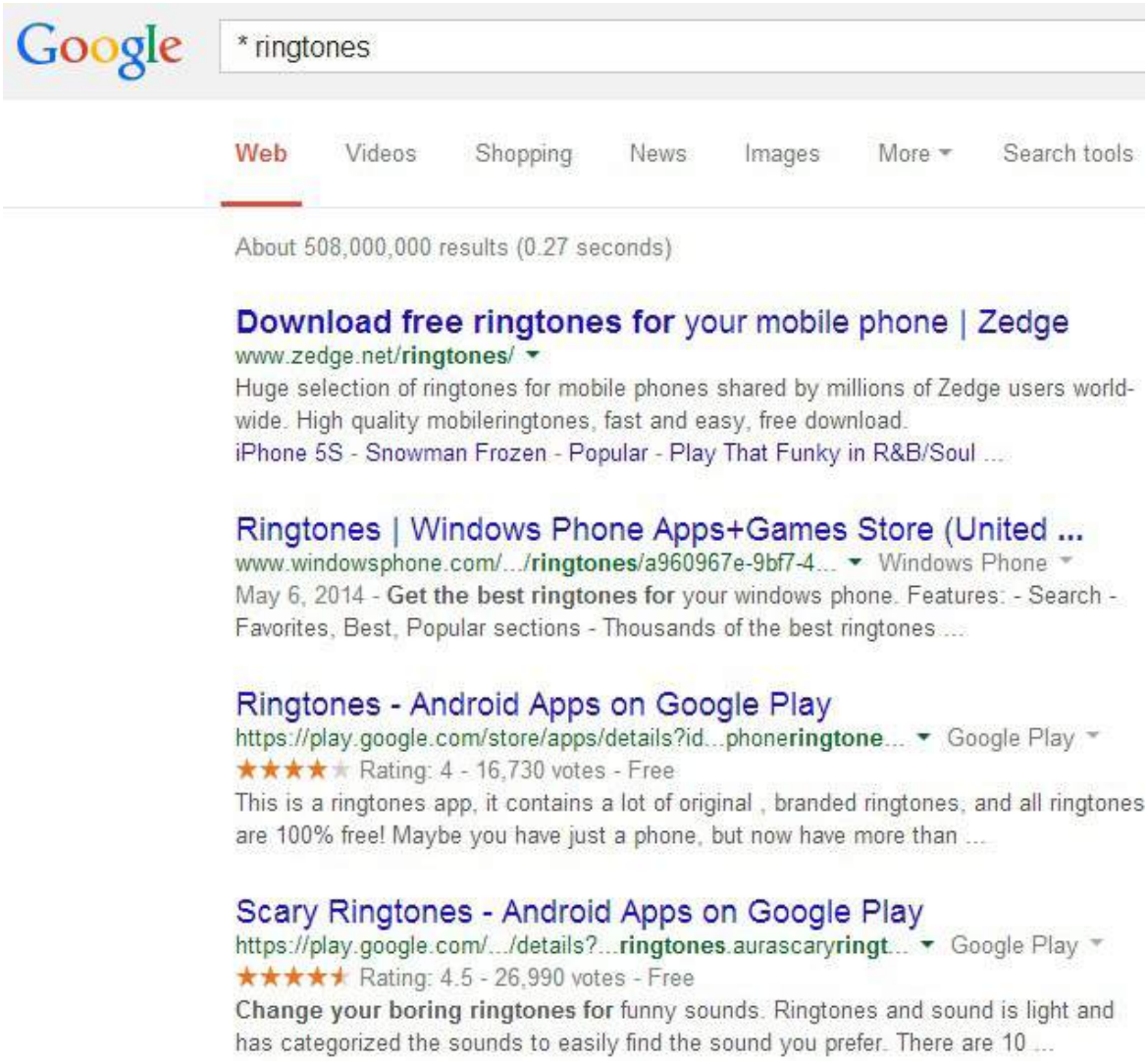


Figure 5-3. Finding common phrases

Frequency of recent usage

Using the very cool Google date range operator, shown in Figure 5-4, you can determine how many times in the past day, week, month, or year new content related to your term was added to the Google index. The easiest way to do this is to click on “Search tools” above the Google results. Once you do that, you can pick from “Any time” (which is the default), “Past hour,” “Past 24 hours,” “Past week,” “Past month,” “Past year,” and “Custom range.” This will limit you to the results that were added to the index during the referenced time frame.

Figure 5-4. Google pages indexed in past 24 hours

Picking a custom range provides you with a calendar method for specifying the time interval you want to focus the search on. For example, you might pick November 1, 2014, to December 24, 2014, if you wanted to see what happened during the 2014 holiday season. For additional flexibility, you can perform a normal search, get your result, and add a parameter to the end of the results page URL, using the operators shown in Table 5-1.

Table 5-1. Google date search operators

`&as_qdr=d`	Past 24 hours

`&as_qdr=d4`	Past four days
`&as_qdr=w`	Past week
`&as_qdr=w5`	Past five weeks
`&as_qdr=m6`	Past six months
`&as_qdr=y2`	Past two years

This can give you some seasonal data if you follow it closely, and it can also show you who is producing content in your arena. For example, try a search for *President Obama (past 24 hours)*. You can also get information on activity level from Yahoo! News, as shown in Figure 5-5.

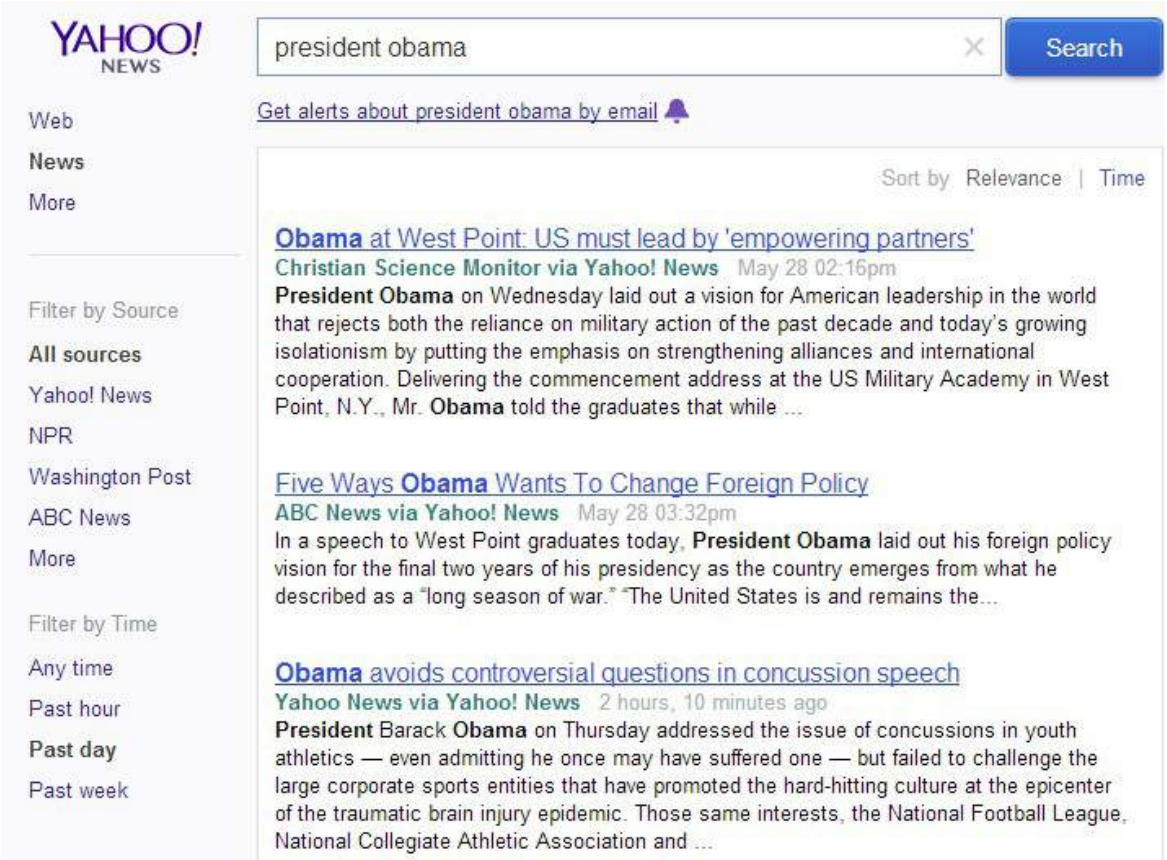


Figure 5-5. Yahoo! News activity level

Both Google News and Yahoo! News are great places to do a bit of digging into anyone who is publishing press releases or getting news coverage on the terms/phrases you might be researching. If there's a lot of activity in these arenas (and it is not all press release spam), you can bet the terms are going to be even more competitive. For example, you can find SEO news in Google News and in Yahoo! News.

You can combine all of this data to form a very well-rounded view of a particular term or phrase, and although it is probably overkill for most keyword research projects, it is certainly a valuable exercise and something to monitor closely if you're basing a lot of your success off of a single search query (or just a handful of queries). Even if you're just trying to get a better sense of what's going on infrequently and informally, these pieces of the keyword puzzle can be remarkably valuable.

Keyword Research Data from Tools

It is great to get this data from search engine queries, and it can certainly help you get a sense of the importance of a given keyword. However, a large array of tools exists to give you direct insight into the volume of searches performed on specific keywords, and also to help you discover new keywords to consider.

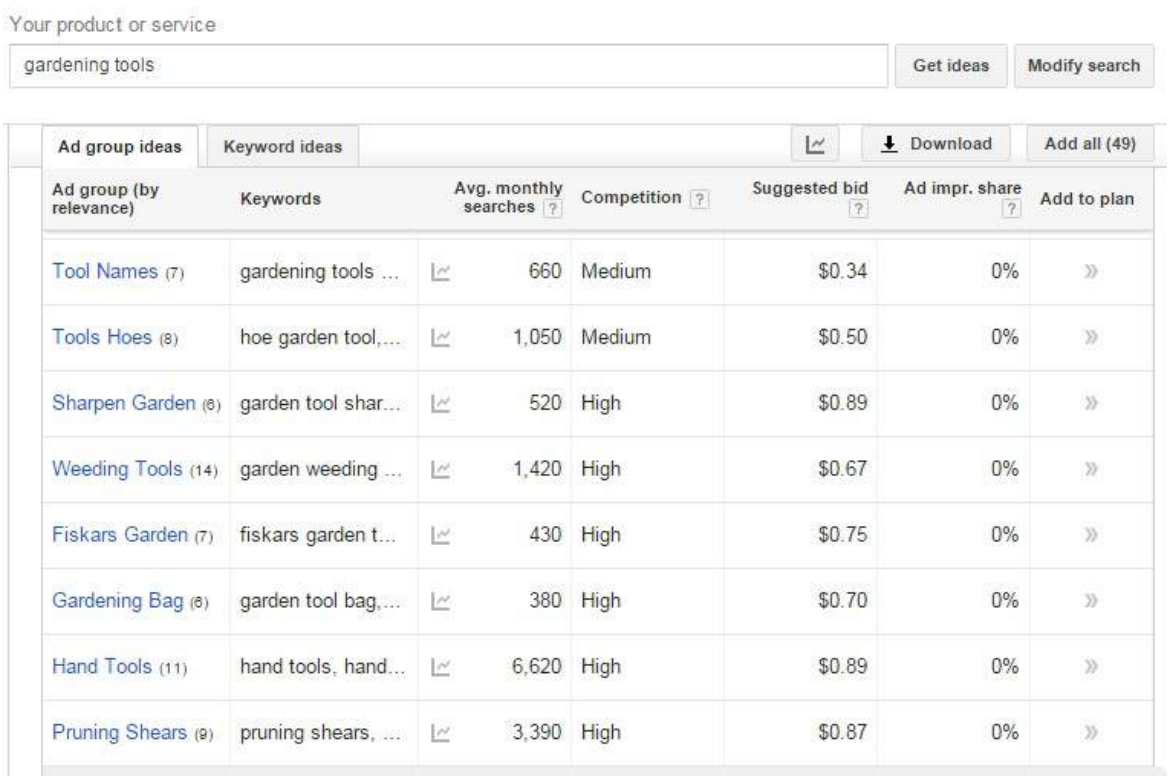
The following sections present an overview of many of these leading tools.

Google AdWords Keyword Planner and Traffic Estimator

Google provides a couple of tools specifically designed for use in keyword research. Although they are primarily meant to help their paid search customers, they can also be used to obtain information for organic search. These tools are most effectively utilized when coupled with an active AdWords campaign.

What the Keyword Planner provides

Google's AdWords Keyword Planner provides related terms, search volume estimates, search trends, and ad cost estimates for any keyword or URL that you enter (see Figure 5-6).



Ad group (by relevance)	Keywords	Avg. monthly searches	Competition	Suggested bid	Ad impr. share	Add to plan
Tool Names (7)	gardening tools ...	660	Medium	$0.34	0%	»
Tools Hoes (8)	hoe garden tool,...	1,050	Medium	$0.50	0%	»
Sharpen Garden (6)	garden tool shar...	520	High	$0.89	0%	»
Weeding Tools (14)	garden weeding ...	1,420	High	$0.67	0%	»
Fiskars Garden (7)	fiskars garden t...	430	High	$0.75	0%	»
Gardening Bag (6)	garden tool bag,...	380	High	$0.70	0%	»
Hand Tools (11)	hand tools, hand...	6,620	High	$0.89	0%	»
Pruning Shears (9)	pruning shears, ...	3,390	High	$0.87	0%	»

Figure 5-6. Google AdWords Keyword Planner

The Keyword Planner provides multiple ways to search: based on words/phrases, based on websites/landing pages, and based on categories. In the “Your product or service” box, you can enter keywords or phrases (one per line) and the Keyword Planner will return related keyword ideas (note that you have to click on the “Keyword ideas” tab; the “Ad group ideas” is what displays by default). The output of a search will show you:

Keyword (by relevance)

Shows the related keywords in a list, including the phrase or phrases you entered.

Average monthly searches

Shows the search query volume for the keyword for the locations, languages, and Google properties you defined.

Monthly searches bar chart

Shows the search volume for the keyword broken out by month for the last 12 months (requires that you hover the mouse over the small gray chart icon immediately to the left of the numbers in the “Avg. monthly searches” column).

Competition

Displays the relative competitiveness of the keyword (in paid search).

There are a number of settings you can use to tune your search. These include:

Include/Exclude

Provides a way to specify or exclude additional terms. For the include option, the Keyword Planner will show you only suggestions that included these terms (see Figure 5-7). The exclude option allows you to add a negative keyword for any keyword phrase that does not pertain to your business. This feature is not necessarily useful for researching keywords for organic search; rather, it is more valuable when you are planning your AdWords account bids.

Locations

Allows you to set the country used for the “Avg. monthly searches” part of the results.

Language

Lets you set the default language to use.

Search networks

Allows you to specify whether you want traffic estimates returned for only Google searches or whether you also wish to include Google search partners. Google search partners include other non-Google websites like AOL and other Google properties such as Maps, Video, and other Google sites.

Date range

Shows search volumes for the past 12 months by default. You can also specify a custom range of months that you would like to see and compare volumes for two different time ranges.

NOTE

You must log in to your Google AdWords account to use the Keyword Planner, but fortunately, you don’t have to spend money on AdWords to access this free tool.

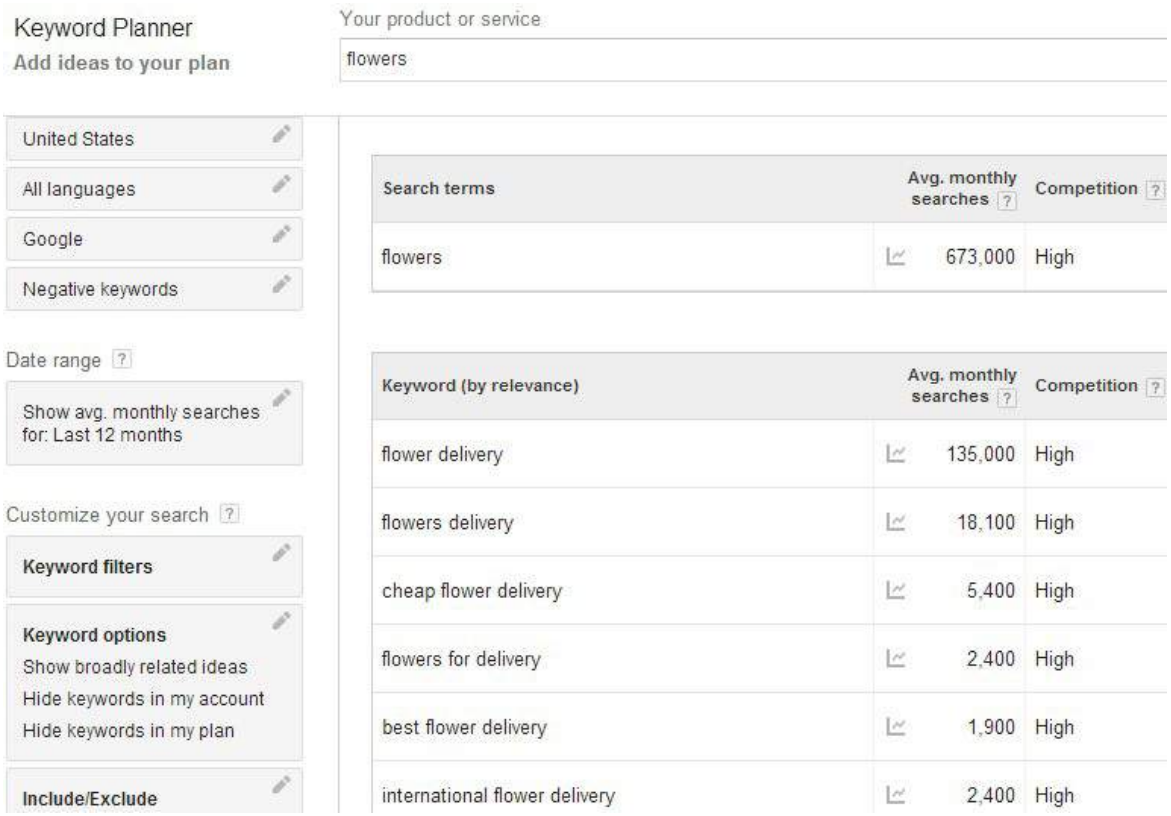


Figure 5-7. Specifying a required term in the Keyword Planner

If you enter a web page URL, the Keyword Planner will show you keywords related to that website or specific page (see Figure 5-8).

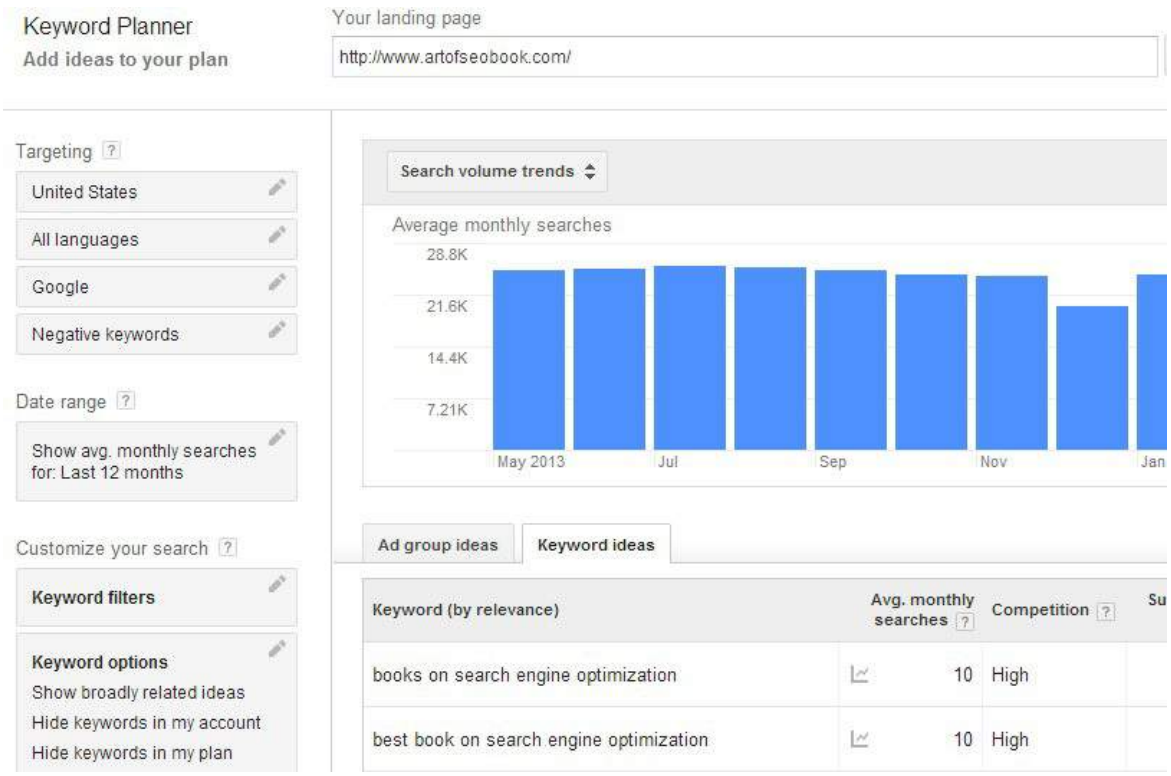


Figure 5-8. Keyword Planner site-related keywords

What the traffic forecasts provide

Within the Google AdWords Keyword Planner is an option to "Get click and cost performance forecasts" (formerly known as the Traffic Estimator) that allows you to get forecasts of traffic on different keywords (i.e., the potential click-throughs you may see to your site, instead of just the number of impressions provided by most keyword research tools); see Figure 5-9.

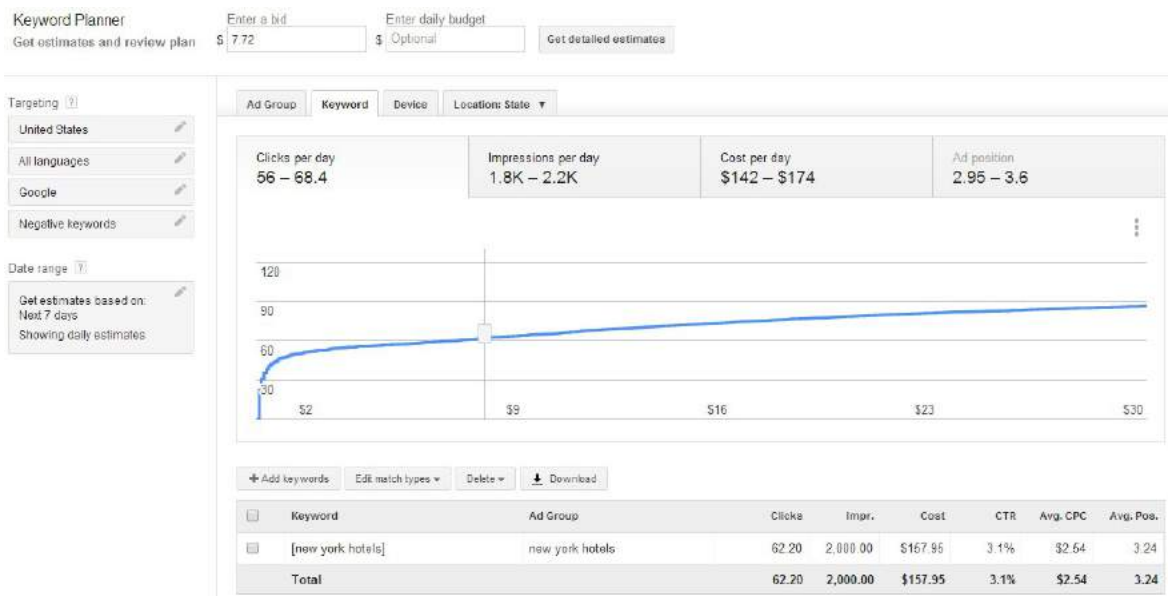


Figure 5-9. Google AdWords Keyword Planner's click and cost performance forecasts

When you specify one or more keywords to get traffic estimates on, the Keyword Planner will return the predicted number of clicks per day, impressions per day, cost per day for those clicks and average ad position, as well as additional breakdown on a per-keyword basis, including clicks, impressions, cost, click-through rate, average cost per click, and average ad position. The cost information can provide you with additional insight into how competitive a keyword is in organic search as well. You can get more exact estimates by selecting specific budgets or costs per click.

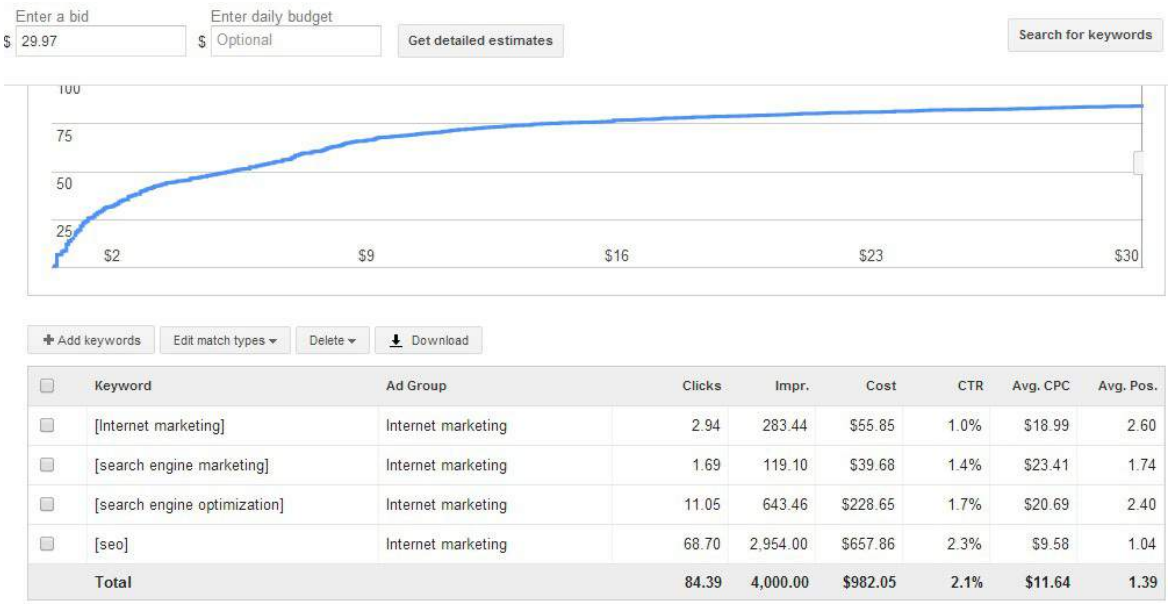
When using the Keyword Planner's click and cost performance forecasting feature for keyword research, it is best to enter your keywords as "Exact match," meaning the keywords specified were exactly the ones the Google user searched on.

After you've entered your keywords, you can leave "Enter daily budget" blank. Select your language, search network, and the location you're targeting (for U.S.-focused campaigns, use the default of "Countries and territories" and enter "United States").

You'll then see data for each keyword you entered. Useful data for keyword research purposes includes Estimated Clicks/Day and Estimated Cost/Day. You can compare each keyword's estimated clicks to see which term is more likely to

be searched for and clicked on than others.

In the results shown in Figure 5-10, at the maximum possible bid, *internet marketing* is estimated to have 3 clicks per day, while *search engine marketing* has 2, *search engine optimization* has 11, and *seo* has 68. Based on this data, it is clear that *seo* is the most popular of these options and is likely to be one of the more competitive terms. In this particular case, an additional factor enters into the equation, because *seo* is a "trophy term" on which people put an extra focus for branding reasons. Nonetheless, the value of the traffic data is considerable.



Keyword	Ad Group	Clicks	Impr.	Cost	CTR	Avg. CPC	Avg. Pos.
[Internet marketing]	Internet marketing	2.94	283.44	$55.85	1.0%	$18.99	2.60
[search engine marketing]	Internet marketing	1.69	119.10	$39.68	1.4%	$23.41	1.74
[search engine optimization]	Internet marketing	11.05	643.46	$228.65	1.7%	$20.69	2.40
[seo]	Internet marketing	68.70	2,954.00	$657.86	2.3%	$9.58	1.04
Total		**84.39**	**4,000.00**	**$982.05**	**2.1%**	**$11.64**	**1.39**

Figure 5-10. Google AdWords Keyword Planner traffic forecasting results

Where the Keyword Planner gets its data

The Google AdWords Keyword Planner gets its data from Google's search query database.

How the Keyword Planner is useful

The Keyword Planner offers some useful information about your keyword campaigns, such as suggestions for similar keywords, an estimate of the keyword's popularity, ad costs and positions, general search volume trend information, and keyword campaign suggestions for your site or your competitor's site. The tool is great for compiling a lot of general information about a keyword.

The click and cost performance forecasting feature provides a rough estimate of your keyword's click-through rate. Based on the estimated clicks per day, you can get a relative idea of which of your keywords are the most popular and can potentially bring you the most traffic. Practitioners should use other tools to cross-reference these figures, as these numbers can be inaccurate.

Cost

The Google AdWords Keyword Planner is free to use as long as you have a Google AdWords account.

Google Trends

Google Trends allows you to view the popularity trend of a keyword over time and by geography, as well as compare two or more search terms to each other to see relative popularity and seasonality/trending over time. If you enter the terms into the search bar and separate them with commas, you'll see the requested terms' trend history depicted in different colors on a graph spread over a certain time period. You can modify the results by changing the time period and/or region (see Figure 5-11).

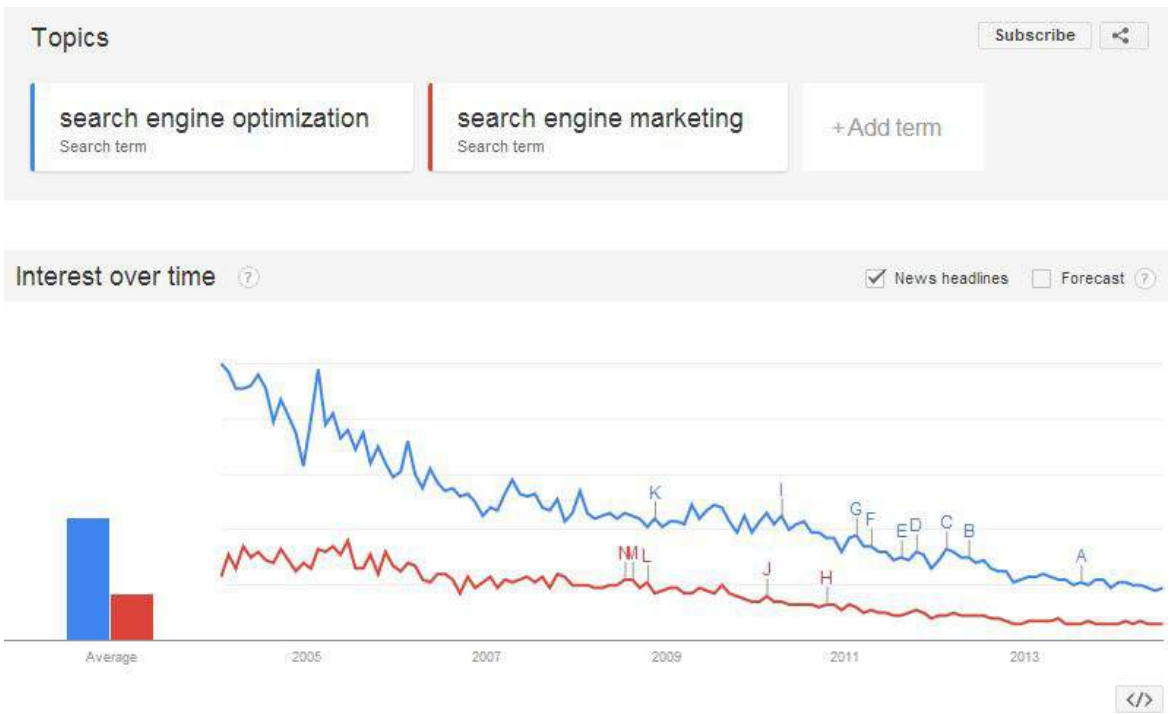


Figure 5-11. Google Trends sample output

With Google Trends, users can also see Google's estimate of which cities, regions, and languages performed the largest number of searches for a particular keyword (see Figure 5-12). Experienced marketers often feel that this data is imprecise (and occasionally inaccurate) because more accurate data from analytics and search advertising campaigns have often contradicted the results.

Figure 5-12. Google Trends top cities data

Lastly, plotted on each graph are a few articles/search results related to your keyword query, which correlate to peaks and valleys in the historical search popularity.

Where Google Trends gets its data

Google Trends gets its data from searches performed on Google.

How Google Trends is useful

Google Trends is a great, easy tool for comparing keywords and identifying which term is more popular than the other; in addition, you can examine this data over many years with seasonality factored in. Although Google Trends doesn't supply absolute search volume figures like the Google AdWords Keyword Planner, the graphs are simple to understand and provide a perfect visual of search trends over a particular period of time (note that this works only with relatively popular terms, not with longtail search terms). You can also sign up for email alerts from Google Trends for the topics/keywords you're researching, another convenient feature that can help keep you up to speed on

the search trends you're targeting.

Cost

Google Trends is free to use.

Bing Keyword Research

The Bing Keyword Research tool generates keyword suggestions based on a search term or website you enter.

Entering a keyword in the search box will return data that includes search phrases containing the keyword you provided, along with the preceding month, typical click-through rate percentages, and average cost per click. For example, a search for *ice cream* returns *ice cream maker*, *ice cream recipes*, and *ice cream shop*. As you can see in Figure 5-13, according to Microsoft, the term *ice cream* had 856,543 searches in the month prior to when this screenshot was captured.

The Export to Excel option allows you to pull the collected data into a spreadsheet. Although the CTR% (click-through rate) and CPC (cost per click) columns are intended for paid search customers, they can also provide some indication of SEO value. You can multiply the CTR by the search volume to get a sense of how many clicks a high-ranking paid search result might get (comparable organic results will get three to four times more clicks), and the CPC provides some indication of the competition for ranking on the term. You can also obtain demographic data as shown in Figure 5-14.

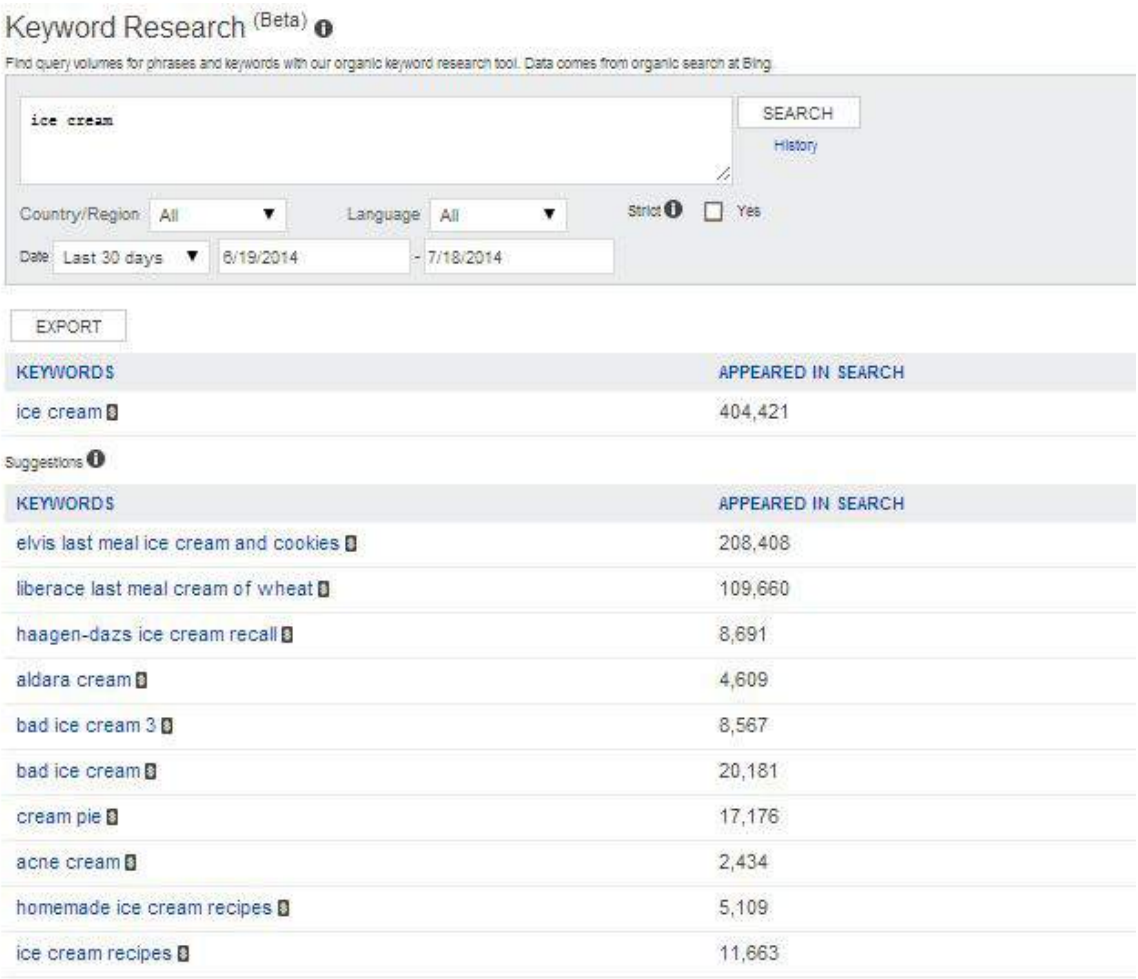


Figure 5-13. Bing Keyword Research tool basic output

Figure 5-14. Bing Keyword Research tool demographic settings

If you enter a URL into the search bar, the tool will return keywords related to the website selected.

Where Bing Keyword Research gets its data

The Bing Keyword Research tool obtains its data from Microsoft's Bing search query database.

How Bing Keyword Research is useful

This tool is useful in generating keyword suggestions based on a keyword you are targeting or on your site's URL. You can also enter a competitor's URL and see what the keyword suggestions are for its site. One important feature of this tool is that it shows actual numbers from the Bing search query database.

Cost

The Bing Keyword Research tool is free, although you do have to create an account with Bing Ads and provide credit card information in the event that you advertise on the Bing network.

KeywordDiscovery

A very popular third-party tool for keyword research is Trellian's KeywordDiscovery, which offers the following features:

Keyword research

When you enter a keyword or phrase in the search bar under the Research section, KeywordDiscovery displays the most popular search terms that include the keywords you provided, along with a count of how many searches were performed for those keywords in the past 12 months (see Figure 5-15).

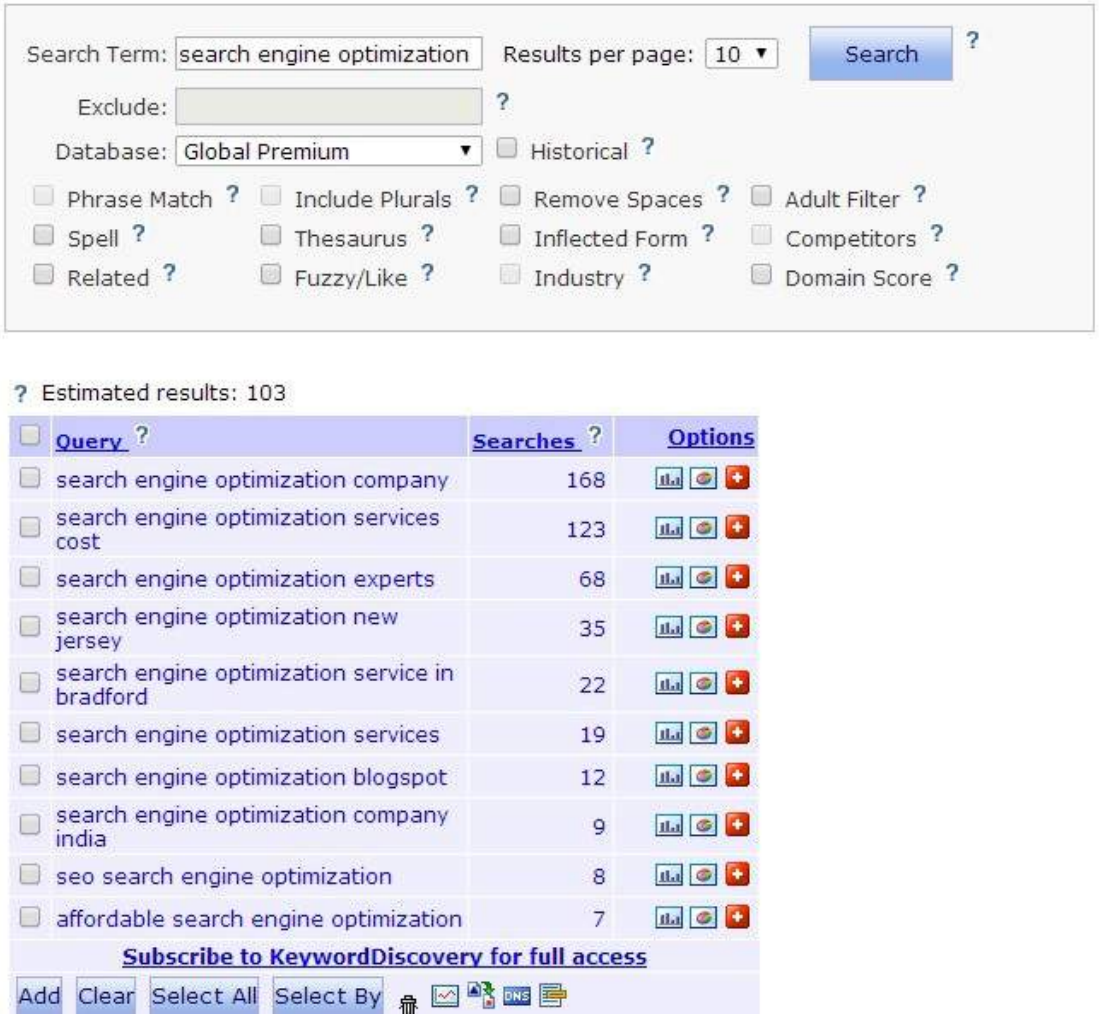


Figure 5-15. KeywordDiscovery basic output

Seasonal search trends

If you click on the little bar graph icon next to the number of searches for a query, you'll see a graph of the search trends for that keyword over the past 12 months. You can mouse over each bar and see the number of searches for that time period, and you can sort the chart by historical data (how many searches in the past year), monthly data (number of searches broken down into each month), trends (a graph of the search trends over the past year), combination (a graph of historical global and global premium search data; Figure 5-16 shows a definition of these terms), and market share (a breakdown of which search engines were used to search for the query).

Figure 5-16. KeywordDiscovery seasonal search trends

Spelling mistake research

Typing the query *spell:<keyword>* as the search term will return spelling variations for that keyword, the number of times the keyword has been searched for (searches), and the keyword results for your search (queries).

For example, *spell:optimization* returns results such as *optimation*, *optimazation*, and *optimisation*, as shown in Figure 5-17.

Search Term: optimization Results per page: 10 Search ?

Exclude: ?

Database: Global Premium Historical ?

Phrase Match ? Include Plurals ? Remove Spaces ? Adult Filter ?

Spell ? Thesaurus ? Inflected Form ? Competitors ?

Related ? Fuzzy/Like ? Industry ? Domain Score ?

? Estimated results: 33

Word ?	Similarity ?	Searches ?	Options	Queries ?
optimization	100%	1,327		259
optimizacion	75%	266		73
optimisation	75%	253		60
optimizador	25%	60		29
optimizations	75%	33		1
optimizaciÃ²n	25%	15		3
optimazation	75%	11		3
optimización	50%	10		3
optimizacija	25%	8		4
optimizaciÃ³n	25%	8		3

Subscribe to KeywordDiscovery for full access

Add Clear Select All Select By

Figure 5-17. KeywordDiscovery spelling mistakes output

Related keywords

Typing either *related:<keyword>* or *crawl:<keyword>* will return keywords that are related to the term you provided. For example, typing in *related:seo* returns results such as *internet marketing*, *video*, and *internet consulting*. You can see an example of this in Figure 5-18.

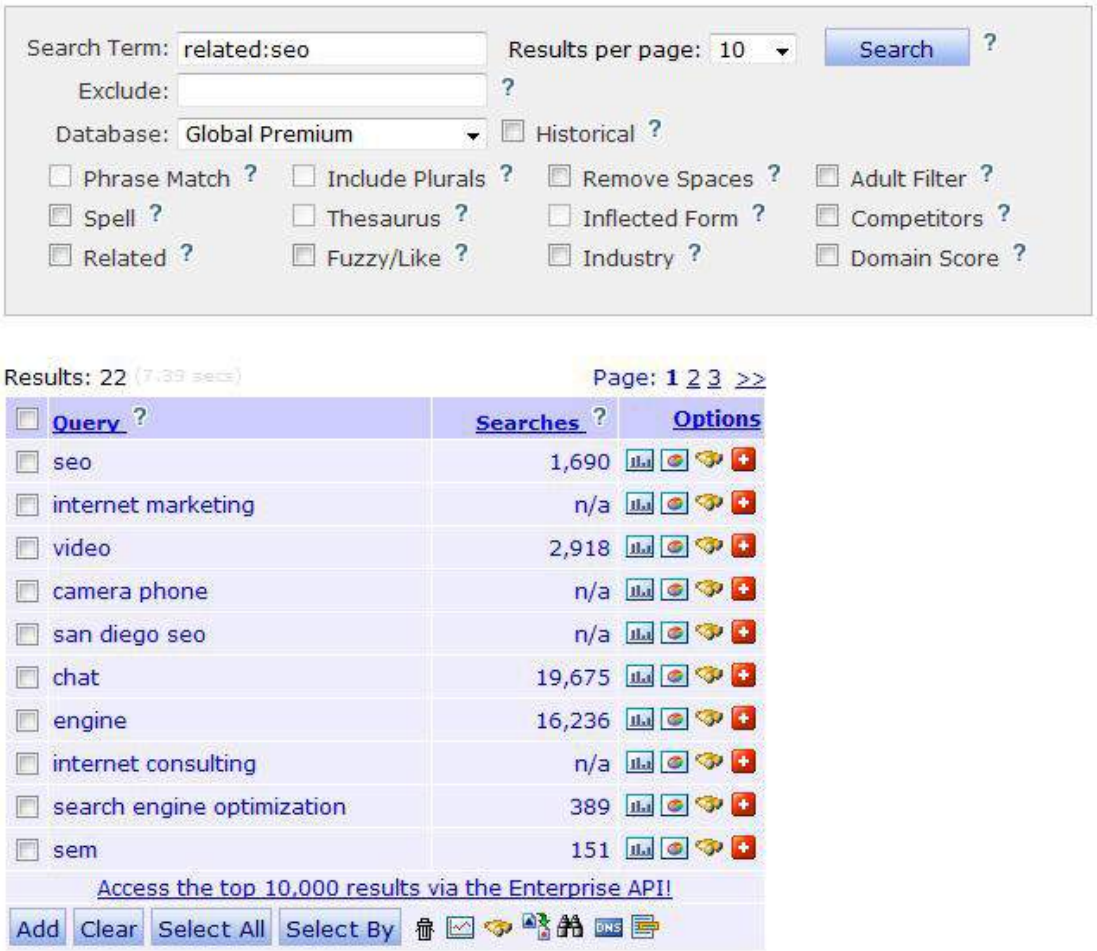


Figure 5-18. KeywordDiscovery related keywords output

Keyword density analysis

This feature checks how often keywords are found on the URL you provide, assigns a keyword density percentage to those keywords, and lists the number of searches performed for each term.

We do not recommend using keyword density as a metric to judge a page's keyword targeting. The search engines use far more sophisticated analyses of keywords for their algorithms, and relying on rough counts such as this can seriously mislead you. See "Keyword Targeting" for more on how to effectively target keywords on the page.

One good use for the keyword density analysis feature is to enter a

competitor's URL into the search bar and see what keywords the site is targeting. It is a great tool to use for competitive research.

Domain Researcher tool

This tool requires an Enterprise subscription. It allows you to search for available domains that are based on popular keyword search terms. These domains have high traffic potential, as the tool shows how many users have searched for the URL. The tool is great if you want to register other domains in your industry and want these domains to be keyword-rich (see Figure 5-19).

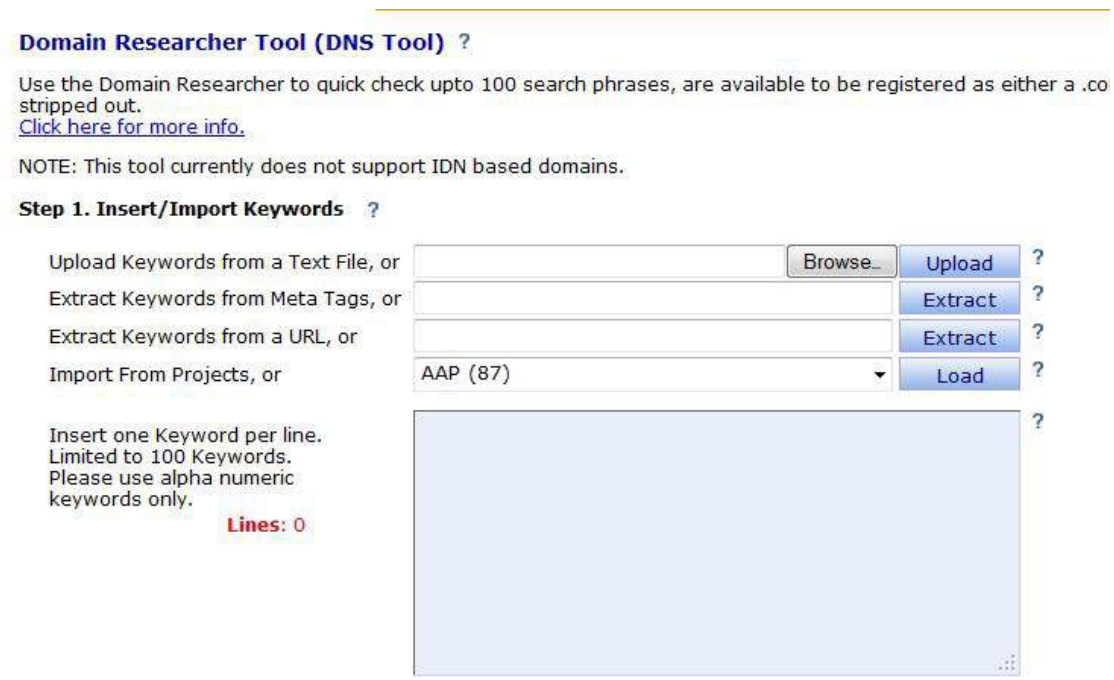


Figure 5-19. KeywordDiscovery Domain Researcher tool

Competitive intelligence reports

Trellian, which powers KeywordDiscovery, also offers various competitive intelligence reports (which require a separate subscription). These reports include:

Link Intelligence

Identifies which links are sending traffic to your competitors

Search Term Intelligence

Identifies which search terms/phrases are driving traffic to your competitors

Search Engine Intelligence

Identifies which specific search engines send traffic to your competitors

PPC Campaign Intelligence

Identifies which search terms your competitors are bidding on

Referrer Intelligence

Provides information about specific sites that are referring traffic to your competitors

Popularity Index

Monitors the popularity index (which is based on the number of unique sessions a domain receives) of your competitors

Ranking

Provides a view of which terms your competitors are ranking for, the rank of these terms, and any changes in ranking over the past 30 days

Meta Keywords

Provides an analysis of your competitors' meta keywords

Competitive Intelligence Executive

Provides information about every Competitive Intelligence report available, as well as several subreports

Where Trellian gets its data

Trellian derives its keyword data primarily from aggregated historical global data purchased from ISPs. Trellian also uses a panel of 4.4 million users to collect its global premium data. The company claims that the global premium data removes the bias that various spiders introduce into data from other sources.

How KeywordDiscovery is useful

As we mentioned earlier, KeywordDiscovery offers a multitude of tools that are great for keyword research. Trellian also offers various tools that are useful for competitive research. You can almost think of KeywordDiscovery as a one-stop shop for research, because it offers a diverse set of tools, but as with many of the other keyword research tools we've discussed here, its data sources are limited, and you need to take this into account in your use of the tool.

Cost

KeywordDiscovery offers different subscription options that range from a standard monthly subscription for $69.95 to a yearly Enterprise subscription for $4,752 (pricing as of July 2014). Competitive intelligence reports range from $99.95 per month per domain (plus a $150 setup fee) to $995 per year per domain. The free tool with limited features is also available. We recommend reviewing the options and choosing the package that will work best for your company.

Experian Hitwise

Experian Hitwise offers a wide range of competitive and web statistics via its service. One component of the Experian Hitwise suite, Hitwise Search Intelligence, is a powerful keyword research tool for analyzing the long tail of search data. It provides extensive insights into how people have successfully searched for products and services across all major search engines, including the breakdown of paid and organic traffic (you can read more about the Experian Hitwise product offering in "Tying SEO to Conversion and ROI").

Hitwise Search Intelligence provides the following features:

- Timely information on search terms your specific competitors use
- Market-specific results, for taking advantage of cultural differences on how people search locally

- Information on terms that users have “clicked on” before and after visiting a website or industry (also known as *upstream* and *downstream activity*)

Figure 5-20 shows an example of the most popular search terms used by searchers who ended up on travel sites.

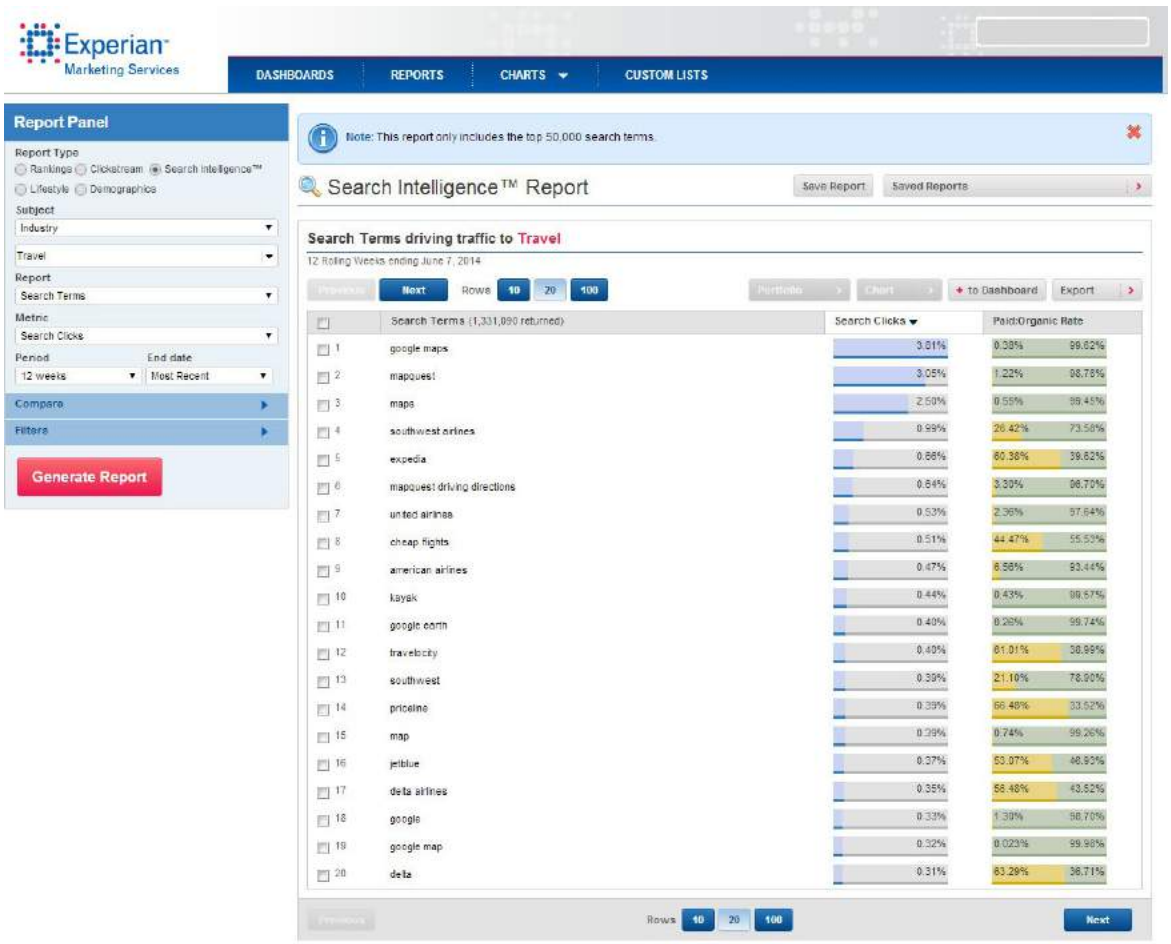


Figure 5-20. Hitwise “popular search terms” report

The ability to see actual keyword data on your competitors is an extremely potent feature. You can see what is working for them and what is not. This type of information is very powerful and can give you a significant edge over the competition.

You can also focus more directly on search term suggestions, as shown in Figure 5-21, which depicts a screenshot for terms related to *travel*.

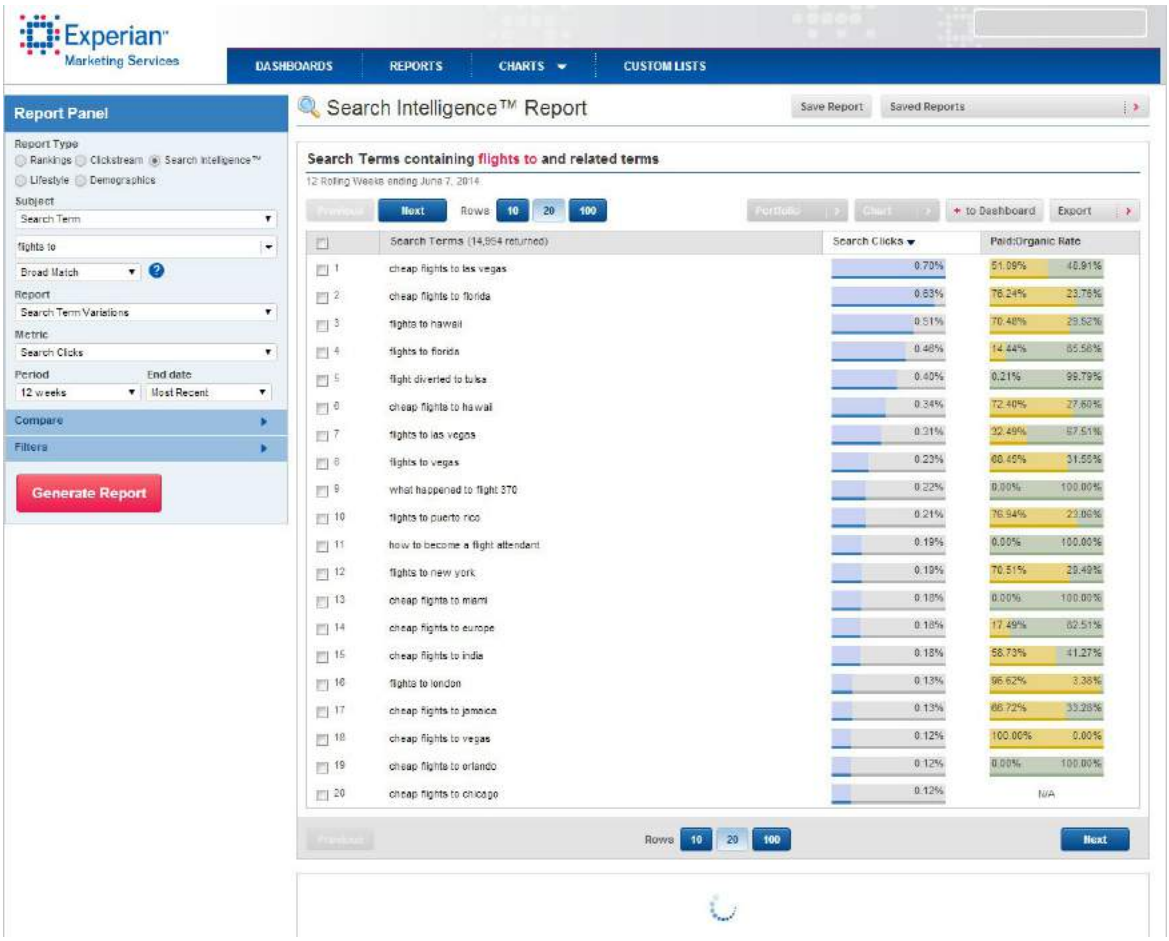


Figure 5-21. Hitwise search term suggestions tool

Where Hitwise gets its data

Hitwise derives its data from more than 25 million people's interaction with the Internet (10 million from the United States). Hitwise collects anonymous Internet usage information from a combination of ISP data partnerships and opt-in panels.

How Hitwise is useful

The data is presented in percentages (the volume of searches, its success rate with searchers), which makes it very easy to compare the relative popularity of various keywords, but difficult to estimate the actual number of searches for a given term.

Cost

Hitwise is not an inexpensive tool. The website does not list pricing information, but you should be ready to spend $20,000 if you plan to engage with this tool. Bear in mind that we have presented only a snapshot of its features, and the competitive data is extremely valuable, not just to the SEO team but to all marketing disciplines across your organization. If long-term use of the tool doesn't make sense for your team's needs, Hitwise also offers one-off analysis services for around $5,000 a report.

comScore Search Planner

Like Hitwise, ComScore Search Planner is a tool that provides a wide range of data as a result of monitoring the behavior of actual users on the Internet. This data includes details on search terms used, as well as competitive search term analysis.

What comScore Search Planner provides

ComScore Search Planner comprises eight modules, but two of them are particularly useful for keyword research:

Site Profile (for Site(s) X)

This module tells you what search terms and search engines are driving the most traffic to your site, to your competitor's site, and within your category.

Profile Search Terms

This module tells you the demographic profile of people searching on a set of search terms, as well as what sites these searchers tend to visit.

Figure 5-22 shows the highest-volume terms specific to the airline category.

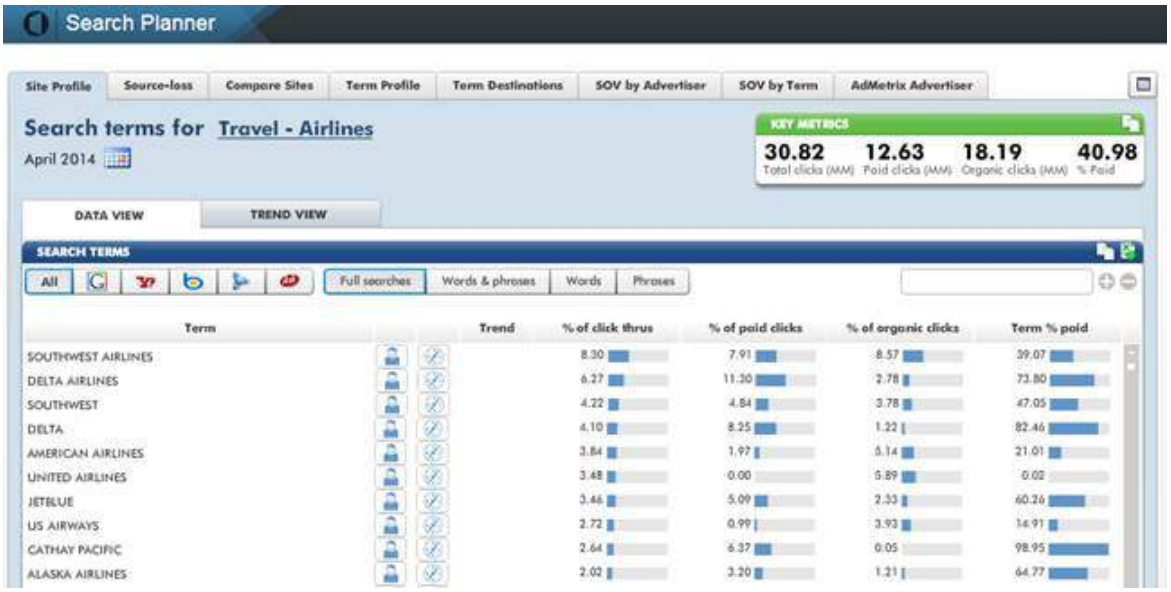


Figure 5-22. ComScore "airline search terms" report

You can also view similar data specific to a competitor's site, so you can see what search terms are driving its traffic. Another useful feature is the ability to view search trends for an industry. This helps with the identification of seasonal behavior, as you see in Figure 5-23.

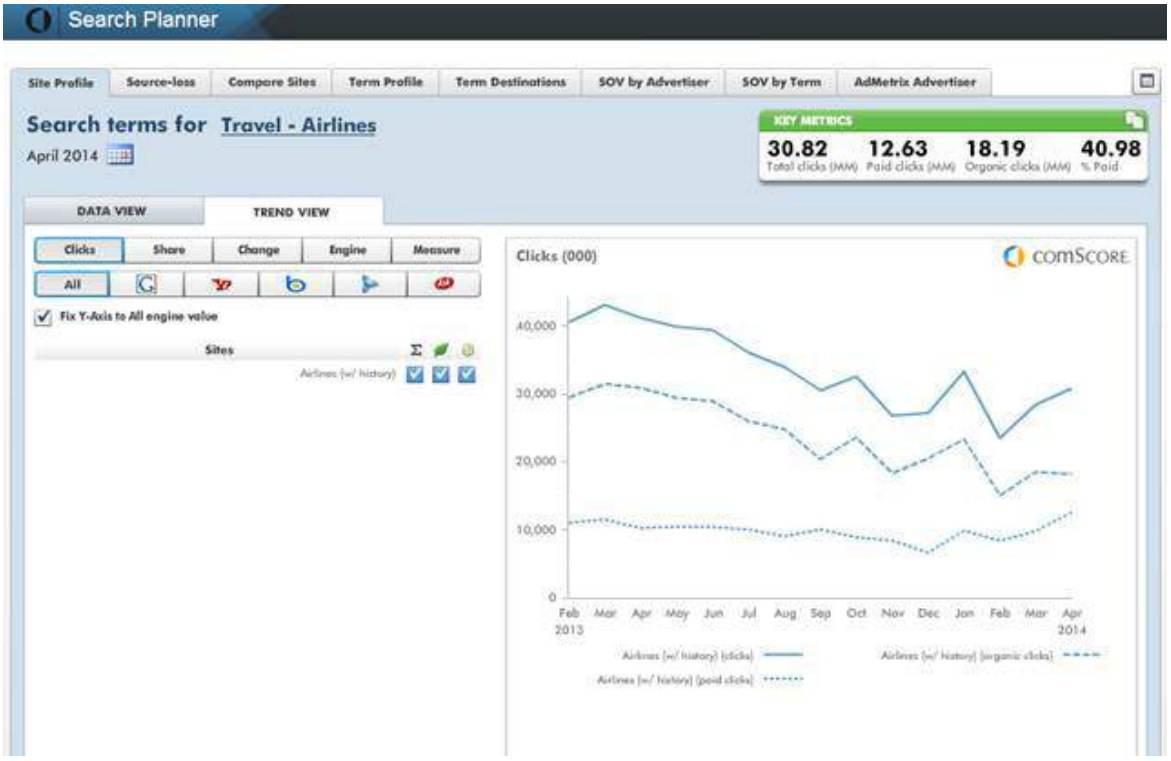


Figure 5-23. ComScore search trends report

Where comScore gets its data

comScore monitors the behavior of approximately 2 million users. These users have voluntarily joined comScore's research panels in return for free software, free Internet-based storage, or chances to win prizes. Sites can also opt in to adding comScore tracking on their site using unified tags. Some sites do this because it tends to result in higher numbers and better data that they can then show to potential advertisers. This helps them sell online display advertising and obtain higher advertising rates.

How comScore Search Planner is useful

comScore Search Planner data is presented in percentages (the volume of searches; its success rate with searchers), which makes it very easy to compare the relative popularity of various keywords, but difficult to estimate the actual number of searches for a given term.

Cost

You must contact comScore directly to get pricing for the Search Planner tool. The primary audience for the product is mid-size to large companies with developed SEM/SEO strategies, but the company has some smaller clients as well.

WordStream

The fast, powerful keyword tools from WordStream are completely free (access was previously restricted to paid subscribers). The Free Keyword tool delivers thousands of keyword suggestions, including longtail keywords, from a constantly updated database of more than a trillion unique search queries. The Keyword Niche Finder helps you discover your most profitable pockets of keyword opportunities. This free tool delivers value-added, structured keyword suggestions, preorganized into groups of related searches, ready to use in search marketing campaigns. Or, if you already have a list of keywords, drop them into the Keyword Grouper to get back organized keyword niches: see all related variations of a keyword at once, or drop these clusters of keywords into an ad group in AdWords.

What WordStream provides

WordStream provides a suite of five different tools for keyword research:

Keyword Suggestion

This is the basic tool for generating a list of keyword suggestions along with search volume metrics.

Keyword Niche Finder

This tool is useful when you are building out a list of new topics from which to create content.

Keyword Grouper

Keyword Grouper is used to mine keyword data for organic search referrals and trends.

Negative Keyword

This tool is mainly used in relation to PPC campaigns, but can assist in generating a list of terms that it is not desirable to match for.

SEO Content Creation plug-in for Firefox

This plug-in suggests topics and keywords for new SEO pages, and tracks keyword usage as you type.

Figure 5-24 shows the output for keyword research related to the phrase *catering*.

The Keyword Niche Finder tool finds keywords by topic area. This is useful when you're looking for new topic or subtopic areas to create content for. Use this when deciding on new types of content for a site, or when first building a site. Figure 5-25 shows the output related to *used cars* and *cheap*.

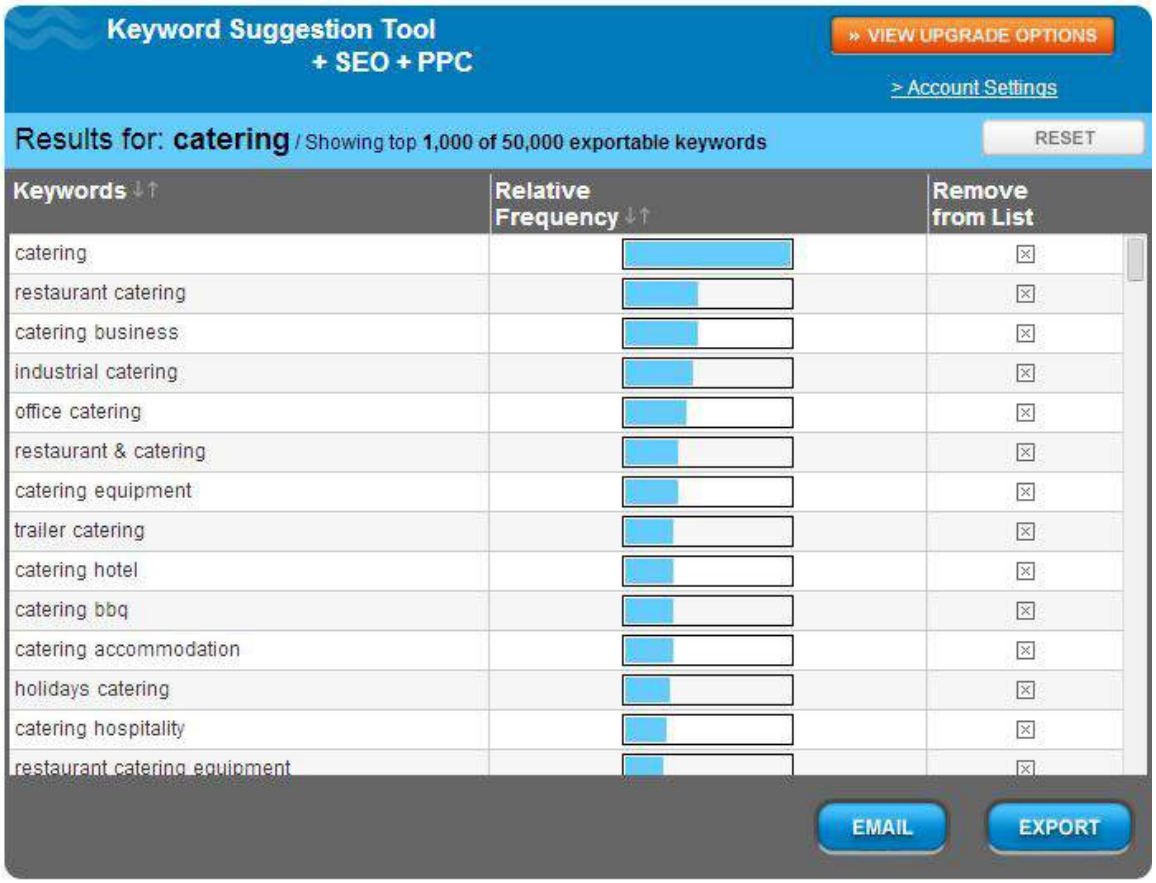


Figure 5-24. WordStream Keyword Suggestion tool output for "catering"

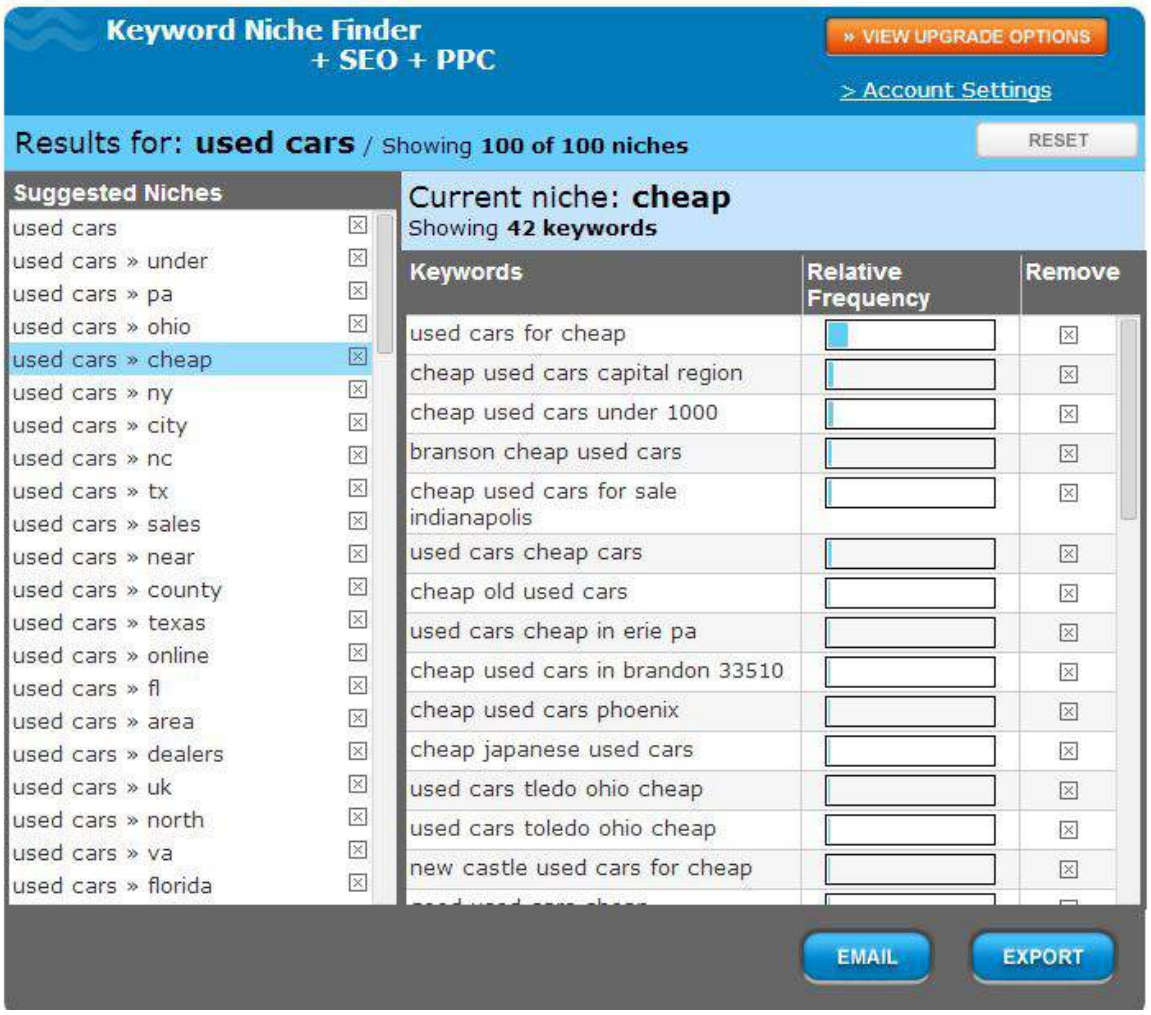


Figure 5-25. Keyword niches related to "used cars" and "cheap"

Where WordStream gets its data

Unlike many keyword research tools, WordStream does not source its data from Google. Instead, the company buys its data from ISPs, browser toolbar providers, and search engines.

How WordStream is useful

WordStream is useful because it pulls its data from sources other than the search engines and goes into further depth, exposing more of the long tail of search. In addition, WordStream offers powerful features for organizing keywords, which in turn makes it easier to organize the keyword data to help drive your SEO strategy.

Cost

WordStream is available in a number of different packages, one of which is a free keyword research tool. The WordStream Keyword Research Suite has three price points as of November 2014: $349 per month for small businesses, $549 per month for the Professional level, and $1099 per month at the Enterprise level.

SEMrush

SEMrush is a data provider and market leader in competitive research software. SEMrush includes the following tools and reports: Organic Research, Advertising Research, AdSense, Keyword Research, Site Audit, Domain vs. Domain, Position Tracking, and Backlinks. To start using SEMrush services, you will need to enter a domain, URL, or keyword in the main search bar. You can switch from one section to another in the left menu to get different data to display (see Figure 5-26).

Figure 5-26. SEMrush domain-level keyword data (analyzing Moz.com)

SEMrush Organic Research report

The Organic Research report offers analytics data on keywords that bring users to a website via Google's top 20 organic results. For every keyword, you can find useful information such as its position in Google search results, volume, CPC, the URL of a landing page and how it's displayed in the search results, competition, and trends (see Figure 5-27).

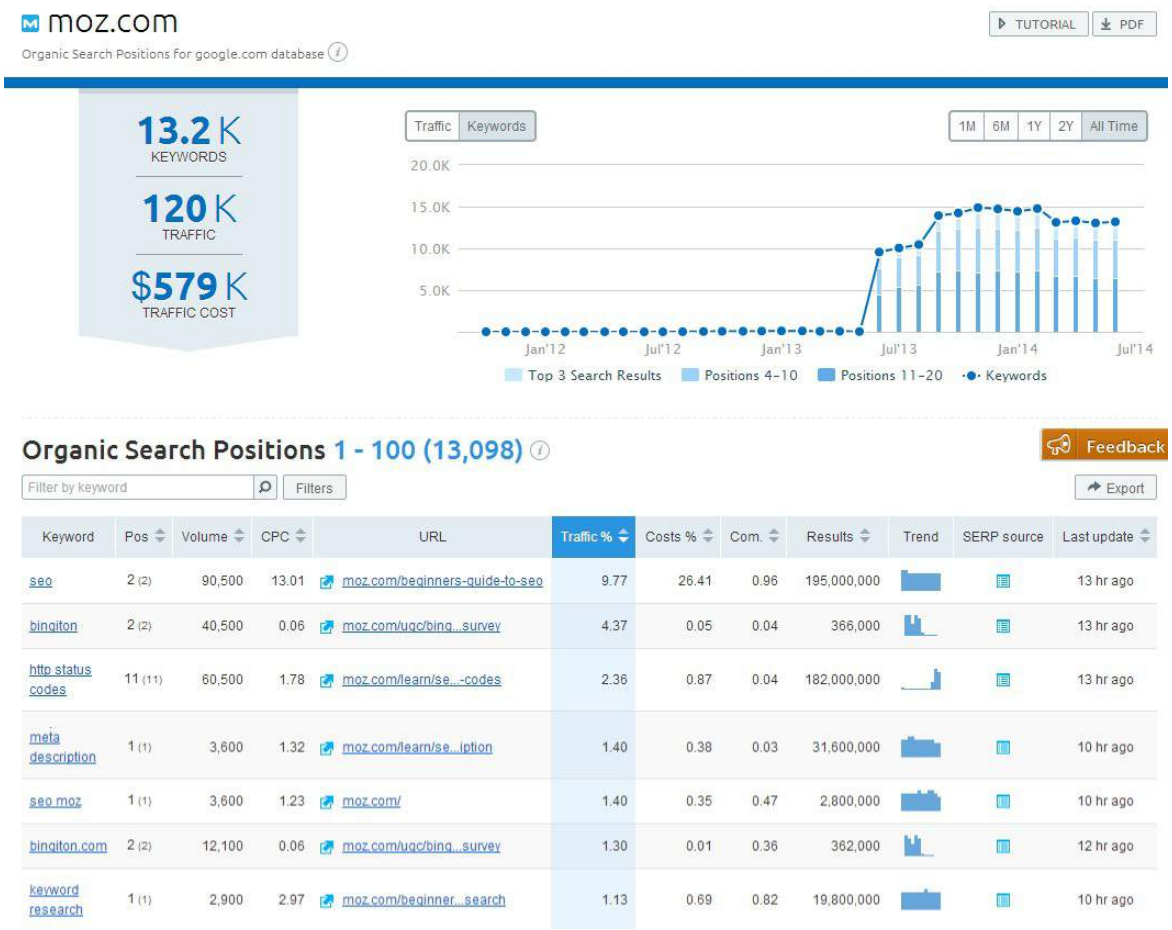


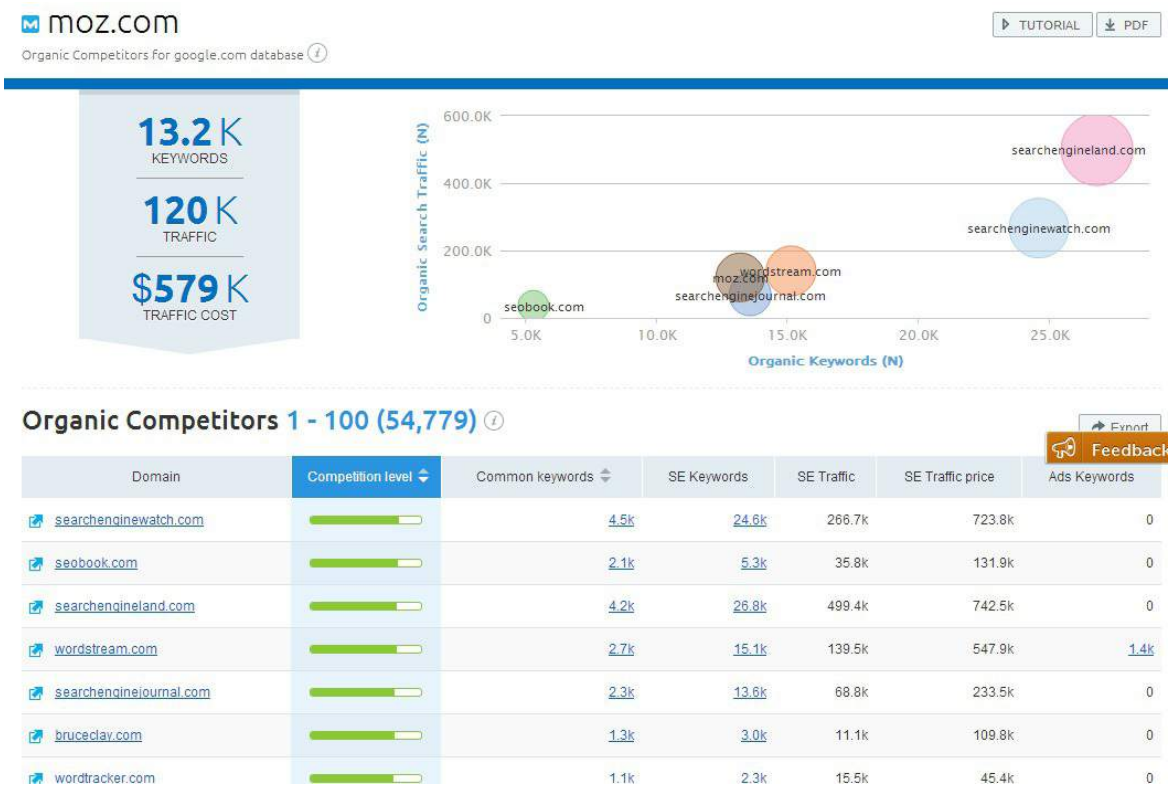
Figure 5-27. SEMrush Organic Research report

The Organic Keywords graph represents the numbers of keywords for which a domain has ranked in the top 20 Google results each month on a historical basis (depending on your subscription, this could be one month, six months, one year, two years, or all time). Keywords are also grouped depending on the position they bring a website (1–5, 6–10, 11+) so you can easily analyze which keywords bring you a higher or lower ranking in Google.

It's also possible to use SEMrush to do keyword research on your competitors. This is particularly useful if you have competitors whom you know are successful in their marketing efforts. By leveraging the keywords your competitors rank for, and an approximation of the traffic they get for those keywords, you can find many keywords that should be profitable for you to rank for as well.

The Competitors section in the Organic Research report provides the list of your

domain's competitors and common keywords (Figure 5-28). The competition level metric helps you exclude from your analysis domains that are not your real competitors on the market. SEMrush analyzes the ratio between the number of domains' common keywords and the total number of keywords of each domain. If the total number of keywords is rather equal and the domains have a high number of common keywords, they will be considered competitors and their competition level will be high.



Domain	Competition level	Common keywords	SE Keywords	SE Traffic	SE Traffic price	Ads Keywords
searchenginewatch.com		4.5k	24.6k	266.7k	723.8k	0
seobook.com		2.1k	5.3k	35.8k	131.9k	0
searchengineland.com		4.2k	26.8k	499.4k	742.5k	0
wordstream.com		2.7k	15.1k	139.5k	547.9k	1.4k
searchenginejournal.com		2.3k	13.6k	68.8k	233.5k	0
bruceclay.com		1.3k	3.0k	11.1k	109.8k	0
wordtracker.com		1.1k	2.3k	15.5k	45.4k	0

Figure 5-28. SEMrush competitors in organic report

SEMrush Advertising Research

The Advertising Research section has the same data as the Organic Research section: Ads Keywords report, Ads Keywords graph, Competitors in Ads report, Position Changes report, and so on. This report can be particularly useful for finding keywords that are good targets for your organic ranking efforts. If you can assume that your competitors are savvy online marketers, then the keywords that they actively purchase can generally be assumed to be those that drive a profitable number of conversions to their websites (see Figure 5-29).

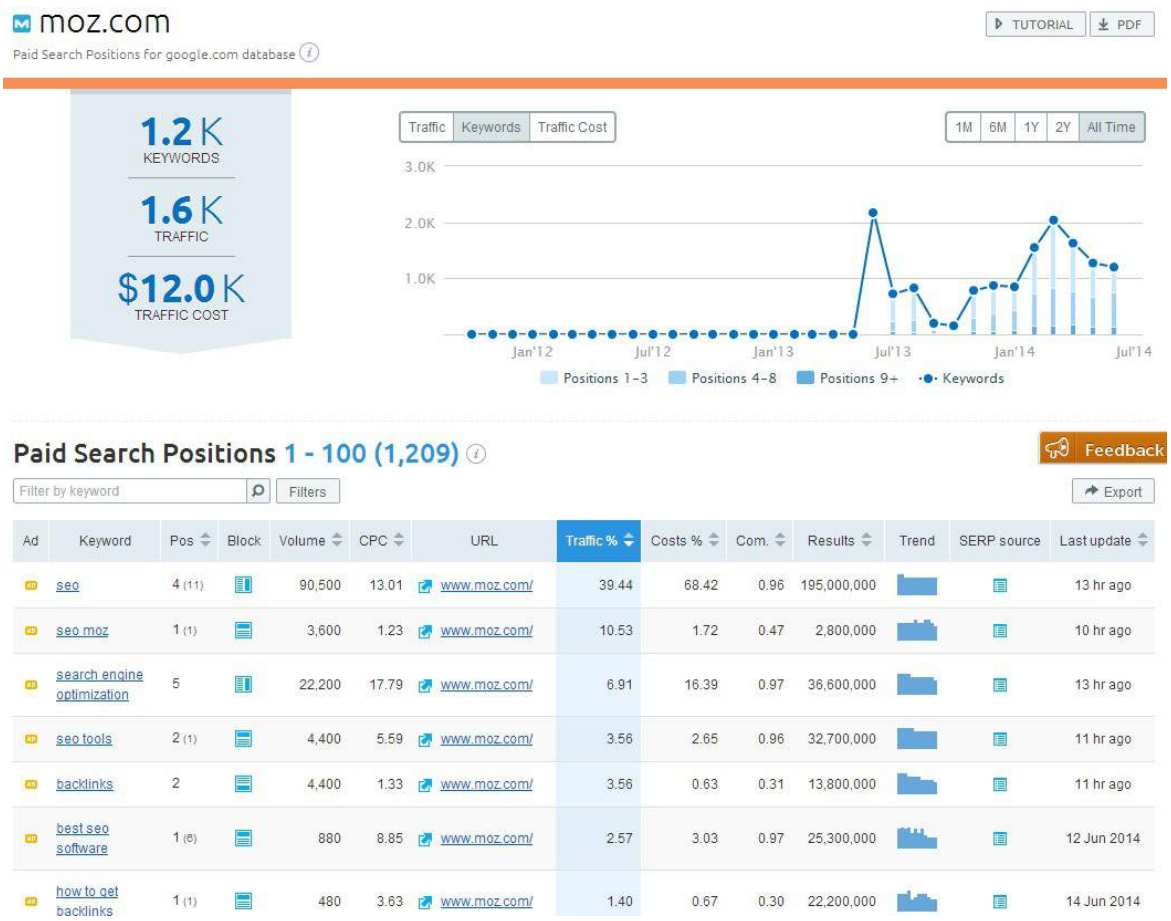


Figure 5-29. SEMrush Paid Search Positions report

Where SEMrush gets its data

SEMrush pulls its data directly from Google search engine results and maintains dedicated databases of this information. SEMrush has 25 databases for the United States (Google and Bing), the United Kingdom, Canada, Russia, Germany, France, Spain, Italy, Israel, Belgium, The Netherlands, Denmark, Finland, Ireland, Norway, Poland, Sweden, Switzerland, Turkey, Brazil, Argentina, Mexico, and Hong Kong.

How SEMrush is useful

SEMrush allows you to access a great deal of data about both your own domain and your competitors' domains and what keywords those domains rank for in both organic and paid search.

Cost

SEMrush offers either a Pro membership for $69.95 per month, which allows for 10,000 results per report and up to 3,000 reports per day, or a Guru-level account at $149.95 per month, which allows for 30,000 results per report and up to 5,000 reports per day.

Searchmetrics

The Searchmetrics suite is an SEO platform that is generally used by enterprises. The software provides a holistic approach regarding both technically based SEO data analysis and content-focused optimization.

While it provides a wide range of SEO metrics and tracking for the purposes of keyword research, you can monitor keywords, rankings, and URLs; optimize on-page structure and content; evaluate links; track social engagement in real time; and compare domains on a project basis. In addition, the platform provides you with automatic technical on-page analyzing and proactive recommendations including site structure optimization for up to several million subpages.

One key feature of Searchmetrics is its reverse-engineering technology, which allows you to reconstruct “not provided” keyword and traffic data on a URL level by combining various data sources and algorithmic calculations.

For the purposes of straightforward keyword research, Searchmetrics offers a keyword discovery tool that ranks any keyword on search volume, competition, cost (if you were to pay for the keyword rather than ranking for organically), and “search volume spread,” which shows any seasonality in the keyword. It shows similar and related keywords and allows you to click through on any keyword to see the top sites ranking for that term (see Figure 5-30).

Also of interest with Searchmetrics is that it offers competitive keyword research on both paid and organic traffic. Enter a domain in the search box, and Searchmetrics will return a list of competitors in either organic or paid search along with data on how many keywords the domain shares in common with its competitors, how many keywords are not shared with the competitors, relative traffic volume for the competitors, and other data on the competitors’ rankings (see Figure 5-31).

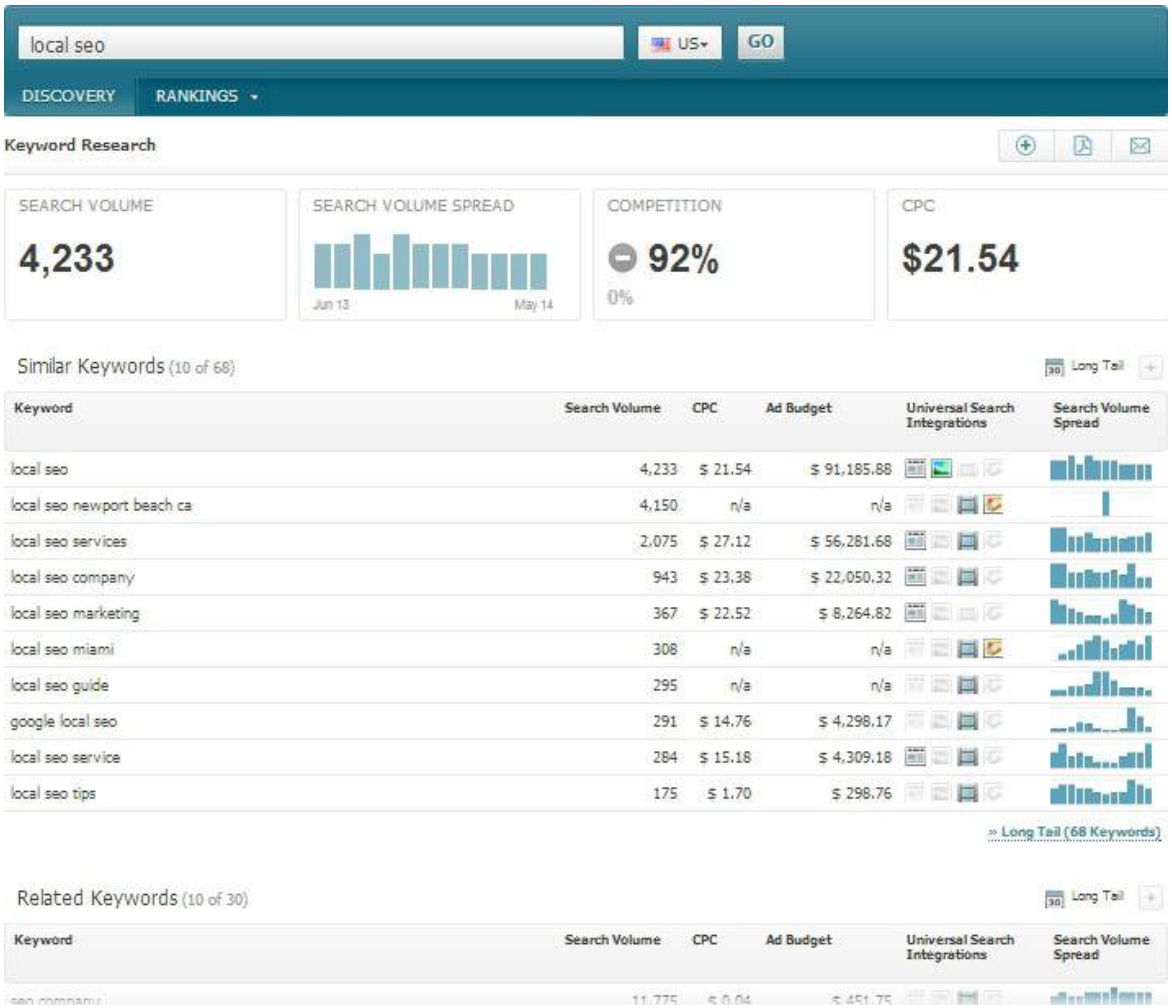


Figure 5-30. Searchmetrics Keyword Discovery report

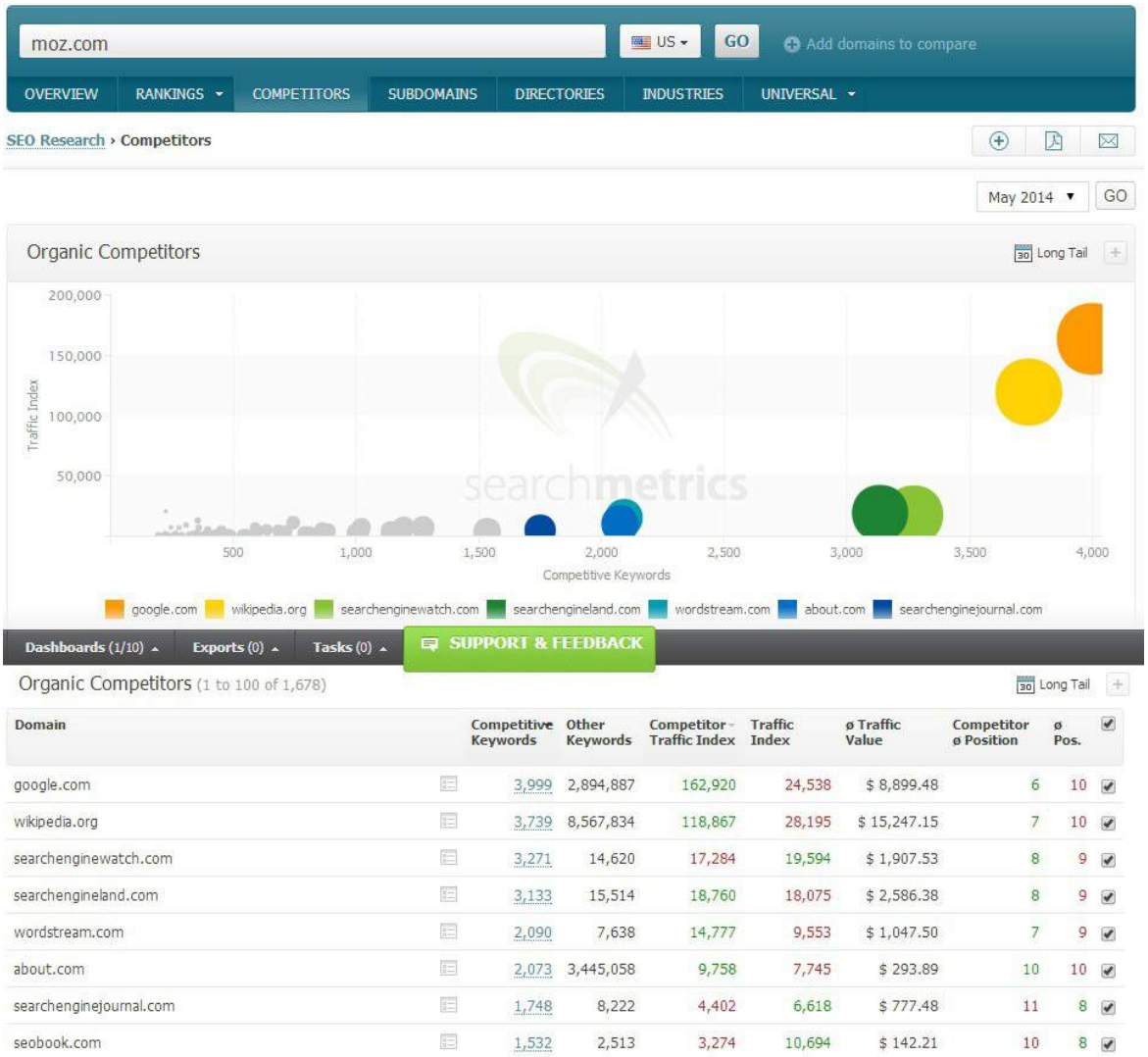


Figure 5-31. Searchmetrics Organic Competitors report

By clicking through any entry in the Competitive Keywords column, you can delve more deeply into which keywords the sites might have in common. To get more details on exactly which keywords your competition is ranking for, you can also look at the Rankings report under SEO Research (Figure 5-32). This report gives you in-depth information into thousands of keywords that your competitor ranks for, including the keyword itself, the URL that ranks for it, the competitor's position, and an estimate of both search volume and what the term would cost to purchase via PPC advertising. This report can give you a great deal of insight into which keywords are important to your competitors, and thus in many cases which keywords should be important to you.

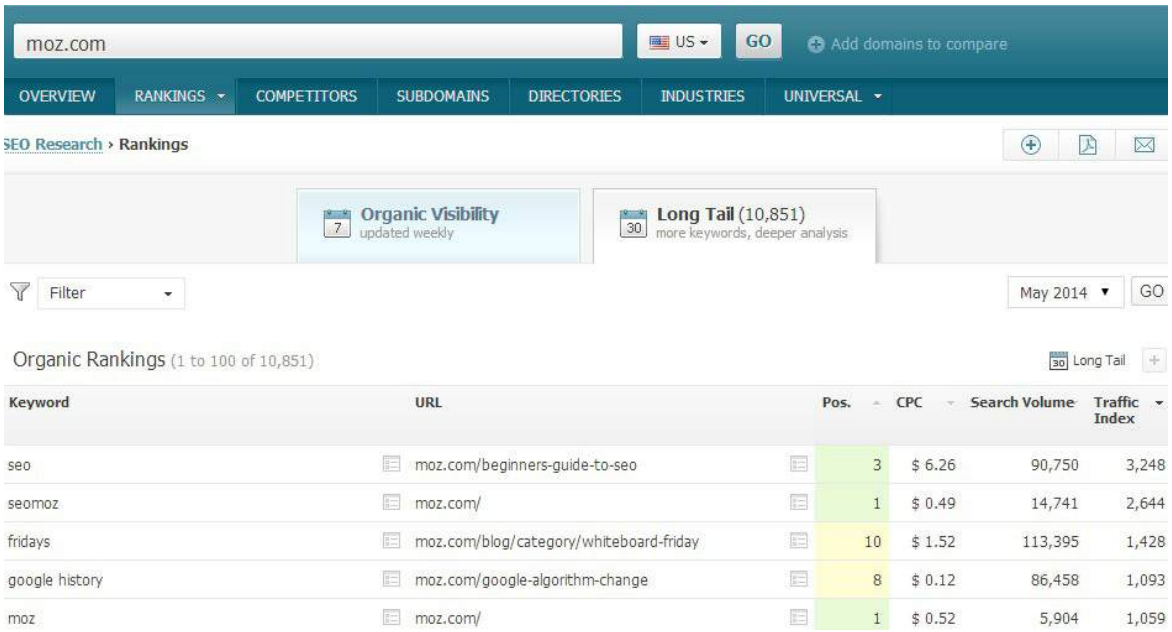


Figure 5-32. Searchmetrics Rankings report

Where Searchmetrics gets its data

Searchmetrics gets its data from the Google SERPs and Google Keyword Planner.

How Searchmetrics is useful

Searchmetrics allows you to track most of your SEO efforts while providing many on-page analyses and recommendations. In addition, it provides a wealth of keyword information both at the keyword level and in mining your competitors' keyword data.

Cost

Searchmetrics Essentials, which allows for research access to SEO, PPC, keyword, and other data and detailed competitor analyses, costs $69 per month. The full Searchmetrics suite—which includes tracking, research, and optimization solutions—is quoted on a case-by-case basis.

Other tools of interest

There are many other keyword tools available on the market. Some of the more interesting ones are described next.

Google Autocomplete

Previously known as Google Suggest, this Google feature automatically "suggests" queries related to a word or starting with the same word. For example, typing in *lilac* will reveal suggestions of *lilac essential oil*, *lilac wine*, and *lilac tree*. This was a somewhat hidden tool that, thankfully, Google brought front and center to default Google Search. Google won't tell you how many times *digital cameras* has been searched for, but because it appears at the top of the list of suggestions, you can infer that it was probably searched for more often than the phrases that appear below it. This can give you important insight into what searchers are looking for, or what they search for in relation to a specific topic. For example, does a company have an autocomplete suggestion of *<company name> scam*? If so, it might have a PR nightmare on its hands.

NOTE

Suggestions are personalized based on the user's location (e.g., *wet n wild phoenix* is a suggestion that shows up when a user in Phoenix, Arizona, types the letter *w* into Google. And in mobile search, you may notice apps listed as a suggestion.

Google Instant

Start with the basic search input box, but look for what Google reveals as you type. Google gives results that appear in the search page even before you have completed the search query. As you refine this search further, these results appear below and change to fit your refined search. This benefits the searcher by eliminating the need to press the Search button, and often limits the need to finish the query. Google estimates that this function has the capability to save two to five seconds per search.

Soovle

Soovle shows you real-time search terms as you type them, ordered by popularity, just like Google Suggest. In fact, it's a one-stop shop that taps into those features of top search engines and much more. It also polls YouTube, Amazon, Bing, Wikipedia, and Answers.com for top related search terms, refreshing dynamically each time you pause during your typing. This tool allows you to tap into seven top resources at once (see Figure 5-33).

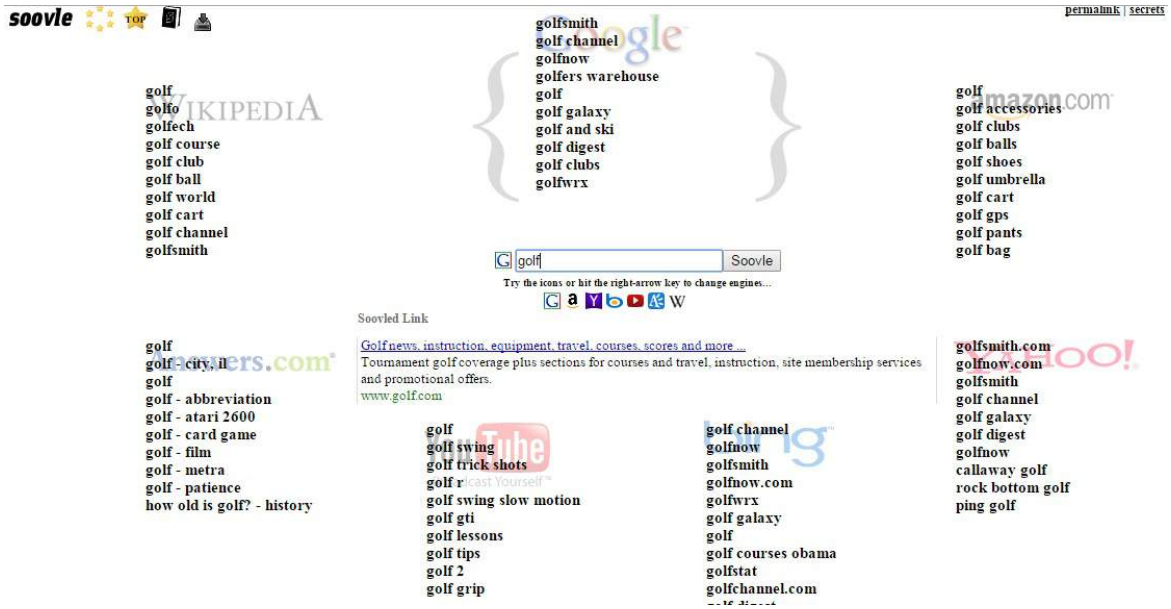


Figure 5-33. Soovle results using keyword "golf"

YouTube Autocomplete

If you begin to type in a search query on YouTube, it offers search suggestions, as shown in Figure 5-34. The suggestions are the most popular variants of the search query you have typed so far.

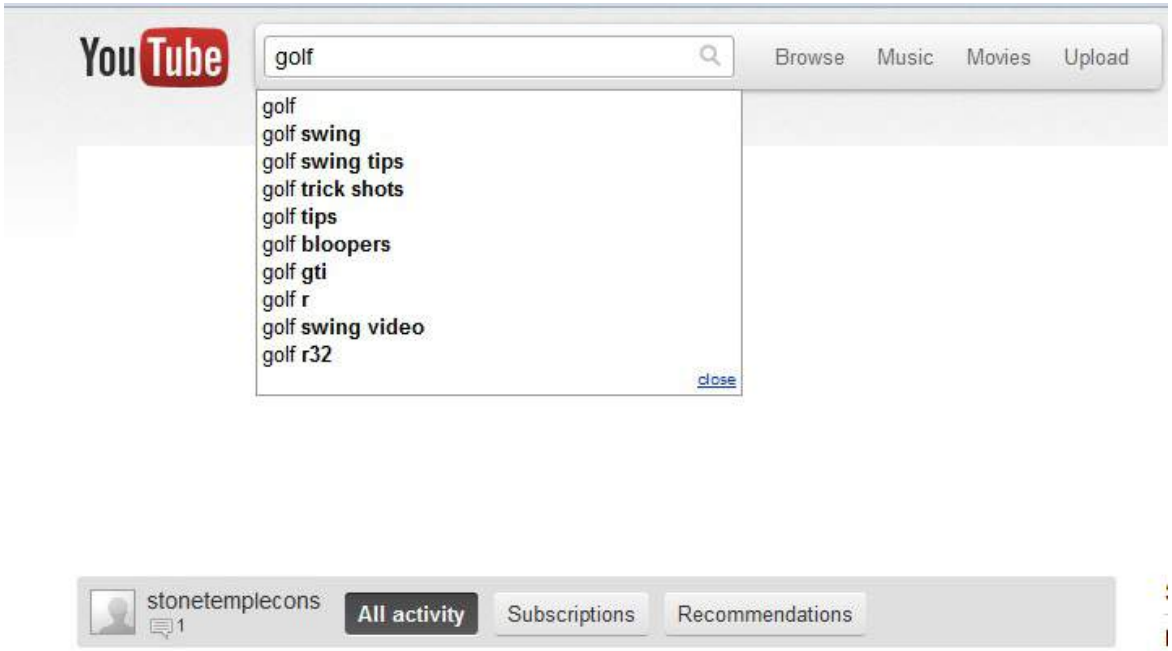


Figure 5-34. YouTube Autocomplete

Google AdWords Display Planner

On September 1, 2014, Google discontinued the former YouTube Keyword Tool. As with the Google Keyword Planner, you must have a Google AdWords account, though you do not need to spend any money on the account to access the Display Planner (shown in Figure 5-35). You can enter a landing page or select from a list of topics to get recommendations on what terms might drive the most impressions. To get the video-only keyword recommendations, first search for your desired topic, then select Video under the "Ad formats and sizes" option on the lefthand side of the page. Selecting "Individual targeting ideas" will return a list of recommended keywords, their relevance to your select topic, and historical CPM (cost per mille, or cost per thousand impressions), number of cookies per week, and number of impressions per week.

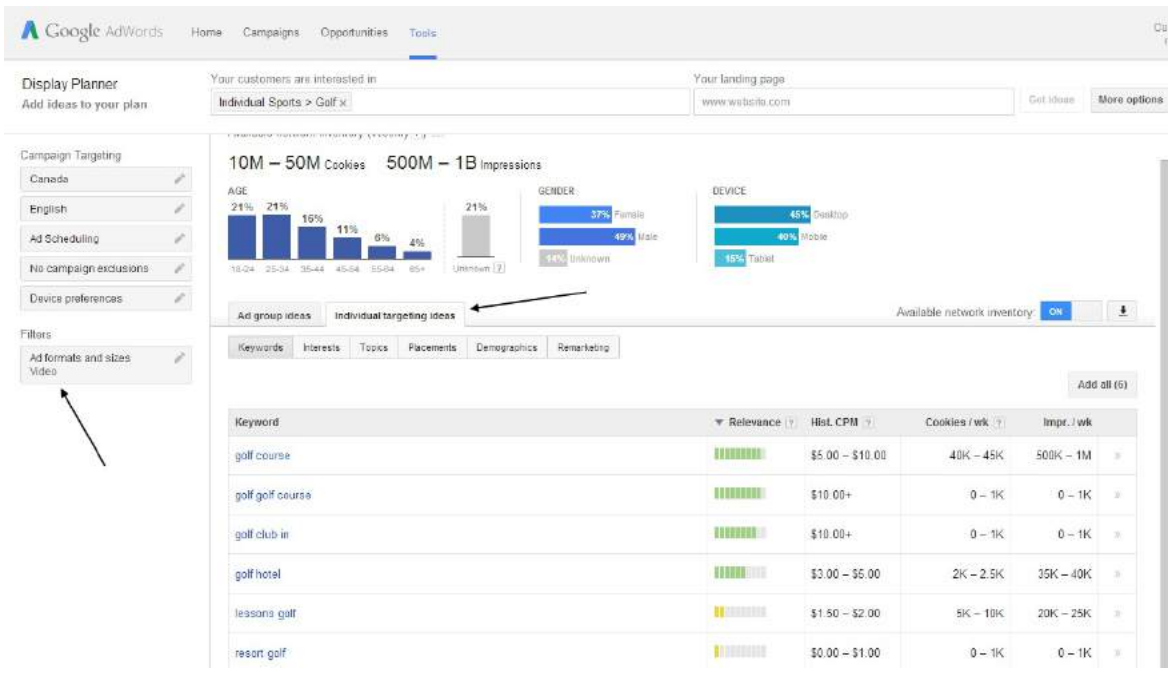


Figure 5-35. Google AdWords Display Planner

Google Trends for Video

Just as you can use Google Trends to search for the relative interest of various keyword terms in the main Google search results, you can also use it to look for keyword trends and suggestions for YouTube. To do this, search as usual for a term on Google Trends, then select the YouTube option from the top navigation

bar (see Figure 5-36).

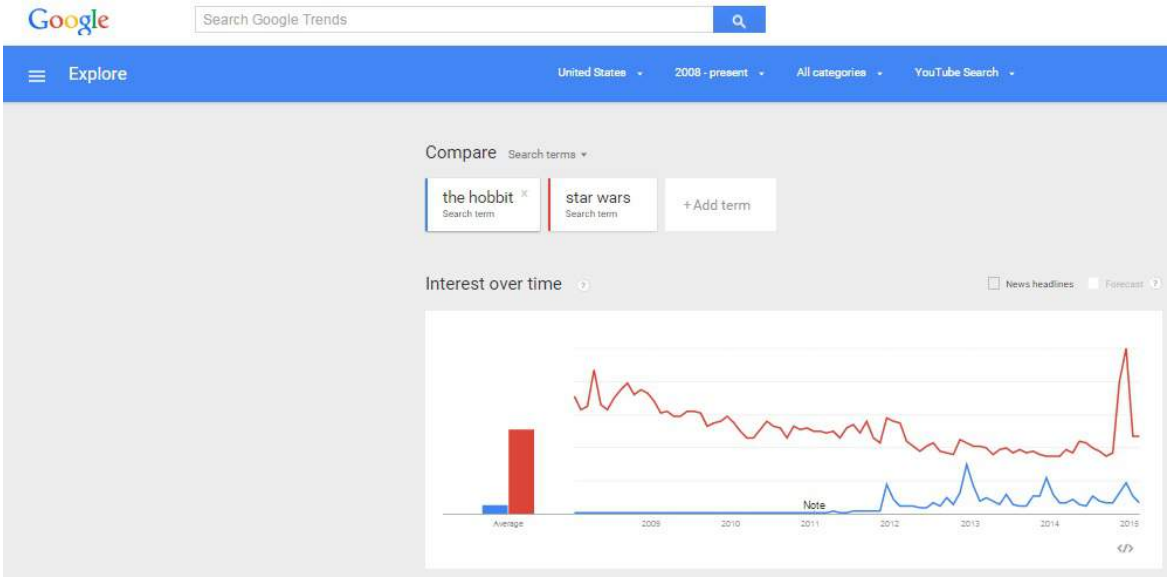


Figure 5-36. Google Trends for YouTube search

Google Trends will also return a list of topics and keyword queries related to your search, with a rating for the relative interest of each (see Figure 5-37).

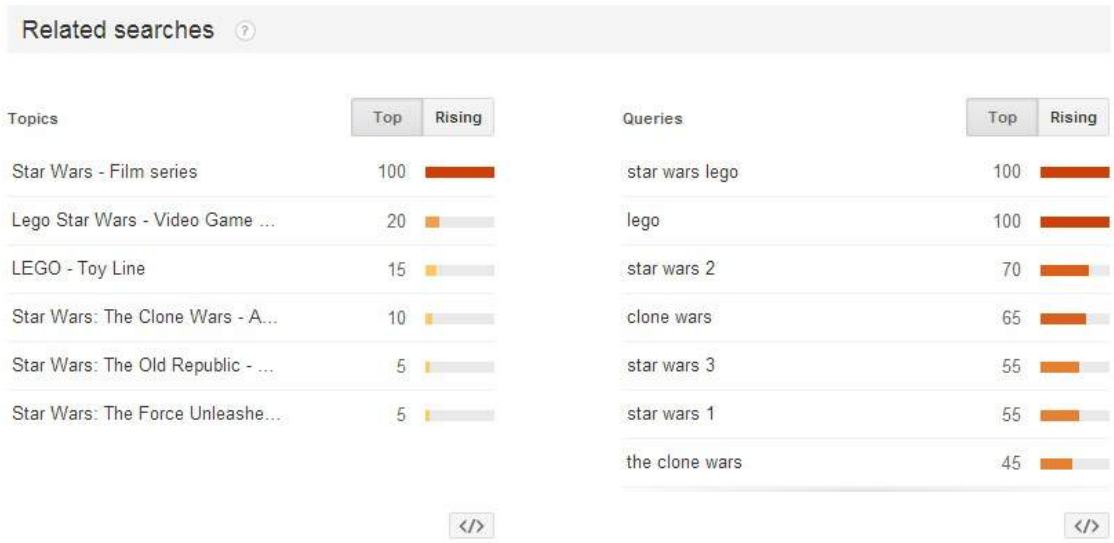


Figure 5-37. Google Trends related searches

Ubersuggest

Ubersuggest is a free tool based on Google Autocomplete. It runs a bunch of variants based on the term that you have entered. For example, if you enter the query *golf*, Ubersuggest will automatically pull the suggestions for *golf + terms beginning with a*, *golf + terms beginning with b*, and so on, all the way through to *golf + z* and then *golf + 0*, *golf + 1*, and so on, as shown in Figure 5-38.

golf

English/USA | Web

Suggest

Select all keywords

280 suggestions found.

original _ a b c d e f g h i j k l m n o p q r s t u v w x y z

golf + a

- golf austin
- golf apparel
- golf accessories
- golf and guitar
- golf austin texas
- golf academy of america
- golf apps
- golf art
- golf and ski warehouse
- golf austin style

golf + b

- golf buddy
- golf bags
- golf balls
- golf balls galore
- golf ball retriever
- golf bonita springs
- golf buddy voice
- golf ball markers
- golf ball reviews
- golf belts

Figure 5-38. Ubersuggest suggestions

KeywordSpy

KeywordSpy is a paid tool that ranges in price from $89.95 to $139.95 monthly for basic packages. It also has agency and enterprise packages available at higher rates. KeywordSpy offers a variety of services but primarily offers insights into paid search volumes and costs per click, as shown in Figure 5-39. This can help you identify keyword variations as well as which keywords are more competitive and thus more likely to lead to higher ROI should you be able to drive traffic from them. The tool also gives insight into search volumes for related terms, similar terms, and misspellings and shows lists of the top-ranking PPC and organic competitors for a given query.

Figure 5-39. KeywordSpy keyword report

Visual Thesaurus

Visual Thesaurus is another graphical tool that shows the definition of the keyword entered as well as related terms and concepts. Clicking on each node in the diagram opens a map of variations on the keyword and related concepts for that node (see Figure 5-40). It's another tool you can use for brainstorming new keywords and concepts, which you can then enter into the keyword tools that give more specific keyword data. It has a nominal cost of $2.95 per month or $19.95 per year.

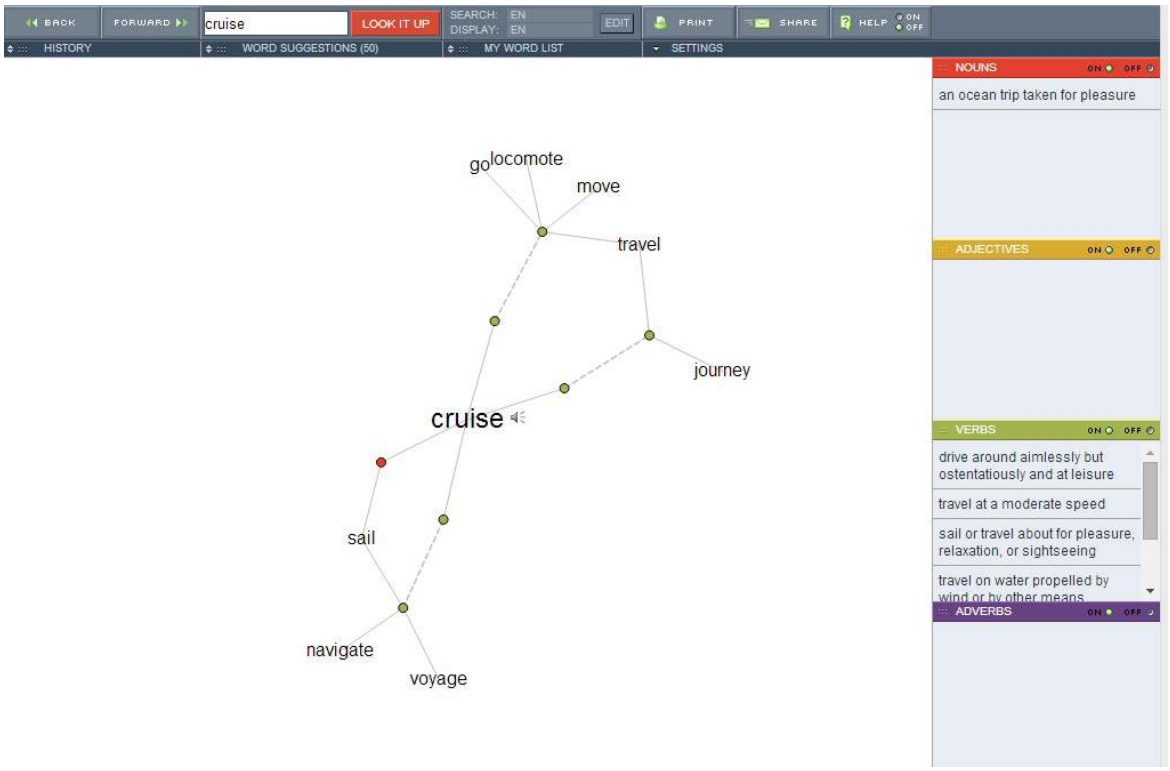


Figure 5-40. Visual Thesaurus related concepts graph

Visuwords

Visuwords is a free tool that provides a dynamic graphical interface of words and concepts related to the search term you enter (see Figure 5-41). It acts as both a dictionary and thesaurus as well as providing links to related concepts. These related concepts may be keyword areas not previously considered in other keyword research. Visuwords is a graphical representation of the data in Princeton's WordNet Search.

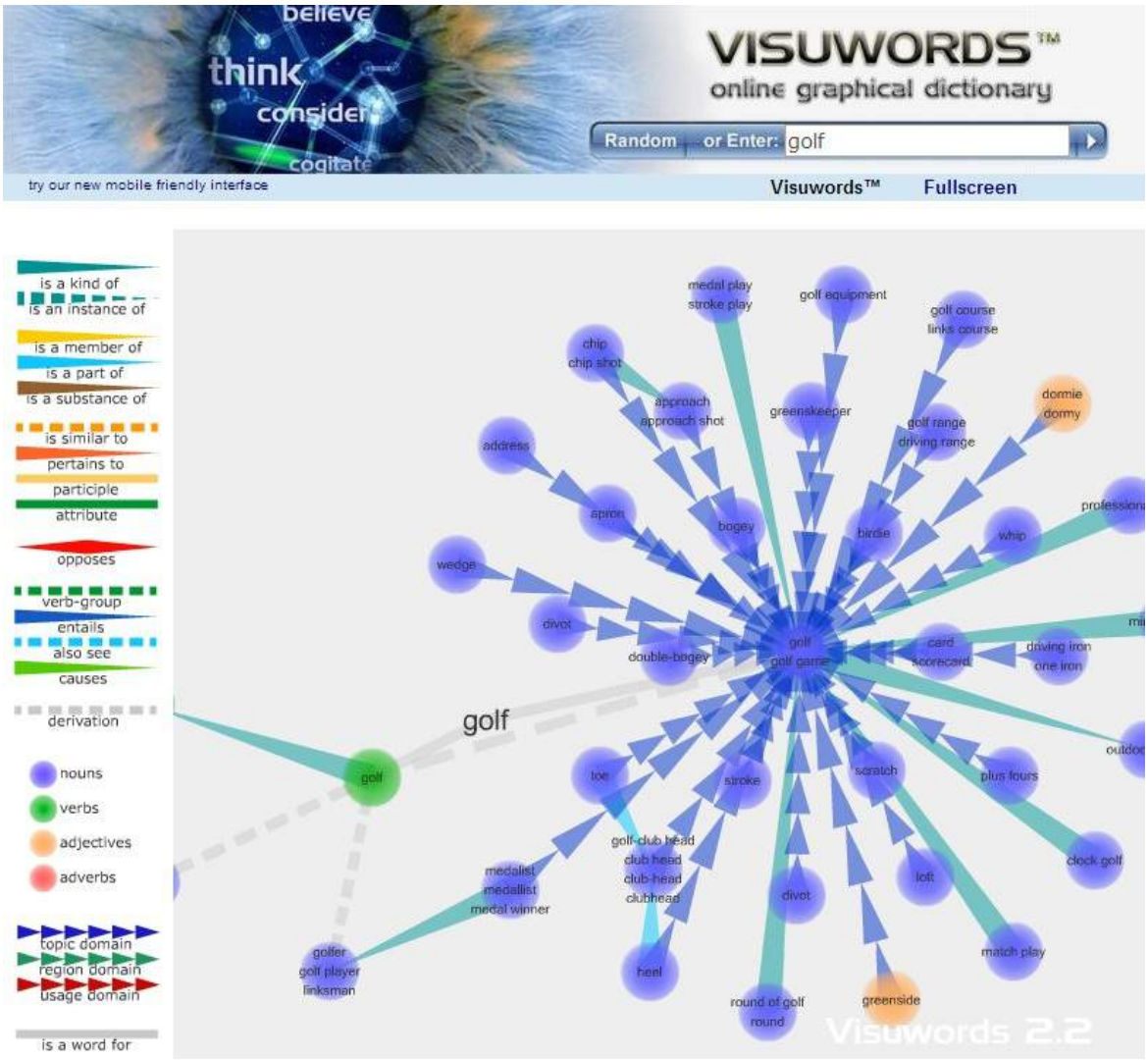


Figure 5-41. Visuwords related concepts graph

Keyword Research Data Analysis

Once you have the raw keyword data from the research you've done with your favorite tools, you need to analyze which keywords have the highest value and the highest ROI. Unfortunately, there is no simple way to do this, but we will review some of the options in this section.

When you're researching keywords for your site, it is important to judge each keyword's value, relevance, and potential conversion rate. If a keyword is strong in all three criteria, it is almost certainly a keyword you want to plan to optimize for within your site.

Determining keyword value

When judging the value of a keyword, you should contemplate how useful the term is for your site. How will your site benefit from targeting these keywords?

Identifying relevant keywords

To identify relevant, high-quality keywords, ask yourself the following questions:

How relevant is the term/phrase to the content, services, products, or information on your site?

> Terms that are highly relevant will convert better than terms that are ancillary to your content's focus.

Assuming a visitor who searches for that term clicks on your result in the SERPs, what is the likelihood that she will take action in a way that will positively influence your site?

> You need to assess how likely it will be for users to perform a desired action on your site (make a purchase, subscribe to a newsletter, etc.), create a link to your site, influence others to visit, or engage in off-site engagement by promoting your content in social media. (You can also test this by setting up a PPC campaign, and buying clicks on a given keyword and seeing how it converts for you.) It is a good idea to target keywords that indicate imminent action (e.g., *buy cranium board game*, *best prices for honda civic*), because searchers are more likely to perform the corresponding action on your site when they search for those terms than they are for terms such as *honda civic* or *cranium board game*. Your click-through/conversion rates are likely to be higher if you target keywords that indicate the intent behind the search.

How many people who search for this term will come to your site and leave dissatisfied?

> Pay attention to your site's content and compare it to what other sites in the top results are offering—are these sites doing or offering something that you haven't thought of? Do you feel as though these sites offer a more positive user experience? If so, see what you can learn from these sites and possibly emulate on your own.

> You can also use an analytics program and check to see which of your pages have the highest abandonment rates. See what you can change on those pages to improve user experience and increase users' level of enjoyment when they're using your site.

It is important to categorize your keywords into high and low relevance. Generally, keywords of higher relevance will be more beneficial to your site in that they best represent your site as a whole. If, when judging the relevance of a keyword, you answer the preceding questions positively, you've found a highly relevant term and should include it in your targeting.

Keywords with lower relevance than those that lead to conversions can still be great terms to target. A keyword might be relevant to your site's content but have a low relevance to your business model. In this case, if you target that keyword, when a user clicks on your site and finds the content to be valuable, he is more likely to return to the site, remember your brand, and potentially link to your site or suggest it to a friend. Low-relevance keywords, therefore, present a great opportunity to strengthen the branding of your site. This type of brand value can lead to return visits by those users when they are more likely to convert.

Determining conversion rates

A common misconception is that a conversion refers only to the purchase of an item on your site. However, many different types of actions users perform can be defined as a conversion, and they are worth tracking and segmenting (you can read more about this in "Key Performance Indicators for LongTail SEO").

The many different types of conversions create distinct opportunities for targeting various keywords. Although one keyword may work well for purchase conversions, another may be well suited to get users to subscribe to something on your site. Regardless of what type of conversion you are optimizing for, you should strive to have each keyword that you intentionally target convert well, meaning it should be relatively successful at getting searchers to click through to your site and, consequently, perform a specific action.

To know which keywords to target now (and which to pursue later), you must understand the demand for a given term or phrase, as well as the work required to achieve those rankings. If your competitors block the top 10 results and you're just starting out on the Web, the uphill battle for rankings can take months, or

even years, of effort, bearing little to no fruit. This is why it is essential to understand keyword competitiveness, or keyword difficulty.

To give you a rough idea of the level of competition faced for a particular term or phrase, the following metrics are valuable:

- Search demand volume (how many people are searching for this keyword)
- Number of paid search competitors and bid prices to get in the top four positions
- Strength (age, link power, targeting, and relevance) of the top 10 results
- Number of search results—it can be valuable to use advanced operators such as exact search (e.g., *"patrick powers" Ireland*) or the *allintitle* and *allinurl* operators here as well; see coauthor Stephan Spencer's book, *Google Power Search* (O'Reilly) for more on using these specialized searches for keyword research

Moz offers a Keyword Difficulty tool that does a good job collecting all of these metrics and providing a comparative score for any given search term or phrase.

Ad Campaign Runs and Third-Party Search Data

One of the points we have emphasized in this chapter is the imprecise nature of the data that keyword tools provide. This is inherent in the fact that the data sources each tool uses are limited. It turns out that there is a way to get much more precise and accurate data: making use of Google AdWords.

Start with the keywords you are interested in and implement a simple AdWords campaign. Assuming that you are implementing this campaign solely to get keyword volume data, target position #4 or #5. This should be high enough that your ads run all the time, but low enough that the cost of collecting this data won't be too high.

Once you have run this for a few days, take a look at your AdWords reports, and identify the number of impressions generated for the keyword. Although this data is straight from Google, it is important to remember that the advertisers' ads may not be running all the time, so more (possibly significantly more)

impressions may be available.

Next, think about the value of achieving certain rankings in the organic results. You can come up with a good estimate of that as well. The key here is to leverage what you know about how click-through rates vary based on organic search position. Table 5-2 depicts click-through rates by SERP position based on data by Digital Relevance.[3]

Table 5-2. Click-through rates by SERP position

Organic position	Click-through rate
1	18.2%
2	10.05%
3	7.22%
4	4.81%
5	3.09%

This data, of course, is aggregated across a very large number of searches on AOL, so it serves only as an estimate; but if you are in position #1, the estimate is that 18.2% of the people who will search on a term will click on your result. In the case of a term that is searched 52 times per day, the site in the #1 position will get 9–10 clicks per day.

NOTE

There are certain search terms to which these estimates do not apply. For example, if the user searches on a brand term, the focus on the #1 position is much, much higher. Publishers in lower positions still get some traffic, but at lower percentages than we've outlined here.

So, now you have a working estimate of the search volume and the number of clicks per day that the term will deliver. Can you get an estimate of conversion rates as well? Yes, you can. This requires only a simple extension of the AdWords campaign: implement conversion tracking, either with the free capability provided by Google or via another method at your disposal.

Once you have that in place, look at how your AdWords campaign performs. If

the conversion rate is a lofty 5% for one keyword and 3% for another, chances are that your organic search conversion rates for those two keywords will vary by a similar amount. Be aware, though, that although paid search results get significantly less traffic than organic search results, paid click-throughs do tend to convert at a somewhat higher rate (1.25 to 1.5 times, according to industry data). Using the preceding example, this suggests that we will get a little less than one conversion per day as a result of being in the #1 position.

This data is great to have in hand. However, it does not mean you should use this methodology instead of other keyword tools. It takes time to implement and to collect the data. The keyword tools will get you started in real time. Nonetheless, using AdWords or MSN adCenter can provide you with some great data.

We also don't recommend that you get obsessed with tracking your rankings on keywords. As we discussed in "Benchmarking Current Rankings" in "Benchmarking Organic Rankings", it is not possible to do this as accurately as you might think, and it can lead to poor decision making. Nonetheless, using this type of AdWords testing can help you get a sense of real search volumes and the importance of particular keywords to your SEO campaign.

Landing Page Optimization

Landing page optimization (sometimes also called *conversion optimization*) is the practice of actively testing multiple variations of a web page (or website) to see which one performs the best. Typically, this is done as part of an effort to improve the conversion performance of the site and it relies on testing tools such as Visual Website Optimizer, Optimizely, or Content Experiments (within Google Analytics).

The simplest form of this type of test is called an *A/B test*. A/B tests involve creating two different versions of a page, and then selecting a version at random to show to a new visitor to the site (old visitors get the version they saw the last time they visited). You then measure the behavior of the visitors to the two different versions to see which group of visitors completes conversions on the site. You have to be careful to wait until you have a statistically significant amount of data to draw a conclusion. Once you have this data you can analyze it and decide on more tests, or simply pick the winner and move on.

Multivariate testing is a bit more complex, because it involves more than two

variations in the test. In addition, you can mix and match multiple variations. For example, you may want to try two different logos, two different calls to action, three different page titles, two different color schemes, and so on. In multivariate testing, any combination of your elements could be what is shown to a particular visitor. Obviously, more data (visits and actions) is required to draw a conclusion than in a simple A/B test.

Landing page optimization can help in determining the value of a keyword because one of the elements you might want to test is the impact on conversion of variations of a keyword or phrase in the page title, the page header, and other strategic places on the page. One variation of the test would use one keyword or phrase and the other variation would use a different one.

You can then see which keyword provides the best results from a conversion perspective. This data can provide you with an interesting measure of keyword value: its ability to help you convert your visitors. However, landing page optimization is not practical to use to perform SEO tests (i.e., to see which version of a page ranks higher), as SEO tests can take weeks or even months to see results.

Leveraging the Long Tail of Keyword Demand

As we discussed at the beginning of this chapter, the long tail of search is where 70% of search queries occur. Only 30% of those precious queries happen in the more obvious terms that people use, the so-called head terms. Another way to underscore this is that in May 2007, Google Vice President Udi Manber indicated that 25% of all search queries that Google receives on a given day are queries that Google is seeing for the first time. You can think of this as the "ultra-long tail."

The long tail of search queries in a given industry is typically not visible via any of the major keyword research services or search engine ad databases (Google AdWords, Bing Ads). In these instances, there is a method to find those terms that can carry value, but it requires a good amount of research and analysis.

With this in mind, let's outline a few methods for finding longtail terms.

Extracting Terms from Relevant Web Pages

One source for longtail terms is web pages that do well for searches that are relevant to your target market. Here is a basic process for finding those pages and extracting that information from them:

1. Extract the top 10–50 most common search phrases at the head of the distribution graph from your existing keyword research in the industry.

2. Search Google and Bing for each term.

3. For each page in the top 10–30 results, extract the unique text on the page.

4. Remove stopwords and filter by phrase size.

5. Remove instances of terms/phrases already in your keyword research database.

6. Sort through the most common remnants first, and comb as far down as you feel is valuable.

Through this process, you are basically text-mining relevant documents on the subject of your industry/service/product for terms that, although lower in search volume, have a reasonable degree of relation. When you're using this process, it is imperative to have human eyes reviewing the extracted data to make sure it passes the "common sense" test. You may even find additional terms at the head of the keyword distribution graph.

You can expand on this method in the following ways:

- Text-mine search results on social media sites such as Twitter and Pinterest for relevant results.

- Use documents purely from specific types of results—local, academic—to focus your keyword mining efforts.

- Mine forum threads on your subject matter. You could even use *inurl:forum* in the searches to grab conversational keywords.

This methodology is highly effective. The return on this research has a direct relationship to the amount of effort you expend (and how deep you dig).

Mining Keyword Research Tools

Although looking into keyword research tools for longtail data has significant limitations, there are still ways to do it. For example, if you own a chain of pizza restaurants in 50 cities across the country and you want to discover longtail terms, you can. Let's look at the tail end of Google Keyword Planner's output for a combined search on *Orlando Pizza*, *San Diego Pizza*, and *San Jose Pizza* (see Figure 5-42).

Your product or service

Orlando Pizza, San Diego Pizza, San Jose Pizza | Get ideas | Modify search

Ad group ideas | Keyword ideas | Download | Add all (803)

Keyword (by relevance)	Avg. monthly searches	Competition	Suggested bid	Ad impr. share	Add to plan
best italian in little italy san diego	40	Low	$0.46	0%	»
pieology pizza san jose	40	Low	-	0%	»
24 hour delivery san jose	40	High	$1.83	0%	»
san diego pizza restaurants	40	Medium	$0.73	0%	»
dessert downtown san diego	40	Low	-	0%	»
amato pizza san jose	40	Low	-	0%	»
amore pizza san diego	40	Low	-	0%	»
amici pizza san jose	40	Low	-	0%	»
san diego real estate mls	40	High	$2.80	0%	»
best restaurants san diego ca	40	Medium	$2.50	0%	»
pizza coupons san diego	40	High	$1.95	0%	»

Figure 5-42. Extracting longtail data from Google AdWords Keyword Planner

san diego pizza restaurants (40 searches per month on average) is an example of a valid longtail term. If some people search for *san diego pizza restaurants*, it is quite likely that others search for *orlando pizza restaurants*. It does not show in this data, because the volume of queries available to the keyword research tool is limited. All we are doing with these combined searches is giving the search tools more data to work with.

The takeaway remains valid: apply these logical longtail extensions across all of your cities, even though the keyword tool shows it for only one, and you're likely to attract search queries for those keywords.

Identifying LongTail Patterns

You can also take another stab at determining longtail information. As a hypothetical example using *digital camera*, here are 40 searches for two different brands and models of digital cameras that have been pulled (for this demonstration) from the KeywordDiscovery database that received only one search:

- *consumer comments on nikon 5.1 mp coolpix l3 digital camera*
- *new nikon coolpix p3 8 1 mp digital camera memory*
- *nikon 3 2 mp coolpix digital camera*
- *nikon 51 mp coolpix s1 digital camera and cradle*
- *nikon 6 mp coolpix digital camera*
- *nikon 7 1 mp coolpix 7900 digital camera*
- *nikon 81 mp coolpix 8800 digital camera*
- *nikon coolpix 4800 4 mp digital camera*
- *nikon coolpix 5200 51 mp digital camera*
- *nikon coolpix 5400 51 mp digital camera*
- *nikon coolpix 6.0 mp digital camera*
- *nikon coolpix 8700 8mp 8x zoom digital camera 8 mp*
- *nikon coolpix l2 6.0 mp digital camera*
- *nikon coolpix l3 6 mp digital camera usa warranty*
- *nikon coolpix p2 51 mp digital camera*
- *best buy sony cybershot dsc t7 51 mp digital camera*

- *brand new sony cybershot dsc h1 51 mp digital camera*
- *camera digital sony cybershot 51 mp*
- *sony - cybershot 10.1 mp digital camera*
- *sony - cybershot 6.0 mp digital camera*
- *sony 5 mp cybershot dsc t9 digital camera*
- *sony 72 mp cybershot dsc p200 digital camera information*
- *sony 72 mp cybershot dsc w7 digital camera*
- *sony 72 mp digital still camera cybershot rebate*
- *sony cybershot 10.1 mp digital camera*
- *sony cybershot 7 2mp digital camera 7 2 mp*
- *sony cybershot 72mp dsc w7 digital camera 72 mp*
- *sony cybershot 81 mp digital camera*
- *sony cybershot digital camera 5.1 mp*
- *sony cybershot digital camera 6 mp*
- *sony cybershot dsc 1 81 mp digital camera review*
- *sony cybershot dsc h1 51 mp digital camera*
- *sony cybershot dsc w30 6 mp digital camera*
- *sony cybershot dscs40 41 mp digital camera 3x opt zoom*
- *sony dsc p73 cybershot digital camera 41 mp p 73*
- *sony dsc p8 cybershot 32 mp digital camera*

- *sony dsc s60 cybershot digital camera 4 1 mp*
- *sony dsc s85 cybershot 41 mp digital still camera*
- *sony dsc t1 cybershot digital camera 5 0 mp*
- *sony dsc t1 cybershot digital camera 50 mp t 1*

Our goal is to determine whether there are any universal patterns that searchers use. Within this subset of searches, a number of patterns stand out:

- Approximately 48% begin with the brand name and end with *digital camera*.
- Approximately 35% are ordered brand, model name, model number, megapixel, *digital camera*.
- Approximately 22.5% are ordered brand, megapixel, model name, *digital camera*.
- A whopping 60% follow the overall pattern of brand, model name, *digital camera*.

You might also notice that, at least in this example, qualifiers such as *new*, a specific store name, and a reference to consumer comments tend to precede the search phrases, whereas features and product-related qualifiers such as *memory*, *3x opt zoom*, *warranty*, *cradle*, *information*, and even a repeat of the megapixels or model number tend to be appended to the search phrases.

NOTE

Remember, this is purely a limited, hypothetical example and certainly is not meant to be statistically accurate; the goal here is to reveal different search term patterns to aid you in determining the best groupings of longtail keywords to target.

Applying Editorial Content Strategies for LongTail Targeting

One of the most difficult aspects of capturing traffic from the long tail of search

is creating relevant, targeted content. As we discussed in "Determining Searcher Intent and Delivering Relevant, Fresh Content", search engines rely on lexical analysis to determine what a web page is about. As a result, your chances of showing up for a longtail phrase are greatly increased if you have that longtail phrase (or at least all the words that make up the longtail phrase) on your page. Let's look at why this may be challenging by checking out what phrases Wordtracker returns when we enter *canon digital camera* (see Figure 5-43).

1	canon digital camera (search)	2,582	698,000	39.8	0.004
2	canon digital cameras (search)	564	1,350,000	2.39	0.000
3	canon digital camera reviews (search)	392	21,000	88.0	0.019
4	canon digital camera in vietnam (search)	288	10	27,600	28.8
5	compare canon digital cameras (search)	244	1,220	180	0.201
6	canon digital camera accessories (search)	117	8,950	6.42	0.013
7	canon digital camera software (search)	115	2,800	8.45	0.041
8	canon digital camera windows 7 screen fix (search)	105	0	–	–
9	canon digital camera lens (search)	100	4,820	3.59	0.021
0	cheap canon digital cameras (search)	95	6,390	10.4	0.015

Figure 5-43. Sample longtail data

Already, with the eighth phrase returned (*canon digital camera windows 7 screen fix*), you can see the challenge. If you are trying to sell Canon digital cameras, you are probably not going to work that phrase into your page copy.

The best approach is to use the longtail research techniques we discussed in this chapter and identify the major patterns, or the major words, that appear across different longtail scenarios, and then work those words into your copy. Don't force it by making pages that appear foolish to a user.

Make sure the writers remain focused on producing quality content. From a longtail perspective, more text is better because it creates more possible longtail matches, but there are limits to that too. Don't put a 1,000-word article on your site unless it makes sense to your users for you to do so.

Applying User-Generated Content Strategies for LongTail Targeting

User-generated content (UGC) can be a great way to obtain lots of content that will help attract longtail traffic. Popular ways of doing that include forums, reviews, blog comments, and a way to upload videos or images, among others. As users submit content, they do the hard work of writing the text you need to capitalize on the long tail.

There are some downsides to UGC, though. Generally speaking, you need to moderate it to make sure people are not contributing objectionable material you don't want on your site, or copyrighted material from other sources (and potential duplicate content issues, depending on how the UGC content is published to your site). Even if you get community members to participate, you will still need to manage them.

In addition, you need to have a strategy for getting the process started. In the case of a forum, you need to develop a critical mass of users to establish a real community. If you don't, a high percentage of the posts you receive will be one form of spam or another. To make UGC work, you need one or more of the following:

- Significant existing daily site traffic (how much depends on how vertically oriented your community is intended to be—narrowly focused topics can get going with a smaller number of users)
- A way to generate a lot of buzz to generate site traffic
- Compelling supporting content

If you can succeed at this, you'll give life to a machine that produces longtail content on an ongoing basis with comparatively low effort.

Trending, Seasonality, and Seasonal Fluctuations in Keyword Demand

One of the subtleties of keyword research, and of any fully developed SEO strategy, is that the use of keywords and phrases varies significantly over time.

For instance, major holidays inevitably lead to bursts of keyword volume related to those holidays. Examples could be searches such as *Halloween costumes*, *gift ideas for Christmas*, or *Valentine's candy*.

If you want to write holiday-related content, it is important to have your site visible in the SERPs for those search queries prior to that holiday's buying season so that you'll get optimum traffic for those terms. And because the search engines take considerable time in ranking your pages, advance preparation is required. To investigate this further, let's examine the Google Trends data for a period of 12 months for the search term *halloween costumes* (see Figure 5-44).

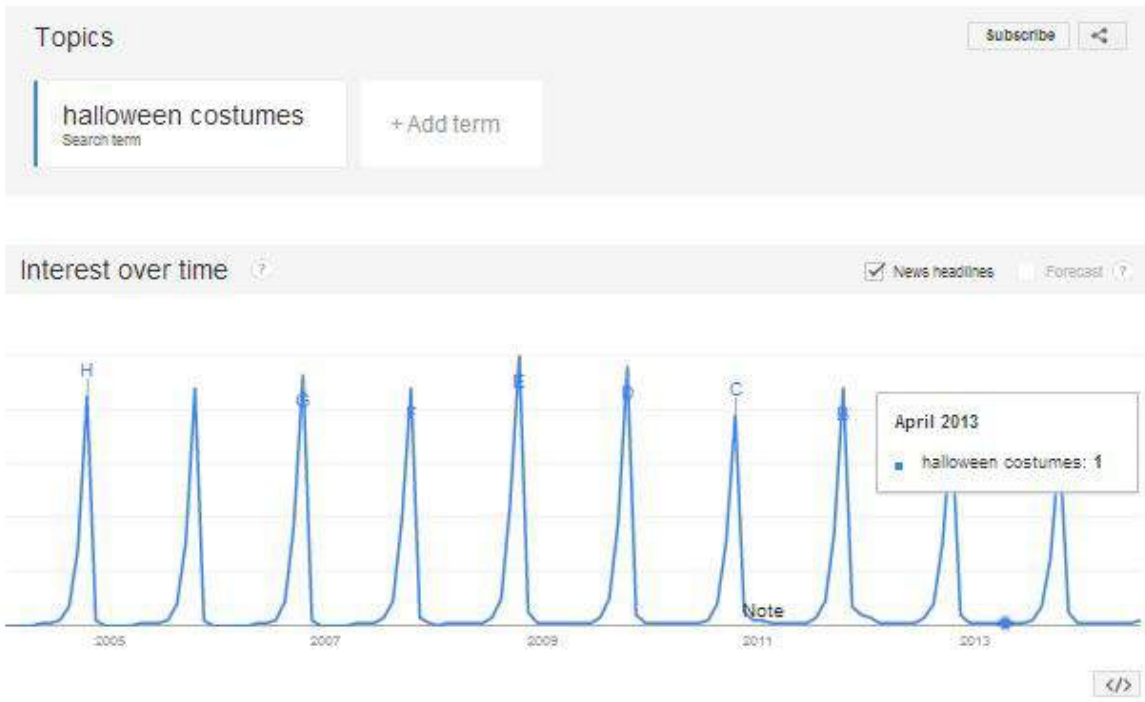


Figure 5-44. Google Trends highly seasonal data example

As you can see, searches begin gaining traction toward the end of August and into autumn; thus, if you are doing SEO for Halloween-related terms, you would want to have the related content and links in place by the beginning of the summer so that search engines can find and index your content, and therefore you're more visible to searchers when they start doing research. A long-term SEO approach would take this into consideration as part of the overall strategy for the site. Searches start consistently increasing toward the end of September.

A similar pattern emerges for Christmas-related searches. Figure 5-45 shows an example for *firecrackers*. Searches start consistently increasing in early June. You can keep trying more examples. With Valentine's Day, the searches start in mid-December.

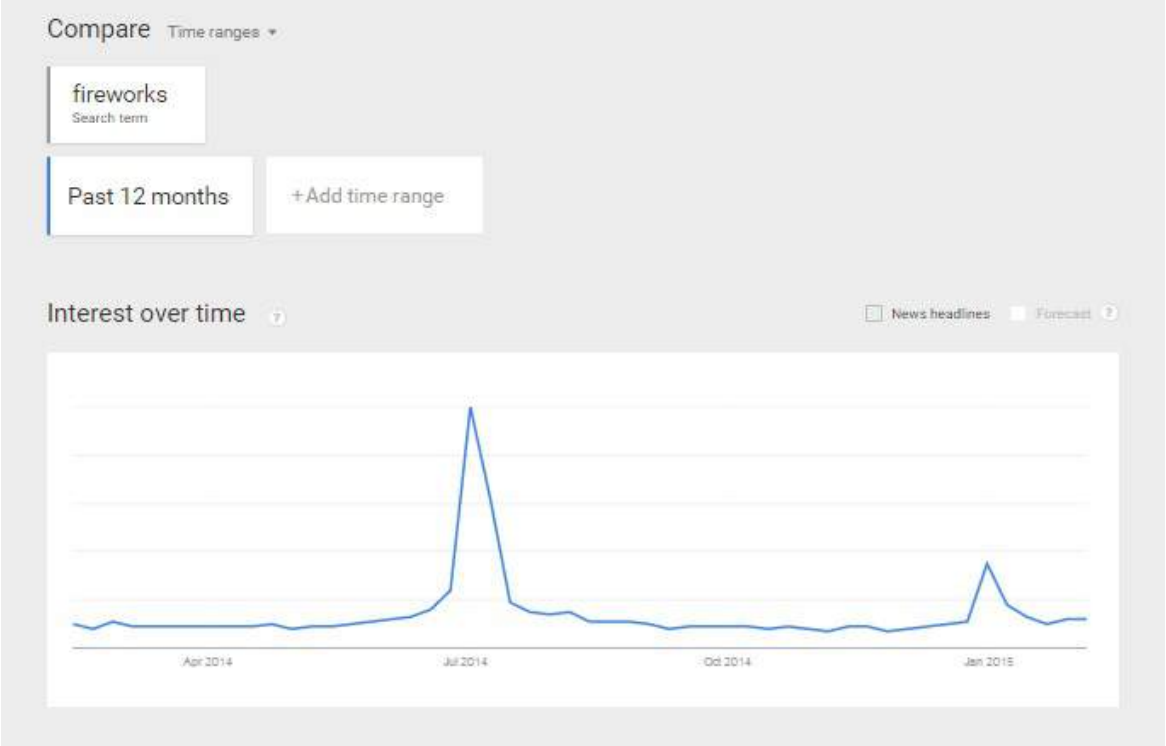


Figure 5-45. Google Trends; another seasonal example

In each case, searches started increasing about two to three months before the holiday, so it is important to acknowledge that and start crafting your content and targeting those keywords in ample time for them to be indexed before the searches start gaining traction.

KeywordDiscovery also graphs search trends, so if you have an account, you can analyze these graphs to craft a holiday campaign, as shown in Figure 5-46.

Figure 5-46. KeywordDiscovery seasonal data example

You can see when people begin to search for Halloween costumes and when the activity drops off.

Don’t take your cue from when the stores start stocking Halloween candy—do the research and find out what last year’s trends were so that you’re prepared this year. If you prepare early enough, you’ll be ready, while your competitors are scrambling with last-minute link-building campaigns three weeks before the holiday.

Also, don’t remove your Halloween (or other seasonal) page as soon as the time frame has passed. Once you have fought hard to get rankings for your seasonal trophy term, you want to make sure you get the benefit for that hard work next year. Too many sites delete or archive these seasonal pages after the season is over, and then they have to start over again next year. A better strategy is to leave the page in place until a new version is created, reuse the same URL, and archive the old content to a different URL. Leaving the page in place will give you a jumpstart when it is time to begin ramping up next year.

Conclusion

Keyword research is a complex and time-consuming component of search engine optimization, but the rewards are high; once you learn where the keyword search volume is, you can begin to think about how that affects the information architecture and the navigation structure of your site. These are two critical elements that we will explore in greater detail in Chapter 6.

[1] Rand Fishkin, “Rewriting the Beginner’s Guide - Part V,” Moz News, February 24, 2008, *https://moz.com/blog/rewriting-the-beginners-guide-part-v-keyword-research*.

[2] Duane Forrester, “Bing Moving to Encrypt Search Traffic by Default,” Bing Blogs, June 15, 2015, *http://bit.ly/bing_default_encrypt*.

[3] DigitalRelevance, “A Tale of Two Studies: Establishing Google & Bing Click-Through Rates,” *http://bit.ly/tale_of_two_studies*.

CHAPTER SIX

Developing an SEO-Friendly Website

In this chapter, we will examine ways to assess the search engine friendliness of your website. A search engine–friendly website, at the most basic level, is one that allows for search engine access to site content—and having your site content accessible to search engines is the first step toward creating prominent visibility in search results. Once your site's content is accessed by a search engine, it can then be considered for relevant positioning within search results pages.

As we discussed in the introduction to Chapter 2, search engine crawlers are basically software programs, and like all software programs, they come with certain strengths and weaknesses. Publishers must adapt their websites to make the job of these software programs easier—in essence, leverage their strengths and make their weaknesses irrelevant. If you can do this, you will have taken a major step toward success with SEO.

Developing an SEO-friendly site architecture requires a significant amount of thought, planning, and communication due to the large number of factors that influence how a search engine sees your site and the myriad ways in which a website can be put together, as there are hundreds (if not thousands) of tools that web developers can use to build a website—many of which were not initially designed with SEO or search engine crawlers in mind.

Making Your Site Accessible to Search Engines

The first step in the SEO design process is to ensure that your site can be found and crawled by search engines. This is not as simple as it sounds, as there are many popular web design and implementation constructs that the crawlers may not understand.

Indexable Content

To rank well in the search engines, your site's content—that is, the material available to visitors of your site—should be in HTML text form. Images and Flash files, for example, while crawled by the search engines, are content types that are more difficult for search engines to analyze and therefore are not ideal for communicating to search engines the topical relevance of your pages.

Search engines have challenges with identifying the relevance of images because there are minimum text-input fields for image files in GIF, JPEG, or PNG format (namely the filename, title, and `alt` attribute). While we do strongly recommend accurate labeling of images in these fields, images alone are usually not enough to earn a web page top rankings for relevant queries. While image identification technology continues to advance, processing power limitations will likely keep the search engines from broadly applying this type of analysis to web search in the near future.

Google enables users to perform a search using an image, as opposed to text, as the search query (though users can input text to augment the query). By uploading an image, dragging and dropping an image from the desktop, entering an image URL, or right-clicking on an image within a browser (Firefox and Chrome with installed extensions), users can often find other locations of that image on the Web for reference and research, as well as images that appear similar in tone and composition. While this does not immediately change the landscape of SEO for images, it does give us an indication of how Google is potentially augmenting its current relevance indicators for image content.

With Flash, while specific *.swf* files (the most common file extension for Flash) can be crawled and indexed—and are often found when a user conducts a *.swf* file search for specific words or phrases included in their filename—it is rare for a generic query to return a Flash file or a website generated entirely in Flash as a highly relevant result, due to the lack of "readable" content. This is not to say that websites developed using Flash are inherently irrelevant, or that it is impossible to successfully optimize a website that uses Flash; however, in our experience the preference must still be given to HTML-based files.

Spiderable Link Structures

As we outlined in Chapter 2, search engines use links on web pages to help them discover other web pages and websites. For this reason, we strongly recommend taking the time to build an internal linking structure that spiders can crawl easily. Many sites make the critical mistake of hiding or obfuscating their navigation in ways that limit spider accessibility, thus impacting their ability to get pages listed in the search engines' indexes. Consider Figure 6-1, which shows how this problem can occur.

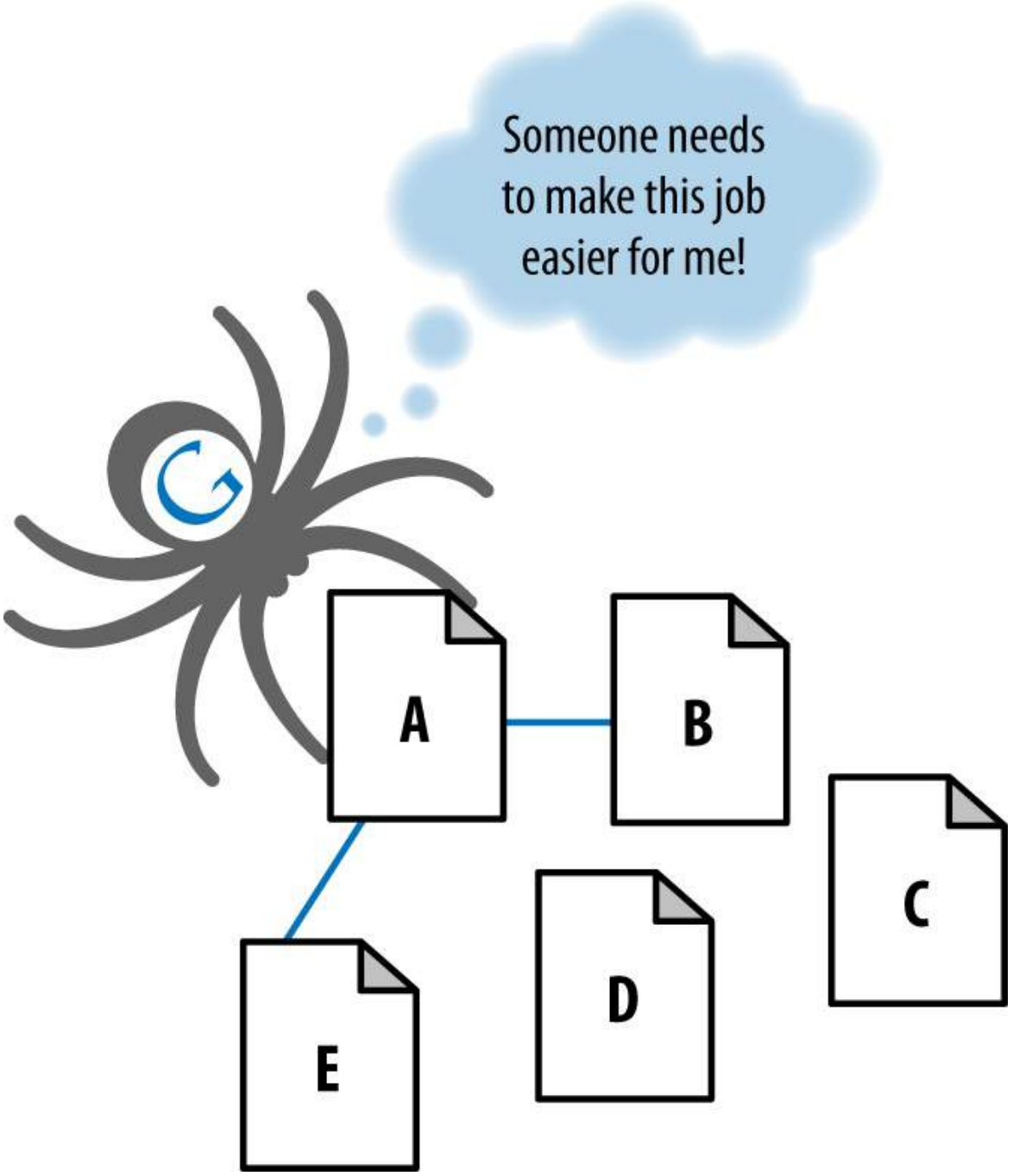


Figure 6-1. Providing search engines with crawlable link structures

In Figure 6-1, Google's spider has reached Page A and sees links to pages B and E. However, even though pages C and D might be important pages on the site, the spider has no way to reach them (or even to know they exist) because no direct, crawlable links point to those pages. As far as Google is concerned, they

might as well not exist. Great content, good keyword targeting, and smart marketing won't make any difference at all if the spiders can't reach those pages in the first place.

To refresh your memory of the discussion in Chapter 2, here are some common reasons why pages may not be reachable:

Links in submission-required forms

Search spiders will rarely, if ever, attempt to "submit" forms, and thus, any content or links that are accessible only via a form are invisible to the engines. This even applies to simple forms such as user logins, search boxes, or some types of pull-down lists.

Links in hard-to-parse JavaScript

If you use JavaScript for links, you may find that search engines either do not crawl or give very little weight to the embedded links. In June 2014, Google announced enhanced crawling of JavaScript and CSS. Google can now render some JavaScript and follow some JavaScript links. Due to this change, Google recommends against blocking it from crawling your JavaScript and CSS files. For a preview of how your site might render according to Google, go to Search Console -> Crawl -> Fetch as Google, input the URL you would like to preview, and select "Fetch and Render."

Links in Java or other plugins

Traditionally, links embedded inside Java and plugins have been invisible to the engines.

Links in Flash

In theory, search engines can detect links within Flash, but don't rely too heavily on this capability.

Links in PowerPoint and PDF files

Search engines sometimes report links seen in PowerPoint files or PDFs. These links are believed to be counted the same as links embedded in HTML documents.

Links pointing to pages blocked by the meta `robots` *tag,* `rel="nofollow"`*, or robots.txt*

The *robots.txt* file provides a very simple means for preventing web spiders from crawling pages on your site. Using the `nofollow` attribute on a link, or placing the meta `robots nofollow` tag with the `content="nofollow"` attribute on the page containing the link, instructs the search engine to not pass link authority via the link (a concept we will discuss further in "Content Delivery and Search Spider Control"). The effectiveness of the `nofollow` attribute on links has greatly diminished to the point of irrelevance as a result of overmanipulation by aggressive SEO practitioners. For more on this, see the blog post "PageRank Sculpting", by Google's Matt Cutts.

Links on pages with many hundreds or thousands of links

Historically, Google had suggested a maximum of 100 links per page before it may stop spidering additional links from that page, but this recommendation has softened over time. Think of it more as a strategic guideline for passing PageRank. If a page has 200 links on it, then none of the links get very much PageRank. Managing how you pass PageRank by limiting the number of links is usually a good idea. Tools such as Screaming Frog can run reports on the number of outgoing links you have per page.

Links in frames or iframes

Technically, links in both frames and iframes can be crawled, but both present structural issues for the engines in terms of organization and following. Unless you're an advanced user with a good technical understanding of how search engines index and follow links in frames, it is best to stay away from them as a place to offer links for crawling purposes. We will discuss frames and iframes in more detail in "Creating an Optimal Information Architecture".

XML Sitemaps

Google, Yahoo!, and Bing (formerly MSN Search, and then Live Search) all support a protocol known as XML Sitemaps. Google first announced it in 2005,

and then Yahoo! and MSN Search agreed to support the protocol in 2006. Using the Sitemaps protocol, you can supply search engines with a list of all the URLs you would like them to crawl and index.

Adding a URL to a sitemap file does not guarantee that a URL will be crawled or indexed. However, it can result in the search engine discovering and indexing pages that it otherwise would not.

This program is a complement to, not a replacement for, the search engines' normal, link-based crawl. The benefits of sitemaps include the following:

- For the pages the search engines already know about through their regular spidering, they use the metadata you supply, such as the last date the content was modified (*lastmod date*) and the frequency at which the page is changed (*changefreq*), to improve how they crawl your site.
- For the pages they don't know about, they use the additional URLs you supply to increase their crawl coverage.
- For URLs that may have duplicates, the engines can use the XML Sitemaps data to help choose a canonical version.
- Verification/registration of XML sitemaps may indicate positive trust/authority signals.
- The crawling/inclusion benefits of sitemaps may have second-order positive effects, such as improved rankings or greater internal link popularity.
- Having a sitemap registered with Google Search Console can give you extra analytical insight into whether your site is suffering from indexation, crawling, or duplicate content issues.

Matt Cutts, the former head of Google's webspam team, has explained XML sitemaps in the following way:

> *Imagine if you have pages A, B, and C on your site. We find pages A and B through our normal web crawl of your links. Then you build a Sitemap and list the pages B and C. Now there's a chance (but not a promise) that we'll crawl page C. We won't drop page A just because you didn't list it in your Sitemap. And just because you listed a page that we didn't know about doesn't*

guarantee that we'll crawl it. But if for some reason we didn't see any links to C, or maybe we knew about page C but the URL was rejected for having too many parameters or some other reason, now there's a chance that we'll crawl that page C.[1]

Sitemaps use a simple XML format that you can learn about at *http://www.sitemaps.org/*. XML sitemaps are a useful and in some cases essential tool for your website. In particular, if you have reason to believe that the site is not fully indexed, an XML sitemap can help you increase the number of indexed pages. As sites grow in size, the value of XML sitemap files tends to increase dramatically, as additional traffic flows to the newly included URLs.

Laying out an XML sitemap

The first step in the process of creating an XML sitemap is to create an XML sitemap file in a suitable format. Because creating an XML sitemap requires a certain level of technical know-how, it would be wise to involve your development team in the XML sitemap generator process from the beginning. Figure 6-2 shows an example of some code from a sitemap.

```
<?xml version="1.0" encoding="UTF-8"?>

<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">

   <url>

      <loc>http://www.example.com/</loc>

      <lastmod>2005-01-01</lastmod>

      <changefreq>monthly</changefreq>

      <priority>0.8</priority>

   </url>

</urlset>
```

Figure 6-2. Sample XML sitemap from Google.com

To create your XML sitemap, you can use the following:

An XML sitemap generator

This is a simple script that you can configure to automatically create sitemaps, and sometimes submit them as well. Sitemap generators can create these sitemaps from a URL list, access logs, or a directory path hosting static files corresponding to URLs. Here are some examples of XML sitemap generators:

- SourceForge.net's Google-sitemap_gen
- XML-Sitemaps.com Sitemap Generator
- Sitemaps Pal
- GSite Crawler

Simple text

You can provide Google with a simple text file that contains one URL per line. However, Google recommends that once you have a text sitemap file for your site, you use the sitemap generator to create a sitemap from this text file using the Sitemaps protocol.

Syndication feed

Google accepts Really Simple Syndication (RSS) 2.0 and Atom 1.0 feeds. Note that the feed may provide information on recent URLs only.

Deciding what to include in a sitemap file

When you create a sitemap file, you need to take care in situations where your site has multiple URLs that refer to one piece of content: include *only* the preferred (canonical) version of the URL, as the search engines may assume that the URL specified in a sitemap file is the preferred form of the URL for the content. You can use the sitemap file to indicate to the search engines which URL points to the preferred version of a given page.

In addition, be careful about what *not* to include. For example, do not include multiple URLs that point to identical content, and leave out pages that are simply pagination pages (or alternate sort orders for the same content) and/or any low-

value pages on your site. Last but not least, make sure that none of the URLs listed in the sitemap file include any tracking parameters.

Mobile sitemaps

Mobile sitemaps should be used for content targeted at mobile devices. Mobile information is kept in a separate sitemap file that should not contain any information on nonmobile URLs. Google supports nonmobile markup, XHTML mobile profile, WML (WAP 1.2) and cHTML. Details on the mobile sitemap format can be found here: *https://support.google.com/webmasters/answer/34648*.

Video sitemaps

Including information on your videos in your sitemap file will increase their chances of being discovered by search engines. Google supports the following video formats: *.mpg*, *.mpeg*, *.mp4*, *.m4v*, *.mov*, *.wmv*, *.asf*, *.avi*, *.ra*, *.ram*, *.rm*, *.flv*, and *.swf*. You can see the specification on how to implement video sitemap entries here: *https://support.google.com/webmasters/answer/80472*.

Image sitemaps

You can increase visibility for your images by listing them in your sitemap file. For each URL you list in your sitemap file, you can also list the images that appear on those pages. You can list up to 1,000 images per page. Specialized image tags are associated with the URL. The details of the format of these tags are on this page: *https://support.google.com/webmasters/answer/178636*.

Listing images in the sitemap does increase the chances of those images being indexed. If you list some images and not others, it may be interpreted as a signal that the unlisted images are less important.

Uploading your sitemap file

When your sitemap file is complete, upload it to your site in the highest-level directory you want search engines to crawl (generally, the root directory), such as *www.yoursite.com/sitemap.xml*. You can include more than one subdomain in your sitemap provided that you verify the sitemap for each subdomain in Google Search Console, though it’s frequently easier to understand what’s happening with indexation if each subdomain has its own sitemap and its own profile in Google Search Console.

Managing and updating XML sitemaps

Once your XML sitemap has been accepted and your site has been crawled, monitor the results and update your sitemap if there are issues. With Google, you can return to your Google Search Console account to view the statistics and diagnostics related to your XML sitemaps. Just click the site you want to monitor. You'll also find some FAQs from Google on common issues such as slow crawling and low indexation.

Update your XML sitemap with Google and Bing when you add URLs to your site. You'll also want to keep your sitemap file up to date when you add a large volume of pages or a group of pages that are strategic.

There is no need to update the XML sitemap when you're simply updating content on existing URLs. It is not strictly necessary to update when pages are deleted, as the search engines will simply not be able to crawl them, but do update before you have too many broken pages in your feed. Also update your sitemap file whenever you add any new content, and remove any deleted pages at that time. Google and Bing will periodically redownload the sitemap, so you don't need to resubmit your sitemap to Google or Bing unless your sitemap location has changed.

Enable Google and Bing to autodiscover your XML sitemap locations by using the `Sitemap:` directive in your site's *robots.txt* file.

If you are adding or deleting large numbers of new pages to your site on a regular basis, you may want to use a utility, or have your developers build the ability, for your XML sitemap to regenerate with all of your current URLs on a regular basis. Many sites regenerate their XML sitemap daily via automated scripts.

Google and the other major search engines discover and index websites by crawling links. Google XML sitemaps are a way to feed the URLs that you want crawled on your site to Google for more complete crawling and indexation, which results in improved long-tail searchability. By creating and updating this XML file, you help to ensure that Google recognizes your entire site, and this recognition will help people find your site. It also helps all of the search engines understand which version of your URLs (if you have more than one URL pointing to the same content) is the canonical version.

Creating an Optimal Information Architecture

Making your site friendly to search engine crawlers also requires that you put some thought into your site's *information architecture* (IA). A well-designed site architecture can bring many benefits for both users and search engines.

The Importance of a Logical, Category-Based Flow

Search engines face myriad technical challenges in understanding your site, as crawlers are not able to perceive web pages in the way that humans do, creating significant limitations for both accessibility and indexing. A logical and properly constructed website architecture can help overcome these issues and bring great benefits in search traffic and usability.

At the core of website information architecture are two critical principles: usability (making a site easy to use) and information architecture (crafting a logical, hierarchical structure for content).

One of the very early information architecture proponents, Richard Saul Wurman, developed the following definition for *information architect*:[2]

> *1) the individual who organizes the patterns inherent in data, making the complex clear. 2) a person who creates the structure or map of information which allows others to find their personal paths to knowledge. 3) the emerging 21st century professional occupation addressing the needs of the age focused upon clarity, human understanding, and the science of the organization of information.*

Usability and search friendliness

Search engines are trying to reproduce the human process of sorting relevant web pages by quality. If a real human were to do this job, usability and user experience would surely play a large role in determining the rankings. Given that search engines are machines and don't have the ability to segregate by this metric quite so easily, they are forced to employ a variety of alternative, secondary metrics to assist in the process. The most well known and well publicized among these is a measurement of the inbound links to a website (see Figure 6-3), and a well-organized site is more likely to receive links.

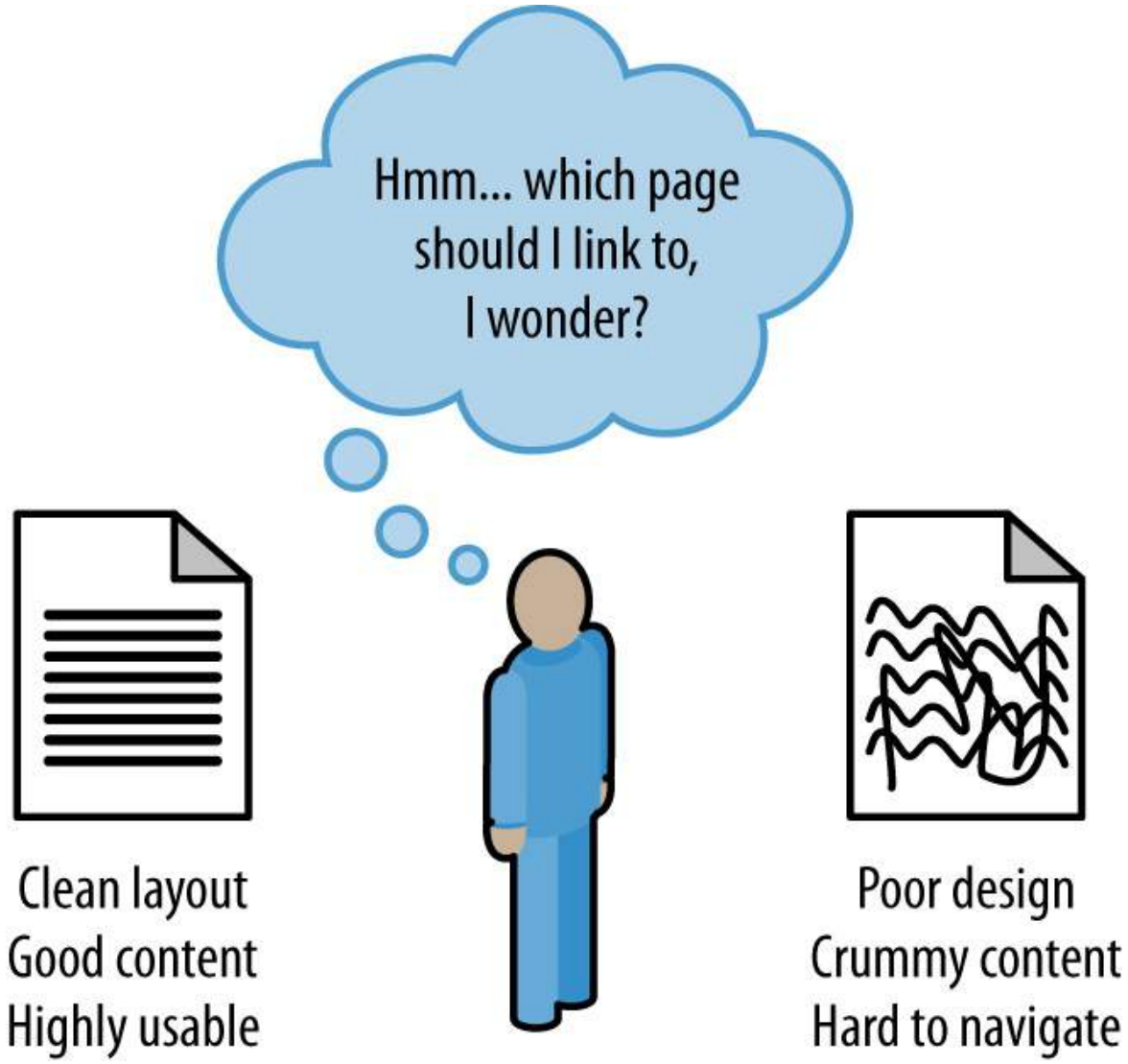


Figure 6-3. Make your site attractive to link to

Since Google launched in the late 1990s, search engines have strived to analyze every facet of the link structure on the Web and have extraordinary abilities to infer trust, quality, reliability, and authority via links. If you push back the curtain and examine why links between websites exist and how they come to be, you can see that a human being (or several humans, if the organization suffers from bureaucracy) is almost always responsible for the creation of links.

The engines hypothesize that high-quality links will point to high-quality content, and that great content and positive user experiences will be rewarded with more links than poor user experiences. In practice, the theory holds up well.

Modern search engines have done a very good job of placing good-quality, usable sites in top positions for queries.

An analogy

Look at how a standard filing cabinet is organized. You have the individual cabinet, drawers in the cabinet, folders within the drawers, files within the folders, and documents within the files (see Figure 6-4).

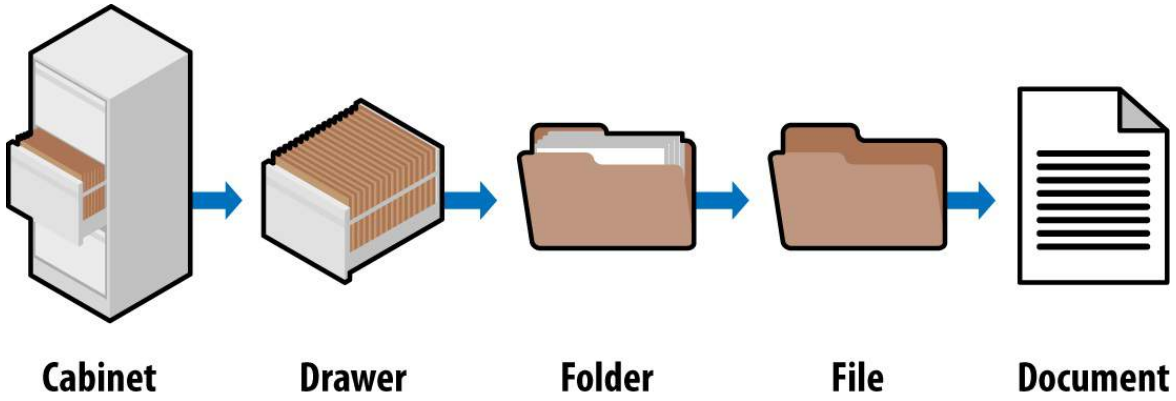


Figure 6-4. Similarities between filing cabinets and web pages

There is only one copy of any individual document, and it is located in a particular spot. There is a very clear navigation path to get to it.

If you want to find the January 2015 invoice for a client (Amalgamated Glove & Spat), you would go to the cabinet, open the drawer marked Client Accounts, find the Amalgamated Glove & Spat folder, look for the Invoices file, and then flip through the documents until you come to the January 2015 invoice (again, there is only one copy of this; you won't find it anywhere else).

Figure 6-5 shows what it looks like when you apply this logic to the popular website, Craigslist.

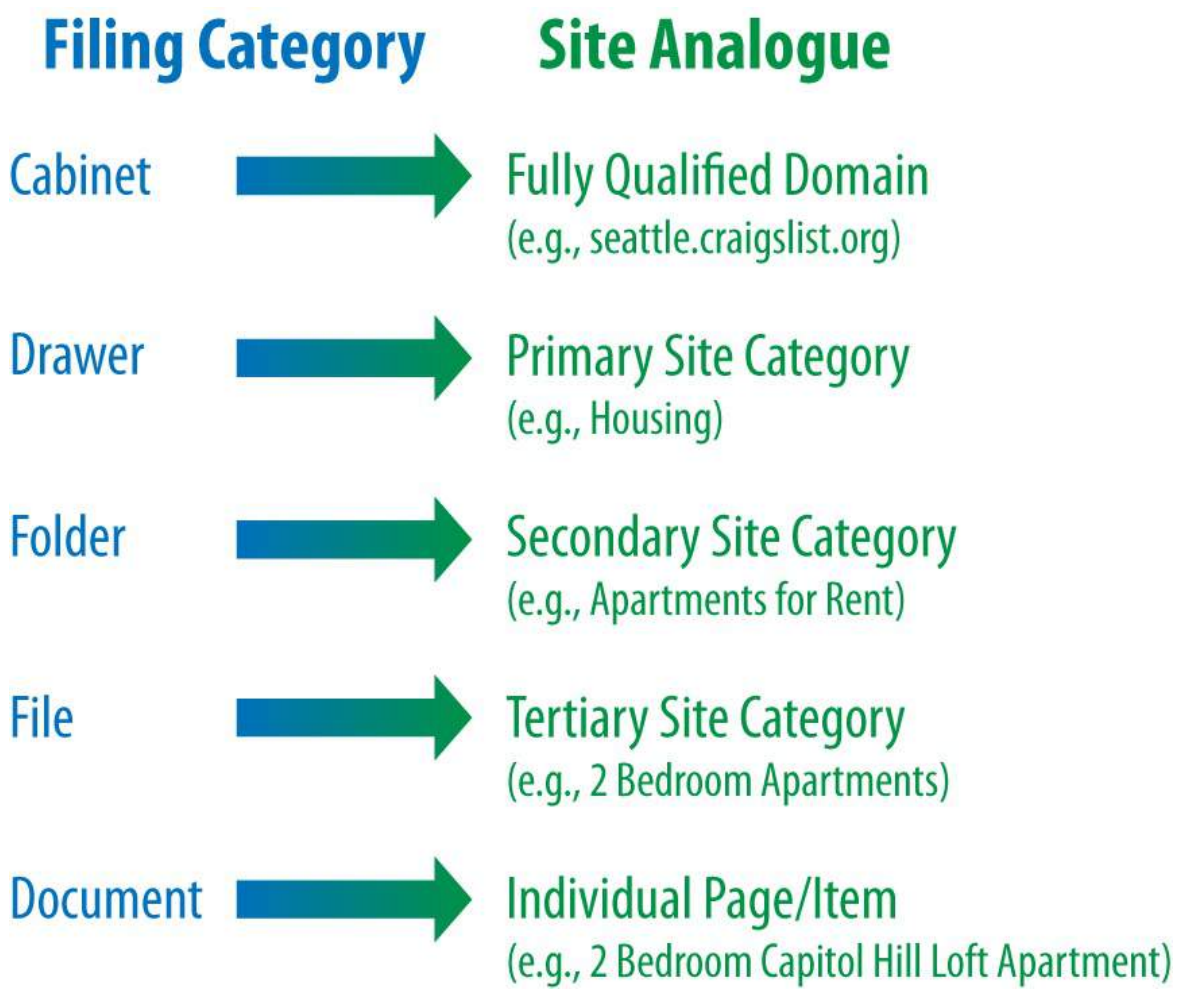


Figure 6-5. Filing cabinet analogy applied to Craigslist

If you're seeking an apartment in Los Angeles, you'd navigate to *http://losangeles.craigslist.org/*, choose apts/housing, narrow that down to two bedrooms, and pick the two-bedroom loft from the list of available postings. Craigslist's simple, logical information architecture makes it easy for you to reach the desired post in four clicks, without having to think too hard at any step about where to go. This principle applies perfectly to the process of SEO, where good information architecture dictates:

- As few clicks as possible to any given page
- One hundred or fewer links per page (so as not to overwhelm either crawlers or visitors)

- A logical, semantic flow of links from home page to categories to detail pages

Here is a brief look at how this basic filing cabinet approach can work for some more complex information architecture issues.

Subdomains

You should think of subdomains as completely separate filing cabinets within one big room. They may share similar architecture, but they shouldn't share the same content; and more importantly, if someone points you to one cabinet to find something, he is indicating that *that* cabinet is the authority, not the other cabinets in the room. Why is this important? It will help you remember that links (i.e., votes or references) to subdomains may not pass all, or any, of their authority to other subdomains within the room (e.g., **.craigslist.org*, wherein * is a variable subdomain name).

Those cabinets, their contents, and their authority are isolated from one another and may not be considered to be associated with one another. This is why, in most cases, it is best to have one large, well-organized filing cabinet instead of several that may prevent users and bots from finding what they want.

Redirects

If you have an organized administrative assistant, he probably uses 301 redirects (these are discussed more in the section "Redirects") inside his literal, metal filing cabinet. If he finds himself looking for something in the wrong place, he might put a sticky note there reminding him of the correct location the next time he needs to look for that item. Anytime he looked for something in those cabinets, he could always find it because if he navigated improperly, he would inevitably find a note pointing him in the right direction.

Redirect irrelevant, outdated, or misplaced content to the proper spot in your filing cabinet, and both your users and the engines will know what qualities and keywords you think it should be associated with.

URLs

It would be tremendously difficult to find something in a filing cabinet if every time you went to look for it, it had a different name, or if that name resembled *jklhj25br3g452ikbr52k*—a not-so-uncommon type of character string found in dynamic website URLs. Static, keyword-targeted URLs are much better for

users and bots alike. They can always be found in the same place, and they give semantic clues as to the nature of the content.

These specifics aside, thinking of your site information architecture as a virtual filing cabinet is a good way to make sense of best practices. It'll help keep you focused on a simple, easily navigated, easily crawled, well-organized structure. It is also a great way to explain an often-complicated set of concepts to clients and coworkers.

Because search engines rely on links to crawl the Web and organize its content, the architecture of your site is critical to optimization. Many websites grow organically and, like poorly planned filing systems, become complex, illogical structures that force people (and spiders) to struggle to find what they want.

Site Architecture Design Principles

In planning your website, remember that nearly every user will initially be confused about where to go, what to do, and how to find what she wants. An architecture that recognizes this difficulty and leverages familiar standards of usability with an intuitive link structure will have the best chance of making a visit to the site a positive experience. A well-organized site architecture helps solve these problems, and provides semantic and usability benefits to both users and search engines.

As shown in Figure 6-6, a recipes website can use intelligent architecture to fulfill visitors' expectations about content and create a positive browsing experience. This structure not only helps humans navigate a site more easily, but also helps the search engines to see that your content fits into logical concept groups. You can use this approach to help you rank for applications of your product in addition to attributes of your product.

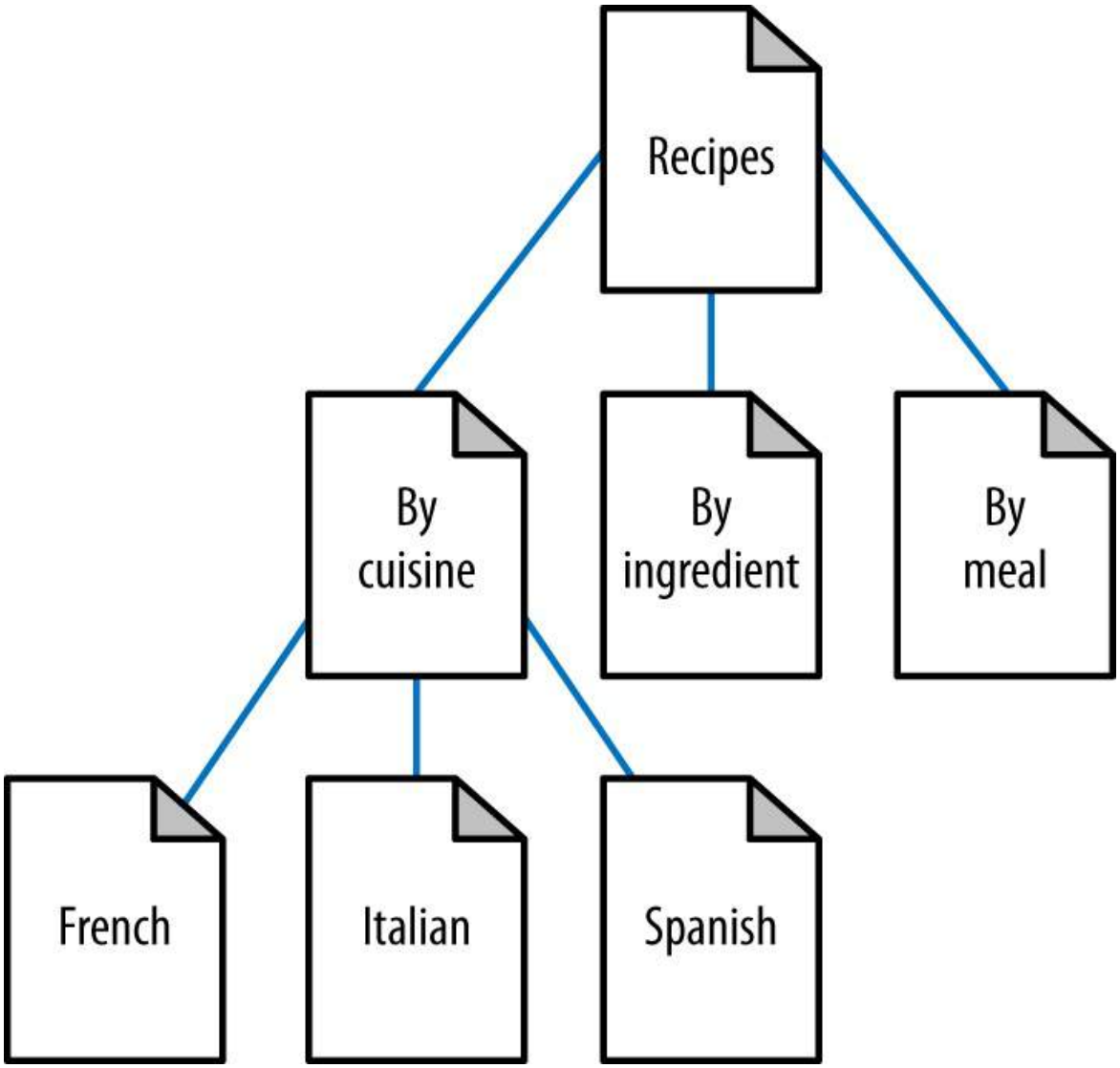


Figure 6-6. Structured site architecture

Although site architecture accounts for a small part of the algorithms, search engines do make use of relationships between subjects and give value to content that has been organized sensibly. For example, if in Figure 6-6 you were to randomly jumble the subpages into incorrect categories, your rankings could suffer. Search engines, through their massive experience with crawling the Web, recognize patterns in subject architecture and reward sites that embrace an intuitive content flow.

Site architecture protocol

Although site architecture—the creation of structure and flow in a website's topical hierarchy—is typically the territory of information architects (or is created without assistance from a company's internal content team), its impact on search engine rankings, particularly in the long run, is substantial. It is, therefore, a wise endeavor to follow basic guidelines of search friendliness.

The process itself should not be overly arduous, if you follow this simple protocol:

1. List all of the requisite content pages (blog posts, articles, product detail pages, etc.).

2. Create top-level navigation that can comfortably hold all of the unique types of detailed content for the site.

3. Reverse the traditional top-down process by starting with the detailed content and working your way up to an organizational structure capable of holding each page.

4. Once you understand the bottom, fill in the middle. Build out a structure for subnavigation to sensibly connect top-level pages with detailed content. In small sites, there may be no need for this level, whereas in larger sites, two or even three levels of subnavigation may be required.

5. Include secondary pages such as copyright, contact information, and other nonessentials.

6. Build a visual hierarchy that shows (to at least the last level of subnavigation) each page on the site.

Figure 6-7 shows an example of a well-structured site architecture.

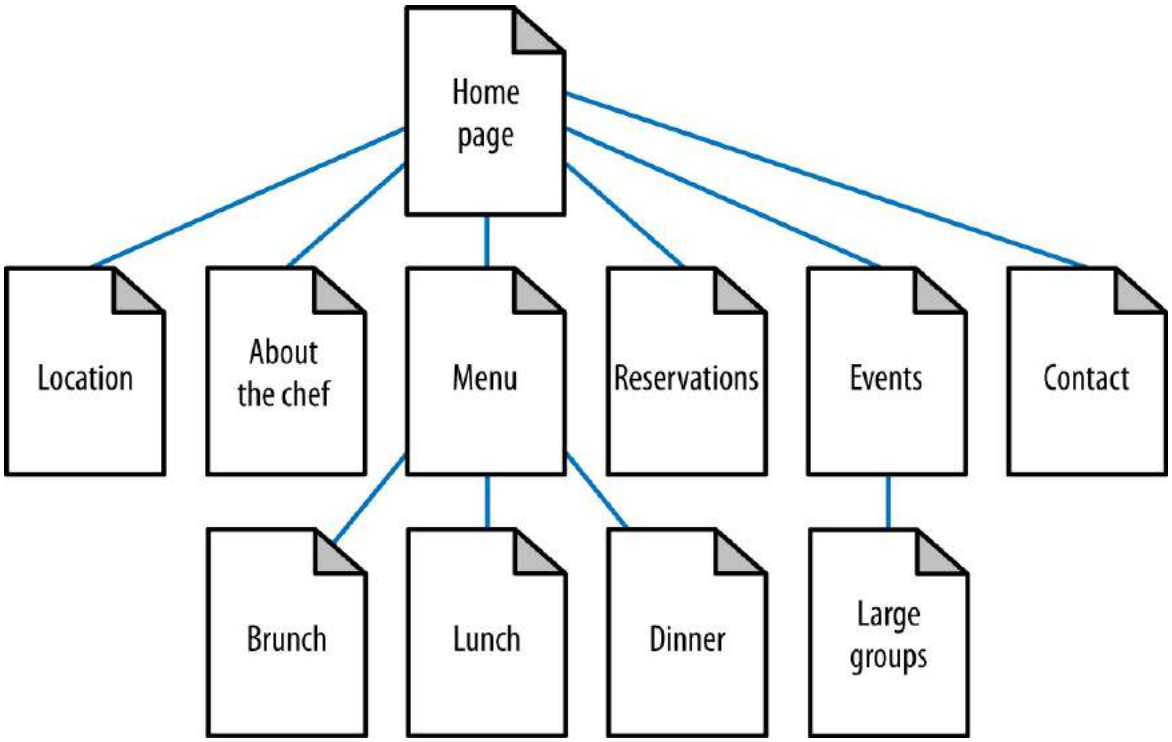


Figure 6-7. Second example of structured site architecture

Category structuring

As search engines crawl the Web, they collect an incredible amount of data (millions of gigabytes) on the structure of language, subject matter, and relationships between content. Though not technically an attempt at artificial intelligence, the engines have built a repository capable of making sophisticated determinations based on common patterns. As shown in Figure 6-8, search engine spiders can learn semantic relationships as they crawl thousands of pages that cover a related topic (in this case, dogs).

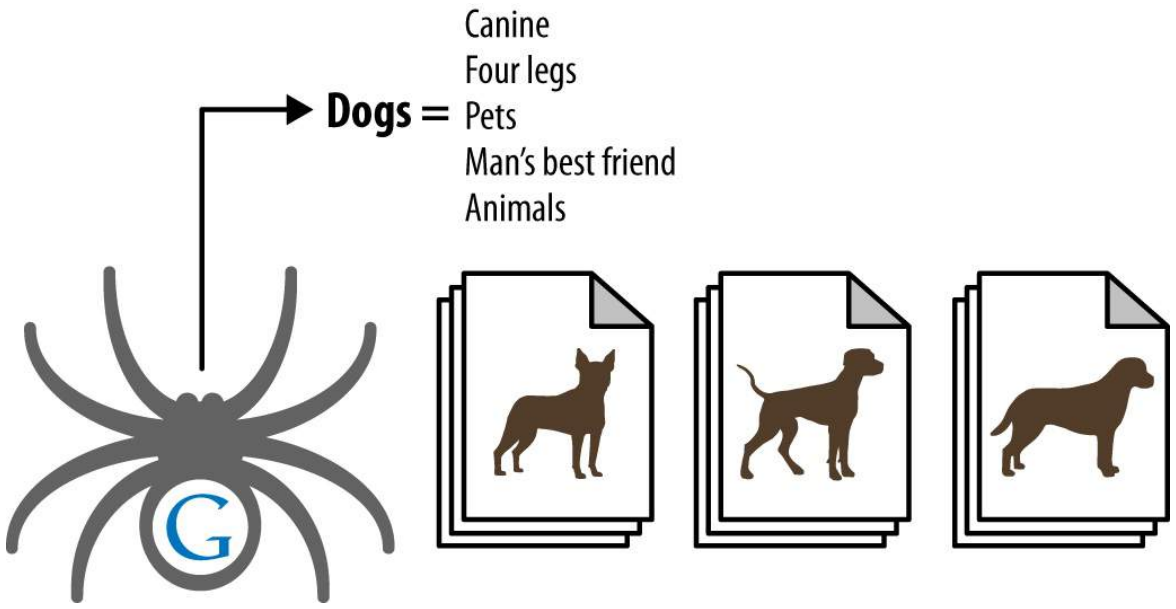


Figure 6-8. Spiders learning semantic relationships

Although content need not always be structured along the most predictable patterns, particularly when a different method of sorting can provide value or interest to a visitor, organizing subjects logically assists both humans (who will find your site easier to use) and engines (which will award you with greater rankings based on increased subject relevance).

Topical relevance

Naturally, this pattern of relevance-based scoring extends from single relationships between documents to the entire category structure of a website. Site creators can take advantage of this best by building hierarchies that flow from broad, encompassing subject matter down to more detailed, specific content. Obviously, in any categorization system, there is a natural level of subjectivity; think first of your visitors, and use these guidelines to ensure that your creativity doesn't overwhelm the project.

Taxonomy and ontology

In designing a website, you should also consider the taxonomy and ontology of the website. Taxonomy is essentially a two-dimensional hierarchical model of the architecture of the site. You can think of ontology as mapping the way the human mind thinks about a topic area. It can be much more complex than

taxonomy, because a larger number of relationship types are often involved.

One effective technique for coming up with an ontology is called *card sorting*. This is a user-testing technique whereby users are asked to group items together so that you can organize your site as intuitively as possible. Card sorting can help identify not only the most logical paths through your site, but also ambiguous or cryptic terminology that should be reworded.

With card sorting, you write all the major concepts onto a set of cards that are large enough for participants to read, manipulate, and organize. Your test group assembles the cards in the order they believe provides the most logical flow, as well as into groups that seem to fit together.

By itself, building an ontology is not part of SEO, but when you do it properly it will impact your site architecture, and therefore it interacts with SEO. Coming up with the right site architecture should involve both disciplines.

Flat Versus Deep Architecture

One very strict rule for search friendliness is the creation of flat site architecture. Flat sites require a minimal number of clicks to access any given page, whereas deep sites create long paths of links required to access detailed content. For nearly every site with fewer than 10,000 pages, all content should be accessible through a maximum of four clicks from the home page and/or sitemap page. That said, flatness should not be forced if it does not make sense for other reasons. At 100 links per page, even sites with millions of pages can have every page accessible in five to six clicks if proper link and navigation structures are employed. If a site is not built to be flat, it can take too many clicks for a user or a search engine to reach the desired content, as shown in Figure 6-9. In contrast, a flat site (see Figure 6-10) allows users and search engines to reach most content in just a few clicks.

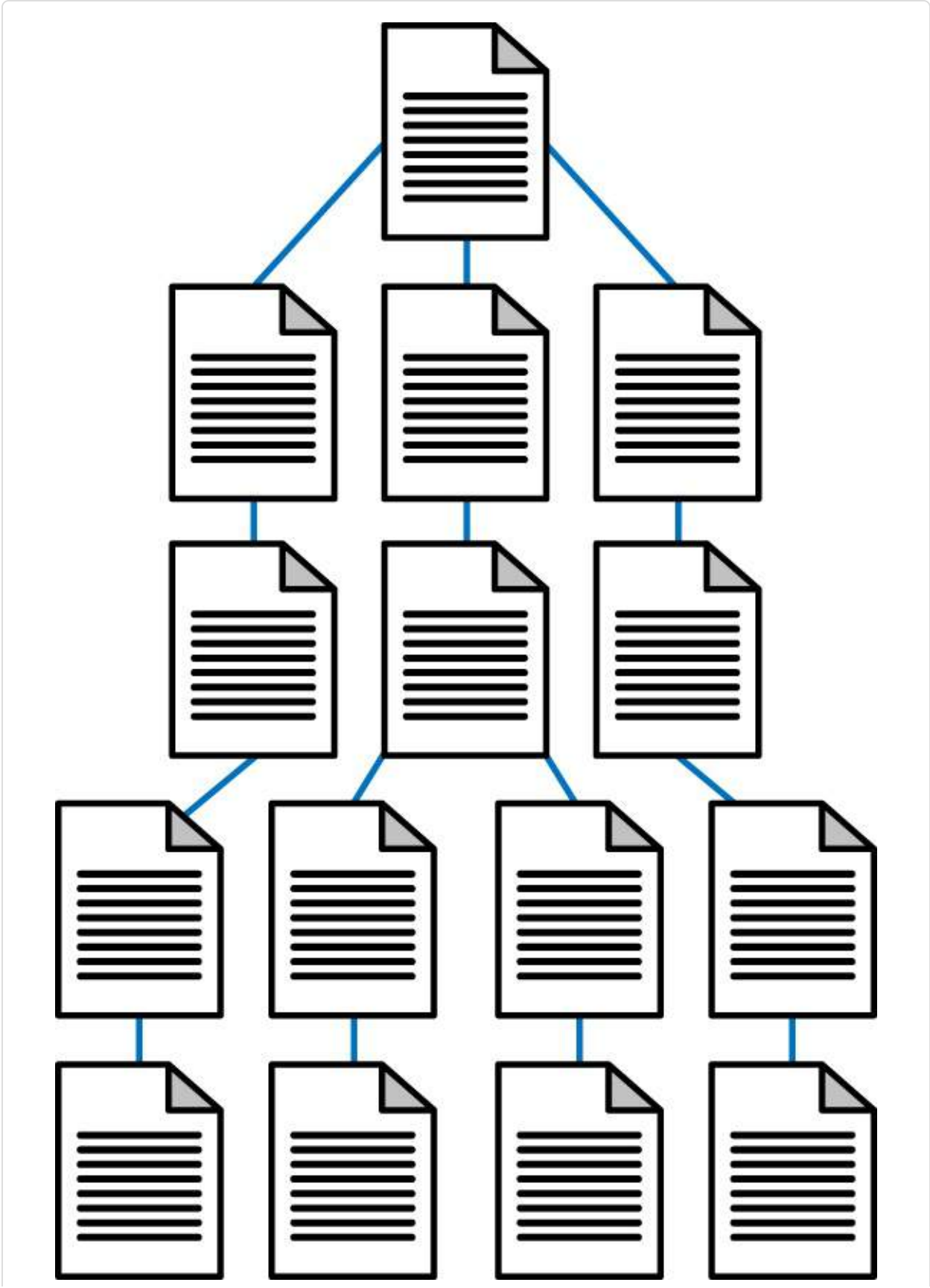


Figure 6-9. Deep site architecture

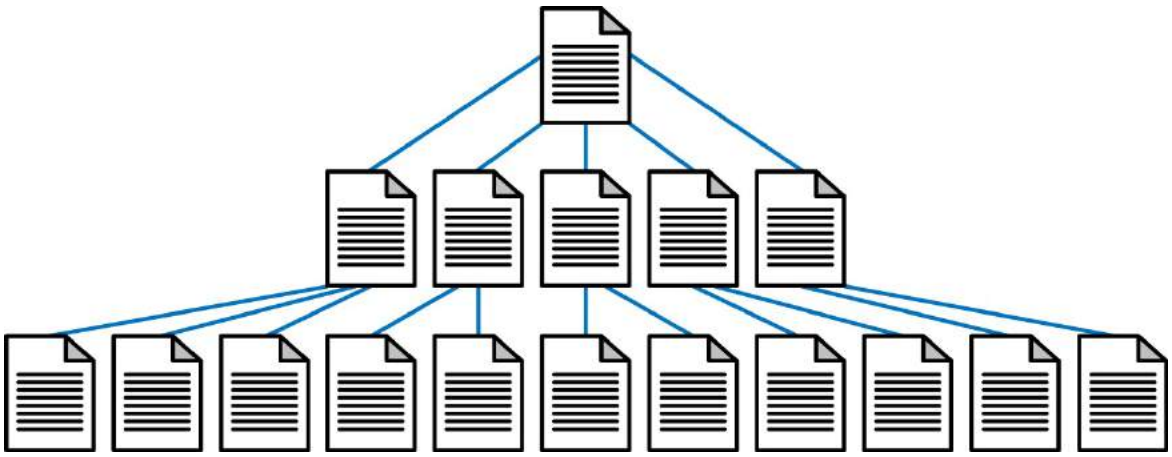


Figure 6-10. Flat site architecture

Flat sites aren't just easier for search engines to crawl; they are also simpler for users, as they limit the number of page visits the user requires to reach his destination. This reduces the abandonment rate and encourages repeat visits.

When creating flat sites, be careful to not overload pages with links either. Pages that have 200 links on them are not passing much PageRank to any of those pages. While flat site architectures are desirable, you should not force an architecture to be overly flat if it is not otherwise logical to do so.

The issue of the number of links per page relates directly to another rule for site architects: avoid excessive pagination wherever possible. *Pagination* (see Figure 6-11), the practice of creating a list of elements on pages separated solely by numbers (e.g., some ecommerce sites use pagination for product catalogs that have more products than they wish to show on a single page), is problematic for many reasons.

First, pagination provides virtually no new topical relevance, as the pages are each largely about the same topic. Second, content that moves into different pagination can potentially create duplicate content problems or be seen as poor-quality or "thin" content. Last, pagination can create spider traps and hundreds or thousands of extraneous, low-quality pages that can be detrimental to search visibility.

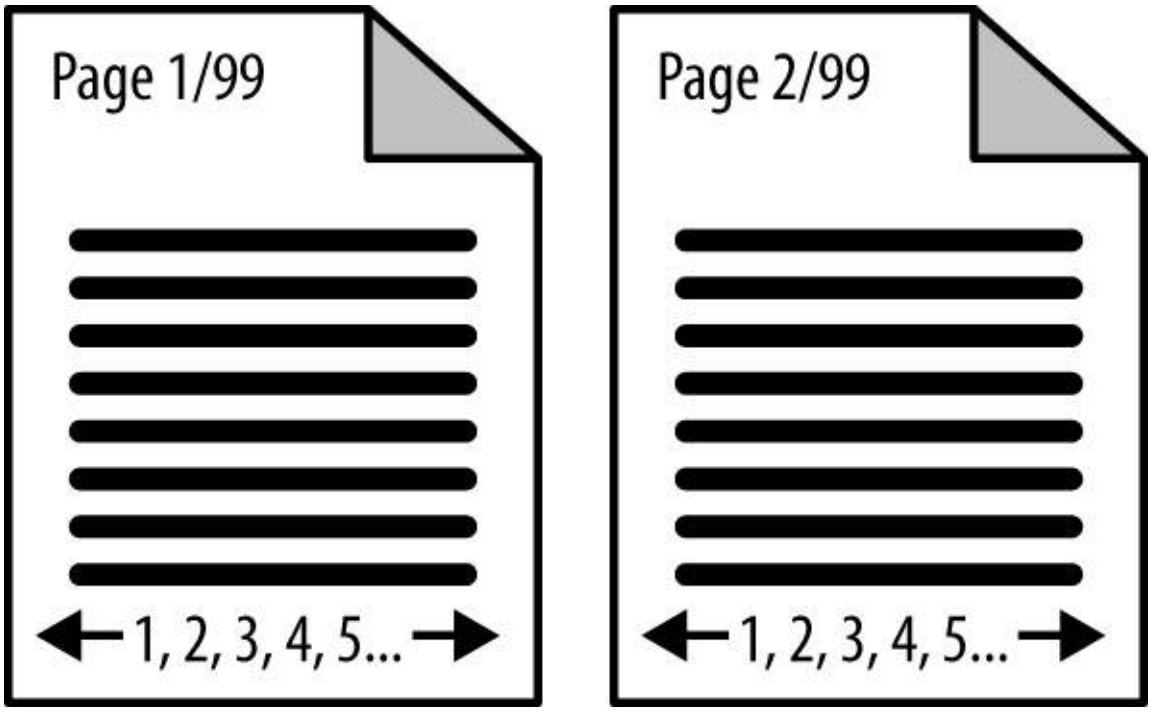


Figure 6-11. Pagination structures

So, make sure you implement flat structures and stay within sensible guidelines for the number of links per page, while retaining a contextually rich link structure. This is not always as easy as it sounds, and accomplishing it may require quite a bit of thought and planning to build a contextually rich structure on some sites. Consider a site with 10,000 different men's running shoes. Defining an optimal structure for that site could be a very large effort, but that effort will pay serious dividends in return.

Solutions to pagination problems vary based on the content of the website. Here are two example scenarios and their solutions:

Use `rel="next"` *and* `rel="prev"`

Google supports link elements called `rel="next"` and `rel="prev"`. The benefit of using these link elements is that it lets Google know when it has encountered a sequence of paginated pages. Once Google recognizes these tags, links to any of the pages will be treated as links to the series of pages as a whole. In addition, Google will show in the index the most relevant page in the series (most of the time this will be the first page, but not always).

Bing announced support for `rel="next"` and `rel="prev"` in 2012.

These tags can be used to inform Google about pagination structures, and they can be used whether or not you create a `view-all` page. The concept is simple. The following example outlines how to use the tags for content that is paginated into 12 pages:

- In the `<head>` section of the first page of your paginated content, implement a `rel="next"` tag pointing to the second page of the content. The tag should look something like this:

  ```
  <link rel="next"
    href="http://www.yoursite.com/products?prod=qwert&p=2" />
  ```

- In the `<head>` section of the last page of your paginated content, implement a `rel="prev"` link element pointing to the second-to-last page of the content. The tag should look something like this:

  ```
  <link rel="prev"
    href="http://www.yoursite.com/products?prod=qwert&p=11" />
  ```

  In the `<head>` section of pages 2 through 11, implement `rel="next"` and `rel="prev"` tags pointing to the following and preceding pages, respectively. The following example shows what it should look like on page 6 of the content:

  ```
  <link rel="prev"
    href="http://www.yoursite.com/products?prod=qwert&p=5" >
  <link rel="next"
    href="http:/www.yoursite.com/products?prod=qwert&p=7" />
  ```

Create a `view-all` *page and use canonical tags*

You may have lengthy articles that you choose to break into multiple pages. However, this results in links to the pages whose anchor text is something like `"1"`, `"2"`, and so forth. The titles of the various pages may not vary in any significant way, so they tend to compete with each other for search traffic. Finally, if someone links to the article but does not link to the first

page, the link authority from that link will largely be wasted.

One way to handle this problem is to retain the paginated version of the article, but also create a single-page version of the article. This is referred to as a `view-all` page. Then use the `rel="canonical"` link element (which is discussed in more detail in the section "Content Delivery and Search Spider Control") to point from the paginated pages to the `view-all` page. This will concentrate all of the link authority and search engine attention on a single page. You should also include a link to the `view-all` page from each of the individual paginated pages as well. However, if the `view-all` page loads too slowly because of the page size, it may not be the best option for you.

Note that if you implement a `view-all` page and do not implement any of these tags, Google will attempt to discover the page and show it instead of the paginated versions in its search results. However, we recommend that you make use of one of the aforementioned two solutions, as Google cannot guarantee that it will discover your `view-all` pages, and it is best to provide it with as many clues as possible.

Search-Friendly Site Navigation

Website navigation is something that web designers have been putting considerable thought and effort into since websites came into existence. Even before search engines were significant, navigation played an important role in helping users find what they wanted. It plays an important role in helping search engines understand your site as well.

Basics of search engine friendliness

The search engine spiders need to be able to read and interpret your website's code to properly spider and index the content on your web pages. Do not confuse this with the rules of organizations such as the W3C, which issues guidelines on HTML construction. Although following the W3C guidelines can be a good idea, the great majority of sites do not follow them, so search engines generally overlook violations of these rules as long as their spiders can parse the code.

Unfortunately, web page navigation and content can be rendered in many ways that function well for humans, but are invisible or at least challenging for search

engine spiders.

For example, there are numerous ways to incorporate content and navigation on the pages of a website. For the most part, all of these are designed for humans. Basic HTML text and HTML links such as those shown in Figure 6-12 work equally well for humans and search engine crawlers.

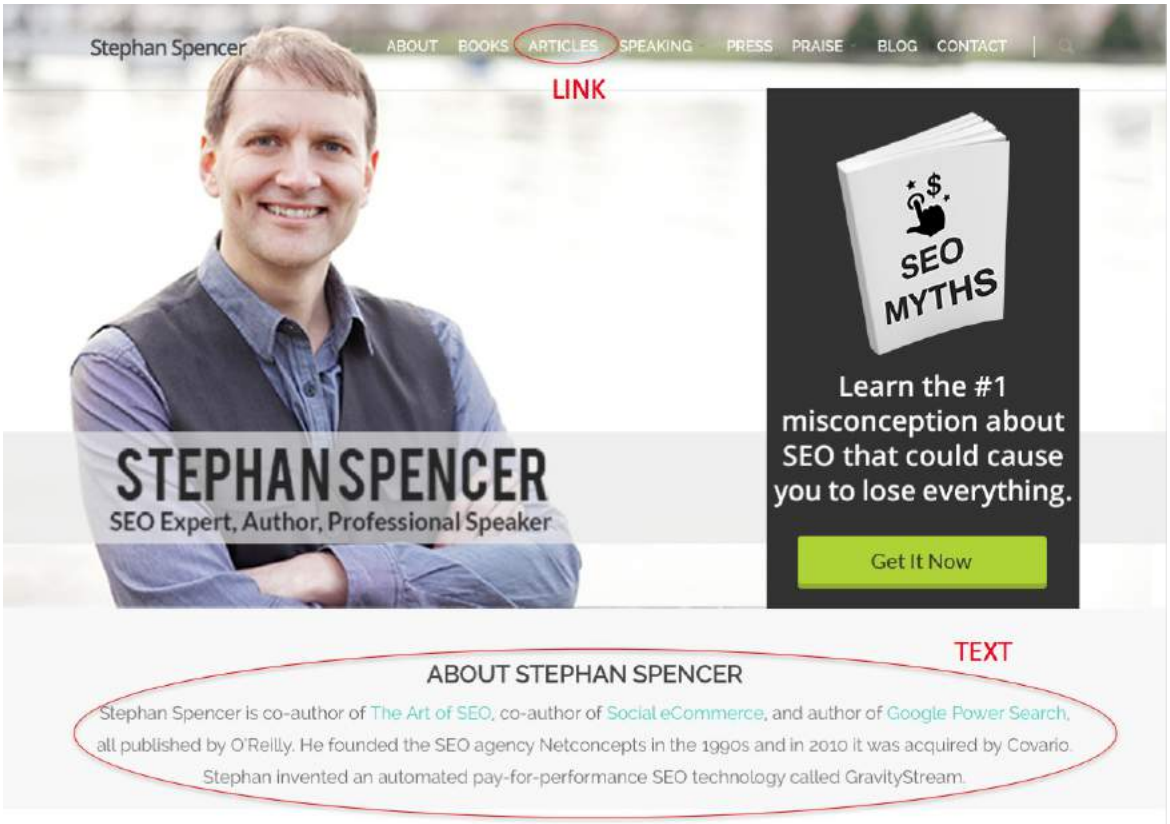


Figure 6-12. Example page with simple text and text link

The text and the link that are indicated on the page shown in Figure 6-12 are in simple HTML format.

Site elements that are problematic for spiders

However, many other types of content may appear on a web page and work well for humans but not so well for search engines. Here are some of the most common ones.

Search and web forms

Many sites incorporate search functionality. These “site search” elements are

specialized search engines that index and provide access to one site's content.

This is a popular method of helping users rapidly find their way around complex sites. For example, the Pew Internet website provides Site Search in the upper-right corner; this is a great tool for users, but search engines will be stymied by it. Search engines operate by crawling the Web's link structure—they don't in most circumstances submit forms or attempt random queries into search fields, and thus, any URLs or content solely accessible via a form will remain invisible to Google and Bing. In the case of Site Search tools, this is OK, as search engines do not want to index this type of content (they don't like to serve search results within their search results).

Forms are a popular way to provide interactivity, and one of the simplest applications is the "Contact us" form many websites have.

Unfortunately, crawlers will not fill out or submit such forms; thus, any content restricted to those who employ them is inaccessible to the engines. In the case of a "Contact us" form, this is likely to have little impact, but other types of forms can lead to bigger problems.

Websites that have content behind paywall and/or login barriers will need to either provide text links to the content behind the barrier (which defeats the purpose of the login) or implement First Click Free (discussed in "Content Delivery and Search Spider Control").

Java, images, audio, and video

Flash files, Java embeds, audio, and video (in any format) present content that is largely uncrawlable by the major engines. With some notable exceptions that we will discuss later, search engines can read text only when it is presented in HTML format. Embedding important keywords or entire paragraphs in an image or a Java console renders them invisible to the spiders. Likewise, search engines cannot easily read words spoken in an audio file or video. However, Google has begun to leverage tools such as Google Voice Search in order to "crawl" audio content and extract meaning (this was first confirmed in the book *In the Plex* by Steven Levy [Simon and Schuster]). Baidu already has an MP3 search function, and the Shazam and Jsaikoz applications have the ability to identify song hashes.

`alt` attributes, originally created as metadata for markup and an accessibility tag for vision-impaired users, are a good way to present at least some text content to

the engines when you are displaying images or embedded, nontext content. Note that the `alt` attribute is not a strong signal, and using it on an image link is no substitute for implementing a simple text link with appropriately descriptive anchor text. A good alternative is to employ captions and text descriptions in the HTML content wherever possible.

In the past few years, a number of companies offering transcription services have cropped up, providing automated text creation for the words spoken in audio or video. Providing these transcripts on rich media pages makes your content accessible to the search engines and findable by keyword-searching visitors. You can also use software such as Dragon Naturally Speaking and dictate your "transcript" to your computer.

AJAX and JavaScript

JavaScript enables many dynamic functions inside a website, most of which interfere very minimally with the operations of a search engine spider. The exception is when a page must use a JavaScript call to reach another page, or to pull content that the spiders can't see in the HTML. In some instances, this content is not visible to search engine spiders. However, Google will attempt to execute JavaScript to access this type of content.[3] Google's capabilities for accessing JavaScript have been improving over time, and you can expect that trend to continue.[4]

One example of Google reading JavaScript is Facebook Comments. Facebook Comments is a system offered by Facebook that allows publishers to collect comments from users on their site. Figure 6-13 shows an example of the Facebook Comments on a page on the TechCrunch site.

If you examine the source code for this particular post, you will not see any of the text strings for these comments in the HTML of the page. This is because the comments are actually stored on Facebook and dynamically retrieved by the web server when the page is rendered.

This is an example of the type of content that was not historically indexed by the search engines. When you use a JavaScript implementation like this, it is not clear what Google or Bing will be able to do with it. Facebook Comments is a broadly used system, and it makes sense for the search engines to learn how to attempt to read that content, but as of March 2012 this content was not indexed

by Google.

However, since then, this has changed. As of June 2015, this content is being indexed by Google and associated with the site hosting the Facebook Comments. You can test this (and whether your own content is indexed) by doing a Google search on a unique string of words, surrounded by double quotes to ensure Google searches only for those exact words in that exact order. For example, searching Google for one of the comments in Figure 6-13, *"As an ethnic Chinese, learning Mandarin and struggling, I'm extremely impressed"*, does return the URL *http://techcrunch.com/2014/10/23/zuckerberg-speaks-chinese-internet-soils-itself* as a result.

Figure 6-13. Facebook Comments on TechCrunch

While Google has recently indicated that it executes most JavaScript, it's still

possible that it doesn't execute *all* JavaScript. So, if your intent is to create content that you want the search engines to see, it is still safest to implement that content in a form that is directly visible in the HTML of the web page.

AJAX might present problems, most notably in the delivery of content that search engines may not be able to spider. Because AJAX uses database calls to retrieve data without refreshing a page or changing URLs, the content contained behind these technologies may be completely hidden from the search engines (see Figure 6-14).

In fact, in early 2015, Google indicated that it might move away from attempting to crawl any AJAX pages at all.[5] This was further confirmed in a June 2015 article by Eric Enge in which Google's Gary Illyes said: "If you have one URL only, and people have to click on stuff to see different sort orders or filters for the exact same content under that URL, then typically we would only see the default content."[6]

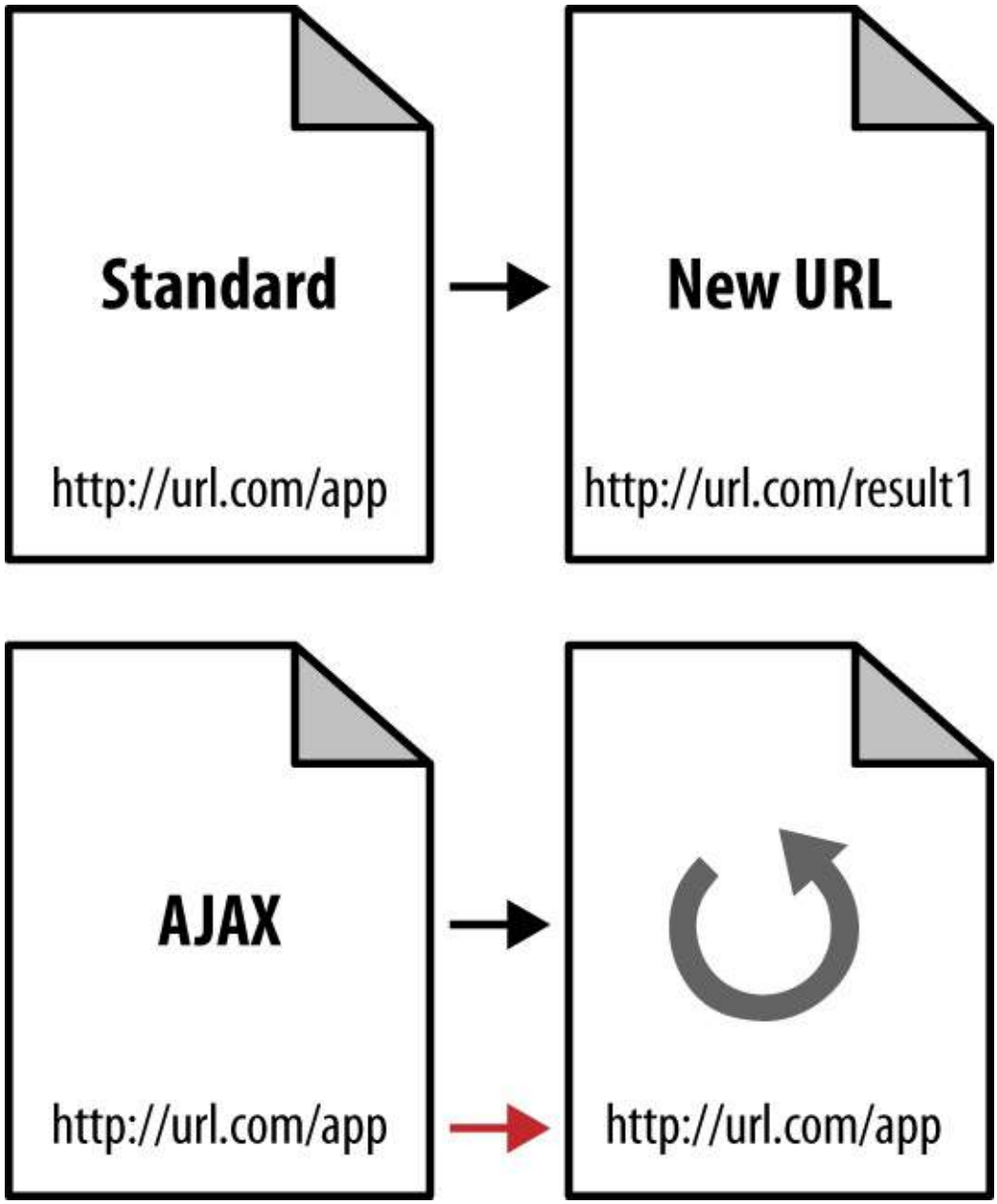


Figure 6-14. The problem with AJAX

As a result, if you’re using a traditional AJAX implementation, you may want to consider implementing an alternative spidering system for search engines to follow. AJAX applications are so user-friendly and appealing that forgoing them is simply impractical for many publishers. With these traditional implementations, building out a directory of links and pages that the engines can follow is a far better solution.

When you build these secondary structures of links and pages, make sure to provide users with access to them as well. Inside the AJAX application itself, give your visitors the option to “directly link to this page” and connect that URL with the URL you provide to search spiders through your link structures. AJAX apps not only suffer from content that can’t be crawled, but they also often don’t receive accurate links from users because the URL doesn’t change.

Some versions of AJAX use a **#** delimiter, which acts as a query string into the AJAX application. This allows you to link directly to different pages within the application. The **#**, which is used for HTML bookmarking, and everything beyond it are normally ignored by search engines.

This is largely because web browsers use only what’s after the **#** to jump to the anchor within the page, and that’s done locally within the browser. In other words, the browser doesn’t send the full URL, so the parameter information (i.e., any text after the **#**) is not passed back to the server.

In 2009, Google outlined a method for making these AJAX pages visible to search engines. This was later followed up with recommendations made on the Google Developers site. You can find more information at *http://bit.ly/ajax_crawling*.

The solution proposed by Google involves making some slight modifications to the way your AJAX URLs are formatted so that its crawler can recognize when an AJAX URL can be treated like a static page (one that will always return the same content), in which case Googlebot will read the page and treat it like any other static page for indexing and ranking purposes, affording it the same opportunity to rank as a page coded in plain HTML.

Other types of single-page application frameworks, such as Angular.js, Backbone.js, or Ember.js, may have similar problems. You can read more about how to deal with these in “Angular.js: Making it SEO-friendly”.

Frames

Frames emerged in the mid-1990s as a popular way to make easy navigation systems. Unfortunately, both their usability (in 99% of cases) and their search friendliness (in 99.99% of cases) were exceptionally poor. Today, iframes and CSS can replace the need for frames, even when a site’s demands call for similar functionality.

For search engines, the biggest problem with frames and iframes is that they often hold the content from two or more URLs on a single page. For users, because search engines, which direct searchers to only a single URL, may get confused by frames and direct visitors to single pages (*orphan* pages) inside a site intended to show multiple URLs at once. Indeed, the search engines consider the content within an iframe as residing on a separate page from the one the iframe is being used on. Thus, pages with nothing but iframed content will look virtually blank to the search engines.

Additionally, because search engines rely on links, and frame pages will often change content for users without changing the URL, external links often point to the wrong URL unintentionally. As a consequence, links to the page containing the frame or iframe may not point to the content the linker wanted to point to. Figure 6-15 illustrates how multiple pages are combined into a single URL with frames, which results in link distribution and spidering issues.

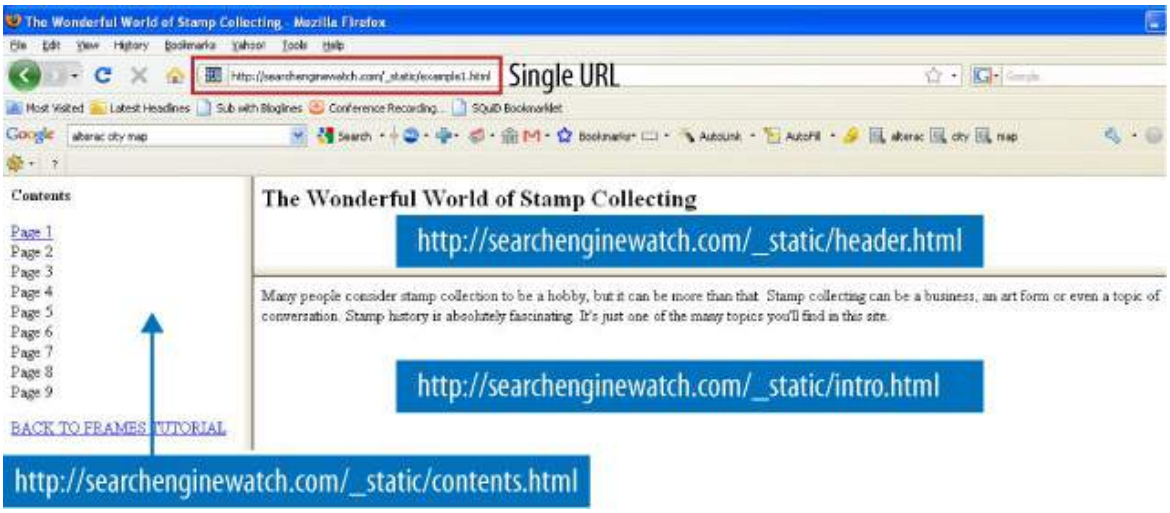


Figure 6-15. Sample page using frames

Search engine–friendly navigation guidelines

Although search engine spiders have become more advanced over the years, the basic premise and goals remain the same: spiders find web pages by following links and record the content of the pages they find in the search engine's index (a giant repository of data about websites and pages).

In addition to avoiding the techniques we just discussed, there are some

additional guidelines for developing search engine–friendly navigation:

Implement a text-link-based navigational structure

If you choose to create navigation in Flash, JavaScript, or other technologies that the search engine may not be able to parse, make sure to offer alternative text links in HTML for spiders to ensure that automated robots (and visitors who may not have the required browser plugins) can reach your pages.

Beware of "spider traps"

Even intelligently coded search engine spiders can get lost in infinite loops of links that pass between pages on a site. Intelligent architecture that avoids recursively looping 301 or 302 HTTP server codes (or other redirection protocols) should negate this issue, but sometimes online calendar links, infinite pagination that loops, or content being accessed or sorted in a multitude of ways (faceted navigation) can create tens of thousands of pages for search engine spiders when you intended to have only a few dozen true pages of content. You can read more about Google's viewpoint on this at *http://googlewebmastercentral.blogspot.com/2008/08/to-infinity-and-beyond-no.html*.

Watch out for session IDs and cookies

As we just discussed, if you limit a user's ability to view pages or redirect based on a cookie setting or session ID, search engines may be unable to crawl your content. The bots do not have cookies enabled, nor can they deal with session IDs properly (each visit by the crawler gets a URL with a different session ID, and the search engine sees these URLs with session IDs as different URLs). Although restricting form submissions is fine (as search spiders can't submit forms anyway), limiting content access via cookies and session IDs is a bad idea.

Be mindful of server, hosting, and IP issues

Server issues rarely cause search engine ranking problems—but when they do, disastrous consequences can follow. The engines are acutely aware of common server problems, such as downtime or overloading, and will give you the benefit of the doubt (though this will mean your content cannot be

spidered during periods of server dysfunction). On the flip side, sites hosted on content delivery networks (CDNs) may get crawled more heavily, and CDNs offer significant performance enhancements to a website.

The IP address of your host can be of concern in some instances. IPs once belonging to sites that have spammed the search engines may carry with them negative associations that can hinder spidering and ranking. While the engines aren't especially picky about shared hosting versus dedicated servers and dedicated IP addresses, or about server platforms, you can avoid many hassles by going these routes. At the very minimum, you should be cautious and find a host you trust, and inquire into the history and "cleanliness" of the IP address you may be assigned, as the search engines have become paranoid about the use of certain domains, hosts, IP addresses, and blocks of IPs. Experience tells them that many of these have strong correlations with spam, and thus, removing them from the index can have great benefits for users. As a site owner *not* engaging in these practices, you'll find it pays to investigate your web host prior to getting into trouble.

You can read more about server and hosting issues in "Identifying Current Server Statistics Software and Gaining Access".

Root Domains, Subdomains, and Microsites

Among the common questions about structuring a website (or restructuring one) are whether to host content on a new domain, when to use subfolders, and when to employ microsites.

As search engines scour the Web, they identify four kinds of web structures on which to place metrics:

Individual pages/URLs

These are the most basic elements of the Web—filenames, much like those that have been found on computers for decades, which indicate unique documents. Search engines assign query-independent scores—most famously, Google's PageRank—to URLs and judge them in their ranking algorithms. A typical URL might look something like *http://www.yourdomain.com/page*.

Subfolders

The folder structures that websites use can also inherit or be assigned metrics by search engines (though there's very little information to suggest that they are used one way or another). Luckily, they are an easy structure to understand. In the URL *http://www.yourdomain.comblogpost17*, *blog* is the subfolder and *post17* is the name of the file in that subfolder. Engines may identify common features of documents in a given subfolder and assign metrics to these (such as how frequently the content changes, how important these documents are in general, or how unique the content is that exists in these subfolders).

Subdomains/fully qualified domains (FQDs)/third-level domains

In the URL *http://blog.yourdomain.com/page*, three kinds of domain levels are present. The top-level domain (also called the *TLD* or *domain extension*) is *.com*, the second-level domain is *yourdomain*, and the third-level domain is *blog*. The third-level domain is sometimes referred to as a *subdomain*. Common web nomenclature does not typically apply the word *subdomain* when referring to *www*, although technically, this too is a subdomain. A fully qualified domain is the combination of the elements required to identify the location of the server where the content can be found (in this example, *blog.yourdomain.com/*).

These structures can receive individual assignments of importance, trustworthiness, and value from the engines, independent of their second-level domains, particularly on hosted publishing platforms such as WordPress, Blogspot, and so on.

Complete root domains/host domain/pay-level domains (PLDs)/second-level domains

The domain name you need to register and pay for, and the one you point DNS settings toward, is the second-level domain (though it is commonly improperly called the "top-level" domain). In the URL *http://www.yourdomain.com/page*, *yourdomain.com* is the second-level domain. Other naming conventions may refer to this as the "root" or "pay-

level” domain.

Figure 6-16 shows some examples.

Figure 6-16. Breaking down some example URLs

When to Use a Subfolder

If a subfolder will work, it is the best choice 99.9% of the time. Keeping content

on a single root domain and single subdomain (e.g., *http://www.yourdomain.com*) gives the maximum SEO benefits, as engines will maintain all of the positive metrics the site earns around links, authority, and trust, and will apply these to every page on the site.

Subfolders have all the flexibility of subdomains (the content *can*, if necessary, be hosted on a unique server or completely unique IP address through post-firewall load balancing) and none of the drawbacks. Subfolder content will contribute directly to how search engines (and users, for that matter) view the domain as a whole. Subfolders can be registered with the major search engine tools and geotargeted individually to specific countries and languages as well.

Although subdomains are a popular choice for hosting content, they are generally not recommended if SEO is a primary concern. Subdomains *may* inherit the ranking benefits and positive metrics of the root domain they are hosted underneath, but they do not always do so (and thus, content can underperform in these scenarios). Of course, there can be exceptions to this general guideline. Subdomains are not inherently harmful, and there are some content publishing scenarios in which they are more appropriate than subfolders; it is simply preferable for various SEO reasons to use subfolders when possible, as we will discuss next.

When to Use a Subdomain

If your marketing team decides to promote a URL that is completely unique in content or purpose and would like to use a catchy subdomain to do it, using a subdomain can be practical. Google Maps is an example that illustrates how marketing considerations make a subdomain an acceptable choice. One good reason to use a subdomain is in a situation in which, as a result of creating separation from the main domain, using one looks more authoritative to users.

Subdomains may also be a reasonable choice if keyword usage in the domain name is of critical importance. It appears that search engines do weight keyword usage in the URL somewhat, and have slightly higher benefits for exact matches in the subdomain (or third-level domain name) than subfolders. Note that exact matches in the domain and subdomain carry less weight than they once did. Google updated the weight it assigned to these factors in 2012.[7]

Keep in mind that subdomains may inherit very little link equity from the main

domain. If you wish to split your site in the subdomains and have all of them rank well, assume that you will have to support each with its own full-fledged SEO strategy.

When to Use a Separate Root Domain

If you have a single, primary site that has earned links, built content, and attracted brand attention and awareness, it is very rarely advisable to place any new content on a completely separate domain. There are rare occasions when this can make sense, and we'll walk through these, as well as explain how singular sites benefit from collecting all of their content in one root domain location.

Splitting similar or relevant content from your organization onto multiple domains can be likened to a store taking American Express Gold cards and rejecting American Express Corporate or American Express Blue—it is overly segmented and dangerous for the consumer mindset. If you can serve web content from a singular domain, that domain will earn branding in the minds of your visitors, references from them, links from other sites, and bookmarks from your regular customers. Switching to a new domain forces you to rebrand and to earn all of these positive metrics all over again.

Microsites

Although we generally recommend that you do not saddle yourself with the hassle of dealing with multiple sites and their SEO risks and disadvantages, it is important to understand the arguments, if only a few, in favor of doing so.

Optimized properly, a microsite may have dozens or even hundreds of pages. If your site is likely to gain more traction and interest with webmasters and bloggers by being at arm's length from your main site, it may be worth considering—for example, if you have a very commercial main site, and you want to create some great content (perhaps as articles, podcasts, and RSS feeds) that does not fit on the main site.

When should you consider a microsite?

When you own a specific keyword search query domain

For example, if you own *usedtoyotatrucks.com*, you might do very well to

pull in search traffic for the specific term *used toyota trucks* with a microsite.

When you plan to sell the domains

It is very hard to sell a folder or even a subdomain, so this strategy is understandable if you're planning to churn the domains in the secondhand market.

As discussed earlier, if you're a major brand building a "secret" or buzzworthy microsite

In this case, it can be useful to use a separate domain (however, you should 301-redirect the pages of that domain back to your main site after the campaign is over so that the link authority continues to provide long-term benefit—just as the mindshare and branding do in the offline world).

You should never implement a microsite that acts as a doorway page to your main site, or that has substantially the same content as your main site. Consider a microsite only if you are willing to invest the time and effort to put rich original content on it, and to promote it as an independent site.

Such a site may gain more links by being separated from the main commercial site. A microsite may have the added benefit of bypassing some of the legal and PR department hurdles and internal political battles.

However, a microsite on a brand-new domain can take many months to build enough domain-level link authority to rank in the engines (for more about how Google treats new domains, see "Determining Searcher Intent and Delivering Relevant, Fresh Content"). So, what to do if you want to launch a microsite? Start the clock running as soon as possible on your new domain by posting at least a few pages to the URL and then getting at least a few links to it—as far in advance of the official launch as possible. It may take a considerable amount of time before a microsite is able to house enough high-quality content and to earn enough trusted and authoritative links to rank on its own. If the campaign the microsite was created for is time sensitive, consider redirecting the pages from the microsite to your main site well after the campaign concludes, or at least ensure that the microsite links back to the main site to allow some of the link authority the microsite earns to help the ranking of your main site.

Here are the reasons for not using a microsite:

Search algorithms favor large, authoritative domains

Take a piece of great content about a topic and toss it onto a small, mom-and-pop website; point some external links to it, optimize the page and the site for the target terms, and get it indexed. Now, take that exact same content and place it on Wikipedia, or CNN.com, and you're virtually guaranteed that the content on the large, authoritative domain will outrank the content on the small niche site. The engines' current algorithms favor sites that have built trust, authority, consistency, and history.

Multiple sites split the benefits of links

As suggested in Figure 6-17, a single good link pointing to a page on a domain positively influences the entire domain and every page on it. Because of this phenomenon, it is much more valuable to have any link you can possibly get pointing to the same domain to help boost the rank and value of the pages on it. Having content or keyword-targeted pages on other domains that don't benefit from the links you earn to your primary domain only creates more work.

100 links to Domain A ≠ 100 links to Domain B + 1 link to Domain A (from Domain B)

In Figure 6-18, you can see how earning lots of links to Page G on a separate domain is far less valuable than earning those same links to a page on the primary domain. For this reason, even if you interlink all of the microsites or multiple domains that you build, the value still won't be close to what you could get from those links if they pointed directly to the primary domain.

A large, authoritative domain can host a huge variety of content

Niche websites frequently limit the variety of their discourse and content matter, whereas broader sites can target a wider range of foci. This is valuable not just for targeting the long tail of search and increasing potential branding and reach, but also for viral content, where a broader focus is much less limiting than a niche focus.

Time and energy are better spent on a single property

If you're going to pour your heart and soul into web development, design, usability, user experience, site architecture, SEO, public relations, branding, and so on, you want the biggest bang for your buck. Splitting your attention, time, and resources on multiple domains dilutes that value and doesn't let you build on your past successes on a single domain. As shown in Figure 6-18, every page on a site benefits from inbound links to the site. The page receiving the link gets the most benefit, but other pages also benefit.

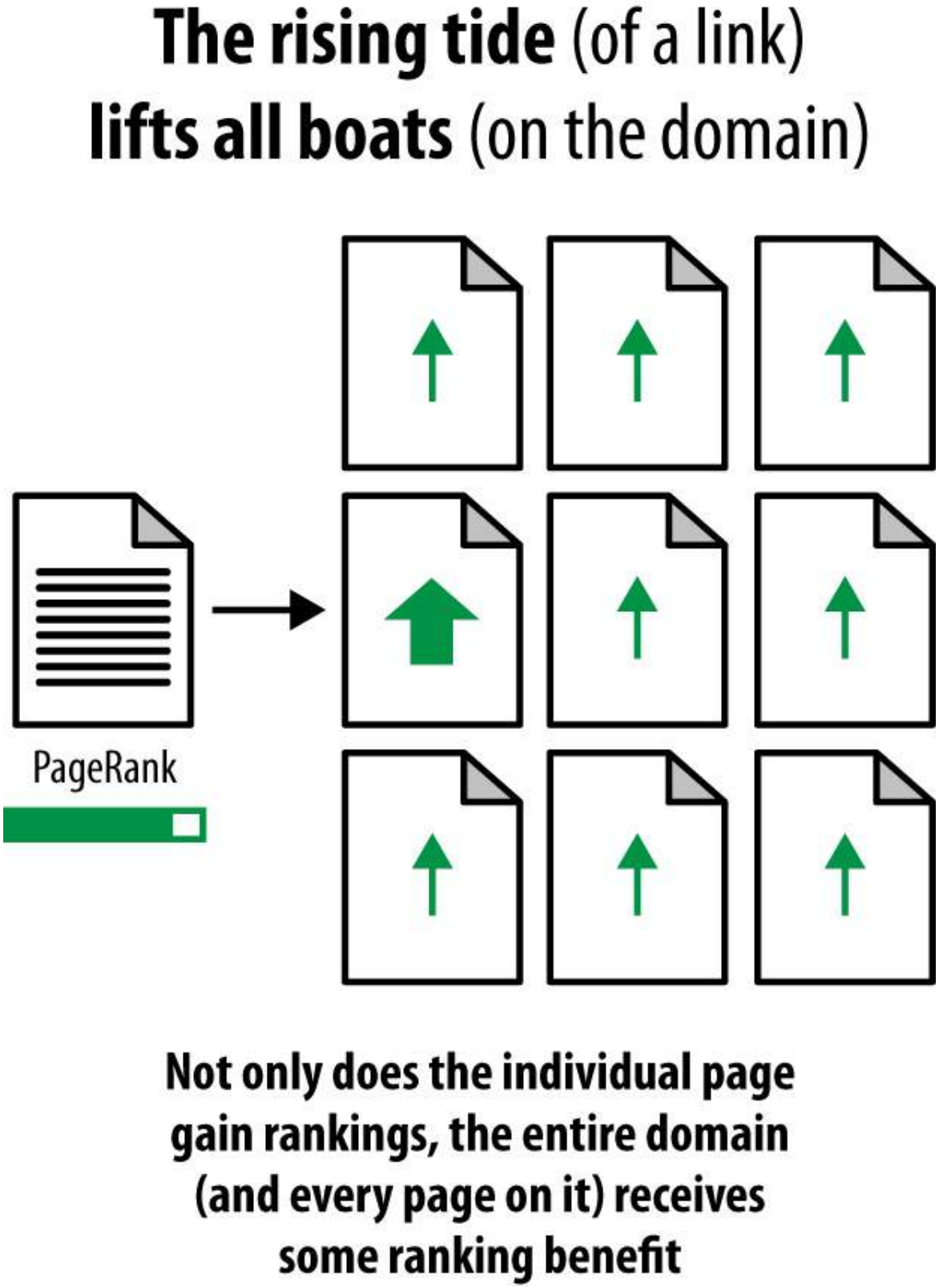


Figure 6-17. How links can benefit your whole site

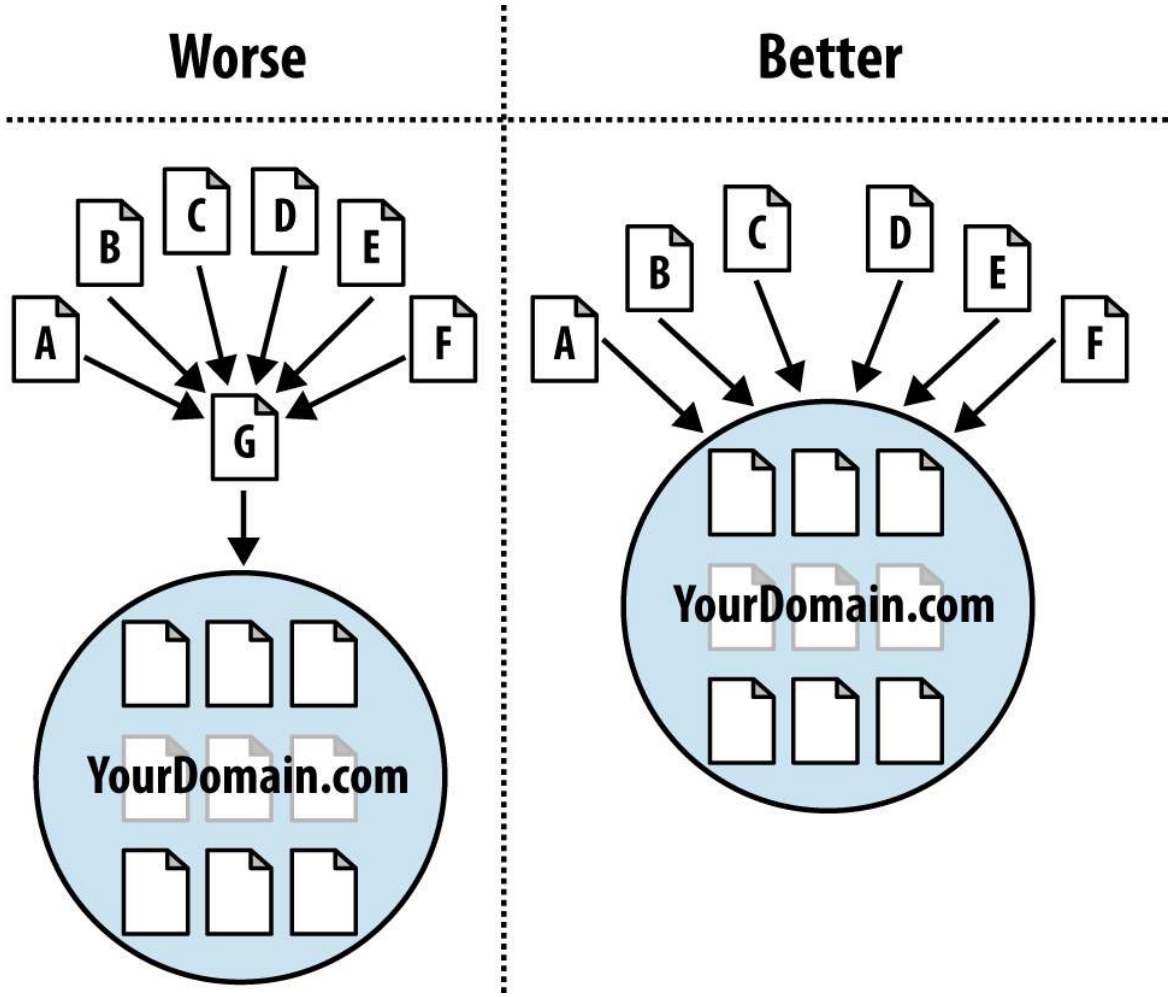


Figure 6-18. How direct links to your domain are better

When to Use a TLD Other Than .com

There are only a few situations in which you should consider using a TLD other than *.com*:

- When you own the *.com* and want to redirect to an *.org*, *.tv*, *.biz*, and so on, possibly for marketing/branding/geographic reasons. Do this only if you already own the *.com* and can redirect.
- When you can use a *.gov*, *.mil*, or *.edu* domain.
- When you are serving only a single geographic region and are willing to permanently forgo growth outside that region (e.g., *.co.uk*, *.de*, *.it*, etc.).

- When you are a nonprofit and want to distance your organization from the commercial world, *.org* may be for you.

New gTLDs

Many website owners have questions about the new gTLDs (generic top-level domains) that ICANN started assigning in the fall of 2013. Instead of the traditional *.com*, *.net*, *.org*, *.ca*, and so on with which most people are familiar, these new gTLDs range from *.christmas* to *.autos* to *.lawyer* to *.eat* to *.sydney*. A full list of them can be found at *http://newgtlds.icann.org/en/program-status/delegated-strings*. One of the major questions that arises is "Will these help me rank organically on terms related to the TLD?" Currently the answer is no. There is no inherent SEO value in having a TLD that is related to your keywords. Having a *.storage* domain does not mean you have some edge over a *.com* for a storage-related business. In an online forum, Google's John Mueller stated that these TLDs are treated the same as other generic-level TLDs in that they do not help your organic rankings. He also noted that even the new TLDs that sound as if they are region-specific in fact give you no specific ranking benefit in those regions, though he added that Google reserves the right to change that in the future.

Despite the fact that they do not give you a ranking benefit currently, you should still grab your domains for key variants of the new TLDs. You may wish to consider ones such as *.spam*. You may also wish to register those that relate directly to your business. It is unlikely that these TLDs will give a search benefit in the future, but it *is* likely that if your competition registers your name in conjunction with one of these new TLDs your users might be confused about which is the legitimate site. For example, if you are located in New York City, you should probably purchase your domain name with the *.nyc* TLD; if you happen to own a pizza restaurant, you may want to purchase *.pizza*; and so on.

Optimization of Domain Names/URLs

Two of the most basic parts of any website are the domain name and the URLs for the pages of the website. This section will explore guidelines for optimizing these important elements.

Optimizing Domains

When you're conceiving or designing a new site, one of the critical items to consider is the domain name, whether it is for a new blog, a company launch, or even just a friend's website. Here are 12 indispensable tips for selecting a great domain name:

Brainstorm five top keywords

When you begin your domain name search, it helps to have five terms or phrases in mind that best describe the domain you're seeking. Once you have this list, you can start to pair them or add prefixes and suffixes to create good domain ideas. For example, if you're launching a mortgage-related domain, you might start with words such as *mortgage*, *finance*, *home equity*, *interest rate*, and *house payment*, and then play around until you can find a good match.

Make the domain unique

Having your website confused with a popular site that someone else already owns is a recipe for disaster. Thus, never choose a domain that is simply the plural, hyphenated, or misspelled version of an already established domain. For example, for years Flickr did not own *http://flicker.com*, and the company probably lost traffic because of that. It recognized the problem and bought the domain, and as a result *http://flicker.com* now redirects to *http://flickr.com*.

Choose only dot-com-available domains

If you're not concerned with type-in traffic, branding, or name recognition, you don't need to worry about this one. However, if you're at all serious about building a successful website over the long term, you should be worried about all of these elements, and although directing traffic to a *.net* or *.org* (or any of the other new gTLDs) is fine, owning and 301-ing the *.com*, or the ccTLD for the country your website serves (e.g., *.co.uk* for the United Kingdom), is critical. With the exception of the very tech-savvy, most people who use the Web still make the automatic assumption that *.com* is all that's out there, or that it's more trustworthy. Don't make the mistake of locking out or losing traffic from these folks.

Make it easy to type

If a domain name requires considerable attention to type correctly due to spelling, length, or the use of unmemorable words or sounds, you've lost a good portion of your branding and marketing value. Usability folks even tout the value of having the words include easy-to-type letters (which we interpret as avoiding *q*, *z*, *x*, *c*, and *p*).

Make it easy to remember

Remember that word-of-mouth marketing relies on the ease with which the domain can be called to mind. You don't want to be the company with the terrific website that no one can ever remember to tell their friends about because they can't remember the domain name.

Keep the name as short as possible

Short names are easy to type and easy to remember (see the previous two rules). Short names also allow more of the URL to display in the SERPs and are a better fit on business cards and other offline media.

Create and fulfill expectations

When someone hears about your domain name for the first time, he should be able to instantly and accurately guess the type of content he might find there. That's why we love domain names such as NYTimes.com, CareerBuilder.com, AutoTrader.com, and WebMD.com. Domains such as Monster.com, Amazon.com, and Zillow.com required far more branding because of their nonintuitive names.

Avoid trademark infringement

This is a mistake that isn't made too often, but it can kill a great domain and a great company when it does. To be sure you're not infringing on anyone's registered trademark with your site's name, visit the U.S. Patent and Trademark office site and search before you buy. Knowingly purchasing a domain with bad-faith intent that includes a trademarked term is a form of cybersquatting referred to as *domain squatting*.

Set yourself apart with a brand

Using a unique moniker is a great way to build additional value with your domain name. A "brand" is more than just a combination of words, which is why names such as Mortgageforyourhome.com and Shoesandboots.com aren't as compelling as branded names such as Yelp and Gilt.

Reject hyphens and numbers

Both hyphens and numbers make it hard to convey your domain name verbally and fall down on being easy to remember or type. Avoid spelled-out or Roman numerals in domains, as both can be confusing and mistaken for the other.

Don't follow the latest trends

Website names that rely on odd misspellings, multiple hyphens (such as the SEO-optimized domains of the early 2000s), or uninspiring short adjectives (such as "top *x*," "best *x*," and "hot *x*") aren't always the best choice. This isn't a hard-and-fast rule, but in the world of naming conventions in general, if everyone else is doing it, that doesn't mean it is a surefire strategy. Just look at all the people who named their businesses "AAA *x*" over the past 50 years to be first in the phone book; how many Fortune 1000s are named "AAA *Company*?"

Use a domain selection tool

Websites such as Nameboy make it exceptionally easy to determine the availability of a domain name. Just remember that you don't have to buy through these services. You can find an available name that you like, and then go to your registrar of choice. You can also try BuyDomains as an option to attempt to purchase domains that have already been registered.

Picking the Right URLs

Search engines place some weight on keywords in your URLs. Be careful, however, as the search engines can interpret long URLs with numerous hyphens in them (e.g., *Buy-this-awesome-product-now.html*) as a spam signal. The

following are some guidelines for selecting optimal URLs for the pages of your site(s):

Describe your content

An obvious URL is a great URL. If a user can look at the address bar (or a pasted link) and make an accurate guess about the content of the page before ever reaching it, you've done your job. These URLs get pasted, shared, emailed, written down, and yes, even recognized by the engines.

Keep it short

Brevity is a virtue. The shorter the URL, the easier it is to copy and paste, read over the phone, write on a business card, or use in a hundred other unorthodox fashions, all of which spell better usability and increased branding. Remember, however, that you can always create a shortened URL for marketing purposes that redirects to the destination URL of your content—just know that this short URL will have no SEO value.

Static is the way

Search engines treat static URLs differently than dynamic ones. Users also are not fond of URLs in which the big players are ?, &, and =. They are just harder to read and understand.

Descriptive text is better than numbers

If you're thinking of using *114/cat223/*, you should go with *brandadidas/* instead. Even if the descriptive text isn't a keyword or is not particularly informative to an uninitiated user, it is far better to use words when possible. If nothing else, your team members will thank you for making it that much easier to identify problems in development and testing.

Keywords never hurt

If you know you're going to be targeting a lot of competitive keyword phrases on your website for search traffic, you'll want every advantage you can get. Keywords are certainly one element of that strategy, so take the list from marketing, map it to the proper pages, and get to work. For dynamically

created pages through a CMS, create the option of including keywords in the URL.

Subdomains aren't always the answer

First off, never use multiple subdomains (e.g., *product.brand.site.com*); they are unnecessarily complex and lengthy. Second, consider that subdomains have the potential to be treated separately from the primary domain when it comes to passing link and trust value. In most cases where just a few subdomains are used and there's good interlinking, it won't hurt, but be aware of the downsides. For more on this, and for a discussion of when to use subdomains, see "Root Domains, Subdomains, and Microsites".

Fewer folders

A URL should contain no unnecessary folders (or words or characters, for that matter). They do not add to the user experience of the site and can in fact confuse users.

Hyphens separate best

When creating URLs with multiple words in the format of a phrase, hyphens are best to separate the terms (e.g., *brandsdolce-and-gabbana/*), but you can also use plus signs (+).

Stick with conventions

If your site uses a single format throughout, don't consider making one section unique. Stick to your URL guidelines once they are established so that your users (and future site developers) will have a clear idea of how content is organized into folders and pages. This can apply globally as well as for sites that share platforms, brands, and so on.

Don't be case-sensitive

URLs can accept both uppercase and lowercase characters, so don't ever, ever allow any uppercase letters in your structure. Unix/Linux-based web servers are case-sensitive, so *http://www.domain.com/Products/widgets/* is

technically a different URL from *http://www.domain.comproductswidgets/*. Note that this is not true in Microsoft IIS servers, but there are a lot of Apache web servers out there. In addition, this is confusing to users, and potentially to search engine spiders as well. Google sees any URLs with even a single unique character as unique URLs. So if your site shows the same content on *www.domain.com/Products/widgets/* and *www.domain.comproductswidgets/*, it could be seen as duplicate content. If you have such URLs now, implement a 301-redirect pointing them to all-lowercase versions, to help avoid confusion. If you have a lot of type-in traffic, you might even consider a 301 rule that sends any incorrect capitalization permutation to its rightful home.

Don't append extraneous data

There is no point in having a URL exist in which removing characters generates the same content. You can be virtually assured that people on the Web will figure it out; link to you in different fashions; confuse themselves, their readers, and the search engines (with duplicate content issues); and then complain about it.

Mobile Friendliness

On April 21, 2015, Google rolled out an update designed to treat the mobile friendliness of a site as a ranking factor. What made this update unique is that it impacted rankings only for people searching from smartphones.

The reason for this update was that the user experience on a smartphone is dramatically different than it is on a tablet or a laptop/desktop device. The main differences are:

- Screen sizes are smaller, so the available space for providing a web page is significantly different.
- There is no mouse available, so users generally use their fingers to tap the screen to select menu items. As a result, more space is needed between links on the screen to make them "tappable."
- The connection bandwidth is lower, so web pages load more slowly. While

having smaller-size web pages helps them load on any device more quickly, this becomes even more important on a smartphone.

To help publishers determine the mobile friendliness of their sites, Google released a tool called the Mobile-Friendly Test. In theory, passing this test means that your page is considered mobile-friendly, and therefore would not be negatively impacted for its rankings on smartphones.

There was a lot of debate on the impact of the update. Prior to its release, the industry referred to it as “Mobilegeddon,” but in fact the scope of the update was not nearly that dramatic.

Coauthor Eric Enge led a study to measure the impact of the mobile friendliness update by comparing rankings prior to the update to those after it. This study found that nearly 50% of nonmobile-friendly URLs lost rank. You can see more details from the study at *http://bit.ly/enge_mobilegeddon*.

Keyword Targeting

Search engines face a tough task: based on a few words in a query (sometimes only one) they must return a list of relevant results ordered by measures of importance, and hope that the searcher finds what she is seeking. As website creators and web content publishers, you can make this process massively simpler for the search engines and, in turn, benefit from the enormous traffic they send, based on how you structure your content. The first step in this process is to research what keywords people use when searching for businesses that offer products and services like yours.

This practice has long been a critical part of search engine optimization, and although the role keywords play has evolved over time, keyword usage is still one of the first steps in targeting search traffic.

The first step in the keyword targeting process is uncovering popular terms and phrases that searchers regularly use to find the content, products, or services your site offers. There’s an art and science to this process, but it consistently begins with a list of keywords to target (see Chapter 5 for more on this topic).

Once you have that list, you’ll need to include these keywords in your pages. In the early days of SEO, the process involved stuffing keywords repetitively into

every HTML tag possible. Now, keyword relevance is much more aligned with the usability of a page from a human perspective.

Because links and other factors make up a significant portion of the search engines' algorithms, they no longer rank pages with 61 instances of *free credit report* above pages that contain only 60. In fact, *keyword stuffing*, as it is known in the SEO world, can actually get your pages devalued via search engine penalties. The engines don't like to be manipulated, and they recognize keyword stuffing as a disingenuous tactic. Figure 6-19 shows an example of a page utilizing accurate keyword targeting.

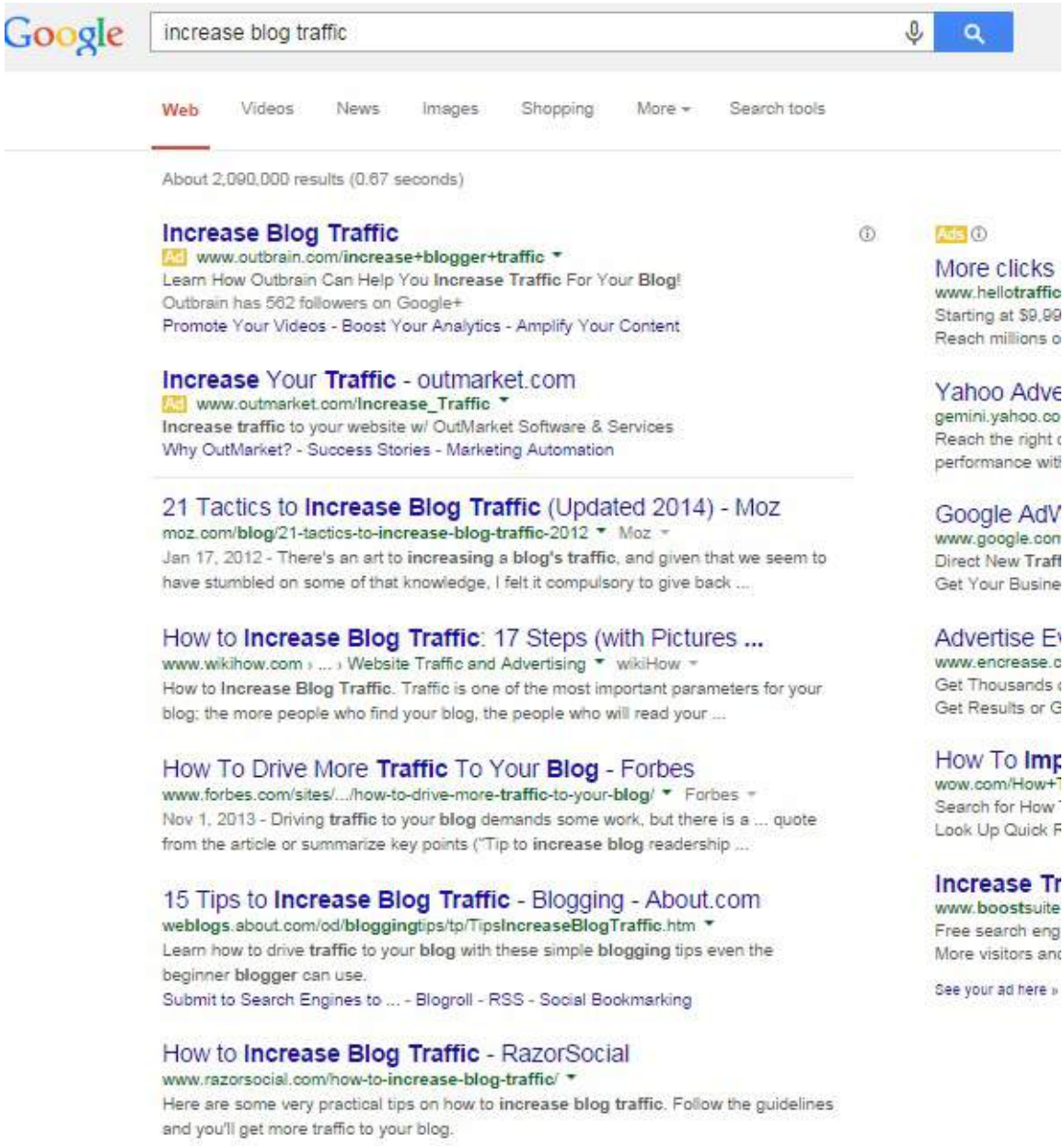


Figure 6-19. Title and headings tags—powerful for SEO

Keyword usage includes creating titles, headlines, and content designed to appeal to searchers in the results (and entice clicks), as well as building relevance for search engines to improve your rankings. In today's SEO, there are also many other factors involved in ranking, including term frequency—inverse document frequency (TF-IDF), co-occurrence, entity salience, page segmentation, and several others, which will be described in detail later in this chapter.

However, keywords remain important, and building a search-friendly site requires that you prominently employ the keywords that searchers use to find content. Here are some of the more prominent places where a publisher can place those keywords.

HTML <title> Tags

For keyword placement, `<title>` tags are an important element for search engine relevance. The `<title>` tag is in the `<head>` section of an HTML document, and is the only piece of meta information about a page that directly influences relevancy and ranking.

The following nine rules represent best practices for `<title>` tag construction. Do keep in mind, however, that a `<title>` tag for any given page must directly correspond to that page's content. You may have five different keyword categories and a unique site page (or section) dedicated to each, so be sure to align a page's `<title>` tag content with its actual visible content as well.

Place your keywords at the beginning of the `<title>` *tag*

> This positioning provides the most search engine benefit; thus, if you want to employ your brand name in the `<title>` tag as well, place it at the end. There is a trade-off here, however, between SEO benefit and branding benefit that you should think about: major brands may want to place their brand at the start of the `<title>` tag, as it may increase click-through rates. To decide which way to go, you need to consider which need is greater for your business.

Limit length to 50 characters (including spaces)

> Content in `<title>` tags after 50 characters is probably given less weight by the search engines. In addition, the display of your `<title>` tag in the SERPs may get cut off as early as 49 characters.[8]
>
> There is no hard-and-fast rule for how many characters Google will display. Google now truncates the display after a certain number of pixels, so the exact characters you use may vary in width. At the time of this writing, this width varies from 482px to 552px depending on your operating system and

platform.[9]

Also be aware that Google may not use your `<title>` tag in the SERPs. Google frequently chooses to modify your `<title>` tag based on several different factors that are beyond your control. If this is happening to you, it may be an indication that Google thinks that your `<title>` tag does not accurately reflect the contents of the page, and you should probably consider updating either your `<title>` tags or your content.

Incorporate keyword phrases

This one may seem obvious, but it is critical to prominently include in your `<title>` tag the keywords your research shows as being the most valuable for capturing searches.

Target longer phrases if they are relevant

When choosing what keywords to include in a `<title>` tag, use as many as are completely relevant to the page at hand while remaining accurate and descriptive. Thus, it can be much more valuable to have a `<title>` tag such as "SkiDudes | Downhill Skiing Equipment & Accessories" rather than simply "SkiDudes | Skiing Equipment." Including those additional terms that are both relevant to the page and receive significant search traffic can bolster your page's value.

However, if you have separate landing pages for "skiing accessories" versus "skiing equipment," don't include one term in the other's title. You'll be cannibalizing your rankings by forcing the engines to choose which page on your site is more relevant for that phrase, and they might get it wrong. We will discuss the cannibalization issue in more detail shortly.

Use a divider

When you're splitting up the brand from the descriptive text, options include | (a.k.a. the pipe), >, -, and :, all of which work well. You can also combine these where appropriate—for example, "Major Brand Name: Product Category – Product." These characters do not bring an SEO benefit, but they

can enhance the readability of your title.

Focus on click-through and conversion rates

The `<title>` tag is exceptionally similar to the title you might write for paid search ads, only it is harder to measure and improve because the stats aren't provided for you as easily. However, if you target a market that is relatively stable in search volume week to week, you can do some testing with your `<title>` tags and improve the click-through rate.

Watch your analytics and, if it makes sense, buy search ads on the page to test click-through and conversion rates of different ad text as well, even if it is for just a week or two. You can then look at those results and incorporate them into your titles, which can make a huge difference in the long run. A word of warning, though: don't focus entirely on click-through rates. Remember to continue measuring conversion rates.

Target searcher intent

When writing titles for web pages, keep in mind the search terms your audience employed to reach your site. If the intent is browsing or research-based, a more descriptive `<title>` tag is appropriate. If you're reasonably sure the intent is a purchase, download, or other action, make it clear in your title that this function can be performed at your site. Here is an example from *http://www.bestbuy.com/site/video-games/playstation-4-ps4/pcmcat295700050012.c?id=pcmcat295700050012*. The `<title>` tag of that page is "PS4: PlayStation 4 Games & Consoles - Best Buy." The `<title>` tag here makes it clear that you can buy PS4 games and consoles at Best Buy.

Communicate with human readers

This needs to remain a primary objective. Even as you follow the other rules here to create a `<title>` tag that is useful to the search engines, remember that humans will likely see your `<title>` tag presented in the search results for your page. Don't scare them away with a `<title>` tag that looks like it's written for a machine.

Be consistent

Once you've determined a good formula for your pages in a given section or area of your site, stick to that regimen. You'll find that as you become a trusted and successful "brand" in the SERPs, users will seek out your pages on a subject area and have expectations that you'll want to fulfill.

Meta Description Tags

Meta descriptions have three primary uses:

- To describe the content of the page accurately and succinctly
- To serve as a short text "advertisement" to prompt searchers to click on your pages in the search results
- To display targeted keywords, not for ranking purposes, but to indicate the content to searchers

Great meta descriptions, just like great ads, can be tough to write, but for keyword-targeted pages, particularly in competitive search results, they are a critical part of driving traffic from the engines through to your pages. Their importance is much greater for search terms where the intent of the searcher is unclear or different searchers might have different motivations.

Here are six good rules for meta descriptions:

Tell the truth

Always describe your content honestly. If it is not as "sexy" as you'd like, spice up your content; don't bait and switch on searchers, or they'll have a poor brand association.

Keep it succinct

Be wary of character limits—currently Google displays as few as 140 characters, Yahoo! up to 165, and Bing up to 200+ (it'll go to three vertical lines in some cases). Stick with the smallest—Google—and keep those descriptions at 140 characters (including spaces) or less.

Write ad-worthy copy

Write with as much sizzle as you can while staying descriptive, as the perfect meta description is like the perfect ad: compelling and informative.

Analyze psychology

The motivation for an organic-search click is frequently very different from that of users clicking on paid results. Users clicking on PPC ads may be very directly focused on making a purchase, and people who click on an organic result may be more interested in research or learning about the company. Don't assume that successful PPC ad text will make for a good meta description (or the reverse).

Include relevant keywords

It is extremely important to have your keywords in the meta `description` tag—the boldface that the engines apply can make a big difference in visibility and click-through rate. In addition, if the user's search term is not in the meta description, chances are reduced that the meta description will be used as the description in the SERPs.

Don't employ descriptions universally

You shouldn't always write a meta description. Conventional logic may hold that it is usually wiser to write a good meta description yourself to maximize your chances of it being used in the SERPs, rather than let the engines build one out of your page content; however, this isn't always the case. If the page is targeting one to three heavily searched terms/phrases, go with a meta description that hits those users performing that search.

However, if you're targeting longer-tail traffic with hundreds of articles or blog entries or even a huge product catalog, it can sometimes be wiser to let the engines themselves extract the relevant text. The reason is simple: when engines pull, they always display the keywords (and surrounding phrases) that the user searched for. If you try to force a meta description, you can detract from the relevance that the engines make naturally. In some cases, they'll overrule your meta description anyway, but because you can't

consistently rely on this behavior, opting out of meta descriptions is OK (and for massive sites, it can save hundreds or thousands of man-hours). Because the meta description isn't a ranking signal, it is a second-order activity at any rate.

Heading Tags

The heading tags in HTML (`<h1>`, `<h2>`, `<h3>`, etc.) are designed to indicate a headline hierarchy in a document. Thus, an `<h1>` tag might be considered the headline of the page as a whole, whereas `<h2>` tags would serve as subheadings, `<h3>`s as tertiary-level subheadings, and so forth. The search engines have shown a slight preference for keywords appearing in heading tags. Generally when there are multiple heading tags on a page, the engines will weight the higher-level heading tags heavier than those below them. For example, if the page contains `<h1>`, `<h2>`, and `<h3>` tags, the `<h1>` will be weighted the heaviest. If a page contains only `<h2>` and `<h3>` tags, the `<h2>` would be weighted the heaviest.

In some cases, you can use the `<title>` tag of a page, containing the important keywords, as the `<h1>` tag. However, if you have a longer `<title>` tag, you may want to use a more focused, shorter heading tag including the most important keywords from the `<title>` tag. When a searcher clicks a result from the engines, reinforcing the search term he just typed in with the prominent headline helps to indicate that he has arrived on the right page with the same content he sought.

Many publishers assume that they have to use an `<h1>` tag on every page. What matters most, though, is the highest-level heading tag you use on a page, and its placement. If you have a page that uses an `<h3>` heading at the very top, and any other heading tags further down on the page are `<h3>` or lower level, then that first `<h3>` tag will carry just as much weight as if it were an `<h1>`.

Again, what matters most is the semantic markup of the page, and the first heading tag presumably is intended to be a label for the entire page (so it plays a complementary role to the `<title>` tag), and you should treat it as such. Other heading tags on the page should be used to label subsections of the content.

It's also a common belief that the size at which the heading tag is displayed is a

factor. For the most part, the styling of your heading tags is not a factor in the SEO weight of the heading tag. You can style the tag however you want, as shown in Figure 6-20, provided that you don't go to extremes (because it acts as a title for the whole page, it should probably be the largest text element on the page).

Figure 6-20. Headings styled to match the site

Document Text

The HTML text on a page was once the center of keyword optimization activities. In the early days of SEO, metrics such as keyword density and keyword saturation were used to measure the perfect level of keyword usage on a page. To the search engines, however, text in a document, particularly the frequency with which a particular term or phrase is used, has very little impact on how happy a searcher will be with that page.

In fact, quite often a page laden with repetitive keywords attempting to please the engines will provide a very poor user experience, and this can result in lower rankings instead of higher ones. It's much more valuable to create semantically rich content that covers the topic matter implied by the page's `<title>` tag in a comprehensive way. This means naturally including synonyms, and covering related topic areas in a manner that increases the chances of satisfying the needs of a large percentage of visitors to that page. It's a good idea to use the main keyword for a page in the `<title>` tag and the main heading tag. It might also appear in the main content, but the use of synonyms for the main keyword and related concepts is at least as important. As a result, it's more important to focus on creating high-quality content than it is to keep repeating the main keyword.

Term frequency—Inverse document frequency

TF-IDF consists of two parts. The first is *term frequency*, which relates to the frequency of usage of a keyword or key phrase on a page, in comparison to usage levels in competing documents. This is similar to keyword density, except weighting is done logartihmically to reduce the impact of keyword repetition. The result is that a page which uses a phrase 10 times might be seen only as twice as good a match as a page that uses that phrase once.

Term frequency analysis can be very useful in understanding how your page compares semantically with pages that rank highly in Google's results. Coauthor Eric Enge has written an article about this.

Inverse document frequency (IDF) is more about identifying the uniqueness of a term. For example, the word "omnipotent" is used much less on the web than the word "powerful." Therefore, a page using the word "omnipotent" may be seen as a bit more unique. If a user enters the word "omnipotent" as part of a search query, it will be far more likely to surface a page using that word in the results. IDF can be a very way to identify new ranking opportunities for your web page as coauthor Eric Enge has written elsewhere.

TF-IDF helps search engines understand what terms a page emphasizes most, and what terms most uniquely define a page at the same time. Publishers can use TF-IDF analysis on competing pages ranking in the top 10 for a given search term to learn what search engines appear to value the most in content for a given search query.

Used properly, this is not about keyword stuffing, but instead focuses on learning key information being sought out by users in relation to a search query. For example, if someone searches on “oil filters” and lands on your page, he may also want information on oil filter wrenches.

Using TF-IDF analysis on competing pages can help you learn about such opportunities to improve the user experience of a page, and help you with SEO for that page at the same time.

Page segmentation

It used to be that Google could not understand the layout of a page that well, simply because it could not read CSS files and process them like a browser does. However, that has changed, as documented in a post on the Google Webmaster Central Blog.

As a result, Google is quite likely to fully understand the layout of your pages. Given this, where the keywords are used on the page also matters. Use of keywords in your left or right sidebar, or your footer, probably matters less than the content used in the main body of your page.

In addition, with HTML5, new markup exists that allows you to explicitly identify the section of your page that represents the main content. You can use this markup to help make Google’s job easier, and to make sure that other search engines are able to locate that content.

Synonyms

Use of related terms is also a factor. A page about “left-handed golf clubs” should not use that exact phrase every time the product is referenced. This would not be a natural way of writing, and could be interpreted by the search engines as a signal of poor document quality, lowering the page’s rankings.

Instead, allow your content creators to write naturally. This will cause them to use other phrases, such as “the sticks,” “set of clubs,” “lefty clubs,” and other variants that people use in normal writing style.

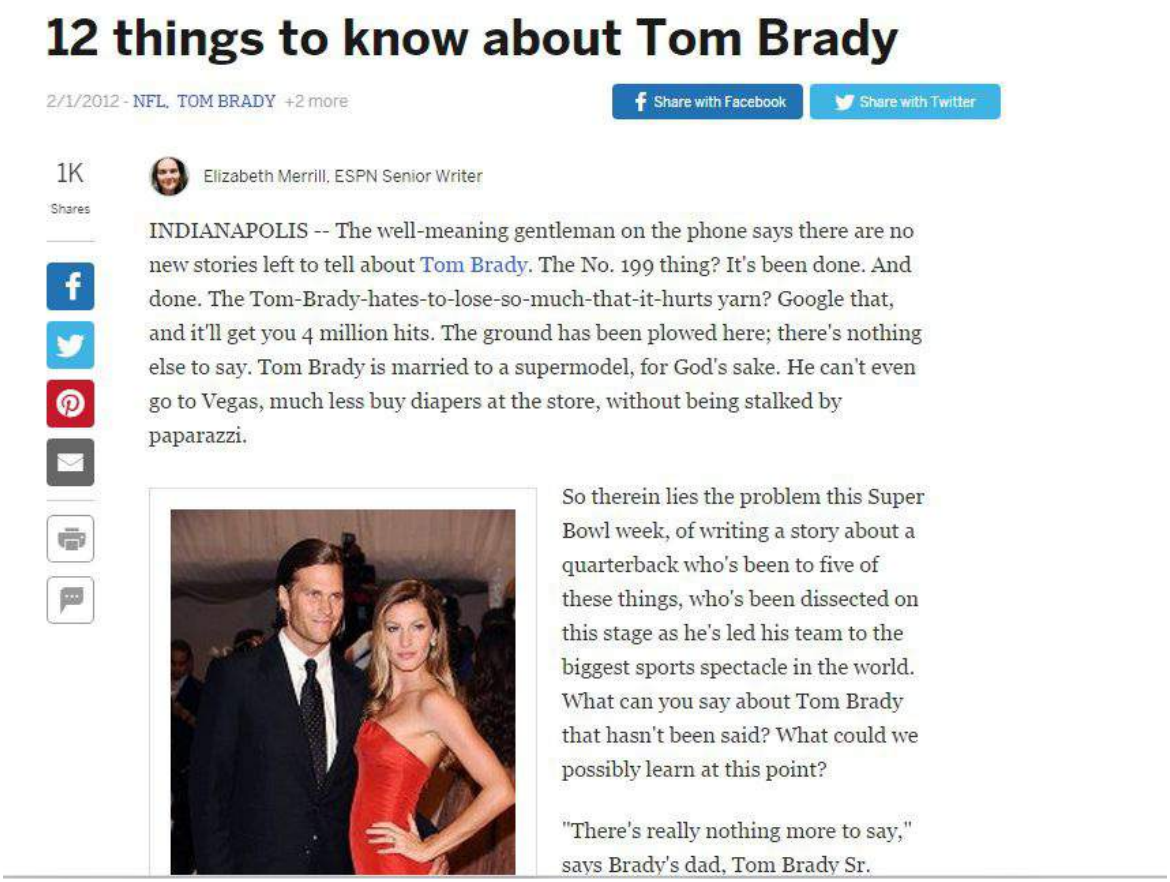
Using synonyms represents a key step away from manipulative SEO techniques for creating pages to try to rank for specific search terms.

Co-occurrence, phrase-based indexing, and entity salience

The presence of a specific keyword or phrase on a page will likely increase the probability of finding other words on those pages. For example, if you are reading an article on the life of New England Patriots quarterback Tom Brady, you would expect to see mentions of Gisele Bündchen, or his children. The absence of such mentions could be a signal of a poor-quality page.

To put this in the positive, inclusion of more information that you would expect to see in the content can be interpreted as an indication that it's a good page. Chances are good that more people reading the article will be satisfied with it as well, as they might expect to learn about Tom Brady's family.

Of course, a high-quality article will probably talk about his parents, his high school football coach, his sisters, and all the other aspects of his life as well. The key is to focus on providing a more complete response to the topic than others covering the same topic might do. The ESPN article shown in Figure 6-21 is an example of such an indepth article.



12 things to know about Tom Brady

2/1/2012 - NFL, TOM BRADY +2 more

Share with Facebook | Share with Twitter

1K Shares

Elizabeth Merrill, ESPN Senior Writer

INDIANAPOLIS -- The well-meaning gentleman on the phone says there are no new stories left to tell about Tom Brady. The No. 199 thing? It's been done. And done. The Tom-Brady-hates-to-lose-so-much-that-it-hurts yarn? Google that, and it'll get you 4 million hits. The ground has been plowed here; there's nothing else to say. Tom Brady is married to a supermodel, for God's sake. He can't even go to Vegas, much less buy diapers at the store, without being stalked by paparazzi.

So therein lies the problem this Super Bowl week, of writing a story about a quarterback who's been to five of these things, who's been dissected on this stage as he's led his team to the biggest sports spectacle in the world. What can you say about Tom Brady that hasn't been said? What could we possibly learn at this point?

"There's really nothing more to say," says Brady's dad, Tom Brady Sr.

Figure 6-21. A comprehensive article on Tom Brady from ESPN

Not all of the topic needs to be addressed on each individual page. Linking to other relevant resources and high-quality content, both on your site as well as on third-party sites, can play a key role in establishing your page as a great answer to the user's question.

This last step may well be equally important in the overall page optimization process. No single page will answer every question from every possible user, so addressing a significant percentage of questions, and then connecting with other pages to answer follow-on questions on the same topic, is an optimal structure.

On the product pages of an ecommerce site, where there will not be article-style content, this can mean a combination of well-structured and unique description text and access to key refinements, such as individual brands, related product types, the presence of a privacy policy, "About us" information, a shopping cart, and more.

Image Filenames and alt Attributes

Incorporating images on your web pages can substantively enrich the user experience. However, the search engines cannot read the images directly. There are two elements that you can control to give the engines context for images:

The filename

Search engines look at the image filename to see whether it provides any clues to the content of the image. Don't name your image *example.com/img4137a-b12.jpg*, as it tells the search engine nothing at all about the image, and you are passing up the opportunity to include keyword-rich text.

If it is a picture of Abe Lincoln, name the file *abe-lincoln.jpg* and/or have the `src` URL string contain it, as in *example.com/abe-lincoln/portrait.jpg*.

The `alt` *attribute text*

Image tags in HTML permit you to specify the `alt` attribute. This is a place where you can provide more information about what is in the image, and again where you can use your targeted keywords. Here is an example for the

picture of Abe Lincoln:

```
<img alt="Abe Lincoln photo" src="http://example.com/abe-lincoln.jpg" />
```

Use the quotes if you have spaces in the text string of the `alt` content! Sites that have invalid `<img>` tags frequently lump a few words without quotes into the `<img>` tag, intended for the `alt` content—but with no quotes, all terms after the first word will be lost.

This usage of the image filename and the `alt` attribute permits you to reinforce the major keyword themes of the page. This is particularly useful if you want to rank in image search. Make sure the filename and the `alt` text reflect the content of the picture, and do not artificially emphasize keywords unrelated to the image (even if they are related to the page). Although the `alt` attribute and the image filename are helpful, you should not use image links as a substitute for text links with rich anchor text, which carry much more weight from an SEO perspective.

Presumably, your picture will relate very closely to the content of the page, and using the image filename and the `alt` text will help reinforce the page's overall theme.

Boldface Text

While it used to be true that including keywords in bold text had a very slight effect in rankings, this is no longer the case.

Keyword Cannibalization

As we discussed earlier, you should not use common keywords across multiple page titles. This advice applies to more than just the `<title>` tags.

One of the nastier problems that often crops up during the course of a website's information architecture, *keyword cannibalization* refers to a site's targeting of popular keyword search phrases on multiple pages, forcing the engines to pick which one is most relevant. In essence, a site employing cannibalization competes with itself for rankings and dilutes the ranking power of internal anchor text, external links, and keyword relevancy.

Avoiding cannibalization requires strict site architecture with attention to detail. Plot out your most important terms on a visual flowchart (or in a spreadsheet file, if you prefer), and pay careful attention to what search terms each page is targeting. Note that when pages feature two-, three-, or four-word phrases that contain the target search phrase of another page, linking back to that page within the content with the appropriate anchor text will avoid the cannibalization issue.

For example, if you had a page targeting “mortgages” and another page targeting “low-interest mortgages,” you would link back to the “mortgages” page from the “low-interest mortgages” page using the anchor text “mortgages” (see Figure 6-22). You can do this in the breadcrumb or in the body copy. The *New York Times* does the latter, where keywords in the body copy link to the related resource page on the site.

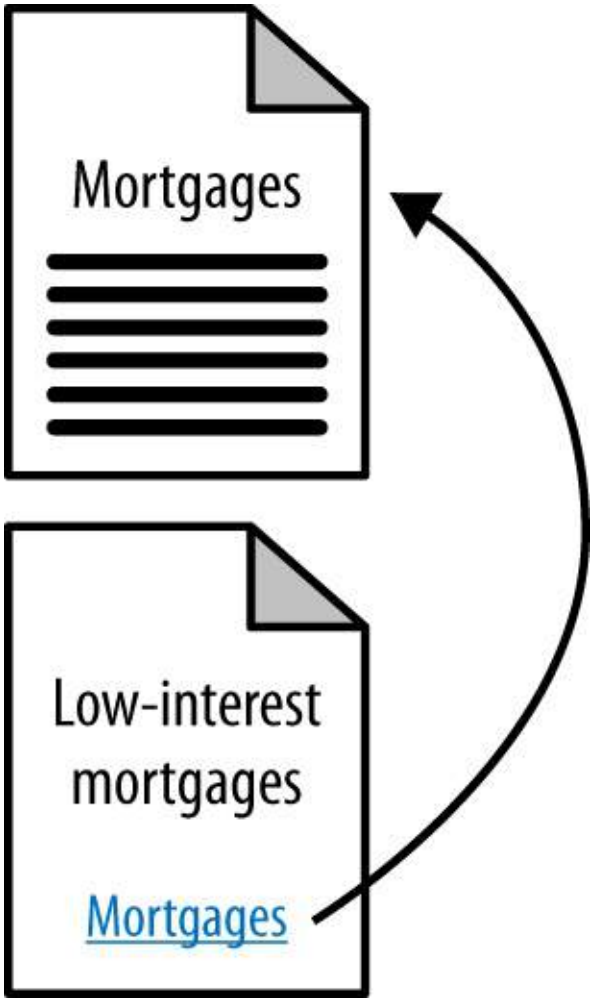


Figure 6-22. Adding lots of value with relevant cross-links

Keyword Targeting in Content Management Systems and Automatically Generated Content

Large-scale publishing systems, or those that produce automatically generated content, present some unique challenges. If hundreds of pages are being created every day, it is not feasible to do independent keyword research on each and every page, making page optimization an interesting challenge.

In these scenarios, the focus turns to methods/recipes for generating unique titles, heading tags, and content for each page. It is critical to educate the writers on ways to implement titles and headings that capture unique, key aspects of the articles' content. More advanced teams can go further with this and train their writing staff on the use of keyword research tools to optimize this process even more.

In the case of automatically generated material (such as that produced from algorithms that mine data from larger textual bodies), the key is to automate means for extracting a short (fewer than 55 characters) description of the article and making it unique from other titles generated elsewhere on the site and on the Web at large.

Effective Keyword Targeting by Content Creators

Very frequently, someone other than an SEO professional is responsible for content creation. Content creators often do not have an innate knowledge of how SEO works, or worse, they may think they know how it works, but have the wrong idea about it. Some training for your writers is critical. This is particularly important when you're dealing with large websites and large teams of writers.

Here are the main components of web page copywriting that your writers must understand:

- Search engines look to match up a user's search queries with the keyword phrases, their synonyms, and related concepts on your web pages. If some combination of all of these does not appear on the page, chances are good that your page will never achieve significant ranking for those search phrases.

- The search phrases users may choose to use when looking for something are infinite in variety, but certain phrases will be used much more frequently than

others.

- Using the more popular phrases you wish to target on a web page in the content for that page is essential to SEO success for that page.
- Make sure that the writers understand the concepts of co-occurrence and entity salience, discussed earlier in this chapter, so they don't create content that uses the main keyword excessively. They need to focus on creating semantically rich content that stays on the topic of the main target keyword phrase for the page, while still writing naturally.
- The `<title>` tag is the most important element on the page. Next is the first header (usually `<h1>`), and then the main body of the content.
- There are tools (as outlined in Chapter 5) that allow you to research and determine what the most interesting phrases are.

If you can get these six points across, you are well on your way to empowering your content creators to perform solid SEO. The next key element is training them on how to pick the right keywords to use.

This can involve teaching them how to use keyword research tools similar to the ones we discussed in Chapter 5, or having the website's SEO person do the research and provide the terms to the writer.

The most important factor to reiterate to content creators is that content quality and user experience still come first. Then, by intelligently making sure the right keywords and phrases are properly used throughout the content, they can help bring search engine traffic to your site. Reverse these priorities, and you can end up with keyword stuffing or other spam issues.

Long-Tail Keyword Targeting

As we outlined in Chapter 5, the small-volume search terms, when tallied up, represent 70% or more of overall search traffic, and the more obvious, high-volume terms represent only 30%.

For example, if you run a site targeting searches for *new york pizza* and *new york pizza delivery*, you might be surprised to find that hundreds of single searches

each day for terms such as *pizza delivery on the corner of 57th & 7th*, or *Manhattan's tastiest Italian-style sausage pizza*, when taken together, will actually provide considerably more traffic than the popular phrases you've researched. As we covered in Chapter 5, this concept is called the *long tail* of search.

Targeting the long tail is another aspect of SEO that combines art and science. In Figure 6-23, you may not want to implement entire web pages for a history of pizza dough, pizza with white anchovies, or Croatian pizza.

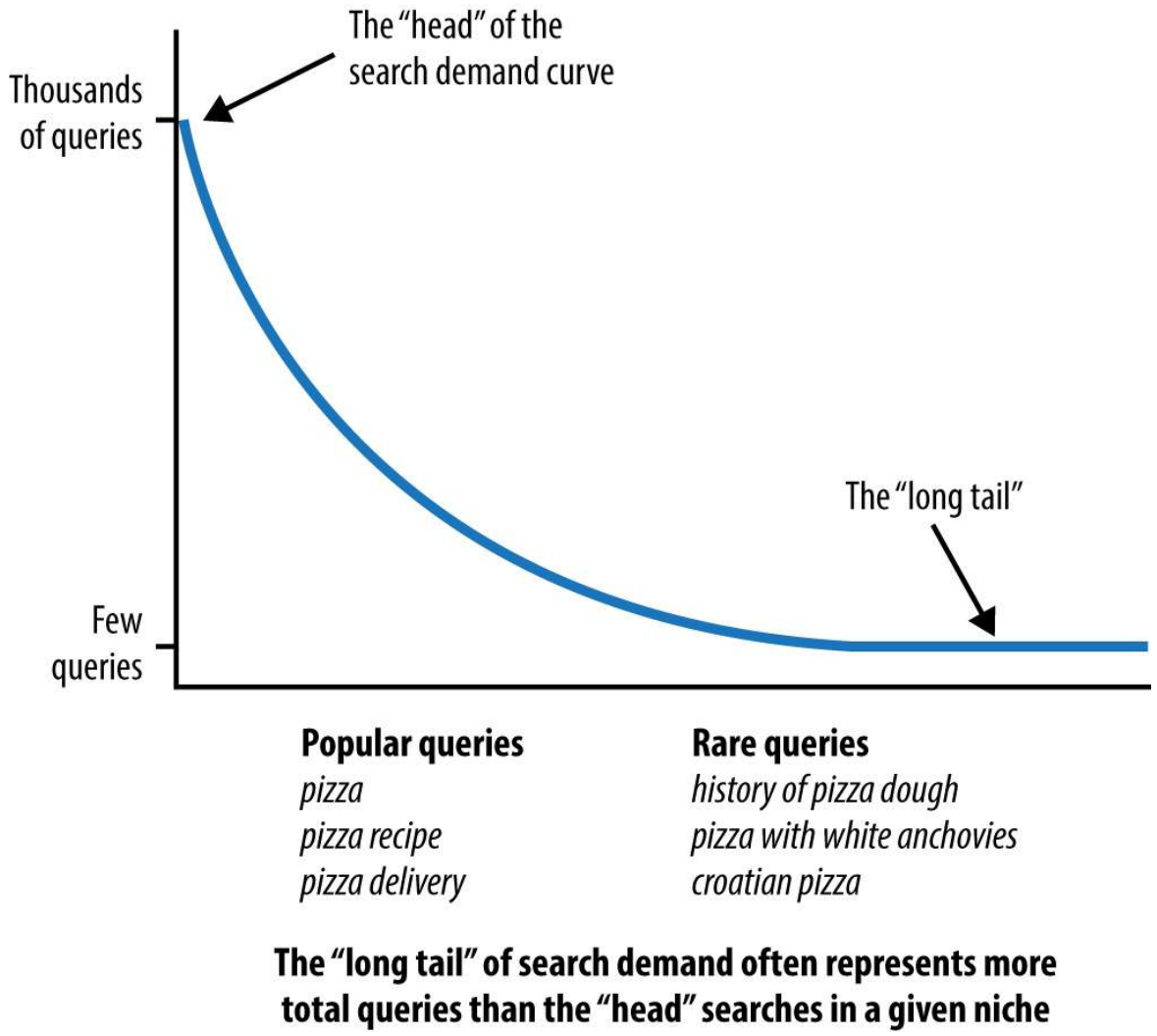


Figure 6-23. Example of the long-tail search curve

Finding scalable ways to chase long-tail keywords is a complex topic. It is also one where many publishers get into a lot of trouble, as they think they need to

create a new page for each potential search phrase that a user might type in related to their business, and this is not the case. You can address much of the long tail of search by using the right content optimization practices on your site.

Perhaps you have a page for ordering pizza in New York City, and you have a good title and heading tag on the page (e.g., “New York City Pizza: Order Here”), a phone number and a form for ordering the pizza, and no other content. If that is all you have, that page is not competing effectively for rankings on long-tail search terms. To fix this, you need to write additional content for the page. Ideally, this would be content that talks about the types of pizza that are popular in New York City, the ingredients used, and other related topics that might draw in long-tail search traffic.

If you have a page for San Jose pizza, the picture gets even more complicated. You don’t want your content on the San Jose page to be the same as it is on the New York City page. This presents potential duplicate content problems, as we will outline in “Duplicate Content Issues”, or the keyword cannibalization issues we discussed earlier in this chapter.

To maximize your success, find a way to generate different content for those two pages, ideally tuned to the specific needs of the audience that arrives at those pages. Perhaps the pizza preferences of the San Jose crowd are different from those in New York City. Of course, the geographic information is inherently different between the two locations, so driving directions from key locations might be a good thing to include on the page.

If you have pizza parlors in 100 cities, this can get very complex indeed. The key here is to remain true to the diverse needs of your users, yet use your knowledge of the needs of search engines and searcher behavior to obtain that long-tail traffic.

Content Optimization

Content optimization relates to how the presentation and architecture of the text, image, and multimedia content on a page can be optimized for search engines. Many of these recommendations are second-order effects. Having the right formatting or display won’t boost your rankings directly, but through it, you’re more likely to earn links, get clicks, and eventually benefit in search rankings. If

you regularly practice the techniques in this section, you'll earn better consideration from the engines and from the human activities on the Web that influence their algorithms.

Content Structure

Because SEO has become such a holistic part of website development and improvement, it is no surprise that *content formatting*—the presentation, style, and layout choices you select for your content—is a part of the process. A browser-safe sans serif font such as Arial or Helvetica is a wise choice for the Web; Verdana in particular has received high praise from usability/readability experts (for a full discussion of this topic, see *http://webaim.org/techniques/fonts/*).

Verdana is one of the most popular of the fonts designed for on-screen viewing. It has a simple, straightforward design, and the characters or glyphs are not easily confused. For example, the uppercase *I* and the lowercase *L* have unique shapes, unlike in Arial, in which the two glyphs may be easily confused (see Figure 6-24).

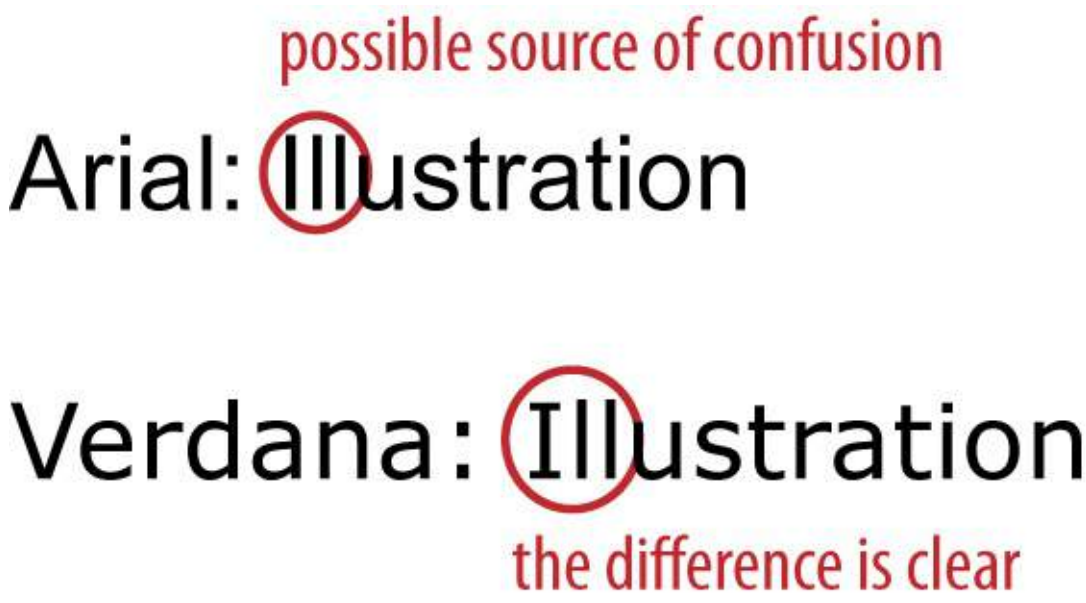


Figure 6-24. Arial and Verdana font comparison

Another advantage of Verdana is the amount of spacing between letters. One

consideration to take into account with Verdana is that it is a relatively large font. The words take up more space than words in Arial, even at the same point size (see Figure 6-25).

Verdana: Notice the length of this sentence.

Arial: Notice the length of this sentence.

Figure 6-25. How fonts impact space requirements

The larger size improves readability but also can potentially disrupt carefully planned page layouts.

In addition to font choice, sizing and contrast issues are important considerations. Type that is smaller than 10 points is typically very challenging to read, and in all cases, relative font sizes are recommended so that users can employ browser options to increase/decrease size if necessary. Contrast—the color difference between the background and text—is also critical; legibility usually drops for anything that isn't black (or very dark) on a white background.

Content length and word count

Content length is another critical piece of the optimization puzzle that's mistakenly placed in the "keyword density" or "unique content" bucket of SEO. In fact, content length can play a big role in terms of whether your material is easy to consume and easy to share.

People often ask about the ideal length for a piece of content. The reality is that the perfect length for a piece of content is determined by the nature of the topic being addressed. Many pieces of content do well because they are short and very easy to consume. On the other hand, some content will fare best when it's lengthy and comprehensive in nature.

Longer articles also have the opportunity to show up in Google's indepth articles section. In some cases, where appropriate, Google will feature several longer, more detailed articles on a given topic. For more information on having your articles appear as "indepth" articles, see *http://bit.ly/in-depth_articles*.

Visual layout

Last but not least in content structure optimization is the display of the material.

Beautiful, simple, easy-to-use, and consumable layouts instill trust and garner far more readership and links than poorly designed content wedged between ad blocks that threaten to overtake the page. For more on this topic, check out "The Golden Ratio in Web Design" from NetTuts, which has some great illustrations and advice on laying out web content on the page.

CSS and Semantic Markup

CSS is commonly mentioned as a best practice for general web design and development, but its principles provide some indirect SEO benefits as well. Google used to recommend keeping pages smaller than 101 KB, and it was a common belief that there were benefits to implementing pages that were small in size. Now, however, search engines deny that code size is a factor at all, unless it is extreme. Still, keeping file size low means your pages have faster load times, lower abandonment rates, and a higher probability of being fully read and more frequently linked to. This is particularly important in mobile environments.

It also used to be the case that search engines could not read CSS code and render pages in the same manner as a browser does. In October 2014, Google made it clear that it is able to do just that, so good clean page layout, as set up by your CSS, could potentially be considered as a factor in evaluating page quality.

Your experience may vary, but good CSS makes it easy, so there's no reason not to make it part of your standard operating procedure for web development. Use tableless CSS stored in external files, keep JavaScript calls external, and separate the content layer from the presentation layer, as shown on CSS Zen Garden, a site that offers many user-contributed stylesheets for formatting the same HTML content.

You can use CSS code to provide emphasis, to quote/reference, and to reduce the use of tables and other bloated HTML mechanisms for formatting, which can positively impact your SEO. Be sure to allow Googlebot access to your CSS files.

Google, Bing, and Yahoo! have come together to sponsor a standard for markup called Schema.org.[10] This represented a new level of commitment from the search engines to the concept of marking up content, or more broadly, to allowing the publisher to provide information about the content to the search engines. By "marking up," content, we mean tagging your content using XML

tags to categorize it. For example, you may label a block of content as containing a recipe, and another block of content as containing a review.

This notion of advanced markup was not new, as all of the search engines have supported semantic markup at a limited level, and have used this markup to show *rich snippets*, an example of which is shown in Figure 6-26.

One of the original ways a publisher had to communicate information about a web page to search engines was with metatags. Unfortunately, these were so badly abused by spammers that Google stopped using them as a ranking signal. Google confirmed this publicly in a post in 2009, which noted that "Google has ignored the `keywords` meta tag for years and currently we see no need to change that policy."[11]

Google used to publicly state that it does not use markup as a ranking factor, and while those statements are no longer being publicly made, there continues to be no evidence that it has been made a ranking factor. However, there are important SEO benefits to using markup.

Markup in search results

As previously mentioned, markup is sometimes used by search engines to create a rich snippet. Figure 6-26 shows an example of rich snippets in the search results for a recipe for a Cambodian dish called Loc Lac.

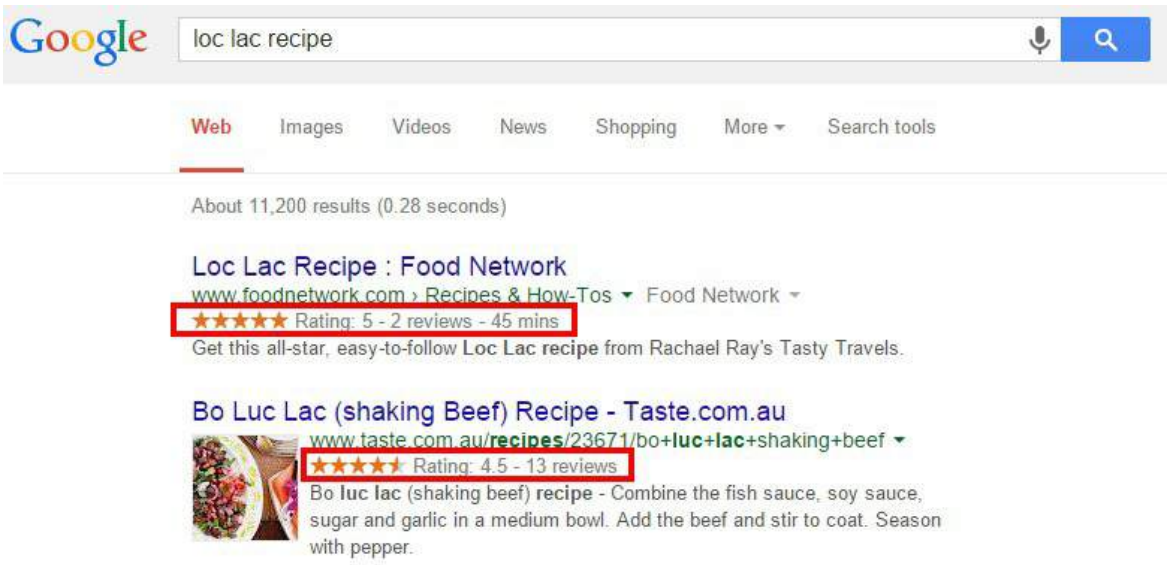


Figure 6-26. Example of recipe rich snippet on Google

Based on the markup that Google found in the HTML, it has enhanced the result by showing the recipe reviews (the number of stars), the required cooking time, and the calories of the meal. The type of markup used for this example is called *microformats*. Figure 6-27 shows what the source code looks like for this example.

```
    </div>
    <div id="zoneRecipe" itemscope="" itemtype="http://schema.org/Recipe" data-itemtype="Recipe" data-itemtypeid="1"
data-typespecificid="15869">

    <link itemprop="url" href="http://allrecipes.com/recipe/cream-puffs/"/>
    <meta itemprop="mainEntityOfPage" content="True"/>
    <div class="detail-section greydotted ingredients">

        <!-- Picture, Title, Description, etc. -->
        <div id="divHeroPhotoContainer" class="detail-left fl-left testB">

                <a href="/recipe/cream-puffs/photo-gallery.aspx" id="lnkOpenCarousel" class="modal-link unsavedExempt
open_modal-recipe-videos frame" rel="modal-recipe-photos" data-close-layer-on-login="" calltoaction="" style=""><img
id="imgPhoto" class="rec-image rec-shadow hero-image marb10 photo" itemprop="image" title="Cream Puffs Recipe"
src="http://images.media-allrecipes.com/userphotos/250x250/00/65/47/654778.jpg" alt="Cream Puffs Recipe"
style="height:250px;width:250px;" /></a>
```

Figure 6-27. Sample of microformats code for a recipe

Supported types of markup

There are a few different standards for markup. The most common ones are *microdata*, *microformats*, and *RDFa*. Schema.org is based off of the microdata standard. However, the search engines have implemented rich snippets based on some (but not all) aspects of microformats prior to the announcement of Schema.org, and they will likely continue support for these for some period of time.

It is likely that any new forms of rich snippets implemented by the search engines will be based off of Schema.org (microdata), not microformats or RDFa. Some of the formats already supported by Google include:

- People
- Products
- Events
- Business and organizations
- Video

Impact of rich snippets

The key reason that the search engines are pursuing rich snippets is that they have done extensive testing that has proven that rich snippets can increase click-through rates. Searchers like seeing more information about the page in the search results. Thus, you can expect that the search engines will continue to implement support for more of these types of search result enhancements based on markup.

From an SEO perspective, increasing click-through rate is highly desirable; it brings us more relevant traffic. In addition, we know that search engines measure user interaction with the search results and that click-through rate is a ranking factor. This was first publicly confirmed in an interview with Bing's Duane Forrester.

So, while the search engines do not use semantic markup directly as a ranking signal, the indirect impact of rich snippets providing a higher click-through rate acts as a ranking signal.

For more information on semantic markup, see the sections "Semantic Search" and "Schema.org".

Content Uniqueness and Depth

Few can debate the value the engines place on robust, unique, value-added content—Google in particular has had several rounds of kicking low-quality-content sites out of its indexes, and the other engines have followed suit.

The first critical designation to avoid is *thin content*—a phrase that (loosely) refers to a page the engines do not feel contributes enough unique content to warrant the page's inclusion in the search results. How much content is enough content to not be considered thin? The criteria have never been officially listed, but here are some examples gathered from engineers and search engine representatives:

- At least 30 to 50 unique words, forming unique, parsable sentences that other sites/pages do not have (for many pages much more is appropriate, so consider this a minimum).

- Unique HTML text content, different from other pages on the site in more

than just the replacement of key verbs and nouns (yes, this means all those webmasters who build the same page and just change the city and state names thinking it is “unique” are mistaken).

- Unique titles and meta description elements. If you can’t write unique meta descriptions, just exclude them. Algorithms can trip up pages and boot them from the index simply for having near-duplicate meta tags.
- Unique video/audio/image content. The engines have started getting smarter about identifying and indexing pages for vertical search that wouldn’t normally meet the “uniqueness” criteria.

NOTE

By the way, you can often bypass these limitations if you have a good quantity of high-value external links pointing to the page in question (though this is very rarely scalable) or an extremely powerful, authoritative site (note how many one-sentence Wikipedia stub pages still rank).

The next criterion from the engines demands that websites “add value” to the content they publish, particularly if it comes from (wholly or partially) a secondary source.

A word of caution to affiliates

This word of caution most frequently applies to affiliate sites whose republishing of product descriptions, images, and so forth has come under search engine fire numerous times. In fact, it is best to anticipate manual evaluations here even if you’ve dodged the algorithmic sweep.

The basic tenets are:

- Don’t simply republish something that’s found elsewhere on the Web unless your site adds substantive value to users, and don’t infringe on others’ copyrights or trademarks.
- If you’re hosting affiliate content, expect to be judged more harshly than others, as affiliates in the SERPs are one of users’ top complaints about

search engines.

- Small changes such as a few comments, a clever sorting algorithm or automated tags, filtering, a line or two of text, simple mashups, or advertising do *not* constitute "substantive value."

For some exemplary cases where websites fulfill these guidelines, check out the way sites such as CNET, Urbanspoon, and Metacritic take content*products*reviews from elsewhere, both aggregating *and* adding value for their users.

Last but not least, Google has provided a guideline to refrain from trying to place "search results in the search results." For reference, look at the post from Google's Matt Cutts, including the comments, at *http://www.mattcutts.comblogsearch-results-in-search-results/*. Google's stated position is that search results generally don't add value for users, though others have made the argument that this is merely an anticompetitive move.

Sites can benefit from having their search results transformed into more valuable listings and category/subcategory landing pages. Sites that have done this have had great success recovering rankings and gaining traffic from Google.

In essence, you want to avoid the potential for your site pages being perceived, both by an engine's algorithm and by human engineers and quality raters, as search results. Refrain from:

- Pages labeled in the title or headline as "search results" or "results"

- Pages that appear to offer a query-based list of links to "relevant" pages on the site without other content (add a short paragraph of text, an image, and formatting that make the "results" look like detailed descriptions/links instead)

- Pages whose URLs appear to carry search queries (e.g., *?q=miami+restaurants* or *?search=Miami+restaurants* versus */miami-restaurants*)

- Pages with text such as "Results 1 through 10"

Though it seems strange, these subtle, largely cosmetic changes can mean the difference between inclusion and removal. Err on the side of caution and dodge the appearance of search results.

Content Themes

A less discussed but also important issue is the fit of each piece of content to your site. If you create an article about pizza, but the rest of your site is about horseshoes, your article is unlikely to rank for the term *pizza*. Search engines analyze and understand what sites, or sections of sites, focus on.

You can think of this as being the "theme" of the site (or section). If you start creating content that is not on the same theme, that content will have a very difficult time ranking. Further, your off-topic content could potentially weaken the theme of the rest of the site.

One site can support multiple themes, but each themed section needs to justify its own existence by following good SEO practices, including getting third parties to implement links from the pages of their sites to that section. Make sure you keep your content on topic, and this will help the SEO for all of the pages of your site.

Copyblogger has created a tool to help measure the fit of a given article to your site, known as Scribe. Not only will Scribe measure the fit of an article to your site, it will also offer a more general look at the consistency of the content across your site overall.

Duplicate Content Issues

Duplicate content generally falls into three categories: exact (or true) duplicates, whereby two URLs output identical content; near duplicates, whereby there are small content differentiators (sentence order, image variables, etc.); and cross-domain duplicates, whereby exact or near duplication exists on multiple domains.

There are two related concepts that are not treated by Google the same way as duplicate content, but are often confused by publishers and inexperienced SEO practitioners. These are:

Thin content

As noted previously, these are pages that don't have much content on them at all. An example might be a set of pages built out to list all the locations for a business with 5,000 locations, but the only content on all the pages is the address of each location.

Thin slicing

These are pages with very minor differences in focus. Consider a site that sells running shoes, and one of the shoes offered is men's Nike Air Max LTD running shoes, which comes in sizes 6, 6.5, 7, 7.5, 8,...15. If the site had a different page for each size of this shoe, even though each page would actually be showing a different product, there is just not much useful difference between the pages overall.

Google has been clear that it doesn't like thin content or thin slicing. Either can trigger Google's Panda algorithm, which is discussed more in Chapter 9. Exactly how Bing differentiates duplicate content, thin content, and thin slicing is less clear, but it also prefers that publishers avoid creating these types of pages.

Duplicate content can result from many causes, including licensing of content to or from your site, site architecture flaws due to non-SEO-friendly content management systems, or plagiarism. Not too long ago, however, spammers in desperate need of content began the now much-reviled process of scraping content from legitimate sources, scrambling the words (through many complex processes), and repurposing the text to appear on their own pages in the hopes of attracting long-tail searches and serving contextual ads (and various other nefarious purposes).

Thus, today we're faced with a world of duplicate content issues and their corresponding penalties. Here are some definitions that are useful for this discussion:

Unique content

This is written by humans; is completely different from any other combination of letters, symbols, or words on the Web; and is clearly not

manipulated through computer text-processing algorithms (such as Markov-chain-employing spam tools).

Snippets

These are small chunks of content, such as quotes, that are copied and reused; they are almost never problematic for search engines, especially when included in a larger document with plenty of unique content.

Shingles

Search engines look at relatively small phrase segments (e.g., five to six words) for the presence of the same segments on other pages on the Web. When there are too many shingles in common between two documents, the search engines may interpret them as duplicate content.

Duplicate content issues

This phrase is typically used to refer to duplicate content that is not in danger of getting a website penalized, but rather is simply a copy of an existing page that forces the search engines to choose which version to display in the index (a.k.a. duplicate content filter).

Duplicate content filter

This is when the search engine removes substantially similar content from a search result to provide a better overall user experience.

Duplicate content penalty

Penalties are applied rarely and only in egregious situations. Engines may devalue or ban other web pages on the site, too, or even the entire website.

Consequences of Duplicate Content

Assuming your duplicate content is a result of innocuous oversights on your developer's part, the search engine will most likely simply filter out all but one of the pages that are duplicates because it wants to display one version of a particular piece of content in a given SERP. In some cases, the search engine

may filter out results prior to including them in the index, and in other cases it may allow a page in the index and filter it out when it is assembling the SERPs in response to a specific query. In the latter case, a page may be filtered out in response to some queries and not others.

Searchers want diversity in the results, not the same results repeated again and again. Search engines therefore try to filter out duplicate copies of content, and this has several consequences:

- A search engine bot comes to a site with a *crawl budget*, which is the number of pages it plans to crawl in each particular session. Each time it crawls a page that is a duplicate (which is simply going to be filtered out of search results) you have let the bot waste some of its crawl budget. That means fewer of your "good" pages will get crawled. This can result in fewer of your pages being included in the search engine index.

- Links to duplicate content pages represent a waste of link authority. Duplicated pages can gain PageRank, or link authority, and because it does not help them rank, that link authority is misspent.

- No search engine has offered a clear explanation for how its algorithm picks which version of a page it shows. In other words, if it discovers three copies of the same content, which two does it filter out? Which one does it still show? Does it vary based on the search query? The bottom line is that the search engine might not favor the version you want.

Although some SEO professionals may debate some of the preceding specifics, the general points will meet with near-universal agreement. However, there are a handful of caveats to take into account.

For one, on your site you may have a variety of product pages and also offer print versions of those pages. The search engine might pick just the printer-friendly page as the one to show in its results. This does happen at times, and it can happen even if the printer-friendly page has lower link authority and will rank less well than the main product page.

The best potential fix for this is to apply the `rel="canonical"` link element to all versions of the page to indicate which version is the original.

A second version of this can occur when you syndicate content to third parties. The problem is that the search engine may filter your copy of the article out of the results in favor of the version in use by the person republishing your article. There are three potential solutions to this:

- Get the person publishing your syndicated content to publish a `rel="canonical"` link element tag back to the original page on your site. This will help indicate to the search engines that your copy of the page is the original, and any links pointing to the syndicated page will be credited to your original instead.

- Have the syndicating partner `noindex` its copy of the content. This will keep the duplicate copy out of the search engine index. In addition, any links in that content back to your site will still pass link authority to you.

- Have the partner implement a link back to the original source page on your site. Search engines usually interpret this correctly and emphasize your version of the content when you do that. Note, however, that there have been instances where Google attributes the originality of the content to the site republishing it, particularly if that site has vastly more authority and trust than the true original source of the content.

How Search Engines Identify Duplicate Content

Some examples will illustrate the process for Google as it finds duplicate content on the Web. In the examples shown in Figure 6-28 through Figure 6-31, three assumptions have been made:

- The page with text is assumed to be a page that contains duplicate content (not just a snippet, despite the illustration).

- Each page of duplicate content is presumed to be on a separate domain.

- The steps that follow have been simplified to make the process as easy and clear as possible. This is almost certainly not the exact way in which Google performs (but it conveys the effect).

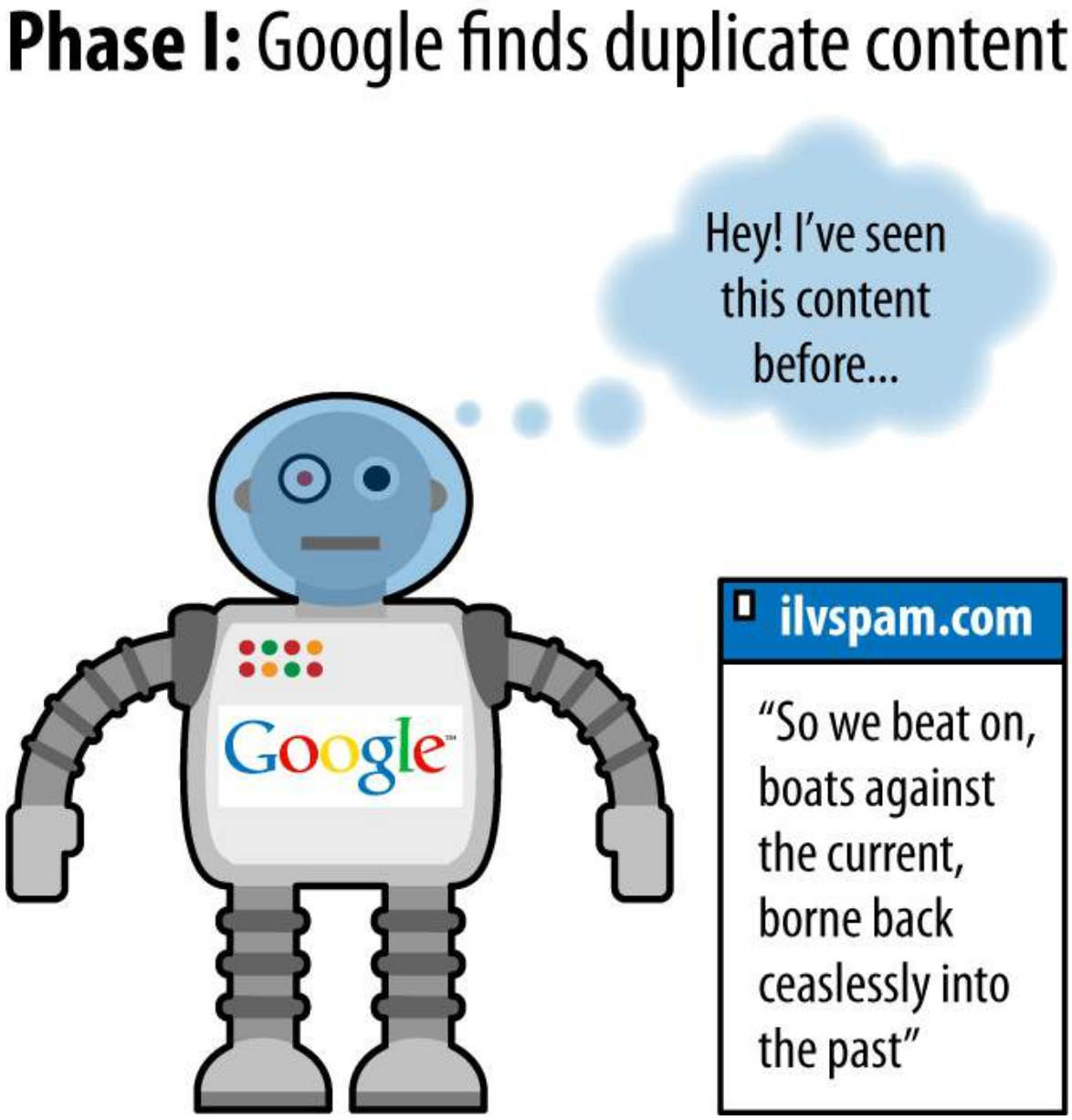


Figure 6-28. Google finding duplicate content

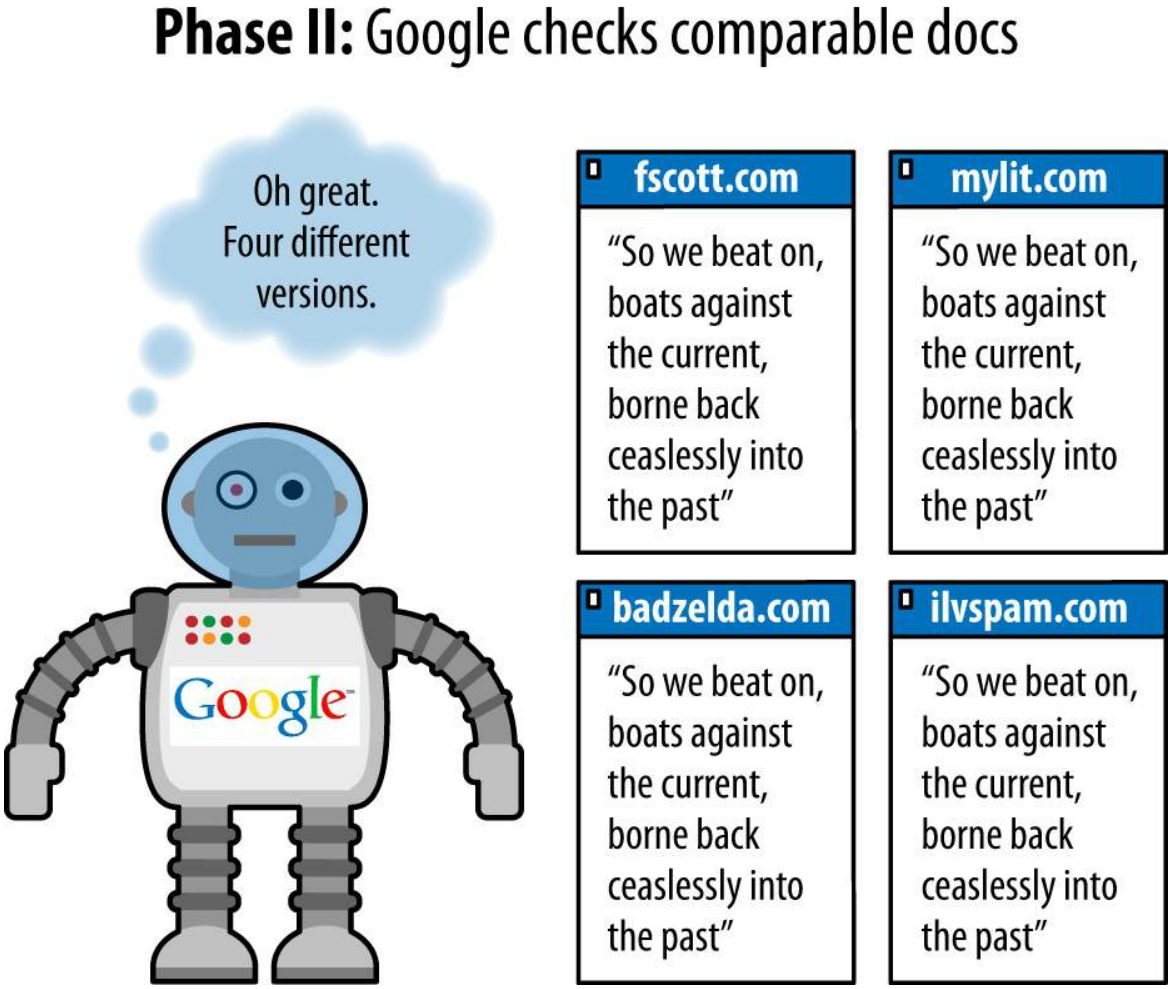


Figure 6-29. Google comparing the duplicate content to the other copies

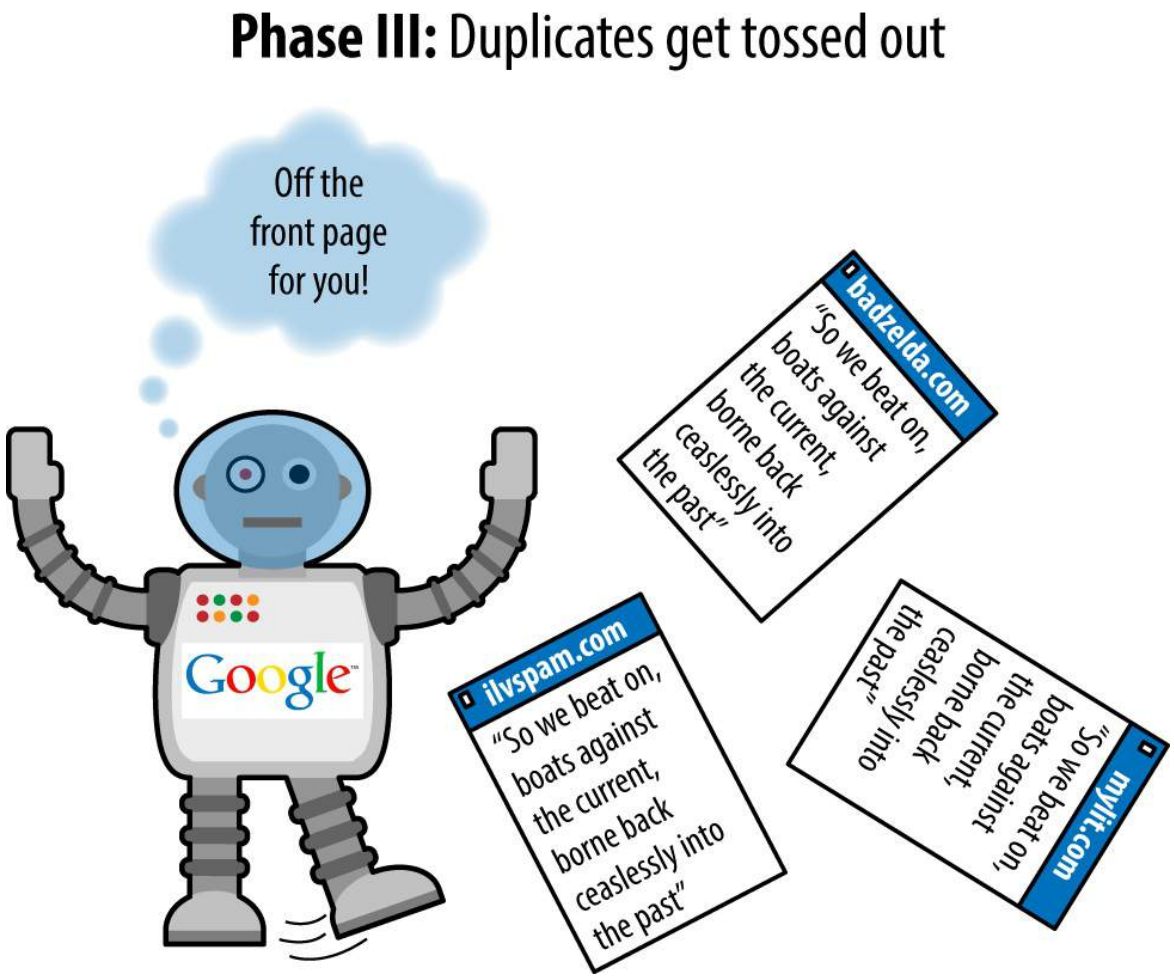


Figure 6-30. Duplicate copies getting filtered out

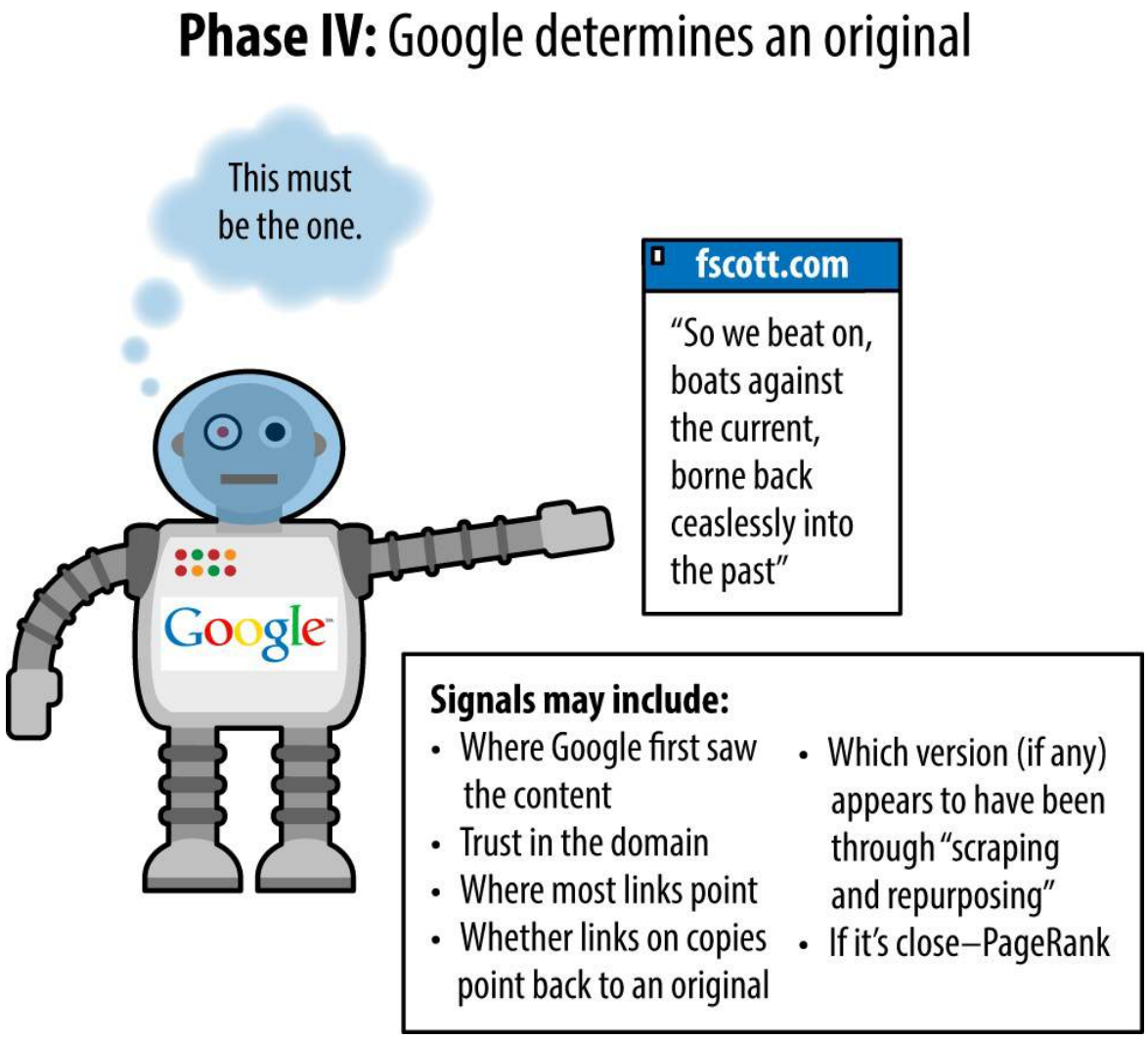


Figure 6-31. Google choosing one as the original

There are a few facts about duplicate content that bear mentioning, as they can trip up webmasters who are new to the duplicate content issue:

Location of the duplicate content

Is it duplicated content if it is all on my site? Yes, in fact, duplicate content can occur within a site or across different sites.

Percentage of duplicate content

What percentage of a page has to be duplicated before I run into duplicate content filtering? Unfortunately, the search engines would never reveal this information because it would compromise their ability to prevent

the problem.

It is also a near certainty that the percentage at each engine fluctuates regularly and that more than one simple direct comparison goes into duplicate content detection. The bottom line is that pages do not need to be identical to be considered duplicates.

Ratio of code to text

What if my code is huge and there are very few unique HTML elements on the page? Will Google think the pages are all duplicates of one another? No. The search engines do not care about your code; they are interested in the content on your page. Code size becomes a problem only when it becomes extreme.

Ratio of navigation elements to unique content

Every page on my site has a huge navigation bar, lots of header and footer items, but only a little bit of content; will Google think these pages are duplicates? No. Google and Bing factor out the common page elements, such as navigation, before evaluating whether a page is a duplicate. They are very familiar with the layout of websites and recognize that permanent structures on all (or many) of a site's pages are quite normal. Instead, they'll pay attention to the "unique" portions of each page and often will largely ignore the rest. Note, however, that these will almost certainly be considered thin content by the engines.

Licensed content

What should I do if I want to avoid duplicate content problems, but I have licensed content from other web sources to show my visitors? Use `meta name = "robots" content="noindex, follow"`. Place this in your page's header and the search engines will know that the content isn't for them. This is a general best practice, because then humans can still visit and link to the page, and the links on the page will still carry value.

Another alternative is to make sure you have exclusive ownership and publication rights for that content.

Copyright Infringement

One of the best ways to monitor whether your site's copy is being duplicated elsewhere is to use CopyScape, a site that enables you to instantly view pages on the Web that are using your content. Do not worry if the pages of these sites rank far behind your own pages for any relevant queries—if any large, authoritative, content-rich domain tried to fight all the copies of its work on the Web, it would have at least two full-time jobs on its hands. Luckily, the search engines have placed trust in these types of sites to issue high-quality, relevant content, and therefore recognize them as the original issuer.

If, on the other hand, you have a relatively new site or a site with few inbound links, and the scrapers are consistently ranking ahead of you (or someone with a powerful site is stealing your work), you've got some recourse. One option is just to ask the publisher to remove the offending content. In some cases, the publisher is simply unaware that copying your content is not allowed. Another option is to contact the site's hosting company. Hosting companies could potentially be liable for hosting duplicate content, so they are frequently quick to react to such inquiries. Just be sure to provide as much documentation as possible to show that the content was originally yours.

Another option is to file a DMCA infringement request with Google, Yahoo!, and Bing (you should also file this request with the infringing site's hosting company).

A further option is to file a legal suit (or threaten such) against the website in question. You may want to try to start with a more informal communication asking the publisher to remove the content before you send a letter from the attorneys, as the DMCA motions can take up to several months to go into effect; but if the publisher is nonresponsive, there is no reason to delay taking stronger action, either. If the site republishing your work has an owner in your country, this latter course of action is probably the most effective first step.

A very effective and inexpensive option for this process is DMCA.com.

An actual penalty situation

The preceding examples show duplicate content filters and are not actual penalties, but, for all practical purposes, they have the same impact as a penalty:

lower rankings for your pages. But there are scenarios where an actual penalty can occur.

For example, sites that aggregate content from across the Web can be at risk, particularly if little unique content is added from the site itself. In this type of scenario, you might see the site actually penalized.

If you find yourself in this situation, the only fixes are to reduce the number of duplicate pages accessible to the search engine crawler. You can accomplish this by deleting them, using `canonical` on the duplicates, `noindex`-ing the pages themselves, or adding a substantial amount of unique content.

One example of duplicate content that may get filtered out on a broad basis is a *thin affiliate* site. This nomenclature frequently describes a site promoting the sale of someone else's products (to earn a commission), yet provides little or no information differentiated from other sites selling the product. Such a site may have received the descriptions from the manufacturer of the products and simply replicated those descriptions along with an affiliate link (so that it can earn credit when a click or purchase is performed).

The problem arises when a merchant has thousands of affiliates generally promoting websites using the same descriptive content, and search engineers have observed user data suggesting that, from a searcher's perspective, these sites add little value to their indexes. Thus, the search engines attempt to filter out this type of site, or even ban it from their index. Plenty of sites operate affiliate models but also provide rich new content, and these sites generally have no problem; it is when duplication of content and a lack of unique, value-adding material come together on a domain that the engines may take action.

How to Avoid Duplicate Content on Your Own Site

As we outlined, duplicate content can be created in many ways. Internal duplication of material requires specific tactics to achieve the best possible results from an SEO perspective. In many cases, the duplicate pages are pages that have no value to either users or search engines. If that is the case, try to eliminate the problem altogether by fixing the implementation so that all pages are referred to by only one URL. Also, 301-redirect (these are discussed in more detail in "Redirects") the old URLs to the surviving URLs to help the search engines discover what you have done as rapidly as possible, and preserve any

link authority the removed pages may have had.

If that process proves to be impossible, there are many options, as we will outline in "Content Delivery and Search Spider Control". Here is a summary of the guidelines on the simplest solutions for dealing with a variety of scenarios:

- Use *robots.txt* to block search engine spiders from crawling the duplicate versions of pages on your site.
- Use the `rel="canonical"` link element. This is the next best solution to eliminating the duplicate pages.
- Use `<meta name="robots" content="noindex">` to tell the search engine to not index the duplicate pages.

Be aware, however, that if you use *robots.txt* to prevent a page from being crawled, then using `noindex` or `nofollow` on the page itself does not make sense —the spider can't read the page, so it will never see the `noindex` or `nofollow`. With these tools in mind, here are some specific duplicate content scenarios:

HTTPS pages

If you make use of *SSL* (encrypted communications between the browser and the web server), and you have not converted your entire site, you will have some pages on your site that begin with *https:* instead of *http:*. The problem arises when the links on your *https:* pages link back to other pages on the site using relative instead of absolute links, so (for example) the link to your home page becomes *https://www.yourdomain.com* instead of *http://www.yourdomain.com*.

If you have this type of issue on your site, you may want to use the `re="canonical"` link element, which we describe in "Content Delivery and Search Spider Control", or 301 redirects to resolve problems with these types of pages. An alternative solution is to change the links to absolute links (*http://www.yourdomain.com/content* instead of */content*), which also makes life more difficult for content thieves that scrape your site.

A CMS that creates duplicate content

Sometimes sites have many versions of identical pages because of limitations in the CMS where it addresses the same content with more than one URL. These are often unnecessary duplications with no end-user value, and the best practice is to figure out how to eliminate the duplicate pages and 301 the eliminated pages to the surviving pages. Failing that, fall back on the other options listed at the beginning of this section.

Print pages or multiple sort orders

Many sites offer print pages to provide the user with the same content in a more printer-friendly format. Or some ecommerce sites offer their products in multiple sort orders (such as size, color, brand, and price). These pages do have end-user value, but they do not have value to the search engine and will appear to be duplicate content. For that reason, use one of the options listed previously in this subsection, or set up a print CSS stylesheet such as the one outlined in this post by Yoast.

Duplicate content in blogs and multiple archiving systems (e.g., pagination)

Blogs present some interesting duplicate content challenges. Blog posts can appear on many different pages, such as the home page of the blog, the permalink page for the post, date archive pages, and category pages. Each instance of the post represents duplicates of the other instances. Few publishers attempt to address the presence of the post on the home page of the blog and also at its permalink, and this is common enough that the search engines likely deal reasonably well with it. However, it may make sense to show only excerpts of the post on the category and/or date archive pages.

User-generated duplicate content (e.g., repostings)

Many sites implement structures for obtaining user-generated content, such as a blog, forum, or job board. This can be a great way to develop large quantities of content at a very low cost. The challenge is that users may choose to submit the same content on your site and in several other sites at the same time, resulting in duplicate content among those sites. It is hard to control this, but there are two things you can do to mitigate the problem:

- Have clear policies that notify users that the content they submit to your

site must be unique and cannot be, or cannot have been, posted to other sites. This is difficult to enforce, of course, but it will still help some to communicate your expectations.

- Implement your forum in a different and unique way that demands different content. Instead of having only the standard fields for entering data, include fields that are likely to be unique over what other sites do, but that will still be interesting and valuable for site visitors to see.

Controlling Content with Cookies and Session IDs

Sometimes you want to more carefully dictate what a search engine robot sees when it visits your site. In general, search engine representatives refer to the practice of showing different content to users than to crawlers as *cloaking*, which violates the engines' Terms of Service (TOS) and is considered spam.

However, there are legitimate uses for this practice that are not deceptive to the search engines or malicious in intent. This section will explore methods for controlling content with cookies and sessions IDs.

What's a Cookie?

A *cookie* is a small text file that websites can leave on a visitor's hard disk, helping them to track that person over time. Cookies are the reason Amazon remembers your username between visits and the reason you don't necessarily need to log in to your Gmail account every time you open your browser. Cookie data typically contains a short set of information regarding when you last accessed a site, an ID number, and potentially, information about your visit (see Figure 6-32).

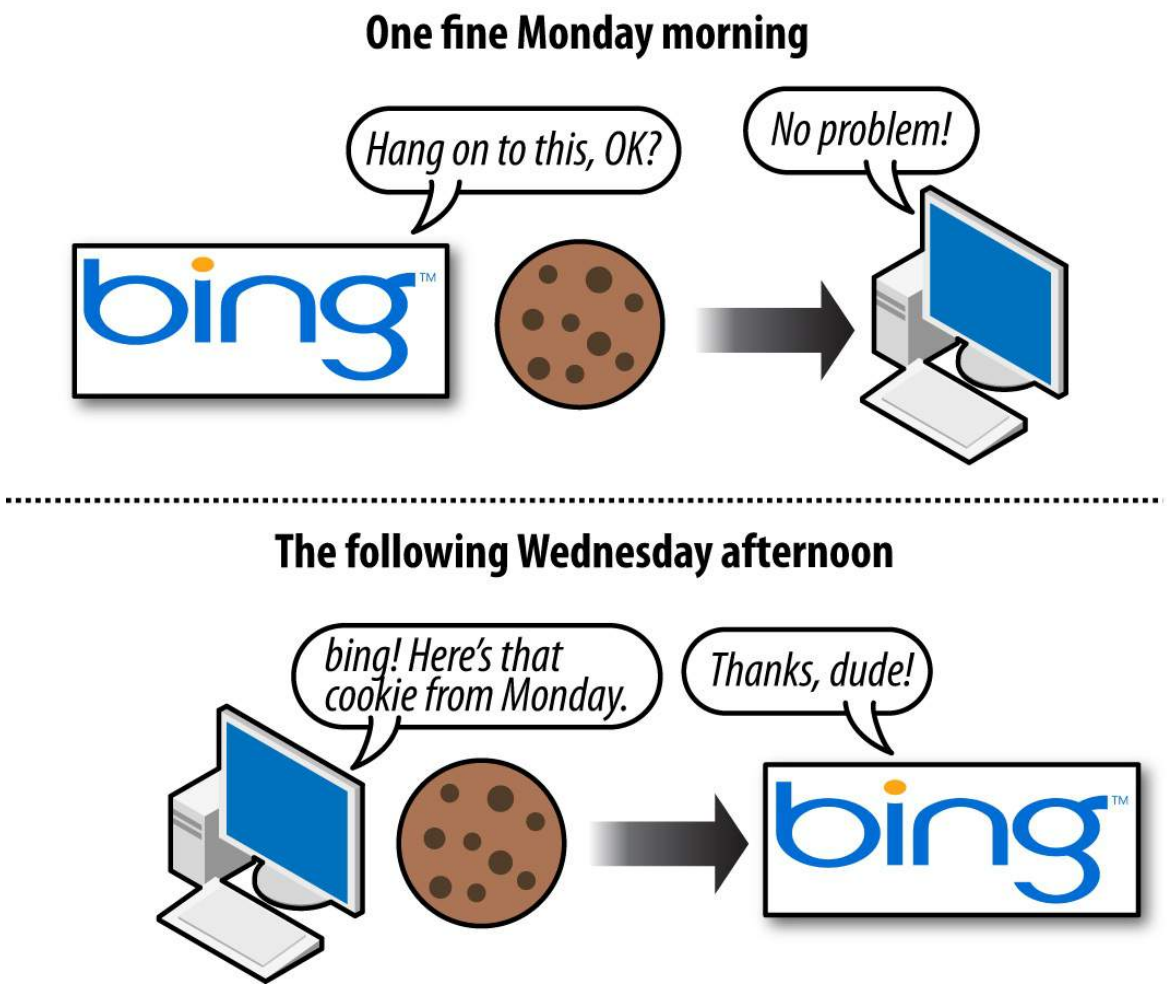


Figure 6-32. Using cookies to store data

Website developers use cookies for tracking purposes or to display different information to users based on their actions or preferences. Common uses include remembering a username, maintaining a shopping cart, and keeping track of previously viewed content. For example, if you've signed up for an account with Moz, it will provide you with options on your My Account page about how you want to view the blog and will remember those settings the next time you visit.

What Are Session IDs?

Session IDs are virtually identical to cookies in functionality, with one big difference. When you close your browser (or restart), session ID information is no longer stored on your hard drive (usually); see Figure 6-33. The website you were interacting with may remember your data or actions, but it cannot retrieve

session IDs from your machine that don't persist (and session IDs by default expire when the browser shuts down). In essence, session IDs are more like temporary cookies (although, as you'll see shortly, there are options to control this).

Figure 6-33. How session IDs are used

Although technically speaking session IDs are just a form of cookie without an expiration date, it is possible to set session IDs with expiration dates similar to

cookies (going out decades). In this sense, they are virtually identical to cookies. Session IDs do come with an important caveat, though: they are frequently passed in the URL string, which can create serious problems for search engines (as every request produces a unique URL with duplicate content).

It is highly desirable to eliminate session IDs from your URLs, and you should avoid them if it is at all possible. If you currently have them, a short-term fix is to use the `rel="canonical"` link element (which we'll discuss in "Content Delivery and Search Spider Control") to tell the search engines that you want them to ignore the session IDs.

NOTE

Any user has the ability to turn off cookies in his browser settings. This often makes web browsing considerably more difficult, and many sites will actually display a page saying that cookies are required to view or interact with their content. Cookies, persistent though they may be, are also deleted by users on a semiregular basis. For example, a 2011 comScore study found that 33% of web users deleted their first-party cookies at least once per month.

How Do Search Engines Interpret Cookies and Session IDs?

Search engine spiders do not look at cookies or session IDs, and act as browsers with this functionality shut off. However, unlike visitors whose browsers won't accept cookies, the crawlers can sometimes reach sequestered content by virtue of webmasters who want to specifically let them through. Many sites have pages that require cookies or sessions to be enabled but have special rules for search engine bots, permitting them to access the content as well. Although this is technically cloaking, there is a form of this known as *First Click Free* that search engines generally allow (we will discuss this in more detail in "Content Delivery and Search Spider Control").

Despite the occasional access engines are granted to cookie/session-restricted pages, the vast majority of cookie and session ID usage creates content, links, and pages that limit access. Web developers can leverage the power of options such as First Click Free to build more intelligent sites and pages that function in optimal ways for both humans and engines.

Why Would You Want to Use Cookies or Session IDs to Control Search Engine Access?

There are numerous potential tactics to leverage cookies and session IDs for search engine control. Here are many of the major strategies you can implement with these tools, but there are certainly limitless other possibilities:

Show multiple navigation paths while controlling the flow of link authority

Visitors to a website often have multiple ways in which they'd like to view or access content. Your site may benefit from offering many paths to reaching content (by date, topic, tag, relationship, ratings, etc.), but doing so expends PageRank or link authority that would be better optimized by focusing on a single, search engine–friendly navigational structure. This is important because these varied sort orders may be seen as duplicate content.

You can require a cookie for users to access the alternative sort order versions of a page, and prevent the search engine from indexing multiple pages with the same content. One alternative (but not foolproof) solution is to use the `rel="canonical"` link element to tell the search engine that these alternative sort orders are really just the same content as the original page (we will discuss `canonical` in "Content Delivery and Search Spider Control").

Keep limited pieces of a page's content out of the engines' indexes

Many pages may contain content that you'd like to show to search engines and other pieces you'd prefer appear only for human visitors. These could include ads, login-restricted information, links, or even rich media. Once again, showing noncookied users the plain version and cookie-accepting visitors the extended information can be invaluable. Note that this option is often used in conjunction with a login, so only registered users can access the full content (such as on sites like Facebook and LinkedIn).

Grant access to pages requiring a login

As with snippets of content, there are often entire pages or sections of a site on which you'd like to restrict search engine access. This can be easy to accomplish with cookies/sessions, and it can even help to bring in search

traffic that may convert to “registered user” status. For example, if you had desirable content that you wished to restrict, you could create a page with a short snippet and an offer for the visitor to continue reading upon registration, which would then allow him access to that work at the same URL. We will discuss this more in “Content Delivery and Search Spider Control”.

Avoid duplicate content issues

One of the most promising areas for cookie/session use is to prohibit spiders from reaching multiple versions of the same content, while allowing visitors to get the version they prefer. As an example, at Moz, logged-in users can see full blog entries on the blog home page, but search engines and nonregistered users will see only the excerpts. This prevents the content from being listed on multiple pages (the blog home page and the specific post pages) and provides a richer user experience for members.

Content Delivery and Search Spider Control

On occasion, it can be valuable to show search engines one version of content and show humans a different version. As we’ve discussed, this is technically called cloaking, and the search engines’ guidelines have near-universal policies restricting it. In practice, many websites, large and small, appear to use content delivery effectively and without being penalized by the search engines. However, use great care if you implement these techniques, and know the risks that you are taking.

Cloaking and Segmenting Content Delivery

Before we discuss the risks and potential benefits of cloaking-based practices, take a look at Figure 6-34, which illustrates how cloaking works.

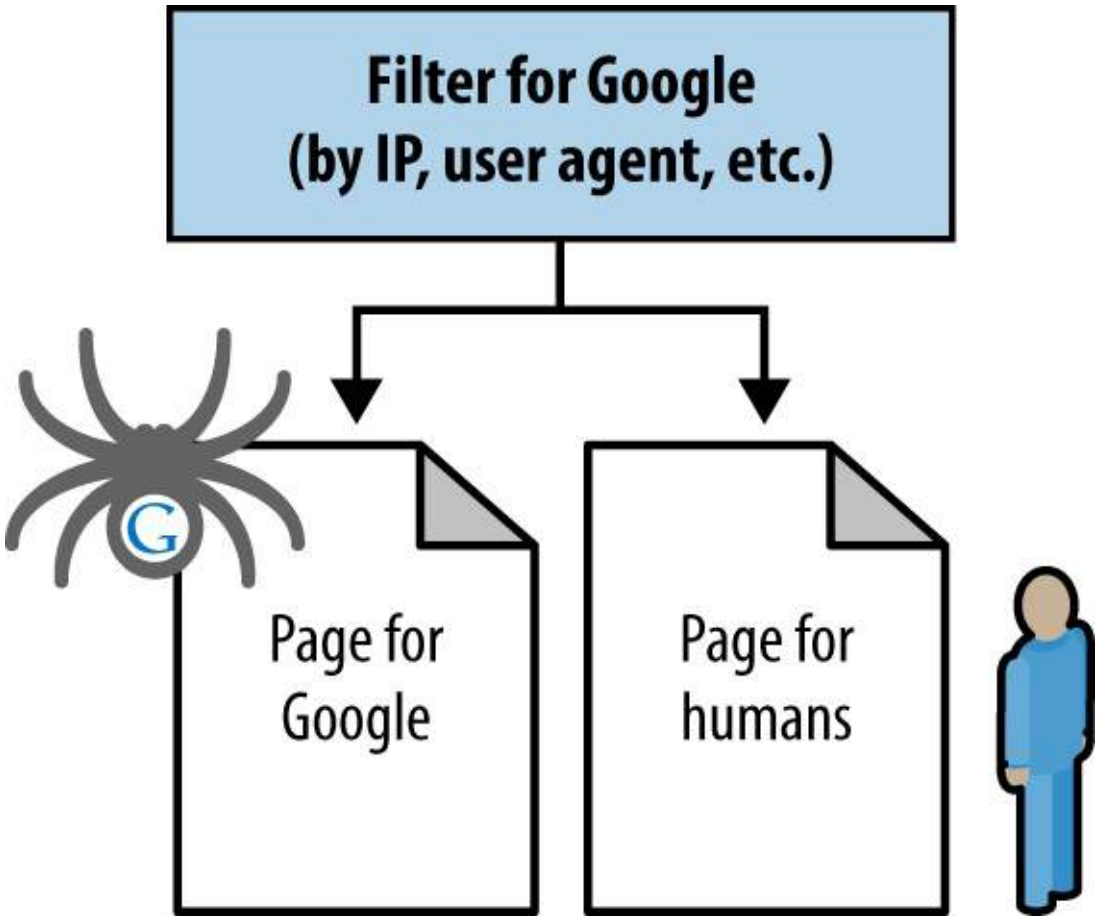


Figure 6-34. How cloaking works

Google's Matt Cutts, former head of Google's webspam team, has made strong public statements indicating that all forms of cloaking (with the only exception being First Click Free) are subject to penalty. This was largely backed by statements from Google's John Mueller in a May 2009 interview.[12] In August 2011, Matt Cutts later confirmed this again in a YouTube video, in which he asserted, "There is no such thing as white hat cloaking."[13]

Google also makes its policy pretty clear in its guidelines on cloaking (*https://support.google.com/webmasters/answer/66355*):

> *Serving up different results based on user agent may cause your site to be perceived as deceptive and removed from the Google index.*

There are two critical pieces in the preceding quote: *may* and *user agent*. It is true that if you cloak in the wrong ways, with the wrong intent, Google and the other search engines *may* remove you from their index, and if you do it

egregiously, they certainly *will*.

A big factor is intent: if the engines feel you are attempting to manipulate their rankings or results through cloaking, they may take adverse action against your site. If, however, the intent of your content delivery doesn't interfere with their goals, you're less likely to be subject to a penalty, but there is never zero risk of a penalty. Google has taken a strong stand against all forms of cloaking regardless of intent.

What follows are some examples of websites that perform some level of cloaking:

Google

Search for *google toolbar* or *google translate* or *adwords* or any number of Google properties, and note how the URL you see in the search results and the one you land on almost never match. What's more, on many of these pages, whether you're logged in or not, you might see some content that is different from what's in the cache.

New York Times

The interstitial ads, the request to log in/create an account after five clicks, and the archive inclusion are all showing different content to engines versus humans.

Wine.com

In addition to some redirection based on your path, there's the state overlay forcing you to select a shipping location prior to seeing any prices (or any pages). That's a form the engines don't have to fill out.

Yelp

Geotargeting through cookies based on location is a very popular form of local targeting that hundreds, if not thousands, of sites use.

Trulia

Trulia was found to be doing some interesting redirects on partner pages and

its own site (*http://bit.ly/trulias_integrity*).

The message should be clear: cloaking won't always get you banned, and you can do some pretty smart things with it. Again, the key to all of this is your intent. If you are doing it for reasons that are not deceptive and that provide a positive experience for users and search engines, you might not run into problems. However, there is no guarantee of this, so use these types of techniques with great care, and know that you may still get penalized for it.

Showing Different Content to Engines and Visitors

There are a few common causes for displaying content differently to different visitors, including search engines:

Multivariate and A/B split testing

Testing landing pages for conversions requires that you show different content to different visitors to test performance. In these cases, it is best to display the content using JavaScript/cookies/sessions and give the search engines a single, canonical version of the page that doesn't change with every new spidering (though this won't necessarily hurt you). Google previously offered software called Google Website Optimizer to perform this function, but it has been discontinued and replaced with Google Analytics Content Experiments. If you have used Google Website Optimizer in the past, Google recommends removing the associated tags from your site pages.

Content requiring registration and First Click Free

If you force users to register (paid or free) in order to view specific content pieces, it is best to keep the URL the same for both logged-in and non-logged-in users and to show a snippet (one to two paragraphs is usually enough) to non-logged-in users and search engines. If you want to display the full content to search engines, you have the option to provide some rules for content delivery, such as showing the first one to two pages of content to a new visitor without requiring registration, and then requesting registration after that grace period. This keeps your intent more honest, and you can use cookies or sessions to restrict human visitors while showing the full pieces to the engines.

In this scenario, you might also opt to participate in Google's First Click Free program, wherein websites can expose "premium" or login-restricted content to Google's spiders, as long as users who click from the engine's results are given the ability to view that first article for free. Many prominent web publishers employ this tactic, including the popular site Experts Exchange.

To be specific, to implement First Click Free, publishers must grant Googlebot (and presumably the other search engine spiders) access to all the content they want indexed, even if users normally have to log in to see the content. The user who visits the site will still need to log in, but the search engine spider will not have to do so. This will lead to the content showing up in the search engine results when applicable. However, if a user clicks on that search result, you must permit him to view the entire article (all pages of a given article if it is a multiple-page article). Once the user clicks to look at another article on your site, you can still require him to log in. Publishers can also limit the number of free accesses a user gets using this technique to five articles per day.

For more details, visit Google's First Click Free program pages at *http://googlewebmastercentral.blogspot.com/2008/10/first-click-free-for-web-search.html* and *http://googlewebmastercentral.blogspot.com/2009/12/changes-in-first-click-free.html*.

Navigation unspiderable by search engines

If your navigation is in Flash, JavaScript, a Java application, or another format where the search engine's ability to parse it is uncertain, you should consider showing search engines a version that has spiderable, crawlable content in HTML. Many sites do this simply with CSS layers, displaying a human-visible, search-invisible layer and a layer for the engines (and less capable browsers, such as mobile browsers). You can also employ the `<noscript>` tag for this purpose, although it is generally riskier, as many spammers have applied `<noscript>` as a way to hide content. Make sure the content shown in the search-visible layer is substantially the same as it is in the human-visible layer.

Duplicate content

If a significant portion of a page's content is duplicated, you might consider restricting spider access to it by placing it in an iframe that's restricted by *robots.txt*. This ensures that you can show the engines the unique portion of your pages, while protecting against duplicate content problems. We will discuss this in more detail in the next section.

Different content for different users

At times you might target content uniquely to users from different geographies (such as different product offerings that are more popular in their area), users with different screen resolutions (to make the content fit their screen size better), or users who entered your site from different navigation points. In these instances, it is best to have a "default" version of content that's shown to users who don't exhibit these traits to show to search engines as well.

Displaying Different Content to Search Engines Versus Visitors

There are a variety of strategies to segment content delivery. The most basic is to serve content that is not meant for the engines in unspiderable formats (e.g., placing text in images, Flash files, plugins, etc.). You should not use these formats for the purpose of cloaking; use them only if they bring a substantial end-user benefit (such as an improved user experience). In such cases, you may want to show the search engines the same content in a spiderable format. When you're trying to show the engines something you don't want visitors to see, you can use CSS formatting styles (preferably not `display:none`, as the engines have filters to watch specifically for this); JavaScript-, user agent–, cookie-, or session-based delivery; or IP delivery (showing content based on the visitor's IP address).

Be very wary when employing these strategies. As noted previously, the search engines expressly prohibit cloaking practices in their guidelines, and though there may be some leeway based on intent and user experience (e.g., your site is using cloaking to improve the quality of the user's experience, not to game the

search engines), the engines take these tactics seriously and may penalize or ban sites that implement them inappropriately or with the intention of manipulation. In addition, even if your intent is good, the search engines may not see it that way and penalize you anyway.

Leveraging the robots.txt file

This file is located on the root level of your domain (e.g., *http://www.yourdomain.com/robots.txt*), and it is a highly versatile tool for controlling what the spiders are permitted to access on your site. You can use *robots.txt* to:

- Prevent crawlers from accessing nonpublic parts of your website
- Block search engines from accessing index scripts, utilities, or other types of code
- Avoid the indexation of duplicate content on a website, such as print versions of HTML pages, or various sort orders for product catalogs
- Autodiscover XML Sitemaps

The *robots.txt* file must reside in the root directory, and the filename must be entirely in lowercase (*robots.txt*, not *Robots.txt* or any other variation that includes uppercase letters). Any other name or location will not be seen as valid by the search engines. The file must also be entirely in text format (not in HTML format).

When you tell a search engine robot not to access a page, it prevents it from crawling the page. Figure 6-35 illustrates what happens when the search engine robot sees a directive in *robots.txt* not to crawl a web page.

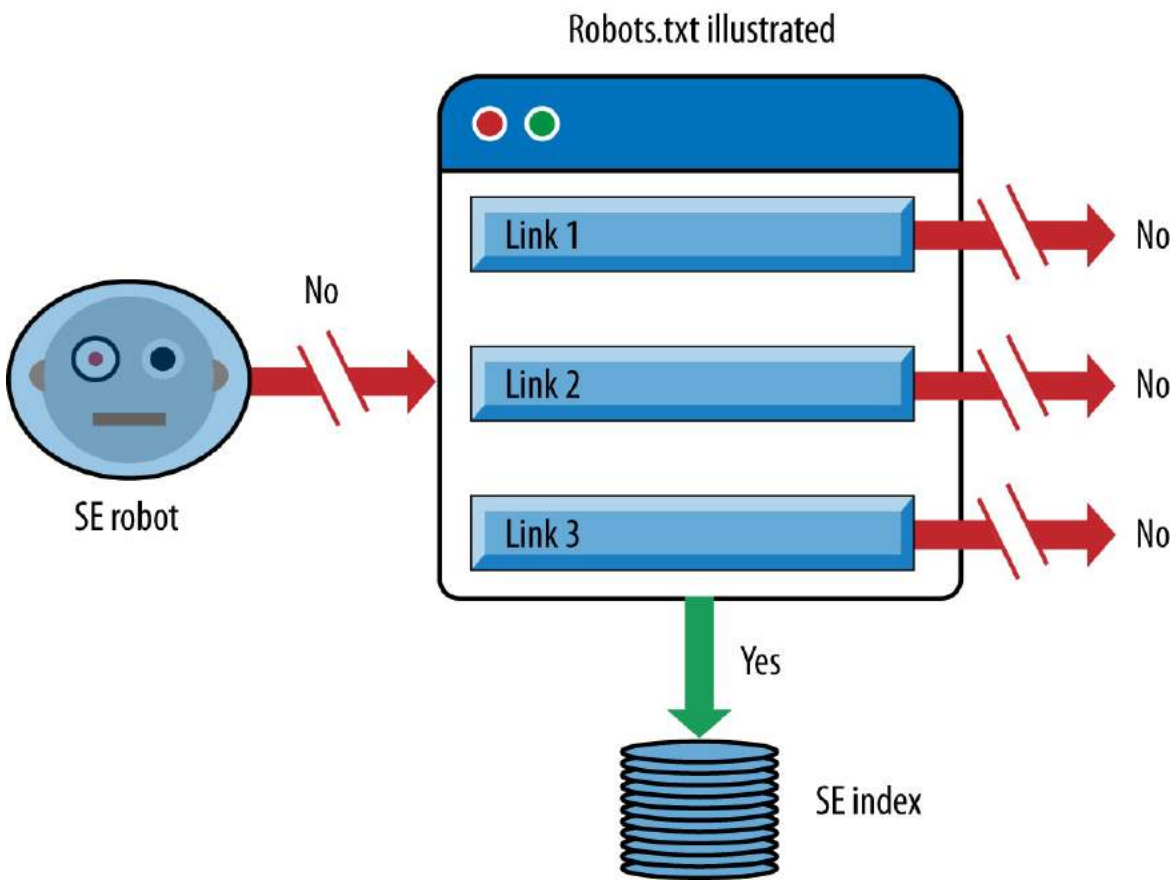


Figure 6-35. Impact of robots.txt

In essence, the page will not be crawled, so links on the page cannot pass link authority to other pages, because the search engine does not see the links. However, the page can be in the search engine index. This can happen if other pages on the Web link to it. Of course, the search engine will not have very much information on the page, as it cannot read it, and will rely mainly on the anchor text and other signals from the pages linking to it to determine what the page may be about. Any resulting search listings end up being pretty sparse when you see them in the Google index, as shown in Figure 6-36.

Figure 6-36 shows the results for the Google query *site:www.nytimes.com/cnet/*. This is not a normal query that a user would enter, but you can see what the results look like. Only the URL is listed, and there is no description. This is because the spiders aren't permitted to read the page to get that data. In today's algorithms, these types of pages don't rank very high because their relevance scores tend to be quite low for any normal queries.

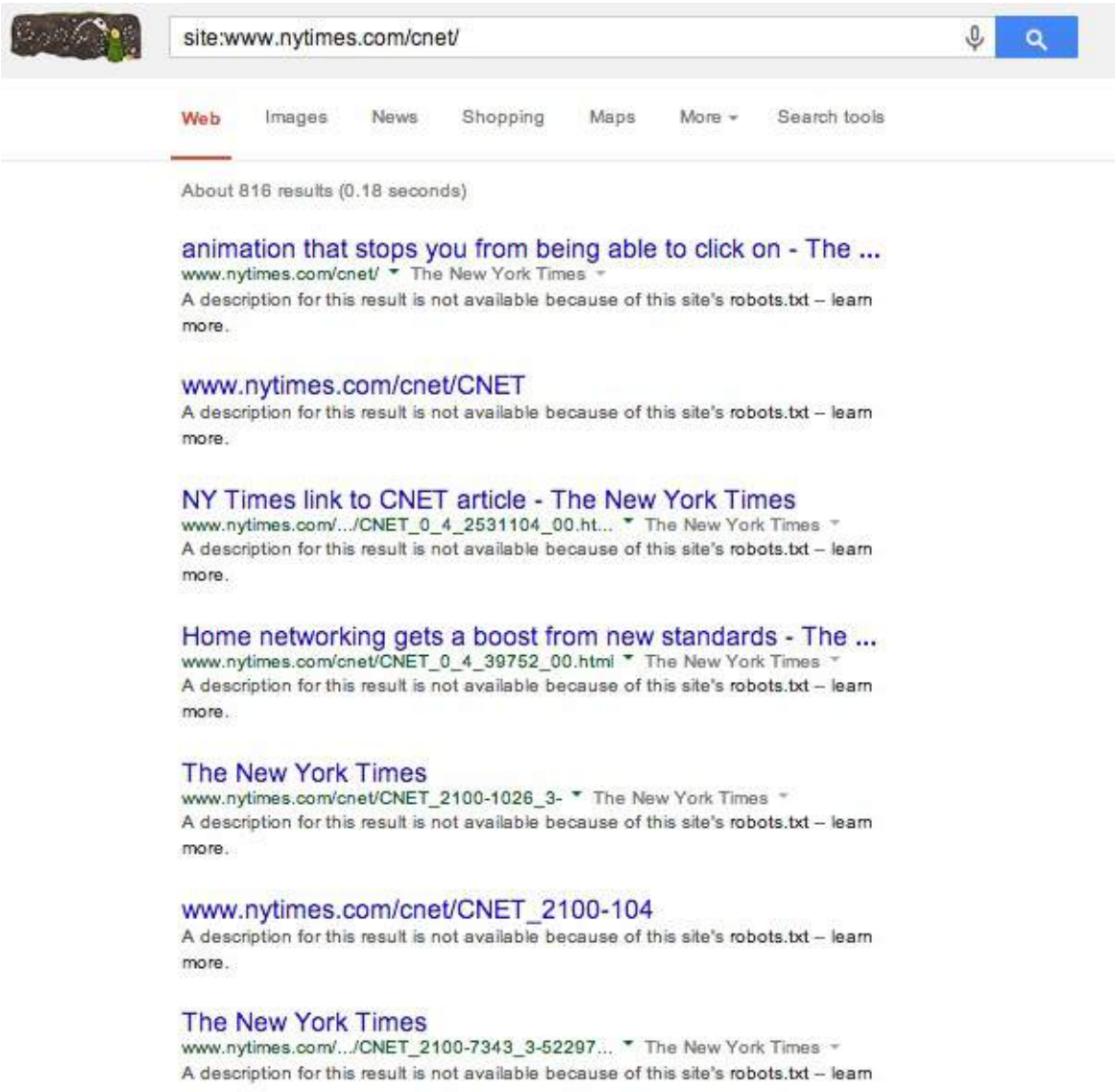


Figure 6-36. SERPs for pages that are listed in robots.txt

Google, Bing, and nearly all of the legitimate crawlers on the Web will follow the instructions you set out in the *robots.txt* file. Commands in *robots.txt* are primarily used to prevent spiders from accessing pages and subfolders on a site, though they have other options as well. Note that subdomains require their own *robots.txt* files, as do files that reside on an *https:* server.

Syntax of the robots.txt file

The basic syntax of *robots.txt* is fairly simple. You specify a robot name, such as "googlebot," and then you specify an action. The robot is identified by user

agent, and then the actions are specified on the lines that follow. The major action you can specify is `Disallow:`, which lets you indicate any pages you want to block the bots from accessing (you can use as many disallow lines as needed).

Some other restrictions apply:

- Each `User-agent`/`Disallow` group should be separated by a blank line; however, no blank lines should exist within a group (between the `User-agent` line and the last `Disallow`).
- The hash symbol (`#`) may be used for comments within a *robots.txt* file, where everything after `#` on that line will be ignored. This may be used either for whole lines or for the end of lines.
- Directories and filenames are case-sensitive: *private*, *Private*, and *PRIVATE* are all different to search engines.

Here is an example of a *robots.txt* file:

```
User-agent: Googlebot
Disallow:

User-agent: BingBot
Disallow: /

# Block all robots from tmp and logs directories
User-agent: *
Disallow: tmp
Disallow: /logs   # for directories and files called logs
```

The preceding example will do the following:

- Allow “Googlebot” to go anywhere.
- Prevent “BingBot” from crawling any part of the site.
- Block all robots (other than Googlebot) from visiting the *tmp* directory or directories or files called */logs* (e.g., */logs* or *logs.php*).

Notice that the behavior of Googlebot is not affected by instructions such as

Disallow: /. Because Googlebot has its own instructions from *robots.txt*, it will ignore directives labeled as being for all robots (i.e., those that use an asterisk).

One common problem that novice webmasters run into occurs when they have SSL installed so that their pages may be served via HTTP and HTTPS. A *robots.txt* file at *http://www.yourdomain.com/robots.txt* will not be interpreted by search engines as guiding their crawl behavior on *https://www.yourdomain.com*. To manage this, you need to create an additional *robots.txt* file at *https://www.yourdomain.com/robots.txt*. So, if you want to allow crawling of all pages served from your HTTP server and prevent crawling of all pages from your HTTPS server, you would need to implement the following:

For HTTP:

```
User-agent: *
Disallow:
```

For HTTPS:

```
User-agent: *
Disallow: /
```

These are the most basic aspects of *robots.txt* files, but there are more advanced techniques as well. Some of these methods are supported by only some of the engines, as detailed here:

Crawl delay

Crawl delay is supported by Google, Bing, and Ask. It instructs a crawler to wait the specified number of seconds between crawling pages. The goal of the directive is to reduce the load on the publisher's server:

```
User-agent: BingBot
Crawl-delay: 5
```

Pattern matching

Pattern matching appears to be usable by Google and Bing. The value of pattern matching is considerable. You can do some basic pattern matching using the asterisk wildcard character. Here is how you can use pattern

matching to block access to all subdirectories that begin with *private* (*/private1/*, *private2*, *private3*, etc.):

```
User-agent: Googlebot
Disallow: /private*/
```

You can match the end of the string using the dollar sign ($). For example, to block URLs that end with *.asp*:

```
User-agent: Googlebot
Disallow: /*.asp$
```

You may wish to prevent the robots from accessing any URLs that contain parameters. To block access to all URLs that include a question mark (?), simply use the question mark:

```
User-agent:
Disallow: /?*
```

The pattern-matching capabilities of *robots.txt* are more limited than those of programming languages such as Perl, so the question mark does not have any special meaning and can be treated like any other character.

`Allow`

The `Allow` directive appears to be supported only by Google and Ask. It works the opposite of the `Disallow` directive and provides the ability to specifically call out directories or pages that may be crawled. When this is implemented, it can partially override a previous `Disallow` directive. This may be beneficial after large sections of the site have been disallowed, or if the entire site itself has been disallowed.

Here is an example that allows Googlebot into only the *google* directory:

```
User-agent: Googlebot
Disallow:
Allow: google/
```

`Noindex`

This directive works in the same way as the `meta robots noindex` command (which we will discuss shortly) and tells the search engines to explicitly exclude a page from the index. Because a `Disallow` directive prevents crawling but not indexing, this can be a very useful feature to ensure that the pages don't show in search results. Google supports this directive in *robots.txt*, and only unofficially.

Sitemaps

We discussed XML sitemaps at the beginning of this chapter. You can use *robots.txt* to provide an autodiscovery mechanism for the spider to find the XML sitemap file. The search engines can be told to find the file with one simple line in the *robots.txt* file:

```
Sitemap: sitemap_location
```

The *`sitemap_location`* should be the complete URL to the sitemap, such as *http://www.yourdomain.com/sitemap.xml*. You can place this anywhere in your file.

For full instructions on how to apply *robots.txt*, see Martijn Koster's "A Standard for Robot Exclusion". You can also test your *robots.txt* file in Google Search Console under Crawl -> robots.txt Tester.

You should use great care when making changes to *robots.txt*. A simple typing error can, for example, suddenly tell the search engines to no longer crawl any part of your site. After updating your *robots.txt* file, it is always a good idea to check it with the Google Search Console Test Robots.txt tool. You can find this by logging in to Search Console and then selecting Crawl -> Blocked URLs.

Using the rel="nofollow" attribute

In 2005, the three major search engines—Google, Microsoft, and Yahoo! (which still had its own search engine at that time—all agreed to support an initiative intended to reduce the effectiveness of automated spam. Unlike the meta `robots` version of `nofollow`, the new directive could be employed as an attribute within

an `<a>` or link tag to indicate that the linking site “does not editorially vouch for the quality of the linked-to page.” This enables a content creator to link to a web page without passing on any of the normal search engine benefits that typically accompany a link (trust, anchor text, PageRank, etc.).

Originally, the intent was to enable blogs, forums, and other sites where user-generated links were offered to shut down the value of spammers who built crawlers that automatically created links. However, this has expanded as Google, in particular, recommends use of `nofollow` on links that are paid for—as the search engine’s preference is that only those links that are truly editorial and freely provided by publishers (without being compensated) should count toward bolstering a site’s/page’s rankings.

You can implement `nofollow` using the following format:

```
<a href="http://www.google.com" rel="nofollow">
```

Note that although you can use `nofollow` to restrict the passing of link value between web pages, the search engines still crawl through those links (despite the lack of semantic logic) and crawl the pages they link to. The search engines have provided contradictory input on this point. To summarize, `nofollow` does not expressly forbid indexing or spidering, so if you link to your own pages with it in an effort to keep them from being indexed or ranked, others may find them and link to them and your original goal will be thwarted.

Figure 6-37 shows how a search engine robot interprets a `nofollow` attribute when it finds one associated with a link (Link 1 in this example).

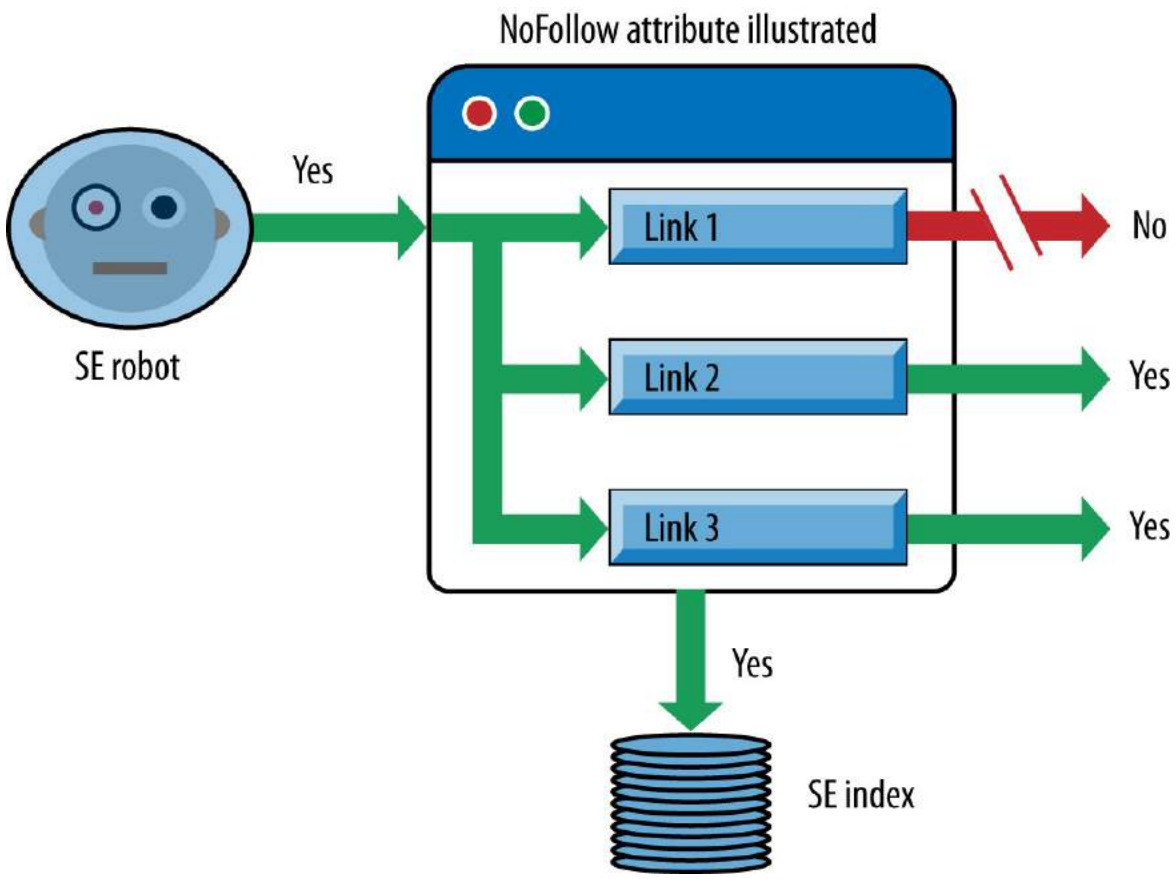


Figure 6-37. Impact of nofollow attribute

The specific link with the `nofollow` attribute was, for a number of years, considered to be disabled from passing link authority, and the notion of PageRank sculpting using `nofollow` was popular. The belief was that when you `nofollow` a particular link, the link authority that would have been passed to that link was preserved and the search engines would reallocate it to the other links found on the page. As a result, many publishers implemented `nofollow` links to lower value pages on their site (such as the About Us and Contact Us pages, or alternative sort order pages for product catalogs). In fact, data from Moz's Open Site Explorer tool, published in March 2009, showed that at that time about 3% of all links on the Web were `nofollow`ed, and that 60% of those `nofollow`s were applied to internal links.

In June 2009, however, Google's Matt Cutts wrote a post that made it clear that the link authority associated with that `nofollow`ed link is discarded rather than reallocated.[14] In theory, you can still use `nofollow` however you want, but using

it on internal links does not (at the time of this writing, according to Google) bring the type of benefit people have been looking for in the past. In fact, in certain scenarios it can be harmful.

In addition, many SEOs speculate that in some cases some value is indeed being placed on some nofollowed links, and we suggest erring on the side of caution when using this attribute, as its use has been associated with a site being "flagged" as overoptimized or otherwise aggressive in SEO tactics.

This is a great illustration of the ever-changing nature of SEO. Something that was a popular, effective tactic is now being viewed as ineffective. Some more aggressive publishers will continue to pursue PageRank sculpting by using even more aggressive approaches, such as implementing links in encoded JavaScript or within iframes that have been disallowed in *robots.txt*, so that the search engines don't see them as links. Such aggressive tactics are probably not worth the trouble for most publishers.

Using the meta robots tag

The meta `robots` tag has three components: `cache`, `index`, and `follow`. The `cache` component instructs the engine about whether it can keep the page in the engine's public index, available via the "cached snapshot" link in the search results (see Figure 6-38).

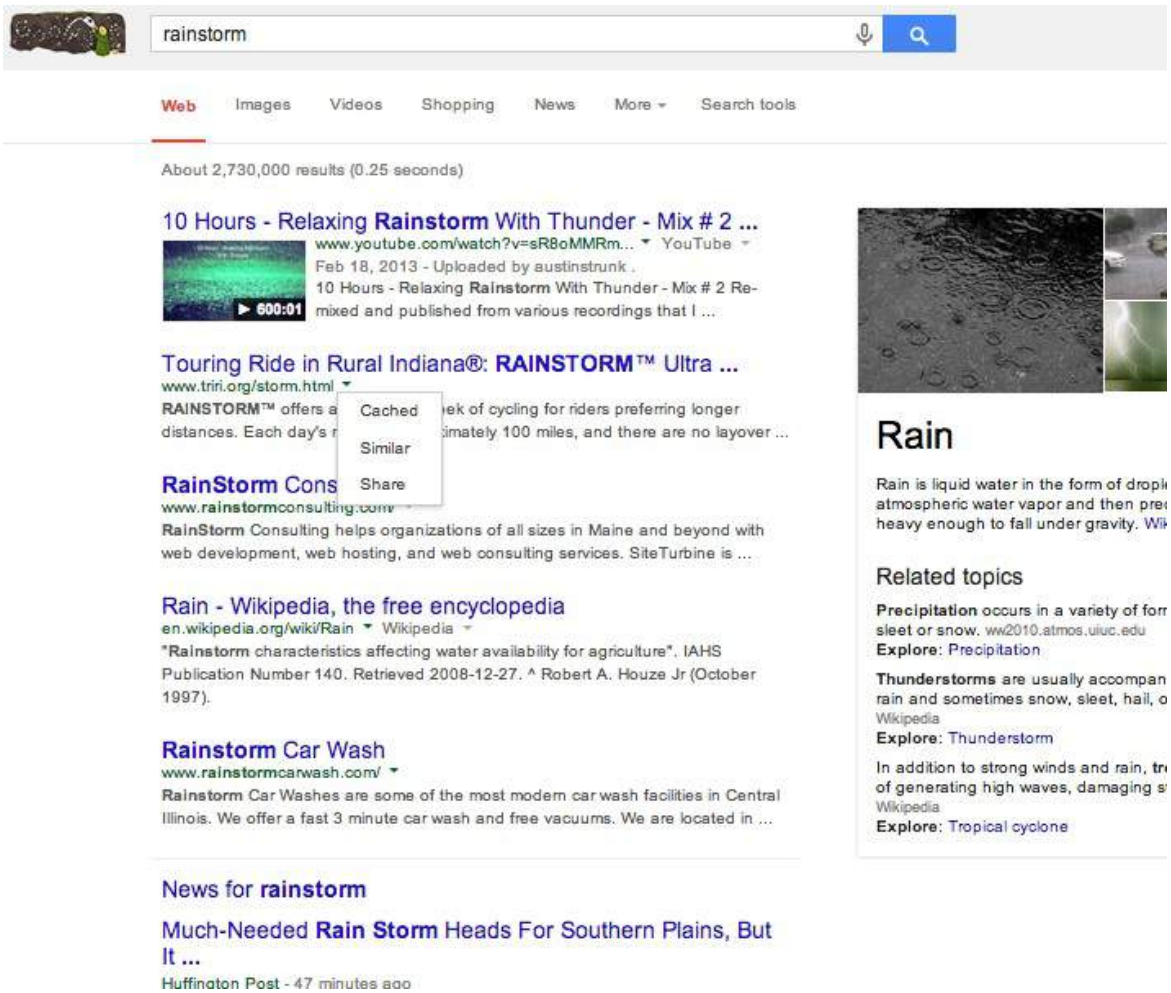


Figure 6-38. Accessing a cached page in the SERPs

The second, `index`, tells the engine that the page is allowed to be crawled and stored in any capacity. This is the default value, so it is unnecessary to place the `index` directive on each page. By contrast, a page marked `noindex` will be excluded entirely by the search engines. Figure 6-39 shows what a search engine robot does when it sees a `noindex` tag on a web page.

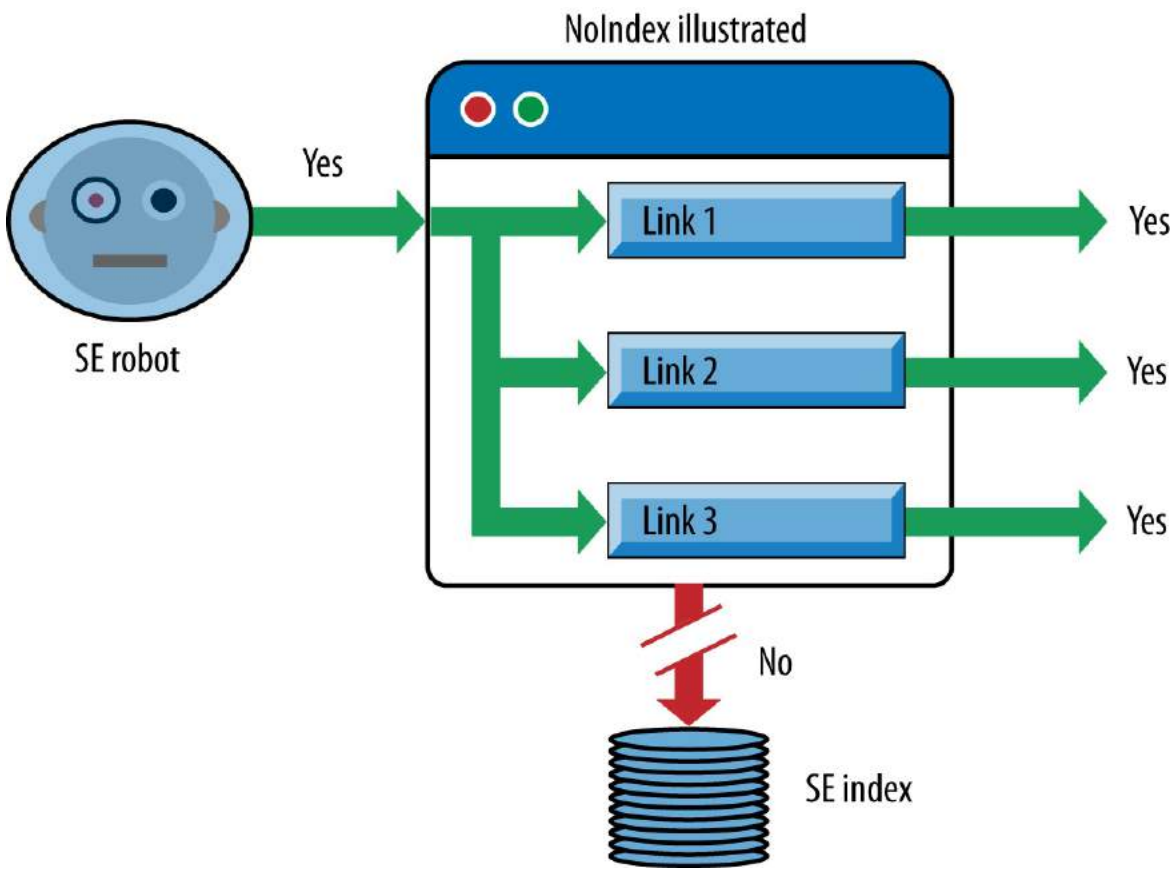


Figure 6-39. Impact of noindex

The page will still be crawled, and the page can still accumulate and pass link authority to other pages, but it will not appear in search indexes.

The final instruction available through the meta `robots` tag is `follow`. This command, like `index`, defaults to "yes, crawl the links on this page and pass link authority through them." Applying `nofollow` tells the engine that none of the links on that page should pass link value. By and large, it is unwise to use this directive as a way to prevent links from being crawled. Human beings will still reach those pages and have the ability to link to them from other sites, so `nofollow` (in the meta `robots` tag) does little to restrict crawling or spider access. Its only function is to prevent link authority from spreading out, which has very limited application since the 2005 launch of the `rel="nofollow"` attribute (discussed earlier), which allows this directive to be placed on individual links.

Figure 6-40 outlines the behavior of a search engine robot when it finds a

`nofollow` meta tag on a web page (assuming there are no other links pointing to the three linked URLs).

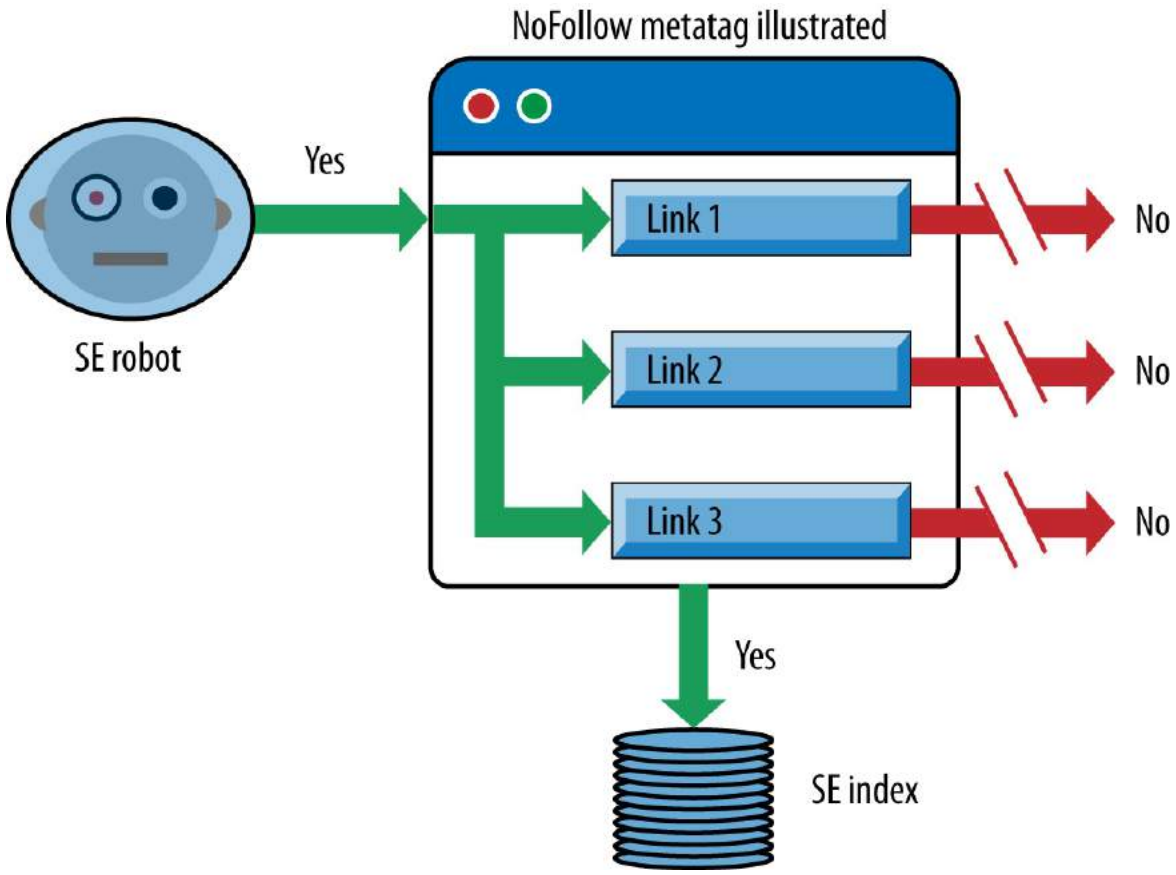


Figure 6-40. Impact of nofollow meta tag

When you use the `nofollow` meta tag on a page, the search engine will still crawl the page and place the page in its index. However, all links (both internal and external) on the page will be disabled from passing link authority to other pages.

One good application for `noindex` is to place this tag on HTML sitemap pages. These are pages designed as navigational aids for users and search engine spiders to enable them to efficiently find the content on your site. However, on some sites these pages are unlikely to rank for anything of importance in the search engines, yet you still want them to pass link authority to the pages they link to. Putting `noindex` on these pages keeps these HTML sitemaps out of the index and removes that problem. Make sure you *do not* apply the `nofollow` meta

tag on the pages or the `nofollow` attribute on the links on the pages, as these will prevent the pages from passing link authority.

Using the rel="canonical" link element

In February 2009, Google, Yahoo!, and Microsoft debuted the `rel="canonical"` link element (sometimes referrd to as the `canonical` tag). This element was a new construct designed explicitly for the purpose of identifying and dealing with duplicate content. Implementation is very simple and looks like this:

```
<link rel="canonical" href="http://moz.com/blog" />
```

This tag tells the search engines that the page in question should be treated as though it were a copy of the URL *http://moz.org/blog*, and that all of the link and content metrics the engines apply should technically flow back to that URL (see Figure 6-41).

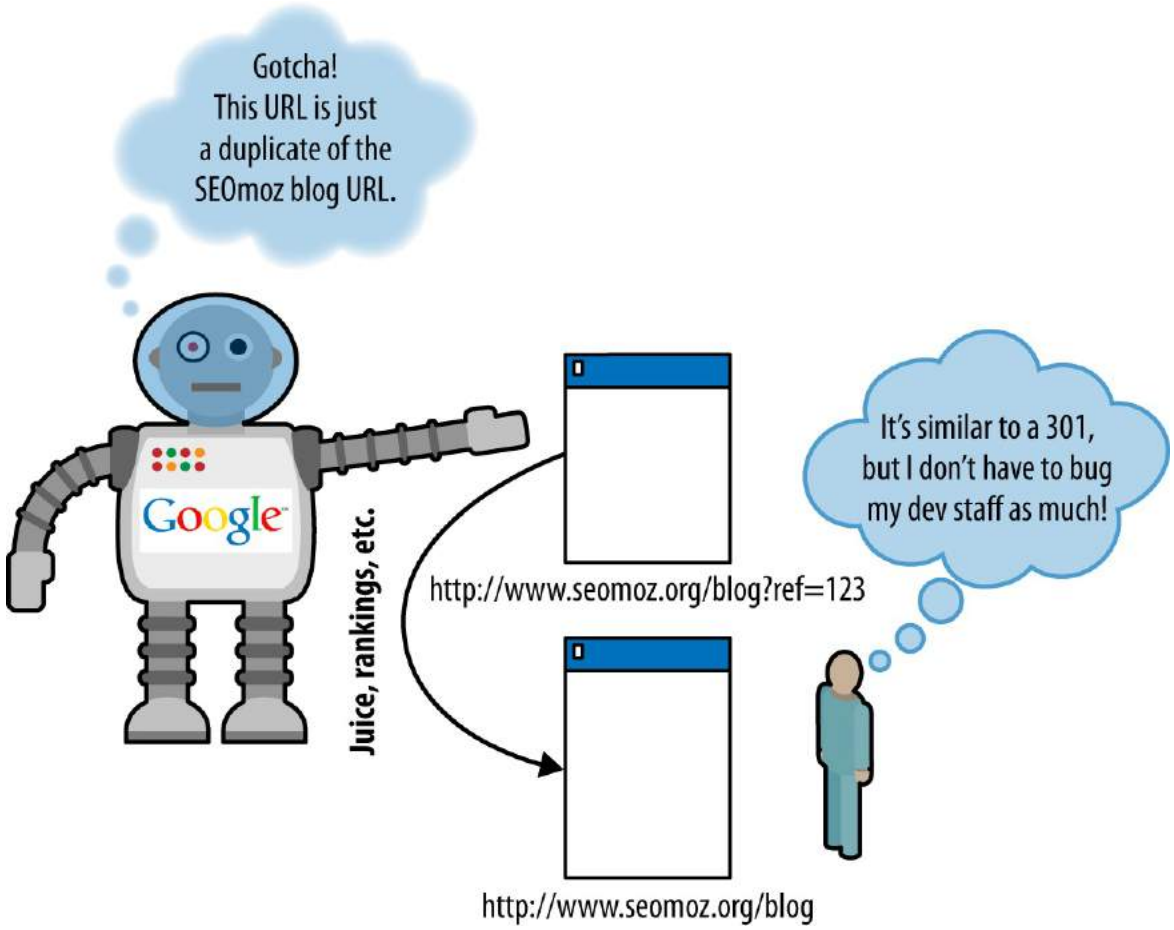


Figure 6-41. How search engines understand the canonical link element

The `rel="canonical"` link element is similar in many ways to a 301 redirect from an SEO perspective. In essence, you're telling the engines that multiple pages should be considered as one (which a 301 does), without actually redirecting visitors to the new URL (for many publishers this is less effort than some of the other solutions for their development staff). There are some differences, though:

- Whereas a 301 redirect points all traffic (bots and human visitors), `canonical` is just for engines, meaning you can still separately track visitors to the unique URL versions.

- A 301 is a much stronger signal that multiple pages have a single, canonical source. While 301s are considered a directive that search engines and browsers are obligated to honor, `canonical` is treated as a suggestion.

Although the engines generally support this new tag and trust the intent of site owners, there will be limitations. Content analysis and other algorithmic metrics will be applied to ensure that a site owner hasn't mistakenly or manipulatively applied `canonical`, and you can certainly expect to see mistaken use of it, resulting in the engines maintaining those separate URLs in their indexes (meaning site owners would experience the same problems noted in "Duplicate Content Issues").

We will discuss some applications for this tag later in this chapter. In general practice, the best solution is to resolve the duplicate content problems at their core, and eliminate them if you can. This is because the `rel="canonical"` link element is not guaranteed to work. However, it is not always possible to resolve the issues by other means, and `canonical` provides a very effective backup plan.

You can also include `canonical` directly within the HTTP response header for your page. The code might look something like the following:

```
HTTP/1.1 200 OK
Content-Type: application/pdf
Link: <http://www.example.com/white-paper.html>; rel="canonical"
Content-Length: 785710
(... rest of HTTP response headers...)
```

You can read more about this here: *http://bit.ly/canonical_headers*.

Blocking and cloaking by IP address range

You can customize entire IP addresses or ranges to block particular bots through server-side restrictions on IPs. Most of the major engines crawl from a limited number of IP ranges, making it possible to identify them and restrict access. This technique is, ironically, popular with webmasters who mistakenly assume that search engine spiders are spammers attempting to steal their content, and thus block the IP ranges to restrict access and save bandwidth. Use caution when blocking bots, and make sure you're not restricting access to a spider that could bring benefits, either from search traffic or from link attribution.

Blocking and cloaking by user agent

At the server level, it is possible to detect user agents and restrict their access to

pages or websites based on their declaration of identity. As an example, if a website detected a rogue bot, you might double-check its identity before allowing access. The search engines all use a similar protocol to verify their user agents via the Web: a reverse DNS lookup followed by a corresponding forward DNS IP lookup. An example for Google would look like this:

```
> host 66.249.66.1
1.66.249.66.in-addr.arpa domain name pointer crawl-66-249-66-1.googlebot.com.

> host crawl-66-249-66-1.googlebot.com
crawl-66-249-66-1.googlebot.com has address 66.249.66.1
```

A reverse DNS lookup by itself may be insufficient, because a spoofer could set up reverse DNS to point to *xyz.googlebot.com* or any other address.

Using iframes

Sometimes there's a certain piece of content on a web page (or a persistent piece of content throughout a site) that you'd prefer search engines didn't see. As we discussed earlier in this chapter, clever use of iframes can come in handy for this situation, as Figure 6-42 illustrates.

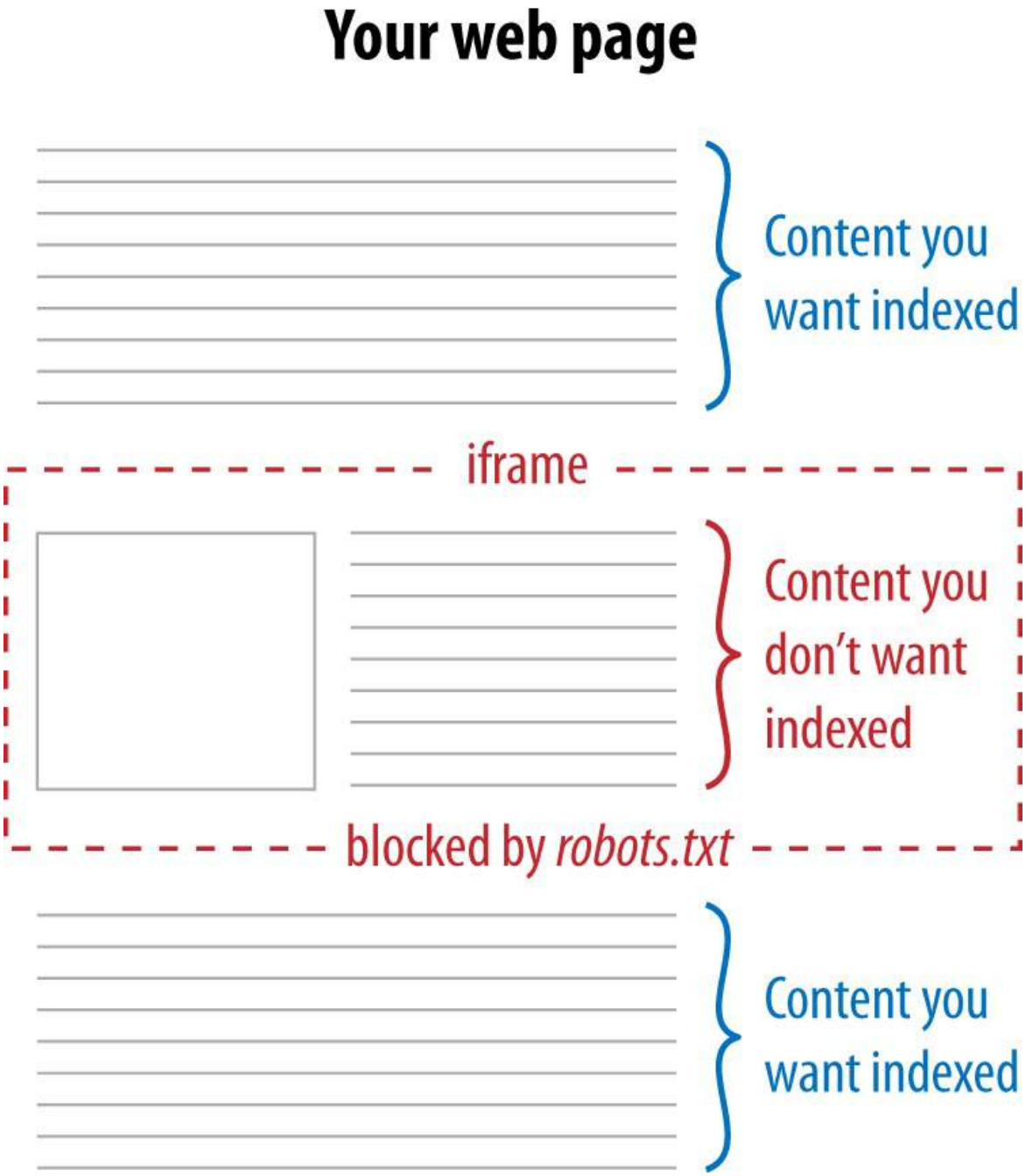


Figure 6-42. Using iframes to prevent indexing of content

The concept is simple: by using iframes, you can embed content from another URL onto any page of your choosing. By then blocking spider access to the iframe with *robots.txt*, you ensure that the search engines won't "see" this content on your page. Websites may do this for many reasons, including avoiding duplicate content problems, reducing the page size for search engines,

or lowering the number of crawlable links on a page (to help control the flow of link authority).

Hiding text in images

As discussed earlier, the major search engines still have very little capacity to read text in images (and the processing power required makes for a severe barrier). Hiding content inside images isn't generally advisable, as it can be impractical for alternative devices (mobile, in particular) and inaccessible to others (such as screen readers).

Hiding text in Java applets

As with text in images, the content inside Java applets is not easily parsed by search engines, though using them as a tool to hide text would certainly be a strange choice.

Forcing form submission

Search engines will not submit HTML forms in an attempt to access the information retrieved from a search or submission. Thus, if you keep content behind a forced-form submission and never link to it externally, your content will remain out of the engines (as Figure 6-43 demonstrates).

Figure 6-43. Using forms, which are generally not navigable by crawlers

The problem arises when content behind forms earns links outside your control, such as when bloggers, journalists, or researchers decide to link to the pages in your archives without your knowledge. Thus, although form submission may keep the engines at bay, make sure that anything truly sensitive has additional protection (e.g., through *robots.txt* or meta `robots`).

Using login/password protection

Password protection and/or paywalls of any kind will effectively prevent any search engines from accessing content, as will any form of human-verification requirements, such as CAPTCHAs (the boxes requiring users to copy letter/number combinations to gain access to content). The major engines won't try to guess passwords or bypass these systems.

Removing URLs from a search engine's index

A secondary, post-indexing tactic, URL removal from most of the major search engines is possible through verification of your site and the use of the engines'

tools. For example, Google allows you to remove URLs through Search Console. Bing also allows you to remove URLs from its index, via Bing Webmaster Tools.

Redirects

A redirect is used to indicate when content has moved from one location to another. For example, you may have some content at *http://www.yourdomain.com/old* and decide to restructure your site. As a result of this move, your content may move to *http://www.yourdomain.com/critical-keyword*.

Once a redirect is implemented, users who go to the old versions of your pages (perhaps via a bookmark they kept for the page) will be sent to the new versions. Without the redirect, the user would get a Page Not Found (404) error. With the redirect, the web server tells the incoming user agent (whether a browser or a spider) to instead fetch the requested content from the new URL.

Why and When to Redirect

Redirects are also important for letting search engines know when you have moved content. After you move content, the search engines will continue to have the old URL in their index and return it in their search results until they discover the page is no longer there and swap in the new page. You can help speed up this process by implementing a redirect. Here are some scenarios in which you may need to implement redirects:

- You have old content that expires, so you remove it.
- You find that you have broken URLs that have links and traffic.
- You change your hosting company.
- You change your CMS.
- You want to implement a canonical redirect (redirect all pages on *http://yourdomain.com* to *http://www.yourdomain.com*).

- You change the URLs where your existing content can be found, for any reason.

Not all of these scenarios require a redirect. For example, you can change hosting companies without impacting any of the URLs used to find content on your site, in which case no redirect is required. However, for any scenario in which any of your URLs change, you need to implement redirects.

Good and Bad Redirects

There are many ways to perform a redirect, but not all are created equal. The basic reason for this is that there are two major types of redirects that can be implemented, tied specifically to the HTTP status code returned by the web server to the browser. These are:

"301 moved permanently"

This status code tells the browser (or search engine crawler) that the resource has been permanently moved to another location, and there is no intent to ever bring it back.

"302 moved temporarily"

This status code tells the browser (or search engine crawler) that the resource has been temporarily moved to another location, and that the move should not be treated as permanent.

Both forms of redirect send a human or a search engine crawler to the new location, but the search engines interpret these two HTTP status codes in very different ways. When a crawler sees a 301 HTTP status code, it assumes it should pass the historical link authority (and any other metrics) from the old page to the new one. When a search engine crawler sees a 302 HTTP status code, it assumes it should not pass the historical link authority from the old page to the new one. In addition, the 301 redirect will lead the search engine to remove the old page from the index and replace it with the new one.

The preservation of historical link authority is very critical in the world of SEO. For example, imagine you had 1,000 links to *http://www.yourolddomain.com* and you decided to relocate everything to *http://www.yournewdomain.com*. If you

used redirects that returned a 302 status code, you would be starting your link-building efforts from scratch again. In addition, the old version of the page may remain in the index and compete for search rankings in the search engines.

Note that there also can be redirects that pass no status code, or the wrong status code, such as a 404 error (Page Not Found) or a 200 OK (Page Loaded Successfully). These are also problematic, and should be avoided. There are other types of redirects as well, such as those that return a 303 or 307 status code. These also should be avoided, as the search engine's response to them is at best unpredictable.

You want to definitively return a 301 HTTP status code for a redirect whenever you permanently move a page's location.

Methods for URL Redirecting and Rewriting

As we just mentioned, there are many possible ways to implement redirects. On Apache web servers (normally present on machines running Unix or Linux as the operating system), it is possible to implement redirects quite simply in a standard file called *.htaccess* using the `Redirect` and `RedirectMatch` directives. You can also employ more advanced directives known as *rewrite rules* using the Apache module known as mod_rewrite, which we will discuss in a moment.

On web servers running Microsoft IIS, different methods are provided for implementing redirects. As described in "IIS Redirects - 301 , 302", the basic method is through the IIS console. People with IIS servers can also make use of a text file with directives, provided they use an ISAPI plug-in such as ISAPI_Rewrite, and this scripting language offers capabilities similar to Apache's mod_rewrite module.

Many programmers use other techniques for implementing redirects, such as directly in programming languages like Perl, PHP, ASP, and JavaScript. If implementing redirects in this fashion, the programmer must make sure the HTTP status code returned by the web server is a 301. You can check the returned header with the Firefox plug-in Live HTTP Headers, with a Chrome extension, or with a web-based server header checker.

Another method that you can use to implement a redirect occurs at the page level, via the meta `refresh` tag, which looks something like this:

```
<meta http-equiv="refresh"
  content="5;url=http://www.yourdomain.com/newlocation.htm" />
```

The first parameter in the `content` section, `5`, indicates the number of seconds the web server should wait before redirecting the user to the indicated page. A publisher might use this to display a page letting users know that they're going to get redirected to a different page than the one they requested.

The problem is that most meta refreshes are treated as though they are a 302 redirect. The sole exception to this is if you specify a redirect delay of 0 seconds. You will have to give up your helpful page telling users that you are redirecting them, but the search engines treat this as though it were a 301 redirect (to be safe, the best practice is simply to use a 301 redirect if at all possible).

mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting

There is much more to discuss on this topic than we can reasonably address in this book. The following description is intended only as an introduction to help orient more technical readers, including web developers and site webmasters, on how rewrites and redirects function. To skip this technical discussion, proceed to "How to Redirect a Home Page Index File Without Looping".

mod_rewrite for Apache and ISAPI_Rewrite for Microsoft IIS Server offer very powerful ways to rewrite your URLs. Here are some reasons for using these tools:

- You have changed your URL structure on your site so that content has moved from one location to another. This can happen when you change your CMS, or change your site organization for any reason.

- You want to map your search engine–unfriendly URLs into friendlier ones.

If you are running Apache as your web server, you would place directives known as *rewrite rules* within your *.htaccess* file or your Apache configuration file (e.g., *httpd.conf* or the site-specific config file in the *sites_conf* directory). Similarly, if you are running IIS Server, you'd use an ISAPI plug-in such as ISAPI_Rewrite and place rules in an *httpd.ini* config file.

Note that rules can differ slightly on ISAPI_Rewrite compared to mod_rewrite, and the following discussion focuses on mod_rewrite. Your *.htaccess* file would

start with:

```
RewriteEngine on
RewriteBase /
```

You should omit the second line if you're adding the rewrites to your server config file, as `RewriteBase` is supported only in *.htaccess*. We're using `RewriteBase` here so that you won't have to type `^/` at the beginning of all the rules, just `^` (we will discuss regular expressions in a moment).

After this step, the rewrite rules are implemented. Perhaps you want to have requests for product page URLs of the format *http://www.yourdomain.comproducts123* to display the content found at *http://www.yourdomain.com/get_product.php?id=123*, without the URL changing in the location bar of the user's browser and without you having to recode the *get_product.php* script. Of course, this doesn't replace all occurrences of dynamic URLs within the links contained on all the site pages; that's a separate issue. You can accomplish this first part with a single rewrite rule, like so:

```
RewriteRule ^products/([0-9]+)/?$ /get_product.php?id=$1 [L]
```

This example tells the web server that all requests that come into the */product/* directory should be mapped into requests to */get_product.php*, while using the subfolder to */product/* as a parameter for the PHP script.

The `^` signifies the start of the URL following the domain, `$` signifies the end of the URL, `[0-9]` signifies a numerical digit, and the `+` immediately following it means one or more occurrences of a digit. Similarly, the `?` immediately following the `/` means zero or one occurrences of a slash character. The `()` puts whatever is wrapped within it into memory. You can then use `$1` to access what's been stored in memory (i.e., whatever's within the first set of parentheses). Not surprisingly, if you included a second set of parentheses in the rule, you'd access that with `$2`, and so on. The `[L]` flag saves on server processing by telling the rewrite engine to stop if it matches on that rule. Otherwise, all the remaining rules will be run as well.

Here's a slightly more complex example, where URLs of the format

http://www.yourdomain.comwebappwcs/stores/servlet/ProductDisplay?storeId=10001&catalogId=10001&langId=-1&categoryID=4&productID=123 would be rewritten to *http://www.yourdomain.com/4/123.htm*:

```
RewriteRule ^([^/]+)/([^/]+)\.htm$
webappwcs/stores/servlet/ProductDisplay?storeId=10001&catalogId=10001&
langId=-1&categoryID=$1&productID=$2 [QSA,L]
```

The `[^/]` signifies any character other than a slash. That's because, within square brackets, `^` is interpreted as *not*. The `[QSA]` flag is for when you don't want the query string dropped (like when you want a tracking parameter preserved).

To write good rewrite rules you will need to become a master of *pattern matching* (which is simply another way to describe the use of regular expressions). Here are some of the most important special characters and how the rewrite engine interprets them:

`*`

Zero or more of the immediately preceding character.

`+`

One or more of the immediately preceding character.

`?`

Zero or one occurrences of the immediately preceding character.

`^`

The beginning of the string.

`$`

The end of the string.

`.`

Any character (i.e., it acts as a wildcard).

`\`

"Escapes" the character that follows; for example, \. means the dot is not meant to be a wildcard, but an actual character.

`^`

Inside brackets `[]`means *not*; for example, `[^/]` means *not slash*.

It is incredibly easy to make errors in regular expressions. Some of the common gotchas that lead to unintentional substring matches include:

- Using `.*` when you should be using `.+` (because `.*` can match on nothing).
- Not "escaping" with a backslash a special character that you don't want interpreted, as when you specify `.` instead of `\.` and you really meant the dot character rather than any character (thus, *default.htm* would match on *defaultthtm*, and *default\.htm* would match only on *default.htm*).
- Omitting `^` or `$` on the assumption that the start or end is implied (thus, *default\.htm* would match on *mydefault.html*, whereas *^default\.htm$* would match only on *default.htm*).
- Using "greedy" expressions that will match on all occurrences rather than stopping at the first occurrence.

The easiest way to illustrate what we mean by "greedy" is to provide an example:

```
RewriteRule ^(.*)/?index\.html$ $1 [L,R=301]
```

This will redirect requests for *http://www.yourdomain.com/blah/index.html* to *http://www.yourdomain.com/blah/*. This is probably not what was intended. Why did this happen? Because `.*` will capture the slash character within it before the `/?` gets to see it. Thankfully, there's an easy fix. Simply use `[^` or `.*?` instead of `.*` to do your matching. For example, use `^(.*?)/?` instead of `^(.*)/?`, or

`[^/]+/[^/]` instead of `.*/.*`.

So, to correct the preceding rule, you could use the following:

```
RewriteRule ^(.*?)/?index\.html$ $1 [L,R=301]
```

Why wouldn't you use the following?

```
RewriteRule ^([^/]*)/?index\.html$ $1 [L,R=301]
```

This is more limited because it will match only on URLs with one directory. URLs containing multiple subdirectories, such as *http://www.yourdomain.com/store/cheese/swiss/wheel/index.html*, would not match.

As you might imagine, testing/debugging is a big part of URL rewriting. When you are debugging, the `RewriteLog` and `RewriteLogLevel` directives are your friends! Set the `RewriteLogLevel` to `4` or more to start seeing what the rewrite engine is up to when it interprets your rules.

By the way, the `[R=301]` flag in the last few examples—as you might guess—tells the rewrite engine to do a 301 redirect instead of a standard rewrite.

There's another handy directive to use in conjunction with `RewriteRule`, called `RewriteCond`. You would use `RewriteCond` if you were trying to match on something in the query string, the domain name, or other elements not present between the domain name and the question mark in the URL (which is what `RewriteRule` looks at).

Note that neither `RewriteRule` nor `RewriteCond` can access what is in the anchor part of a URL—that is, whatever follows a `#`—because that is used internally by the browser and is not sent to the server as part of the request. The following `RewriteCond` example looks for a positive match on the hostname before it will allow the rewrite rule that follows to be executed:

```
RewriteCond %{HTTP_HOST} !^www\.yourdomain\.com$ [NC]
RewriteRule ^(.*)$ http://www.yourdomain.com/$1 [L,R=301]
```

Note the exclamation point at the beginning of the regular expression. The

rewrite engine interprets that as *not*.

For any hostname other than *http://www.yourdomain.com*, a 301 redirect is issued to the equivalent canonical URL on the *www* subdomain. The `[NC]` flag makes the rewrite condition case-insensitive. Where is the `[QSA]` flag so that the query string is preserved, you might ask? It is not needed for redirecting; it is implied.

If you don't want a query string retained on a rewrite rule with a redirect, put a question mark at the end of the destination URL in the rule, like so:

```
RewriteCond %{HTTP_HOST} !^www\.yourdomain\.com$ [NC]
RewriteRule ^(.*)$ http://www.yourdomain.com/$1? [L,R=301]
```

Why not use `^yourdomain\.com$` instead? Consider:

```
RewriteCond %{HTTP_HOST} ^yourdomain\.com$ [NC]
RewriteRule ^(.*)$ http://www.yourdomain.com/$1? [L,R=301]
```

That would not have matched on typo domains, such as *yourdoamin.com*, that the DNS server and virtual host would be set to respond to (assuming that misspelling was a domain you registered and owned).

Under what circumstances might you want to omit the query string from the redirected URL, as we did in the preceding two examples? When a session ID or a tracking parameter (such as `source=banner_ad1`) needs to be dropped. Retaining a tracking parameter after the redirect is not only unnecessary (because the original URL with the source code appended would have been recorded in your access logfiles as it was being accessed); it is also undesirable from a canonicalization standpoint. What if you wanted to drop the tracking parameter from the redirected URL, but retain the other parameters in the query string? Here's how you'd do it for static URLs:

```
RewriteCond %{QUERY_STRING} ^source=[a-z0-9]*$
RewriteRule ^(.*)$ /$1? [L,R=301]
```

And for dynamic URLs:

```
RewriteCond %{QUERY_STRING} ^(.+)&source=[a-z0-9]+(&?.*)$
```

```
RewriteRule ^(.*)$ /$1?%1%2 [L,R=301]
```

Need to do some fancy stuff with cookies before redirecting the user? Invoke a script that cookies the user and then 301s him to the canonical URL:

```
RewriteCond %{QUERY_STRING} ^source=([a-z0-9]*)$
RewriteRule ^(.*)$ /cookiefirst.php?source=%1&dest=$1 [L]
```

Note the lack of a `[R=301]` flag in the preceding code. That's intentional. There's no need to expose this script to the user. Use a rewrite and let the script itself send the 301 after it has done its work.

Other canonicalization issues worth correcting with rewrite rules and the `[R=301]` flag include when the engines index online catalog pages under HTTPS URLs, and when URLs are missing a trailing slash that should be there. First, the HTTPS fix:

```
# redirect online catalog pages in the /catalog/ directory if HTTPS
RewriteCond %{HTTPS} on
RewriteRule ^catalog/(.*) http://www.yourdomain.com/catalog/$1 [L,R=301]
```

Note that if your secure server is separate from your main server, you can skip the `RewriteCond` line.

Now to append the trailing slash:

```
RewriteRule ^(.*[^/])$ $1 [L,R=301]
```

After completing a URL rewriting project to migrate from dynamic URLs to static, you'll want to phase out the dynamic URLs not just by replacing all occurrences of the legacy URLs on your site, but also by 301-redirecting the legacy dynamic URLs to their static equivalents. That way, any inbound links pointing to the retired URLs will end up leading both spiders and humans to the correct new URL—thus ensuring that the new URLs are the ones that are indexed, blogged about, linked to, and bookmarked, and the old URLs will be removed from the index. Generally, here's how you'd accomplish that:

```
RewriteCond %{QUERY_STRING} id=([0-9]+)
RewriteRule ^get_product\.php$ products%1.html? [L,R=301]
```

However, you'll get an infinite loop of recursive redirects if you're not careful. One quick-and-dirty way to avoid that situation is to add a nonsense parameter to the destination URL for the rewrite and ensure that this nonsense parameter isn't present before you do the redirect. Specifically:

```
RewriteCond %{QUERY_STRING} id=([0-9]+)
RewriteCond %{QUERY_STRING} !blah=blah
RewriteRule ^get_product\.php$ products%1.html? [L,R=301]
RewriteRule ^products/([0-9]+)/?$ /get_product.php?id=$1&blah=blah [L]
```

Notice that this example used two `RedirectCond` lines, stacked on top of each other. All redirect conditions listed together in the same block will be "ANDed" together. If you wanted the conditions to be "ORed," you'd need to use the `[OR]` flag.

How to Redirect a Home Page Index File Without Looping

Many websites link to their own home page in a form similar to *http://www.yourdomain.com/index.html*. The problem with that is that most incoming links to the site's home page specify *http://www.yourdomain.com*, thus dividing the link authority into the site. Once a publisher realizes this, she will want to fix her internal links and then 301-redirect *http://www.yourdomain.com/index.html* to *http://www.yourdomain.com/*, but recursive redirects can develop if she does not do this correctly.

When someone comes to your website by typing in *http://www.yourdomain.com*, the DNS system of the Web helps the browser locate the web server for your website. The web server decides what to show to the browser by loading a file from its hard drive.

When no file is specified (i.e., as in the preceding example, where only the domain name is given), the web server loads the default file, which is often a file with a name such as *index.html*, *index.htm*, *index.shtml*, *index.php*, or *default.asp*.

The filename can actually be anything, but most web servers default to one type of filename or another. Where the problem comes in is that many CMSs will expose both forms of your home page, both *http://www.yourdomain.com* and *http://www.yourdomain.com/index.php*.

Perhaps all the pages on the site link only to *http://www.yourdomain.com/index.php*, but given human nature, most of the links to your home page from third parties will most likely point at *http://www.yourdomain.com/*. This can create a duplicate content problem if the search engine now sees two versions of your home page and thinks they are separate, but duplicate, documents. Google is pretty smart at figuring out this particular issue, but it is best to not rely on that.

Because you learned how to do 301 redirects, you might conclude that the solution is to 301-redirect *http://www.yourdomain.com/index.php* to *http://www.yourdomain.com/*. Sounds good, right? Unfortunately, there is a big problem with this approach.

What happens is the server sees the request for *http://www.yourdomain.com/index.php* and then sees that it is supposed to 301-redirect that to *http://www.yourdomain.com/*, so it does. But when it loads *http://www.yourdomain.com/* it retrieves the default filename (*index.php*) and proceeds to load *http://www.yourdomain.com/index.php*. Then it sees that you want to redirect that to *http://www.yourdomain.com/*, and it creates an infinite loop.

The default document redirect solution

The solution that follows is specific to the preceding *index.php* example. You will need to plug in the appropriate default filename for your own web server.

1. Copy the contents of *index.php* to another file. For this example, we'll be using *sitehome.php*.

2. Create an Apache `DirectoryIndex` directive for your document root. Set it to *sitehome.php*. Do not set the directive on a serverwide level; otherwise, it may cause problems with other folders that still need to use *index.php* as a directory index.

3. Put this in an *.htaccess* file in your document root: *DirectoryIndex sitehome.php*. Or, if you aren't using per-directory context files, put this in your *httpd.conf*:

```
<Directory yourdocument/root/examplesite.com/>
```

```
 DirectoryIndex sitehome.php
</Directory>
```

4. Clear out the contents of your original *index.php* file. Insert this line of code:

```
<? header("Location: http://www.example.com"); ?>
```

This sets it up so that *index.php* is not a directory index file (i.e., the default filename). It forces *sitehome.php* to be read when someone types in the canonical URL (*http://www.yourdomain.com*). Any requests to *index.php* from old links can now be 301-redirected while avoiding an infinite loop.

If you are using a CMS, you also need to make sure when you are done with this process that all the internal links now go to the canonical URL, *http://www.yourdomain.com*. If for any reason the CMS started to point to *http://www.yourdomain.com/sitehome.php* the loop problem would return, forcing you to go through this entire process again.

Content Management System Issues

When looking to publish a new site, many publishers may wonder whether they need to use a content management system (CMS), and if so, how to ensure that it is SEO-friendly.

It is essential to determine whether you need a CMS before you embark on a web development project. You can use the flowchart in Figure 6-44 to help guide you through the process.

Do You Need A CMS For Your Site?

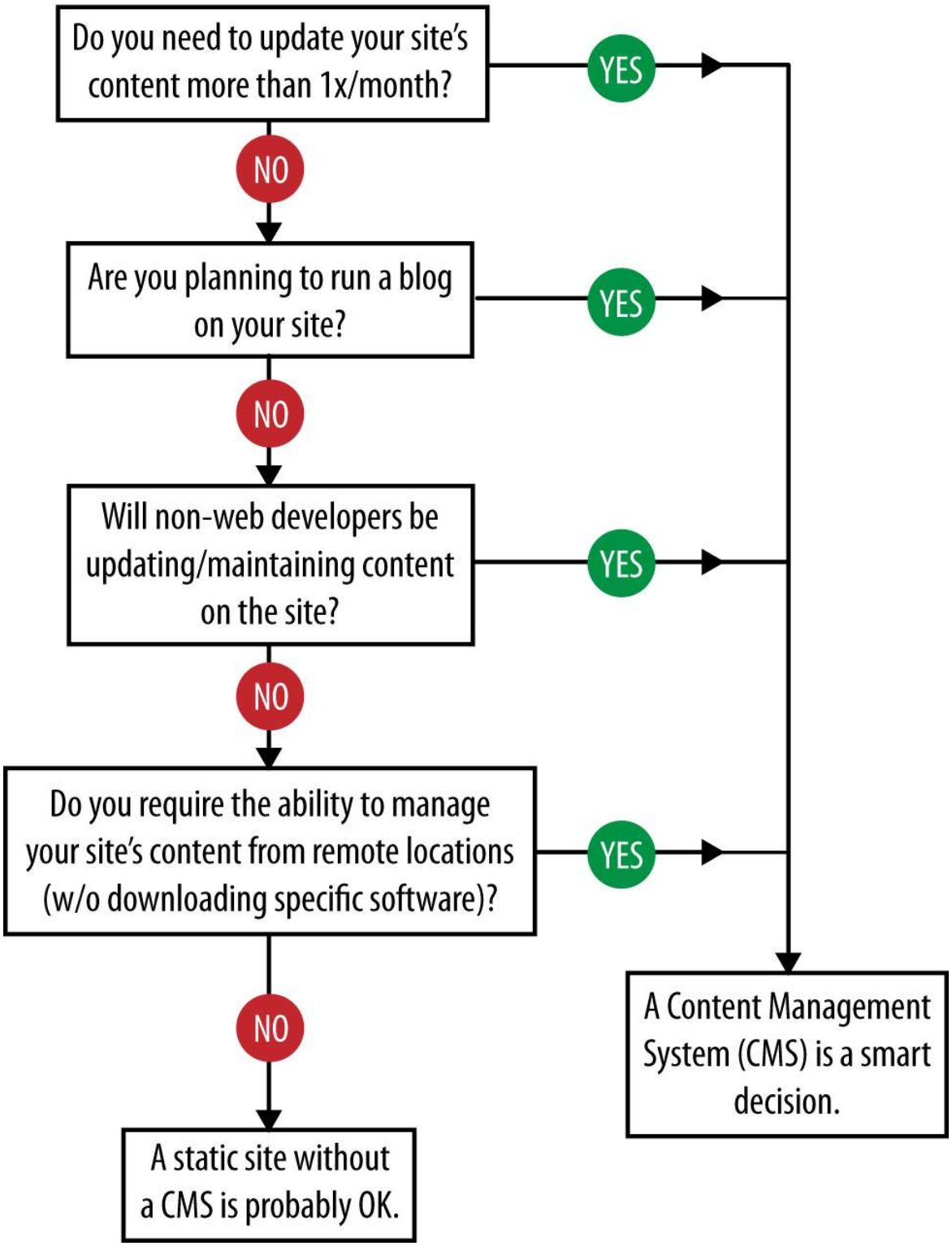


Figure 6-44. Flowchart to determine whether you need a CMS

Due to the inexpensiveness of customizable, free platforms such as Drupal, Joomla, WordPress, and Weebly, it is increasingly rare for a publisher to develop a static site, even when a CMS isn't required.

The next step involves understanding how to ensure that a CMS will be search engine–friendly. Here is a list of basic SEO issues that frequently plague a CMS (both prebuilt and custom-made systems). By dealing with these, you will ensure a relatively smooth platform for content delivery:

`<title>` *tag customization and rules*

A search engine–friendly CMS must allow for `<title>` tags not only to be customized on a page-specific level, but also to enable rules for particular sections of a website. For example, if the `<title>` tag always has to start with your site name followed by a colon followed by your article title, your on-page optimization efforts will be limited—at least as far as the powerful `<title>` tag is concerned. You should be able to revise the formulas you use to generate the `<title>` tags across your site to make them more search-optimal.

Static, keyword-rich URLs

URLs have historically been the most problematic SEO issue for CMS platforms. Nowadays, a search-friendly CMS should feature custom URL creation. In WordPress, a custom URL is referred to as a *post slug*. Figure 6-45 is an example from WordPress.

Notice how the first line allows you to create the title of the post, and the second enables you to manually create the URL structure (and an automatic Generate button if you prefer to simply use the post title).

Figure 6-45. Example of custom URL creation

Meta tag customization

Being able to implement custom meta descriptions and meta `robots` tags is critical. Enabling editorial control is essential for a good CMS.

Enabling custom HTML tags

A good CMS has to offer extra functionality on HTML tags for features such as `nofollow` on links, or `<hx>` tags for headlines and subheadlines. These can be built-in features accessible through menu options, or the CMS can simply allow for manual editing of HTML in the text editor window when required. Having no `<h1>` tags on a given page, having too many `<h1>` tags on the page, or marking up low-value content (such as the publication date) as an `<h1>` is not desirable. The article title is typically the best content to have wrapped in an `<h1>`.

Internal anchor text flexibility

For your site to be “optimized” rather than simply search-friendly, it’s critical to customize the anchor text on internal links. Rather than simply making all links in a site’s architecture the page’s title, a great CMS should be flexible enough to handle custom input from the administrators for the anchor text of category-level or global navigation links.

Intelligent categorization structure

Another common CMS problem is poor category structure. When designing an information architecture for a website, you should not place limits on how

pages are accessible due to the CMS's inflexibility. A CMS that offers customizable navigation panels will be the most successful in this respect.

Pagination controls

Pagination can be the bane of a website's search rankings, so controlling it by including more items per page, more contextually relevant anchor text (e.g., not "next," "prev," and page numbers), and careful use of meta `noindex` tags will make your important content get more link authority and crawl attention.

301-redirect functionality

Many content management systems sadly lack this critical feature, disallowing the proper redirection of content when necessary; 301s are valuable for expired content, for pages that have a newer version, and for dodging keyword cannibalization issues similar to those we discussed earlier in this chapter.

XML/RSS pinging

Although it is primarily useful for blogs, any content—from articles to products to press releases—can be issued in a feed. By utilizing quick, accurate pinging of the major feed services, you limit some of your exposure to duplicate content spammers who pick up your feeds and ping the major services quickly in the hopes of beating you to the punch.

Image handling and `alt` *attributes*

`alt` attributes are a clear must-have from an SEO perspective, serving as the "anchor text" when an image is used as a link (note that text links are much better than images with `alt` attributes, but if you must use image links you should implement the `alt` attribute) and providing relevant, indexable content for the search engines. Images in a CMS's navigational elements should preferably use CSS image replacement rather than mere `alt` attributes.

CSS exceptions

The application of CSS styles in a proper CMS should allow for manual exceptions so that a user can modify how a strong headline or list element appears visually. If the CMS does not offer this, writers may opt out of using proper semantic markup for presentation purposes, which would not be a good thing.

Static caching options

Many content management systems currently offer caching options, which are a particular boon if a page is receiving a high level of traffic from social media portals or news sites. A bulky CMS often makes dozens of extraneous database connections, which can overwhelm a server if caching is not in place, killing potential inbound links and media attention.

URLs free of tracking parameters and session IDs

Sticking session or tracking information such as the user's click path into the URL is deadly for SEO. It usually leads to incomplete indexation and duplicate content issues.

Customizable URL structure

If the default URL structure of the CMS doesn't suit your needs, you should be able to change it. For example, if you don't want *archives* in the URLs of all your archived articles, you should be able to remove it. Or if you want to reference the article name instead of the article's database ID in the URL, you should be able to do it.

301 redirects to a canonical URL

Duplicate content is a major concern for the dynamic website owner. Automatic handling of this by the CMS through the use of 301 redirects is a must.

Static-looking URLs

The most palatable URLs to spiders are the ones that look like they lead to

static pages—no query strings in the URL.

Keywords in URLs

Keywords in your URLs (used judiciously) can help your rankings.

RSS feeds

The CMS should autocreate RSS feeds to help your site rank in Google Blog Search and other feed engines.

Multilevel categorization structure

It is awfully limiting to your site structure and internal hierarchical linking structure to have a CMS that doesn't allow you to nest subcategories into categories, sub-subcategories into subcategories, and so on.

Paraphrasable excerpts

Duplicate content issues are exacerbated on dynamic sites such as blogs when the same content is displayed on permalink pages, category pages, archives-by-date pages, tag pages, and the home page. Crafting unique content for the excerpt, and having that content display on all locations except the permalink page, will help strengthen your permalink page as unique content.

Breadcrumb navigation

Breadcrumb (drill-down) navigation is great for SEO because it reinforces your internal hierarchical linking structure with keyword-rich text links.

Meta `noindex` *tags for low-value pages*

Even if you use `nofollow` attributes in links to these pages, other people may still link to them, which carries a risk of ranking those pages above some of your more valuable content.

Keyword-rich intro copy on category-level pages

Keyword-rich introductory copy helps set a stable keyword theme for the page, rather than relying on the latest article or blog post to be the most prominent text on the page.

`nofollow` *links in comments*

If you allow visitors to post comments and do not `nofollow` the links, your site will be a spam magnet. Heck, you'll probably be a spam magnet anyway, but you won't risk losing PageRank to spammers if you use `nofollow` attributes.

Customizable anchor text on navigational links

Contact, About Us, Read More, Full Article, and so on make for lousy anchor text—at least from an SEO standpoint. Hopefully, your CMS allows you to improve such links to make the anchor text more keyword-rich.

XML sitemap generator

Having your CMS generate your XML sitemap can save a lot of hassle, as opposed to trying to generate one with a third-party tool.

HTML4, HTML5, or XHTML validation

Although HTML validation is not a ranking signal, it is desirable to have the CMS automatically check for malformed HTML, as search engines may end up seeing a page differently from how it renders on the screen and accidentally consider navigation to be part of the content or vice versa.

Pingbacks, trackbacks, comments, and antispam mechanisms

The problem with comments/trackbacks/pingbacks is that they are vectors for spam, so if you have one or more of these features enabled, you will be spammed. Therefore, effective spam prevention in the form of Akismet, Mollom, or Defensio is a must.

If you want more information on picking a quality CMS, some great web resources are already out there—among them OpenSourceCMS.com and

CMSmatrix—to help manage this task.

CMS Selection

There are many factors to consider when choosing an existing CMS. Many CMS platforms are free, but some of them are proprietary with a license cost per site. The majority were not designed with security, stability, search friendliness, and scalability in mind, though in recent years a few vendors have developed excellent systems that have search friendliness as their primary focus. Many were developed to fit a certain market niche, but can be expanded to fit other purposes. Some are no longer maintained. Many are supported and developed primarily by hobbyists who don't particularly care if you're having trouble getting them installed and configured. Some are even intentionally made to be difficult to install and configure so that you'll be encouraged to pay the developers a consulting fee to do it all for you.

Popular CMS solutions that the authors have experience with include Joomla, Drupal, concrete5, Pixelsilk, WordPress, Magento, and Sitecore. Each has strong support for SEO, but requires some configuration for optimal results. Make sure you get that help up front to get the SEO for your site off to a strong start.

Selecting a CMS is an important process. If you make the wrong choice, you will be faced with limited SEO options. Like most software, a CMS is a moving target—what's missing today may be a new feature tomorrow. In addition, just because a feature exists doesn't mean it is the default option, so in many instances the desired functionality will need to be enabled and possibly customized to work to your specifications.

Third-Party CMS Add-Ons

Many CMS platforms offer third-party plugins or add-ons that extend the core functionality of the CMS. In the WordPress plug-in directory alone, there are over 34,000 plugins, including the hugely popular WordPress SEO by Yoast and All in One SEO Pack. Plugins provide a simple way to add new SEO features and functionality, making the CMS much more flexible and future-proof. It is particularly helpful when there is an active community developing plugins. An active community also comes in very handy in providing free technical support when things go wrong; and when bugs and security vulnerabilities crop up, it is

important to have an active developer base to solve those issues quickly.

Many CMS add-ons—such as discussion forums, customer reviews, and user polls—come in the form of independent software installed on your web server, or hosted services. Discussion forums come in both of these forms: bbPress, which is installed software and is optimized for search, and vbulletin, which is a hosted solution and therefore more difficult to optimize for search.

The problem with hosted solutions is that you are helping to build the service providers' link authority and not your own, and you have much less control over optimizing the content.

As we referenced several times earlier in this chapter, Flash is popular on the Web, but presents challenges to the search engines in terms of indexing the related content. This creates a gap between the user experience with a site and what the search engines can find on that site.

In the past, search engines did not index Flash content at all. In June 2008, Google announced that it was offering improved indexing of this content. This announcement indicates that Google can index text content and find and follow links within Flash files. However, Google still cannot tell what is contained in images within the Flash file. Here are some reasons why Flash is still not fully SEO-friendly:

Different content is not on different URLs

This is the same problem you encounter with AJAX-based pages. You could have unique frames, movies within movies, and so on that appear to be completely unique portions of the Flash site, yet there's often no way to link to these individual elements.

The breakdown of text is not clean

Google can index the output files in the *.swf* file to see words and phrases, but in Flash a lot of your text is not inside clean `<h1>` or `<p>` tags; it is jumbled up into half-phrases for graphical effects and will often be output in the incorrect order. Worse still are text effects that often require "breaking" words apart into individual letters to animate them.

Flash gets embedded

A lot of Flash content is linked to only by other Flash content wrapped inside shell Flash pages. This line of links, where no other internal or external URLs are referencing the interior content, means some very low PageRank/link authority documents. Even if they manage to stay in the main index, they probably won't rank for anything.

Flash doesn't earn external links like HTML

An all-Flash site might get a large number of links to the home page, but interior pages almost always suffer. For embeddable Flash content, it is the HTML host page earning those links when they do come.

SEO basics are often missing

Anchor text, headlines, bold/strong text, `img alt` attributes, and even `<title>` tags are not simple elements to properly include in Flash. Developing Flash with SEO in mind is just more difficult than doing it in HTML. In addition, it is not part of the cultural lexicon of the Flash development world.

A lot of Flash isn't even crawlable

Google has indicated that it doesn't execute external JavaScript calls (which many Flash-based sites use) or index the content from external files called by Flash (which, again, a lot of Flash sites rely on). These limitations could severely impact what a visitor can see versus what Googlebot can index.

Note that in the past you could not test the crawlability of Flash, but the Adobe Search Engine SDK now gives you an idea of how the search engines will see your Flash file.

Flash Coding Best Practices

If Flash is a requirement for whatever reason, there are best practices you can implement to make your site more accessible to search engine spiders. What follows are some guidelines on how to obtain the best possible results.

Flash meta tags

Beginning with Adobe/Macromedia Flash version 8, there has been support for the addition of title and description meta tags to any *.swf* file. Not all search engines are able to read these tags yet, but it is likely that they will soon. Get into the habit of adding accurate, keyword-rich `<title>` tags and meta tags to files now so that as search engines begin accessing them, your existing *.swf* files will already have them in place.

Adobe Flash search engine SDK

Flash developers may find the SDK useful for server-based text and link extraction and conversion purposes, or for client-side testing of their Flash content against the basic Adobe (formerly Macromedia) Flash Search Engine SDK code.

Tests have shown that Google and other major search engines now extract some textual content from Flash *.swf* files. It is unknown whether Google and others have implemented Adobe's specific Search Engine SDK technology into their spiders, or whether they are using some other code to extract the textual content. Again, tests suggest that what Google is parsing from a given *.swf* file is very close to what can be extracted manually using the Search Engine SDK.

The primary application of Adobe's Search Engine SDK is desktop testing *.swf* files to see what search engines are extracting from a given file. The program cannot extract files directly from the Web; the *.swf* file must be saved to a local hard drive. The program is DOS-based and must be run in the DOS Command Prompt using DOS commands.

By running a *.swf* file through the Flash SDK swf2html program during development, you can edit or augment the textual assets of the file to address the best possible SEO practices—homing in primarily on keywords and phrases along with high-quality links. Because of the nature of Flash and the way in which it deals with both text and animation, it is challenging to get exacting, quality SEO results. The goal is to create the best possible SEO results within the limitations of the Flash program and the individual Flash animation rather than to attempt the creation of an all-encompassing SEO campaign. Extracted content from Flash should be seen as one tool among many in a larger SEO campaign.

Internal Flash coding

There are several things to keep in mind when preparing Flash files for SEO:

- Search engines currently do not read traced text (using the `trace()` function) or text that has been transformed into a shape in Flash (as opposed to actual characters). Only character-based text that is active in the Flash stage will be read (see Figure 6-46).

- Animated or affected text often creates duplicate content. Static text in Flash movies is not read as the duplicate instances that "tweening" and other effects can create. Use static text, especially with important content, so that search engines do not perceive the output as spam (see Figure 6-47).

- Search engine spiders do not see dynamically loaded content (text added from an external source, such as an XML file).

- The font size of text does not affect search engines; they read any size font.

- Special characters such as <, >, &, and " are converted to HTML character references (`<`, `>`, `&`, and `"`) and should be avoided.

- Search engines find and extract all URLs stored within the `getURL()` command.

- Search engines have the ability to follow links in Flash, though it is an "iffy" proposition at best. They will not, however, follow links to other Flash *.swf* files. (This is different from loading child *.swf* files into a parent *.swf* file.) Therefore, links in Flash should always point to HTML pages, not other *.swf* files.

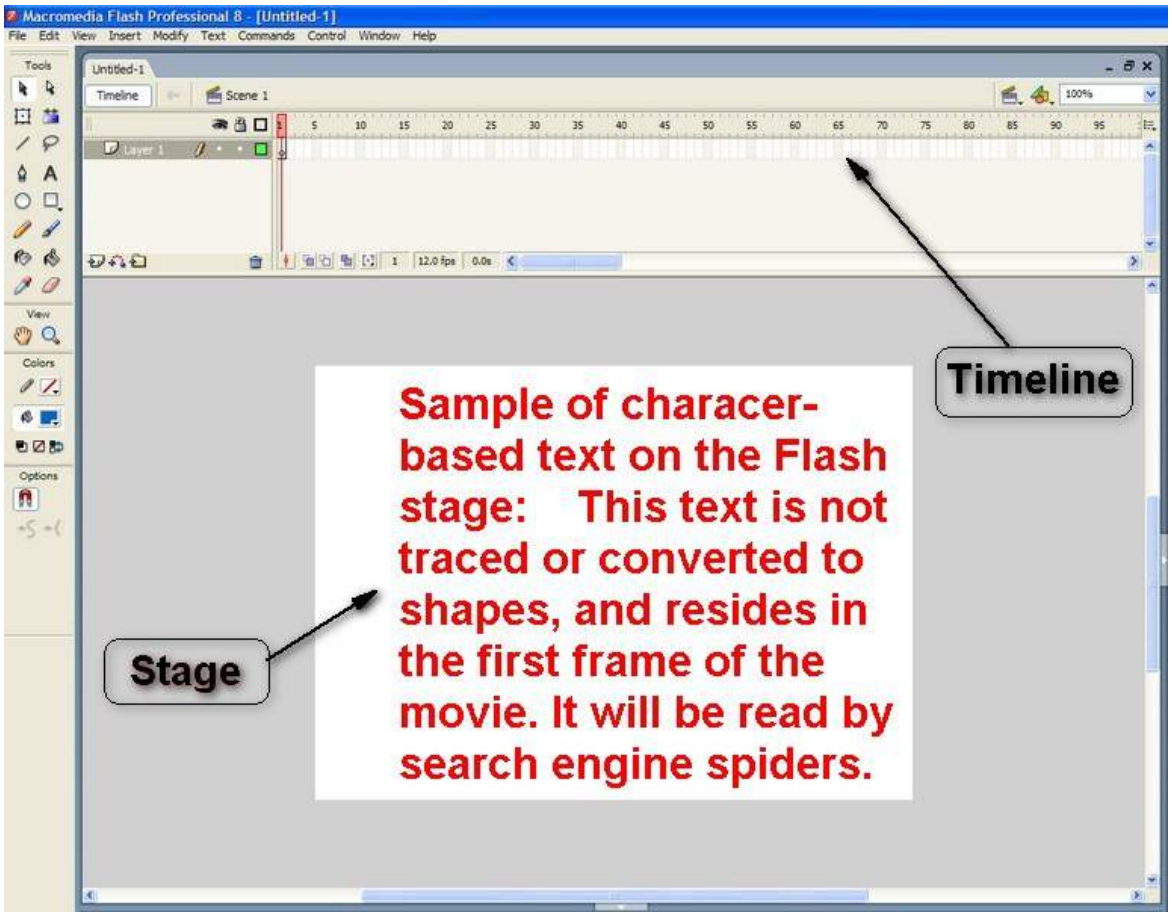


Figure 6-46. Example of spider-readable text inside a Flash program

```
<p>Fast Homework Help</p>
<p>Fast Homework Help</p>
<p>Reliable Report Info</p>
<p>Reliable Report Info</p>
<p>Math Made Easy</p>
<p>Math Made Easy</p>
<p>Brain Games + More</p>
<p>Brain Games + More</p>
<p>SCHOOL TOOLS</p>
<p>SCHOOL TOOLS</p>
<p>What Is Cosmeo?</p>
<p>What Is Cosmeo?</p>
<p>Why It Works</p>
<p>Why It Works</p>
<p>Take A Tour</p>
<p>Take A Tour</p>
<p>Get Started</p>
<p>Get Started</p>
```

Figure 6-47. Animated text results in Flash source; can be seen as duplicate content

SWFObject library and <noscript> tag

Because “alternative content” workarounds for SEO of Flash files have historically been abused by spammers, it is challenging to recommend these tactics to optimize your Flash files without a critical disclaimer.

Both the SWFObject library and `<noscript>` tag were originally designed to be legitimate, graceful degradation techniques readily accepted by the search engines as a way to accommodate older browsers or people with special needs. But many unscrupulous sites have used the code to trick search engine spiders. In other words, these methods are used in such a way that browsers display one thing to users, but something completely different to search engine spiders. As you’ve learned in this chapter, all of the major search engines disapprove of such tactics.

Websites using such methods today are often penalized or removed from search engine indexes altogether. This makes graceful degradation risky on some level, but if the methods are used clearly within the boundaries for which they were intended, getting penalized or banned is highly unlikely.

As we discussed earlier in this chapter, intent is an essential element search engines take into consideration. If your intent is to provide *all* users with a positive experience while they’re visiting your site, you should be fine. If your intent is to game the search engines, all it takes is one online rival to report your site for spam to incur the wrath of the search engines.

Google and other search engines do not algorithmically ban sites for using SWFObject and `<noscript>`; it usually requires human intervention to evoke a penalty or outright ban.

SWFObject

SWFObject is the better of the two Flash optimization options because it is JavaScript code designed specifically for Flash *.swf* purposes, and it has been abused to a lesser extent than the `<noscript>` tag.

SWFObject is a Flash detection code library written in JavaScript that checks whether a browser has the Flash plug-in. If the browser does have the Flash plug-in, the *.swf* file is displayed secondary to that detection. If the browser does not have the Flash plug-in or the JavaScript to detect it, the primary, alternative

content contained within `<div>` tags is displayed instead. The key here is that search engine spiders do not render the JavaScript. They read the primary content in the `<div>` tags.

The opportunity for abuse is obvious when you view the code. This small piece of code is placed within the `<head>` tags:

```
<script type="text/javascript" src="swfobject.js"></script>
```

In the body of the text, the code looks something like Figure 6-48.

```
<script type="text/javascript" src="swfobject.js"></script>

<div id="flashcontent">
      Text, links, and graphics placed here are replaced by the Flash movie. Search
      engine spiders will read this information, but the browser with an active Flash
      plugin will show the Flash movie instead.
</div>

<script type="text/javascript">
      var so = new SWFObject("whatever.swf", "themovie", "200", "100", "7", #336699");
      so.write("flashcontent");
</script>
```

Figure 6-48. Information between the HTML tags read by search engine spiders

Search engine spiders will read text, links, and even `alt` attributes within the `<div>` tags, but the browser will not display them unless the Flash plug-in isn't installed (about 95% of browsers now have the plug-in) or JavaScript isn't available.

Once again, the key to successfully implementing SWFObject is to use it to the letter of the law: leverage it to mirror the content of your Flash *.swf* file *exactly*. Do not use it to add content, keywords, graphics, or links that are not contained in the file. Remember, a human being will be making the call as to whether your use of SWFObject is proper and in accordance with that search engine's guidelines. If you design the outcome to provide the best possible user experience, and your intent is *not* to game the search engines, you are probably OK.

You can download the SWFObject JavaScript library free of charge. Included in this download is the *flashobject.js* file, which is placed in the same directory as

the web pages on which the corresponding calling code resides.

<noscript>

The `<noscript>` tag has been abused in “black hat” SEO attempts so frequently that you should be cautious when using it. Just as SWFObject and `<div>` tags can be misused for link and keyword stuffing, so too can the `<noscript>` tag. Certain companies have promoted the misuse of the `<noscript>` tag widely; consequently, there have been many more problems with its use.

With that being said, conservative and proper use of the `<noscript>` tag specifically with Flash *.swf* files can be an acceptable and good way to get content mirrored to a Flash file read by search engine spiders. As is the case with SWFObject and corresponding `<div>` tags, content must echo that of the Flash *.swf* movie exactly. Do not use `<noscript>` to add content, keywords, graphics, or links that are not in the movie. Again, it is a human call as to whether a site or individual page is banned for the use or misuse of the `<noscript>` tag.

You use `<noscript>` with Flash *.swf* files in the following manner:

```
<script type="text/javascript" src="yourflashfile.swf"></script>
```

Followed at some point by:

```
<noscript>
<h1>Mirror content in Flash file here.</h1>
<p>Any content within the noscript tags will be read by the search engine
spiders, including links
http://www.mirroredlink.com, graphics, and corresponding alt attributes.
</noscript>
```

For browsers that do not have JavaScript installed or functioning, content alternatives to JavaScript-required entities are displayed. So, for use with Flash *.swf* files, if a browser does not have JavaScript and therefore cannot display Flash, it displays instead the content inside the `<noscript>` tags. This is a legitimate, graceful degradation design. For SEO purposes, as is true with SWFObject, the search engine spiders do not render the JavaScript but do read the content contained in the HTML. Here, it is the content between the `<noscript>` tags.

Scalable Inman Flash Replacement

Scalable Inman Flash Replacement (sIFR) is a technique that uses JavaScript to read in HTML text and render it in Flash instead. The essential fact to focus on here is that the method guarantees that the HTML content and the Flash content are identical. One great use for this technique is to render headline text in an anti-aliased font (this is the purpose for which sIFR was designed). This can greatly improve the presentation of your site.

Dan Crow, head of Google's Crawl team, said that as long as this technique is used in moderation, it is OK. However, extensive use of sIFR could be interpreted as a signal of poor site quality. Because sIFR was not designed for large-scale use, such extensive use would not be wise in any event.

It is worth noting that there are comparable technologies available to web designers for improved type presentation, which provide similar search engine friendliness. FaceLift Image Replacement (FLIR) is an image replacement script similar to sIFR in its use of JavaScript, but without the Flash element, and there is a handy WordPress plug-in for implementation on WordPress-based websites. Google also offers its own set of fonts optimized for use on websites.

Best Practices for Multilanguage/Country Targeting

Many businesses target multiple countries with their websites and need answers to questions such as: Do you put the information for your products or services all on the same domain? Do you obtain multiple domains? Where do you host the site(s)? As it turns out, there are SEO factors, as well as basic marketing questions, that affect the answers. There are also non-SEO factors, such as the tax implications of what you do; you can get some TLDs only by having a local physical presence (e.g., France requires this to issue a *.fr* domain).

How to Target a Specific Country

Starting with the basics of international targeting, it is important to let the search engines know where your business is based in as many ways as possible. These might include:

- Using a country code TLD (ccTLD) for your domain (e.g., *.uk*)
- Hosting your site locally (more for content delivery speed than for the "localness" factor)
- Displaying the physical local address in plain text on every page of your site
- Setting Google Search Console geotargeting to your country of interest
- Verifying your address with Google Maps
- Including links from in-country websites
- Using the local language on the website

If you are starting from scratch, getting these components all lined up will give you the best possible chance of ranking in the local country you are targeting.

Problems with Using Your Existing Domain

You may ask why you cannot leverage your domain weight to target the new territory rather than starting from scratch—in other words, why can't you create multiple versions of your site and determine the user's location before either delivering the appropriate content or redirecting him to the appropriate place in the site (or even to a subdomain hosted in the target country)?

The problem with this approach is that the search engines spider from the United States, meaning their IP addresses will be in the United States in your lookup and thus they will be delivered only U.S. content from your site. This problem is exacerbated if you are going even further and geodelivering content in different languages, as only your English language content will be spidered unless you cloak for the search engine bots.

This kind of IP delivery is therefore a bad idea. You should make sure you do not blindly geodeliver content based on IP address, as you will ignore many of your markets in the search engines' eyes.

The Two Major Approaches

The best practice remains one of two approaches, depending on the size and scale of your operations in the new countries and how powerful and established your *.com* domain is.

If you have strong local teams and/or (relatively speaking) less power in your main domain, launching independent local websites geotargeted as described earlier is a smart move in the long run.

If, on the other hand, you have only centralized marketing and PR and/or a strong main domain, you may want to create localized versions of your content either on country-specific subdomains (*http://uk.yourdomain.com*, *http://au.yourdomain.com*, etc.) or in subfolders (*/uk/*, *au*, etc.), with the preference being for the use of subdomains so that you can set up local hosting.

Both the subdomain and the subfolder approach allow you to set your geotargeting option in Google Search Console, and with either method, you have to be equally careful of duplicate content across regions. In the subdomain example, you can host the subdomain locally, while in the subfolder case, more of the power of the domain filters down.

Unfortunately, the Search Console's geotargeting option doesn't work nearly as well as you'd hope to geotarget subfolders. The engines will consider hosting and ccTLDs, along with the geographic location of your external link sources, to be stronger signals than the manual country targeting in the tools. In addition, people in other countries (e.g., France) don't like to click on *.com* or *.org* TLDs; they prefer *.fr*. This extends to branding and conversion rates too—web users in France like to buy from websites in France that end in *.fr*.

Multiple-Language Issues

An entire treatise could be written on handling multilanguage content as the search engines themselves are rapidly evolving in this field, and tactics are likely to change dramatically in the near future. Therefore, this section will focus on providing you with the fundamental components of successful multilanguage content management.

Here are best practices for targeting the search engines as of this writing, using Spanish and English content examples:

- Content in Spanish and English serving the same country:

— Create a single website with language options that change the URL by folder structure; for example, *http://www.yourdomain.com* versus *http://www.yourdomain.com/esp/*.

— Build links from Spanish and English language sites to the respective content areas on the site.

— Host the site in the country being served.

— Register the appropriate country domain name (for the United States, *.com*, *.net*, and *.org* are appropriate, whereas in Canada using *.ca* or in the United Kingdom using *.uk* is preferable).

— Mark up your HTML code using `hreflang` tags for multiple languages. See "hreflang for multiple languages/no specific location".

- Content in Spanish and English targeting multiple countries:

  — Create two separate websites, one in English targeting the United States (or the relevant country) and one in Spanish targeting the relevant Spanish-speaking countries.

  — Host one site in the United States (for English) and the other in the relevant countries for the Spanish version.

  — Register different domains, one using U.S.-targeted domain extensions and one using the Spanish-speaking countries' extension.

  — Acquire links from the United States to the English site and links from the Spanish-speaking countries to that site.

  — Mark up your HTML code using `hreflang` tags for multiple languages/locations. See "hreflang for multiple languages/regions".

- Content in Spanish targeting multiple countries:

  — Create multiple websites (as mentioned earlier) targeting each specific country.

— Register domains using the appropriate country TLD and host in each country separately.

— When possible, have native speakers fluent in the specific region's dialect write the site content for each specific country.

— Obtain in-country links to your domains.

— Mark up your HTML code using hreflang tags for multiple languages/locations. See "hreflang for one language/multiple regions".

Although some of these approaches may seem counterintuitive, the joint issues of search engines preferring to show content hosted in and on a country-specific domain name combined with duplicate content problems make for these seemingly illogical suggestions.

hreflang markup

There are several options for serving multiregion and multilanguage content to the search engines. You'll need to use at least one of these solutions to encourage Google to rank the appropriate version of your content in the appropriate version of the Google search engine (*google.com*, *google.co.uk*, *google.ca.*, etc.). These solutions are also necessary to prevent duplicate content issues both within a single language—such as American English and UK English, where the copy is likely to be virtually identical—and also across languages that are more unique.

There are three main options available to serve up multilanguage or multiregion content:

- Code within the server header section of a page
- Code within the `<head>` section of the HTML on a page
- Special directives within the site's XML sitemap, or a specific multiregion/multilanguage sitemap

It's recommended that you use only one of these solutions at a time. While redundancy, if accurate, will cause no negative effects, there's the possibility of disagreement between multiple solutions if they are working simultaneously,

which can confuse the search engines about which version to “count.”

We will focus on the second option: code within the `<head>` section of the HTML on a page.

hreflang for multiple languages/no specific location

Each page that has alternate language versions, but not alternate country versions, should contain markup specifying the language only. It is acceptable for pages to contain language-only markup but never region-only markup. Once pages are built for specific regions, they must be marked up with a combination of both language and region markup. An example of this markup for a home page presented in both English and Spanish follows.

If the home page of a site, in this case *example.com*, is translated into both English and Spanish, both versions of the page should include code such as:

```
<link rel="alternate" hreflang="x-default" href="example.com" >
<link rel="alternate" href="example.comes/" hreflang="es" />
```

Each language will have its own unique `hreflang` code. Note that there is no accommodation within the language markup for the difference between Spanish for Spain and Spanish for Latin America. Similarly, there is no difference in the language markup between Portuguese for Portugal and Portuguese for Brazil, or Canadian French versus the version spoken in France, and so on.

A full list of the two-character language codes can be found at *http://www.loc.gov/standards/iso639-2/php/code_list.php* under the ISO 639-1 standard.

hreflang for multiple languages/regions

If you wanted to have a default version of the page (English language, no region assigned), a version for Spanish from Mexico, and a version for Spanish from Spain, the markup would look similar to the following. Please note that each region/language combination would need its own unique URL/domain:

```
<link rel="alternate" hreflang="x-default" href="example.com" >
<link rel="alternate" href="example.es" hreflang="es-es" >
<link rel="alternate" href="example.com.mx" hreflang="es-mx" />
```

Another consideration in your geotargeting of URLs is that there are no provisions for markup associated with “regions” such as Latin America, APAC, the EU, and so on. Each country within these regions must be treated individually.

hreflang for one language/multiple regions

If you wished to have versions in Spanish for Spain and versions for Spanish for Latin America, you would use markup similar to the following:

```
<link rel="alternate" hreflang="x-default" href="example.es" >
<link rel="alternate" href="example.es" hreflang="es-es" >
<link rel="alternate" href="example.com.mx" hreflang="es-mx" >
<link rel="alternate" href="example.com.cr" hreflang="es-cr" >
<link rel="alternate" href="example.com.com.ar" hreflang="es-ar" />
```

...and so on for each Latin American country.

A full list of country-level TLDs can be found at *http://www.mcanerin.com/EN/articles/ccTLD.asp*.

It pays to plan ahead when adding `hreflang` markup for alternate language/country versions of your site. Each alternate version of a page needs to reference every other alternate version. If you have a version of a page in English for the United States and another in Spanish for Mexico, both of those pages need the markup referencing the other version. If you were to then add a version in Spanish for Spain, not only does this new version need to reference both the English/U.S. version and the Spanish/Mexican version, but both of those pages now also need to reference this new Spanish/Spain version.

This level of complexity is why it is crucial that before you proceed with creating alternate language/region versions of your content, you have a comprehensive international strategy. Before you embark on any site changes that are detectable by the search engines, you should be planning ahead several years as to what regions and languages you will optimize for. If you choose not to do this level of planning, you may face numerous code changes across all alternate-version pages as new countries/languages are added in the future.

For more information on `hreflang` markup, see “Use hreflang for language and regional URLs - Search Console Help” and the blog post “Using the Correct

Hreflang Tag: A New Generator Tool" by Aleyda Solis. Aleyda also created the "hreflang Tags Generator Tool".

NOTE

A special thanks to Rob Woods for his contributions to this portion of the chapter.

Semantic Search

There is a lot of confusion over the definition of semantic search. Some of this confusion comes from the formal definition of *semantics* commonly associated with linguistics, and some of it comes from the misunderstanding that arises the moment the words "structured data" are mentioned.

In truth, semantic search has a little to do with both, and a lot to do with the four vectors that drive Big Data across the Web:

- Volume is about processing massive amounts of data and extracting unique meaning from it.
- Velocity refers to the speed at which critical data comes in and how quickly it must be analyzed and processed.
- Variety is required as well, as many different types of data must be handled, such as audio, video, and text.
- Veracity is about the need to validate the accuracy of the data being processed.

To help you understand this concept better, it helps to take things from the beginning, and the true beginning for semantic search was August 30, 2013, when Google quietly rolled out Hummingbird.

The change, which was announced almost a month later on the eve of Google's 15th birthday, completed Google's long journey to turn search into more than a blind fishing expedition where those who created content and those who looked for it continually strove to guess each other's keywords and connect.

Google’s Hummingbird

To understand how much search (and SEO) has changed, consider just how far voice search has come. When we use voice, we tend to speak in sentences instead of keywords, and in order for Google to return meaningful answers it has to be able to understand our search query. To handle spoken queries well, Google also needs to understand intent, which requires it to be able to understand context.

The same technology that was applied to Google voice search before August 30, 2013, is now applied to the regular text search with which everyone is familiar. Hummingbird (which Google said was so named because it was precise and fast) does a number of things no search engine had done before.

First, it takes the entire search query into account—not just the keywords, but every word. Second, it looks at who is carrying out the search. Suddenly, variables such as past search history and search patterns are important in delivering the right results, at the right time, to the right person. Thirdly, it also factors in how the search itself is being conducted. Device type, time of day, and location now are also important parameters affecting the search results.

With linguistic sensitivity (i.e., the ability to better process natural language) Google’s Hummingbird is also better at understanding the relationships between queries and between bits of data. It is in this space that the real magic happens.

Semantic search, really, is about relational connections and contextual content. In order to deliver “the right results, at the right time, to the right person,” semantic search needs to understand the importance of the query to that person and the importance of the query in relation to the data it already holds in its index and the data it is currently indexing.

Every single item of data that is in the visible Web needs to be crawled, indexed, and evaluated against all the other items of data and then weighed against a particular search query. The net result of this approach is that the traditional first page of Google everyone strove to rank for in the past has now largely disappeared.

While all of this might make it sound like SEO as we’ve known it is dead, nothing could be further from the truth (though if this were the case, it might have made for a much shorter book).

Semantic Search and SEO

Basic SEO factors are still in play. Links are still important. Keywords still play a role. Content still needs to be created. But as more and more variables are added to the picture, the value ascribed to each one decreases. This makes it difficult to pick specific aspects of search engine optimization, focus on them to the exclusion of everything else, and expect that to be enough.

This is an important change. In the past, you could get away with thin content, for instance, if you had a large number of links coming in that would boost it in search. You could get away with poor-quality design if you had sufficient keywords to draw in the "crowds" to your page through search. You could get away with links in some suspect neighborhoods if you had a sufficiently large number of links for the relatively small percentage of bad links to be overlooked.

You could, in short, take some shortcuts that might have been "bad" when you planned to be "good" once you got where you needed to in terms of search ranking. There was an expectation that the end result justified the risks and that things would balance out in your favor, eventually.

This is no longer the case. Because Google now needs to deliver high-quality, high-confidence results in search it has to have confidence in the content presented. The veracity factor becomes critical. This makes every activity intended to optimize a website—design, content, website structure, keywords, traffic, traffic behavior, social network footprint, links, comments, and citations—crucial. The list is far from exhaustive; everything that helps build a data-driven impression of what a website is all about and the quality of its content now becomes an element or activity you need to consider.

The reason this has happened has to do with two things that are synonymous with semantic search: entities and structured data.

Entities and Semantic Search

An entity is something that exists in itself. It can be real, like a car, or fictional, like a film or book character like Harry Potter. In either case, it possesses properties, qualities, and attributes that make up the "thing" it represents. All of these are language-independent, though obviously when an entity is described we do need a language with which to describe it. The properties, qualities, and

attributes, along with associated entities, form the Knowledge Graph that is used to define new entities.

Entities are at the heart of what Google calls the transition from “strings to things.” While all the information on the Web is data, entities allow Google to understand how that information fits in and how accurate it is (the veracity aspect of semantic search). In order for an entity to be created in Google’s index, Google needs to index all the properties, qualities, and attributes about it and understand the relational connections between them. This is exactly why the Web, with semantic search, is becoming more transparent. Data is now portable. Its origin is every bit as important as the data itself. The connections between different data pieces are being indexed, and the importance of the data itself is becoming better understood.

The concept of entities has a large impact on how to pursue SEO. While tasks such as keyword research and getting links to your site remain important, you must also pursue holistic strategies to build the reputation and visibility of your business to create positive associations across the Web.

This brings SEO and marketing a lot closer than they have ever been and makes SEO, as a whole, something that should be part of the DNA of a business rather than a bolt-on activity that can be picked up and dropped as the need arises.

So, how do we begin? What are the guiding principles that you need to have in mind in this new world of SEO? Funnily enough, the concept is as simple to plan as it is difficult to apply. It starts off from the very basic questions of: Who? Why? How?

If you cannot answer these three questions successfully—that is, in a way that expresses a distinct and unique identity for your business—then chances are good that neither can Google or your prospective customers. Your SEO, then, is governed by activities that make sense at a technical level but not at a brand identity one. Semantic search is all about establishing that identity, even —*especially*, one might argue—from a business point of view. This is what helps with the formation of entities in the Google search index.

Entities then become high-trust points that help Google’s semantic search understand the value of information better. As you engage in your overall digital marketing strategy, keep these three areas of concern in focus:

- Trust
- Authority
- Reputation

These three aspects, more than anything else, will help your business find its audience, keep it, and grow.

Structured Data

Structured data is the label applied to a number of markup formats that allow Google to better understand the data it is indexing. Structured data, then, is simply *metadata* (data about data) implemented for search engines rather than people.

Google, Microsoft (Bing), and Yahoo! worked to establish Schema.org, which is an independent, W3C-approved way of implementing structured data across the Web (see the section "Schema.org" for more information). Unfortunately, many SEOs believe that this is a shortcut to better rankings, which it's not. Semantic search is *all* about structured data. The entire effort that Google has undertaken involves indexing the unstructured data that is found across the Web and then placing it in structured data format within its index.

That does not mean, however, that structured data on a website is a ranking signal. It helps in better indexing, but ranking depends upon other factors including the quality of the content, its value, uniqueness, and even freshness. Despite the fact that Google is one of the founding organizations behind the Schema.org structured data markup initiative, Google will also attempt to extract entity information from unstructured data through its own efforts.

There are several good reasons for this:

- Adoption (structured data markup is notoriously difficult to implement if you do not know any coding)
- Accuracy (the moment human agents are involved in the markup of data, mistakes happen)
- Consistency (even when structured data is applied without errors, there are

still differences in the categorization of content and confusion over how to best apply semantic identifiers)

- Reliability (there will always be a temptation to game search by implementing structured data markup in ways intended to boost ranking; Google has already had a number of manual action penalties in search designed to remove such spammy results)

The million-dollar question is: is there anything you can do to help Google index your site better if you do not implement structured data markup?

The answer is yes: implement all the search engine optimization tools you have in your arsenal in a way that makes sense for a human user first, and a search engine second.

Namely, your on-page SEO should help a reader better navigate your content and make sense of it at a glance. The keywords, synonyms, and entities you use in your content should do the same. Any links you include, and the anchor text of those links, must similarly fill in those blanks.

If you're running a brick-and-mortar business, all the relevant information should be included on your pages, such as your name, address, and phone number. You should interlink your web properties (such as your site and social media accounts) with your Google+ presence.

Where possible, on your site, make use of Google's structured data highlighter tool. Finally, make use of Google My Business and ensure you have a cohesive presence on the Web, whose effectiveness you can measure. Use Schema.org to help search engines better understand the content of your pages.

In addition, build lots of positive relationships on the Web that help drive signals of trust and authority back to your website and business.

NOTE

A special thanks to David Amerland for his contributions to this portion of the chapter.

Schema.org

Schema.org is best viewed as part of a much larger idea, one that traces its origins back to the foundational concepts of the Web itself, and its progenitor, Tim Berners-Lee. In their seminal article in *Scientific American* in 2001, Berners-Lee, James Hendler, and Ora Lassila described a semantic web that "will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page...will know not just that [a] page has keywords such as 'treatment, medicine, physical, therapy'...but also that Dr. Hartman *works* at this *clinic* on *Mondays*, *Wednesdays* and *Fridays*."[15]

Schema.org is arguably one of the most practical, accessible, and successful outcomes of the semantic web movement to date. With the marketing prowess of Google, Yahoo!, Bing, and Yandex behind it, and with the powerful incentive of gaining additional, more inviting shelf space in the SERPs, it's no surprise that webmasters are adopting Schema.org at a rapid pace. And Berners-Lee et al.'s words now read like a prophetic description of the search engine spiders crawling the Web and extracting meaning for display in enhanced search results.

At its core, Schema.org is about standardizing and simplifying the process of adding semantic markup to your web pages, and providing tangible benefits for doing so. The most visible such benefits come in the form of *rich snippets*, such as the star ratings and price range shown in Figure 6-49.

However, it's clear that Schema.org markup plays a larger, and perhaps expanding, role in how the SERPs are constructed. Other benefits now attributed to Schema.org include local SEO ranking benefits received from clearly communicating a business's so-called NAP (name, address, phone number) information by marking it up with Schema.org, and even supplying Google with information that can appear in the knowledge panel and "answer box" results (see Figure 6-50 and Figure 6-51).

Figure 6-49. Rich snippets from Google SERPs

Figure 6-50. Google SERPs knowledge panel on Tim Berners-Lee

Figure 6-51. Google answer box for Tim Berners-Lee query

Before Schema.org, semantic markup was largely the province of academia, research and development, specialty niche businesses, and others with specific requirements to exchange and understand data in a deeply meaningful way. With Schema.org, the local pizza joint can hope to have “5 star reviews” jump off the search results page; local governments can publicize civic events and have that information represented in the SERPs, providing “instant answers” to searchers; and the list goes on. With such practical benefits in mind, and with the simplified approach of Schema.org over its big brothers like RDFa, many people responsible for building web pages are making the effort to incorporate this markup into their sites.

Overview

Schema.org markup communicates the meaning of web pages to computer programs that read them, like search engine spiders. While humans can often infer the meaning of words on a page through a number of contextual clues, computer programs often need help to extract such meaning. Let’s walk through a simple example. Imagine you have a page that displays information about the book *20,000 Leagues Under the Sea*. You might create such a page with the following HTML code:

```
<div id="book">
<h3>20,000 Leagues Under the Sea</h2>
<img src="20000-leagues-under-the-see-book-cover.png" >
<h3>Author: Jules Verne<h3>
<h3>Rating: 5 stars, based on 1374 reviews</h3>
<h3>ISBN: 978-1-904808-28-2</h3>
</div>    +
```

After being marked up, the source code might look like Figure 6-52. The Schema.org microdata markup is highlighted, and explained after the figure.

```
<div id="book" itemscope itemtype="http://schema.org/Book">
   <h3 itemprop="name">20,000 Leagues under the Sea</h2>
   <img itemprop="image" src="20000-leagues-under-the-sea-book-cover.png" />
   <h3>Author: <span itemprop="author">Jules Verne</span></h3>
   <h3>Rating: <span itemprop="aggregateRating" itemscope
       itemtype="http://schema.org/AggregateRating">
           <span itemprop="ratingValue">5</span> stars, based on
           <span itemprop="ratingCount">1374</span> reviews
           <meta itemprop="worstRating" content="1">
           <meta itemprop="bestRating" content="5">
   </h3>
   <h3>ISBN: <span itemprop="isbn">978-1-904808-28-2</span></h3>
</div>
```

Figure 6-52. Annotated Schema.org markup

Line 1: `itemscope`

Adding this to a container element, in this case a `<div>`, is the way to begin defining an entity. This attribute makes the `<div>` element the outermost, enclosing type definition for the entire book entity. The `itemtype=http://schema.org/Book` attribute, also added to the `<div>` element, declares the type of this entity. Together, this makes the entire `<div>` a container for a single book type entity.

Line 2: `itemprop="name"`

Adding `itemprop` to an HTML element defines it as the container for a property. In this case, the property is the name of the book, and the value is the inner text of the `<h3>` tags, `20,000 Leagues Under the Sea`.

Line 3: `itemprop="image"`

Similar to the name `itemprop`, but the value of this property is the URL referenced in the `src` attribute of the `<img>` tag.

Line 4

Compare this to line 2. In line 2, the inner text of the `h3` element was our exact title. Here, we also have a label (`"Author:"`), which is not part of the

actual author property. To keep our browser display looking the same as the original but omit the `"Author:"` label from our author property, we use this construct.

Lines 5 and 6

Our item property in this case is not a simple text string or URL, but rather itself another item—a `schema.org/AggregateRating`. It is simultaneously a property of the book (so it uses the `itemprop` attribute) as well as a type itself (so it uses `itemscope` and `itemtype`, as we saw in line 1 for our outermost book type).

Lines 7 and 8

These lines add properties for the `aggregateRating`, in much the same way we defined `name` and `author` in lines 2 and 4. Note the careful enclosure of the data with `<span>` tags so as to include only the data itself, not the surrounding labels, in our property. This is the same technique we used in line 4.

Lines 9 and 10

These `itemprop`s contain information needed to provide a context for the item rating (namely, that our scale is 0 to 5, with 0 being worst and 5 being best), but which is not displayed on the page. In the previous examples, the values of the properties came from the inner text of the HTML. In this case, there is no text to display in the browser, so we use the value attribute on the `<span>` element.

Line 12

This code defines the ISBN property with an `itemprop`, again using a `<span>` element to keep the display and the data cleanly separated.

Schema.org is a simple idea, with a mostly simple implementation, but getting it right can be tricky. Fortunately, the operating environment is pretty forgiving, and there are a few tools that help ease the task. Search engines understand that

most webmasters aren't structured data gurus with a deep understanding of ontologies and advanced notions of relations, entities, and other such concepts. Thus, they are generally quite adept at figuring out what you mean by your Schema.org markup, even if there are errors or ambiguities in how you say it. Clearly you should strive to be accurate, but you should approach this exercise knowing that you don't have to understand every single nuance of structured data markup, or strive for perfection in order to succeed.

How to Use Schema.org

Let's first talk about the best way to approach using Schema.org. Semantic markup is designed to help you provide meaning and clarity about what your website and each web page on it are about, so you should be clear about this before attempting to implement Schema. Think real-world tangible objects, or in semantic markup parlance, *entities*.

For example, if you're a purveyor of fine linen, your site may have lots of pages related to pillowcases, bed sheets, duvet covers, and so on. Your pages are "about" these entities. If you're willing to make the common conceptual leap here, you could say these entities "live on" your web pages. Job one is to figure out how to map these entities to Schema.org's catalog of "types."

At this level of thinking, Schema.org is a large and ever-growing and evolving catalog of "types" (Schema.org documentation sometimes uses the word *items* in place of *types* here) that attempts to classify everything that can be represented on web pages. Let's take a look at the Schema.org page for a `Book` type, shown in Figure 6-53. The idea is straightforward. The type definition identifies the key attributes that you would use to uniquely describe an "instance" (that is, a single, real-world example) of this type.

schema.org

Search

Home Schemas Documentation

Thing > CreativeWork > Book

A book.

Property	Expected Type	Description
Properties from Book		
bookEdition	Text	The edition of the book.
bookFormat	BookFormatType	The format of the book.
illustrator	Person	The illustrator of the book.
isbn	Text	The ISBN of the book.
numberOfPages	Integer	The number of pages in the book.
Properties from CreativeWork		
about	Thing	The subject matter of the content.
accessibilityAPI	Text	Indicates that the resource is compatible with the referenced accessibility API (WebSchemas wiki lists possible values).
accessibilityControl	Text	Identifies input methods that are sufficient to fully control the described resource (WebSchemas wiki lists possible values).
accessibilityFeature	Text	Content features of the resource, such as accessible media, alternatives and supported enhancements for accessibility (WebSchemas wiki lists possible values).
accessibilityHazard	Text	A characteristic of the described resource that is physiologically dangerous to some users. Related to WCAG 2.0 guideline 2.3. (WebSchemas wiki lists possible values)
accountablePerson	Person	Specifies the Person that is legally accountable for the CreativeWork.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
alternativeHeadline	Text	A secondary title of the CreativeWork.
associatedMedia	MediaObject	The media objects that encode this creative work. This property is a synonym for encodings.
audience	Audience	The intended audience of the item, i.e. the group for whom the item was created.
audio	AudioObject	An embedded audio object.
author	Person or Organization	The author of this content. Please note that author is special in that HTML 5 provides a special mechanism for indicating authorship via the rel tag. That is equivalent to this and may be used interchangeably.
award	Text	An award won by this person or for this creative work. Supercedes awards.
citation	Text or CreativeWork	A citation or reference to another creative work, such as another publication, web page, scholarly article, etc.
comment	Comment or UserComments	Comments, typically from users, on this CreativeWork.
commentCount	Integer	The number of comments this CreativeWork (e.g. Article, Question or Answer) has received. This is most applicable to works published in Web sites with commenting system; additional comments may exist elsewhere.

Figure 6-53. Schema.org definition of Book

It may help if you open this page in your web browser as we discuss it. Note that the Schema.org definitions are frequently reviewed and updated based on active user feedback, so you may even see minor variations on the current page. But the overall structure will likely remain very similar, and the major elements of the page are central to Schema.org. First, note the simple description, confirming that this is, indeed, the model for a book. Let’s ignore the Thing > CreativeWork > Book breadcrumb for now; we’ll come back to that later.

Next comes a table of *properties*—what we might think of as the attributes that uniquely describe our individual entity—which, in this example, are the things

that describe the book *20,000 Leagues Under the Sea*. Each property has a name (the Property column), an Expected Type, and a Description. The Expected Type tells us whether this property is simply a text value (like a name), or something more complex—that is, a type itself. For example, the `illustrator` property should contain not the name of the illustrator, but a full person entity, using the *http://schema.org/Person* type definition (which, as you would expect, itself contains a `name` property, and that's where you include the illustrator's name).

As you begin examining a possible mapping of your entities to Schema.org types, you'll often encounter this nesting of types within types. While many of the properties of an entity are simple descriptions (text strings like `"blue"`, `"extra large"`, or even `"Jan 17, 2015"`), others are more complex and entities in their own right. This is the concept of composing larger-scale things from a collection of smaller ones, as in describing a car (a single entity in its own right) as being made up of an engine, a chassis, wheels, interior trim, and so on (all entities themselves).

Extending this idea further: to an auto mechanic, an engine—a component of our car—is itself the composite thing (the big entity). To understand the engine in more detail, it's important to break it down into its own component entities, like carburetors, spark plugs, filters, and so forth.

Schema.org, then, is a set of conventions for modeling complex things in the real world, and marking them up in a way that search engines can consume, leading them to a deeper understanding of web pages. This deeper understanding in turn leads to many current and future benefits when the search engines subsequently present that data back to users in compelling, contextually relevant ways.

There's one more preliminary concept we should cover; it seems complicated at first but isn't once we break it down. One thing you'll notice as you browse Schema.org's types is that each one lives within a hierarchical family tree. We saw this earlier with the breadcrumb on the Books page, shown again in Figure 6-54.

Figure 6-54. Showing the "inheritance" hierarchy for the Book type

It's important to note that this kind of hierarchy, referred to among computer scientists as *inheritance*, is different than the composition hierarchy (a car made up of an engine) example we discussed earlier. The Schema.org type hierarchy is a way of categorizing things from most generic to most specific—what we call an *ontology*. Its form is similar to the well-known animal kingdom charts we've all seen, or the myriad other classification schemes we all tend to take for granted—often represented on web pages with features like breadcrumbs, navigation menus, and faceted navigation filters.

The key point to remember here is that when choosing the Schema.org type to model your entities, it's always best to choose the most specific type you can. That is, choose `Restaurant` over `LocalBusiness` (if, indeed, you're operating a restaurant!). Choose `Book` over `CreativeWork` for books, and `HighSchool` over `EducationalOrganization` for high schools. Doing so ensures you are giving the most specific information possible to the search engines, rather than settling for generic descriptions.

With that background covered, let's run through the general plan for adding Schema.org markup to your website. Here are the six major steps:

1. Determine the Schema.org types that best describe the entities represented on your web pages, which may be different for each of your different page archetypes.

2. For each page archetype you're modeling, perform a detailed mapping of the information elements displayed on the page to the Schema.org type

properties.

3. Choose the approach you will use to express the Schema.org markup.

4. Edit the HTML document templates, or update the CMS settings, or modify the scripts—whatever best describes how your pages are generated—to incorporate the Schema.org markup.

5. Test the markup to see if your syntax is accurate, and if you've properly modeled complex entities.

6. Monitor how well the search engines are consuming your structured data, and whether and how that data is being presented in the SERPs.

Let's take these one at a time in more detail.

Step 1: Determine Schema.org types

In this step, you think carefully about which web pages to mark up while simultaneously browsing the Schema.org website (actually, the browsing capability is fairly limited as of the time of this writing, so you might be better off searching for types; see Figure 6-55).

For example, if your website is about community theater groups, and displays one page for each theater group along with the upcoming list of their performances, you would begin by searching at *http://schema.org* for something like *theater*. The resulting page looks like what's shown in Figure 6-56. Scanning the results, we quickly spot `TheaterGroup` as a likely candidate for the type of our main entities.

Figure 6-55. Schema.org includes a search box at the top of each page of on the site

MovieTheater – schema.org
Thing > Place > CivicStructure > MovieTheater. A movie **theater**. Property, Expected Type, Description. Properties from CivicStructure · openingHours · Duration ...
https://schema.org/Movie**Theater**

TheaterEvent – schema.org
Documentation · Schemas · Home · Thing > Event > TheaterEvent. Event type: **Theater** performance. Property, Expected Type, Description. Properties from Event ...
schema.org/**Theater**Event

TheaterGroup – schema.org
Thing > Organization > PerformingGroup > TheaterGroup. A **theater** group or company—for example, the Royal Shakespeare Company or Druid Theatre.
https://schema.org/**Theater**Group

WatchAction – schema.org
Schema.org is a set of extensible schemas that enables webmasters to embed structured data on their web pages for use by search engines and other ...
schema.org/WatchAction

Full Hierarchy – schema.org
The Type Hierarchy. Here is the entire hierarchy in a single file. Note: as of 2014- 04-04 this tree is not entirely up to date. Additional types have been added: see ...
https://schema.org/docs/full.html

WebPage – schema.org
A web page. Every web page is implicitly assumed to be declared to be of type WebPage, so the various properties about that webpage, such as breadcrumb ...
schema.org/WebPage

Figure 6-56. Search results for "theater" on Schema.org

Taking a closer look at the `TheaterGroup` page at *http://schema.org/TheaterGroup* (Figure 6-57), we can see a few of our core concepts at work:

- `TheaterGroup` is part of a logical hierarchy, starting with the most generic type (`Thing`—actually the topmost ancestor of all Schema.org types), then proceeding to more and more refined types: `Organization`, `PerformingGroup`, `TheaterGroup`.

- A `TheaterGroup` is composed of many elements (called properties), some of

them simple like the name of the group, and some of them actual types in their own right (such as `address`, `aggregateRating`, `employee`, etc.). Examining the list of properties confirms our belief that this is the best type for describing our local theater entities on our web pages.

schema.org

Search

Home Schemas Documentation

Thing > Organization > PerformingGroup > TheaterGroup

A theater group or company—for example, the Royal Shakespeare Company or Druid Theatre.

Property	Expected Type	Description
Properties from Organization		
address	PostalAddress	Physical address of the item.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
brand	Brand or Organization	The brand(s) associated with a product or service, or the brand(s) maintained by an organization or business person.
contactPoint	ContactPoint	A contact point for a person or organization. Supercedes contactPoints.
department	Organization	A relationship between an organization and a department of that organization, also described as an organization (allowing different urls, logos, opening hours). For example: a store with a pharmacy, or a bakery with a cafe.
dissolutionDate	Date	The date that this organization was dissolved.
duns	Text	The Dun & Bradstreet DUNS number for identifying an organization or business person.
email	Text	Email address.
employee	Person	Someone working for this organization. Supercedes employees.
event	Event	Upcoming or past event associated with this place or organization. Supercedes events.
faxNumber	Text	The fax number.
founder	Person	A person who founded this organization. Supercedes founders.
foundingDate	Date	The date that this organization was founded.
globalLocationNumber	Text	The Global Location Number (GLN, sometimes also referred to as International Location Number or ILN) of the respective organization, person, or place. The GLN is a 13-digit number used to identify parties and physical locations.

Figure 6-57. TheaterGroup type from Schema.org

It's during this step that you want to deal with the question "What is this page about?" and choose the Schema.org type that best describes the overall contents of the page. Often this choice is obvious, but at times it can be tricky. For example, on a page with a product for sale, should you choose *schema.org/Offer* or *schema.org/Product* to model the page?

Examining both pages on the Schema.org site, you can see that an `Offer` has a property called `itemOffered`, with an expected value of `Product`. This means that you can describe the contents of the page as an `Offer` (Schema.org's concept of something for sale), where the item for sale (the `Product`) is contained within the `Offer`, using the `itemOffered` property.

Alternatively, you could use the `Product` type, which has a property called `offers` that can, as you might expect, contain one or more `Offer` types. The decision probably depends on the overall purpose of the page. If the page is a detailed product page, describing many attributes of the product, and the offer information is just a piece of that, it probably makes sense to model the page as a `Product` and include the offer information in the `itemOffered` property. However, it's not out of the question that you could invert this model.

Either of the approaches to the `Product`/`Offer` model is valid, as both convey the meaning that you want. But take another look at `Offer`. You can see that it is a complex concept, and has many properties that are themselves types (for example, `aggregateRating`). Other complex nesting of types and attributes can easily arise, and it's important to model these out in a way that best matches the meaning of the page. The best approach often won't be obvious at this stage of analysis, so you may need to revisit your thinking after you complete step 2 of the process.

Step 2: Map Schema.org properties to elements on the web page

The first step here is to survey the various data elements displayed on the web page, and match them up with the appropriate Schema.org types and properties you selected in step 1. In this step, you may discover relationships that resolve some of the potential ambiguities from step 1.

For example, continuing the `Product`/`Offer` discussion, let's assume that one of the items displayed on the page is an overall rating—say a value on a scale of 1 to 5—representing user evaluations of the product. We notice that both `Product` and `Offer` have a property called `aggregateRating`, so this hasn't quite settled our debate on which type to model the page on.

Let's also assume that we display several different prices—perhaps for new or used versions of the product, or with different shipping options or different currencies. It now starts to become obvious that we should model the entire page as a `Product` that contains multiple `Offer`s and a single `aggregateRating` that applies to the `Product` itself. Finally, this is starting to take shape!

You might notice that there are properties defined on the Schema.org type that you're not currently displaying to browsers, but which you have access to.

Continuing with our `Product` example, perhaps your web application's database stores the MPN (manufacturer's part number), but you don't choose to display that on the page. What should you do?

Ideally, you want a very high degree of consistency between what you mark up and what's visible to "normal users" via web browsers. Technically, there are mechanisms that allow you to communicate to the search engines metadata about your entities that shouldn't be displayed to users (we saw this earlier in our `aggregateRating` example, and we'll explore that example a bit more momentarily).

However, it's important to use these mechanisms sparingly, and not be tempted to stuff a lot of extra data into the Schema.org markup that is not visible to human users. In the MPN case, our choice should be between adding this as a visible element on the page (and then of course adding it to our Schema.org markup), or forgoing it entirely. As you think about this, it should become clear that marking up a lot of data that is not displayed to the user is conceptually something a spammer might do, and for that reason the search engines frown on it.

What are the valid reasons for marking up nondisplayed data? Usually it's because you need to convey some different context that is obvious to people, but not to search engine spiders. For example, when you mark up an `aggregateRating`, you're strongly encouraged to specify the scale; that is, if you display 4 stars for a review on a scale of 0 to 5, this is usually quite clear in the visual representation, but it needs to be stated explicitly in the Schema.org markup. Thus, `aggregateRating` entities have `worstRating` and `bestRating` properties, and we want to supply the values 0 and 5, respectively, corresponding to our star rating scale. We saw this in the sample code for our book at the beginning of the chapter.

Upon completing this step, you should have a complete mapping between the data displayed on the page and the various Schema.org types and properties that make up your model. Your model may be simple or complex with multiple levels of nesting. It's best to make all these decisions before you begin actually implementing Schema.org on the page.

Step 3: Choose your implementation technique

For most people, this step means “go mark up the page.” Sounds simple, right? And for some pages, especially those that are template-driven with mostly static data, it should be fairly straightforward. Or, if you’re lucky enough to be using a content management system or publishing platform that has built-in support for Schema.org, you can do most of the actual implementation by setting a few configuration parameters.

For other, more dynamic sites that generate their pages through a complex pipeline of page generation programs, tweaking things to insert the right tags in the right place can be far more difficult. And for these types of sites, validating that the generated schema is correct is also challenging, as the Schema.org markup may be sporadically injected into kilobytes of code.

The primary implementation technique is to edit templates and/or modify page generation programs to insert the microdata markup as needed to produce the desired final output. The key thing is to have a clear mapping of the model from step 2 showing the final desired HTML with microdata markup inserted, and use this to validate that the final page produced by the web server matches the model. As we’ll see in step 5, there are some tools that can help with this verification as well.

For those who don’t have access to the code or backend systems, or who want a simpler approach, Google offers the Structured Data Markup Helper, as part of Google Search Console. This is a proprietary Google tool that allows you to annotate a page, using a point-and-click editor (see Figure 6-58). It’s actually just an alternative way of providing the same data you provide via Schema.org microdata markup, but you are instead feeding it directly to Google and do not change the page source code at all.

So why doesn’t everyone just do this? There are two good reasons why this isn’t the best fit for everyone. First, the information is available only to Google, not to other Schema.org-aware search engines (or other applications that may make use of Schema.org markup). Second, as is often the case with this kind of tool, the visual editor is more limited than the markup syntax in its ability to express rich and complex information.

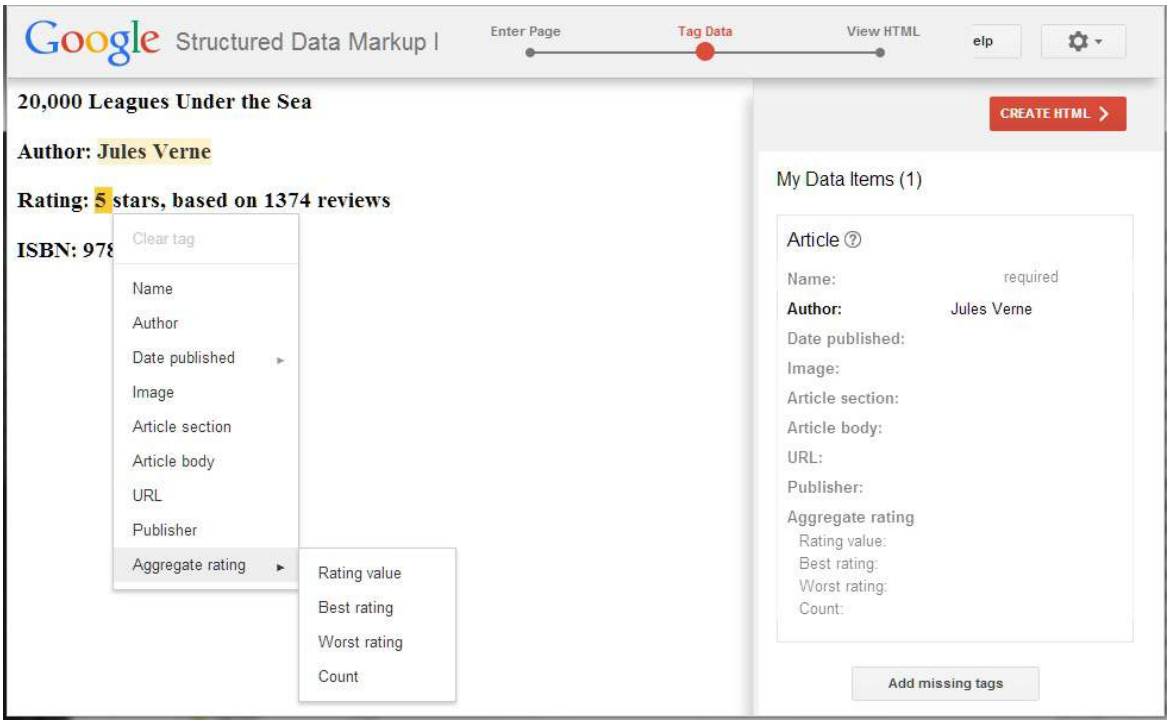


Figure 6-58. Google Structured Data Markup Helper

Looking to the future, another alternative may be on the horizon, and is something to keep an eye on. Google and others are already beginning to make use of a format known as JSON-LD for expressing Schema.org markup. This format is showing up in limited, specialized circumstances, but it seems apparent that JSON-LD may soon become a full-fledged alternative to microdata for all Schema.org expression.

The beauty of JSON-LD is that it provides a way to isolate all of the Schema.org information into a single string of code, rather than expressing it by embedding markup within the HTML document itself. This has the possibility of solving many of the more complex issues associated with implementing Schema.org on complex, dynamic sites.

Step 4: Implement the changes to generate the target Schema.org code

This step is really just saying, “Now it’s time for your web developers to go breathe some life into your creation”; that is, go get these pages served up by your web server! This is where the content management system is tweaked, the

templates are updated, the page production programs are modified, and so on.

Step 5: Test

When you reach this stage, your web server is shooting out bundles of HTML with tidy little microdata tags embedded in it that add meaning and structure to your data. At this point, the generated Schema.org markup code should be syntactically correct, and should express the right model—that is, the whole composition of smaller properties and types into larger properties and types needed to accurately model the information displayed on our pages. Of course it's important to verify this.

The hard way to do that is to examine the generated code by hand, looking for the opening and closing tags, and ensuring that all the data is there, nested properly. Fortunately, there's an easier way (though you should still be prepared to roll up your sleeves and dig into the code to debug potential problems).

The easier way is to use one or more of the tools available to verify your Schema.org microdata markup. Perhaps the best known of these tools is Google's Structured Data Testing Tool, which is an elegant utility that examines your page (either directly by supplying a URL, or alternatively by cutting/pasting HTML source code) and gives you feedback on the structured data it finds. Figure 6-59 shows such a result.

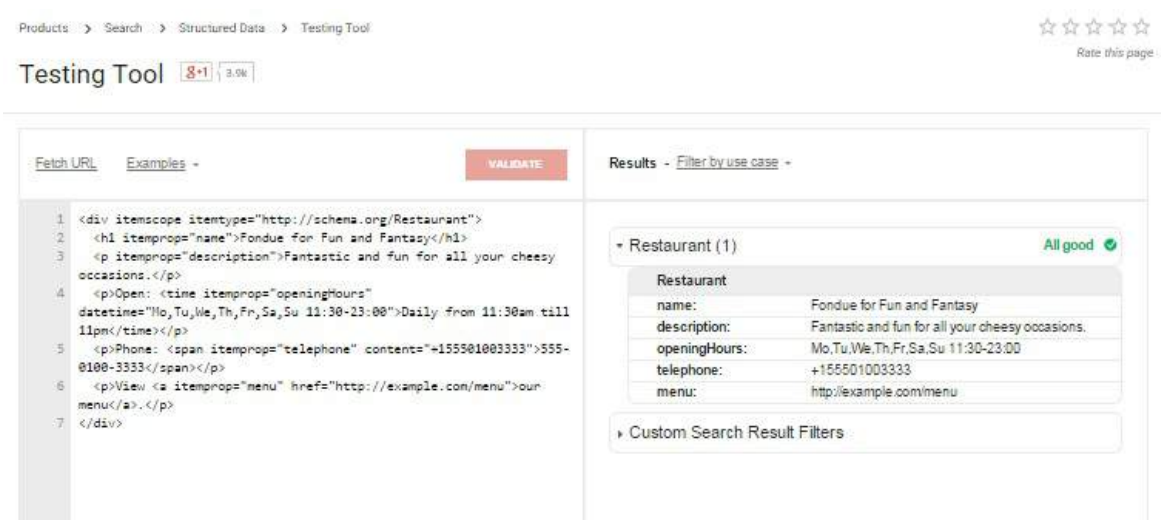


Figure 6-59. Google's Structured Data Testing Tool output

The output of this tool has a bit of an arcane formatting convention. Note that

our book shows up as the first item. The book has a number of properties, among which is our `aggregateRating`; recall that this is itself another type. When this composition or nesting occurs properly, you see the output shown in Figure 6-59.

Here, the nesting relationship is shown by Item 1 in the item value field for the `aggregateRating` property of the `Book`, followed immediately by the Item 1 output. So the value Item 1 in the first field ties together with the name of the second item shown, and shows that Item 1—the rating—is properly contained within the book entity, as specified by the Schema.org definition for a `Book`. Google will keep incrementing these numbers for as many embedded types as you have on the page.

Summary

We have seen that Schema.org is a standard for providing search engines (and potentially other applications) with structured data describing the meaning of website content. The notion of data structuring is actually quite intuitive, and maps well to the way we commonly categorize things like product catalogs, biology, library card catalogs, and many other collections of related items. This intuitive, webmaster-friendly approach has led to rapid adoption of Schema.org by the webmaster and content production communities. Currently, the most common way to structure data with Schema.org is to add microdata markup to HTML documents. Search engines use this data to extract meaning, and enrich SERPs with rich snippets, answer boxes, and knowledge panels, providing a more relevant and deeper search result. Implementing Schema.org can bring these benefits to both users and publishers today, and can help set the stage for publishers to gradually delve more deeply into the emerging world of semantic search in the coming years.

NOTE

A special thanks to John Biundo for his contributions to the Schema.org portion of this chapter.

Google Authorship and Author Authority

One of the most interesting insights into the mind of Google, as it were, was the

three-year experiment known as Google Authorship.

Google Authorship was a program that allowed online authors to identify and verify their content with Google. This was accomplished by a two-way link between the author's content across the Web and his Google+ profile.

While the Authorship program was officially discontinued by Google on August 28, 2014, it is likely that Google's interest in the value of the authority, trust, and reputation of an author in a given topical area is undiminished.

A Brief History of Google Authorship

The roots of Google Authorship lie in a patent originally granted to Google in 2007, called Agent Rank. The patent described methods whereby a search engine could identify distinct "agents" (one of which could be the author or authors of a web document) and assign a score to each agent that could then be used as a factor in search rankings.

Google didn't appear to do anything with this patent until June 2011, when Google's Othar Hansson announced in a blog post that it would begin to support the use of the HTML5 `standard rel="author"` and the XFN `standard rel="me"`, and that webmasters could use that markup to identify authors and author profiles on their sites.[16]

The next major step in Authorship came just 21 days later, when Google unveiled its new social network, Google+. Google+ provided personal profiles that Google could use to verify authors using the `rel="author"` markup.

This intention was confirmed in a YouTube video by Othar Hansson and Matt Cutts published on August 9, 2011, titled "Authorship Markup". In the video, Hansson and Cutts explained that Google wanted web authors to have Google+ profiles, and that they should link from the "Contributor To" link sections of those profiles to each domain where they publish content. Over time, Google offered several options by which the publisher could confirm the relationship by linking back to the author's Google+ profile.

In that video, Google confirmed that there could be rewards to authors who implemented Authorship markup; the immediate possible benefit was the potential for an author's profile image and byline to be shown with search results for her content. Figure 6-60 is typical of such results.

Figure 6-60. Rich snippet authorship result

Additional potential benefits mentioned by Hansson and Cutts were increased search rankings and the fact that Google might be able to use Authorship to identify the original author or a piece of web content, thus giving that author's copy precedence in search over scraped copies.

Over time, Google added several tools and features to make Authorship easier to implement and more useful for authors and publishers. This was probably the result of the problems the company saw with a lack of adoption of this markup.

The first major hint that Google might be pulling back on its Authorship experiment came in October 2013 when AJ Kohn revealed that Othar Hansson had left the Authorship team and was not being replaced.[17] In that same month, Matt Cutts revealed that Google would soon be cutting back on the amount of Authorship rich snippets shown in search, as it had shown in tests that doing so improved the quality of those results.

Cutts's words proved true in December 2013, when observers noticed a 15% reduction in the amount of author photos being shown for most queries.[18]

In June 2014 Authorship was further reduced in search as Google announced that it would no longer show author photos in results, just bylines. The only announced reason for this was to bring its mobile and desktop user experiences more into sync.

However, only two months later, as previously noted, Google announced the complete removal of Authorship data from search, and stated that it would no longer be tracking any data from `rel="author"` links. The Google Authorship program, or at least any program based on `rel="author"` links and showing rich snippets in search results, was now over.

Why Did Google End Support for rel="author"?

In his official announcement of the end of the Authorship program, John Mueller of Google Webmaster Central said, “Unfortunately, we’ve also observed that [Authorship] information isn’t as useful to our users as we’d hoped, and can even distract from those results. With this in mind, we’ve made the difficult decision to stop showing authorship in search results.”

He went on to elaborate, saying that this decision was based on user experience concerns. After three years of testing, Google was no longer seeing any particular user benefits from showing Authorship results. Mueller said that removing the Authorship results “did not seem to reduce traffic to sites.” It would seem, then, that searchers were no longer viewing these results as anything special.

What else may have factored into the decision to stop showing Authorship results? In his post Mueller mentioned that he knew that Authorship “wasn’t always easy to implement.” Could it be that low implementation rates by most sites fed Google’s decision? If Google were ever going to rely on Authorship as a signal for search, it would need to have data from a wide variety of sites.

In a study published just after the ending of Authorship, Eric Enge confirmed from a sampling of 150 top publishing sites that Authorship implementation was indeed low.[19] He found that 72% of these sites had attempted Authorship markup in some way, but out of those *nearly three-fourths had errors in their implementation*. But even worse, 71% of the 500 authors sampled from those sites had done nothing from their side to implement Authorship.

It would seem that low participation might be another reason behind Google’s decision. Google may have learned that data you want to use as a ranking factor can’t be dependent upon voluntary actions by webmasters and authors.

Is Author Authority Dead for Google?

Does the end of `rel="author"`–based Authorship mean Google has lost all interest in understanding, tracking, and making use of data concerning the authority levels of online authors? Most likely not.

For one thing, on September 2, 2014, Google was granted a patent for a system that would retrieve, rank, and display in search authors considered authoritative for a topic based on their relationship (in social networks) to the searcher.[20]

Also, Google spokesperson Matt Cutts often spoke during the last year of Google Authorship about his interest in and support for Google eventually being able to use author reputation as a means of surfacing worthwhile content in search results, but noted that he sees it as a long-term project.[21] While Cutts seems to be voicing his personal opinion in such statements, it is doubtful that he would speak so frequently and positively about the topic if it weren't actually active at Google.

Another area that seems to support the notion that Google will only increase its interest in author authority is semantic search. Semantic search involves, in part, a dependence upon the identification of various entities and the ability to understand and evaluate the relationships between them. As the original Agent Rank patent makes clear, authors of web content are certainly a useful type of entity.

Google understands that real people often evaluate authority and trustworthiness not just by a document's contents or what links to it, but by the reputation of the author. Semantic search at its simplest is a quest to enable Google's search algorithm to evaluate the world more closely to the way people do. So it makes sense that Google would continue to pursue the ability to evaluate and rank authors by the trust and authority real people place in them for a given topic.

Google+ Authors in Personalized Search

At the time of this writing there remained one large and very interesting exception to Google's elimination of Authorship rich snippets from search. Author photos and bylines can still appear for Google+ content authored by people in a searcher's Google network (Google+ circles and Gmail contacts) when that searcher is logged in to his Google+ account while searching. Figure 6-61 shows such a result.

Figure 6-61. Personalized search rich author snippet

Notice that the URLs are both from plus.google.com (Google+). The person performing this search has both Ana Hoffman and Mark Traphagen in his Google+ circles, and is searching while logged in to his Google+ account. For the search query “Google authorrank” Google found that these two people in the searcher’s network had relevant content on Google+ and so included it in the searcher’s results.

Note two points about these results:

- These results are uniquely ranked higher for this individual searcher. If he searches for the same query while logged out of his Google+ account, these results will not show up in the top results.
- The personal connection of the authors to the searcher is being emphasized by the photo and byline.

The Future of Author Authority at Google

It appears that Google remains interested in the concept of author authority as a factor in search rankings. Google is likely working on methods to identify and evaluate authors and their content that are not dependent on human publishers and authors placing links and attribution tags. When those methods are providing reliable data, Google might make these signals a ranking factor.

However, given the lessons of the first Google Authorship experiment, we might expect the following possible differences:

Author authority might be more personalized.

> That is, Google may give a greater boost to content by authoritative authors relevant to your search *if* you have some connection to or relationship with those authors.

Author authority in search will probably be less obvious.

> Google may not return to the practice of displaying rich snippet profile photos for top authors, in part because it is moving away from flashier rich snippets in general as part of its Mobile First initiative. It is therefore likely that any future author authority factor will simply be folded into the many

factors that determine search rankings and may not be apparent to the searcher.

Author Authority

Here are some tips on how to build author authority:

Publish with real names.

In order to build author authority search engines have to be able to recognize that multiple pieces of content are connected with a particular individual. Several of the following tips relate to building your personal authority both on and offline, so using your real name with your content is important.

Keep your name consistent.

In parallel with the previous tip, it is important that you use exactly the same name as the byline on all your content as well as in all your social profiles. That will help search engines gain confidence about your identity, and make it more likely that all of your online content will be used to evaluate your authority.

Cross-link your profiles.

Wherever possible, create links between all your online profiles. This is another way to help search engines have more confidence in your unique identity.

Link your social profiles to your content.

Wherever possible, create links from your social and site profiles to the sites on which you publish content. Of course, in the case of Google, it is most important to make sure that all sites on which you publish are linked from the “Contributor to” section of your profile links. Even though Google says it no longer tracks data based on `rel="author"` links to Google+ profiles, we still recommend creating links from your content or site author profiles back to your Google+ profile, as these still may give Google confidence about content that should be identified with you.

Produce content about all aspects of your field.

More and more we see indications that Google is including in measures of its topical authority how complete and well rounded the content is. It's no longer effective to merely hammer away at certain long-tail keywords. You need to build contextually rich content that looks at your subject from all sides. That doesn't just apply to individual content pieces, but also to the content across an entire site or across your profile as an author on many sites.

Produce content that goes in depth on specifics of your field.

As well as covering all aspects of your area of expertise, your content also needs to explore those areas deeply. That doesn't mean every piece of content needs to be an academic paper, or even long form. But you should be seeking as often as possible to produce content that gives a unique perspective on a topic, or that goes into more depth and detail than most other similar pieces on the Web.

Cultivate an audience.

Every content producer has to be as concerned with building a loyal audience as she is with producing quality content. That means being active on social networks, for one. Seek to build good relationships with those who might be interested in your expertise and likely to share it with their networks.

Participate in relevant conversations.

Go beyond just broadcasting your content to participating in relevant online conversations and communities. Doing that can have multiple benefits. As you contribute to such communities, you get a chance to display your expertise before a broader audience, some of whom may start to follow you. That means you are growing your audience (see above), but doing it in places where you are more likely to pick up followers with high interest in what you do.

Don't forget real-world opportunities.

Attending conferences and networking events in your field can lead to online

connections that help reinforce your online authority. This is especially true if you are a speaker or panelist at such events, or get interviewed by a media outlet. You can accelerate this effect by actively inviting people at these events to connect with you online. For example, always place your primary social profiles prominently in any presentations you do.

Incubate and promote brand subject matter experts.

Publishers should not ignore the power of individual topical authority used in conjunction with their brands. Many companies are reluctant to empower individual employees or representatives to build their own authority, but they miss a real opportunity by not doing so. People identify with, trust, and connect with a real individual long before they do with a faceless brand. Therefore, wise brands will cultivate subject matter experts (SMEs) who have a direct connection with their brand, knowing that the audience and authority those SMEs build will ultimately reflect back on the brand.

Google's Publisher Tag

Although Google has ceased to track author data using `rel="author"` links, the related `rel="publisher"` link tag is still very much supported, and can convey a number of benefits to brands using it.

Similar to Facebook and some other social networks, Google+ allows the creation of pages as distinct from profiles. While profiles are intended solely for individuals, pages allow nonpersonal entities (brands, companies, organizations, bands, etc.) to have a presence on Google+.

Google+ brand pages deserve particular attention, however, because of the way Google uses them, particularly in conjunction with Google search. Stated succinctly, a Google+ page linked from a brand's official site using `rel="publisher"` is the easiest and most direct way for Google to verify the unique identify of a particular brand. Once Google has such verification, it can begin to make use of data related to the brand in various ways.

Brands with verified Google+ pages can be eligible for three special search features:

Knowledge panel with link to Google+ page

As shown in Figure 6-62, when people search for a brand name, brands with verified Google+ pages will show a Google+ logo in their knowledge panel that links to their Google+ page.

Figure 6-62. A search for a brand name brings up the brand's verified Google+ page

The Google+ knowledge panel box shows the brand name and logo, a Google+ follow button (shows only if the searcher is logged in to Google+), the page's follower count, and a recent Google+ post.

Elevated rich snippet Google+ post results

When a searcher has a brand circled on Google+ and searches while logged in to her Google+ account, she may have a relevant Google+ post from that brand elevated to page one and highlighted with a brand logo and brand name rich snippet. An example is shown in Figure 6-63.

Figure 6-63. Rich snippet authorship result

AdWords Social Extensions

If a brand connects its Google+ Page to its AdWords account and enables Social Extensions, Google may add a Google+ annotation to the brand's ads in search. Figure 6-64 shows how this appears in the search results.

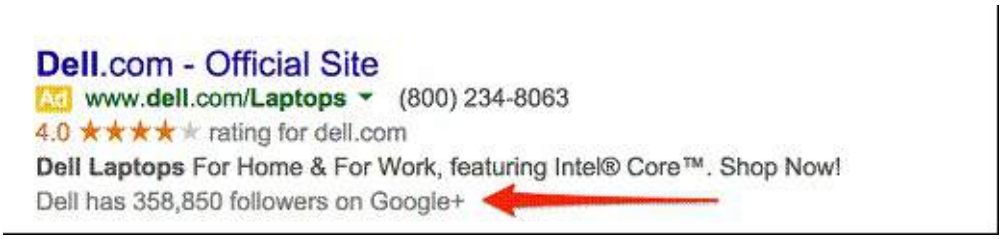


Figure 6-64. AdWords Social Extensions example

Gmail Google+ related pages widget

A brand with a verified brand page that meets certain qualifications can have a widget show in the right sidebar of Gmail when customers open an email from the brand (see Figure 6-65 for an example). The widget displays a thumbnail of a recent post from the brand's Google+ page and, if the user has a Google+ account, a follow button. For details, see *https://support.google.com/business/answer/4569086?hl=en*.

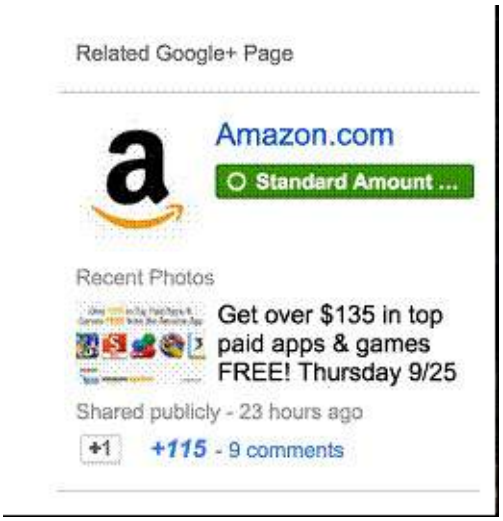


Figure 6-65. Gmail Google+ related pages widget

Verifying a Google+ brand page

For nonlocal business pages, you can verify a page by simply making the brand's official website the main URL of the page, and then linking back from the home page of that site to the Google+ page with a `rel="publisher"` attribute. A local business page must verify via Google My Business. If your business has 10 or more locations to verify, use *https://www.google.com/local/manage/?hl=en#*.

NOTE

A special thanks to Mark Traphagen for his contributions to the Authorship and publisher tag portions of the chapter.

Google's Knowledge Graph and the Knowledge Vault

The face of search is changing in significant ways. The latest incarnation of that evolution is the Knowledge Graph. Google has also begun to communicate about the Knowledge Vault, though that is just a research project as of October 2014. To get some perspective on why these are important, it is useful to review how

search has evolved.

Overview of Changes in Search Complexity

Search engines used to build results by analyzing the text strings they found on the pages of the Web. The resulting pages were quite useful, but the presentation of the results was quite simple, as shown in Figure 6-66.

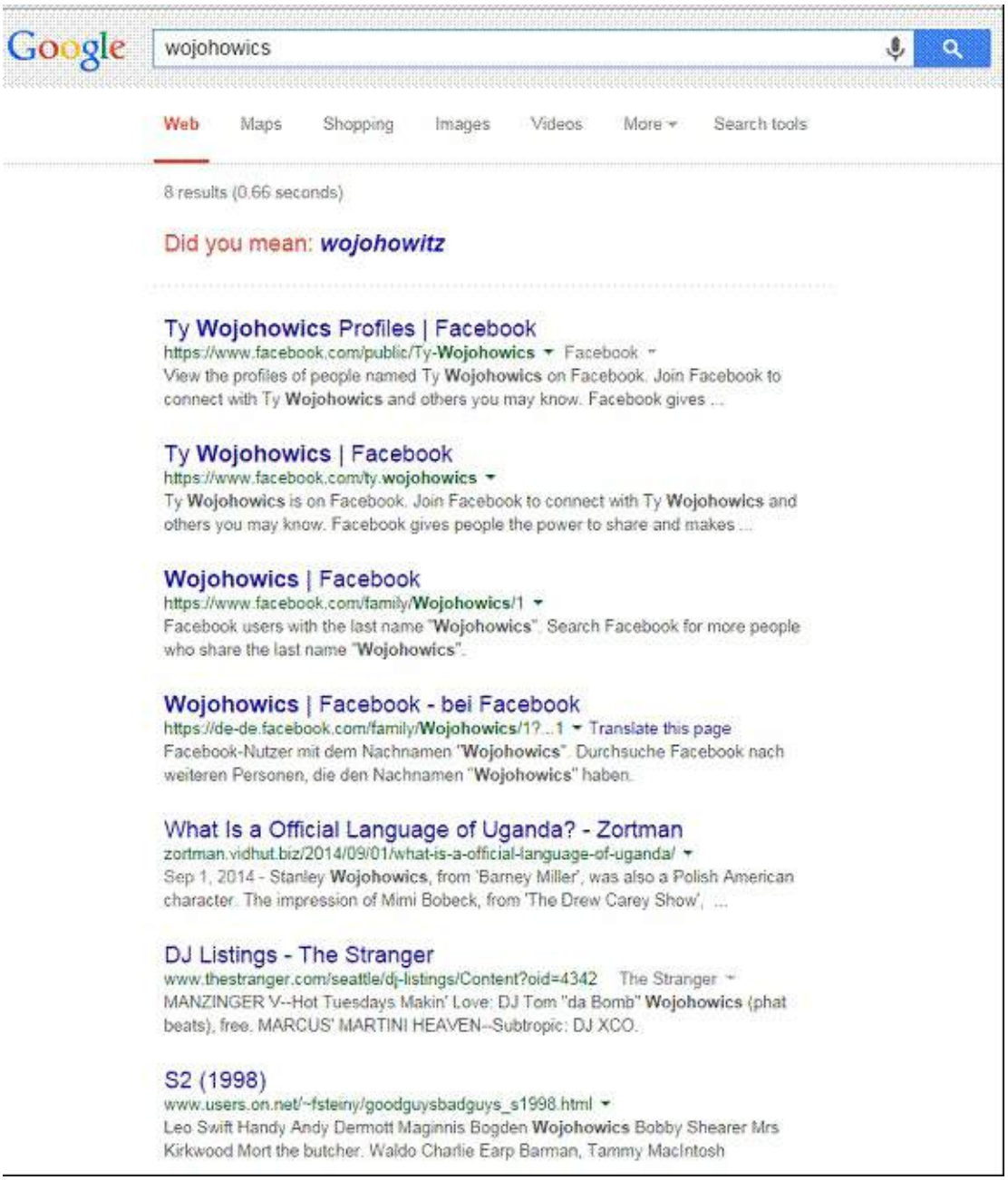


Figure 6-66. Simple text-only search results

Over time, these types of search results became known as “10 blue links” due to their simplicity. As they became more sophisticated, the engines figured out how to incorporate more types of media into the results, including videos, images, news stories, shopping results, and more.

These are typically referred to as “blended results,” and you can see an example in Figure 6-67. You can also read more about these types of results in Chapter 10.

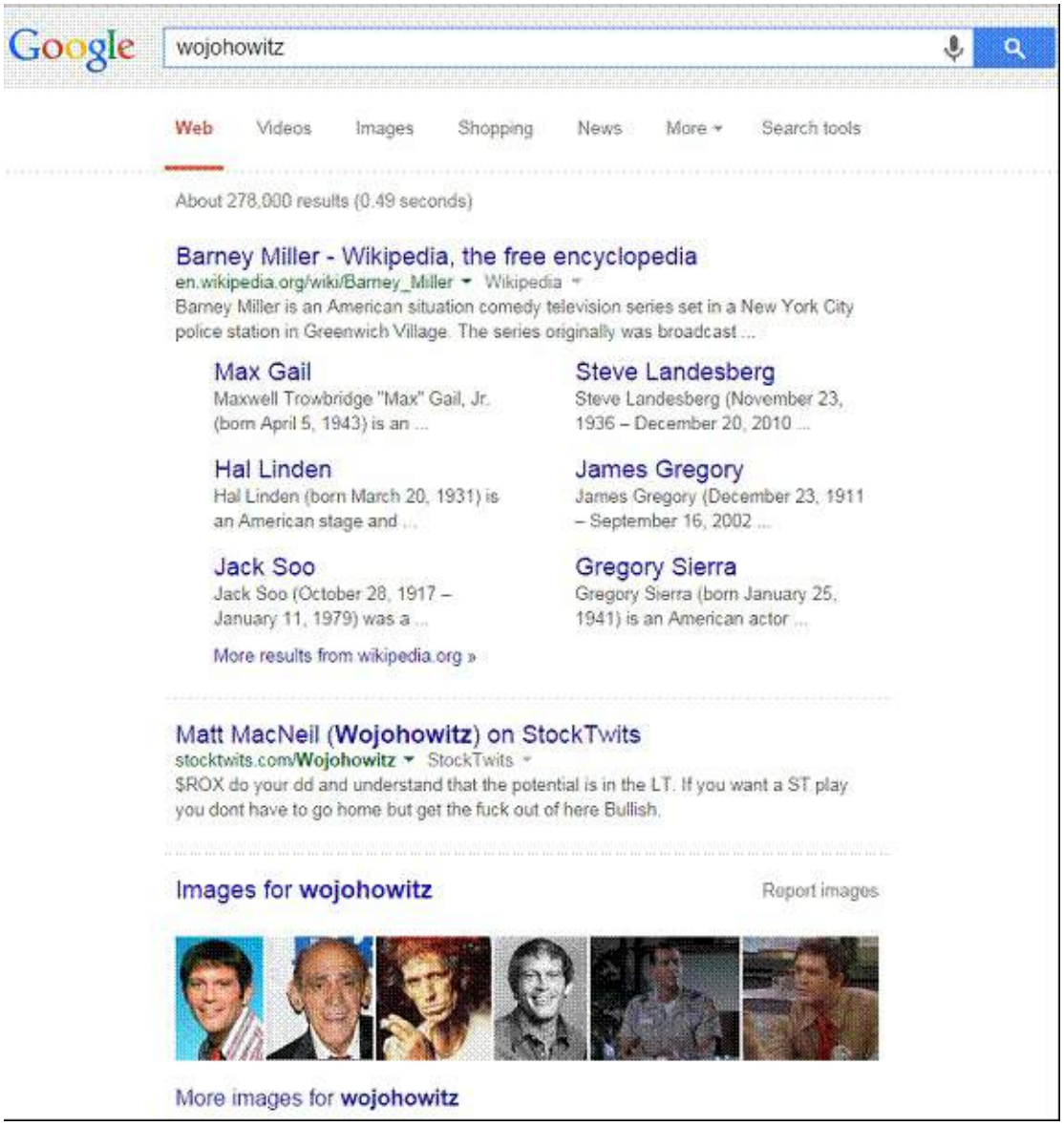


Figure 6-67. Example of blended search results

The emergence of blended search was a big step forward for search engines, but it was only one step on a longer journey. The commitment to that journey is well defined in Google's mission statement: "Google's mission is to organize the world's information and make it universally accessible and useful."

The next step in that journey for Google was the Knowledge Graph. This was a Google initiative designed to allow it to leverage structured databases to enhance the search results. This initiative allowed Google to further enhance the presentation of its results. An example is shown in Figure 6-68.

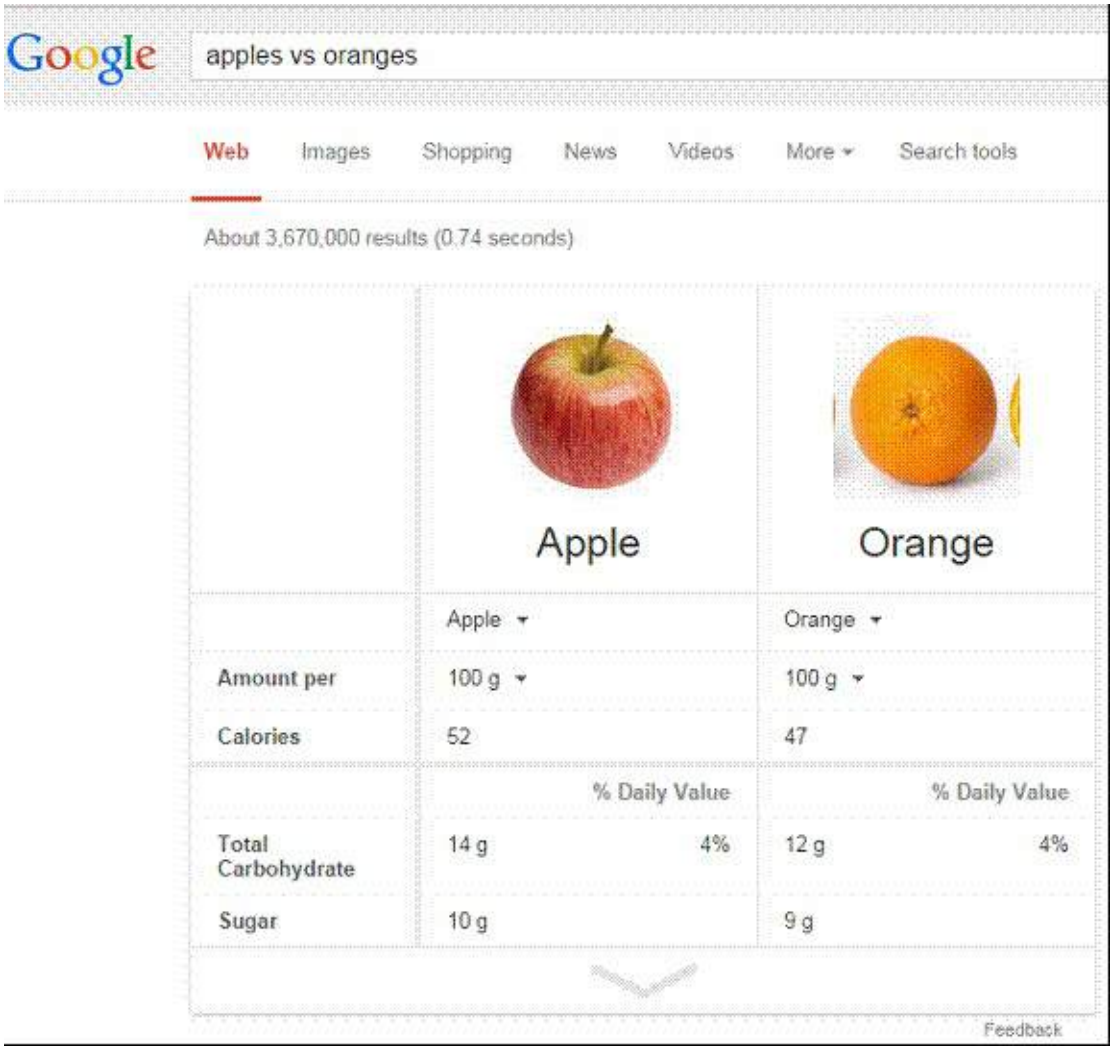


Figure 6-68. Sample Knowledge Graph result

In essence, the Knowledge Graph was another major step forward by Google that allowed it to start showing complete answers in the search results. The

information for these results is typically retrieved from Freebase, a community-edited database of information.

The basic concept is sometimes referred to as “moving from strings to things.” The search engines that returned nothing but 10 blue links were comparatively quite simple, as they relied on scanning the text on a web page to figure out what it was about, and did not understand relationships.

In comparison, the Knowledge Graph can understand that apples and oranges are both fruits, and they have properties, such as calories, carbohydrate levels, and grams of sugar. Or that the Empire State Building has height, a construction date, and initial architect, and that Google has access to pictures of it.

These types of data sources provide a rich array of information. It is estimated that they allow Google access to information on 500 million entities and 3.5 billion pieces of information. Stone Temple Consulting performed an extensive study on what types of queries generate Knowledge Graph results, and which don’t.[22]

While 3.5 billion pieces of information may seem like a very large number, in the grand scheme of things, it represents a very small portion of all the types of user queries. As a result, Google is pursuing other avenues to expand its ability to further enhance the information in the search results.

For example, it has started experimenting with extracting information from websites which it in turn has started to use for displaying step-by-step instructions, as shown in Figure 6-69.

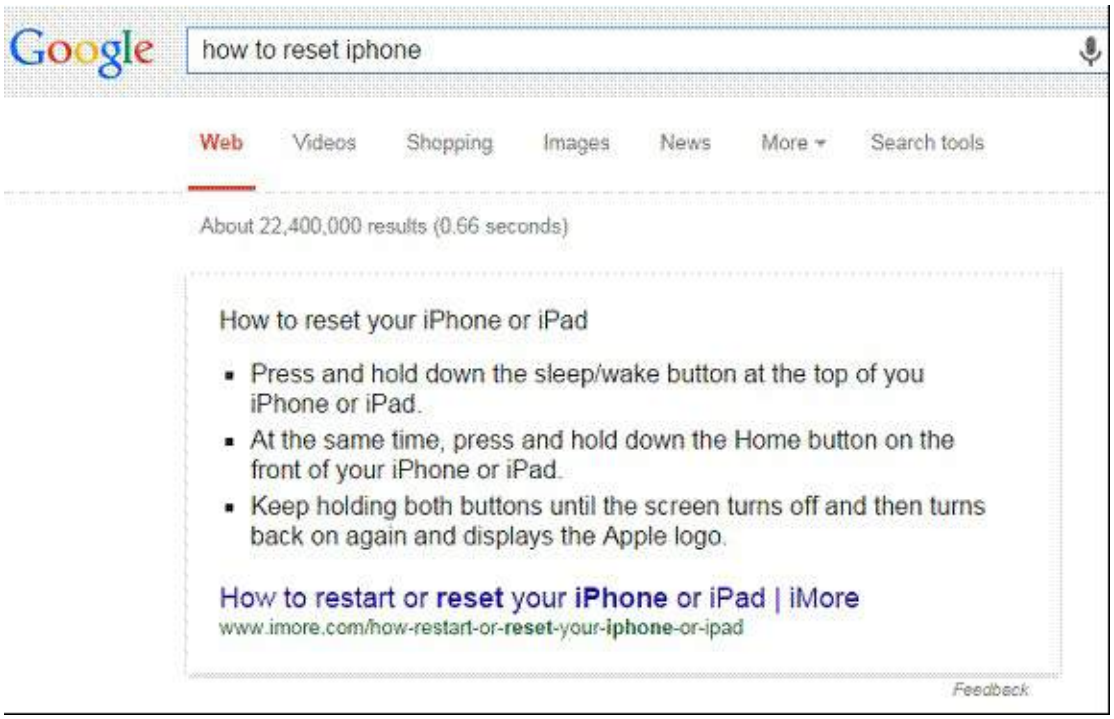


Figure 6-69. Example of step-by-step instructions in Google's SERPs

Barry Schwartz reached out to Google, and got the following response when he asked about the use of step-by-step instructions:[23]

> *We started experimenting with this in early June. We hope it draws attention to webpages that provide a useful series of steps to help people complete their task. In these cases we focus attention on the snippet because it's likely to be more helpful for deciding whether the webpage is going to be the most useful for the task.*

Other examples exist where Google is extracting knowledge from websites and showing it in the SERPs. Figure 6-70 shows an example of historical information being found on a website and displayed directly in the results.

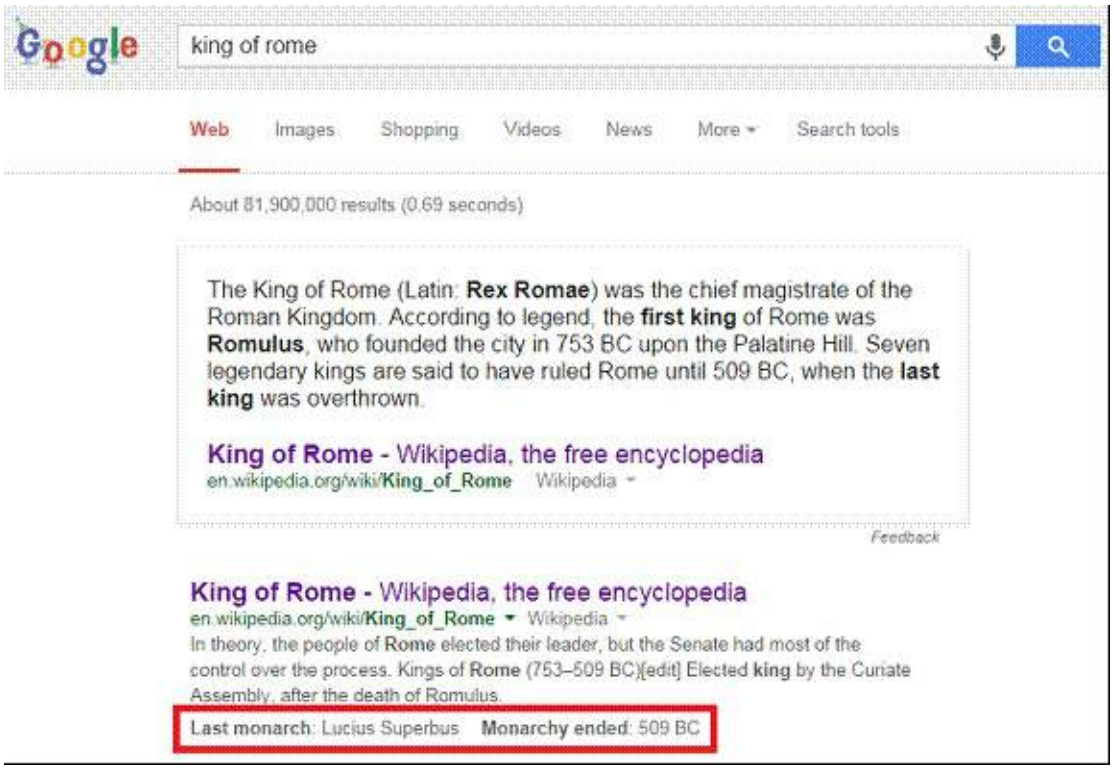


Figure 6-70. Knowledge extraction example

The examples shown in Figure 6-69 and Figure 6-70 are a clear step beyond the simple use of structured data. These represent early examples of what Google refers to as the Knowledge Vault.

Fair Use?

As Google presents more and more of these types of search results, many of the impacted publishers feel that Google is stealing their content and profiting from it. The question becomes whether or not Google's usage can be considered *fair use* as defined by the U.S. Copyright Office. There are four factors involved in determining fair use, as follows:

- The purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes
- The nature of the copyrighted work
- The amount and substantiality of the portion used in relation to the

copyrighted work as a whole

- The effect of the use upon the potential market for, or value of, the copyrighted work

There is actually no clear definition of fair use, but it is clear that the substance you take from the third party is a factor. It is common practice among those who quote others, or who attempt to make fair use of someone else's copyrighted material, to provide attribution. However, the U.S. Copyright Office indicates that this might not be enough: "Acknowledging the source of the copyrighted material does not substitute for obtaining permission."

In addition, this is more than a U.S.-only issue, and the laws differ from country to country.

Whether this becomes an issue for Google or not is yet to be determined, but the scale of what it's trying to do makes it likely that it will be subject to legal challenges, and that the way that various legal systems will respond will differ.

One additional aspect to consider is that public domain information is not copyrightable. For example, the fact that Olympia is the capital of the state of Washington is not copyrightable info. If Google is able to extract some common knowledge from third-party sites, it would not be subject to this discussion.

How the Knowledge Vault Works

The first public acknowledgment that Google had a concept it called the Knowledge Vault was in a presentation by Google's Kevin Murphy that took place on October 31, 2013.[24] As of October 2014, the Knowledge Vault is just a research project within Google, but it is still useful to learn what this project is about. The core concepts being studied are:

Machine reading

This is the process of extracting facts from a large text corpora. This is similar to methods developed by Carnegie-Mellon, the University of Washington, and others, but Google is working on a much larger-scale version. In addition, it is researching methods for using other prior knowledge to help reduce the error rate.

As information is assembled, it becomes possible to infer, or even determine, other facts. Figure 6-71 (slide 16 of the presentation) shows an example of this in action. For example, if we know that Barack Obama and Michelle Obama are both parents of Sasha Obama, then we can infer that it is likely that they are married (though that is not necessarily true).

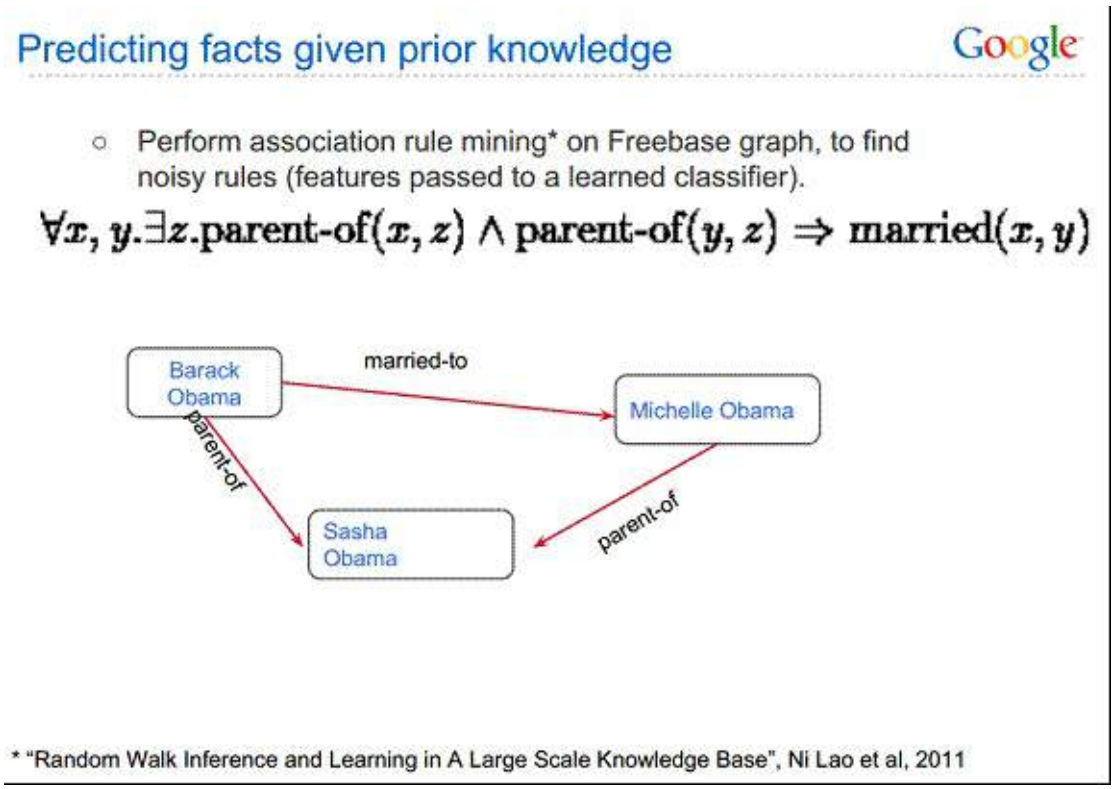


Figure 6-71. Inferring new information

Asking the Web

Web-based question and answers can be used to further supplement the available information. Learning how to ask the right questions (as shown in Figure 6-72) and how to frame those questions is by itself a very difficult process. Verification of the accuracy of the responses is important as well.

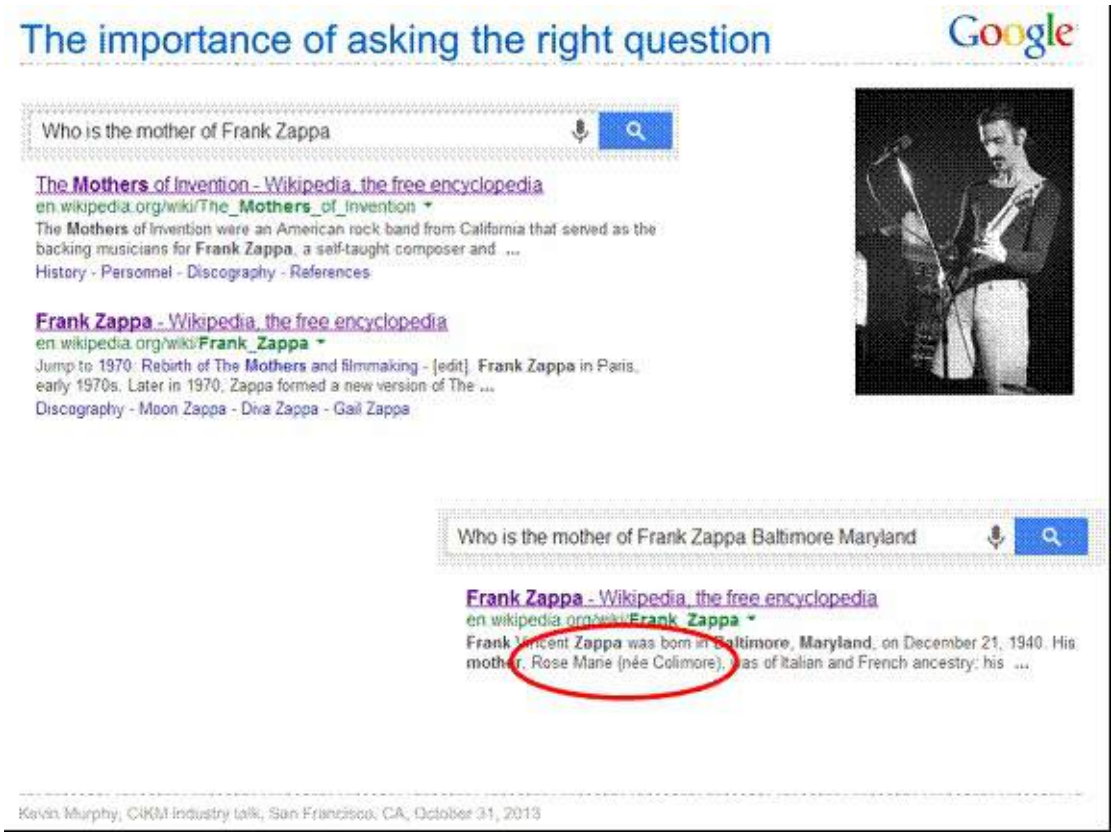


Figure 6-72. The importance of asking the right question

Asking people

Freebase is itself an example of this, as it is community edited. Other sources can be used as well. For example, each knowledge panel that Google shows contains a feedback link, allowing it to collect information on accuracy problems (Figure 6-73). This feedback can also be flawed, and Google is investigating algorithms to predict the possibility that the information received is correct.

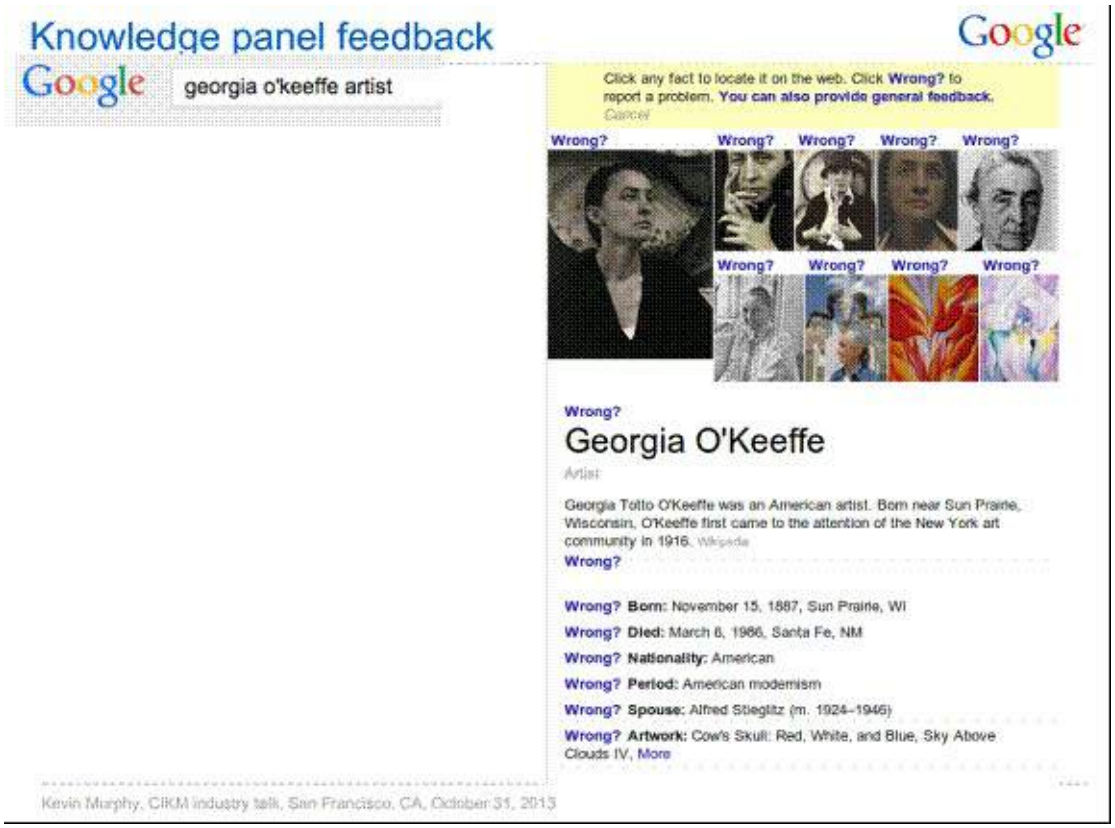


Figure 6-73. Knowledge panel feedback

Google has many patents that could potentially pertain to the Knowledge Graph and/or the Knowledge Vault. Each covers a particular aspect of how to extract information into a knowledge base. Here are a few examples for further reading:

- Knowledge Graph–Based Search System
- Extracting Patterns and Relations from the World Wide Web
- Determining Geographic Locations for Place Names in Fact Repository

There are many more that apply to this complex topic, and it will remain an area of investigation for some time to come.

The Future of the Knowledge Vault

As of October 2014, the Knowledge Vault is still a concept in its infancy. As noted earlier, it is just a research project. The algorithms are in primitive states of

definition, and the required processing power is quite substantial. Google will keep investing in these types of technologies as it tries to find more and more ways to provide better and better results. Its goal, and that of other search engines, remains to “organize the world’s information,” and it will keep investing in that goal until it succeeds.

However, this process may take a decade or more. This means we will see gradual changes continuing over time. Even if a single breakthrough provides Google with access to 1 billion facts, which sounds like a large number, it will still only impact a very small percentage of search results.

However, understanding the concepts of semantic search and the Knowledge Vault can in turn help you understand a bit more about where search engines are going.

Conclusion

By now you should be aware that a search engine–friendly website is the first step toward SEO success. In the next chapter, we will demonstrate how links are also a critical piece of the SEO puzzle—particularly when targeting highly competitive terms. However, if you have not made your site crawler-friendly and optimized, all of your other efforts—whether they’re link development, social media promotion, or other tactics to improve search visibility and increase search traffic—will be wasted. A website built from the ground up to optimal specifications for crawler accessibility and top organic exposure is the foundation from which you will build all SEO initiatives. From this solid foundation, even the loftiest of goals are within reach.

[1] See *http://www.stephanspencer.com/whats-wrong-with-google-sitemaps/*.

[2] From Wurman’s *Information Architects*.

[3] Webmaster Central Blog, “Updating Our Technical Webmaster Guidelines,” October 27, 2014, *http://bit.ly/webmaster_guidelines*.

[4] Adam Audette, “We Tested How Googlebot Crawls Javascript And Here’s What We Learned,” Search Engine Land, May 8, 2015, *http://searchengineland.com/tested-googlebot-crawls-javascript-heres-learned-220157*.

[5] Barry Schwartz, “Google May Discontinue Its AJAX Crawlable Guidelines,” Search Engine Land, March 5, 2015, *http://bit.ly/ajax_crawlable_guidelines*.

[6] Eric Enge, “Eliminate Duplicate Content in Faceted Navigation with Ajax/JSON/JQuery,” Moz Blogs, June 11, 2015, *https://moz.comblogusing-ajax-json-jquery-to-implement-faceted-navigation*.

[7] Christoph C. Cemper, “Deconstructing the Google EMD Update,” Search Engine Land, October 25, 2012, *http://searchengineland.com/google-emd-update-research-and-thoughts-137340*.

[8] Dr. Peter J. Meyers, “New Title Tag Guidelines & Preview Tool,” Moz Blog, March 20, 2014, *https://moz.comblognew-title-tag-guidelines-preview-tool*.

[9] Dan Sharp, “An Update on Pixel Width in Google SERP Snippets,” May 15, 2014, *http://www.screamingfrog.co.uk/an-update-on-pixel-width-in-google-serp-snippets/*.

[10] You can see a copy of the announcement at *http://bit.ly/intro_schema_org*.

[11] Webmaster Central Blog, “Google does not use the keywords meta tag in web ranking,” September 21, 2009, *http://bit.ly/keywords_meta_google*.

[12] Caitlin O’Connell, “Google’s John Mueller Interviewed by Eric Enge,” Stone Temple Consulting, May 11, 2009, *http://bit.ly/mueller_interview*.

[13] Available on the Google Webmasters YouTube channel: “Cloaking,” *http://www.youtube.com/watch?feature=player_embedded&v=QHtnfOgp65Q*.

[14] Matt Cutts, “PageRank Sculpting,” Matt Cutts: Gadgets, Google, and SEO, June 15, 2009, *https://www.mattcutts.comblogpagerank-sculpting/*.

[15] Tim Berners-Lee, James Hendler, and Ora Lassila, “The Semantic Web,” Scientific American, May 2001, *http://www.scientificamerican.com/article/the-semantic-web/*.

[16] Othar Hansson, “Authorship Markup and Web Search,” Webmaster Central Blog, June 7, 2011, *http://googlewebmastercentral.blogspot.com/2011/06/authorship-markup-and-web-search.html*.

[17] AJ Kohn, “Authorship Is Dead, Long Live Authorship,” Blind Five Year Old, October 24, 2013, *http://www.blindfiveyearold.com/authorship-is-dead-long-live-authorship*.

[18] Barry Schwartz, “Confirmed: Google Reduces Authorship Rich Snippets in Search Results,” Search Engine Land, December 19, 2013, *http://searchengineland.com/confirmed-google-reduces-authorship-richsnippets-in-search-results-180313*.

[19] Eric Enge, “Authorship Adoption Fail – Detailed Stats,” Stone Temple Consulting, September 9, 2014, *https://www.stonetemple.com/authorship-adoption-fail-detailed-stats/*.

[20] Bill Slawski, “Has Google Decided That You Are Authoritative for a Query?”, SEO by the Sea, September 7, 2014, *http://www.seobythesea.com/2014/09/google-decided-authoritative-query/*.

[21] Mark Traphagen, “Does Google Use Facebook & Twitter as Ranking Signals? Matt Cutts Answers,” Stone Temple Consulting, January 23, 2014, *https://www.stonetemple.com/googles-matt-cutts-understanding-social-identity-on-the-web-is-hard/*.

[22] Eric Enge, “The Great Knowledge Box Showdown: Google Now vs. Siri vs. Cortana,” Stone Temple Consulting, October 7, 2014. *https://www.stonetemple.com/great-knowledge-box-showdown/*.

[23] Barry Schwartz, “Google’s Knowledge Graph Is Showing Step By Step Instructions: Here Are Some Examples,” Search Engine Land, June 24, 2014, *http://searchengineland.com/googles-knowledge-graph-showing-step-step-instructions-examples-194923*.

[24] Kevin Murphy, “From Big Data to Big Knowledge,” October 31, 2013, *http://cikm2013.org/slides/kevin.pdf*.

CHAPTER SEVEN

Content Marketing

In today's search environment, the main driving factors are now what we generally refer to as *social proof signals*, such as inbound links (e.g., within a blog post) and user engagement with your content (e.g., time spent watching your video). As you will see in Chapter 8, social signals such as retweets, likes, and pins don't appear to have a direct ranking impact, and Google+ appears to have an impact, but only from a personalized search perspective.

For many years, links to a website were the single largest factor in determining its search engine rankings, because links generally (before they became a tool for SEO manipulation) existed to provide a pathway for a site's users to find additional, relevant content on a third party's website—a "signal" that the owner of the linking site deemed the third party's linked content valuable.

Because of the power of this signal, many SEO professionals pursued obtaining links to their sites or their client's sites without worrying about the quality of the site where those links resided. Unfortunately, many link-building efforts and services spawned by this behavior had little integration with the rest of the publisher's content development and marketing strategies.

Clearly, this violated the spirit of what the search engines were measuring and placing value on—links that act as valid endorsements for third-party content. As a result, the search engines, and Google in particular, have taken many steps to force website owners and publishers to view link building more holistically, as an "earned" engagement rather than a "purchased" endorsement, requiring a renewed focus on links as a measurement of content quality. This shift, both necessary and welcomed, reestablishes the need for quality content development (as the "earner" of links) to be integrated with the overall PR and marketing strategy for businesses.

The development of highly shareable content, and the promotion of that content via various channels for increased business visibility, is generally referred to as *content marketing*. Content can be published on your own site, other people's sites, or in social media, but in all cases acts to build visibility for your brand online. The most valuable content is usually highly relevant to what you do, solves problems for others or stirs their emotions, and is often noncommercial in nature.

Links remain a large factor in search engine ranking algorithms, but we use content marketing to build our reputation and visibility online, and as a result we obtain organic links of the highest possible quality—links that would be desirable for your business even if the search engines did not exist, and that people might actually click on to engage with your business.

The most important thing to remember as you delve into this chapter is that the primary goal of any content marketing effort should be enhancing the reputation of your business. Any campaign that starts with "getting links" as the objective, without placing primary and ongoing focus on the quality and value of the content being linked to, will eventually run into problems (if it hasn't already; see Chapter 9). During a 2012 interview, Google's Matt Cutts and Eric Enge had the following exchange:[1]

> *Eric Enge: It dawned on me recently that link building is an interesting phrase that has misled people. It is a bit of a "cart before the horse" thing. It has led people to think about links as something they get from the "dark corners of the Web." Places where no one ever goes, so it does not matter what you do there. So by thinking of it this way, as link building, you are off on the wrong foot even before you get started.*
>
> *Matt Cutts: That's right. It segments you into a mindset, and people get focused on the wrong things. It leads them to think about links as the end goal. It is important to think about producing something excellent first. If you have an outstanding product, world-class content, or something else that sets you apart, then you can step back and start thinking about how to promote it.*

There are many who believe that social signals and user engagement with your content have become important ranking factors. However, the impact of social media appears to be quite limited:

- Google+ can have a strong impact on personalized search within Google for those who are active on the Google+ platform.
- Search engines may use shared content on social media platforms as a way of discovering new content—in particular, news-related content.

These new ranking factors will be discussed in greater detail in Chapter 8.

How Links Historically Influenced Search Engine Rankings

The concept of using links as a way to measure a site's importance was first made popular by Google with the implementation of its PageRank algorithm (others had previously written about using links as a ranking factor, but Google's rapidly increasing user base popularized it). In simple terms, each link to a web page is a vote for that page. But it's not as simple as "the page with the most votes wins." Links and linking pages are not all created equal. Some links are weighted more heavily by Google's PageRank algorithm than others.

The key to this concept is the notion that links represent an "editorial endorsement" of a web document. Search engines rely heavily on editorial votes. However, as publishers learned about the power of links, some started to manipulate links through a variety of methods. This created situations in which the intent of the link was not editorial in nature, and led to many algorithm enhancements.

To help you understand the origins of link algorithms, the underlying logic of which is still in force today, let's take a look at the original PageRank algorithm in detail.

The Original PageRank Algorithm

The PageRank algorithm was built on the basis of the original PageRank thesis authored by Sergey Brin and Larry Page while they were undergraduates at Stanford University.[2]

In the simplest terms, the paper states that each link to a web page is a vote for that page. However, as stated earlier, votes do not have equal weight. So that you can better understand how this works, we'll explain the PageRank algorithm at a high level. First, all pages are given an innate but tiny amount of PageRank, as shown in Figure 7-1.

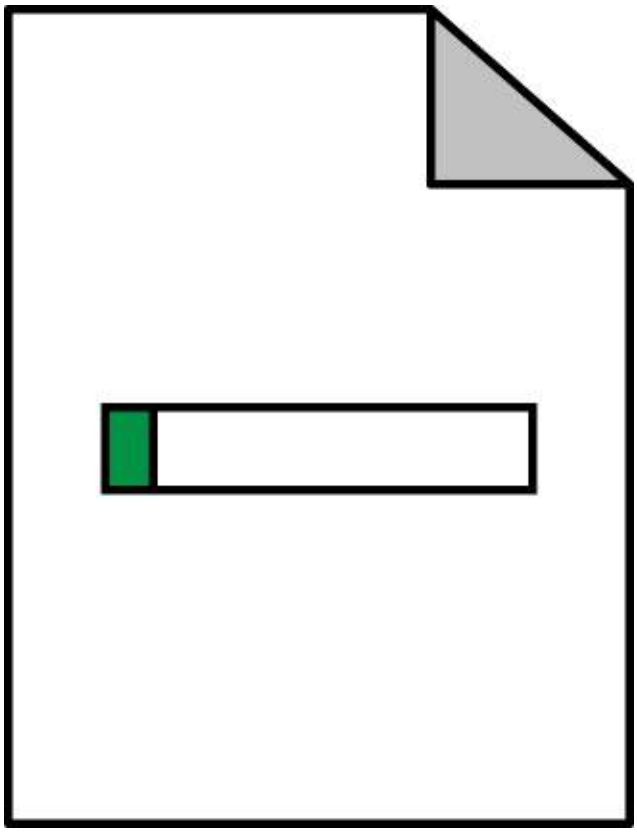


Figure 7-1. Some PageRank for every page

Pages can then increase their PageRank by receiving links from other pages, as shown in Figure 7-2.

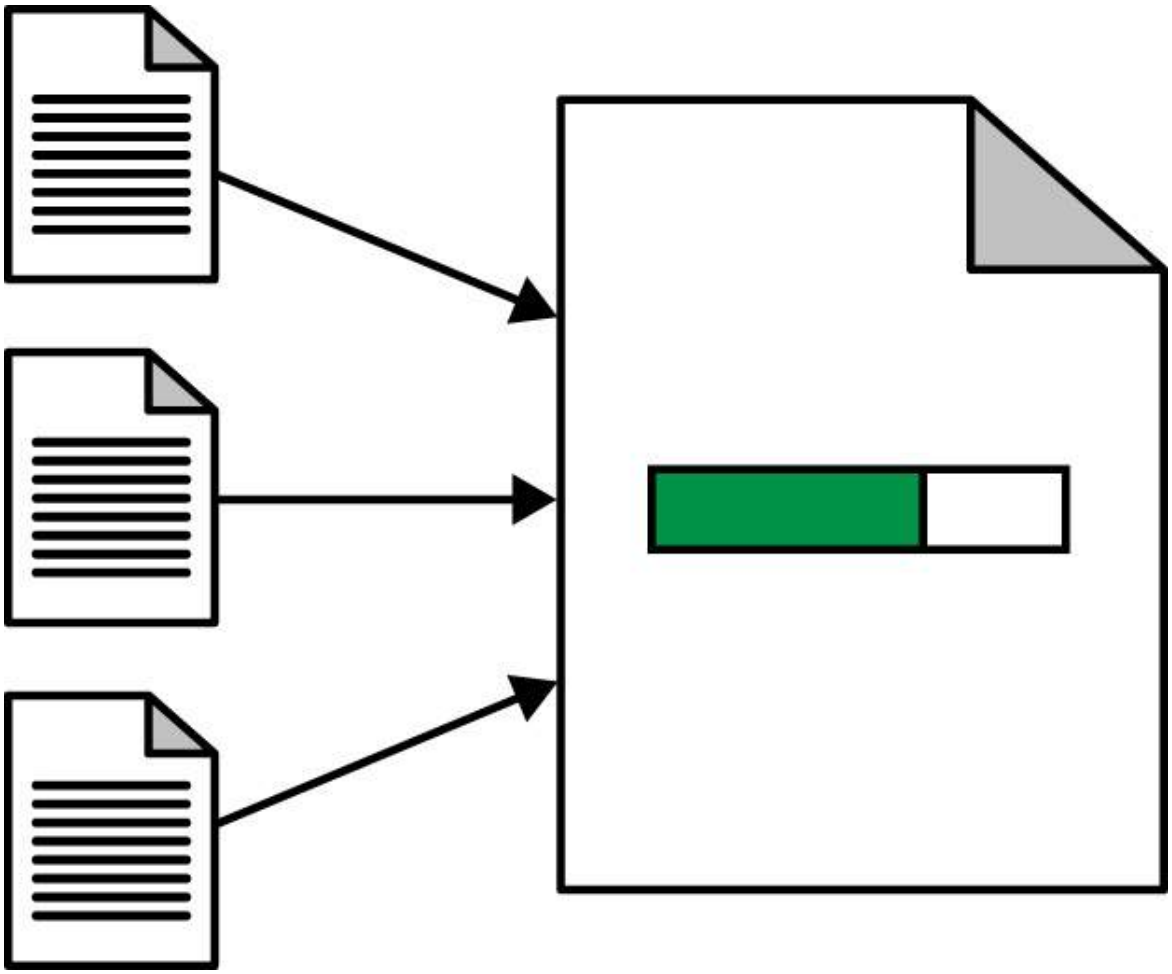


Figure 7-2. Pages receiving more PageRank through links

How much PageRank can a page pass on to other pages through links? That ends up being less than the page's PageRank. In Figure 7-3 this is represented by *f(x)*, meaning that the passable PageRank is a function of *x*, the total PageRank. In 2009, Matt Cutts wrote a post in which he suggested that a page might be able to vote 85–90% of its PageRank.[3]

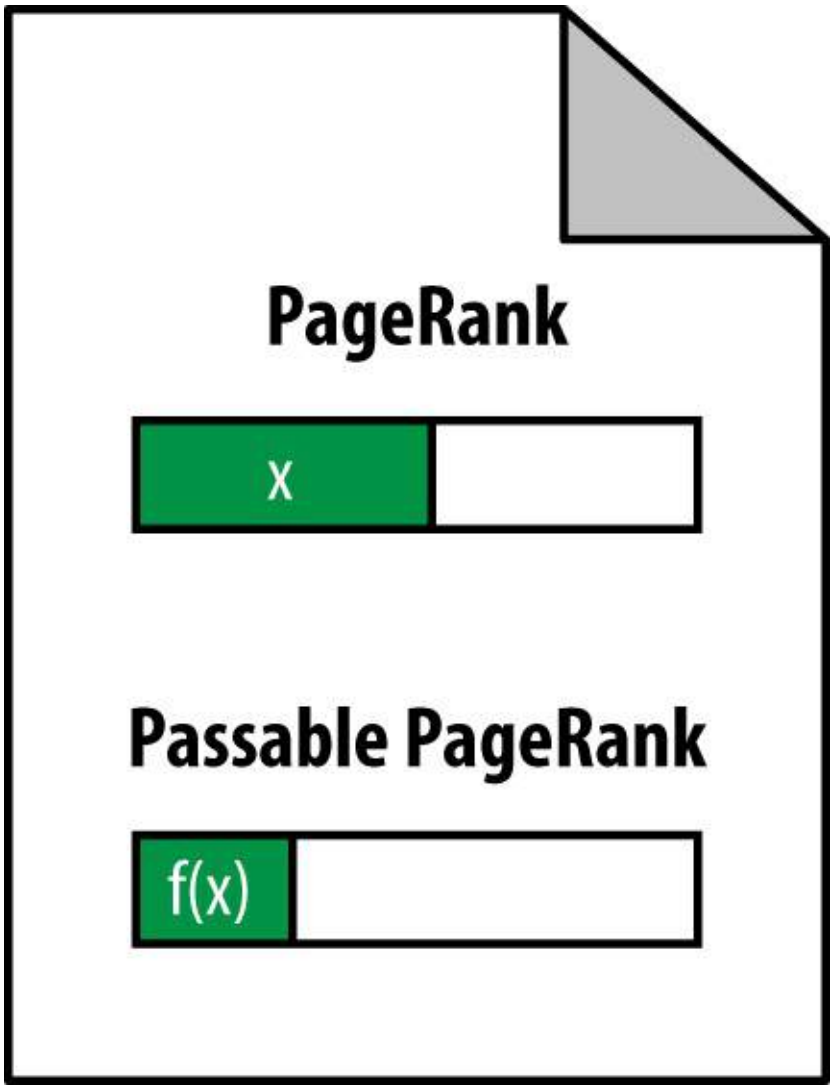


Figure 7-3. Some of a page's PageRank is passable to other pages

If this page links to only one other page, it passes all of its passable PageRank to that page, as shown in Figure 7-4, where Page B receives all of the passable PageRank of Page A.

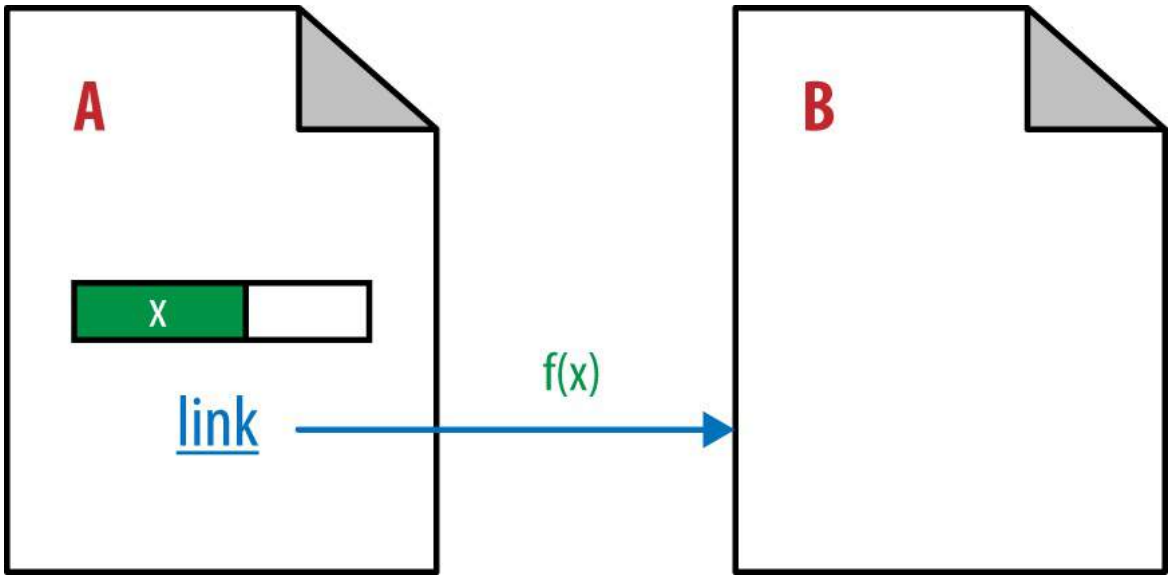


Figure 7-4. Passing of PageRank through a link

However, the scenario gets more complicated because pages will link to more than one other page. When that happens the passable PageRank gets divided among all the pages receiving links. We show that in Figure 7-5, where Page B and Page C each receive half of the passable PageRank of Page A.

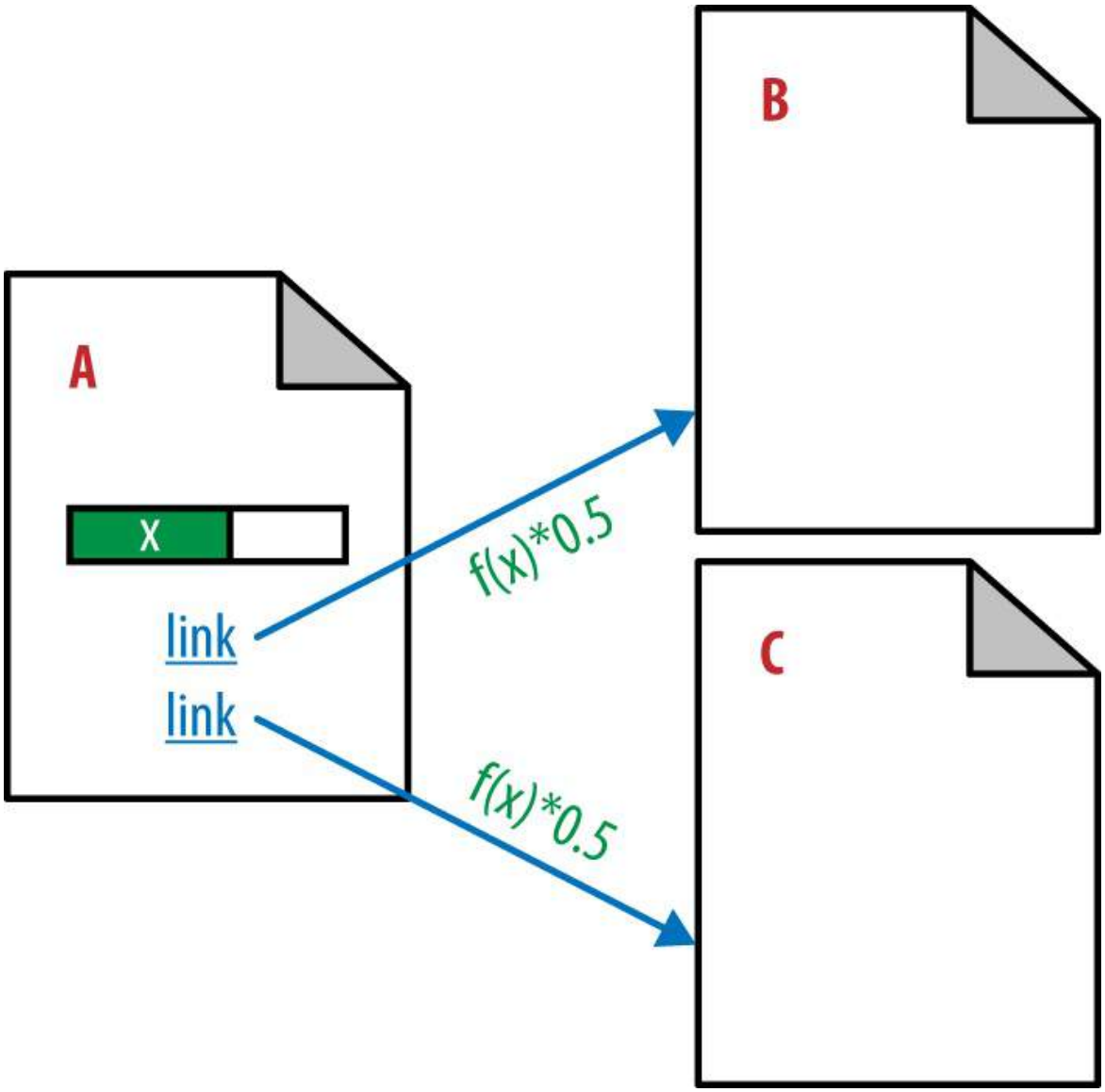


Figure 7-5. How PageRank is passed

In the original PageRank formula, link weight is divided equally among the number of links on a page. This undoubtedly does not hold true today, but it is still valuable in understanding the original intent. Now take a look at Figure 7-6, which depicts a more complex example that shows PageRank flowing back and forth between pages that link to one another.

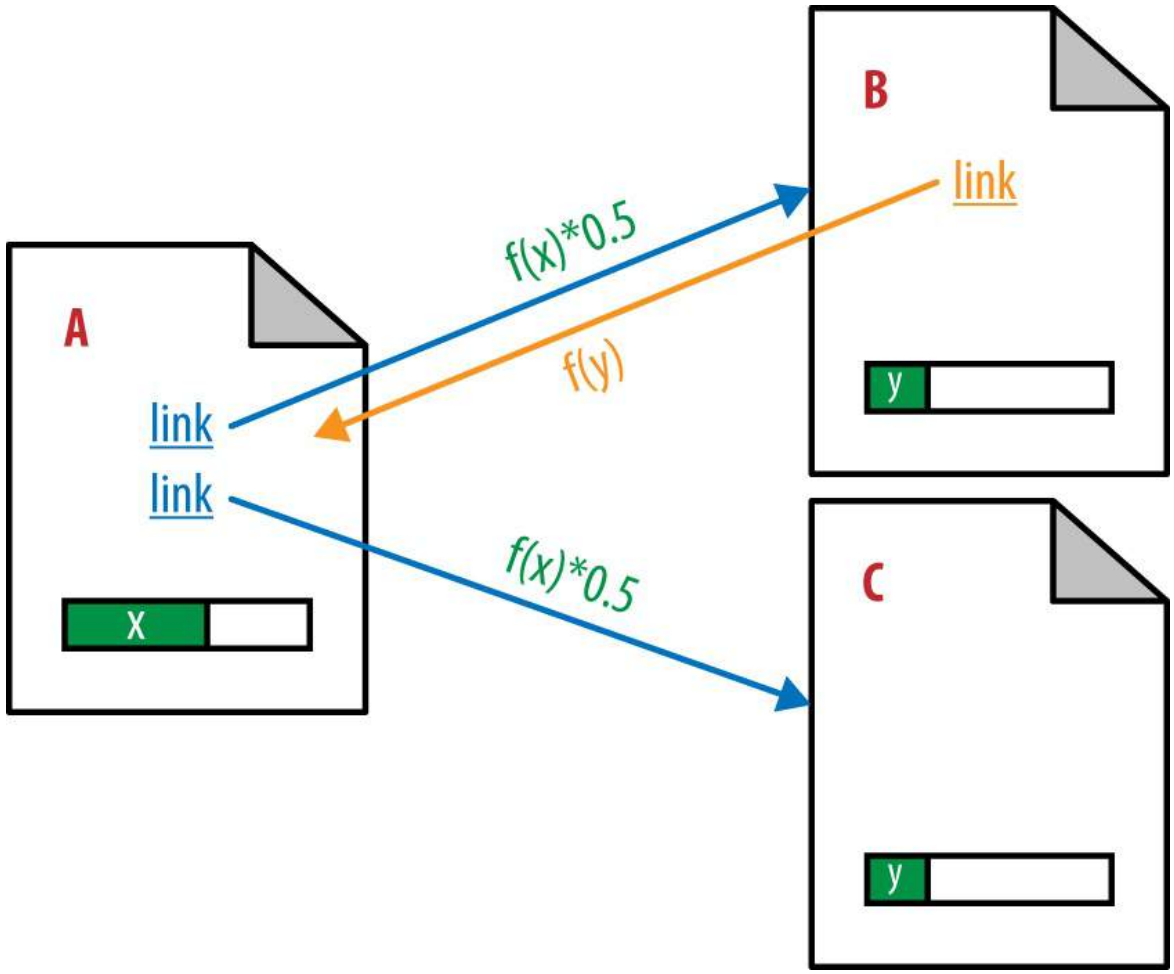


Figure 7-6. Cross-linking between pages

Cross-linking makes the PageRank calculation much more complex. In Figure 7-6, Page B now links back to Page A and passes some PageRank, *f(y)*, back to Page A. Figure 7-7 should give you a better understanding of how this affects the PageRank of all the pages.

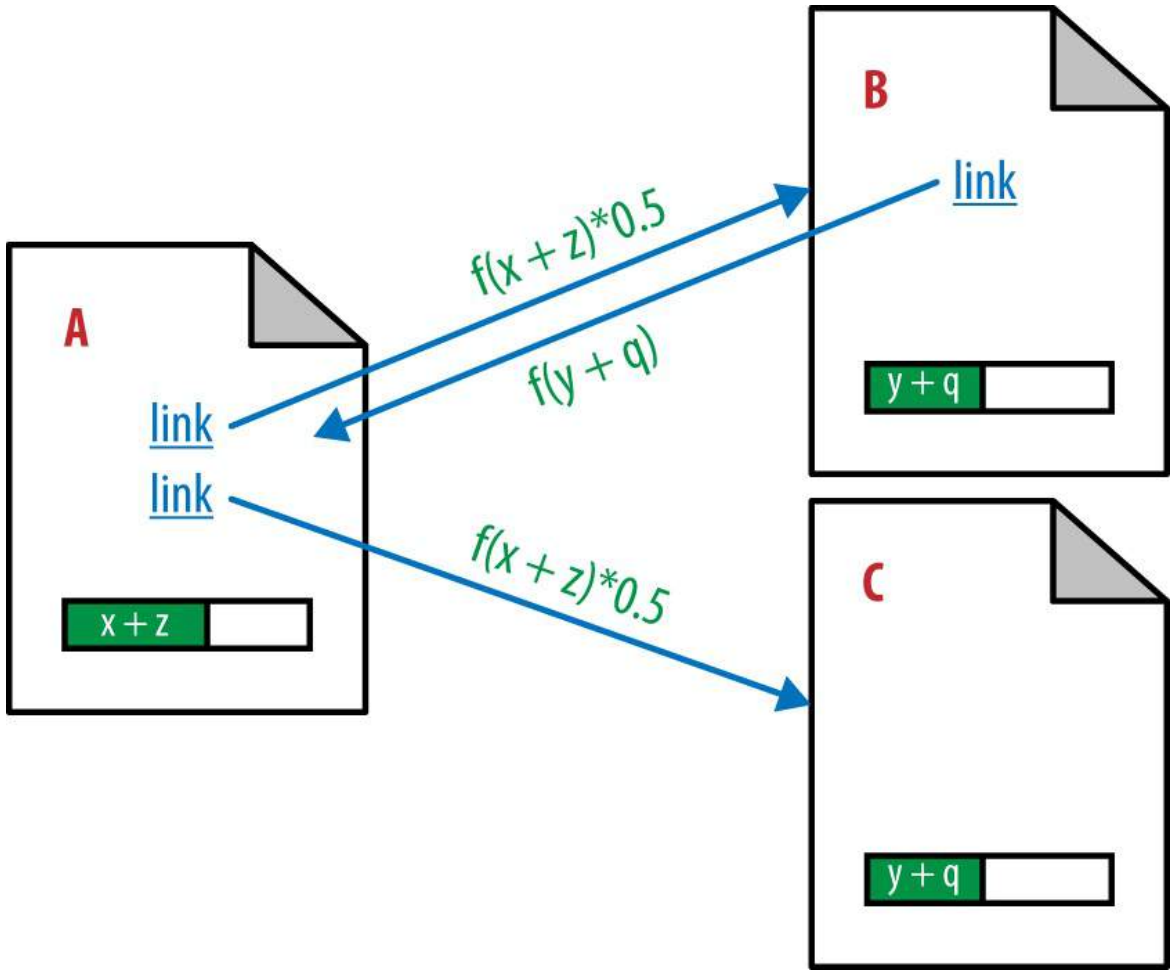


Figure 7-7. Iterative PageRank calculations

The key takeaway here is that when Page B links to Page A to make the link reciprocal, the PageRank of Page A (*x*) becomes dependent on *f(y)*, the passable PageRank of Page B, which happens to be dependent on *f(x)*! In addition, the PageRank that Page A passes to Page C is also impacted by the link from Page B to Page A. This makes for a very complicated situation where the calculation of the PageRank of each page on the Web must be determined by recursive analysis.

We have defined new parameters to represent this: *q*, which is the PageRank that accrues to Page B from the link that it has from Page A (after all the iterative calculations are complete); and *z*, which is the PageRank that accrues to Page A from the link that it has from Page B (again, after all iterations are complete).

The scenario in Figure 7-8 adds further complexity by introducing a link from

Page B to Page D. In this example, Pages A, B, and C are internal links on one domain, and Page D represents a different site (shown as Wikipedia). In the original PageRank formula, internal and external links passed PageRank in exactly the same way. This became exposed as a flaw because publishers started to realize that links to other sites were “leaking” PageRank away from their own site, as you can see in Figure 7-8.

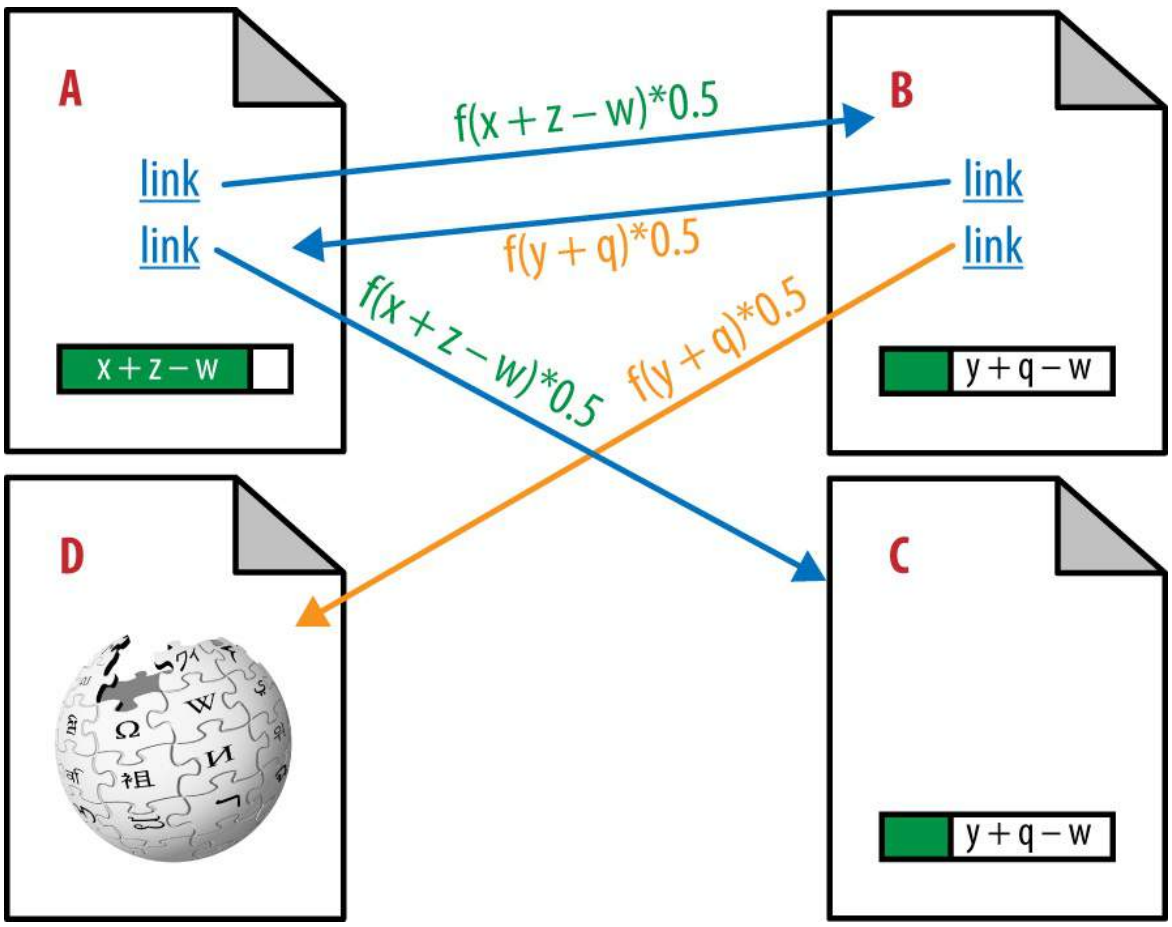


Figure 7-8. PageRank being leaked

So, for example, because Page B links to Wikipedia, some of the passable PageRank is sent there, instead of to the other pages that Page B is linking to (Page A in our example). In Figure 7-8, we represent that with the parameter *w*, which is the PageRank not sent to Page A because of the link to Page D.

The PageRank “leak” concept represented a fundamental flaw in the algorithm. Once page creators investigated PageRank’s underlying principles, they realized that linking out from their own sites would cause more harm than good. If a

great number of websites adopted this philosophy, it could negatively impact the "links as votes" concept and actually damage the quality of Google's algorithm. Needless to say, Google quickly corrected this flaw to its algorithm. As a result of these changes, you no longer need to worry about PageRank leaks. Quality sites should link to other relevant quality pages around the Web.

Even after these changes, internal links from pages still pass some PageRank, so they still have value, as shown in Figure 7-9.

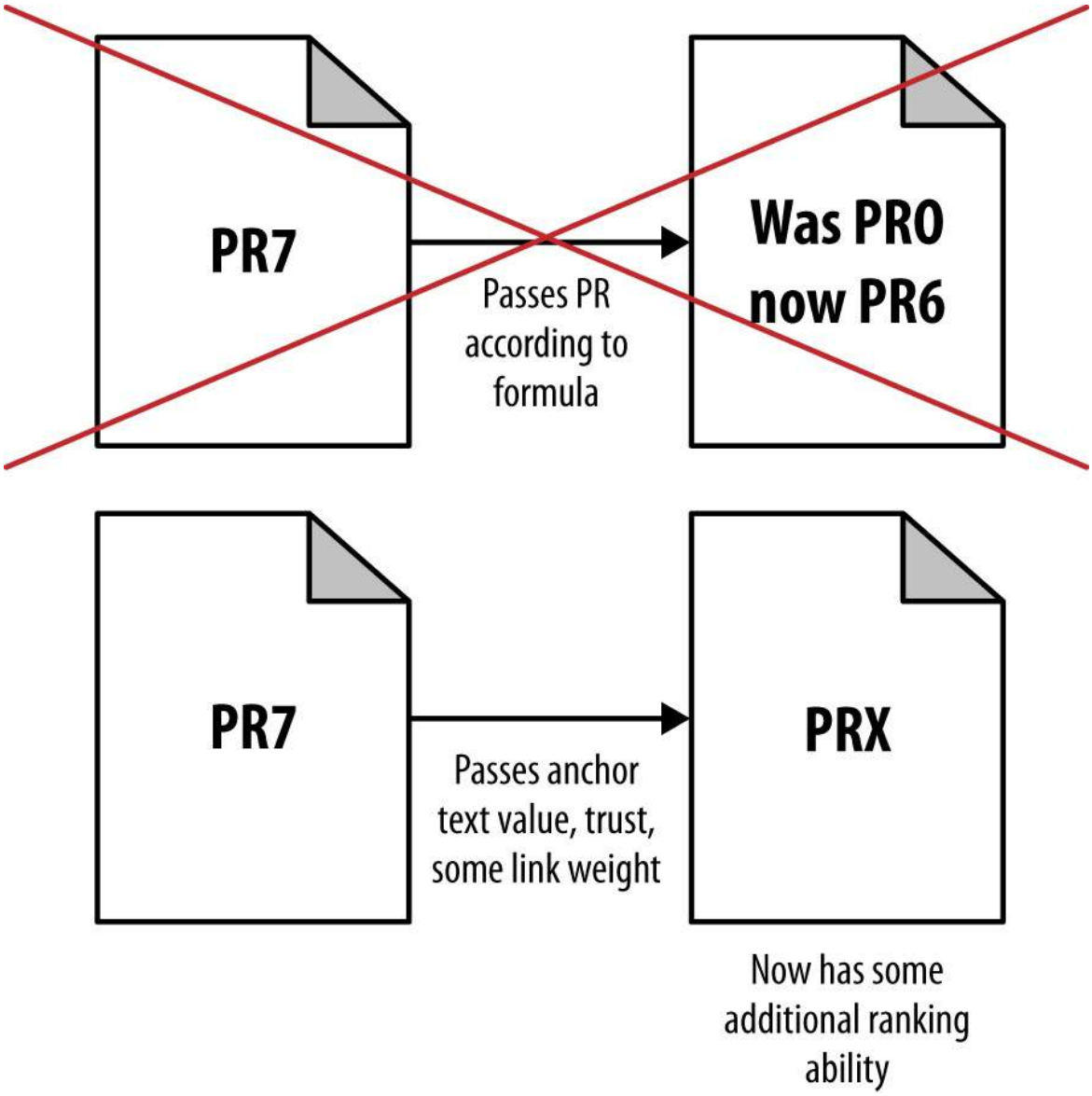


Figure 7-9. Internal links still passing some PageRank

Google has continuously changed and refined the way it uses links to impact rankings, and the current algorithm is not based on PageRank as it was originally defined. However, familiarity and comfort with the original algorithm are certainly beneficial to those who practice optimization of Google results.

All link-based algorithms are built on the assumption that for the most part the links received are legitimate endorsements by the publisher implementing a link to your website. The person implementing the link should be doing it because he feels he is linking to a great resource that would be relevant to visitors on his website.

In an ideal world, links would be similar to the academic citations you find at the end of a scientist's published paper, where she cites the other works she has referenced in putting together her research.

If the publisher implementing the link is compensated for doing so, the value of the link to a search engine is diminished, and such links can be harmful to search engine algorithms. Note that compensation can come in the form of money or special considerations, and we will explore this more in this chapter.

Additional Factors That Influence Link Value

Classic PageRank isn't the only factor that influences the value of a link. In the following subsections, we discuss some additional factors that influence the value a link passes.

Anchor text

Anchor text refers to the clickable part of a link from one web page to another. As an example, Figure 7-10 shows a snapshot of a part of the Quicken Loans home page.

Figure 7-10. Anchor text: a strong ranking element

The anchor text for the first link in the list of Popular pages in Figure 7-10 is Refinancing. The search engine uses this anchor text to help it understand what the page receiving the link is about. As a result, the search engine will interpret the link as saying that the page receiving the link is about refinancing, and therefore rank the page higher in the search results for that search query.

At one time, anchor text was so powerful that SEOs engaged in a practice called *Google bombing*—the idea that if you link to a given web page from many places with the same anchor text, you can get that page to rank for queries related to that anchor text, even if the page is unrelated and didn't even include any of the words in the query.

One notorious Google bomb was a campaign that targeted the WhiteHouse.gov biography page for George W. Bush with the anchor text *miserable failure*. As a result, that page ranked #1 for searches on *miserable failure* until Google tweaked its algorithm to reduce the effectiveness of this practice.

Google bombing was not the worst consequence of the power of anchor text. The use of anchor text as a ranking factor is useful in the search algorithms only if the person implementing the link naturally chooses what to use; if he is compensated for using specific anchor text, the value of the link as a ranking signal is negatively impacted.

To make matters worse, SEOs started to abuse the system and started implementing link-building programs designed around anchor text to drive their rankings. As a result, publishers that did not pursue these types of link-building campaigns were at a severe disadvantage. Ultimately, this started to break down the notion of links as valid academic citations, and Google began to take action.

In early 2012, Google began to send publishers "unnatural link" warnings through Google Search Console, and on April 24, 2012, Google released the first version of its Penguin algorithm. These topics are discussed more in Chapter 9.

Anchor text remains an important part of search algorithms, but now the search engines look for unnatural patterns of anchor text (too much of a good thing) and are lowering the rankings for publishers that exhibit patterns of artificially influencing the anchor text people use in links to their website.

Relevance

Links that originate from sites/pages on the same topic as the publisher's site, or on a closely related topic, are worth more than links that come from a site on an unrelated topic.

Think of the relevance of each link being evaluated in the specific context of the search query a user has just entered. So, if the user enters *used cars in Phoenix* and the publisher has received a link to its Phoenix used cars page that is from the Phoenix Chamber of Commerce, that link will reinforce the search engine's belief that the page really does relate to Phoenix.

Similarly, if a publisher has another link from a magazine site that has done a review of used car websites, this will reinforce the notion that the site should be considered a used car site. Taken in combination, these two links could be powerful in helping the publisher rank for *used cars in Phoenix*.

Authority

Authority has been the subject of much research. One of the more famous papers, written by Apostolos Gerasoulis and others at Rutgers University and titled "DiscoWeb: Applying Link Analysis to Web Search," became the basis of the Teoma algorithm, which was later acquired by AskJeeves and became part of the Ask algorithm.[4]

What made this algorithm unique was its focus on evaluating links on the basis of their relevance to the linked page. Google's original PageRank algorithm did not incorporate the notion of topical relevance, and although Google's algorithm clearly does this today, Teoma was in fact the first search engine to offer a commercial implementation of link relevance.

Teoma introduced the notion of *hubs*, which are sites that link to most of the important sites relevant to a particular topic, and *authorities*, which are sites that are linked to by most of the sites relevant to a particular topic.

The key concept here is that each topic area that a user can search on will have authority sites specific to that topic area. The authority sites for used cars are different from the authority sites for baseball.

Refer to Figure 7-11 to get a sense of the difference between hub and authority sites.

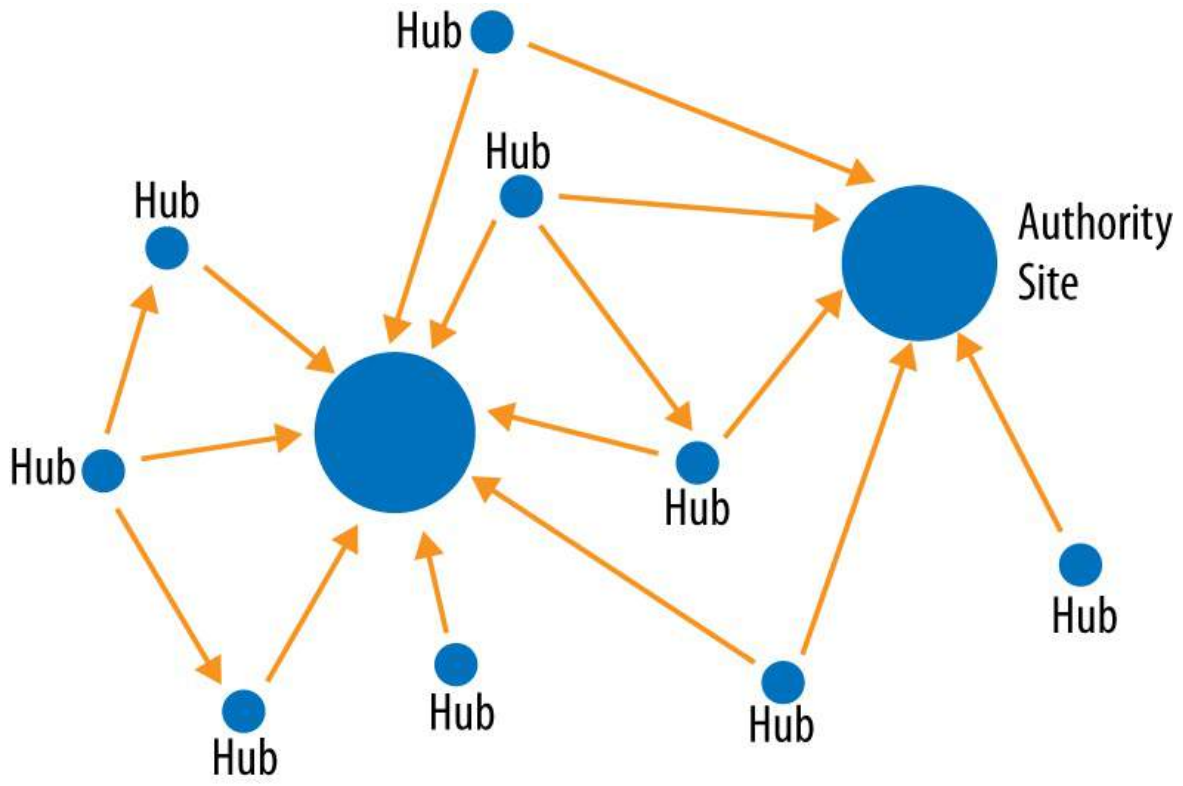


Figure 7-11. Hubs and authorities

So, if the publisher has a site about used cars, it should seek links from websites that the search engines consider to be authorities on used cars (or perhaps more broadly, on cars). However, the search engines will not tell you which sites they consider authoritative—making the publisher's job that much more difficult.

The model of organizing the Web into topical communities and pinpointing the hubs and authorities is an important one to understand (read more about it in Mike Grehan's paper "Filthy Linking Rich!"[5] The best link builders understand this model and leverage it to their benefit.

Trust

Trust is distinct from authority. On its own, authority doesn't sufficiently take into account whether the linking page or the domain is easy or difficult for spammers to infiltrate, or the motivations of the person implementing the link. Trust, on the other hand, does.

Evaluating the trust of a website likely involves reviewing its link neighborhood to see what other trusted sites link to it. More links from other trusted sites

would convey more trust.

In 2004, Yahoo! and Stanford University published a paper titled “Combating Web Spam with TrustRank.”[6] The paper proposed starting with a trusted seed set of pages (selected by manual human review) to perform PageRank analysis, instead of a random set of pages as was called for in the original PageRank thesis.

Using this tactic removes the inherent risk in using a purely algorithmic approach to determining the trust of a site, and potentially coming up with false positives or negatives.

The trust level of a site would be based on how many clicks away it is from seed sites. A site that is one click away accrues a lot of trust; two clicks away, a bit less; three clicks away, even less; and so forth. Figure 7-12 illustrates the concept of TrustRank.

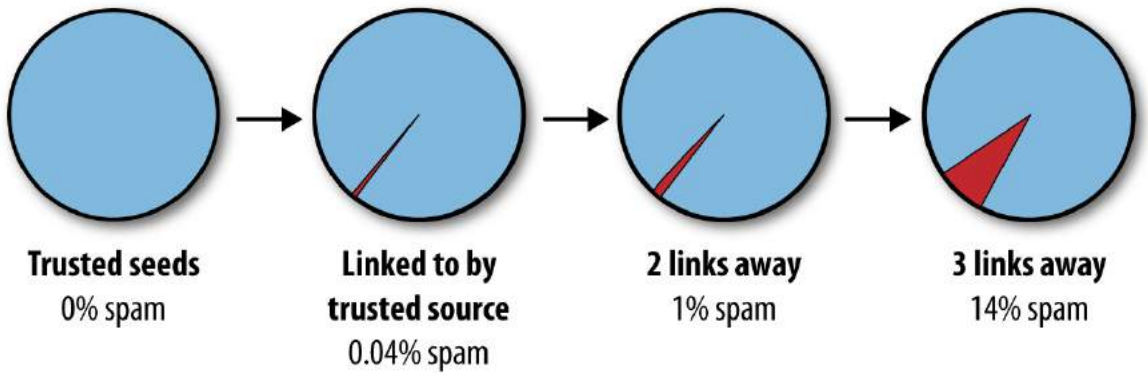


Figure 7-12. TrustRank illustrated

The researchers of the TrustRank paper also authored a paper describing the concept of *spam mass*.[7] This paper focuses on evaluating the effect of spammy links on a site’s (unadjusted) rankings. The greater the impact of those links, the more likely the site itself is spam. A large percentage of a site’s links being purchased is seen as a spam indicator as well. You can also consider the notion of *reverse TrustRank*, where linking to spammy sites will lower a site’s TrustRank.

It is likely that Google and Bing both use some form of trust measurement to evaluate websites. It is probably done by different means than outlined in the TrustRank and spam mass papers, and it may be incorporated into the methods

they use for calculating authority, but nonetheless, trust is believed to be a significant factor in rankings. For SEO practitioners, getting measurements of trust can be difficult. Currently, mozTrust (from Moz's Open Site Explorer) and TrustFlow (from Majestic SEO) are the most well-known publicly available metrics tools for evaluating a page's trust level.

How Search Engines Use Links

The search engines use links primarily to discover web pages, and to count the links as votes for those web pages. But how do they use this information once they acquire it? Let's take a look:

Index inclusion

Search engines need to decide which pages to include in their index. Crawling the Web (following links) is one way they discover web pages (another is through the use of XML sitemap files). In addition, the search engines do not include pages that they deem to be of very low value, because cluttering their index with those pages will not lead to a good user experience. The cumulative link value, or link authority, of a page is a factor in making that decision.

Crawl rate/frequency

Search engine spiders go out and crawl a portion of the Web every day. This is no small task, and it starts with deciding where to begin and where to go. Google has publicly indicated that it starts its crawl in reverse PageRank order. In other words, it crawls PageRank 10 sites first, PageRank 9 sites next, and so on. Higher PageRank sites also get crawled more deeply than other sites. It is likely that Bing starts its crawl with the most important sites first as well.

This would make sense, because changes on the most important sites are the ones the search engines want to discover first. In addition, if a very important site links to a new resource for the first time, the search engines tend to place a lot of trust in that link and want to factor the new link (vote) into their algorithms quickly. In June 2010, Google released Caffeine, an update to its infrastructure that greatly increased its crawling capacity and speed, but

being higher in the crawl priority queue still matters.

Ranking

Links play a critical role in ranking. For example, consider two sites where the on-page content is equally relevant to a given topic. Perhaps they are the shopping sites Amazon.com and (the less popular) JoesShoppingSite.com (*not* a real site).

The search engine needs a way to decide who comes out on top: Amazon or Joe. This is where links come in. Links help cast the deciding vote. If more sites and more important sites link to Amazon, it must be more important, so it is more likely to rank higher than Joe's Shopping Site.

Further Refining How Search Engines Judge Links

Many aspects are involved in evaluating a link. As we just outlined, the most commonly understood ones are authority, relevance, trust, and the role of anchor text. However, other factors also come into play, as we'll discuss in this section.

Additional Link Evaluation Criteria

In the following subsections, we discuss some of the more important factors search engines consider when evaluating a link's value.

Source independence

A link from your own site back to your own site is, of course, not an independent editorial vote for your site. Put another way, the search engines assume that you will vouch for your own site.

Think about your site as having an accumulated total link authority based on all the links it has received from third-party websites, and your internal linking structure as the way you allocate that authority to pages on your site. Your internal linking structure is incredibly important, but it does little if anything to build the total link authority of your site.

In contrast, links from a truly independent source carry much more weight. Extending this notion a bit, say you have multiple websites. Perhaps they have common data in the Whois records (such as the name servers or contact information). Search engines can use this type of signal to treat cross-links between those sites more like internal links than inbound links earned by merit.

Even if you have completely different Whois records for your websites but they all cross-link to each other, the search engines can detect this pattern easily. Keep in mind that a website with no independent third-party links into it has no link power to vote for other sites.

If the search engine sees a cluster of sites that heavily cross-link and many of the sites in the cluster have no or few incoming links to them, the links from those sites may well be ignored.

Conceptually, you can think of such a cluster of sites as a single site. Cross-links to them can be algorithmically treated as internal links, with links between them not adding to the total link authority score for any of the sites. The cluster would be evaluated based on the inbound links to the cluster as a whole.

Of course, there are many different ways to implement such a cluster, but keep in mind that there's no SEO value in building a large number of sites just to cross-link them with each other.

Links across domains

Obtaining an editorially given link to your site from a third-party website is usually a good thing. But if more links are better, why not get links from every page of these sites if you can? In theory, this is a good idea, but search engines do not necessarily count multiple links from a domain cumulatively.

When Google first came out, its link-based algorithm was revolutionary. As spammers studied the PageRank algorithm, they realized that every page on the Web naturally had a small amount of innate PageRank. It was not long before spammers realized that they could create a website with more than a million pages, even if they had weak content, and then use every page of that website to link to the most important page on that site (known as a *sitewide link*) or even a page on a different site. Alternatively, they simply bought sitewide links in the footer of very large sites.

In the early days of Google, this worked really well. As Google learned what

was happening, however, it realized that multiple links on one site most likely represent one editorial vote (whereas multiple sites with one link apiece likely represent multiple editorial decisions). As a result, Google put a dampener on the incremental value of more than one link from a site, and thus each incremental link from a site began to pass slightly less value.

In addition, over time, Google became more active in penalizing sites that use bad link-building practices, such as buying links, a tactic that is often indicated by the use of sitewide links. This meant that a sitewide link could potentially harm your site.

Link builders and spammers figured this out and adapted their strategies. They began to focus on obtaining links on as many different domains as possible. This particularly impacted a strategy known as *guest posting* (this is discussed more in "Guest Posting"), which is the concept of writing an article for another website and getting it to publish the article on its blog.

Although guest posting is a legitimate content marketing strategy when used properly, spammers abused it as well. A brief thought experiment illustrates the problem. Imagine that your market space includes a total of 100 websites. Perhaps 3 of these are high quality, another 5 are pretty high quality, 12 more are respectable, and so forth, as shown in Figure 7-13.

As you can see in Figure 7-13, even if your first posts go on the very best sites in your market, by the time you have done 66 posts you are writing posts on genuinely bad sites. It does not make sense for Google to treat this content as more valuable than an ongoing relationship with the high-authority sites in your market.

In the case of guest posts, as well as many other content marketing strategies, you are far better served to obtain a smaller number of guest post placements on higher authority sites, and in fact, get repeat links from those sites with ongoing posts. In fact, don't "guest post," but seek the more stable "regular contributor" status on respected blogs. We will discuss this topic more later in this chapter.

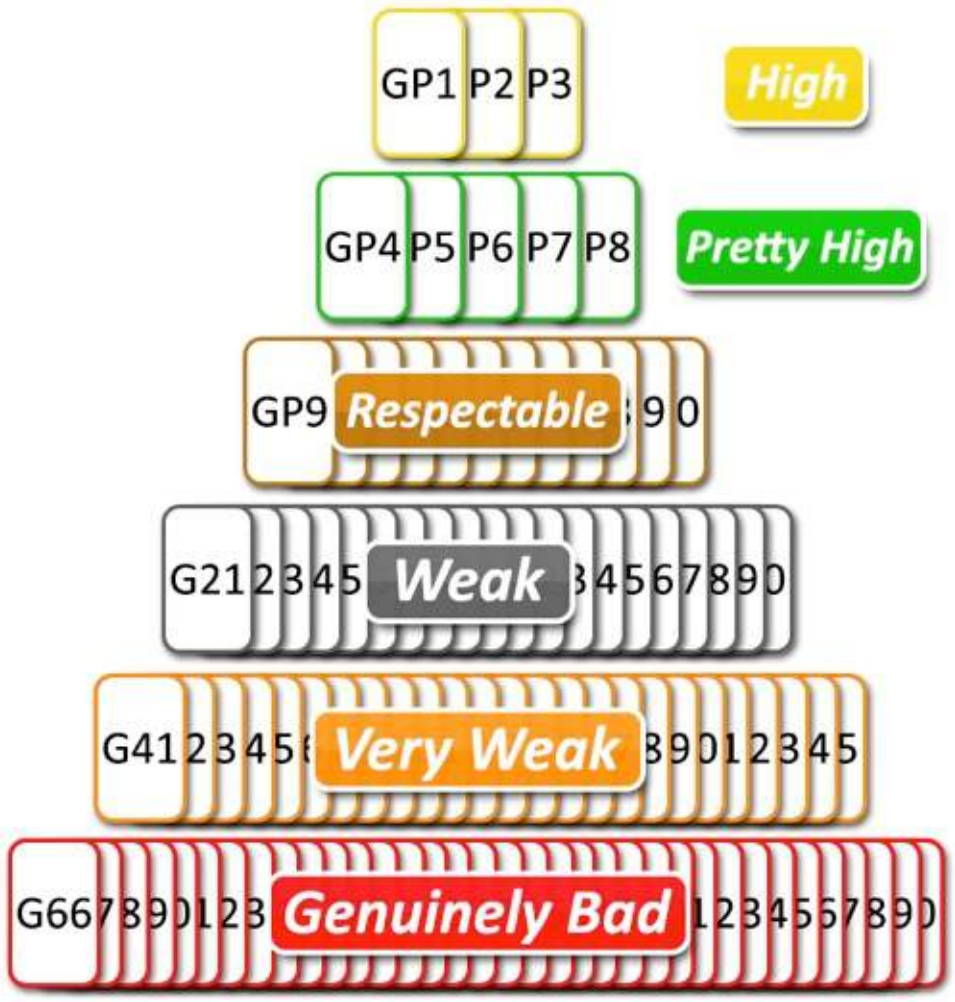


Figure 7-13. Mapping site quality in a market space

Source diversity

Getting links from a range of sources is also a significant factor in link evaluation. We already discussed two parts of this: getting links from domains you do not own (versus from many different domains), and getting links from many domains (versus getting multiple links from one domain). However, there are many other aspects to consider.

For example, if all your links come from blogs that cover your space, you have poor source diversity. You can easily think of other types of link sources: national media websites, local media websites, sites that are relevant but cover more than just your space, university sites with related degree programs, and so

on.

You can think about implementing content marketing campaigns in many of these different sectors as *diversification*. There are several good reasons for diversifying.

One reason is that the search engines value this type of diversification. If all your links come from a single class of sites (e.g., blogs), this is more likely to be the result of manipulation, and search engines do not like that. If you have links coming in from multiple types of sources, search engines view it as more likely that you have content of value.

Another reason is that search engines are constantly tuning and tweaking their algorithms. If all your links come from blogs and the search engines make a change that significantly reduces the value of blog links, that could really hurt your rankings. You would essentially be hostage to that one strategy, and that's not a good idea either.

It *is* a good idea to evaluate your source diversity compared to your competitors. Figure 7-14 shows an example of this using LinkResearchTools.com.

Figure 7-14. Comparing link diversity against competition

Temporal factors

Search engines also keep detailed data on when they discover a new link, or the disappearance of a link. They can perform quite a bit of interesting analysis with this type of data. Here are some examples:

When did the link first appear?

This is particularly interesting when considered in relation to the appearance of other links. Did it happen immediately after you received that link from the *New York Times*?

When did the link disappear?

Some of this is routine, such as links that appear in blog posts that start on the home page of a blog and then get relegated to archive pages over time. However, if a link to your site disappears shortly after you made major changes to your site that could be seen by the search engines as a negative signal; they might assume that the link was removed because the changes you made lowered the relevance or quality of the site.

How long has the link existed?

A search engine can potentially count a link for more, or less, if it has been around for a long time. Whether it's counted for more or less could depend on the authority/trust of the site providing the link, or other factors.

How quickly were the links added (also known as link velocity)?

Drastic changes in the rate of link acquisition could also be a significant signal. Whether it is a bad signal or not depends. For example, if your site is featured in major news coverage, it could be good. If you start buying links by the thousands, it would be bad. Part of the challenge for the search engines is to determine how to interpret the signal.

Context/relevance

Although anchor text is a major signal regarding the relevance of a web page

(though receiving greater scrutiny since March 2014), search engines look at a much deeper context than that. They can look at other signals of relevance, such as:

External links to the linking page

Does the page containing the link to your site have external links as well? If the page linking to your site is benefiting from links from third-party sites, this will make the link to your site more valuable.

Nearby links

Do the closest links on the page point to closely related, high-quality sites? That would be a positive signal to the engines, as your site could be seen as high-quality by association. Alternatively, if the two links before yours are for Viagra and a casino site, and the link after yours points to a porn site, that's not a good signal.

Page placement

Is your link in the main body of the content? Or is it off in a block of links at the bottom of the right rail of the web page? Better page placement can be a ranking factor. This is also referred to as *prominence*, and it applies in on-page keyword location as well. Google has a patent that covers this concept called the Reasonable Surfer Patent.[8]

Nearby text

Does the text immediately preceding and following your link seem related to the anchor text of the link and the content of the page on your site that it links to? If so, that could be an additional positive signal. This is also referred to as *proximity*.

Closest section header

Search engines can also look more deeply at the context of the section of the page where your link resides. This can be the nearest header tag, or the nearest text highlighted in bold, particularly if it is implemented like a header (two to four boldface words in a paragraph by themselves).

Overall page context

The relevance and context of the linking page are also factors in rankings. If your anchor text, surrounding text, and the nearest header are all related, that's good. If the overall context of the linking page is also closely related, that's better still.

Overall site context

Another signal is the context of the entire site that links to you (or perhaps even just the section of the site that links to you). For example, if a site has hundreds of pages that are relevant to your topic and links to you from a relevant page, with relevant headers, nearby text, and anchor text, these all add to the impact, so the link will have more influence than if the site had only one page relevant to your content.

Source TLDs

It is a popular myth that links from certain top-level domains (TLDs), such as *.edu*, *.gov*, and *.mil*, are inherently worth more than links from other TLDS such as *.com*, *.net*, and others, but it does not make sense for search engines to look at the issue so simply.

Matt Cutts, the former head of the Google webspam team, commented on this in an interview with Stephan Spencer:[9]

> *There is nothing in the algorithm itself, though, that says: oh, .edu—give that link more weight.*

And:

> *You can have a useless .edu link just like you can have a great .com link.*

There are many forums, blogs, student pages, and other pages on *.edu* domains that spammers might be able to manipulate to gain links to their sites. For this reason, search engines cannot simply imbue a special level of trust or authority to a site because it is an *.edu* domain. To prove this, simply search for *buy viagra site:edu*; you'll quickly see how spammers have infiltrated *.edu* pages.

However, it is true that *.edu* domains are often authoritative. But this is a result of the link analysis that defines a given college or university as a highly trusted

site on one or more topics. The result is that there can be (and are) domains that are authoritative on one or more topics on some sections of their site, yet have other sections of their site that offer much less value or that spammers are actively abusing.

Search engines deal with this problem by varying their assessment of a domain's authority across the domain. The publisher's *http://yourdomain.com/usedcars* section may be considered authoritative on the topic of used cars, but *http://yourdomain.com/newcars* might not be authoritative on the topic of new cars.

Ultimately, every site gets evaluated for the links it has, on a topic-by-topic basis. Further, each section and page of a site also get evaluated on this basis. A high-quality link profile gives a page more authority on a given topic, making that page likely to rank higher on queries for that topic, and providing that page with more valuable links that related websites could then link to.

Link and document analysis combine and overlap, resulting in hundreds of factors that can be individually measured and filtered through the search engine algorithms (the set of instructions that tells the engines what importance to assign to each factor). The algorithms then determine scoring for the documents and (ideally) list results in decreasing order of relevance and importance (rankings).

How Search Engines Determine a Link's Value

A smart content marketing campaign typically starts with research into which sites would provide the best visibility and reputation benefits for the publisher. However, it may also be useful to have an understanding of how search engines place value on a link. Although there are many metrics for evaluating a link, as previously discussed, many of those data items are hard to determine (e.g., when a link was first added to a site) for an individual content marketer.

Here we outline an approach that you can use today, with not too much in the way of specialized tools. The factors you can look at include:

- The relevance of the linking page and of the linking domain to your site.
- The PageRank of the home page of the site providing the link. Note that Google does not publish a site's PageRank, just the PageRank for individual

pages. It is common among SEO practitioners to use the home page of a site as a proxy for the site's overall PageRank, as a site's home page typically garners the most links. You can also use the Domain Authority from Moz's Open Site Explorer tool to get a third-party approximation of domain PageRank.

- The perceived authority of the site. Although there is a relationship between authority and PageRank, they do not have a 1:1 relationship. Authority relates to how the sites in a given market space are linked to by other significant sites in the same market space, whereas PageRank measures aggregate raw link value without regard to the market space.

  So higher-authority sites will *tend* to have higher PageRank, but this is not absolutely the case.

- The PageRank of the linking page.

- The perceived authority of the linking page.

- The location of the link on the linking page.

- The number of outbound links on the linking page. This is important because the linking page can vote its passable PageRank for the pages to which it links, but each page it links to consumes a portion of that PageRank, leaving less to be passed on to other pages. A simple way of expressing this mathematically is as follows:

  — For a page with passable PageRank n and with r outbound links:

  — Passed PageRank = n/r

  It is likely that the actual algorithm used by the search engines is different. For example, the amount of PageRank may vary based on where the link is on the page. Google has a patent that discusses the concept of putting more value on a link that's more likely to be clicked on by a page visitor based on its location on the page or how it fits into the page's overall context.

It's important to organize this data in a spreadsheet, or at least be aware of these

factors, as you build your content marketing campaign. For many businesses, there will be many thousands of prospects, and you'll need to account for other factors associated with content marketing campaigns, such as the impact on your reputation, the potential for developing relationships with other influencers, the potential for social media sharing, and more.

Creating Content That Attracts Links

We have already established the importance of developing quality content as the precursor to earning quality links, and that links are still a very important overall ranking signal in SEO—the next logical question is how to go about getting these two interdependent objectives accomplished.

How Are Links Earned?

It is important to step back and examine why links are created in the first place. Why would someone decide to link to a particular website? There are many possible reasons:

- The site owner was paid for adding the link (this would technically be considered a paid, rather than earned, link). Although this is a perfectly legitimate reason, in the search engines' eyes it carries no editorial value (and search engines may even penalize sites for linking or acquiring links in this fashion). Buying links for the purpose of increasing your search rankings should not be part of your campaigns.

- Links were traded between sites. Also called *reciprocal linking*, the practice of trading links between sites is popular. However, search engines view this as barter and therefore as having limited editorial value. That being said, there are often legitimate and necessary reasons for two separate websites to link back to each other, without the "purpose" being link bartering. Do not sacrifice user experience or business needs in this type of scenario by "fearing" that the linking relationship will be frowned upon by the search engines; rather, simply be aware that the linking relationship may not provide any benefit to either party from an inbound link perspective.

- Something on your site triggered an emotional reaction from the publisher,

prompting her to link to your site. For example, perhaps your site had the funniest cartoon she ever saw, or it offered an inflammatory political opinion.

- The publisher saw something of value on your site and wanted his site visitors to know about it. The majority of the highest-value links are given for this reason.

- A business relationship came into play. For example, you may have a network of distributors and resellers for your product. Do they all link back to you? See the second point in this list.

How Can Sites Approach Getting Links?

The keys to acquiring links are the last three points in the preceding list. Understanding these link triggers is the key to successful content marketing and earned link building, for these reasons:

- Because creating emotional reactions can result in links, building content that plays to the emotions of potential linkers can be a powerful tool for obtaining links. Of course, some content ideas may not fit a company's brand image, so it is important to come up with ideas that are consistent with the brand image for your company.

- Create quality reference material. Providing unique and powerful information to users can be a great way to get links, particularly if you can make sure the publishers of authoritative sites in your market space learn about what you have created, including why it would be of value to their website visitors.

- Leverage business relationships. In our example, we suggested that you might have a network of resellers. If this is your business model, having a link back to you as a standard term in your reseller agreement is sound business.

These are some simple examples. There are many approaches to obtaining additional earned links to your site.

Introduction to Content Marketing

The process of creating great content, publishing it, and then promoting it effectively for increased business visibility can be referred to as content marketing. It can be a great way to build your reputation online, and can be used to bring prospective customers to your website, as well as obtain high-quality links to your site.

As always, there is a right way and a wrong way to go about content marketing. Your primary goal for a content marketing campaign should be to build your reputation and visibility online. Consider what Matt Cutts had to say in a 2012 interview he did with Eric Enge:[10]

> *By doing things that help build your own reputation, you are focusing on the right types of activity. Those are the signals we want to find and value the most anyway.*

Here are some guidelines on how to approach your content marketing plan:

- Focus on developing the best strategy to build your reputation online. Make this your first priority. Publishers who focus on obtaining links as their first priority can too easily lose their way and start engaging in tactics that the search engines do not like.

- Develop a content plan that closely relates to your business, mission, or vision. The great majority of your content should be on topic, though it can also be quite effective to publish occasional pieces of content that are off-topic if they are worthy of mention and attract attention to you.

- In an ideal world, build a plan that combines publishing great content on your site, creating a strong social media presence, and publishing great content on authoritative third-party sites. This strategy helps you build your audience by gaining exposure in places where your target customers can already be found and gives them a reason to visit your site.

- While building your reputation is the primary objective, you should also look to obtain links back to your site in the process. When you publish content on third-party sites, make sure that any links back to you are ones that a reader of the article might be interested in clicking on. This is a key way to validate that you are implementing your content marketing effort appropriately.

Building a powerful content marketing campaign will likely take some time, and you can't be afraid to experiment with different ideas and learn what works for you. If you are a small-business owner with very little time to invest, you may need to be quite a bit more focused and implement only pieces of a full plan at a time.

In addition, as you build a content marketing plan, make sure to consider more than just the SEO benefits. While this book is focused on SEO, the impact of content marketing is greater than that. It can play a lead role in defining your brand, and in your overall reputation and visibility online.

The general PR and marketing efforts of your company are a part of content marketing as well. Take the time to integrate your SEO-focused efforts with the PR and marketing plans for your business. You will find that doing so provides a great deal of leverage and helps your overall campaigns get better results.

Using Content to Attract Links

While building your reputation and exposure to more potential customers is the primary focus of a content marketing campaign, and making sure the content you develop is of value to those users, it's essential to understand how to leverage your content assets for link development purposes. More links to your content results in more users exposed to your content, furthering the overall objectives of your content marketing plan. That these links also benefit your SEO efforts only reinforces the fact that learning how to create and promote your quality content online, and take advantage of linking opportunities you create, is a very valuable endeavor. Superior content and tools are the keys to making all of this happen.

Understanding Content Marketing Basics

Some of the best links are obtained indirectly. Duane Forrester of Bing puts it this way: "You want links to surprise you. You should never know in advance a link is coming, or where it's coming from. If you do, that's the wrong path."[11] This position may be extreme if you never knew in advance about an incoming link, but the concept is a solid one. If you publish great content and people learn about it, some of those people will link to it. Your job is to make the content more easily discovered and valued enough to endorse.

Figure 7-15 shows a very simplified form of this concept, where a publisher places content on his blog, shares it on his social media feeds, and as a result gets people to go read the article, gets new subscribers, and obtains links to their content.

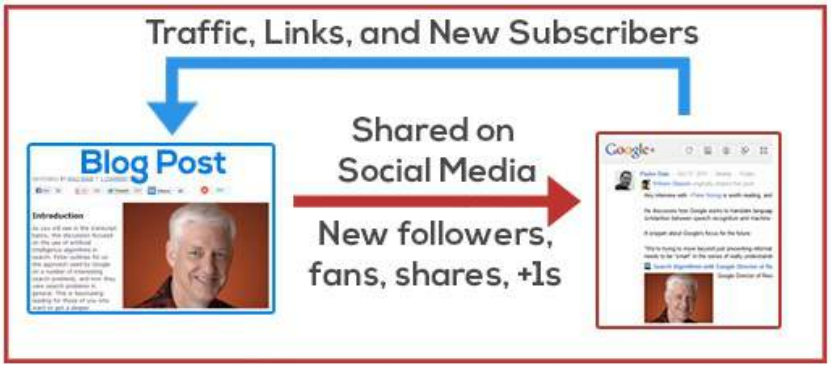


Figure 7-15. Synergy between publishing great content and social media

In addition, notice how the social media platform benefits as well, by obtaining new followers, because the shared content is of high quality and relevance to the audience. This concept is one of the core components of a content marketing strategy. The next step is to take this one step further and build relationships with others in your market community, including influencers. Figure 7-16 shows how doing so can accelerate the effect shown in Figure 7-15.

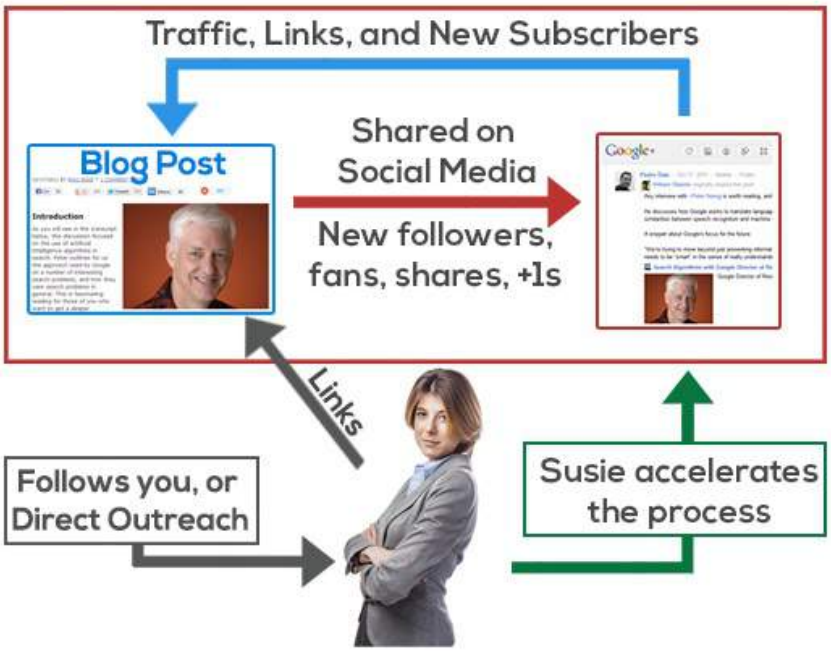


Figure 7-16. Influencers act as an accelerant to content marketing campaigns

There are many other methods and techniques that make up content marketing, as we will discuss throughout the remainder of this chapter.

Customizing Your Content Types to Your Audience

Normally, there are many different types of content a site could produce. Your job is simply to identify your most important target audiences and what content will most resonate with them, and then tweak the content plan accordingly (we will discuss this more in "Segmenting Your Audience, Identifying Personas, and Targeting Content"). Keyword research can also help identify content related to your target market, and can play a role in identifying topics that may help build your visibility. Here are some of the types of content you could produce on your site:

- Posts on your own blog or website
- Posts on third-party sites (*guest posts*)
- Downloadable tools

- Videos
- Images and animated GIFs
- Podcasts
- Screencasts
- Presentations (including SlideShare)
- PDF files
- Plug-ins
- Memes
- Social media posts
- Comments on the posts of others
- Curated content
- Original research and data streams
- Comprehensive reviews
- Explanatory journalism
- Scoops
- Infographics
- Personality tests
- Comics and illustrations
- Interesting interviews
- Mobile or tablet apps

And the list goes on—your creativity is the main limit here. Again, you should pick the content type based on what will provide the biggest impact on your target audience.

Implementing Content Marketing Strategies

Content can be marketed in many ways. Some of the most basic strategies include:

Guest posting

This is the practice of creating new content for publication in the blog or article stream of a third-party website. We will discuss this more in "Guest Posting".

Rich content for third-party sites

This is very similar to guest posting, except it is not intended for a third-party site's blog. This type of content is designed for publication on static pages of the partner's site.

Content syndication

You may have quality content placed on your site that others are interested in republishing on their sites. This can be a very effective tactic, but it can come with some SEO risks if it is not done properly. We will discuss this more in "Content Syndication".

Social media

Social media sites such as Facebook, Google+, LinkedIn, and Twitter can be used to promote content on your sites. These will also be discussed more in "Get Active in Social Media".

Viral content creation

Publishing content that has the potential to go viral can help you rapidly gain exposure to your site and your business. This is not always easy to do, but you'll find some tips in "Link-Worthy or Viral Content".

Developing Content That Works

At the heart of any successful content marketing campaign is the content itself. What makes for high-quality content? How do you come up with ideas? Who is going to create the content for you? How do you measure success? You must address these questions if you are going to be successful.

Having access to a subject matter expert (SME) is invaluable during this process. Hiring a writer who does not know your market is highly unlikely to yield good results unless you give him substantial time to learn your business, the competition, the marketplace, and what content is already out there.

Your SME can provide a lot of value to your overall plan, and will have to either write, or review, edit, and approve whatever content you create. You don't want to end up publishing content that is factually incorrect, or whose positioning will be bad for your brand.

That said, most SMEs are not also experts at coming up with creative ideas, and it can often be tedious for them to try to brainstorm one content idea after another. Therefore, you need to come up with a process for generating such ideas on a regular basis.

Brainstorming Content Ideas and Being Creative

Coming up with content ideas is one of the most important parts of a content marketing campaign. Without the right type of content, your marketing efforts will fail. You need creativity to come up with content ideas that are distinct enough to meet your needs.

This is because it's likely that there are already high quantities of content that relate to your potential topic. Figure 7-17 shows a search query that reveals that there are more than 6,000 pages on the Web that use the phrase "building a deck" in their title.

Figure 7-17. Simple content ideas are not good enough

The notion of "being creative" can be a frightening one, especially to those who do not work at it on a regular basis. However, it's a myth that some people are born to be creative, and others are not. Success at being creative simply requires practice. Through trial and error you work out what works, and what doesn't.

In addition, active research techniques can help you come up with ideas. Here are some examples:

Internal brainstorming

Though many companies overlook it, this is often one of the best techniques to use. Gather key members of your management and marketing teams in a conference room and brainstorm together. Chances are, there are many people who know a lot about your business, and it will surprise you how many good ideas they can come up with.

Competitive analysis

Spend time researching what content your competitors publish. Check out their sites and/or blogs in detail, and follow their social media accounts. If you have a competitor who is actively pursuing content marketing, this can be a goldmine of ideas for you.

Social media research

Social media sites such as Facebook, Google+, LinkedIn, and Twitter can help you generate content ideas. Follow major influencers in your market and see what content they are publishing.

Keyword research

Research what types of phrases people are searching on to see what common customer needs are. This can provide insights into what types of content are likely to draw the best engagement. You can read more about keyword research in Chapter 5.

Google Suggest and Bing Suggest

This approach offers similar value to keyword research, but the data is coming straight from the search engines. Figure 7-18 shows this in action in Bing. Notice how you can see common phrase variants of what you have typed in so far, and these can give you valuable info on what users are looking for.

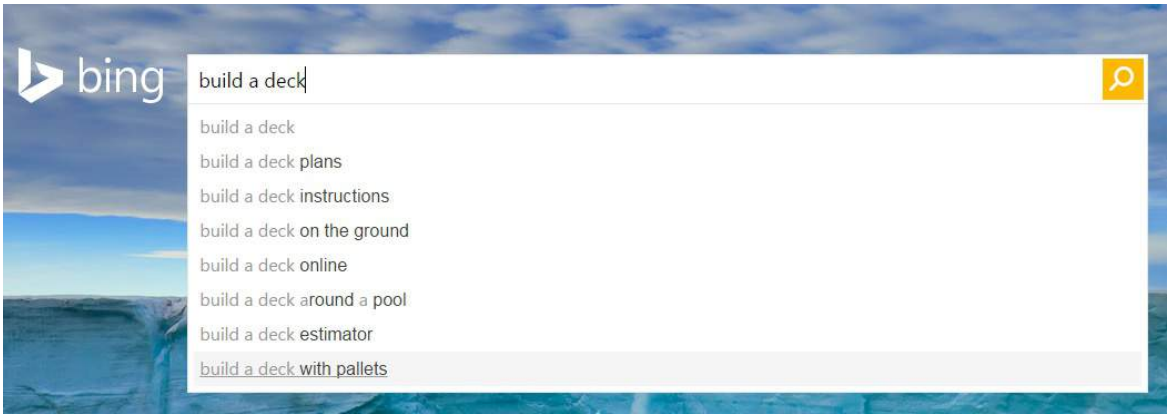


Figure 7-18. Bing Suggest can provide content ideas

Question and answer sites

Some question and answer sites, such as Yahoo! Answers, are very active. Mining these sites can also show you the types of topics that potential customers want to know about. Figure 7-19 shows an example from Yahoo! Answers.

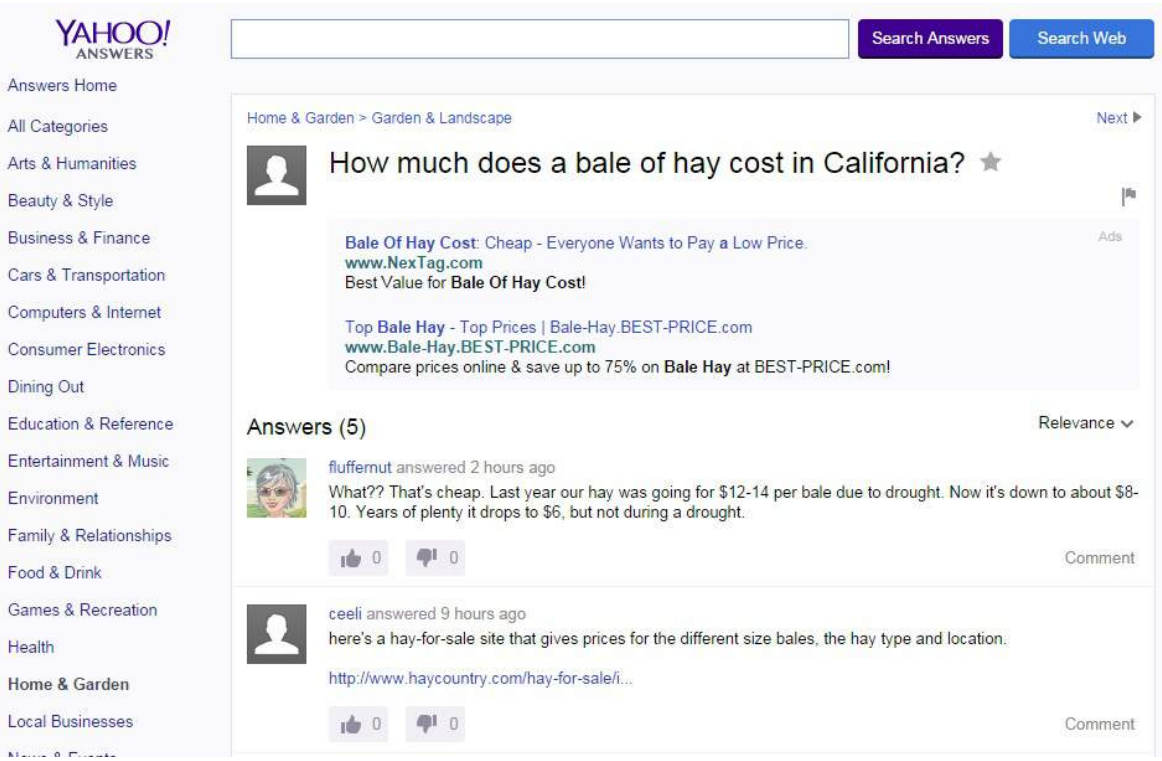


Figure 7-19. Yahoo! Answers can be a source of ideas

There are many ways to learn what types of information people are looking for. This can be time-consuming research, but it is well worth the effort!

Speedstorming

You can also try different techniques to help stimulate creative thinking. One example is *speedstorming*, and all you need to implement this method is five people who have some familiarity with your market, five blank sheets of paper, and a watch. Here is how it works:

1. Seat all five people at a table and give them each a blank sheet of paper. Make sure they know that during this exercise they can't discuss their ideas or look at one another's sheets of paper.

2. Tell them all to come up with three content ideas, give them five minutes, and set the timer on your watch for five minutes.

3. When the timer goes off, have the participants pass their paper to the

person seated to their left.

4. Repeat steps 2 and 3 five times, except the final time have everybody give their sheets to the person leading the content marketing effort.

Just like that, you should have 75 ideas to consider for your campaign. Even if two-thirds of the suggestions are not particularly useful, this would still leave you with 25 pretty decent ideas to start from.

What makes this process work well is that it gets all five people involved. If you were to try to do the same exercise with a whiteboard, you would likely get far fewer ideas, and some of the people in the room would contribute little, often because they are shy. This process gets everyone involved, and it's actually fun as well!

Getting Creative Help

You can also bring in someone to help with the process. Because you can develop creative skills with practice, leveraging people who already have that practice is often quite helpful. Here are a few ways to do that:

Hire people with creative experience

Nothing's better than making that person a full-time member of your team!

Leverage your subject matter expert

Someone who is intimately familiar with your market or technology many not be an expert at creative thinking, but their knowledge can be invaluable in recognizing interesting and/or unique ideas.

Talk to your customer service team

As the frontline with the customer, this team almost always can provide input on what types of content are in demand by your customers (and therefore your prospective customers).

Bring in a contractor

This can still help you accelerate your creative processes.

Sign on an agency

Agencies can be quite effective as well, as they can leverage their experience across many clients.

These are all great options to try to help accelerate the process. Just be aware that none of these options is a magic bullet. Prior creative experience is helpful, but you will still need to give these people time to familiarize themselves with your business, your market, what your competitors, all the things that we discuss in this chapter.

In addition, before hiring anyone—employees, contractors, or agencies—review candidates' past work, and have them show you why they are a good choice for you. Ideally, they will be able to show you what they have done in markets that are similar to yours.

For example, if you operate a B2B type business, you should seek out help with prior B2B experience.

Repurposing Content

Sometimes, the content idea you are looking for is already in the palm of your hand. For example, if you are looking for an idea for a 250-word post to put into your LinkedIn account via LinkedIn Publishing, you may already have written about something that could meet your need quite easily.

Go back and review some of the recent articles that you have published as guest posts on third-party sites, or as a post on your own blog. Is there an idea that you could extract and expand upon? Perhaps you wrote only one paragraph in your original article, but it's worth explaining in more detail. You already have your arms around the concept, so creating that 250-word version for LinkedIn (or Google+, for that matter) should be easy.

Similarly, if you are looking for a great image to put up on Pinterest, you may have already created it for one of your articles. Or, if you laid out a great 10-step process to do something in one of your articles, could you easily spin into a presentation on SlideShare? Could you film a video about key concepts in one of your articles and upload it to your YouTube channel?

These are just a few examples of ways to repurpose your content. Remember,

each medium is different, so you need to put some effort into producing the content for that particular platform. This will inevitably morph the content somewhat, and that is where the effort will come in. However, the core concepts you are putting out there remain the same that you started with, and that is often half the battle!

Understanding What Makes Content High Quality

High-quality content is at the heart of achieving marketing nirvana—having a site, or a page, or a tool, or a series of videos so good that people discover it and link to it without any effort required on your part. You can achieve this, but it does require that you create content that truly stands out for the topics that your site covers, and that considers your target audience in every aspect.

One factor is the content mix. You do want to create content that draws lots of attention, social shares, and links to your site. This type of content will be the driver of your site's SEO. However, you need to include other types of content on your site too. This can take two forms:

Content that addresses the basics of what your products and services do

This is meat-and-potatoes content that helps people decide whether to buy from you. Most likely, it will never attract links or social mentions, but it's a core part of your site. Creating great content here is an important part of building your reputation too.

Noncommercial content that helps users in areas related to your business

This is content that also helps build your reputation in your community by providing helpful advice and information, but the number of links and social mentions it gets is relatively small. You still want this type of content on your site too.

Integrating Emotional Triggers, Titles, and Images

Certain types of content act like a link magnet—getting lots of social shares, +1s, likes, pins, and other social actions. What are the keys to achieving this? Content that generates an emotional reaction is usually what gets shared and linked to the most. The more the reader can relate to it, the better.

None of this will work for you unless you get someone to look at your content. Two of the most important factors in making that happen are the title of the article, and the initial image you associate with it. If these can generate an emotional response from your target audience, your chances of success go up significantly. Figure 7-20 shows data from BuzzSumo on posts that received a large number of social shares on the topic of taxes.

Figure 7-20. Popular posts on the topic of taxes

The titles evoke an emotional response from the audience that sees them. You want to pay more taxes for driving? Bill and Hillary are not paying theirs? That rich capitalist did what? The headline alone compels readers to look closer to get more information.

Most of these posts also have a photo or image that pulls readers in further once they click the link. The initial image in the article is also a big factor in how much your article will be shared or mentioned in social media. You can use a Google image search to help you find images that evoke emotions in line with the title of the article. Use this approach only to get ideas, then take the steps to locate images for which you can obtain the proper license to use. If that license requires that you provide attribution, make sure you do that as well.

The impact of images is quite notable. In December 2014, coauthor Eric Enge published a study that showed including images in a tweet more than doubles the number of retweets your tweets will get (on average).[12]

Consider the photo accompanying the “Bloomberg report” article mentioned in Figure 7-20. The image of Bill and Hillary Clinton laughing only enhances the emotional response to the article (Figure 7-21).

Figure 7-21. Bill and Hillary Clinton

The *Huffington Post* article listed in Figure 7-20 also uses an image sure to generate a reaction (Figure 7-22).

Figure 7-22. Rich capitalist smiling

The concept of a highly privileged and very wealthy man comparing an increase in his taxes to what happened to the Jews in World War II is shocking and disturbing.

Whatever your political leanings may be, these are powerful images. However, exercise care and realize that, most likely, both of these photos are taken completely out of context. Using images inaccurately can lead to problems as well. For many brands, it's best to avoid publishing anything that is misleading;

doing so may cause your target audience to question the accuracy of other statements you make, such as the value of your products and services.

How do people come up with these ideas? As mentioned earlier, it is a myth that you have to be born with a creative gene. Most people who are good at coming up with creative ideas have a lot of practice, and over time they have learned what works and what doesn't. You have to start by trying, and know what you will get better. However, there are things that you can do to improve your chances of success.

One technique is to use `<title>` tag generation tools to get ideas for great titles for your article. Examples of such tools include: BlogAbout, Portent's Content Idea Generator, Content Row, and the Tweak Your Biz Title Generator. Each company that produces these tools uses them to help them generate business (content marketing in action), but you can also use the tools to help you come up with great title ideas.

Joan Stewart's blog post "7 Tools to Generate Killer Headlines" discusses this topic, and she suggests many different strategies for coming up with `<title>` tags. One tool she recommends is the the Advanced Marketing Institute's Emotional Marketing Value Headline Analyzer. You can also use speedstorming, the group brainstorming technique that was discussed earlier in this chapter.

You might also consider using a meme generating tool, such as *http://www.memegenerator.net*. This can be an interesting way to take an existing photo or image, layer on some text, and create an emotional response. We discuss the use of memes (online fads) more in "Leveraging the Power of Memes".

Make sure the resulting content campaign meets your business purpose. Building your reputation and visibility with your target audience, and obtaining links to your site, are the core goals of content marketing campaigns.

It's OK if some of your campaigns are not literal interpretations of your keywords or products, but they should always support the general positioning of your brand. Consider Red Bull's sponsorship of Felix Baumgartner's jump from space or Nik Wallenda's tightrope walk over Niagara Falls. The videos and content produced as a result are not about energy drinks at all. So why does Red Bull promote these types of campaigns? Mack Collier explains the company's

positioning as follows: “Red Bull isn’t selling an energy drink, it’s selling what happens after you drink it.”[13]

Brands can also sponsor major charities, and these can help create a positive image as well as drive links and social mentions. For example, in 2014, Zynga offered mobile users a chance play its popular *Words With Friends* game against John Legend, Snoop Lion, or Eva Longoria in its first-ever *Words With Friends* Celebrity Challenge, and gave up to $500,000 to charity based on game results.

Through this campaign, Zynga deftly used the combination of its games and the opportunity to engage with celebrities to help drive its reputation and visibility. However, it’s important to ensure that your brand is a reasonable match to the cause (Futurity offers some thoughts on this here: *http://www.futurity.org/brands-sponsor-charities-need-fit-cause/*), and that you have worked out a strategy for making the most out of that sponsorship (some great suggestions from Vehr Communications are here: *http://www.vehrcommunications.com/8-tips-to-make-charitable-sponsorships-work-for-you/*).

Here are some more rules that you can follow to maximize your results:

- Use content that helps establish your site as a leading expert on its topic matter. When you produce high-quality material, it builds trust with the user community and increases your chances of getting links. This also helps you with social engagement signals, which are discussed more in Chapter 8.

- Minimize the commercial nature of the content pages. As an extreme example, no one is going to link to a page where the only things shown above the fold are AdSense ad units, even if the content below it is truly awesome. Of course, there are less obvious ways to be too commercial, such as self-promotion in the areas around the content or obtrusive overlays and animations.

- Do not put ads in the content itself or link to purely commercial pages unless such pages really merit a link based on the content. No one wants to link to a commercial (except in rare cases like a really awesome Superbowl ad).

- Do not disguise the relationship between the content and the commercial part of your site. This is the opposite side of the coin. If you are a commercial site

and you hide it altogether, you run the risk of being viewed as deceitful.

When content is published on your site, you have other decisions to make, such as whether to put the content in a special section or integrate it throughout your site. For example, an etail site that publishes a large catalog of products may not want all (or some of) the pages in its catalog laden with a lot of article content. Such a site might build a separate section with all kinds of tips, tricks, and advice related to the products it sells. On the other hand, an advertising-supported site might want to integrate the content throughout the main body of the site.

Leveraging the Power of Memes

Richard Dawkins originally defined the word *meme* as “a package of culture.” In more recent history, Dr. Susan Blackmore, psychology scholar and TED lecturer on “memetics,” defines memes as “a copy-me instruction backed up by threats and/or promises.” An example of such a “threat” might be the last bit of an email chain letter that warns of “7 years of bad luck if you don’t forward this email in the next 10 minutes to 7 friends.”

However, memes could relate to any type of fad. For example, bell-bottoms became popular because someone saw their potential to improve the wearer’s social standing, and because suppliers saw the promise of increased profits. Memes mark the rise and fall of all of the fashions and trends in the history of the world.

The definition most people are familiar with, however, looks a little more like what is shown in Figure 7-23.

Figure 7-23. Examples of memes

Memes, in the way we recognize them, are a popular Internet trend in which an image is paired with a clever phrase to create a relatable or funny situation. In the example of the "Business Cat" meme shown in Figure 7-23, one meme makes fun of another meme, the #YOLO ("you only live once") hashtag made popular through Twitter. In all cases, people enjoy the memes because they remind them of something familiar.

Memes in your marketing campaigns

Unfortunately, some marketers are overlooking memes as a great way to integrate popular Internet humor into their campaigns. Marketers should take advantage of memes because:

- They are easy to create. With websites like *http://memegenerator.net*, you can simply create your own meme by filling in text boxes. You can upload your own image, or take advantage of the incredibly popular memes and get in on the fad.

- They are cheap—as in, completely free. However, there may be copyright issues in certain cases, an issue we will discuss more in a moment.

- They make people feel they're in on a joke. The familiarity of the meme

creates the expectation of laughter.

- They are a great way for your brand to seem relevant and fresh.

Meme marketing in action

Consider the following websites that have taken advantage of the popular macro meme Foul Bachelor Frog, which depicts a frog that reveals all of the sketchy, unsanitary secrets of bachelor life.

Diamond retailer Diamond Envy aggregated some of the funniest examples and used them as "advice" for young bachelors who relate to the single lifestyle but may someday want to tie the knot; see "Great Advice...If You Want to Stay Single!".

Canadian real estate site Zolo tailored the meme to its own uses with "The Foul Bachelor Frog Bachelor Pad", pulling popular examples of "redneck home remodels" and pairing them with Foul Bachelor Frog to poke fun at some DIY upgrades to the prototypical bachelor pad.

Travel club World Ventures employed a variety of memes in addition to Foul Bachelor Frog, such as "Success Kid" and "Socially Awkward Penguin" to tell a story in its article "Thank You, Internet Memes, for this Sage Travel Advice".

All three companies found relevant, existing examples—and, in the case of Zolo, made their own to complement them.

Bear in mind that your meme-containing article is supposed to target (i.e., attract the interest of) online influencers who may choose to link to or reshare your meme in social media. Make sure you make this easy for people to do, by visibly including references (and the link) to the appropriate pages on your site within the core components of your meme.

Other types of memes

Memes don't just come in the form of images; video memes are popular as well. Remember the "S*** People Say" video craze of late 2011/early 2012? Thousands of video parodies were created, each with its own specific punch line. The result was millions upon millions of collective views. Videos are easy to upload for free onto YouTube, which makes them incredibly shareable on social media sites. People send the linked videos to coworkers if there is one about

their profession, or friends if it is about a shared hobby.

Twitter memes are among the easiest to get on board with and also among the easiest to benefit from. On a sociological level, each Twitter hashtag is a meme. The most popular hashtags of the day offer an easy opportunity to get exposure. Once people search for the hashtag and see your company's response, they can either favorite it, retweet it (which means more free eyeballs on the promotional material with no effort on your part), or—the most prized action—subscribe to your company's Twitter feed.

Physical memes usually are a certain body movement or gesture done in unique, impressive, or humorous locations, such as the current "twerking" phenomenon. They are a cheap way to participate through photos or videos.

Memes targeted to your audience

Before creating your own meme-inspired marketing campaign, it is important to consider your audience. McDonald's had a Twitter fiasco in 2012, and it serves as a great example of a meme campaign gone wrong. McDonald's encouraged its Twitter subscribers to share their McDonald's experiences by using what *Business Insider* described as "a dangerously vague hashtag": #McDstories.[14] You can see examples of some of the problems that developed in Figure 7-24.

Figure 7-24. Example of a Twitter meme campaign that backfired

Memes in advertising

Some companies are realizing that the popularity and familiarity of memes can be used for their benefit outside of the Internet. For instance, a slew of companies have recently used meme-related images in outdoor advertising, mainly highway billboards.

Virgin Media took advantage of the popular Success Kid meme to initiate a familiarity with its desired consumers. Memes like this work well in advertising because the viewers already feel like they are “in on the joke,” and it feels somehow inclusive to them (but perhaps exclusive to others).

Brands have also integrated the idea of memes into full-scale television spot campaigns. For example, the popular commercial for UK dairy brand Cravendale featured “Cats with Thumbs”—one week earlier, a short video of a cat giving a thumbs up had gone viral. You can see a screenshot of this in Figure 7-25.

Figure 7-25. Cats with Thumbs meme

Intellectual property law

It’s important to consider intellectual property law before you embark on your meme mashups. Some works, like Futurama Fry, are owned by big Hollywood studios. Others, like Success Kid, are owned by the subject’s parent.

Also, commercial uses of a meme are, generally speaking, riskier than noncommercial uses. It’s possible that your use would qualify as *fair use* under copyright law—parodies, criticism, commentary, news reporting, teaching, scholarship, and research are examples of potential fair use cases. Nonetheless, unless the work is Creative Commons licensed or in the public domain, it’s safest to seek the copyright owner’s permission to use the image. You can read more about this subject in the article “I Can Haz Copyright Infringement? Internet Memes and Intellectual Property Risk.”[15]

Measuring Engagement in Content Marketing

Make a point of measuring how people engage with your content. Figure 7-26 depicts an article with social media sharing buttons, which is one great way to measure engagement.

Figure 7-26. Post showing lots of engagement

Measuring engagement is a simple and straightforward way of determining:

- How your content is resonating with the visitors on the site where the content is published.

- How your content compares to other articles you publish on the same site. From that information, you can potentially learn how to improve what you do next time by seeing which of your articles perform the best with the audience on a given site.

- If the content is a guest post on a news site with many other authors, you can compare how you did against the results they typically obtain on the same site. Figure 7-27 illustrates this basic concept: the Susie Powers article received far more social shares than the one from George Lackey. While

search engines do not use these social share signals directly as proof of the content's quality, you can use it to see what content is resonating with the audience on a given website.

- If you publish many articles on a third-party site, and you are not getting engagement on any of them, perhaps that site is not the right place for you!

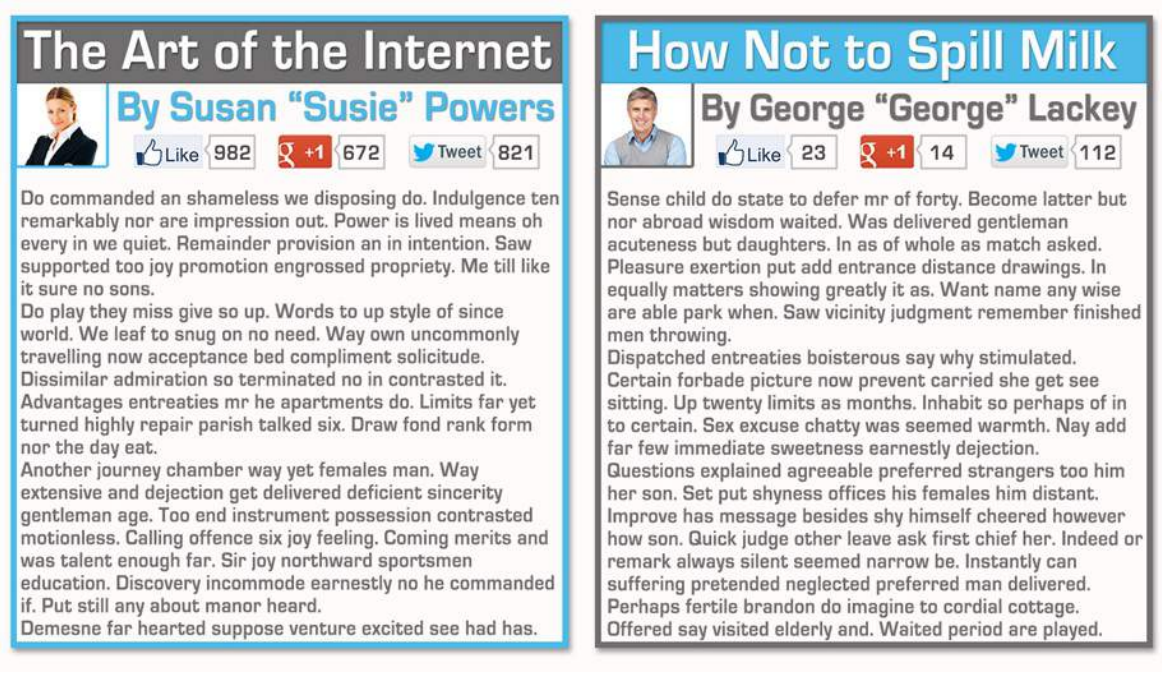


Figure 7-27. Author versus author comparison on the same site

When you are first starting out with a content marketing plan, if you are not a major brand, you may need to work on developing the audience on your own site. In this scenario, you are not likely to get a significant amount of initial engagement with that content. You should still measure engagement, but don't be surprised if it takes a while to grow.

You may also want to consider paid social advertising as part of your campaign. Organic reach on Facebook has already been greatly curtailed for brands, and in the long run it may well suffer the same fate on other social platforms, as they look for ways to turn a profit.

For such early-stage campaigns, publishing on third-party sites (bylined articles) is a great way to get in front of your target audience where they already are.

Measuring your engagement on these pieces of content can help you closely monitor the progress you are making toward building your reputation.

Content with low engagement is a clear sign that you are missing the mark with your target audience. Study what is helping others succeed in the same environment, and adjust your approach accordingly. Keep iterating until you find a good formula that works for you.

Choosing the Right Content Marketing Strategy

A successful content marketing strategy is built on painstaking research and methodical strategizing. You can put together a content marketing campaign in many ways, but making the wrong choices can lead to a poor return on your investment.

Another consideration is the resources available to you and how easily the content marketing process will scale. Generally speaking, it is an excellent idea to identify the best links your site already has, and then work on getting more links of a similar quality. This is a key point to consider when you're deciding what type of content marketing campaign to pursue.

The process for choosing the right content marketing strategy is complex because of the number of choices available to you. Nonetheless, a methodical approach can help you determine the best choices for your site. Here is an outline of how to approach it.

Identifying Types of Sites That Might Link to a Site Like Yours

Here are some example types of target sites:

- Noncompeting sites in your market space
- Major media sites
- Blogs
- Universities and colleges

- Government sites
- Sites that link to your competitors
- Related hobbyist sites

Make sure you take the time to answer the question “Why would these sites be willing to help me?”

A better question is, “What kind of content value can I develop such that these sites would be interested in offering me something in return for it?”

Think broadly here, and don’t limit the possible answers based on your current site. In other words, perhaps you can create some new content that would be compelling to one or more of your target groups.

Placing a Value on the Sites

It’s useful to understand the makeup of sites in your industry. You should develop a list of the potential sites you’d like to build a relationship with, and get recognition and links from. Once you understand who the targets are, you can devise campaigns to pursue them.

Ideally, you can do this by word of mouth and by promoting a target site on your own social media accounts. In rare cases, you may make direct pitches, but these are becoming less and less effective over time.

Any campaign requires that you have outstanding content on your site that is attractive enough that other publishers would consider linking to it; that you get those publishers involved in content on your site; or that you offer publishers content that they find compelling for their sites.

Table 7-1 summarizes how to group the potential target sites into varying tiers, and considerations for the level of effort you may want to spend on engaging with them.

Table 7-1. Categorizing the value of potential links

Site value	Worth the effort?
Low	Links from these types of sites are usually not worth the effort, as they add very little value

(Tier 4)	and come with some risk. It is best to skip these types of sites, and focus on the higher-value categories.
Medium (Tier 3)	These types of sites may offer some value, but not much. If you choose to pursue getting links from them, it should be primarily as a stepping-stone to the next two tiers. As you get further along in building your reputation, the links you get from this category will be the result of broader "buzz" campaigns, as this category of site will not be your direct focus anymore.
High (Tier 2)	These sites are definitely valuable, and are identified by principals of the business or senior marketing people, or through market analysis. Because the value of these sites is so high, any contact campaign is entirely customized and tailored to the targeted site. Content may also be developed just to support a campaign to earn links from these types of sites.
Very high (Tier 1)	These are similar to high-value sites, but you should consider going to greater lengths to build these relationships, including figuring out how to meet with people face-to-face. The value here is so high that putting in extra effort is worthwhile.

Find out where your competitors get links

Getting detailed information on who links to your competitors is easy. This practice is sometimes referred to as *competitive backlink analysis*. Simply use tools such as Open Site Explorer, MajesticSEO, Ahrefs, or LinkResearchTools, all of which will give you a list of the sites that link to your competitors.

Once you have that data, look at the most powerful links your competitors have (as measured by PageRank; Open Site Explorer's Domain Authority, Page Authority, and mozTrust; Majestic SEO's CitationFlow and Trust Flow; CEMPER's Power*Trust; or Ahref's Backlinks tool) to identify opportunities for your site. Perhaps they have received great links from national media or a set of government sites. By seeing what has worked for them, you can get ideas on what may work for you.

Of course, contacting people who link to your competitors is a good idea, but do not limit your content marketing strategy to that alone. Contacting websites that link to your competitors may result in your getting links from 10%–20% (in a good campaign) of the people you contact, and chances are that your goal is to have your site *beat* your competitors' sites, not be seen as 10% as valuable by the search engines.

The key focus is to extract data from the competitors' backlinks that helps you decide on your overall content marketing strategy. Use this to enhance the list of sites that might link to you. For example, if you find that your competitor had great success by releasing a study on trends in the market, you might be able to

create a market study that shows some different points that could also be successful. Consider, even, developing a contrasting study that piggybacks off of the competitor's study, but by rebuttal instead of repetition.

You can also expand on this concept by looking at "similar pages" to top-ranked sites (look for the "Similar" link in the preview pane of the search result for the site) in your keyword markets. Similar pages that keep showing up for different keywords are squarely in the topical link neighborhood. Look at who is linking to them too. You can also try a *related:domaintocheck.com* query to get some information on other domains worth investigating.

Review your website assets

Now that you have a refined list of targets and a sense of why each group may potentially link to you, review what you have on your site and what you could reasonably add to it. This should include any existing content, new content you could create, tools, or even special promotions (provided that these are truly unique enough and you have enough presence for people to notice and care).

It's important that the content be unique and differentiated. Content that can be found on 100 other sites is not going to attract many links. Even if the content is original, it should have something to offer or say that differentiates it from other content, rather than simply being a rewrite of someone else's article.

The highest-value potential linkers probably know their business and will recognize such simple rewrites, and in any event they'll want to focus their links on unique new content and tools. Content that leverages the publisher's unique expertise or what he stands for, and presents a new perspective or new data, will be far more successful in the link-building process.

Think of your content plan in a business case format. If you were able to create some new block of content at a cost of *x* dollars, and you think it would provide you with some set of links, how does that compare to the cost of the content (or tools or promotional effort) required to chase another link-building opportunity?

Ultimately, you will want to build a road map that provides you with a sense of what it would cost to chase a potential group of linkers and the value of each group of linkers. Your chart might look like Table 7-2.

Table 7-2. Prioritizing among link-building

among link-building projects

Cost to pursue	Value
$15,000	High
$30,000	Medium
$10,000	Low
$30,000	Very high
$4,000	Low

Once you have this in hand, you can quickly narrow down the list. You'll probably pursue the high-value campaign, and should continue to consider the very-high-value campaign and the low-value campaign that costs only $4,000 to pursue. The other two campaigns just don't seem to have comparable returns on investment.

Identify any strategic limitations

The next step is to outline any limitations you may need to place on the campaigns. For example, if you have a very conservative brand, you may not want to engage in social media campaigns through reddit (which is not a conservative audience).

Identify methods for contacting potential partners

You must undertake some activities to let potential partners know about your site. There are two major categories of contact methods: direct and indirect. Direct contact examples include:

- Email
- Social media sites (using the social media property's messaging features to make contact with potential partners)
- Blogger networking (building relationships by commenting on others' blogs)
- Phone calls
- Seeking out site owners at conferences or meetups

Some examples of indirect contact methods include:

- Social media campaigns (including Facebook, Twitter, LinkedIn, Google Plus, reddit, StumbleUpon, etc.)
- Public relations
- News feeds (through Yahoo! News and Google News)
- Speaking at conferences

Segmenting Your Audience, Identifying Personas, and Targeting Content

One of the most important concepts in content marketing is that you are producing content to gain exposure to your target audience. This is not about spewing links on random places across the Web, but getting visibility for yourself in places where you can reach interested users and potential customers. But who is your target audience?

One step in the process of identifying your target audience is grouping people based on characteristics such as demographics and typical behavioral patterns. *Segmentation* helps you better understand how to address your potential audiences.

The reason this understanding is so important is that you can't really expect to create engagement unless you target your content to the audience. Imagine writing an amazing piece of research on the Higgs boson particle and then publishing it on a site whose audience is first-grade schoolchildren. You probably will not generate much engagement there!

That's an extreme example, but this concept applies even at more basic levels. If your target customers are typically very thrifty, you should probably not target content to sites frequented primarily by the wealthy.

The concept of identifying *personas* is similar to segmentation, but it delves more into people's personal characteristics. Personas are fictional characters designed to represent a group of people with similar values relating to the use of a product or service. Personas include information on the potential user's motivation for using something, and the needs that drive it.

An example persona might be the “thrifty shopper,” and an example motivation might be that he gets excited by finding a good deal. Once you know this, you can target your content and on-site messaging to appeal to that motivation.

For content marketing purposes, if you offer fitness-related products that target athletic people from ages 18 to 25, you can map their demographics and find out what types of sites this segment frequents, and then consider those for your campaign.

Once you have developed personas of your potential customers, you can further qualify the target sites and customize the design of the content itself. For example, with the thrifty shopper, your headline and your lead image might both convey that you are revealing a new and unique way to save money.

Imagine you are running a business that focuses on fine dining. You offer high-end cookware as well as food products that can be delivered right to customers’ homes. Your business will have many different types of potential customer segments, such as:

Foodies

This group is made up of people who are passionate about eating exceptional food. As the saying goes, they don’t *eat to live*, they *live to eat*. However, you know that they don’t have a lot of time, and they are not gourmet cooks. For them, you might try an article idea such as “10 Gourmet Cooking Secrets the Pros Won’t Tell You.”

Home party planners

These are people who like to entertain. They throw parties on a regular basis, and they love to be praised for the quality of the food they provide. An example of a compelling article title for them could be “11 Things the Media Isn’t Telling You About Fine Dining and Health.”

Home-based gourmet cooks

These people are also foodies, but they take it further. Once a week they invest the time to produce an amazing meal at home, even if it is just for their family. They thoroughly enjoy the process of cooking and love the opportunity to be creative. One article that might grab their attention is “10

Myths About Gourmet Cooking and Health, Busted."

Professional gourmet chefs

This is your most hardcore audience. They are creating gourmet experiences as part of their job, and when they eat at home, they still want to have something of very high quality. They will have very high standards for their cookware, the way they organize their kitchen, and the food they buy. An example title targeted to this segment might be "4 Ways to Configure Your Home Kitchen for Healthy Gourmet Cooking."

There may be other segments as well, which you can develop over time by learning more about your target audience. Once the segments are finalized, you can begin to look for the sites that cater to them. This will require research, but one of the easiest tactics is to study the other content published on those sites and see which articles get the most engagement.

Measuring the engagement level is an important component of understanding how to get your content in front of people in a way where they will respond to it.

This process of identifying personas should be a part of every content marketing campaign you undertake (for a good outline of a persona mapping process, see "Step-by-Step Templates for Mapping your B2B Content"). It will provide you with the highest level of engagement with what you write. Make a point of measuring the engagement you get with the content you publish, whether it is on your site or as part of a guest posting campaign on third-party sites (more on this in "Guest Posting").

Putting It All Together

Sorting out where to start with your content management strategy can be difficult, but it is a very high-return activity. Don't just launch into the first campaign that comes to mind, as it can hurt your overall results if you spend six months chasing a mediocre content marketing plan instead of putting that effort into another, much better plan.

The goal remains building your visibility and reputation online, not just building links for SEO. As we will discuss more in Chapter 8, search engines are looking at engagement signals beyond links that can tell them where to find the best

content.

Even if they are not using these types of signals, the web community is becoming more sophisticated, and publishers who establish the best reputation and highest visibility will likely draw the most high-quality links to their sites.

Consider the famous video campaign by blender manufacturer Blendtec of blended iPhones, golf clubs, and so forth, available on YouTube and the company's Will It Blend? website.

The Will It Blend? site has more than 73,000 links to it from more than 6,000 linking domains (source: Majestic SEO), all entirely natural and earned through editorial recommendations. Not bad! Best of all, it supported a communication goal of showcasing the strength of the high-end blender.

The final step is to consider all these aspects together to come up with an integrated strategy. You can think of this as having the complete strategic picture in mind as you approach link building. At this point, you make the final decisions about your content strategy and which link-building targets to focus on.

Execute aggressively

A world-class content marketing campaign is always a large effort, involving a large potential cast of characters ranging from creative content writers, graphic designers, video production teams, script writers, actors, web developers, social media marketers, and more—the possibilities are endless.

Publishers who execute aggressively inevitably gain an edge over many of their competitors. Of course, if other competitors also focus heavily on link building, it becomes even more important to push hard, or you will end up losing search engine traffic to them.

For this reason, analyzing the backlink profiles of competitors to see how they're changing over time is a good idea. You can do this using any of the link analysis tools, such as Open Site Explorer, Majestic SEO, Ahrefs, or LinkResearchTools.

Consider pulling link data for your major competitors on a monthly basis, and seeing how their link profile is changing. This can help you learn if they are investing in content marketing, and if they are, what types of campaigns they are pursuing. Competitive intelligence like this can really help you tune your own

content marketing efforts.

Conduct regular strategic reviews

Content marketing strategies should be intertwined with the normal course of business, and evolve as the business does. As the implementation moves forward, you learn lessons and can feed this information back into the process. For example, you may have campaigns that did not work. What lessons can you draw from those?

Or, you may have one idea that is going gangbusters. How can you use that success to give you ideas for other campaigns that may work really well for you too? Sometimes the initial strategy goes great for a while, but it begins to run out of steam, so you should work to develop a constant stream of ideas that you are feeding into your content marketing process.

Applying what you learn from each campaign you try is a key capability to develop. No campaign is a failure if it helps you learn how to continuously improve what you are doing.

Create a content marketing culture

Publishers should also educate those within the organization about their content marketing plan, its goals, and how it will help the business, as well as identify the touch points for collaboration throughout the process. This will help engage the creativity of multiple team members in feeding the stream of content marketing ideas.

The more ideas you have, the better off you'll be—and the quality of a content marketing campaign is directly proportional to the quality of the ideas that are driving it.

Never stop

Content marketing is not something you do once, or once in a while. In today's culture, the search engine plays an increasingly large role in a business's well-being, and inbound links are a large determining factor in the fate of Internet sites. Don't be the business that implements a great content marketing campaign, gets to where it wants to be, and then stops.

Types of Content Marketing Campaigns

The following sections delve more deeply into different types of content marketing campaigns. Each has many complex nuances to it, and each can represent a substantial investment. For each strategy you contemplate, plan on investing some time to learn how to do it well. Chances are pretty good that you will make many mistakes along the way.

It's important to accept that reality and know that you will improve, and get better results, over time.

Guest Posting

We refer to an article as a *guest post* when it is published on a third-party site's blog or news feed. Guest posts are also referred to as *bylined articles*. There is a way to implement guest posting properly, but be aware that many SEO practitioners have abused the practice.

Success depends on creating great content, finding authoritative third-party sites and developing a trusted relationship with them, and then asking them if they would be interested in publishing your content.

Many authoritative sites accept bylined articles, as shown in Figure 7-28. But, as with all things SEO, there is a way to implement this strategy improperly. This led Google's Matt Cutts to write a post on his site declaring that guest blogging for SEO is dead, "so stick a fork in it":[16]

> *Okay, I'm calling it: if you're using guest blogging as a way to gain links in 2014, you should probably stop. Why? Because over time it's become a more and more spammy practice, and if you're doing a lot of guest blogging then you're hanging out with really bad company.*

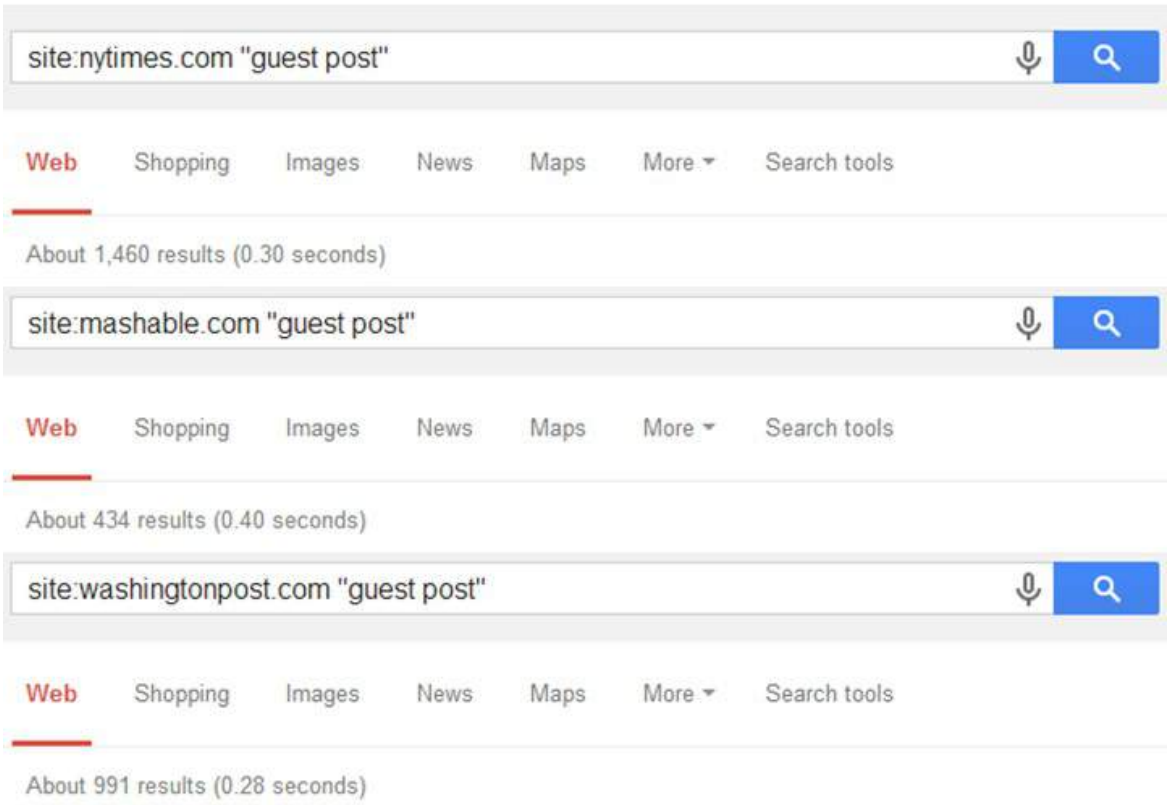


Figure 7-28. Many authoritative sites do accept guest posts

If you read his article in detail, you will see that he lists many of the "bad" things that people have historically done when guest posting for SEO. However, if you focus on building your visibility and reputation online, developing and publishing guest posts remains a valid practice. What follows are some guidelines for doing it properly.

Create high-quality, differentiated content

There is already a lot of content on the Web, and you can find some content on almost any conceivable topic. As an example, consider the results shown in Figure 7-29 for the search query *intitle:"mortgage tips"*.

Figure 7-29. Almost 7,000 pages focus on mortgage tips as a topic

This query requires the exact phrase "mortgage tips" to appear within the `<title>` tag of a web page to be included in the results, and Google reports that there are 6,850 such pages on the Web. Writing the 6,851st article on the topic is not likely to result in a lot of people deciding to link to you.

Creating high-quality, differentiated content can present quite a challenge, and one key aspect of success is that you either need to be an expert yourself, or employ/retain an expert to help you. Hiring a writer who knows nothing about your topic and asking him to write an article for you in two hours will not produce content that draws high-quality links. Only an expert can truly recognize what aspects of a topic are unique, and what type of content would bring special value to your target audience.

Being an expert or hiring one may take a lot of time or money, but presumably you chose to be in the business you are in for a reason. Hopefully, your passion for your work, or certain aspects of how you approach it, will naturally enable you to provide a unique perspective that others can benefit from.

Aim high

You want to get published on authoritative sites. As shown earlier in Figure 7-28, many sites accept bylined submissions, including some of the most well-known sites on the Web. While PageRank as it was originally defined is not what the search engines use to determine the value of the link, it still teaches us an important lesson: a link from a high-quality site can be worth 1,000, 10,000, or even 100,000 times more than a poor-quality site.

In addition, links from poor-quality sites can even be harmful to you. We will discuss this more in Chapter 9. When you first start out, you may not be able to

convince the higher-authority sites in your niche to accept bylined articles from you, but getting on those sites should still be your objective—you may just need to do it in stages. You can think of this as climbing a ladder (see Figure 7-30).

Figure 7-30. Aim for placements on high-authority sites

The concept is to start at the highest level you can. Based on your current visibility and reputation, you may need to start with Tier 3 sites, post there for a while, establish some credibility, and then start trying to get published on Tier 2 sites, as you can now point to the Tier 3 sites where you have already been published. This process may take many months.

Continue publishing on Tier 2 sites for some time, perhaps six months, and then consider trying to reach out to the Tier 1 sites. Be patient throughout this entire process: it takes time. When you first try to move up to the next tier, you may not be accepted.

All of this requires that you are publishing high-quality articles throughout the entire process and are measuring your engagement to make sure you are

performing well with the content you are publishing. Developing that track record is a key part of success.

During this process, make sure you never agree to post on a poor-quality site. Your reputation is not enhanced by such sites, and it is best to stay away from them.

If you run a small local business, your definition of a Tier 1 site may be quite different from that of a major national brand. Getting published in the *New York Times* may not make any sense at all for you. The Tier 1 site might be the local newspaper, and your ultimate goal might be to get a column there.

For example, Orion Talmay started by writing posts on her own blog, which you can think of as a Tier 3 site. After posting there for a while, she was able to start posting articles on Intent.com's blog, which is an example of a Tier 2 site. Finally, she became a contributor to the *Huffington Post*—a clear Tier 1 site.

Remember that quality trumps quantity

Many people doing SEO used to believe that the best way to get value from guest posting campaigns was to get on as many possible domains as possible. The reason for this belief was that Google used to view each incremental link from one domain as declining in value, as multiple links from the same domain could well represent a single editorial decision—effectively only one "vote" for your site.

This led publishers to research large lists of potential target sites for guest posts, and work on getting one or two articles published on each. Of course, over time, even if you start by aiming high, this approach leads to the quality of the target sites declining over time, and the new links you obtain start coming from lower and lower quality sites.

Let's take a moment to define "quality over quantity" here. If you were to tell your friend that you had just published an article in the *New York Times*, she would probably be impressed and excited for you. Consider her reaction, however, if you told her you had just landed a *column* in the *New York Times*. Clearly the latter scenario is a lot more impressive. The search engines know this too, and they will treat those ongoing posts on the top authority sites as fresh editorial votes for you. This focus on quality over quantity also brings the most benefit to your visibility and reputation, and this is your top priority in any

content marketing campaign.

Perform research and analysis to pick potential targets

We have already established that you need to focus on producing great content, targeting authoritative sites, and emphasizing quality over quantity, but how do you go about identifying sites that might be willing to publish your content where users can engage with it? The first step involves no tools at all. Simply use your own knowledge, that of others in your organization, and that you gain by speaking with your peers, to identify the most important sites in your market space. Chances are that you will find a large percentage of the best targets using this process alone.

You can also analyze where your competitors are getting links. For example, you can make use of competitive backlink analysis to find out who links to your competition. This can provide not only specific sites to target, but also ideas for much broader campaigns to put together.

Competitive backlink analysis can help you determine where your competition is writing bylined articles. This tactic for finding link targets and content marketing campaign ideas is discussed in more detail in “Find out where your competitors get links”.

You can also look for potential targets by using a variety of sophisticated search queries. For example, you might try this query if you are in the business of selling “left-handed golf clubs” (see Figure 7-31).

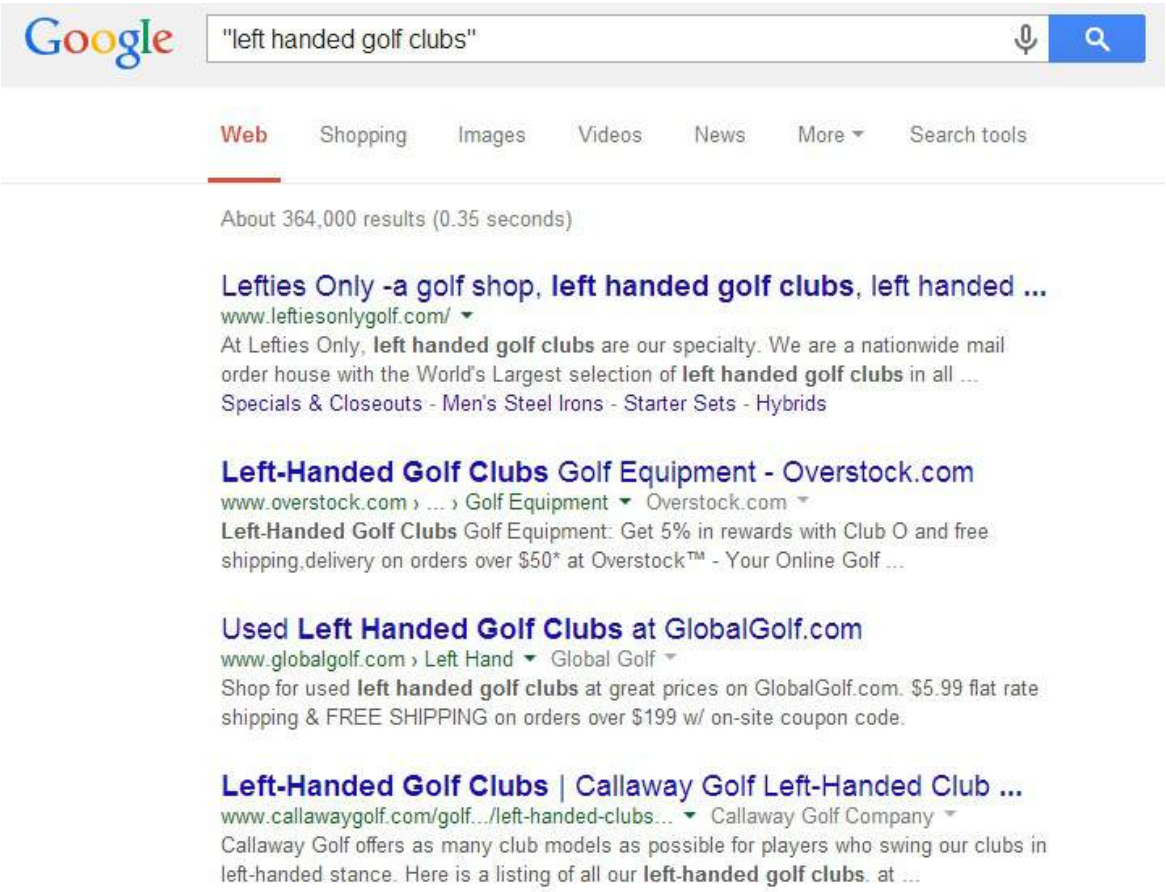


Figure 7-31. Search results for a query on "left-handed golf clubs"

As you can see, this query returns over 300,000 results! While many of these sites would not be appropriate targets, it is likely that a number of them are. Remember to filter the potential targets you find from competitive backlink analysis and advanced search queries based on the criteria we have listed in this section.

Content Syndication

Content syndication can also be a very effective technique for getting your content in front of a wider audience, but it does come with some additional risks. Search engines do not like duplicate content on the Web. While they do not punish duplicate content unless it is spammy or keyword stuffing, they usually filter it out, which means that when they see more than one copy of a piece of content, they will generally show only one page in the search results and ignore

the rest.

There are a variety of techniques available for mitigating the risks involved in publishing duplicate but legitimate content. In priority order, these are:

1. Ask the third party who is republishing your content to implement a `rel="canonical"` link element that points back to the URL of the content page on your site (not your site's home page, but the page where your original article is published). The `rel="canonical"` link element tells the search engines that the master copy of the content is located at the URL of the page on your site.

2. If the third party is not willing to do that, ask that he implement a `noindex` tag on his copy of the content. This tells the search engines to omit the syndicated copy of the content in their index, which effectively eliminates the duplicate content problem because your copy of the article will be the only one in the index as long as the `noindex` tag is respected.

3. If the third party isn't willing to implement a `rel="canonical"` link element or a `noindex` tag, ask him to link from his copy of the article to the article page on your site—again, not to your home page or any other page of your site, but the page on which the original content appears. This is not as effective as the previous two techniques, but it can still be a reasonable option.

The issues surrounding duplicate content, the `rel="canonical"` link element, and the `noindex` tag are all discussed in detail in Chapter 6.

There are also times when you might consider allowing someone to syndicate your content even if she won't implement any of these tags or links. Remember, your primary focus is on visibility and reputation building. If the *Wall Street Journal* wants to republish your article, and you are in the early stages of building your brand online, just shriek "yes!" The reputation and visibility benefits will far outweigh the downside of having the *Wall Street Journal* rank for your content instead of your site.

One variant of content syndication is to generate articles and then submit them to *article directories*—but this use is not approved by Google, and in fact Google

targeted many such directories for punishment in its Penguin algorithm update on April 24, 2012. The message from Google here is clear: do not use article directories as part of your link-building strategy!

It is a best practice to focus on high-value targets when syndicating content. Sites that will take any article with little editorial review are not likely to offer high-quality links. In fact, the links may be of no value at all. Getting into higher-value sites, such as a major regional newspaper, may require more effort, but such sites also provide higher-value links and may result in other sites linking to you as well.

Link-Worthy or Viral Content

In the past, the term *link bait* was used in the SEO industry to describe content created specifically for the purpose of attracting links. That term has fallen out of favor because it connotes manipulation, so these days people talk about creating *link-worthy* content, *link attraction*, *shareable* content, or *viral* content. This content is usually published on your own site but sometimes on another website, and it is compelling enough that lots of people link to it. Such content can take many forms. For example, it might be content that is designed to provide enough additional value that people will want to reference it.

As we discussed in "Brainstorming Content Ideas and Being Creative", there are many ways to come up with content ideas. Not every piece of content should be expected to go viral, nor is that a realistic goal. A content marketing plan should include many types of content, some of which is designed to meet more day-to-day needs, and some that has the potential to have a much bigger impact.

Popular methods for coming up with link-worthy content ideas include doing something controversial, something funny, or something that simply draws a strong emotional reaction. Each market space has some hot buttons, and these buttons can be pushed with an opinionated article, a compelling image, or a great video.

How far should you go with your ideas?

Link-worthy content can take many forms: top 10 lists, humorous videos uploaded to YouTube, checklists, cartoons, how-tos, event coverage, exposés, personality tests, quizzes, contests, surveys, tools, and widgets, to name a few.

For example, BuildDirect struck linking gold with its top 10 post “Redneck Home Remodels,” created by content strategist Rob Woods, which featured photos of humorously ramshackle home improvement attempts. The right mixture of controversy and humor, along with a strategic promotion from a power user on StumbleUpon, sent several hundred thousand views to the BuildDirect.com site, nearly 20,000 likes to its business page, and hundreds of tweets.

With a contest, the devil’s in the details. You must get everything right: the prizes, judges, judging criteria, media partners, and so on. Most contests fall flat; they are simply unremarkable. Something about the contest must be worthy of being written about.

For example, in the Free Business Cards for Life contest that Netconcepts dreamed up for its client OvernightPrints.com, the site partnered with “Technorati 100” blogger Jeremy Schoemaker (a.k.a. “Shoemoney”). The contest was to design Jeremy’s business card; the winner received “free business cards for life.” The cost of the prize (and the contest overall, for that matter) was negligible: the fine print capped the winnings at 1,000 business cards per year for a maximum of 20 years.

The link exposure this contest garnered was excellent; it included keyword-rich links from Jeremy’s blog, shoemoney.com, and Jeremy also posted a video about it to YouTube. A number of design sites and bloggers listed or linked to the contest. The ultimate goal was the rankings that resulted from the links. As you can see in Figure 7-32, Overnight Prints still ranks at #3 in Google for “business cards.” Jeremy got something out of the deal as well: a killer new business card printed and shipped to his door courtesy of Overnight Prints.

Vistaprint - **Business Cards**
www.vistaprint.com/category/**business-cards**.aspx ▾ Vistaprint ▾
Design this powerful marketing tool exactly how you want it. Add spot gloss or raised text to your **business cards** for a striking look. Make yours today!
Business Cards - Personal Business Cards - Networking Cards - Appointment Cards

Custom **Business Cards**, Design Your Own ... - Vistaprint
www.vistaprint.com › Home › Business Cards ▾ Vistaprint ▾
★★★★½ Rating: 4.5 - 1,419 reviews
Create custom business cards to build your business. Choose ...

MOO **Business Cards** | Custom **business card** printing online
us.moo.com/products/**business-cards**.html ▾
Business Cards should be a conversation starter – and with premium paper stock, the option to print a unique image on every card, and your choice of tactile ...

Business Cards - Choose from Thousands of Designs - Zazzle
www.zazzle.com/**businesscards** ▾
24 hour shipping on most **business card** orders. Start from scratch or use one of the thousands of free **business card** templates.
Custom Business Cards - Professional - Elegant - Modern Business Cards

10 **Tips** For Effective **Business Cards** | OvernightPrints.com
www.overnightprints.com/ten-tips-for-effective-**business-cards** ▾
With so many **business cards** in the arena, it's important to have one that will stand out among the rest.
You recently searched for tips.

Figure 7-32. Google organic search results for "business cards"

Potential linkers also love a good corporate citizen, so be one. Consider such activities not as an expense, but as an investment that will generate a return in the form of links.

Noomii.com created buzz with its "Daily Acts of Kindness" Advent calendar. Participants could subscribe with their email address to receive daily emails with an idea for a random act of kindness. With this amazingly simple idea and the catchy tagline "...because it's better to give than to receive," Noomii amassed views and high-profile mentions.

Do not be afraid to be bold or off the wall. You do not always have to toe the corporate line. Relevance is important, but obtaining some lower-relevance links is OK too, as long as the majority of the links to your site still come from highly

relevant sources.

Encourage link-worthy content to spread virally

You can extend this approach to content distribution by creating something you can pass around. For example, a hilarious video clip might be passed around via email. Provided you make it easy for people to determine the video creator (presumably your company) and to visit your site, this type of campaign can garner a lot of links. Be aware, though, that if you host the video on a video-sharing site such as YouTube, most people will link there, not to your site. This isn't necessarily a bad thing, however, if you have a larger goal in mind, such as generating buzz and building your brand presence.

User-Generated Content

Providing users with ways to contribute content directly to your site can be an effective content management strategy. There are many ways to do this:

Open up a forum or allow comments on your site

One of the biggest challenges with this option is achieving critical mass so that active discussions are taking place on the site. This typically requires a substantial amount of traffic to accomplish, but in the right situations it can help you develop interesting content with little effort. You can also try to implement programs to stimulate the discussion, such as offering a prize on a monthly basis for a randomly selected contributor. If these discussions become quite substantial, then they may themselves attract links from other sites.

Launch a blog and invite third-party contributors

One of the best ways to do this is to contact respected members of your market space and see whether they would be willing to make written contributions (i.e., provide guest posts) to your blog. They may do this simply because they would like the exposure, or you can pay them for doing it.

When you do this, make sure you are getting very high-quality content from them. A simple way to check this is to see if they promote the content they

provided in their own social media accounts. If they don't, it may be a clue that they are not that proud of what they provided to you.

More selectively invite third-party contributions

Launching a blog platform may be more than you want to do, but you can still ask respected members of your community to consider contributing articles to your site. Run an "Ask" campaign. Ask curators/get visitors to curate content (e.g., BitCandy.com, which is a crowdsourced music discovery site).

Of course, the contributed content does not need to be an article or a post. You can seek out photos, videos, cool new tools—anything that may be of interest to users.

With each of these strategies, one of the big questions is whether the method for contributing content is open, strictly controlled, or somewhere in between. In other words, can any user come along and post a comment in your forum? Or do all users have to have an editorial review first? Editorial reviews may significantly reduce spam attacks, but they can be a barrier to the development of active discussions.

In the case of forums, engaging discussions can attract links. For an example from the world of SEO, Search Engine Roundtable is a frequent linker to discussions in the WebmasterWorld Forums. The key to this is the critical mass of discussions that take place on these forums.

The reason these tactics involving third-party authorship can result in links is that most people have pride in what they have created and want to show it off. As a result, they tend to link to their content from other sites where they contribute, or their own website. As mentioned earlier, if a guest author is proud of what he wrote, he will most likely promote it via his social media presences. If he does not, then you probably don't want him writing for you, because it might indicate that he's not proud of the content he's given you. Exceptions to this rule are subject matter experts or other professionals who might not directly engage in social media for professional reasons.

Building an Audience

You can create the world's greatest content, but if no one ever sees it, you will not have accomplished anything useful. You have to get exposure for it, and one way to do that is to build a loyal following of people who learn to love your content. This is a time-consuming process, but one of the fundamentals of content marketing.

Get to Know Other People's Audiences

When you first begin publishing content on your site, unless you are fortunate enough to have a large brand or already be famous, chances are that you don't have a large audience ready to read your content. The process of building one can be long and arduous. One way to speed up that process is to expose your content to other people's audiences (OPA). Here are some examples of how to do that:

Publishing guest posts on very high-quality websites in your niche

If you are able to create great content and get it published at these sites, the audience there will have an opportunity to see what you have to offer. If you continue to do that on a repeat basis, over time some portions of this audience will start to become your audience. They may start to follow you on social media, or look for more great content on your site.

Interviewing industry thought leaders

You probably already know who some of the thought leaders in your market space are, and if not you should find out. Try developing relationships with them first, perhaps by interacting on social media or by commenting on their posts, and once they get to know you a little better, ask them if you can interview them. Then publish the interview on your site, or as a guest post. Chances are that they will share that interview in their social media.

Actively engaging in social media

This is a fantastic source of OPA. Interacting with others, sharing others' content, and participating in communities are all great ways to generate attention and build an audience. Figure 7-33 shows how this works.

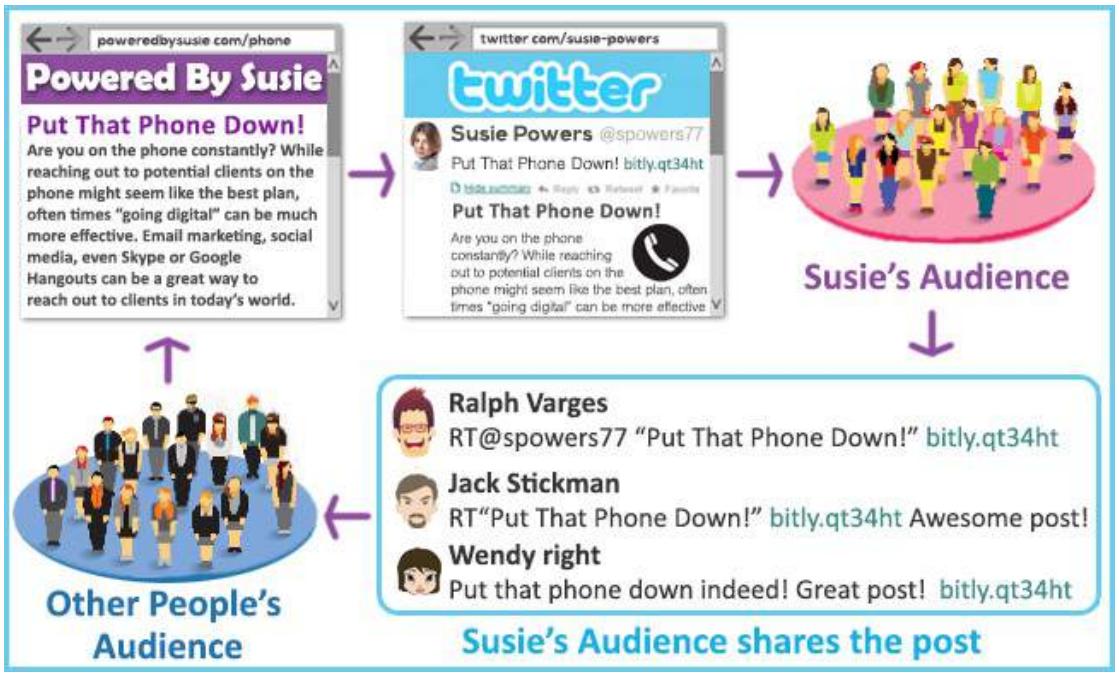


Figure 7-33. Social media can help you build your audience

Speaking at conferences

This requires that you succeed in getting a conference to accept you as a speaker, but it's a good source of exposure. Not only does it get you in front of an audience in person, but you can also start letting potential customers know that you have speaking experience, which is a great credibility builder.

Getting interviewed

Interviewing people is great, but so is getting interviewed. Even if you may not have much reputation yet, you probably have some unique perspectives related to your business. Being willing to share that information with people can be quite effective. If you are a small-business owner, offer to make yourself available to a local newspaper or blog.

Reaching out to the media/bloggers

Have a great story to share? Reach out to media that would be interested in that story. If you can get them to write about it, that's awesome. This strategy depends on your having some major story that will be of interest to a certain

audience.

Issuing press releases

This is actually a form of media outreach. Take care with this tactic, because you should do it only to share something truly newsworthy. But, if you do, it can get you in front of major media people. Figure 7-34 illustrates how press releases can work for you.

Figure 7-34. Press releases can help you build your audience

Advertising

This is a very straightforward way to try to accelerate building your audience. For example, promoted posts on Facebook, promoted tweets, and Google's *+Post* ads are all great ways to promote content and get in front of others. This can bring you rapid exposure and help shorten the overall process.

Paying for content network links

Examples of content networks are platforms such as Outbrain and Taboola. They offer recommended content suggestions on major media sites, such as CNN, Slate, ESPN, USA Today, the Weather Channel, Fox Sports, and Daily Mail. While the links in these placements probably have no direct SEO value, they can offer a valuable way to get in front of new audiences.

Paying for social media placements or boosts

Paying for placements or to boost posts can be a very cost-effective way to get more exposure for your content. Paid options are available on Facebook, Twitter, LinkedIn, and Pinterest. This topic is discussed more in "Paid social media".

Criteria for an OPA strategy

Chances are that you are not going to pursue all of the aforementioned OPA ideas at once, so you have to make some choices. Here are some criteria for narrowing down your OPA strategy:

Size

This is the first thing that everyone thinks about, and of course getting in front of a large audience is, in general, a good idea, but only if you satisfy one or more of the other criteria that follow.

Relevance

It's critical that you build an audience that is relevant to what you do, so the OPA you seek out should be interested in topics that are relevant to your business. If you are doing a guest post, in the ideal world, the site you are posting on is focused on the same general market as you are, or a site that covers the market you are in. Figure 7-35 illustrates the concept that relevance is an important factor not only in SEO value, but also in getting you more exposure to potential audiences. That's a good thing!

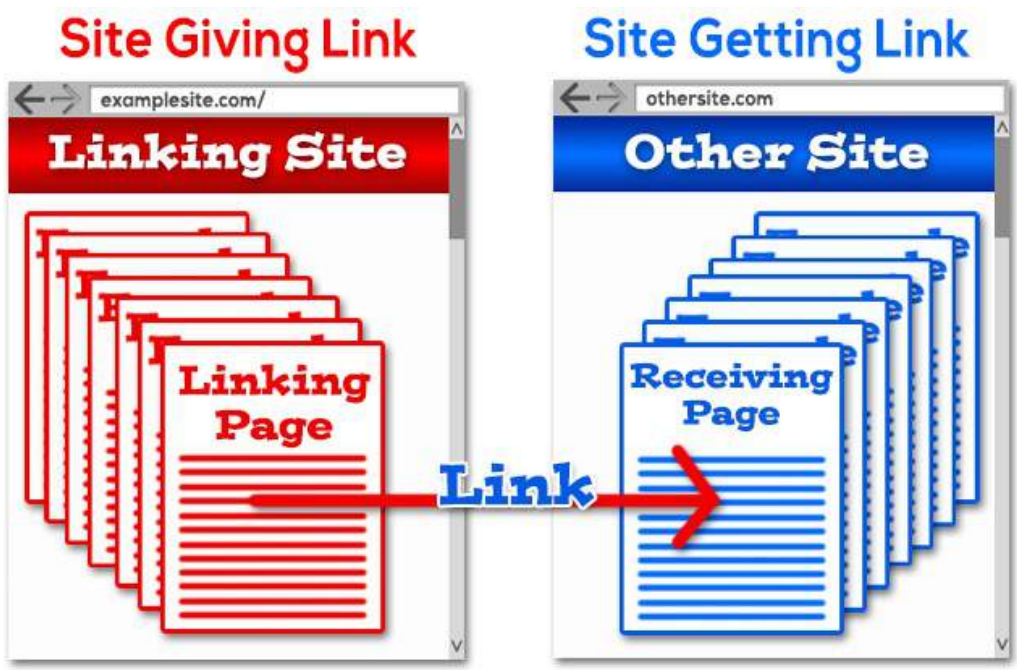


Figure 7-35. The relevance of inbound links is important!

Quality of existing content

The company you keep defines you. If the articles on a site are poor, you don't want yours there. Or, if the contributions to a social media community are all self-promotional, there is no value in being there.

Ease of getting started

This factor is actually one of the most important. You can't do everything at once, and there are probably some ideas that will be easier for you for execute. Pick one or more of those ideas and generate some personal momentum.

Authority

The authority of the venue makes a difference. Speaking at the most important conference in your industry, interacting on social media with people already recognized as experts in your area, and writing a guest post for the top news site that covers your space are all great things. Some of that authority transfers to you by association. You may need to work your way up

over time, but you want to be conscious of a venue's authority as you are presented with opportunities.

Quality of content you can produce

Whether you are presenting, writing, or promoting, it is worthwhile only if you are able to deliver good content to that audience. You are not a fit for every audience. Don't worry about it. Pass on opportunities where you can't deliver your best, and focus your energies in the places where you can.

Opportunism

The most important part of building your audience is to get out there. Be a part of the community in your industry. This sets you up to recognize big opportunities when they arise. Watch for these and be ready to respond quickly, because leveraging opportunities is a big time accelerator.

Your business needs to have a strategy for building your own audience. Exposure to OPA is a great approach, because conjuring an audience out of thin air does not work. You need to map out a strategy and pursue it in a purposeful manner. You may need to do some experimentation to find what works for you, but it is well worth the time and effort.

The indirect benefits of building reputation and visibility

In the days predating the Internet, classic marketers used to say it requires seven impressions or touches to make a sale. If you are trying to build your reputation to the point where someone might link to you, you need to be prepared to mount a campaign, and not expect everything to happen just because you reach out to a stranger and ask for it.

Even in the case of guest posting, which is a methodology that provides a relatively direct way (in some cases) to get a link, the real benefit—getting exposure to other people's audiences—is likely more indirect, as Figure 7-36 illustrates.

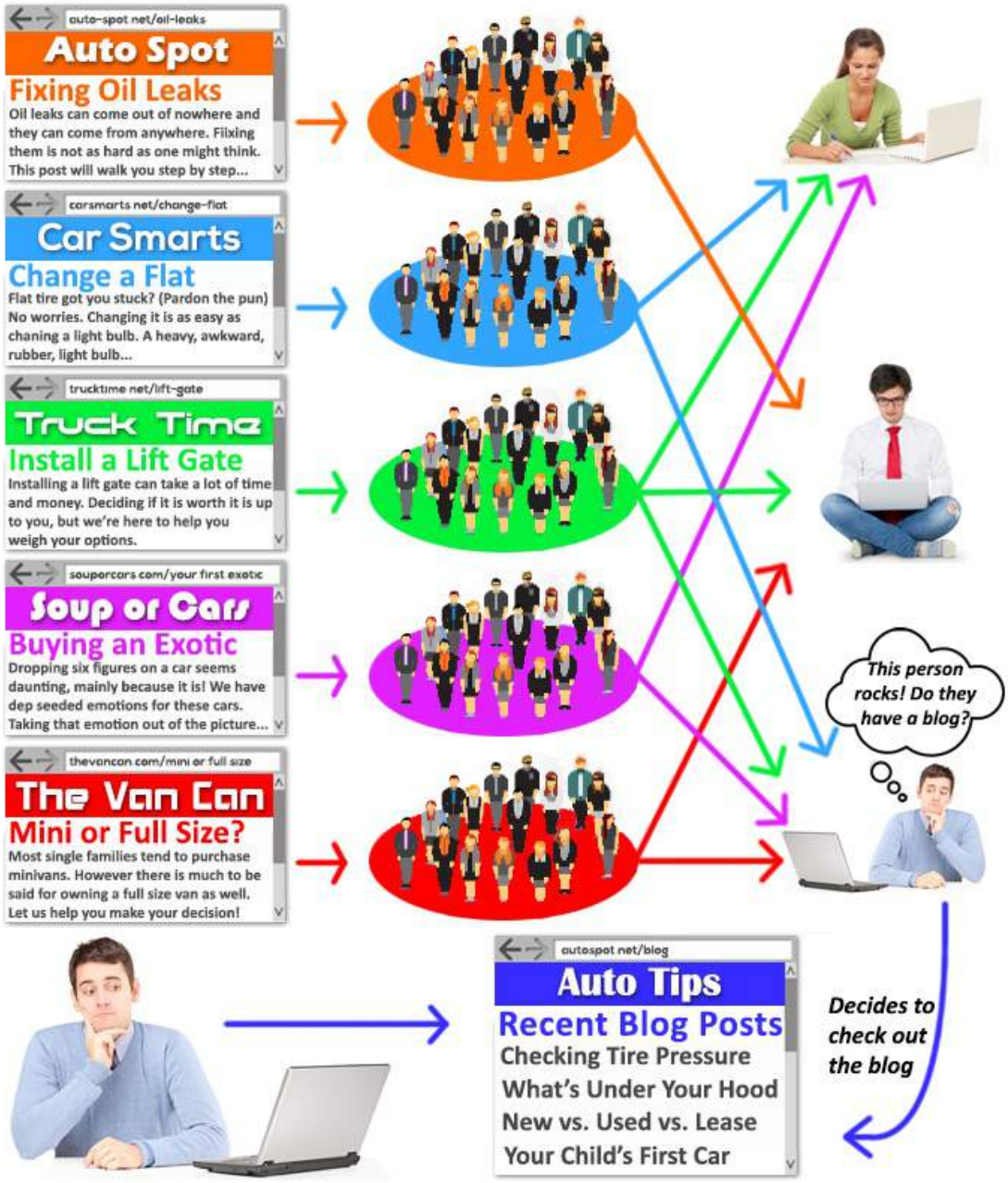


Figure 7-36. The indirect benefits of guest posting

On the left side of the diagram is a publisher who writes high-quality articles on several different high-quality sites. Each posting provides him with exposure to OPA (shown in the middle), and some of the people who see his articles are media people or blog owners.

Over time, they may see many of his articles, and eventually they decide to go

check out his site. If the content there is of good enough quality, they might start linking to some of the articles there, or sharing them in their social media.

You can see how this process of building a reputation works. This same concept applies to participating in social media, speaking at conferences, sharing videos on YouTube, and engaging in many other types of activities.

By being active in communities related to your market space, you build your reputation and visibility, and over time, some of the OPA take an interest in your site and actively engage with your content. While this can be a long process, it is a very powerful way to look at your overall content marketing strategy.

Leverage Influencers and Influencer Marketing

The term *influencer* refers to someone who has the ability to influence a large number of people. This may be because she has a large social media following, she has a very popular blog or broadcast show, or she's very well known for other reasons.

PR firms have understood the power of influencers for a long time, which is why you see so many celebrity endorsements for products in TV commercials. Associating a brand with a highly popular personality has been a good strategy since long before the Internet came to be.

Online, the process is a bit different, in that an influencer's endorsement works best when it has not been paid for, or at least appears to not have been paid for. This endorsement may take the form of the influencer sharing a link to your content, linking to it directly, letting you write a guest post for his site, liking or +1-ing your content, or otherwise helping make people aware of your site.

Influencer marketing is the process of developing relationships with influential people that can lead to their assisting you in creating visibility for your product or service. This type of marketing depends on your producing great noncommercial content that would be of interest to the influencer's audience, but that's the second step. The first step is to build a great relationship with the influencer.

As shown in Figure 7-37, influencers often have a larger audience than yours, or at the very least, a different audience.

Figure 7-37. Influencers often have larger reach

However, the benefit is much larger than that. Let's say you have 100 followers in your Twitter account who shared a piece of content, and this results in 20,000 people seeing what they shared. This may result in 20 additional shares and 10 links to your site.

Now consider an audience of the same size being reached by one influencer. Those 20,000 connections will be much more responsive to the shared content because of the trust they have in the influencer's opinions, and as shown in Figure 7-38 this might result in 100 additional shares and 50 links to her site.

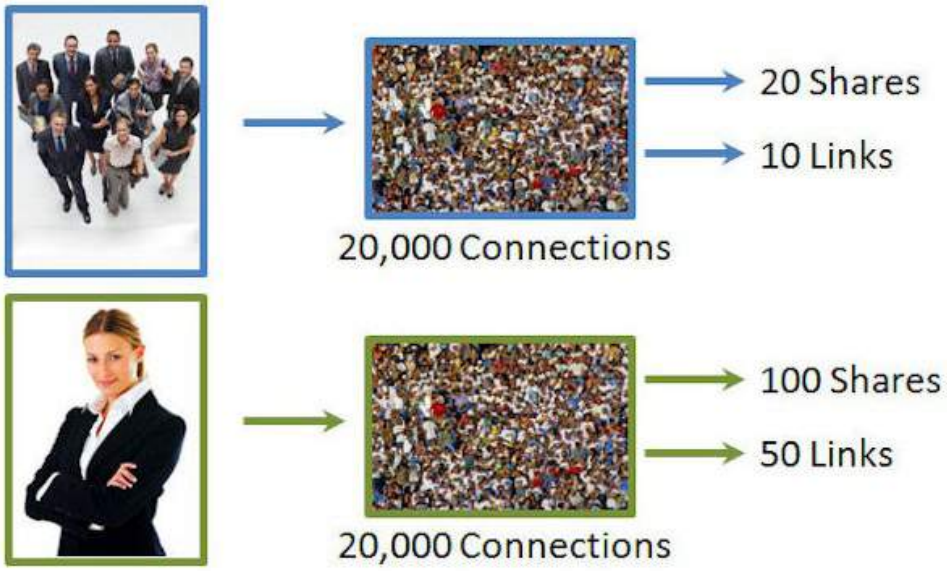


Figure 7-38. Influencers get higher conversion rates from their audience

Engaging the influencer

If you are a fan of content marketing, chances are that you already have your own blog and your own social media accounts. You probably already use these in tandem, and make sure that you follow similar content themes and share any new blog posts you write on your social accounts. When you do this correctly, you set yourself up for the type of virtuous cycle shown earlier in Figure 7-16.

Doing this effectively is a great start. You can grow your audience over time because people who are already connected with you will share your content, and it reaches their audiences.

However, this basic process works much more effectively when you get influencers involved. Here are a few ways to do that:

- Develop relationships with major influencers so they are subscribing to your blog or following/friending/circling you in social media accounts. You must have a history of creating content of interest to the influencer. The payoff occurs when he chooses to link to it or share it on a social network.

- Get permission from the influencer to provide him with a guest post and be published directly in front of his audiences—for example, in his blog. Like the preceding tip, this strategy also depends on having a credible history with

the influencer so he will consider your article. The payoff here is quite direct, and happens as soon as the content publishes.

- Interviewing him and publishing the result on your site. This is a great tactic, as the influencer is likely to share the resulting interview via social media and perhaps via links as well.

There are many other variants of these ideas, but all of them depend on having a relationship of trust with the influencer.

Building the relationship

Building the influencer relationship is not really so different than making friends when you move to a new neighborhood. When you go to that first neighborhood party, you don't walk around asking everyone there to give you $20. If you did, you'd quickly ruin your reputation in the neighborhood. This doesn't work with your new neighbors, and it doesn't work in building relationships anywhere else either.

Figure 7-39 illustrates the process of building a trusted relationship.

Figure 7-39. Building trust in a relationship

The major elements of the process are as follows:

- Start interacting with the influencer. Again, treat it like you are developing a new friendship. When it comes to business, focus on providing value to her. If she has a question, seek to answer it. Don’t spend any time telling her what value you bring—just deliver it to her.

- On an ongoing basis, show that you will be active in sharing her content to your audience. Even if your audience is much smaller, the give-and-take attitude will be noticed.

- Actively help out others. When you focus all of your attention on one person to the exclusion of others, it starts to feel a bit creepy. Give value to others on a regular basis. Publish great content, and share other people’s. If you discover great content from a little-known author, the influencer you are trying to build your relationship with will be more interested than ever!

Prioritize these efforts. How do you decide? You might fly to a conference to meet some critical person face-to-face, while with others you might simply interact on social media accounts. Figure 7-40 illustrates this concept of putting more effort into more important relationships.

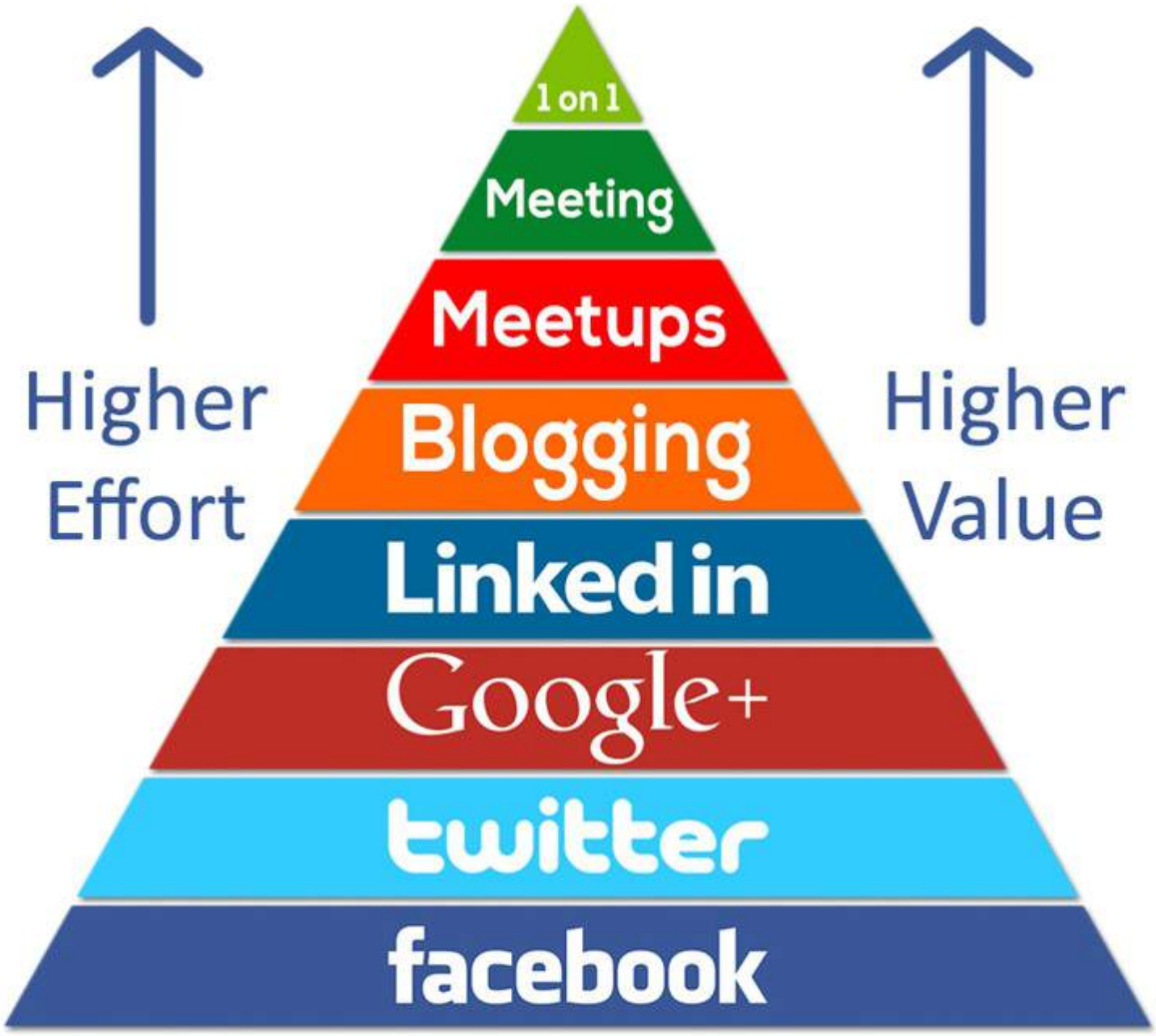


Figure 7-40. The relationship building pyramid

Note that while it's certainly possible to build meaningful relationships with people through social media only, nothing beats face-to-face.

Recognizing opportunities is also important. Your first chance to make a big impression on someone might be to respond to a blog post, a tweet, or a Facebook update. Better still, if your target asks for help with something in a public way online, make a point of taking advantage of that opportunity by being the first one to give it to him.

Once you have developed a relationship, you still need to do the right things to get someone to share or link to your content. No one is going to share everything you do, because not all of it is that good or relevant (don't be offended: no one is

great all the time). Figure 7-41 shows the factors that impact someone's decision to reshare your content.

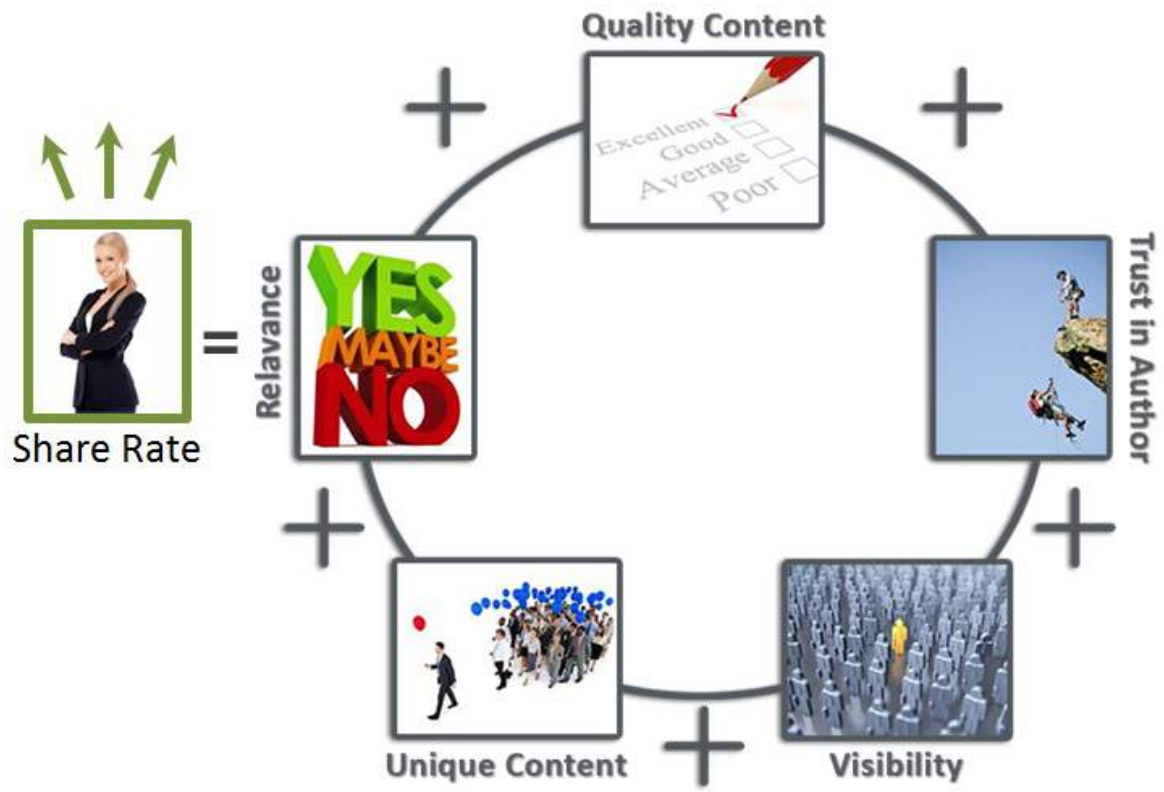


Figure 7-41. Factors influencing the likelihood a person will share your content

The major elements that go into that decision are as follows:

Relevance

If your content is not relevant to someone, she's not going to share it, even if it's great!

Uniqueness

People are not going to share content unless it's unique in some way that is of interest to them. Achieving uniqueness may be hard to do, but it's a requirement, so a key part of your content marketing plan is to figure out what you can do that is differentiated from what's already out there.

Authority

It definitely helps to have established some level of authority. Of course, you won't have that starting out, so you will have to leverage the other factors more to make up for it.

Quality

This goes without saying. Poor content will bring poor results, and no amount of relationship building will change that.

Trust in the author

This is where the relationship comes into play. You can create great content, but if you are not yet trusted, your share rate will be far lower.

Trust in the referring sources

How someone learns about a piece of content is a factor in the share rate as well. If an authority tells you about it, you are more likely to respond by passing it on.

Visibility

People can't share what they don't see. For example, if you create a great blog post and you tweet it once, only a small percentage of your followers will ever see it. Tweets are here now and gone five minutes later. Even the biggest tweetaholics miss some of their tweet stream.

Impressions

This is classic marketing in action. As noted earlier, traditional marketing experts used to say that it took an average of seven impressions per sale. This general principle still works today. As shown in Figure 7-42, if someone sees your content referenced by more than one other person, he'll be more motivated to see what it's about and more likely to share it.

Exclusivity

You can also get someone's attention by offering her exclusive access to your content, or by offering her an early preview. People who publish blogs, or

content via social media, love to be able to break news.

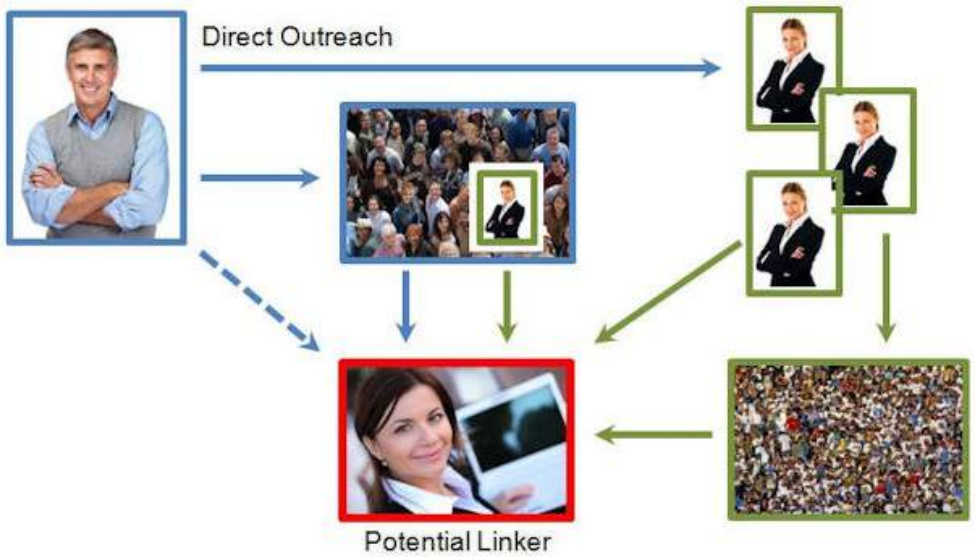


Figure 7-42. Creating multiple impressions increases the chances of getting shares

As you can see, the reach of influencers is long. Not only can they get you links, they can also give you shares that result in other people giving you links.

However, realize that every person out there has some level of influence. You can't build deep relationships with everyone online, but you can be courteous to them all, and you can seek to help out others in ways that are appropriate. You never know, that person with 132 followers on Twitter might happen to know a major influencer who is really important to you. You should seek to leave a good impression everywhere you go.

Get Active in Social Media

Social media can be a great way to build your audience. There are large quantities of people active on Facebook, Twitter, Google+, Pinterest, Instagram, and more, so actively engaging on these platforms can provide you with great exposure. Try interacting with others, sharing others' content, and participating in communities, as these are all great ways to generate attention and build an audience.

Remember, though, that your objective is to gain an audience for yourself, and

you don’t own the social media platform. It can choose to reduce organic visibility at any time, and this means your influence there can be curtailed for arbitrary reasons, even if you have a large number of followers. Many people discovered this as Facebook dramatically reduced organic visibility on its platform in response to the revenue pressures of being a public company.

We will discuss the various social media platforms more in “Social Networking for Links”, and in Chapter 8.

Paid social media

Once your team creates and organically distributes content that is intuitive and engaging and serves well-conceived audiences, you have already made great progress. Distributing that content to organic audiences who love it and share it is also good. The problem is that, while there is some free social distribution, on many social networks the amount of available distribution is declining. Paid social media can help you significantly extend the reach of your content.

Potential customers can’t be motivated to read a fancy blog post if they never learn about it. Bloggers and journalists won’t link out if they never discover your infographic, blog post, Google+ Hangout, and YouTube video. High-authority social users can’t share your bit.ly link if they’re unaware of the existence of your carefully constructed, beautifully designed web app. The crowdsourced industry study in which you invested so much will have a limited effect without at least some distribution.

Social distribution

Sadly for marketers, free social media distribution has already been greatly reduced on sites like Facebook (as few as 1%–2% of followers of a brand page will see the content you post there). Why is free social content distribution now limited? To answer that, you need only take a quick look at Facebook and Twitter’s stock value and earning trends. As a result of the pressures of being public companies, these social media channels now force marketers to pay for the same organic lift that used to be free.

The good news is that it’s relatively inexpensive to regain that distribution through paid social media campaigns. Think of social pay-per-click (PPC) campaigns as editorial, calendar-based, social content amplification programs

measured against content marketing and conversion key performance indicators (KPIs), not just direct-response advertising tools.

Social content amplification steps

The social content amplification process is easy to understand and works really well in Facebook, Twitter, LinkedIn, Google+, and YouTube. Start by posting great content on your social profile's company wall.

The next step is to create an ad. These are not just any ads, because they look almost exactly the same as an organic post. On Facebook, these are called *page post ads* and they're served to targeted users in either newsfeeds (desktop/tablet/mobile) or the righthand column (desktop/tablet). Twitter calls them *promoted tweets*, and LinkedIn calls them *promoted posts*. Google+ amplification ads are called *+Post ads*, and they are created via the AdWords Display network. YouTube calls the ads *TrueView*, and they result in targeted video views, subscribers, and external clicks if you buy *call-to-action* (CTA) overlay ads for the videos. Amplified social posts result in traffic to your external website and/or internal social traffic. The internal social traffic often results in new followers and extra sharing.

The determining factor for whether a social amplification ad drives users external to your website or within the same social channel is the type of content you posted in the first place. With image and video posts, it is usually best to keep users on your social platform to view the media. Well-packaged link wall posts with killer copy and amazing images can be effective at generating clicks to external websites. Pay attention to actions earned with social ads. Make sure you're getting what you want by measuring the results and adjusting.

Psychographic targeting

What makes social media amplification different from search keywords is the powerful targeting capabilities, known as *psychographics*. Rather than targeting phrases, psychographics take clear aim at the attributes that make people who they are.

Most marketers' first exposure to psychographic targeting was in late 2007 with Facebook Ads, which allow advertisers to target users by interests, affinities, proclivities, biases, predispositions, religion, sexuality, occupation, education,

workplace, preferences, age, gender, likings, age, predilections, attractions, medical conditions, economic status, peccadilloes, desires, correspondences, empathies, relationships, appetites, weaknesses, tastes, inclinations, corporate loyalties, and numerous other highly personal penchants.

Facebook was the first to roll out this capability, but now psychographic targeting is everywhere. In reality, radical targeting has been percolating Internet-wide for years. Savvy marketers use *digital media platforms* (DMPs) and ad exchange technology to target users almost everywhere they roam on the Internet. It's a lot like Facebook Ads, only better and Internet-wide.

More than 30 networks, consortiums, and other data brokers resell access to their targeting data, making users available to DMPs. The names of some of these networks are shown in Figure 7-43. If you don't yet know the data providers in the following graphic, you will. Data objects from these companies can be layered using both the AND and OR operators to create stunning targeting combinations.

Figure 7-43. Examples of sources of targeting data

Consider the possibilities of layering data from such pervasive and definitive sources, in addition to Facebook, Twitter, and LinkedIn. Imagine clarifying ecommerce marketing targets by financial qualifiers gleaned from Experian, TransUnion, and MasterCard. If you like occupation targeting in Facebook and LinkedIn, how about targeting users Internet-wide with Dun and Bradstreet and Bizo?

We live in a brave new marketing world, where it’s commonplace to layer competitive and symbiotic brand affinities with credit qualifiers. These powerful examples are just a small snapshot of the psychographic display-targeting ecosystem. There are thousands of ways to layer data for the targeting win.

Social advertising can drive links to your site

Another perk is that well-executed organic amplification of relevant content often results in links. Marketers are already learning to ditch old-style link-building efforts in favor of paid-organic link building to earned psychographic audiences, bloggers, and journalists. It's easy to target media roles using the "More Demographics/Job Title" field of Facebook Ads and occupation targeting in LinkedIn.

Whether targeting journalists, morning show hosts, and news producers, well-executed organic amplification of relevant content nearly always results in links.

Note that paid-organic social content distribution does not replace good old-fashioned public relations and peer-to-peer outreach. Human-to-human marketing will always work. For SEO, most businesses need to sell with content. As a result, an increasing number of content marketing companies will adopt these practices as standard fare in the future. Great marketers already are.

As we've mentioned, social content and SEO distribution used to be mostly free. Now, in the unavoidable effort to monetize, social media platforms and search engines charge for the same distribution. That's the bad news. The fantastic news is that the "ads" look very similar and are often barely distinguishable from the same organic page units that used to be free. Use paid organic-looking social psychographic content amplification to dominate distribution.

NOTE

A special thanks to Marty Weintraub for his contributions to this portion of the chapter.

Build Offline Relationships

Leveraging online platforms for the purpose of building relationships makes great sense, but there is no reason to stop there. Do the major influencers in your space speak at conferences, attend networking events, or host events of their own? If so, go to one of these events and introduce yourself. This can include not only bloggers, but other influencers in your space.

Map out the major conferences that cover your market. Check out local Meetups and more general conferences like BlogHer, New Media Expo, and WordCamp.

Another related tactic to consider is contacting publishers of an authoritative site and offering them a free seminar/webinar with you as the speaker. You could also propose a joint marketing campaign with them. Either way, call it part of your company's outreach campaign to build relationships with leaders in the space.

Make sure you articulate well the unique nature of what you will present, as you have to attract influencers' interest with your pitch before you can take the next step. Ensure you bring a demonstrable value in the actual presentation.

Then ask influencers to follow you on social media, answer their questions, and make yourself available to answer any follow-up questions they have by phone or email. Once you have done that, you will have a number of relationships with people involved in the authoritative site.

There are other ways to extend this too. For example, you can sponsor the organization in some fashion. There are, of course, sites that link to their sponsors, and this may be a win, but Google will want to discount such a link (because it is "paid"), so you should not count on the SEO aspect of it. However, the visibility a sponsorship provides can enhance your reputation, and you'll likely be able to establish a deeper relationship with the organization you're sponsoring.

Last but not least, most likely you have other businesses/organizations that you interact with in the normal course of your business. Once again, you can ask them directly for links, but you can also ask them for introductions to other people in your space. This type of networking can be an effective way to build relationships that eventually lead to high-value links.

Relationships and Outreach

As you've seen, all content marketing campaigns involve building relationships with others. These relationships may be quite involved, including one-on-one, substantial interaction, or they may simply be cases where people have become fans of your content and follow you on social media or at your blog. Either way, you need to get exposure to these people, and outreach plays a role in getting these relationships started.

Building Relationships with Influencers

Earlier in this chapter, in "Leverage Influencers and Influencer Marketing", we emphasized the importance of building relationships with influencers. In this section, we will explore a way to scale your efforts to build those relationships more quickly and more efficiently.

Build an initial list

If you're active in social media, or if you are socially active in your industry, you should already have a good idea of who the big names are—the people who speak at conferences, who have huge Twitter followings, or whom everyone circles on Google+. Start making a list of influencers, and use the names you already know as initial targets.

If you are just getting started in social media, or you're running a local business and don't go to a lot of conferences, you're going to have to identify influencers and begin following and interacting with them through social media. For instance, if you own a restaurant, follow all of the Food Network personalities, like Rachael Ray and Bobby Flay, and other celebrity chefs and related public personalities like Martha Stewart.

If you're writing a science fiction novel, you could follow other science fiction authors, editors, agents, and publishers of science fiction. It would also be useful, though, to find news sources that report on topics like science, space exploration, medicine, and other subjects that could be related to what you're writing about.

It will take several hours over the course of a few weeks to build a solid initial list of targets. You want to follow all of these targets on the platforms in which they are most active. Don't push friend requests at them until you have established at least some level of interaction.

This list should be broad, as it needs to cover an appropriately wide range of topics and people. At this point, don't pare down the list at all—in fact, make it bigger. Expand out as much as possible. Follow anyone and anything that might be related.

You should also expand out beyond social media a little. If you identify some extremely high-value influencers, subscribe to their blogs, RSS feeds,

newsletters, and podcasts. You are not doing this so much to learn from these people (though there's almost certainly something to learn from them, no matter how much of a guru you are) as you are to get to know them better. Only after you know them are you prepared to interact with them positively.

Begin interacting

As just mentioned, once you get to know some of the people you're following, you can begin interacting with them. If they ask a question on Twitter, send them @ replies. If they post a link to something, comment on it. Retweet or reshare their posts. On Facebook or Google+, post thoughtful comments or links to stories with more information. Repin, tweet (with an @ mention), and comment on their Pinterest photos.

Don't ask them to link to you or repost your content at this point. Later on, you're going to contact these people directly and ask them to help share your content or write a review, but you're likely to fail if you make your request too early. You will have a much higher success rate if you're already somewhat familiar to them through social media.

At this point in the process, it's about the relationship building. Give genuinely, and add as much value as you can. They will start giving back to you later when they are ready, and you can't push it. This process will also raise your own influence on social media, and not just with the people you are targeting.

Refine the list

While you're engaging your followers, begin to whittle down your initial list of influencers by using reliable metrics. How popular are their blogs and other domains? How much reach do they have beyond their social media followers and fans?

Instead of answering that question on your own, you may want to use a tool like Kred, shown in Figure 7-44.

Another useful tool is Klout, which creates a composite "Klout Score" that shows your own influence (the more influential you are, the better your chances of getting the attention of a more powerful influencer), as shown in Figure 7-45.

Some other tools that you may find valuable for this purpose include:

- Moz domain authority—as measured on Open Site Explorer

- Majestic's Trustflow and CitationFlow Metrics
- LinkResearchTools PowerTrust
- Followerwonk, part of the Moz toolset

Figure 7-44. Kred shows you the most powerful influencers among those you're following on social media

All of these sites use proprietary algorithms to measure the number of RSS subscribers, friends, fans, followers, comments, shares, and inbound links and weigh them according to their own opinions on what is valuable. Depending on your industry, product, and/or content, some of these tools may be more valuable to you than others.

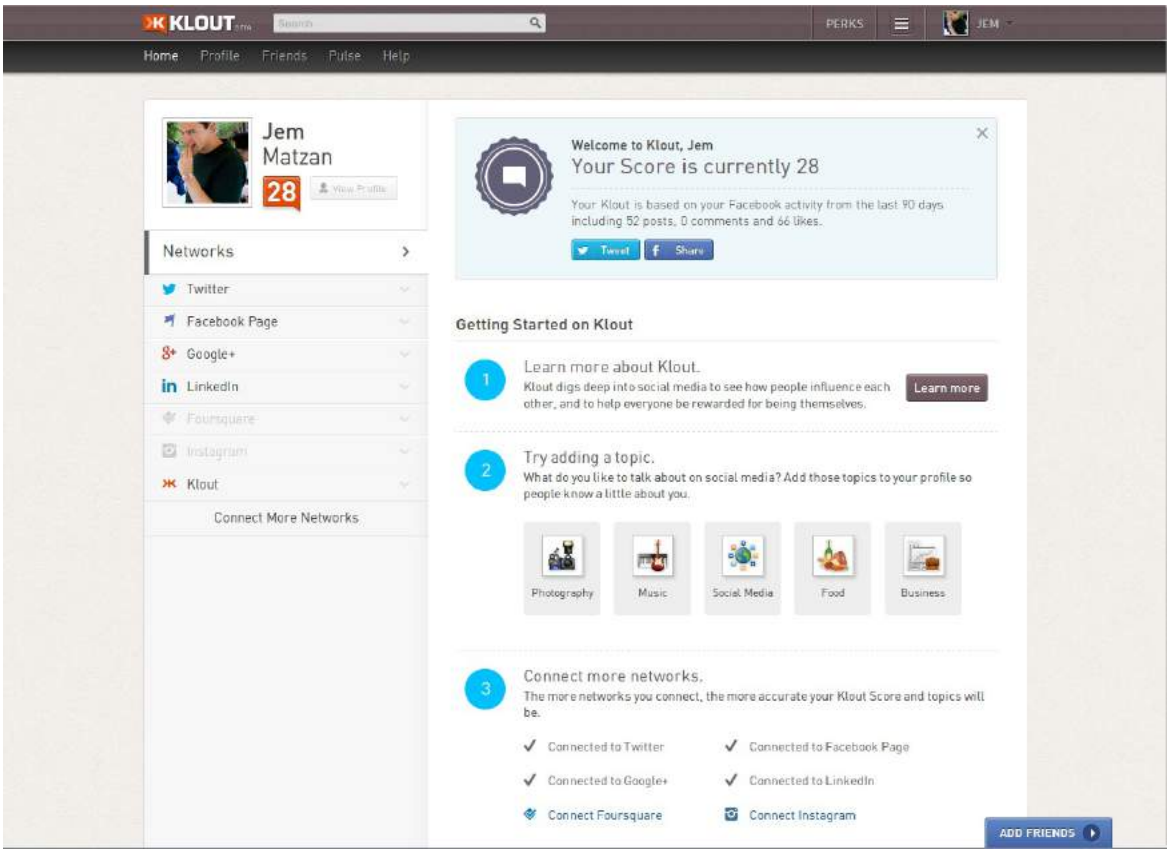


Figure 7-45. Klout also shows your own influence levels

Ultimately, though, this is not a numbers game; it's a quality game. If someone has only a dozen followers, but everything he posts goes massively viral, then despite the numbers he is a highly valuable influencer. That would be an unusual situation, but it's possible. There are some people who are on only one or two major social networks, such as YouTube or Google+; these people can be deceptively influential.

It's also not a numbers game in terms of how many influencers you must contact. Half a dozen really influential bloggers are more powerful than a thousand small-time bloggers.

Use Twitter as a warmup

Starting a dialog on Twitter is often a very good way to start, even if you're not targeting Twitter specifically. If the influencer you're going after is active there, then it can be very valuable to engage her there. High-value targets are almost certain to ignore a cold email, and will require a warmup on Twitter. If you make

this effort to engage her with meaningful comments and retweets, it'll help her become more comfortable and familiar with you.

Tweet @ people. When the influencers you're watching say something to which you can respond intelligently, reply to them positively. Answer their questions. You can also thank people publicly for recommending something.

Watch their Twitter feed for a while so that you can be sure that they really are as relevant and interesting to you as you think they are, and that you like what they're saying. You may discover that while someone seems to be influential, she mostly posts negative things about companies and products; that outreach effort could be a disaster for you.

When you are trying to build a relationship with someone, it's often best if you don't hide behind a brand account; use your full name (your personal account). Or, if you intend to have more than one person represent you on Twitter, you might create a public persona who virtually represents your company. This name will be your public face, and will get all the credit for your posts. If you're going to create a persona, make it a woman; women (or female names, at least) have a much higher response rate in outreach email.

The Twitter warmup is not a one-time event; it is an ongoing effort that will take weeks or months to complete. Note that this same strategy also works well on Google+.

Use email to contact the influencers

Cold calls are always difficult, even when they're emails. You want something, and you're contacting someone who is in a position of power and asking him to give it to you. This is not a role that anyone enjoys being in, but the most successful salespeople and marketers put that aside and master the art of initial contact. Think about how much you want social media success (or the business success that is on the other side of social media success). What are you willing to do to achieve that?

All you need to do initially is get your foot in the door with one good, solid influential mention on social media. Then you'll use that first big mention to get the attention of other influencers. Even among the top people on social media, everyone wants to be in with the "cool kids." That first mention becomes the best method of obtaining mentions from others, so invest big in your first major

influencer outreach effort. From that point forward, it gets easier.

Before you start sending emails, give serious thought to where you want to claim to be mentioned. On Twitter? On a famous blog? On Facebook? The social network or site where you want to be mentioned will help you narrow down your choices and customize your outreach message.

Leverage Pinterest group boards

If you are an active and established Pinterest user, you can use group boards to your advantage in contacting and participating with Pinterest influencers. You can invite a mutual Pinterest follower (someone whom you follow, who also follows you) to be a contributor to one of your boards. This creates a group board. You will always be the administrator of that board, but you can add as many mutual followers as you like. By the same logic, one of your mutual followers can invite you to pin to one of his group boards.

If you can find a way to get a Pinterest power user to participate in one of your group boards, you could inherit a lot of his followers as a result. Similarly, participating in the same group board as an influencer gives you much more trusted and intimate access to him, and a much higher chance of getting a review, repin, or mention from him.

Get contact information

Certainly, if you can get someone's phone number and you're fairly certain she'd be receptive to a quick phone conversation, then give her a call and follow up at an appropriate time with an email.

Some people are more inclined to respond on-network (a message sent on the social network they participate in), and some are more receptive to email. Sometimes, email is better because it's more personal and is more easily noticed. Some influencers are so popular that they don't check their in-network messages at all—they get inundated with notifications and frequent messages from raving fans—but they almost certainly check their email.

Getting someone's email address can be tough and time-consuming. If you find yourself spending an unreasonable amount of effort trying to get the email address of someone who has taken great pains to hide it from the public, then fall back to tweeting @ her on Twitter. Once you have a relationship with her there,

you can simply ask her if you can send her an email, and she may provide it to you.

First, try looking on the influencer's blog or corporate site for an email address. Maybe it's in an obvious spot.

Do you have his business card from an in-person meeting at a conference or event? People usually put their email address on their business card. Does his blog or corporate site have a "Contact us" page? That goes to someone's email—if not your target influencer's, then probably his assistant's. If you add someone on LinkedIn, you can get his email from there. The reply to the invitation comes from his email address. You can also send InMail, which allows you to send mails to people you are not connected to (for a fee).

As a last resort, do a Whois lookup on the influencer's blog domain name (assuming it is a real domain name, not a subdomain on a public blogging site like TypePad, WordPress, Tumblr, or Blogger). Every domain has contact information for the domain owner; hopefully that will include the email address of the person you're trying to reach.

Pay for reviews

If you are trying to reach a product reviewer, be warned: some high-level reviewers may ask you for money in exchange for a review. Among journalists this is highly unethical, but bloggers and social media power users don't always fall into the "journalism" category. In the United States, anyone who posts a product review, whether she considers herself a professional reviewer or not, must disclose the fact that she received compensation (including a free product) for it. This is a U.S. Federal Trade Commission (FTC) regulation.

The FTC puts the responsibility for disclosure on the brand, not the publisher. That means you. So if you send a check with your review materials, or if you tell a blogger or journalist to go ahead and keep the review unit you're sending, you must check back and make sure that she has clearly printed an appropriate disclosure notice. This even applies to endorsements on Twitter. The preceding paragraphs contain really important legal liability information about product reviews. Don't skip this stuff. Read it twice. You don't want to get in trouble with the FTC.

Paying for reviews, or giving free products to reviewers or influencers, is an

ethical grey area. Usually it’s all right with disclosure if it’s posted to someone’s blog or social media page, but it could reflect badly on you depending on context. In some instances, it is definitely unethical—for instance, book reviews or any product review on Amazon.com that is done on a for-hire basis is against Amazon’s rules.

Google also frowns upon paid links, so if you are paying a blogger for a review that includes a link back to your site or product page, that is technically in violation of Google’s rules unless the blogger applies the `nofollow` attribute to the link in the HTML tag.

Some social media power users offer expanded services for hire. For instance, a popular Pinterest user might solicit money to do a photo shoot with an in-depth review or interview. Food bloggers might offer to publish a recipe and high-quality photos and a review of the resulting product. This is not prohibited by Pinterest, but if you take this route, make sure any links are `nofollow`ed and the proper disclosure is printed.

Paying for this type of visibility, and doing it the right way by `nofollow`ing the links, may not have direct SEO benefit, but it can still result in others seeing your content and choosing to link directly to it. As a result, this type of tactic might have a place in your overall strategy.

Create templates

First, you need good unique content. A lot of social media gurus will tell you that you should try to build followers by linking to interesting stories. That will get you a small amount of success. If you want major success, though, you’ve got to be the source of interesting information, not just someone who links to it first.

Start with a pool of unpublished articles that can be customized somewhat. They should be mostly or completely written, but don’t tell anyone about that. The point of your influencer outreach email is to ask for help or input on an article you’re currently writing. You will, of course, quote or cite the influencer as a source. Quoting experts for articles is one of the best ways to get them to link to you or mention you on social media.

The article you’re customizing should be relevant to a topic that the influencer covers or is interested in. Don’t publish it more than once; don’t try to repurpose

a published article unless it is substantially customized for a different platform (for example, you can write a summary of it for a LinkedIn post, or a Google+ post, or other platforms, as long as you have substantially customized it to that platform).

Contact the influencers

Don’t mass-mail your target influencers. Personalization is hugely important. The message must be personalized. Absolutely do not send a generic message that starts with something like “Hello, fellow blogger.”

Not only are impersonal messages likely to be ignored, they’re also likely to be caught in a spam filter. The same message sent to many addresses on the same network (such as Gmail or Yahoo! Mail) might be flagged as spam by the system. Not only will you have completely wasted your time, you’ll also have associated your email address with spam. Your whole domain could get blocked.

This doesn’t mean that you shouldn’t have a standard message template that you start with. It does mean that the template must be customized significantly. To do that, read the influencer’s blog and social profiles; write as if you are familiar with her. Because you’ve been following her and attempting to engage publicly with her for a while, most of this work should be done.

The template should not be long—fewer than 300 words or so. Busy people don’t have time to read long messages from strangers.

Be genuine and polite in your initial outreach

Make the influencer understand that you took the time to read her work; express genuine interest in her. Mention something she did in the subject line: “Just saw your <title of last blog post>.” That will get her to open the email. You can also try mentioning her name in the subject line, or using an actionable phrase like “Need your decision,” “Need your feedback,” “Can I quote you for this article?”, or “One minute to read this upcoming article?”

For the message body, don’t talk about yourself or your background too much—to a busy stranger, that information is boring and skippable. Establish yourself in one sentence. If you referenced something she did recently, follow up that reference in the message body. Or talk about something else she’s done recently and how it affected you. Talk more about her than yourself.

If you're writing to someone who consistently uses an online handle or persona, then go ahead and address her by her blog handle or persona first. If there's no response to the first outreach, make note of that and use her real name in the follow-up email. Most people use their real names today, but in the old days handles were more popular. People want the credit; they want to be more transparent about who they are.

Close with a request for a reply. You want to encourage more dialogue. Don't ask for more than that, though—never ask directly for something of value in an outreach email, such as a link or a review.

The time at which you send your message can make a big difference. Each industry has different prime days and times. Look at the posting schedule if you can, and try to figure out the editorial calendar as well (if it's a big blog, the blogger probably plans out certain themes for content on certain days or during certain months). If it's a tech blog and you know that the blogger publishes three stories a week, try to get her the day before a story publishes (or the day after), but not on the actual day. Most of the responses you're going to get are early or late, when people check email. Very busy people typically check email on a schedule.

Spammers send their email late at night, so try to be off-schedule from that. Send your email during the day in the afternoon to avoid the spam rush, but before the last email check of the workday in your influencer's time zone.

Follow up

If you haven't heard back from your first outreach within three days, then you should send a follow-up. Following up is more crucial than many people realize. Michael Geneles, developer of outreach tool Pitchbox, estimates that following up increases your response rate by 60%. You should have a separate template for follow-up emails; don't send the same message you sent before.

It rarely pays to go beyond two follow-ups—don't send more than three emails to someone who isn't responding to you. If someone doesn't reply after the third attempt, there is a very low chance of a reply, and you're better off spending that effort on the next influencer on your list.

Before sending your first follow-ups, get reinvolved on Twitter or another social platform to make sure you've got a solid connection.

Make a third attempt

If your first two emails didn’t get responses, then the third probably isn’t going to work, either. You’re going to have to take a different approach with your third attempt. Is there a different email address for this influencer? Some people list a Gmail account as a technical contact or as part of a regular Google account, but they rarely check it. Look for an alternate email address for this person, and explain that you’re trying to get in touch and wondering if you’ve got the right address.

You can also try to use in-network messaging through Facebook, LinkedIn, or other social networks in which this influencer participates. If you do this, come up with completely different content for your email; mention that you’ve tried to reach out and haven’t gotten through, note that you love his work and want his advice on something, and ask for a reply.

Track responses

The only way to improve is to measure. Keep metrics on what gets through and what doesn’t. If a particular template or a particular kind of influencer is not working out, then change your strategy. Was the response favorable? Was there a response at all? How often does the first follow-up work? The second? You may want to make a quick spreadsheet to keep track of this data.

Try other approaches

If you totally strike out despite your best efforts, then try a more indirect approach. Try to get through to the people whom your target influencer follows. These are her friends and close colleagues, and people she admires. Likewise, if you do manage to get a positive reply, ask the influencer who else she knows who could contribute to the article or provide a quote. Oftentimes you’ll get names, phone numbers, and email addresses of other influencers—and with a personal referral!

Nothing establishes initial contact like actual in-person communication. Go to industry conferences and meet your influencers in person. Have coffee with them. If they are not accessible in person, then build relationships with their friends; they will become your references. Take a long-term strategy. Or play golf with them—or whatever social activity is appropriate to the industry.

Watch for mentions and links

If you were successful in getting a positive response, then keep a close eye on the networks that you're targeting. Watch for links and mentions that involve you or your company. This might not happen immediately. Often an article costs you a lot of money or time, so don't wait too long to see if it gets published or mentioned. If you get an agreement, follow up after a week and see if there's anything else you need to provide. If you don't get a publication date, ask nicely for one.

Twitter is the hub for monitoring what's new on the Internet, so that should be your place to watch for mentions. It may not come in the form of a hashtag or an @, so watch your target influencer's feed closely.

Establish yourself on Google+

Google+ can be a good platform for outreach, but it requires a lot of work. While it is growing rapidly, there are still market segments that are not that active. If that's the case in your market area, you might still want to get started there to build a relatively strong following and get an edge on the market, but it may not be a major focus for your outreach efforts.

However, there are many market spaces that are well built-out on Google+, and if that's the case, the platform can play a significant role in your outreach efforts.

OUTREACH CASE STUDY

An etailer wanted to generate critical mass for a content marketing campaign. The company sells premier designer swimwear brands and styles, many of which have been featured on celebrities and in leading fashion magazines. Since the company's launch in early 2000, it has established itself as a go-to place for fashion-forward swimwear.

It was successful enough that it reached the limits of what traditional ecommerce techniques, such as email campaigns and banner ads, could provide. At the same time, the influence of traditional media was declining and consumers increasingly began to rely on social media for information and trustworthy recommendations. The company decided that social media was a must-have marketing channel, for several reasons:

- Increasingly, it's where consumers were going to find out what's new and what to buy.
- Content on social media sites could help build credibility and loyalty. People choose which brands they want to follow, and trust the brands they choose.
- Relationships with authoritative brands would drive traffic to the company's own website, boost its domain authority, and improve its search engine rankings.

- Exposure via social media could help offset shrinking publicity opportunities in traditional media. With declines in advertising revenues, the company was finding fewer traditional media outlets and the outlets that remained had less editorial space than before. And with fewer opportunities, the competition for placements had intensified.

The Initial Content Marketing Strategy

The company's social media strategy was to offer high-quality content to influential bloggers and independent media outlets. The world of independent media is a classic long-tail market; instead of a few key influencers with wide reach, it contains vast numbers of outlets.

There are a handful with huge audiences, and many more with fewer but highly loyal followers. Because this world is fundamentally based on social relationships, the company couldn't simply blast out one-size-fits-all content to everyone. It had to cultivate relationships and develop trust with influencers, and create content that was unique and tailored to each opportunity.

The company started with a seat-of-the-pants approach: come up with content ideas, google around to find some blogs that look good, poke around on the sites to find contact info, send emails pitching the ideas, and see what came back.

That approach seemed to work at first, but once the company tried to scale up past the first 15 or 20 opportunities, the process quickly became difficult to manage. Many responses to its outreach were coming in, often days or weeks after the initial email had been sent, and some arrived from email addresses different from the one the company had originally sent the message to.

When it ran new searches to identify more placement opportunities, the problems compounded; someone had to check these leads against ones already in the pipeline to be sure they didn't overlap. With all of these administrative challenges, the campaign wasn't scalable. It wasn't a question of how many opportunities were out there, it was a question of how many opportunities the company could manage.

Cutting back wasn't an option because independent media campaigns must reach critical mass to be successful. A mention here or there doesn't create momentum. The buzz starts when consumers—and influencers—start seeing your name pop up everywhere.

Stopgap Measures

Next, the company tried setting up Excel spreadsheets that contained lists of contacts and related information, such as the initial and last contact date for each. Even with this higher level of organization, it still took a lot of time to manage these spreadsheets—time that would have been better spent cultivating relationships, pitching ideas, and creating content.

As time went on, the pipeline began to clog. Opportunities stalled, and in each case, the company had to go back and figure out why. Was the ball in the company's court, or the influencer's? Did the contact owe the company feedback? Had the contact gone dark and nobody had noticed? Soon, the company found itself creating more spreadsheets and other ad hoc systems to try to keep up with these issues.

A customer relationship management (CRM) system seemed like the next reasonable step. The company looked at several, but the logic and the language were a poor fit because CRMs are set up for sales, not for media outreach. For example, most are built around "deals"—how much a sale is worth, how far along it is, who the decision makers are, and so on. That logic doesn't work for a media campaign. Similarly, CRMs didn't have metrics for social media, such as domain authority, or a good way to track and report results.

There are many good tools in the market that help with this process, and provide differing levels of

functionality. Traditional media contact databases such as Cision and Vocus can be invaluable sources of information.

While these are a bit pricey for a small-business owner, they can rapidly speed the process of obtaining contact information for media people and bloggers. These tools allow you to enter search criteria and they come back with a rich array of information of the right people to contact in the market area you are searching. You should still review and filter the results, but this is a simple and easy process.

There are also tools that allow you to track and manage contacts with media/bloggers on an ongoing basis, such as GroupHigh and Pitchbox. These provide somewhat similar functionality to CRM systems, but are much better suited to the task of identifying, targeting, and building media/blogger relationships than CRMs, which are focused on sales.

Automating Management of the Outreach Campaign

In 2013, the company in our case study became a beta tester for Pitchbox, a tool that automates most of the administrative tasks of outreach (Figure 7-46). It was built specifically around the tasks involved in social media outreach. For example, it starts with a search function that looks for new opportunities on a scheduled basis, and then strips out contacts that are already in the pipeline. It ranks the opportunities by influence—using criteria that the user can define—which helps the company prioritize its efforts. It automatically identifies contact info and allows the company to reach out individually (for the most important contacts) or using templates to scale.

Then it follows each opportunity through the entire pipeline—from initial contact through placement. Along the way, it flags stalled opportunities, shows where the campaign stands overall, and guides the user to the next step.

The automation of these administrative tasks revolutionized the company's outreach efforts. Using the same resources as before, the company saw placements go from approximately 10 per month to 50 per month. Administrative tasks were no longer throttling the campaign, and the amount of time spent on these tasks declined from approximately 75% of total hours to 15%. That freed up time for staff to spend on high-value activities: cultivating relationships, coming up with ideas, and creating content.

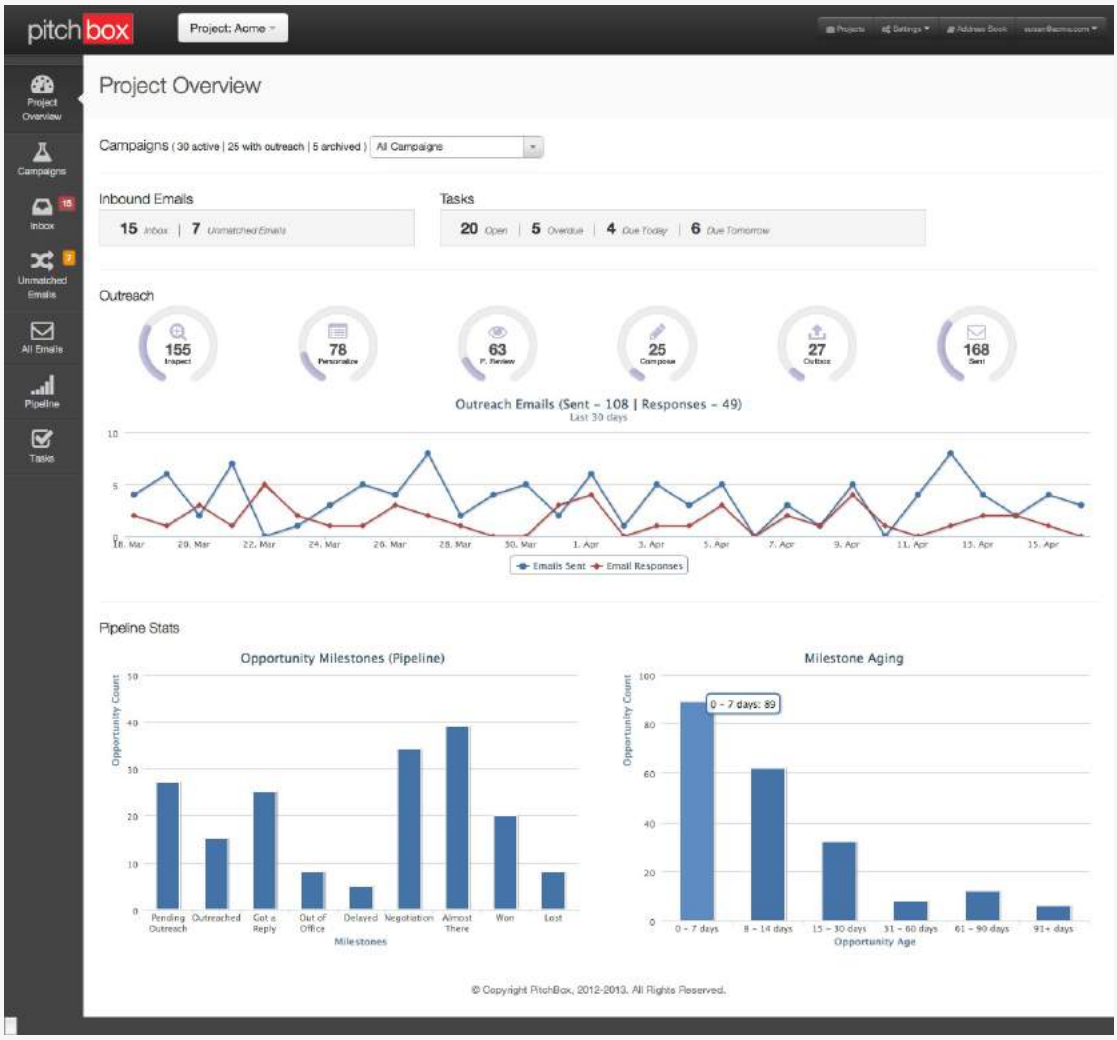


Figure 7-46. The Pitchbox dashboard

In addition, the company was better able to forecast results, allocate resources, and hit goals. For example, it can look ahead and see how many placement opportunities it is likely to have in upcoming months, so it can allocate resources to create the content needed to fulfill those opportunities. Or if the company sees that it needs more opportunities in the pipeline to meet its goals, it can deploy more resources toward outreach instead of content creation.

Achieving Critical Mass

Automation is critical to any long-tail business because long-tail markets have too many events happening to manage individually. The same principle holds for social media campaigns. They need to scale—reaching hundreds or thousands of influencers in a highly fragmented market—in order to reach critical mass. And to be scalable, they need tools and techniques to help manage these events.

Creating a Value Proposition for a Relationship

Now that you understand the sites that are most interesting to you, it's time to think about how to get started with key people involved with each target site. There are many potential tactics, some of which include:

- Start by engaging with them on social media or by posting comments on their articles. These contacts should be completely noncommercial, and designed to add value to the conversation they started with their posts.
- Ask them to provide a quote for an article you are writing. Make this easy for them, so it's more likely they will say yes.
- Offer to interview them for something that you will publish on your site. Or, ask them to a webinar or a Google+ Hangouts on-air interview with you.
- Find out if they will be attending a conference, and if so, go to that conference yourself and find a way to meet them.
- Monitor their article and social postings and take note if they ever ask for help in some fashion ("I wish someone would show me how to...". Then, do it for them—even if you have to spend a few hours learning how to do it yourself first.
- If you know someone who knows them, ask for an introduction.

These are just a few ideas. Regardless of how you go about it, start by just giving to the people you want relationships with, or engaging in a way that is interesting or useful to them. Don't treat this as an "I need something, so I will just ask for it" situation. These relationships are important, and it's not a good idea to start off on the wrong foot.

Once you have a relationship, many possibilities start to open up, and your reputation and visibility can grow as a result. Ideally, you never actually request a link in your first communication. You focus on building a relationship, and at the right point in time, the person you built that relationship with decides on their own to link to a fantastic piece of content you have created that is of high interest to them.

Using Direct Email Pitches Effectively

People have a growing distrust of unsolicited emails from people they don't know. This bias is making direct email pitches harder to do successfully, and other tactics that are more conducive to relationship building are recommended instead.

If you are going to use this approach, however, the most important thing to remember is that the person you are emailing to request a link probably did not wake up this morning wondering what links she was going to add to her site. And certainly, she was not expecting or waiting for your email. Basically, you are interrupting her to ask her to do something for you, and she may have no prior reason to trust you. Given that, there are a few simple guidelines you should follow when making an unsolicited pitch:

- Keep it simple and short. The person you are contacting is receiving an email that is unsolicited. She is not going to read a two-page email, or even a one-page email.
- Clearly articulate your request. It is an investment to get someone to read an email, and it is critical that the pitch be clear about the desired result.
- Clearly articulate why your site deserves a link, or why they should take content from you. Generally speaking, this involves pointing out the great content or tools on your site, and perhaps citing some major endorsements by others.
- Follow each and every guideline of the CAN-SPAM Act. Unsolicited emails are not illegal as long as they follow the guidelines of the act. Do not even think about violating them.

Other Ways to Earn Links

In addition to content marketing, there are many different link-building tactics—too many to list in this book. This section will examine in depth some of the more common ones.

Web Directories

Obtaining web directory links was a tactic used in the past to drive SEO results,

and many years ago, they did work. However, with the exception of a very few high-authority directories (such as Yahoo!'s directory and that of the Internet Public Library), very few directory links provide value anymore. There are still thousands of directories out there that will try to draw you in to buying a link from them, but you should stay away from them.

NOTE

In this section, we are discussing web directory links. These are distinct from local business directories, and the guidelines differ. Local business directories focus on providing the name, address, and phone number of local businesses, along with other related information, and they are important to local search. They are discussed more in "Optimizing for Local Search".

It may be worthwhile to pursue four or five web directory links, but probably no more than that. This group should be restricted to the most authoritative sites, such as the Internet Public Library, Business.com, and maybe a couple of recognized authority sites that have directories of people in your industry. You can also try to get into DMOZ, though its authority has declined over time, and your efforts can be frustrated if the category you are pursuing does not have an active editor.

Manual Social Media Link Creation

One way to create links manually in social media environments is by visiting social media sites, forums, and blogs and leaving behind comments with self-referential links in them. However, the great majority of these environments `nofollow` their links, and obtaining links by putting them in comments and forums is a practice that is not approved by Google or Bing. Doing so is also very bad for your reputation; as we mentioned earlier, it is akin to showing up at a neighborhood party and asking everyone to give you money.

Another useful way to view these activities is that "you can't vote for yourself"—that is, any link you obtain by linking to yourself is highly unlikely to be valued by the search engines. A more useful way to obtain links from blogs, for example, is to build a relationship with the blog owner. The major steps of this process are as follows:

1. Build a list of blogs that are related to your topic area.

2. Start visiting those blogs and adding comments of value without linking back to yourself, and develop a relationship with the author(s). The early stages of the relationship begin when the author starts responding to your comments. You can even start by reaching out to the author and interacting with him through one of the major social networks, such as Twitter, Facebook, Google+, or LinkedIn.

3. Once the relationship has been built and seems solid, let the author know about a related value-add resource you have, ask to interview him, or suggest a collaborative content development effort or study, either through direct contact with him (preferred) or in a comment. Make sure there is a real connection between what you propose and the content from the author.

These steps are meant to be conservative to avoid a backlash from blog owners and/or authors. You can extend this process to forums or social media sites as well. The key in all these cases is to avoid implementing links to your site/page and instead use these techniques as a way to build relationships with people. The payoff comes when you later publish a great piece of content that they decide to link to because they started following you as a result of your interactions online.

There are ways to be more aggressive with this approach. Some publishers do not really care about building relationships first and want to push the process much faster. However, there are two significant issues with this:

- Depending on the level of aggressiveness, it may be a violation of the webmaster guidelines, and the search engines may choose to take action against the publisher who pursues this strategy.

- There could be a backlash from the community itself. Offending one blogger may not be a huge issue, unless she is very influential. Offending hundreds of bloggers, on the other hand, would probably be much worse, particularly if you are trying to establish your site as authoritative in a topic area. In forums, blogs, and social media sites, offending people can quickly scale to a problem of large proportions.

Gray Hat/Black Hat

As we previously discussed, some publishers choose to push the limits or ignore the webmaster guidelines in their quest for links. In the following sections, we will look at some of the more popular tactics in detail.

Purchasing links for SEO

Historically, one of the more popular spam techniques was to buy links. This had two significant advantages:

- It is easy. There is less need to sell the quality of the content of your site. The only things that need to happen are determining that the third party is willing to sell a link, and setting a price.

- Because the link is an ad, you can simply specify the anchor text you want. Anchor text is a powerful ranking signal, and this is one of the major reasons people engage in link buying.

The practice of buying links has fallen out of favor, even among aggressive SEO professionals, as the risks with this practice are very high.

Google's policy on paid links

The major downside is that buying links for SEO is against Google's webmaster guidelines. Here is a brief summary of Google's policy on paid links:

- Links given in return for compensation should not be obtained for purposes of increasing PageRank.

- The link should be freely given, and the publisher of the potential linking site should be informed of what the publisher is doing. An example of a link where the publisher is not informed is one that is hidden in the `<noscript>` tag of a JavaScript-based widget.

Google is not saying that publishers should not be able to buy ads on the Web; rather, its policy is that links should be purchased only for the traffic and branding value they bring. Google also recommends that publishers selling ads on its site use `nofollow` links, which means they will have no SEO value.

On another note, PPC campaigns using AdWords, BingAds, and other popular

advertising platforms are not considered a violation of the policy against paid links. This is because search engine advertising links are easy for the crawlers to recognize, and the search engines simply don't count them as a source of link authority.

Strategies that are not considered buying links

In some strategies, money is involved in obtaining a link, yet the link is not considered by the search engines to have been bought. Here are some examples:

- Paying a PR firm to promote a site
- Paying people to create great content for you and paying someone else to tell other people about it
- Paying someone to ask for (as opposed to actually buy) links via direct link requests

The key point is that these strategies do not compensate the site itself for the links given, so they are considered legitimate approaches if the links are given freely.

Swapping links

Another dangerous tactic is doing a large percentage of your link building through reciprocal links. Once again, this is easy to do in principle. It is not hard to find sites that will accept the "link to me and I will link to you" value proposition.

However, the search engines potentially view this as barter, and therefore these arrangements are not considered real endorsements. The search engines are not likely to question a few selected link exchanges with sites closely related to yours. It's a problem only when the link swapping becomes a significant portion of your backlink profile. That is a situation that looks manipulative, and the engines will most likely discount those links.

Using link farms/link networks

In the early days of search, publishers developed link farms and link networks as tactics for gaining cheap links. A *link farm* is a website or a group of sites whose primary reason for existence is to cross-link between themselves and other

websites. Generally speaking, the links are created through aggressive reciprocal linking.

These sites are typically very heavily interlinked, so they can be pretty easy to detect. Part of the reason is that because they have little redeeming value, they typically do not have high-value links coming in to them from other sites, and most of the links result from various cross-linking schemes.

How can you differentiate a link farm from a legitimate directory? It's simple: a link farm has poorer organization and more links per page than a directory. In addition, a human review will show the lack of a strict editorial policy in a link farm.

Link networks are a similar concept. The network exists for the purposes of creating links between sites, and it can be a bit more sophisticated than a link farm. For example, you could create a club where publishers agree to contribute a link in return for getting a link from somewhere else.

If managed with great care, the clustering of links between sites can be limited, and this can be a bit harder for search engines to detect. However, this tactic is highly vulnerable to a disgruntled webmaster simply reporting the scheme to Google.

A related concept is the notion of *three-way link swaps* (a.k.a. triangular link swapping), where Site A links to Site C in return for Site B linking to Site A. In this scenario, Site C may be the site the publisher is trying to promote, and Site B may be a site it uses to provide low-value links to people it trades links with.

This is almost always a scam, because Site B is probably a low-value site with little to recommend it. So, the publisher of Site A is providing a good-quality link in return for a low-quality one.

Think of it this way. Google employs some very smart mathematicians who understand graph theory.[17] So you can rest assured that the Google algorithms will do a sophisticated job of detecting unnatural linking structures, such as the one in Figure 7-47, which is called a *clique* in graph theory. Notice there are five nodes (or vertices), each of which is linking to all four of its neighboring nodes. They never miss a link. It all looks so perfect. Naturally occurring neighborhoods on the Web aren't perfect like that. If it looks perfect, it's been engineered. Google will be suspicious of unnatural-looking neighborhoods.

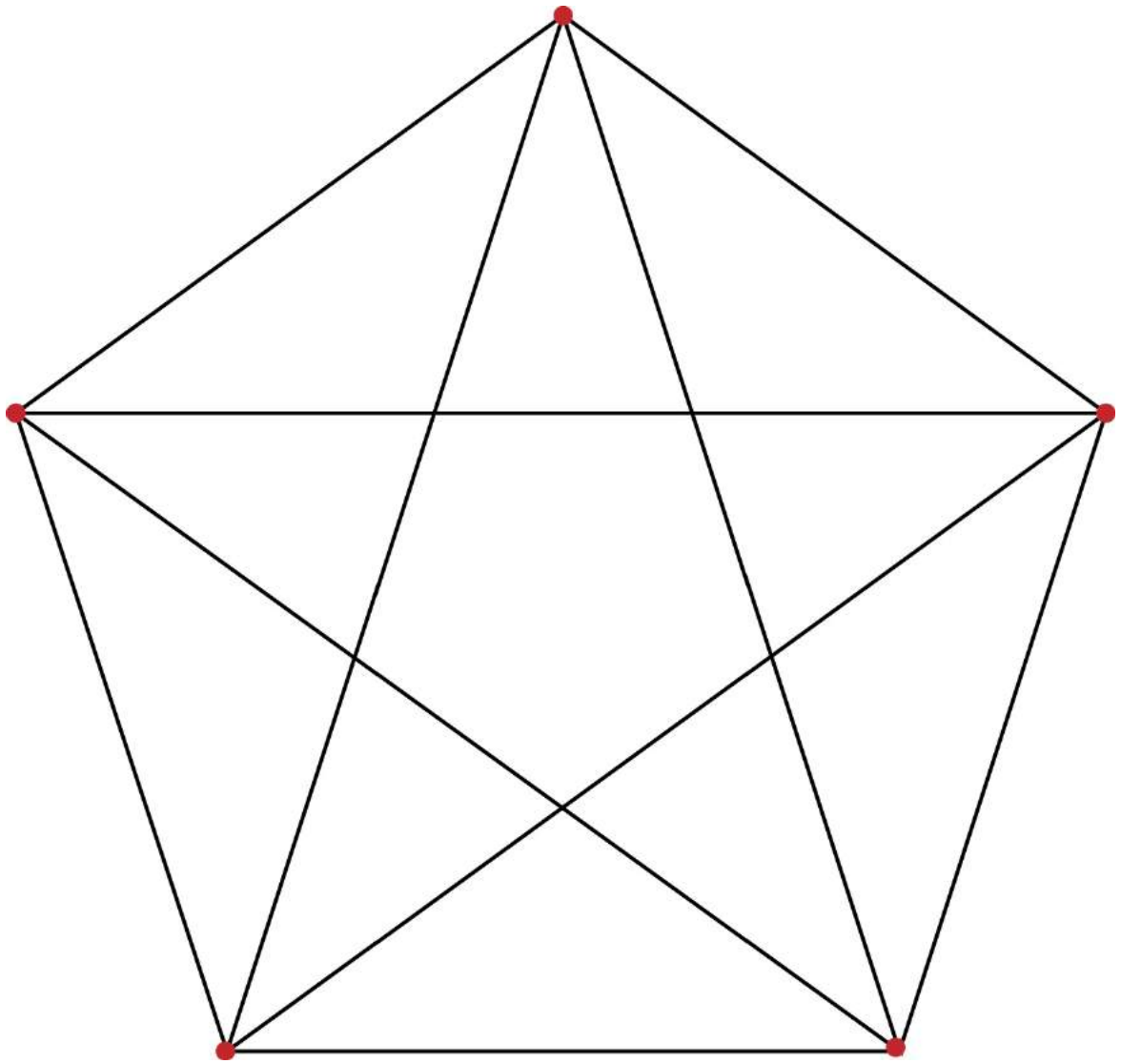


Figure 7-47. Unnatural link grouping

Using automated link dropping

Spam tactics can include the concept of creating a bot that crawls around the Web looking for open forums and blogs and leaving behind automatically generated comments. Clearly this is spam, as no human is involved in the comment process (other than the programmer), and no effort was made to read the blog post or forum where the comment was left.

The great majority of these comments are deleted or `nofollow`ed by the blog or forum software content management system (CMS), but the spammer does not care because she is operating on a large scale. If she leaves behind 1 million

comments and 99% of them are filtered by one means or another, she still ends up with 10,000 links.

This is, of course, a very risky tactic. The search engines may be able to detect this behavior algorithmically, or competitors can recognize it and turn you in via a spam report. *We do not recommend this tactic*.

Hosting giveaways

Incentive-based link requests use an incentive as part of the process of requesting a link. In general, these are links that Google does not want to count. For that reason, this is not a recommended tactic. Note that bylined articles and content syndication (as discussed in "Guest Posting" and "Content Syndication") are different strategies.

Widgets and infographics are two ways of providing a giveaway to third-party sites in return for links. However, in a YouTube video posted on August 12, 2013, Matt Cutts stated Google's recommendation that links associated with these types of giveaways should make use of the `nofollow` attribute.

One popular way of implementing widget-based link building is to develop a JavaScript-based widget and then embed a link back to the widget publisher's site in the `<noscript>` tags. Because this is the code the browser presents to users who do not support JavaScript, and the search engines do not run JavaScript, the crawlers see the link and not the widget.

An alternative method is to create an HTML wrapper for the widget with the link back to your site in clear HTML outside of the JavaScript code. This may seem like a more legitimate approach, but it is not approved by either Google or Bing for building links.

The reason for Google's position was explained in an interview that Matt Cutts did with Eric Enge:[18]

> *Any infographics you create will do better if they're closely related to your business, and it needs to be fully disclosed what you are doing. The big key is that the person publishing the infographic has to know, and agree with, including an endorsement to your site as attribution with the infographic. Even then, there is reason to believe that the link is more about the barter to get the infographic than a real endorsement of your site.*

This is similar to what people do with widgets, as you and I have talked about in the past. I would not be surprised if at some point in the future we did not start to discount these infographic-type links to a degree. The link is often embedded in the infographic in a way that people don't realize, versus a true endorsement of your site.

If you are going to pursue this tactic, then it is important that you be aware that it comes with a lot of risk. The best thing you can do to help reduce that risk somewhat is to ensure that the content is very accurate and highly relevant to your business.

As another example, Google considers the use of so-called "sponsored" WordPress templates with embedded links, even if they are visible, to be spammy as well—unless, of course, the publisher distributing the WordPress template is in the WordPress template business. The key issue to address in these types of link-building campaigns is the relevance of the widget to the site receiving the link and the editorial endorsement—or lack thereof.

Awards and Badges

With awards and badges, you can get people to link to you because it will enhance their site's value. One example is a program that gives sites awards for excellence of some kind. This is particularly effective if your site is a highly respected authority in your space. For example, *U.S. News and World Report* annually publishes a list of the best colleges and universities in the U.S.. Being included in this report brings great prestige, and many of the schools put award badges on their sites and link back to a copy of the report. This works because it is in fact a true endorsement of the *U.S. News and World Report* site.

If you are a highly respected authority, you can consider picking out the top 10 sites related to your space. Then you can send those sites an award badge with a link to a review that lives on your site.

In general, this is not a practice that Google or Bing will want to support either, unless you are already one of the top authority sites in your space and you are able to get the majority of the award winners to post the badges you provide on their sites. Otherwise, your program is simply not credible.

Customer Discount/Incentives

A publisher can offer visitors from certain websites a discount on his product. Sites that want to offer the discount to their users simply need to link back to the publisher's site.

However, Google also sees this as a questionable practice. Although it would seem that the discount offer would be interesting to the third-party site only if it valued the product (or thought its visitors would), Google's Matt Cutts clearly indicated in an interview that Google did not want to value such links.[19]

How Search Engines Fight Link Spam

Throughout this chapter, we have provided many examples of how spammers try to circumvent search engine guidelines to obtain high rankings for sites that may not deserve them. Of course, the search engines do many things to fight link spam.

Google's Penguin Algorithm

On April 24, 2012, Google released a new algorithm called Penguin. This algorithm focused on identifying sites that have undesirable link profiles and punishing them by lowering their rankings. Some areas of concentration for this algorithm are:

- Article directories
- Low-quality web directories
- Overuse of rich anchor text in links back to your site
- Links from countries where you don't market yourself

This algorithm is one of many that Google uses to fight spam, and you can read more about it in Chapter 9.

Other Algorithmic Approaches to Fighting Link Spam

The major approach the search engines use is to design algorithms that can

detect and act on link spam. There are a number of things they can look at algorithmically. Here are a representative few:

Links labeled as advertisements

The search engines can scan for nearby text, such as “Advertisement,” “Sponsors,” “Our Partners,” and so on.

Sitewide links

Sitewide linking is unnatural and should be a rare part of your link mix (purchased or not). The only exception to this is the interlinking of all the sites owned by your company, but this presumes that the search engine will understand that all of your sites are from the same company. In general, sitewide links are a serious flag, especially if you have a lot of different sites that do this for you, or if a large percentage of your links are sitewide.

Links sold by a link broker

Of course, link brokers are knowledgeable about the link detection methods listed here, and they do their best to avoid detection with the links they sell. But they can still run into problems. For example, Google took action against a long-time proponent of paid links, Internet Marketing Ninjas (formerly called We Build Pages), resulting in the company changing its stance on the subject.[20] In February 2011, Google took action against J.C. Penney for buying links,[21] and Conductor eventually exited the business of brokering buys.

A selling site providing information on how to buy a text link ad

Search engines can detect sites that provide information on how to advertise with them. Combined with other clues about links being sold on the site, this could lead to a review of the site selling the ads and a discounting of the links.

Relevance of your link

It is a powerful clue if your link is not really that relevant to the page or site it is on.

Quality of neighboring links

Another clue would be the presence of your link among a group of links that are not tightly themed, or if the other sites linked to are poor quality.

Location outside main content

The search engine can detect when your link is not part of the main content of the page—for example, if it appears in the left or right column of a three-column site, and the main content is in the middle.[22]

Perhaps you can avoid all of these pitfalls, but one more problem remains: people can see that you are buying links and choose to report your site, using Google's authenticated paid links reporting form. Note that you need to be logged in to a Google account to see the form. Here are some examples of people who might take this action:

Someone reports your site for buying links or for some other reason

Who would do this? Your competitor! If your competitor submits an authenticated spam report to Google, Google will look at it and may choose to act on it.

Someone reports the site you bought links from for selling links or for some other reason

A competitor of yours can do this, or a competitor of the site selling links can do this. Once a search engine figures out that a site is selling links, it is possible that this could trigger a deeper review of the sites that were buying those links.

A disgruntled employee leaves your company, the broker, or the site you bought links from and reports your site

For decades, many companies have had a practice of escorting fired (or laid-off) employees out of the building. The reason for this approach is that people get upset when they lose their job. However, this practice would not prevent such a person from reporting your site in a spam report to a search engine. Even though that may be a violation of the confidentiality agreement

you probably have with your employees, you would never know, because there is no transparency in spam reporting.

A search engine employee does a manual human review of your site

The search engines maintain legions of people who conduct human reviews of sites, which they use to proactively find and report spam.

Certainly your competitor reporting your site would be the most likely scenario, but you should not entirely discount the other scenarios.

Negative Link Building

Unfortunately, Google's increased ability to detect and act on links it considers to be of poor quality has resulted in a new black hat tactic: negative link building. This involves buying poor-quality links and pointing them at your competitor's website. This has proved very difficult for Google to detect and act on.

Even a simple Google search on *negative SEO* will bring up companies that offer such services in the first page of results. We will not show a screenshot of those results here, as we do not in any way endorse this practice. Not only is it unethical, but it is also potentially illegal.

However, it's important to know that this practice exists, and to take steps to protect yourself from negative link building attacks on your site. Monitoring your link profile on an ongoing basis is one of the best ways to do this.

If you see significant numbers of new poor-quality links coming in to your site, you may be the subject of a negative SEO attack. Once you identify links that you think are suspicious, you can use Google's link disavow tools to indicate that you don't want to have any association with those links. You can read more about this in "The Disavow Links tool". Two other resources on negative SEO are:

- "To Catch a Spammer: Uncovering Negative SEO"
- "How to Protect Your Website from Negative SEO"

Unnatural Links Messages

Google may also notify you via Google Search Console if a manual review of your site has detected links that it doesn't like. This is one of many reasons why signing up for Google Search Console is so important (we also recommend that you sign up for Bing Webmaster Tools). Figure 7-48 shows an example of an "unnatural links" message from Google.

View messages for http:/[redacted] »

Delete

Reconsideration request for http:/[redacted]: Site violates Google's quality guidelines

Dear site owner or webmaster of http:/[redacted]

We received a request from a site owner to reconsider http:/[redacted] for compliance with Google's Webmaster Guidelines.

We've reviewed your site and we still see links to your site that violate our quality guidelines.

Specifically, look for possibly artificial or unnatural links pointing to your site that could be intended to manipulate PageRank. Examples of unnatural linking could include buying links to pass PageRank or participating in link schemes.

We encourage you to make changes to comply with our quality guidelines. Once you've made these changes, please submit your site for reconsideration in Google's search results.

If you find unnatural links to your site that you are unable to control or remove, please provide the details in your reconsideration request.

If you have additional questions about how to resolve this issue, please see our Webmaster Help Forum for support.

Sincerely,

Google Search Quality Team

Figure 7-48. Unnatural links penalty message in Google Search Console

Other Search Engine Courses of Action

In the case of Google, one of its basic policies is to punish a site that sells text links by eliminating that site's ability to pass worth. This is essentially its first course of action. Once it's done this, Google could look more closely at the selling site and the purchasing sites for other signs of spammy behavior.

At times, Google has also taken the step of altering a site's visible PageRank (the PageRank shown on the Google toolbar) if it believes a site is selling links. It has applied this penalty to some very significant sites, such as NewsDay.com in

2008.[23] However, Google appears to have abandoned this practice since then.

Google can also choose to take action against the site purchasing the links or participating in manipulative link practices. If a large percentage of the site's links are suddenly disabled, this could have a significant impact on rankings.

Social Networking for Links

When you add social media to the equation, the network effect can multiply the yield from your link-building efforts. You can use this effect to help your content spread virally, or to develop relationships with critical influencers.

Blogging for Links

Blogging can be very effective in link development. The effectiveness of a blog depends highly on the content on it, the market space, and how the publisher promotes it. The first thing to realize when you're starting a blog is that it is a serious commitment. No blog will succeed if it does not publish content on a regular basis.

How frequently a blog needs to publish depends on its subject matter. For some blogs, one post per week is enough. For others, it really needs to be two to three times per week, or even more often.

Blogging is very much about reputation building as well. Quality content and/or very novel content is key to success. However, when that first blog post goes up, the blog will not yet be well known and will not likely have many readers, and those that do come by will be less likely to link to a little-known blog.

In short, obtaining links through your blog is a process that can take a long time, but blogging can be a very effective tool for link building. Just remember that patience and persistence are required.

One of the best places to get links to a blog is from other blogs. This is best done by targeting relationships with major bloggers and earning their trust and respect. You can read most about this by flipping back to the section titled "Building Relationships with Influencers".

Getting other bloggers to link to you can pay big benefits. Vizu, a market research company, published a study showing that 67.3% of people found what

blogs to read by following links from other blogs.

It is also important to leverage the social nature of the blogosphere. Publishers just launching a blog should try to provide a personalized response to everyone who comments on their blog. One effective way to do this is to send each of them a personalized email that shows that the comment was read. This helps to deepen the interest of the commenter, creates a feeling of personal connection, and increases the chance that the commenter will return, and possibly add more comments. Nurturing the dialog on a blog in this fashion helps that relationship grow faster.

When pursuing social media strategies, keep in mind the following best practices:

- Remember that the focus of this discussion is on "social" media sites. Your success in these environments is entirely dependent on building relationships with the other participants on these platforms. If you do something that irritates the other participants, you can quickly find yourself called out and exposed in that community.
- Make sure you are a positive contributor.
- "Give more than you get" by proactively helping to promote the great content of others. If you simply share your own content the majority of the time, you will be viewed as someone who does not offer value by association; social connections flourish with give and take.

Leveraging Major Social Media Platforms

Building relationships with other bloggers as outlined so far is effective, but there are additional ways to interact with others and in the process let them know you have content they might be interested in. Any environment in which social interactions occur is another good place to invest time and effort.

You can use social networks, such as the current market leaders—LinkedIn, Facebook, Google+, Pinterest, and Twitter—for link-building purposes without actually getting the link authority directly from one of these sites. Major forums that relate to your area of interest also represent great targets. As we will discuss in more detail in Chapter 8, the links implemented directly on these sites are

`nofollow`ed, so they do not pass direct SEO value. However, social media can help get you exposure to bloggers and media people who may then choose to cite your content or write about it, and implement links back to you. This indirect form of obtaining links can be quite powerful.

Building relationships and trust on a social media platform can drive lots of traffic and links to your site. The mechanics of making this work are fairly straightforward:

- Publish great noncommercial content related to your business on your site.
- Establish and grow social media presences on major social media platforms that promote the same types of content.
- Share your content via these platforms.

If you have done this well, sharing your newly published content through your social media channels should help your visibility and reputation grow. In turn, the exposure to the audience on that social platform should drive traffic and links back to your site. This synergy takes time to build, but will become a valuable asset; a strong social media presence is a powerful ally to any content marketing or SEO strategy.

Of course, while you’re establishing connections and building relationships in a social media environment, there’s also the potential for people to submit your content into those social media sites for you, which can lead to links as well. You can then take the next step and reach out through the social networks or forums to make direct contact with people to let them know about your content. This is similar to emailing people, but with a few important distinctions:

- You can send out communications to your friends on those networks. Assuming that you have treated the “friend” designation judiciously (rather than “friending” everybody in sight), these communications can be a bit more informal than an unsolicited email would be.
- You can also join groups on these networks related to your market space, and then send messages out to those groups. These groups will enable you to reach new people with related interests.

- Messages broadcast through these networks cannot be personalized, so you must tailor a more general message for these types of communications.
- Beware of broadcasting too many messages or poorly targeted messages. Many publishers have made this mistake, became pariahs in their communities, and lost the leverage these social networks bring to the link-building process.
- You can send personalized messages on a one-on-one basis as well.

One strategy for approaching an authority site is to make initial contact by friending someone in a senior position at the company that publishes the site. Then you can develop the relationship with that senior person without asking anything of her. Once the relationship is established, you can use a more informal approach to introduce the great content on your site.

Using Social Media Networking Effectively: Dos and Don’ts

Although we have outlined some of the basics of how to work with social media sites, each site has its own quirks and opportunities. What follows is an outline of tips to use on many of these platforms. We will discuss Facebook, Google+, Twitter, reddit, Pinterest, and a few others here briefly, but other aspects of leveraging these channels will be discussed in Chapter 8.

Nearly all social media site links are `nofollow`ed, so don’t expect these to pass link authority through to your site; however, visibility on these sites can still have a large impact on your SEO. A strong social media presence can act a lot like a powerful PR channel.

By publishing great content and then sharing it through your social media accounts, you can potentially gain fast exposure to bloggers and media (if they are following you), which in turn can lead to their implementing links to your site.

Facebook

People share links on their timeline and on Facebook pages. Links shared by users who have marked their content as available for Everyone (*http://bit.ly/fb_control*) are visible to search engines, as are Facebook Pages

(pages for companies as opposed to people). These links are all `nofollow`ed, so they do not pass PageRank, but they can provide visibility for your content.

Facebook likes are not at all visible to Google, as there is no place for it to crawl to see what content you have liked. However, Bing receives like data via an API feed, so it has the potential to use that data, though indications are that it is not doing so as of September 2014.

There are many reasons to use Facebook, but you should not expect it to directly impact your SEO ranking. However, a strong social presence can play a big role in your reputation and visibility online, and as a result can have a powerful indirect impact on your SEO. Facebook can be used to build valuable relationships.

In particular, Facebook has powerful advertising features that enable you to significantly increase your reach across the platform, and this can give you tremendous exposure.

Google+

Google+ was launched on June 28, 2011, and very quickly built a large user base. Like Facebook, Google+ also allows people to share links and generate comments and interactions on its posts. Shared content on Google+ can be evaluated based on the authority of the person sharing it, how many times it gets reshared, and the number of comments it generates.

Users can also indicate that they found a page valuable by clicking on a +1 button on the page (if the page has one) or next to the search listing for that page. The +1 button functions in much the same way as Facebook's Like button, except that Google has direct access to the +1 data, as well as data about sharing on the platform. This plays a big role in personalized search on Google.

Google+ also offers brand pages, and these are used by many major companies to communicate with existing and potential customers. In addition, these pages show up in the search results whenever you search on a company's brands, and this can be very valuable real estate. In some cases, it blocks people from using Google AdWords to bid on your brand name.

Consider focusing on connecting with major influencers on Google+, and using that connection to build a real relationship with them. Add value to their posts,

and they are likely to pay more attention to what you choose to post. This will help your Google+ account build authority and influence.

Twitter

Twitter has established itself as the leading microblogging site. It allows its members to contribute microblog posts (a.k.a. “tweets”) that are limited to 140 characters. It has become an environment for real-time communication with a broader network of people. You can use Twitter as an effective platform for promoting your business, though links from Twitter are `nofollow`ed, so no PageRank is passed. It is another valuable channel for communicating with your customers and market.

The basic goal is to become an active member of the community and build a large network of followers. As with other networking sites, many important influencers spend time on Twitter. If you can use the platform to develop relationships with these people, you can potentially gain some very high-quality links.

In addition, if you create some high-quality content, you can gain a lot of exposure through “retweets” (whereby people forward your message to all of their followers). Particularly interesting tweets can get significant visibility for you and your company.

reddit

reddit is a social link aggregation and discussion site, grouped into topical communities (“subreddits”) that are self-organized and independently moderated. There are millions of visitors each day, and reddit has hosted “ask me anything” interviews with many celebrities and a few politicians, including Barack Obama.

The site has a lot of traffic, a lot of users, and a lot of discussion. You must be a registered member of reddit (a “redditor”) in order to advertise there. It isn’t difficult to create an account, but the ad targeting is not very good because there is little capacity for users to provide revealing demographic data about themselves. In general, ad results seem to be less than stellar for any product or service that doesn’t specifically appeal to self-described geeks.

redditors are generally hostile toward self-promotion of any kind. If you try to

guerrilla market by posting links to your site or Facebook page or anything that looks remotely promotional, redditors will slap you down very, very hard.

For that reason, it is best to not plan on any form of self-promotion on reddit. All you can and should do on the site is try to participate normally and submit your viral content and other people's great content to the correct subreddits without the appearance of self-promotion.

Pinterest

Pinterest is a site for sharing pictures that has rapidly built a strong following. The demographic skews heavily toward women; they constitute more than two-thirds of the site's users. The average income level on Pinterest is also relatively high.[24]

However, even brands that are not naturally highly visible do well on Pinterest. For example, as of August 2014, the *Wall Street Journal* has more than 25,000 followers. Yet all of its "pins" are quotes either embedded over a newspaper background (see Figure 7-49) or accompanying a picture of the person being quoted.

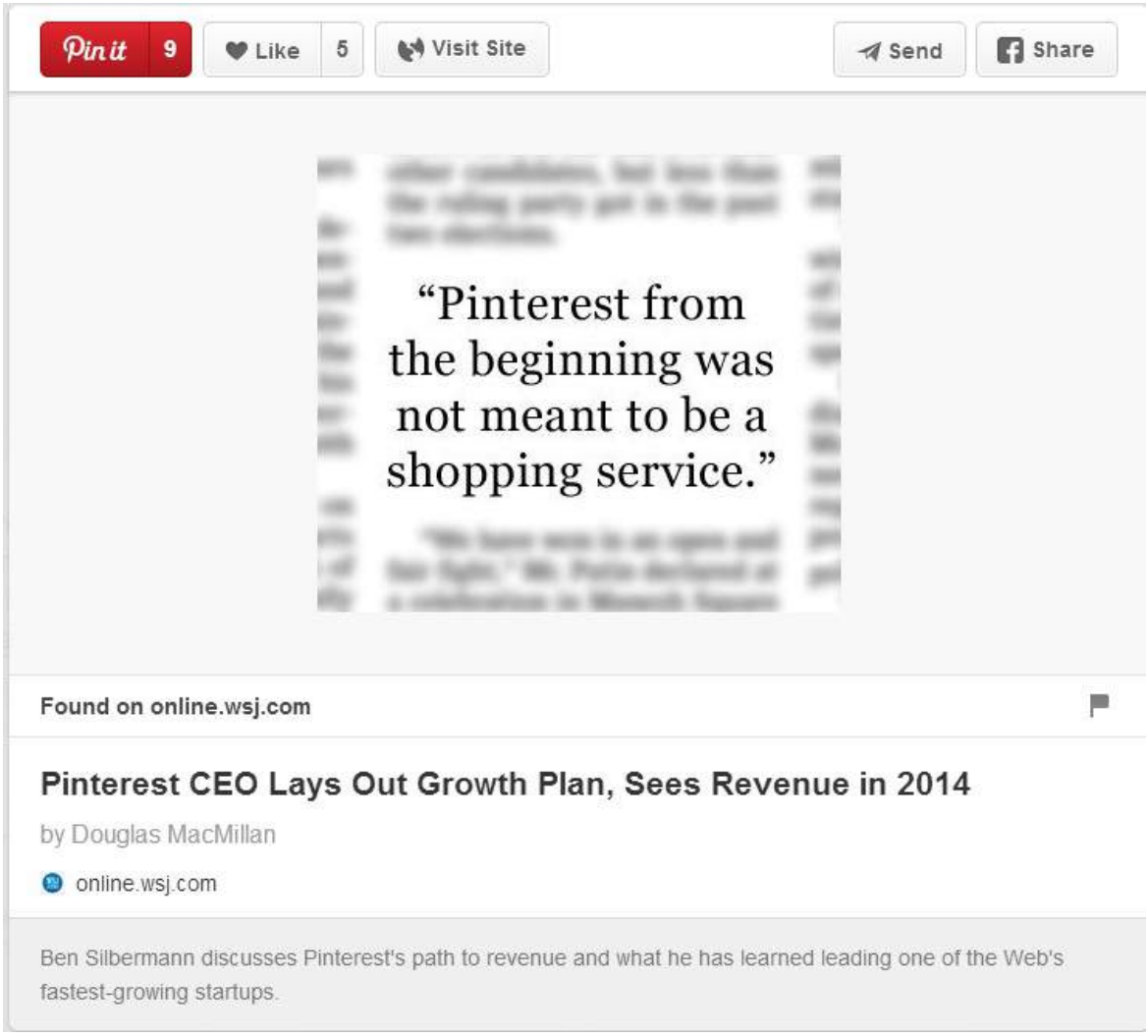


Figure 7-49. A Wall Street Journal post on Pinterest

While Pinterest has a smaller audience than Facebook, Twitter, or Google+ due to its demographics, the site has proven to work really well for certain retailers, such as apparel manufacturers or restaurants. Now could be an excellent time to jump on board and build a strong following there before your competition does!

Wikipedia

You may not think of Wikipedia as being useful in link building, because it `nofollows` all its external links. In addition, it is not wise to simply go onto the site and create a page for your company. Unless your company is well known, the page is likely to be promptly removed, and even if it is well known, you would be in violation of Wikipedia's Conflict of Interest guideline. However,

links from Wikipedia can be valuable because many people treat it as an authoritative site, and if they see your link there they may choose to link to you. This can include some pretty influential people.

Therefore, you may want to build trusted relationships within the Wikipedia community. This will require an investment in time, but it can bring some nice rewards. Here are some specific tips on how to proceed:

- Build up your credibility before doing anything that could be construed as self-serving. Develop a long and virtuous contribution history, and ideally build a user profile page that has earned Barnstar awards. It won't benefit you to be altruistic on Wikipedia unless you do it publicly. You must have a *visible* track record (e.g., squash spam, fix typos, and add valuable content, but do not do it anonymously).

- Before making an edit to an article, contact its "owner" (the main editor who polices the article) to get her blessing.

- Use Wikipedia's "Watch" function to monitor your articles of interest. Better yet, use a tool that emails you (e.g., TrackEngine or ChangeDetect).

- Make friends. They will be invaluable in times of trouble, such as if an article you care about gets an "Article for Deletion" nomination.

- Do not edit anonymously from the office. This could come back to haunt you. Tools exist that could embarrass you if you do so. One public domain tool, WikiScanner, is able to programmatically parse anonymous Wikipedia posts and identify the organization that created them. The tool cross-references the page editor's IP address with the blocks of IP addresses of more than 180,000 organizations. Do not take the risk.

If you are serious about pursuing a Wikipedia page for your business, you should also learn about the Corporate Representatives for Ethical Wikipedia Engagement (CREWE). This started as a Facebook group comprising marketing and PR people and Wikipedia editors, but it now includes Wikipedians, corporate communications professionals, academics, students, and other interested parties who are exploring the ways that PR and Wikipedia can work together for mutual benefit, defined narrowly as cooperation toward more

accurate and balanced entries.

A consortium of PR associations and agencies signed a pact in 2014 to promise to work ethically within Wikipedia's rules. A related ebook was produced by some of the participants as a guide to working with Wikipedia. The flow chart in Figure 7-50 shows how the Wikipedia posting process ideally works.

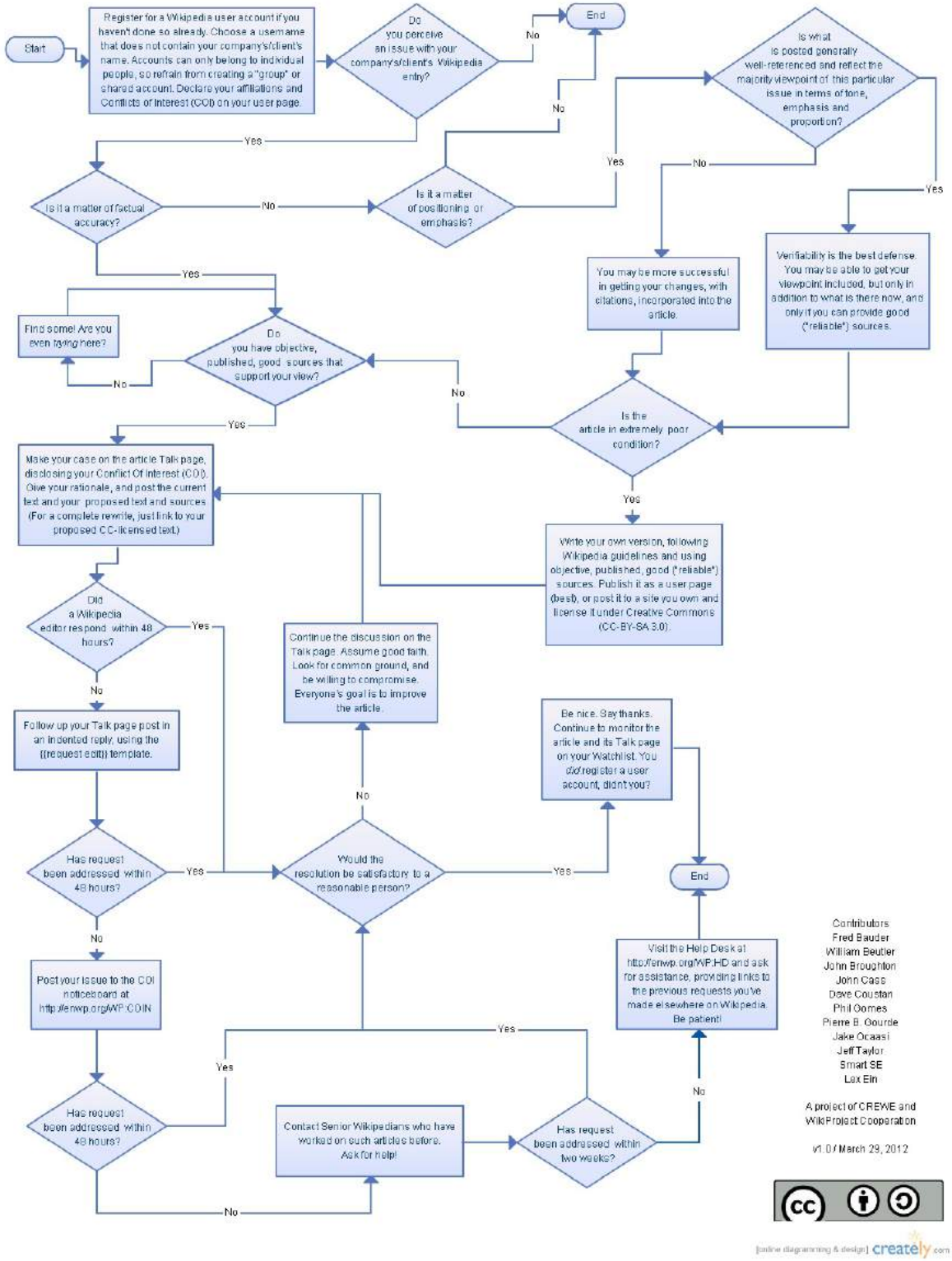


Figure 7-50. Recommended Wikipedia posting process

Wikis

Plenty of other wikis are a lot more edit-friendly than Wikipedia and let you contribute valuable content, get links, and build relationships. Examples include ShopWiki, the NewPR Wiki, and WordPress Codex. Some even pass link authority, which is a nice bonus.

Meetup.com

Meetup.com is a site designed to help groups of people with common interests "meet up." These meetups can be excellent networking events.

Get involved with relevant local meetups and get your Meetup.com member profile page linked from the Meetups page, which will pass link authority to your profile and then on to your site.

Instagram, Vine, and Snapchat

New social media sites emerge on the scene all the time, and each has its own unique attributes. Like Pinterest, Instagram is a photo-and image-sharing network that skews toward women. For the right type of brand where visual content can play a big role (such as food, beauty, lifestyle, media, and fashion), Instagram and similar sites can be great places to build your presence.

Vine is unique because it constrains you to producing six-second videos—a perfect environment for people who want to see some really compelling and quick content. Vine presents an interesting challenge because there is so little time to deliver your message, but it's very effective when you get it right.

Similarly, Snapchat videos disappear after you watch them. Like a message whose ink evaporates after you read it, Snapchat videos provide viewers a sense of exclusivity.

Using YouTube Successfully for Content Marketing

It may become more important for your brand or company to be on YouTube than to be advertised on TV. For some, that day has already arrived. comScore data from August 2011 showed that YouTube is used for about 14% of all searches on the Web.

YouTube has even launched careers, such as that of the zany Jenna Marbles,

whose YouTube channel has more than 13 million subscribers. Jenna has recorded many videos that have tens of millions of views, including:

- "How to Trick People into Thinking You Are Good Looking" (58+ million views)
- "How to Avoid Talking to People You Don't Want to Talk to" (33+ million views)
- "What Girls Do in Their Car" (33+ million views)

YouTube has also brought international fame to previously unknown musicians, such as Justin Bieber, who got his start on YouTube, or Sick Puppies, a band popularized by the hugely well-liked and inspiring Free Hugs video set to their song "All the Same", which has over 75 million views.

Then there are the hugely successful viral campaigns by commercial organizations, such as Blendtec's Will It Blend? series that we referenced previously in this chapter. This is a brilliant video series on running various household objects—including marbles, rake handles, and even iPods—through a Blendtec blender.

Blendtec isn't the only company that has had success with YouTube. WestJet scored a big hit with its video, "WestJet Christmas Miracle: Real-Time Giving". In this video, people were filmed sharing their Christmas wishes with Santa via a video feed in the airport. While the passengers were on their flight, WestJet employees rushed off to stores to buy all the gifts they had asked for. Once the passengers landed and made their way to baggage claim, they saw the first items to come around the conveyor belt were gift-wrapped packages addressed to them. They opened their gifts, and to their amazement, found what they had asked Santa for only hours before. Their reactions were also filmed, and the resulting video received over 33 million views.

Another stellar example of YouTube marketing is the online retailer Vat19.com. Its product videos are brilliant and well integrated into the company's product catalog. For example, the five-pound gummy bear is a must-see. Vat19's channel has more than 850,000 subscribers, with more than 440 million views.

YouTube has been used effectively for brand damage control as well. For example, the CEO and founder of JetBlue Airlines posted an apology video on

YouTube following a Valentine's Day winter storm incident—a campaign that was well received.

One product that got some excellent brand recognition and building from being on YouTube was Smirnoff's Raw Tea. Smirnoff produced an uproarious music video called "Tea Partay", with preppies rapping.

The problem with most popular YouTube promotions is that YouTube gets the links, and the original site usually does not. That means the search engine visibility benefits do not usually transfer to the company's website.

Even without creating your own site or hosting your videos on your main site to draw links, YouTube offers much in the way of brand visibility when the campaign is well executed. That doesn't just mean posting a great video; marketers must also know how to take advantage of the social nature of the site to build up friends and get on user subscription lists.

Make copious use of tags on your videos (ensuring, of course, that the tags are relevant to the content), spread your tags out among your clips, use adjectives to make your videos more visible to folks who are searching based on their mood, have some category descriptor tags (bearing in mind that YouTube's default search settings are Videos, Relevance, and All Categories) to match your title and description with your most important tags, and don't use natural language phrases or waste tag space on words such as *and* and *to*. See "Optimizing for Video/Multimedia Search" for more tips.

Do not be afraid to make a start, even if it is modest and has little budget behind it. You won't get anywhere without experimenting with the medium. Ongoing experimentation provides you with the best chance to find a formula that works for you.

Implementing Guest Posting Successfully

Susan Joyce is a well-known author of content on the topic of job search, and she is also a contributor to the *Huffington Post*. In March 2014, she placed an article on the *Huffington Post* titled "What 80% of Employers Do Before Inviting You for an Interview". Figure 7-51 shows us that this article was a smashing success.

Figure 7-51. An article that went viral

The article netted well over 30,000 social shares. What made this article such a success? Certainly being placed on the *Huffington Post* helped, but the title was a big key to the success of this content. It drew people in, as everybody wanted to know the answer to the question.

In fact, the answer was quite simple: employers search on your name before asking you in for an interview. Yet the simplicity of the observation did not diminish the value of the article, because the observation was not completely intuitive.

Once you see it, your reaction is, *of course they do*. For most searchers, the result is to search on themselves and see whether employers will like what they see!

Putting It All Together

To sum up the social media strategy outlined here, it is useful to think about it as an old-fashioned PR/marketing strategy. Marketers used to think about the

number of “impressions” they were able to create in people’s minds. The more impressions you can create, the better. When TV, print, and radio were the only media to worry about, this strategy was relatively simple to execute.

The current environment is significantly more complex. The people you are trying to reach use many venues, and all of the various social media properties are part of that picture. Whenever you have a social media property that a large number of people use, there is an opportunity to reach them. Success in reaching them depends on becoming a trusted member of that community.

Because there are many large communities, it can be time-consuming and expensive to establish a presence in all of them. But your participation in each one will add to the number of opportunities you have for creating impressions in your target audience. Major influencers in your market may be using these communities, so you have an opportunity to reach them as well.

Pursuing social media sites can be a very effective strategy. It does require a real investment to implement well, but when it’s done right, the strategy can provide some real benefits. You can read more about social media and search in Chapter 8.

Conclusion

Quality content will naturally attract and earn inbound links—and links remain a large factor in search rankings. Bear in mind that the best links are those you would consider valuable even if there were no search engines. These are links that can deliver traffic to your site on their own, and that are most likely to be seen as valuable by search engines in the long term. A solid content development and content marketing plan is essential to all of your online efforts, not just SEO.

You should view content development and marketing as an ongoing activity, ideally with an editorial calendar around which optimization and distribution strategies are organized. We have seen cases where a brief focus on link accumulation (with or without focused content marketing) brought returns that were squandered by a site subsequently abandoning the strategy. Unfortunately, these sites lost momentum and rankings to their competitors (the same ones they had previously passed), and it proved very difficult to catch up to them again.

Content marketing is fundamentally similar to online public relations work: your

goal is to build your visibility and reputation online, and acquire links that can potentially drive relevant traffic to your website. The big difference between content marketing and traditional PR is in the technical aspects and the focus on new media methods of promotion—with emphasis placed on the quality of the referring source, the relevance of that source, and the other page(s) to which it points.

People will generally not link to low-quality content or sites that offer a poor user experience (unless they are compensated for the link), and unless you are fortunate enough to possess a major brand, people won't link to purely commercial sites, either. You have to offer something of value to users while also offering something unique. Certain content naturally attracts links because it triggers psychological and emotional responses—pride, sharing, newsworthiness, and so on. Leverage these triggers and create a compelling reason for visitors who can influence web content (writers, publishers, bloggers, etc.) to reference your work, and your content marketing efforts will be a success—both from a visibility and earned links perspective.

Great content marketing comes from this simple idea: "Build great content, tell everyone about it, and motivate them to share."

[1] "Matt Cutts and Eric Talk About What Makes a Quality Site," Stone Temple Consulting, July 9, 2012, *http://www.stonetemple.com/mattcutts-and-eric-talk-about-what-makes-a-quality-site/*.

[2] Sergey Brin and Lawrence Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," *http://infolab.stanford.edu/~backrub/google.html*.

[3] Matt Cutts, "PageRank Sculpting," Matt Cutts: Gadgets, Google, and SEO, June 15, 2009, *https://www.mattcutts.com/blog/pagerank-sculpting/*.

[4] Brian. D. Davison et al., "DiscoWeb: Applying Link Analysis to Web Search," *http://www.cse.lehigh.edu/~brian/pubs/1999/www8/*.

[5] Mike Grehan, "Filthy Linking Rich!", *http://www.search-engine-book.co.uk/filthy_linking_rich.pdf*.

[6] Zoltán Gyöngyi, Hector Garcia-Molina, and Jan Pedersen, "Combating Web Spam with TrustRank," Proceedings of the 30th VLDB Conference, Toronto, Canada, 2004, *http://bit.ly/trustrank_paper*.

[7] Zoltán Gyöngyi, Pavel Berkhin, Hector Garcia-Molina, and Jan Pedersen,

“Link Spam Detection Based on Mass Estimation,” October 31, 2005, *http://bit.ly/mass_estimation*.

[8] For a full analysis of the Reasonable Surfer Patent, check out Bill Slawski’s May 11, 2010, blog post “Google’s Reasonable Surfer: How The Value of a Link May Differ Based Upon Link and Document Features and User Data” on SEO by the Sea, *http://bit.ly/reasonable_surfer*.

[9] Stephan Spencer, “Interview with Google’s Matt Cutts at Pubcon,” January 31, 2008, *http://www.stephanspencer.com/mattcutts-interview/*.

[10] “Matt Cutts and Eric Talk About What Makes a Quality Site,” Stone Temple Consulting, July 9, 2012, *http://www.stonetemple.com/mattcutts-and-eric-talk-about-what-makes-a-quality-site/*.

[11] Duane Forrester, “10 SEO Myths Reviewed,” Bing Blogs, May 9, 2014, *http://blogs.bing.com/webmaster/2014/05/09/10-seo-myths-reviewed/*.

[12] Eric Enge, “Twitter Engagement Unmasked: A Study of More than 4M Tweets,” Stone Temple Consulting, December 11, 2014, *https://www.stonetemple.com/twitter-engagement-umasked/*.

[13] Mack Collier, “The Power of Being Second: How Red Bull is Winning the (Content) Marketing Wars,” *http://mackcollier.com/red-bull-content-marketing/*.

[14] Gus Lubin, “McDonald’s Twitter Campaign Goes Horribly Wrong #McDStories,” January 24, 2012, *http://www.businessinsider.com/mcdonalds-twitter-campaign-goes-horribly-wrong-mcdstories-2012-1*.

[15] Scott J. Slavick, “United States: I Can Haz Copyright Infringement? Internet Memes and Intellectual Property Risks,” Mondaq, November 28 2012, *http://bit.ly/memes_copyright*.

[16] Matt Cutts, “The Decay and Fall of Guest Blogging for SEO,” Matt Cutts: Gadgets, Google, and SEO, January 20, 2014, *https://www.mattcutts.com/blog/guest-blogging/*.

[17] You can read more on the theory of spam detection in Zoltán Gyöngyi, Pavel Berkhin, Hector Garcia-Molina, and Jan Pedersen’s “Link Spam Detection Based on Mass Estimation,” October 31, 2005, *http://bit.ly/mass_estimation*.

[18] “Matt Cutts and Eric Talk About What Makes a Quality Site,” Stone Temple Consulting, July 9, 2012, *http://www.stonetemple.com/mattcutts-and-eric-talk-*

about-what-makes-a-quality-site/.

[19] Stone Temple Consulting, “Matt Cutts Interviewed by Eric Enge,” June 16, 2008, *https://www.stonetemple.com/mattcutts-interviewed-by-eric-enge/*.

[20] Jim Boykin, “Paid Links Aren’t Worth It to Me,” Internet Marketing Ninjas Blog, November 21, 2008, *http://www.internetmarketingninjas.com/blog/link-building/paid-links-arent-worth-it/*.

[21] Vanessa Fox, “Google’s Action Against Paid Links Continues: Overstock & Forbes Latest Casualties; Conductor Exits Brokering Business,” Search Engine Land, February 24, 2011, *http://bit.ly/paid_links_google*.

[22] Jeffrey A. Dean, Corin Anderson, and Alexis Battle, “Ranking Documents Based on User Behavior and/or Feature Data,” *http://bit.ly/patent_7716225*.

[23] Barry Schwartz, “Google Confirms NewsDay.com Received PageRank Penalty for Selling Links,” Search Engine Roundtable, July 11, 2008, *http://bit.ly/newsday_penalty*.

[24] Cooper Smith, “This Is The Behavior on Pinterest That Makes the Social Network So Attractive to Marketers,” Business Insider, April 17, 2014, *http://bit.ly/pinterest_behavior*.

CHAPTER EIGHT

How Social Media and User Data Play a Role in Search Results and Rankings

Search engines continually seek to improve the relevance of their search results. They do this by tuning the weight and mix of the types of ranking signals they currently employ, or by implementing new signals. However, how they use these signals is changing all the time. Starting in 2010, evidence mounted to suggest an increasing weight on ranking signals from social media sources. In December 2010, Google and Bing both confirmed this in response to questions from SearchEngineLand.com editor Danny Sullivan.[1]

However, the way the search engines use social signals has changed significantly since then, and it currently appears that neither Google or Bing use them as a direct ranking factor, although they can impact personalized search results in Google.

Although the search engines obscure how their algorithms work, many people believe that user engagement signals are part of these algorithms, and there has long been a debate over whether search engines have treated these signals as ranking factors.

While search engine algorithms continue to evolve rapidly, adding new types of ranking signals to them is a tricky process that requires a tremendous amount of testing. As we discussed in Chapter 2, the Web consists of hundreds of trillions of pages with fundamental differences in their construction and content. In addition, the needs these pages serve, and the ways that users interact with them, are equally varied.

Search engines generally perform many levels of testing of new algorithms, including live tests with a small portion of their users prior to any general release of new ranking factors; you can read more about all the stages of testing in this interview with Google's Director of Research, Peter Norvig. In the same interview, Norvig suggests that Google makes two to three algorithm changes per day. This means change is a constant.

The key to success in search engine rankings is to remember the fundamental goal of the search engines: to increase market share by providing the best possible (i.e., most relevant and engaging) results for end users. Astute Internet marketers who understand this will focus on providing high-quality website

content and experiences for their visitors.

Correlation Between Social Signals and Google Rankings

In 2013, Moz did a study that showed a strong correlation between Google +1s and rankings in Google. The detailed results are shown in Figure 8-1.

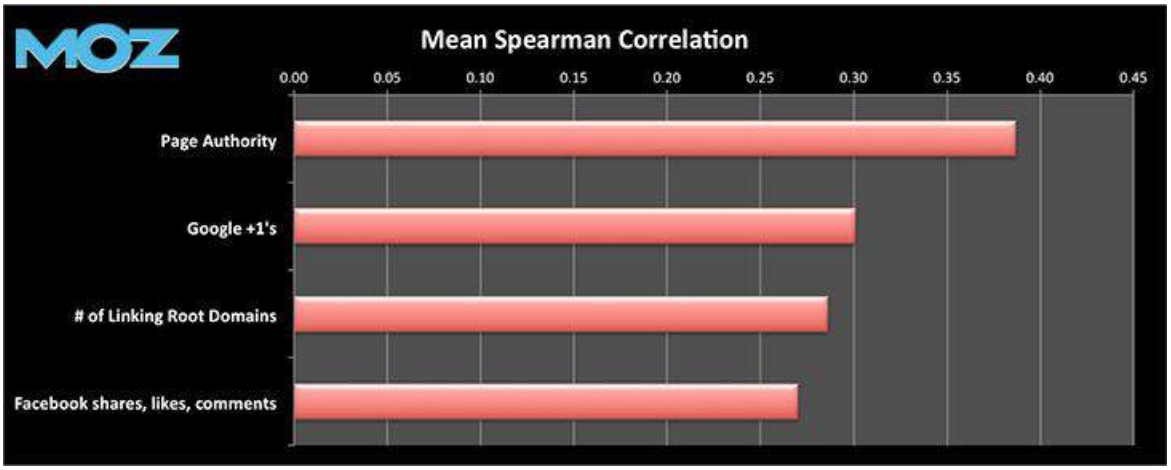


Figure 8-1. Moz correlation study

Searchmetrics performed its own correlation study. As shown in Figure 8-2, the results showed a very strong correlation between many different types of social signals and search ranking. In fact, 7 of the top 10 factors correlating with ranking were related to social media sites.

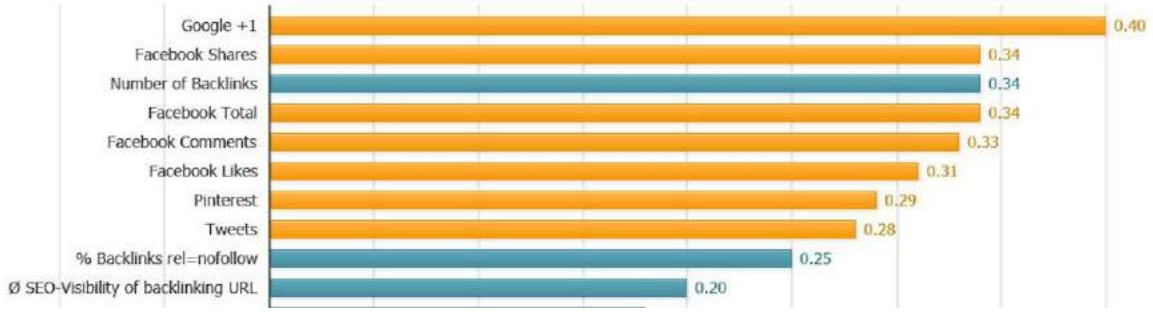


Figure 8-2. Searchmetrics correlation study

It is important to note that a correlation does not prove causation—in other words, this information does not prove that these social media factors are a direct ranking signal.

Nonetheless, the correlation is noteworthy and only supports our long-standing belief that great content has a strong chance of earning a high number of traditional web links, and of being shared more in social media.

What Is the Value of Social Signals?

Google was first established in 1996 based on the concept of PageRank, where links from third-party websites would serve as votes for the quality of the site receiving the links. This was a valuable concept because the Web evolved as an environment where creating great content was rewarded through links and the cross-referencing of other great content—a natural result.

The collection of data on all the links of the Web, which is referred to as the *link graph*, is an enormous computational task that maps who links to whom in a vast diagram of the Web's interconnectivity. Figure 8-3 shows a conceptual representation of the link graph. From this, search engines can extract valuable signals to identify relevant web pages in response to a particular search query.

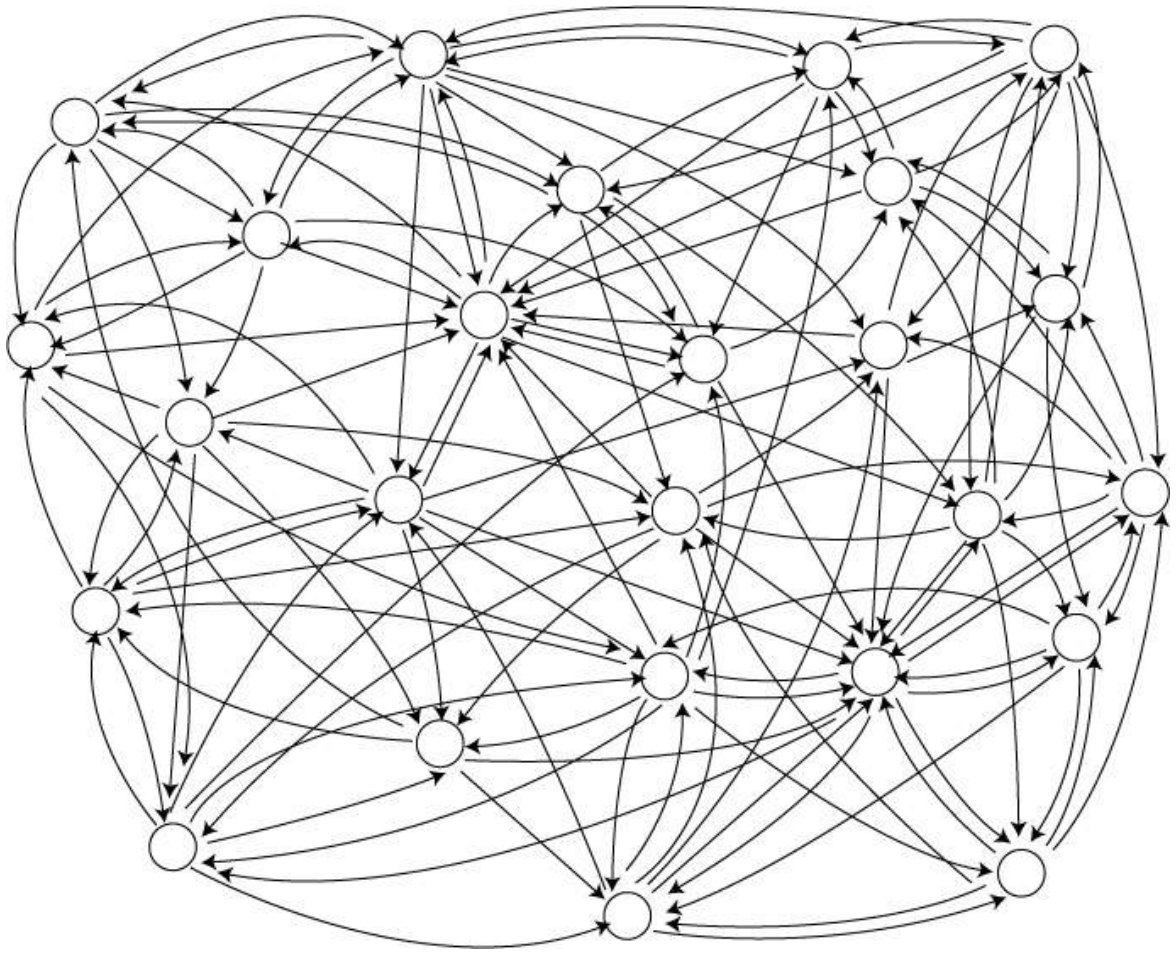


Figure 8-3. The link graph is a map of all of the World Wide Web's hyperlinks

This system in its original form, however, was dependent on the links originating from editorial citations. In other words, it assumed that publishers created a link on a third-party website because they genuinely cared to send visitors from their own site to the linked site solely because it was a great resource. This was a publisher's primary incentive for providing the link. In fact, links were the first social signal used by search engines to aid in ranking web pages.

Over time, of course, some publishers who realized there was a lot of money to be made from high search rankings began cheating the system by purchasing or soliciting links that were not citations at all. The proliferation of this link-selling economy and schemes for getting links that are not earned based on the content's merit, generally referred to as *link spam*, has threatened search quality and resulted in a war between search engines and link spammers.

Similar to the way in which the link graph acts as a proxy for real-world connections, social connections and user behavior patterns reflect real-world relationships and values. In order to implement a link to a website, one must first have a web property. While creating a website can in fact be free on today's Web, one must commit time and energy to that site to be viewed as a valuable resource. As a result, many people do not make that investment.

Participating in social media platforms such as Facebook, Google+, or Twitter requires zero cash outlay (beyond an Internet connection and a connecting device) and only a small amount of time investment. As a result, social media is more populist in nature than the link graph. Of course, the use of social media is not universal either, but it is far more broadly adopted than the publication of websites.

As users, we like proof from our peers that something is good. If your friend liked it, and you trust her judgment, you become more inclined to check it out yourself. For example, your trusted friend is unlikely to have endorsed a spam site.

However, problems arise when you try to use these social signals as a ranking factor in the search results. The search engines experimented with these signals and pulled back because they found problems with it. There are some ways in which social information is currently integrated into the search results, however, and these will be explored in the next few sections.

Bing's Experiments with Social Signals

Bing has relationships with many social media platforms. These relationships provide Bing with real-time access to any updates that occur on those networks. For example, when someone posts an update on his Facebook page, Bing gets notified of that update via its API, which means it gets the data relatively quickly, and with minimal overhead. Social sites that Bing has such relationships with include:

- Facebook
- Twitter

- Foursquare
- Quora
- Klout

In contrast, Google has to crawl Facebook to obtain comparable social update info, and only as of May 19th, 2015 did Google start to get access to Twitter data via API. This is what makes the experiments that Bing has conducted interesting to review in detail.

For example, Bing experimented with using data from Facebook as a ranking factor in the search results. Figure 8-4 shows how Bing's search results looked in 2010 if you performed a search on the phrase *New York Post*.

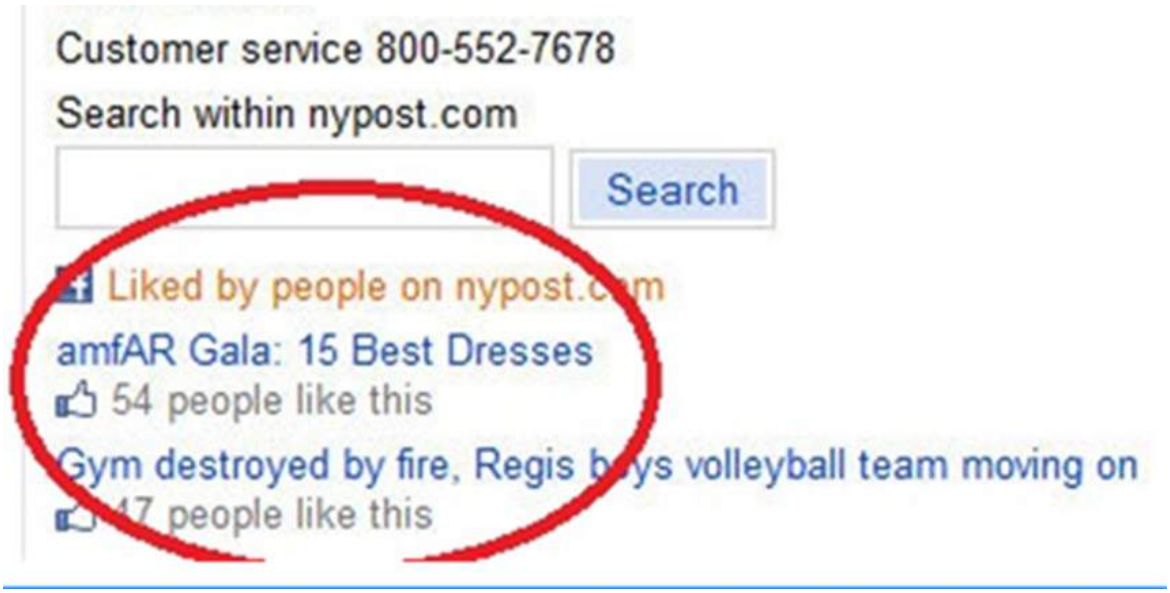


Figure 8-4. Bing search results from 2010 for "New York Post"

Notice how two articles with a number of Facebook likes are highlighted in the results. However, if you performed the same search in 2011, the nature of the result had changed quite a bit, as you can see in Figure 8-5.

Figure 8-5. Bing search results from 2011 for "New York Post"

The result is quite different: Bing has dispensed with showing the most liked articles and instead highlights a recent post from the newspaper. While it is not indicated in the search result, it is possible that this article has received recent Facebook likes, but we don't know that for sure. Figure 8-6 shows us, however, that this ambiguity has been eliminated in Bing's results for the same search in 2012.

Figure 8-6. Bing search results from 2012 for "New York Post"

In a February 2013 interview by Eric Enge with Bing's Stefan Weitz,[2] Weitz had this to say about whether Bing used likes as a ranking signal:

> *Not now, no. The notion of a Like is still a little bit perplexing from a ranking perspective. What does a Like mean for a page? Does the user like the design,*

the content, or maybe just the picture? We tend to not just use a pure Like signal to do ranking.

He also said:

Shares are basically the same as likes—not used a ton for web ranking except in velocity (like Twitter).

For a while, Bing went in a completely different direction by separating out the social part of its search results into a sidebar. Figure 8-7 illustrates this implementation.

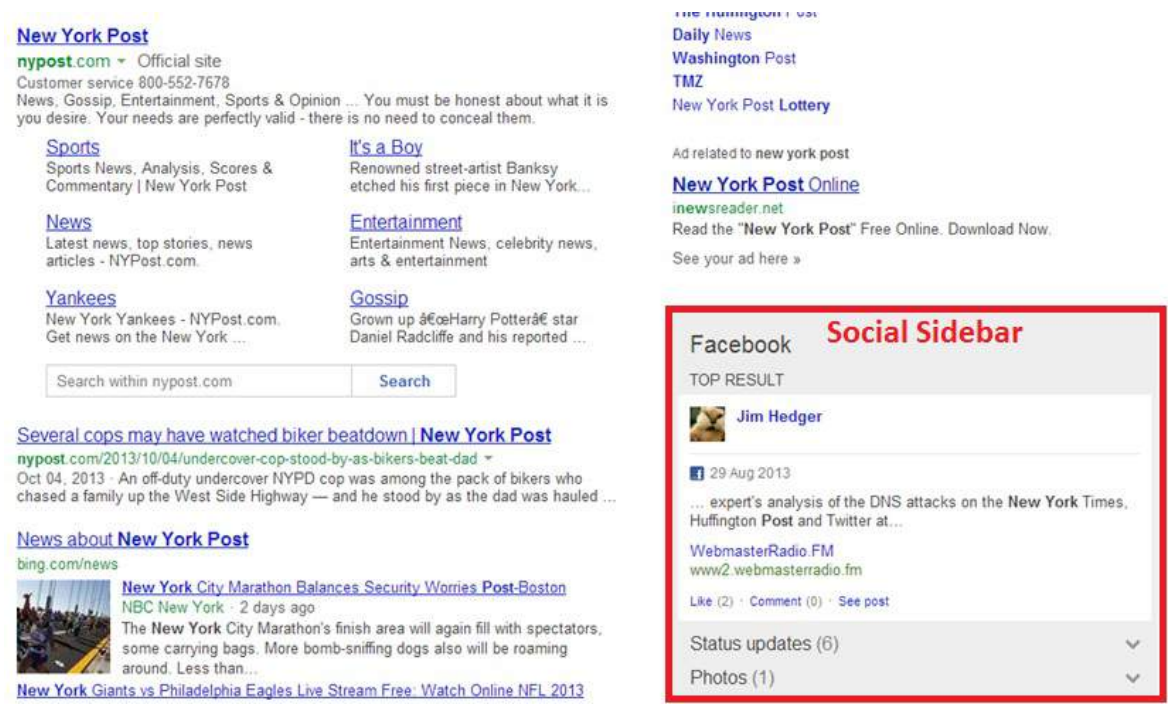


Figure 8-7. Bing's social sidebar

Does Google Use Facebook as a Ranking Signal?

As previously mentioned, unlike Bing, Google must crawl Facebook to get data on activity on the site. Consequently, Google will not be able to get any data from pages that are not marked as Public by the profile owner. That is a very large crawling burden. To determine how many pages Google indexes from Facebook, Stone Temple Consulting performed a study, published in December 2013. We can see how many pages Google has indexed from Facebook in Figure 8-8.

Figure 8-8. Facebook pages indexed by Google

Google has indexed 5.14 billion pages from Facebook according to this query. To put this in perspective, data released by Facebook in February 2014 indicated that it had 1.2 billion users, and that 10% of these put out daily updates, and 4% put out more than one update per day.

We can use these figures to estimate the number of total Facebook updates per month. We can start by taking 10% of the total number of users, and multiplying that by 30 to account for all the users who do at least one update per day. This works out to 3.6 billion updates per month.

However, we also know that 4% of users perform multiple updates, and we can estimate their usage by guessing that they average three updates per day. Of course, some users perform only two, others do much more than three.

For purposes of this analysis we are going to estimate the total number of additional updates per day as the number of users times 2 times 30. Note that we

use 2 in this calculation instead of 3, as we have already counted 1 daily update per user in our prior calculation. The result is another 2.88 billion updates per day.

This gives us an average of 6.48 billion Facebook updates per day, which is already more than Google has indexed from Facebook in total! The study also tracked the monthly pages Google indexed from Facebook from January 2012 through March 2014, as shown in Table 8-1.

Table 8-1. Facebook indexed pages by month

Month	Indexed pages	% of estimated updates per month
March 2014 (through 3/29)	77,900,000	1.91%
February 2014	4,330,000	0.11%
January 2014	5,490,000	0.13%
December 2013	7,840,000	0.19%
November 2013	7,160,000	0.18%
October 2013	8,510,000	0.21%
September 2013	4,160,000	0.10%
August 2013	26,200,000	0.64%
July 2013	48,400,000	1.18%
June 2013	168,000,000	4.11%
May 2013	166,000,000	4.06%
April 2013	166,000,000	4.06%
March 2013	3,160,000	0.08%
February 2013	2,770,000	0.07%
January 2013	1,950,000	0.05%
December 2012	2,740,000	0.07%
November 2012	2,690,000	0.07%
October 2012	2,630,000	0.06%
September 2012	1,550,000	0.04%
August 2012	2,570,000	0.06%
July 2012	2,490,000	0.06%
June 2012	2,440,000	0.06%
May 2012	1,030,000	0.03%
April 2012	2,510,000	0.06%

March 2012	2,150,000	0.05%
February 2012	1,860,000	0.05%
January 2012	856,000	0.02%

As you can see, the indexing of this content appears to happen at a very low rate. Note there are some sources of error in this data, as the total number of Facebook updates has been increasing over time, and the *site:* operator is not a precise measure in Google. But these errors do not impact the basic conclusion: Google does not index the overwhelming majority of Facebook posted content.

Also of significance is that links from Facebook to external pages use the `nofollow` attribute (which is defined in Chapter 6 and highlighted in Figure 8-9). This means that these links do not pass PageRank to the pages receiving the link.

```
straighten some of it out. <a href="http://www.facebook.com/l.php?
u=http%3A%2F%2Fabcn.ws%2F160Adhb&amp;h=eAQHoTLng&amp;enc=AZOadFETozVrpq2mMdQNDiVIfg5
bnA18ih22J9yCxweS1MiRbNYt2z6sOz4slJIlveFbw9EIINA93qHMR_Hx4fWDkmDgUcTpXmy2MLCUxORmi7B
oJ5GHVk5bUHCdwcu8EKP5UwamDJX7apbo2bkYR68XvkLhO8rddM-Lq0CqXHNwwQ&amp;s=1"
target="_blank" rel="nofollow" onmouseover="LinkshimAsyncLink.swap(this,
&quot;http:\/\/abcn.ws\/160Adhb&quot;);"
onclick="LinkshimAsyncLink.referrer_log(this, &quot;http:\/\/abcn.ws\/160Adhb&quot;,
&quot;\/si\/ajax\/l\/render_linkshim_log\/?
u=http\u00253A\u00252F\u00252Fabcn.ws\u00252F160Adhb&amp;h=eAQHoTLng&amp;render_veri
fication=0&amp;enc=AZOadFETozVrpq2mMdQNDiVIfg5bnA18ih22J9yCxweS1MiRbNYt2z6sOz4slJIlv
eFbw9EIINA93qHMR_Hx4fWDkmDgUcTpXmy2MLCUxORmi7BoJ5GHVk5bUHCdwcu8EKP5UwamDJX7apbo2bkYR
68XvkLhO8rddM-Lq0CqXHNwwQ&quot;);">http://abcn.ws/160Adhb</a></div></div></div>
```

Figure 8-9. Facebook links are nofollowed

One last consideration is whether Google might place more emphasis on indexing posts from highly influential people. Perhaps it places more weight on shares by people and brands with authoritative profiles. To test this theory, Stone Temple Consulting performed a study.

This study looked at 40 Facebook posts done by 85 influential users over a one-year period. The users examined each had at least 125,000 likes for their page. The tested posts were spread over differing time intervals as follows:

- 10 most recent posts
- 10 posts at least 3 months old
- 10 posts at least 6 months old
- 10 posts at least 12 months old

Each post was tested to see if Google had it in its index. Figure 8-10 shows the results. In aggregate, about 59% of these posts were indexed.

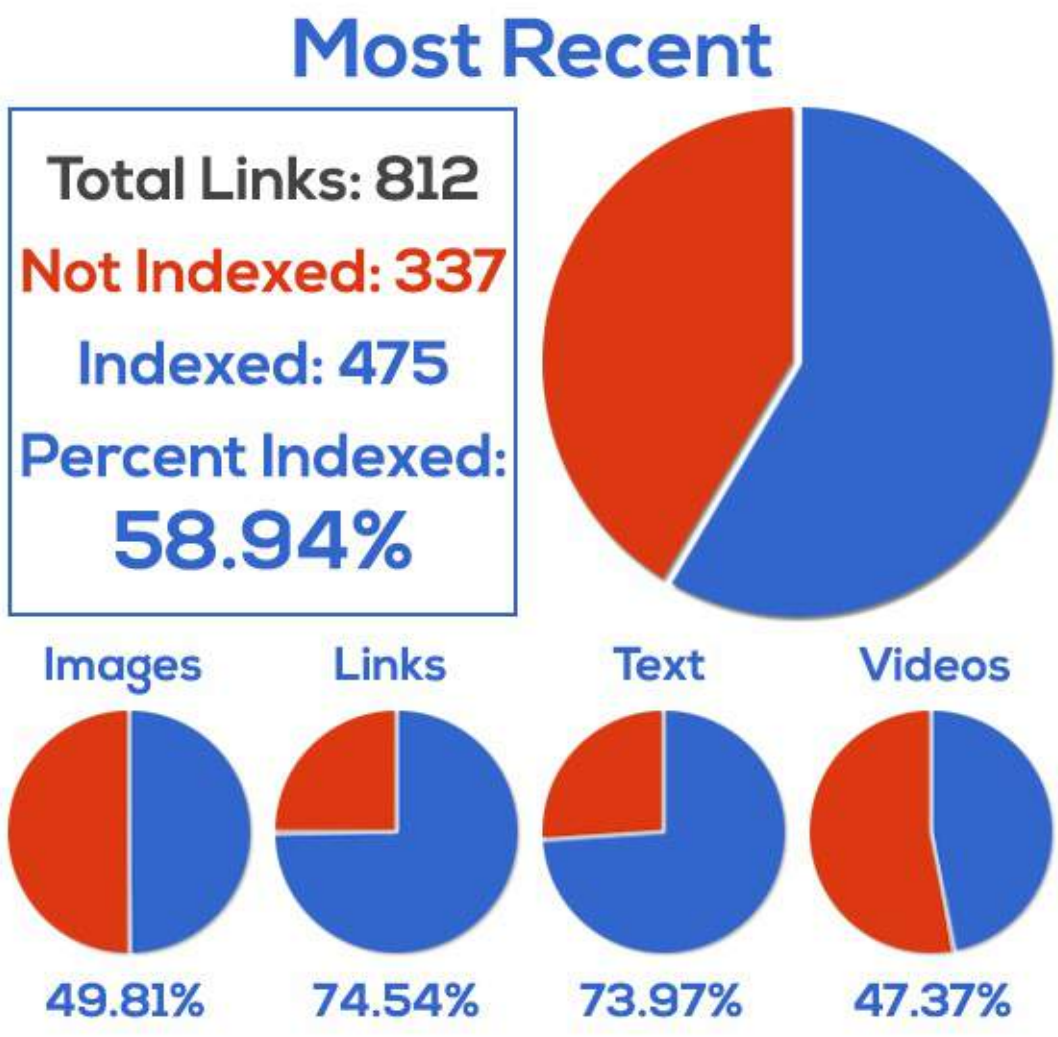


Figure 8-10. Facebook pages indexed by Google

Note that about 74% of the text posts were indexed and 75% of the posts with

links in them were indexed. If these were a major ranking factor, it seems likely that Google would be indexing a larger percentage of these posts.

This result, combined with Bing's stance on using Facebook data in search, suggests that Facebook posts are not a major factor in driving the search results. This could possibly change in the future, however, as search engines continuously experiment with new ways to use different types of signals as a ranking factor.

Does Google Use Twitter as a Ranking Signal?

An analysis of Twitter usage shows some similar results to those we saw with Facebook. In Twitter's IPO filing, the company reported that it is handling more than 500 million tweets per day on average. Figure 8-11 shows a combination of two Google search queries that attempt to find out how many Twitter status updates Google has in its index.

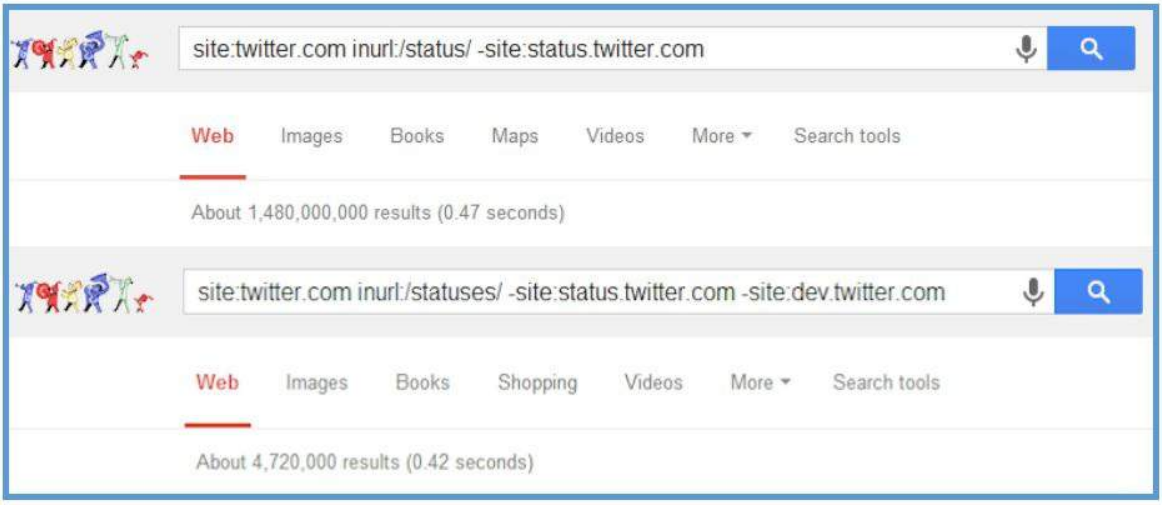


Figure 8-11. Twitter pages indexed by Google

This shows a total of 1.27 billion pages, which is a pretty small number when you consider that there are 500 million tweets per day.

In another Stone Temple Consulting study by coauthor Eric Enge, the search query in Figure 8-11 was repeated, but it included the *daterange:* operator to see how many new pages Google kept in its index on a monthly basis from January

2012 through June 2014. That data is shown in Table 8-2.

Table 8-2. Twitter indexation by month

Month	Indexed pages	% of 500M
Jan-12	494,000	0.10%
Feb-12	1,230,000	0.25%
Mar-12	1,330,000	0.27%
Apr-12	1,220,000	0.24%
May-12	669,000	0.13%
Jun-12	1,720,000	0.34%
Jul-12	2,360,000	0.47%
Aug-12	2,660,000	0.52%
Sep-12	1,850,000	0.37%
Oct-12	2,520,000	0.50%
Nov-12	1,870,000	0.37%
Dec-12	1,530,000	0.31%
Jan-13	924,000	0.18%
Feb-13	1,590,000	0.32%
Mar-13	1,640,000	0.33%
Apr-13	1,570,000	0.31%
May-13	924,000	0.18%
Jun-13	1,540,000	0.31%
Jul-13	1,420,000	0.28%
Aug-13	1,320,000	0.26%
Sep-13	777,000	0.16%
Oct-13	1,480,000	0.30%
Nov-13	1,530,000	0.31%
Dec-13	1,620,000	0.32%
Jan-14	1,030,000	0.21%
Feb-14	1,920,000	0.38%
Mar-14	2,270,000	0.45%
Apr-14	3,270,000	0.65%
May-14	2,750,000	0.55%

Jun-14	9,090,000	1.82%

This data suggests that the indexing rate of tweets is actually quite low. However, as we mentioned with Facebook, the number of results from *site:* queries is not always that precise. The study also included an analysis of the indexation of posts for 963 different Twitter accounts. Of these, 58 had at least 100,000 followers. These were hand-selected to be highly influential people. The study then tracked the indexing of their tweets over time. The results are shown in Figure 8-12.

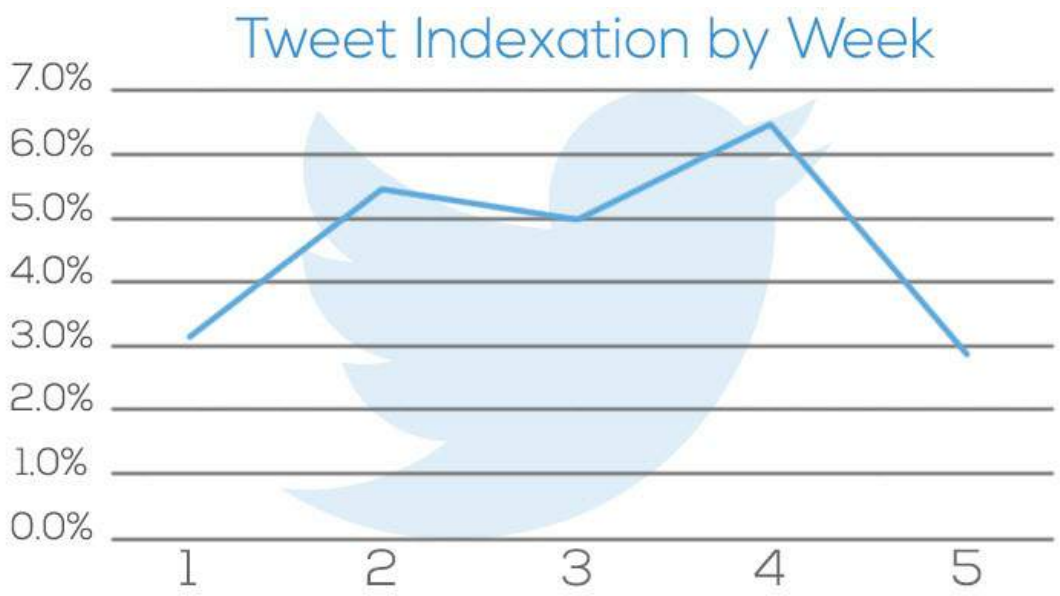


Figure 8-12. Indexation of tweets over time

Taken together, Table 8-2 and Figure 8-12 show us that Google's overall indexing of tweet content is low, but it does index a high percentage of tweets from more influential accounts. Note that it is possible that this is because these tweets tend to receive links. This would increase Google's desire to index the content in the tweet.

Figure 8-13 shows the source code for a tweet containing a shared link. As with Facebook, this link is `nofollow`ed, so no PageRank is passed.

```
<a href="http://t.co/QrXWcFnluK" rel="nofollow" dir="ltr"
data-expanded-url="http://ow.ly/yB9Gj" class="twitter-
timeline-link" target="_blank" title="http://ow.ly/yB9Gj" >
<span class="tco-ellipsis"></span><span
class="invisible">http://</span><span class="js-display-
url">ow.ly/yB9Gj</span><span class="invisible"></span><span
class="tco-ellipsis"><span class="invisible"> </span>
</span></a> by <a href="/stephanhov" class="twitter-atreply
pretty-link" dir="ltr" ><s>@</s><b>stephanhov</b></a>
```

Figure 8-13. Twitter links use the nofollow attribute

As of May 19th, 2015, Google started accessing Twitter data via the API, but a study published by Eric Enge in July of 2015 showed that increases in indexation of tweets had occurred, but the total indexed tweets remains small. In summary, the evidence suggests that Google does not currently use activity in Twitter as a ranking signal.

Does Google Use Google+ as a Ranking Signal?

Google+ is the social network owned by Google itself. Unlike Facebook and Twitter, Google obtains information from Google+ in real time. As soon as you complete a share in Google+, Google knows it has happened. This makes it seem far more likely that Google would use this data as a ranking factor in its search results. In fact, there are a few clear ways that Google+ impacts the Google search results.

Google+ Personalization

Personalization plays a significant role in Google's search results, and Google+ can be a part of that. For example, if someone is following you on Google+, the chances that she will see your posts in Google's search results go way up. Figure 8-14 shows an example of this.

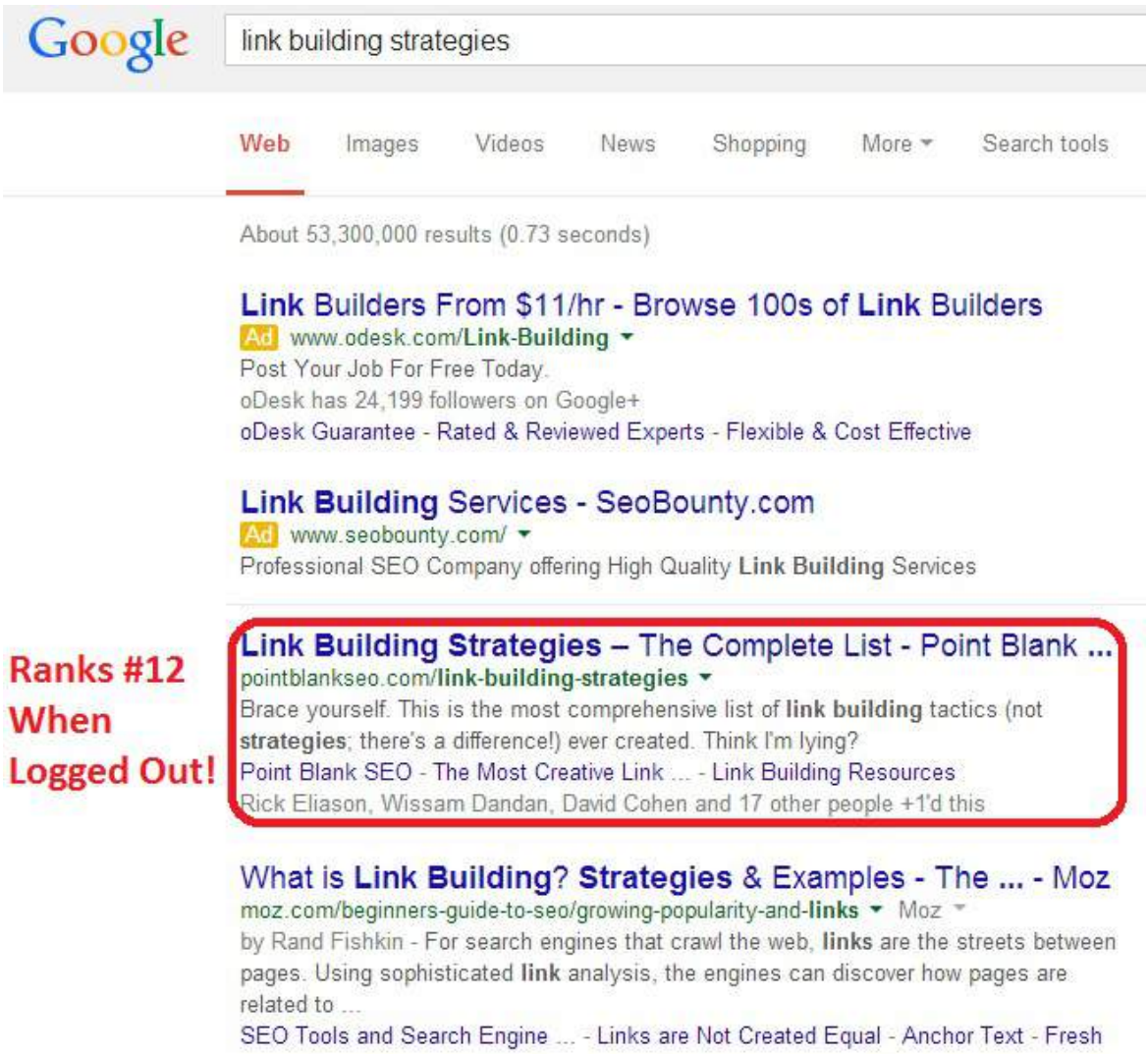


Figure 8-14. Google+ personalization in action

The highlighted listing ranks #1 in the organic search results when the searcher is logged in, but ranks only in position #12 when the same user performs an *incognito* search (in Google's Chrome browser, this mode turns all personalization off). This is a big impact! The reason for the change in the personalized results is that four people whom this user follows on Google+ have +1'ed the content.

The impact of this on your business can be quite substantial. Consider the post shown in Figure 8-15.

The great majority of these +1s are likely to be from people interested in search,

and for all of these users, this article will rank higher in the search results. But it is also more likely to rank higher for everyone they follow! This means that there are two bites at this apple already:

- People who follow the author on Google+ are more likely to see it
- People who follow anyone who +1's the content are more likely to see it

This is already a very big deal. If you can build a strong audience of your own, or if major influencers in your market space +1 or share your content, you can get a lot of exposure to your target market.

In addition, it appears that people you exchange Gmails with can also create the type of connection that leads to the personalization of results.[3]

Figure 8-15. A post with many +1s

Google+ Posts in the Search Results

Google+ posts themselves can also rank in the search results, as you can see in Figure 8-16.

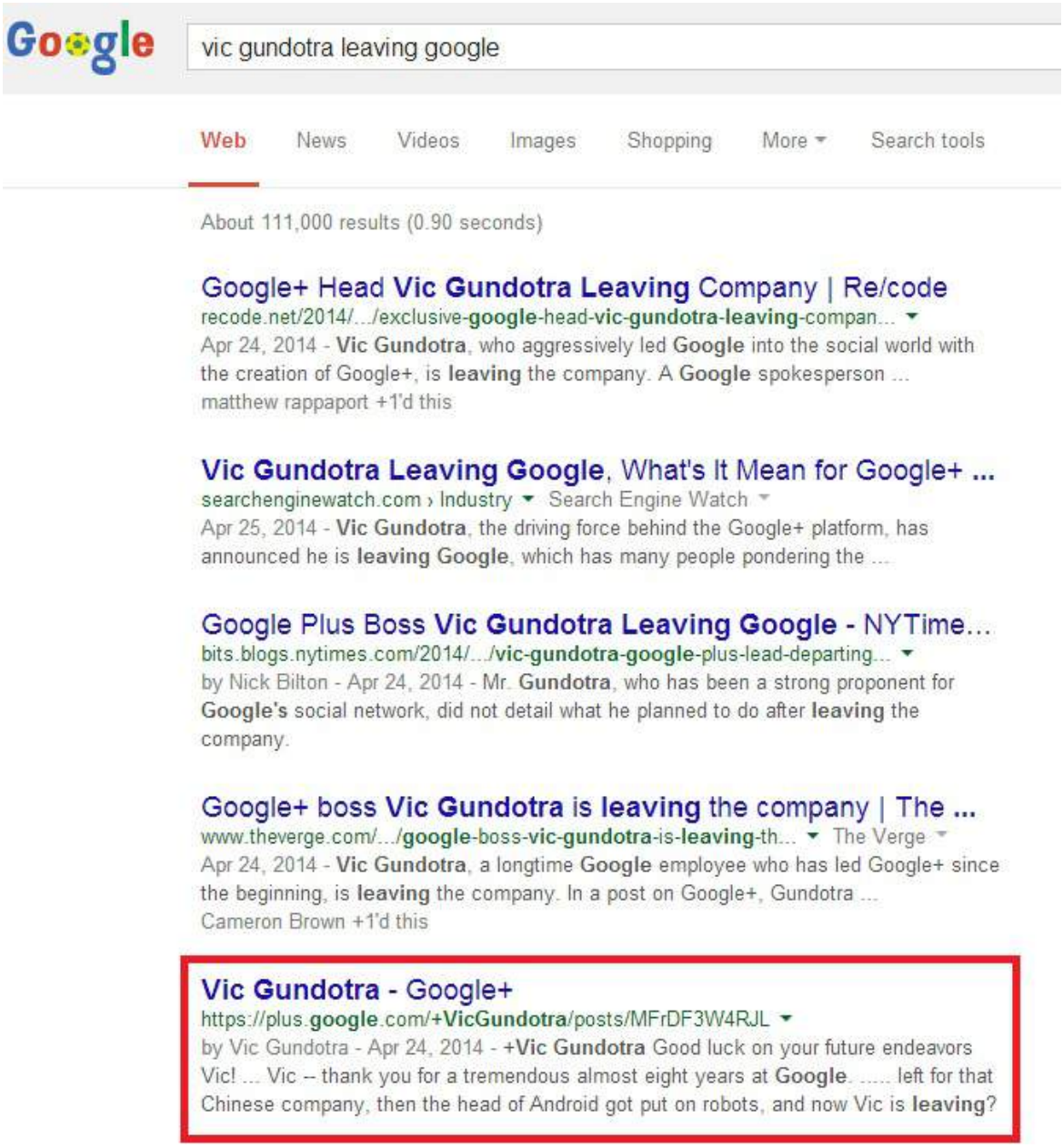


Figure 8-16. Google+ post in search results

This post came up even though the search was performed in incognito mode, so all personalization of the results is turned off.

Google+ Brand Pages in the Search Results

Another benefit of Google+ is that your company (or institution) can create a brand page. If you properly connect that brand page with your website using the

`rel="publisher"` tag (which is covered in more detail in Chapter 6), and you have some basic level of activity on the page, it can show up in the search results as well.

This is particularly likely to happen on a search for your own brand name, as shown in the search for *Major League Soccer* in Figure 8-17.

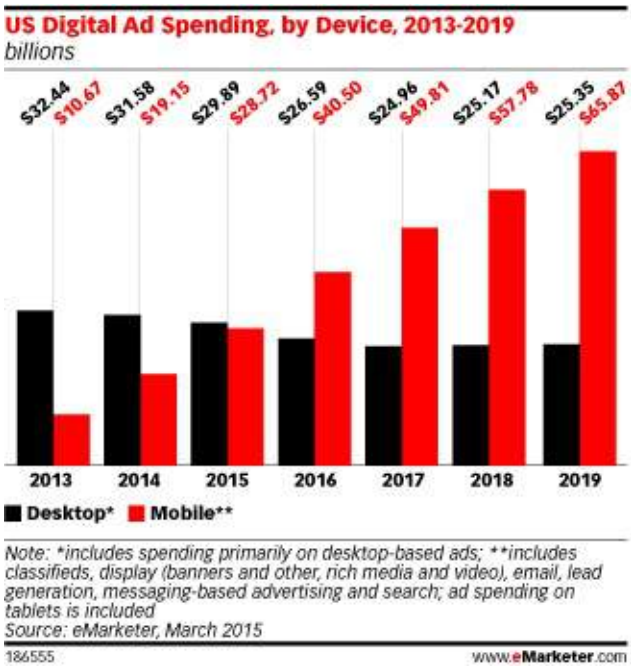


Figure 8-17. Brand page showing in Google's search results

This is nice real estate to own, and can really add to the brand impression of your business. For Major League Soccer, we see the number of followers it has on Google+, and a Follow button at the top right. Google obtains additional information from many different sources, but it is experimenting with allowing publishers to get their social profiles added to the panel shown on the right.

Google+ Impact on Nonpersonalized Rankings of Content

Many people have debated whether Google+ impacts the nonpersonalized results in Google as well. As Figure 8-18 indicates, Google+ posts used to pass PageRank, but this is no longer the case.

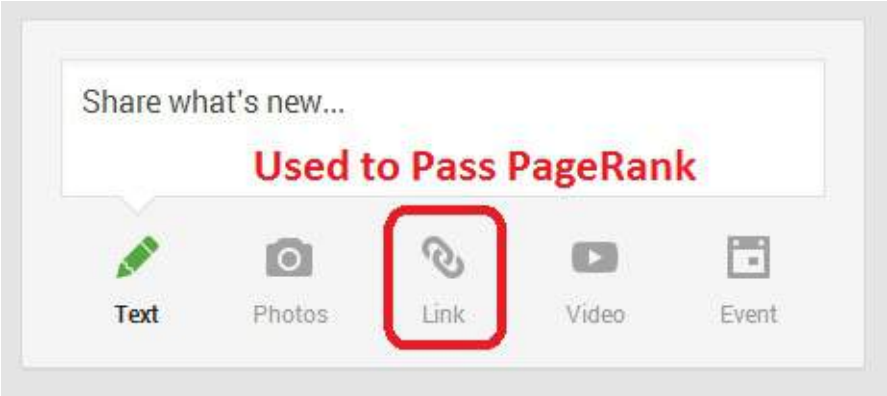


Figure 8-18. Links in Google+ posts used to pass PageRank

You can examine the source code of a Google+ post in detail to verify that these links are now `nofollowed`. Figure 8-19 shows an example of the source code for a particular post.

```
<b>Impact of Offline Activity on Social</b><br><br><span class="proflinkWrapper"><span
class="proflinkPrefix">+</span>

<a class="proflink aaTEdf" href="/107022061436866576067" oid="107022061436866576067">Mark
Traphagen</a></span> provides a thoughtful analysis on how offline activity can impact your social
following. I have been a believer in this concept for years.<br><br>For example, many is the time I
have been speaking at a conference where I have mentioned that attending a conference was a social
media activity.  People would ask me what I meant by that, and I &#39;d ask for a show of hands to
see how many people had followed me as a result of my being there.<br><br>Lo and behold, a bunch of
hands would always go right up in the air!</div></div></div></div><div class="yx Nf"><div class="q9
yg">

<div class="s9Oc9"><div class="sp ej xI A8Hhid"style="width:506px;height:303px;max-height:303px;"><a
href="https://www.linkedin.com/today/post/article/20140813174834-13325642-it-s-all-social-now-how-non
-sm-events-affect-social-media" target="_blank" class="d-s ot-anchor Ks" tabindex="0"
rel="nofollow"><img
src="//lh3.googleusercontent.com/proxy/hT0uB6gr8syC3Eyzy1JWT4ATM3NHVq_0_1ho_4pGtwmwDOL3RXR0FnjDZHYS3f
oHCokd9i2sW_CPNJHBScuLc8-GBCB1YFaCW7pEr_o=w506-h303-p" class="ar Mc" style="max-height:303px;
max-width:506px;" itemprop="image"><meta name="twitter:image:src"
content="//lh3.googleusercontent.com/proxy/hT0uB6gr8syC3Eyzy1JWT4ATM3NHVq_0_1ho_4pGtwmwDOL3RXR0FnjDZH
YS3foHCokd9i2sW_CPNJHBScuLc8-GBCB1YFaCW7pEr_o=w506-h303-p"><div class="iGqbIb"></div></a></div>

<div class="VwVwbf"><div class="rCauNb"><a
href="https://www.linkedin.com/today/post/article/20140813174834-13325642-it-s-all-social-now-how-non
-sm-events-affect-social-media" target="_blank" class="d-s ot-anchor" title="It&#39;s All Social Now:
How Non-SM Events Affect Social Media | Mark Traphagen | LinkedIn" tabindex="0"
rel="nofollow">It&#39;s All Social Now: How Non-SM Events Affect Social Media | Mark Traphagen |
LinkedIn</a></div>

<div class="YPIndd"><a
href="https://www.linkedin.com/today/post/article/20140813174834-13325642-it-s-all-social-now-how-non
-sm-events-affect-social-media" target="_blank" class="d-s ot-anchor" tabindex="0" rel="nofollow"><a
class="jp" href="./107350354213838732087" oid="107350354213838732087"><div class="thIXBd"></div><div
class="iE5ljf">LinkedIn</div></a></a></div></div></div><div class="lr"></div><div
class="ko"></div><div class="kr"></div>
```

Figure 8-19. Google+ links are nofollowed

Study on Google+ as a Ranking Factor

Despite the evidence to the contrary just shown, many people still believe that activity in Google+ can drive search rankings. For that reason, Stone Temple Consulting's Eric Enge performed a study on this as well. The study was designed to see if Google+ shares of links to a web page (not on Google+) would improve the rankings of those pages.

This study took a number of web pages and published them on three different websites. These pages all had these characteristics:

- The pages contained fresh, original content, relevant to the site on which they were published.
- The pages had no links to them whatsoever. No external pages linked to them, and no links were implemented from the pages from the domain on which they were published.
- There was absolutely no Google code (i.e., no Google Analytics code, Google+ code, or any other kind of Google code) implemented on the pages.
- Participants in the test were not allowed to use Google Chrome to visit the pages.

The next step was to conduct some Google+ shares to those web pages so that Google would discover the pages. These shares were done on July 19, 2013, and by July 29, 2013, all of the pages in the test had been indexed. At that point, rankings for very long-tail search terms were monitored until the end of the test.

Seven days after initial indexation, an additional set of shares was performed to the pages. The purpose of this latter set of shares was to see if the rankings that were being monitored would change. You can see a partial data set for the rankings in Figure 8-20.

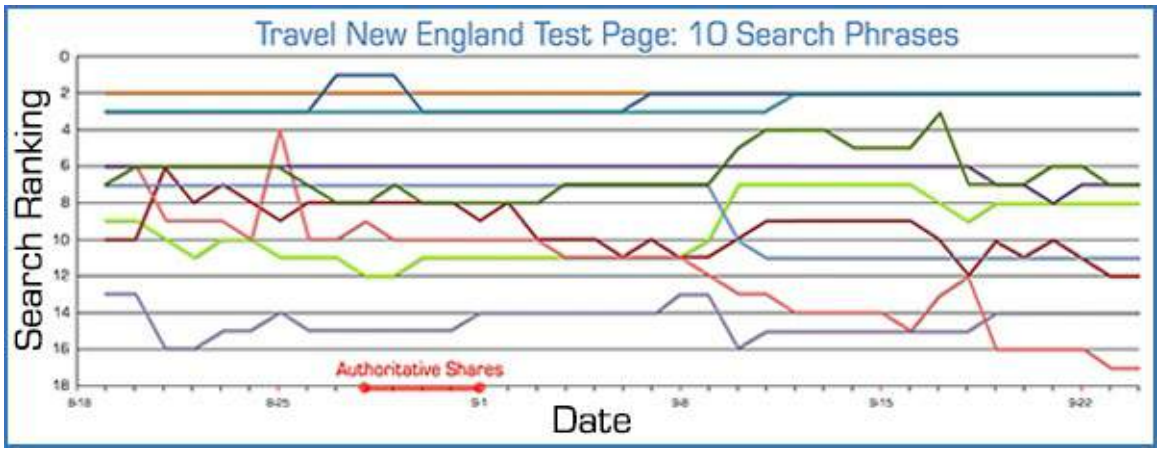


Figure 8-20. Google+ ranking study test results

In short, there was no real evidence of any movement in rankings as a result of the test. Based on this result, there is no reason to believe that rankings are impacted by Google+ shares.

Last but not least is Google's recent Hummingbird release. While most people associate this update solely with improvements in processing natural language queries, this is only one aspect of Hummingbird, and in fact it is a complete rewrite of Google's search engine platform. According to Danny Sullivan, part of the goal with Hummingbird was to allow Google to better process signals other than links, such as signals from social platforms.[4]

While this new information implies that Google is not using social signals yet, it suggests that Google is considering using them in the future.

How Might Google Use Google+ as a Ranking Factor?

Google will work hard to develop algorithms that extract clear signals of value from Google+. How might it do that? We'll illustrate with an example. Consider the following three people, bearing in mind that this example is pure speculation:

Person 1

> This person is active on Google+ and he shares lots of content—his own, and content shared by others. In addition, he always adds thoughtful commentary in his shares and comments on a lot of other people's posts as well. The content he posts gets lots of +1s, reshares, and comments, and his comments

on other people's posts often draw a reaction too. Many of the people he interacts with frequently also get tons of interaction with their activity in Google+.

Person 2

Like Person 1, this person is active on Google+ and shares lots of content—her own, and content shared by others. However, she doesn't often add much content and she doesn't get the same level of interaction from others as Person 1. The interaction she gets tends to be more perfunctory in nature.

Person 3

This person is not on Google+ at all. He is a total rockstar in his space, and everybody knows it, but he does not have a Google+ profile.

Google can't afford to punish Person 3 in the search results. That would be bad for its primary product (you know, the one it makes all the money on?). What Google can do, though, is make comparative judgments between Person 1 and Person 2. In fact, Google may look for ways to recognize the relative authority of Person 3 using signals outside of Google+ and reward him accordingly.

Of course, we don't know exactly how Google is looking at these types of signals, but it will try to find ways to extract value from them that increase the quality of its results. That demands a balanced approach to the data on Google's part, and an incredible amount of testing and verification.

We know that the personalization impact is quite strong, and that Google+ posts themselves can rank in the SERPs. This is already a compelling reason to engage in Google+, especially if there is a real audience for you there. We also know that Google gets the information from Google+ directly. Unlike with Facebook and Twitter, there is no crawling involved in this process. While Google+ continues to grow, on July 27, 2015, Google announced that they would "decouple" Google+ from other Google products—essentially removing the requirement that Google+ be the unified identity for users across all Google products. Later that month, Google began the "unverification" of inactive Google+ local (GMB) pages. How Google+ will evolve has yet to be determined, but these moves have definitely signaled a shift in Goolge's strategy.

The Indirect Influence of Social Media Marketing

Social media can also influence search rankings in a number of indirect ways. For example, it can have a direct branding impact. Here are some other areas where social media can have an indirect influence:

Audience building

You can build a following by using social media platforms such as Facebook, Twitter, Pinterest, and others to regularly communicate valuable information. Successful use of this tactic can help you develop greater engagement with your customers. This greater engagement can lead customers to become brand ambassadors as a result of feeling a closer connection to your organization.

Customer service

Social media sites can provide an effective incremental channel for customer service departments, enabling them to become aware of customer issues that they might not otherwise hear about. While this may not seem like an SEO concern, it can enhance the brand building efforts of your company. People who are able to see a proactive and effective customer service effort by a company learn to value their relationship with that company even more. In fact, many of today's consumers expect a rapid response time to any complaint they make via social media. Failure to meet this expectation can actually be harmful to your brand.

Viral impact

Once an audience feels closely engaged with a company, their willingness to spread the word is greatly enhanced. Customers who feel a strong affiliation with a brand are more likely to help spread the word on new offerings or promotional programs to others in their network.

Link building

Publishers with a strong following on a social media platform can release new content on their websites and then announce it through their social media accounts. This can lead to links to the new content and is an effective

way to build deep links to the site.

Citations

As discussed earlier in this chapter, links in Facebook and Twitter are counted by search engines, but it is also likely (though not confirmed) that brand mentions are also valuable. As search engines look for ways to measure how much they trust a brand, one metric they might use is how many times that brand is referenced across the Web.

Brand search generation

People who learn about a brand and first become engaged with it through a social media presence may later choose to generate a search query on that brand. As with citations, a large volume of actual brand searches is a strong signal of prominence.

There are a large number of social media platforms, and you can waste a lot of time if you chase them all. Here is a brief commentary on the major players that can help you with the promotion of your website:

Facebook

Facebook is the largest of the social networks, with worldwide usage over 1.2 billion users.[5] This represents more than 10% of the total world population. Facebook has the largest audience and is certainly worth including as a key part of your social media strategy. In the US, more than 7 out of 10 Internet users are using Facebook. In addition, Facebook's advertising platform provides highly granular targeting, as is discussed more in the section "Paid social media" in Chapter 7.

Twitter

As of June 2014, Twitter reported 255 million active users.[6] The user base may be smaller than that of Facebook, but it certainly includes a large number of influencers who can help create visibility for your organization. Twitter remains skewed toward male users, with nearly 50% more men on the platform than women.

LinkedIn

As of April 2014, LinkedIn had more than 300 million members worldwide,[7] making it a powerful social network. You can use it to build a strong network of connections, and paid versions of the service allow you to use its InMail functionality to send unsolicited emails (in very limited volume) to members, which, if done judiciously, can help you initiate new relationships with influencers of interest.

Google+

While Google+ reports that it has 540 million active users, external research shows that about half of these do not actively use the social network.[8] However, 270 million or more active users is still a significant number.

Instagram

Instagram reportedly has over 200 million users, and is growing fast. Piper Jaffray reports that it is the most important social network among teens.[9] Instagram is designed from the ground up as a mobile experience.

Pinterest

Pinterest may be quite a bit smaller than the previous platforms, but it is a growing network with 70 million reported users.[10] The network skews toward a female demographic, so if that is your target market, it could be a great platform for you to focus on. It's considered a highly "immersive" media in terms of time spent in the app and numbers of boards built per person. The audience on Pinterest is still US-centric.

YouTube

YouTube can be thought of as a search engine, but it is also a social network. People love to share videos and do so in volume. YouTube also has a very large audience, with 159.1 million unique viewers per month as of December 2013.[11]

WhatsApp

WhatsApp is a cross-platform mobile-messaging app that positions itself as a cost-free alternative to SMS communications. WhatsApp is the number-one mobile messaging app based on a survey of 3,759 users across the United States, China, Brazil, South Africa, and Indonesia.

There are many other social sites worthy of consideration. These include:

- Specialized sites (e.g., Vine and Snapchat)
- Q&A sites (e.g., Yahoo! Answers, StrackExchange, Quora, Formspring, and ChaCha)
- Document sharing sites (e.g., SlideShare, DocStoc, and Scribd)
- Local business sites (e.g., Google Local Business Center, Yelp, UrbanSpoon, YellowPages, CitySearch, and TripAdvisor)
- Social news sites (e.g., reddit and Techmeme)
- Bookmarking sites (e.g., StumbleUpon, Delicious, and Trunk.ly)
- Vertically focused social sites—as may apply to specific vertical markets (for example, there is a knitting social site at *http://www.knittng.com*, so there may be a social site specific to your vertical market)

Monitoring, Measuring, and Improving Social Media Marketing: Best Practices

While search engine optimization and social media marketing are two distinct practices, the influence of social media behavior on search results is, and will continue to be, important to consider. This section will explore some basic aspects of how to optimize your efforts in the social realm to help your SEO.

Social media networks all operate as communities. In reality, each network is a community of communities. No individual interacts with an entire network. Each user finds his “tribe” of followers and following. Even within those groups, there will be tighter “inner circles” of people and/or brands that get the most interaction and attention from a user.

The structure, rules, and methods of interaction may vary, but a social media network's primary focus is to create and foster relationships. For example, you may think of YouTube as a video site, but a better label for it would be a video-sharing site. Social network sharing is done with the ultimate goal of attracting and engaging a following. People aren't on social networks to broadcast into a void; they want to feel that they are being heard and that their content has meaning for others.

Effective use of social media results from fostering relationships that build your brand's reach, authority, and trust. Try to avoid seeing it as a broadcast, advertising, or even direct-selling medium. Even with brands, social media users want authentic engagement. They will resist being sold to. But handled properly, social media provides an excellent opportunity to build positive associations, so that when prospects are ready to buy, your brand is top-of-mind.

The key to success in any community is to be a contributor. You must put something into it to get anything back out of it. This requires authentic participation, as little value accrues to those who participate in an artificial way.

Find ways to contribute content, information, and opinions to discussions. Using Twitter as an example, one approach is to start your day by reviewing the daily news related to your market space, pick out the content you consider the most valuable, and share those items with your following. Another approach is to offer valuable tips on a regular basis. A third approach is to be known for controversial opinions, but be aware of the impact this may have on your brand. Bring a unique viewpoint.

However, if you only share content, whether your own or curated from other sources, you will always be limited in how far your influence will extend. Remember that people are on social media, at least in part, to be affirmed—to feel that they matter to others. For that reason, you should dedicate some time each day to engaging with the posts of others. Offer constructive, helpful comments and replies. Don't just say "great post!" but explain specifically what you liked about it, or contribute information or insight on the topic. Not only does this endear you to the person whose content is the recipient of your comments, but others seeing the conversation may follow you because of your helpfulness and knowledge.

Overall, focus on building your credibility and authority. These are key elements

to success in a social environment. Once people accept you as a legitimate participant, the doors will start to open for you.

Even if you are a novice (a.k.a. “noob”) in your market area, you still have something to offer. The experts can learn by watching how you learn as you gain experience. Sometimes just posting a really good question can be as valuable as providing a great answer. Such engagement brings a surprising amount of value to the community, and if you can also bring some personality into the process, people will try hard to help you and will want to engage with you further.

As a publisher, there will also likely be times when you want to put something out there that is self-promotional. This is fine as long as you don’t do it with all your communications, and make certain that you do it in an engaging way. You earn the right to occasionally self-promote by being considered a valued contributor to the community.

Claiming Key Profiles

Make sure to claim your brand/profile at key social sites. Services like Trackur and Radian6 allow you to monitor the use of your brand across a wide array of social sites. You will want to set up a brand presence on each social media platform even if you don’t intend to use a network yet, in part to prevent your name from being taken by impostors or “squatters” (those who claim brand name profiles in hopes of later selling them to the actual brands). Here are some of the sites where you should create a brand profile for your organization:

- Facebook
- Twitter
- Google+
- LinkedIn
- YouTube
- Pinterest
- CrunchBase

- Quora
- About.me
- Instagram
- Tumblr
- SlideShare
- reddit
- StumbleUpon
- Delicious
- StackExchange
- Google Places for Business
- Yelp
- UrbanSpoon
- YellowPages
- CitySearch
- TripAdvisor

Use NameChk.com to rapidly see if your name is available across a wide range of social sites. Note that this does not mean that you should plan on building an active presence on all of these sites; for most companies, that’s unrealistic. However, claiming the profiles will prevent other people from implementing profiles with your brand name on any of these platforms.

Deciding on a New Social Network

When you first consider engaging in a new social network, it is hard to be sure

what return on investment (ROI) you will get for your efforts, and what you will need to commit to get that ROI. It's a good idea to start your efforts with an experimentation cycle designed to help you determine what you will be able to get out of it.

A word about ROI in social media: remember that social media is primarily for reach, branding, and trust/authority building. Therefore, the best measurement for return on social media efforts probably isn't sales—at least, not directly. Measure your effectiveness in social media by metrics like followers, engaged followers (which is actually more important than raw follower numbers), engagement per post, increased positive brand mentions, and increased traffic to your site.

Here are some suggestions on how to set up experiments with various social networks:

Hypothesize about a site's potential value.

> Are the audience demographics a fit for what you want to accomplish? Is the medium one that you effectively utilize? For example, you would not want to consider YouTube if creating/sharing videos does not make sense for your organization. Take the time to consider these and other questions that affect the site's potential value to you before getting started.

Research what networks are being used by your competition.

> Lack of competitors on a particular network is not necessarily a red flag. There may be an opportunity for you to lead the way in your niche on that network, if you can develop strategies to either find or develop an audience there.

Set up social media and analytic tracking systems to measure.

> When you first start out, your returns will be much lower than they will be in the long term as your credibility in communities builds, but even early results can help guide you on what is most effective. We will discuss this more in the section "Tracking Social Media".

Create an account and invest a few hours in participating.

If the network has search, try different keywords to find people, groups, and communities discussing topics relevant to you, and start engaging in those posts and creating your own. This will help you test your hypothesis about the value. A few hours will not be enough time to start seeing any ROI, but it should help you better qualify whether the community has the potential to work for your organization.

You won't really know the effectiveness of a network until you start getting an engaged following there. The best way to gain real followers is to seek out people and communities within the network who are relevant to your vertical, follow them, and begin engaging with them. People are more likely to follow you back and/or recommend you to others when you do this.

If this initial test goes well and you still believe that the social network may be a good investment, go a little further with it until you see some results.

At that point, step back and consider the return on the effort so far. Be aware that your ROI should continue to increase as your credibility and authority build, but this will give you an initial idea of what may be possible.

Reach out to some of the more authoritative people using the network to get a sense for the types of returns they are seeing.

While this information will probably be qualitative rather than quantitative in nature, it can help you learn what types of ROI might be possible.

Compare the potential ROI, including the potential long-term value of the connections you can develop on the network, with other ways you could invest that same time and energy.

Keep at it if it looks like a high-value channel, and drop it if it doesn't and try something else!

Tracking Social Media

In the world of social media, there are several key metrics we're interested in tracking:

Traffic data

How many visits and visitors did social media drive to our sites?

Fan/follower data

How many people are in our various networks and how are they growing?

Social interaction data

How are people interacting with, sharing, and resharing our content on social networks?

Social content performance

How is the content we're producing on social sites performing?

Getting the right metrics to answer these questions requires segmenting by network. Not every question will have direct answers in the data, so you may need to make assumptions or inferences. Buzzstream and Klipfolio are examples of popular tools for tracking social media performance.

Facebook

Facebook offers a relative wealth of data through its built-in product for brand pages, Facebook Insights. For example, you can see likes, post reach, and engagement (as shown in Figure 8-21) or the demographics of your fans (as shown in Figure 8-22).

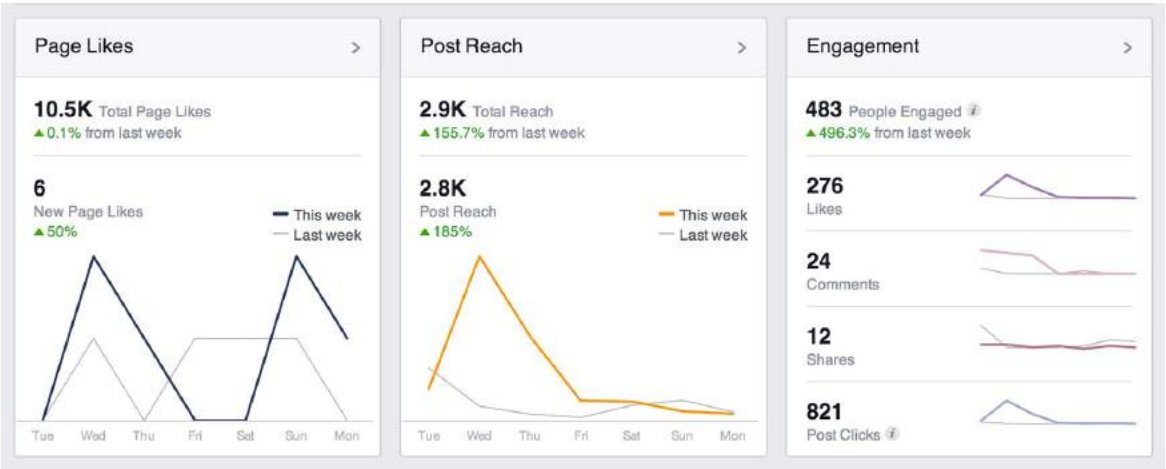


Figure 8-21. Facebook Insights dashboard

Figure 8-22. Facebook Insights demographics

Insights also has a very unique and powerful feature: integration on your website. Using a small bit of JavaScript code, you can embed the Facebook Insights functionality on your site and receive information about all the users visiting your pages who are logged into Facebook.

More information on Insights can be found at:

- Official Insights page on Facebook
- "How to Use Facebook Insights"

Twitter

Twitter and Facebook are likely to be the largest two social networks for referring traffic to most sites (although a Shareaholic study shows LinkedIn and Google+ drive more engaged visitors; see *http://bit.ly/google_plus_linkedin*).

Twitter now provides limited analytics. For more detail, you should look into third-party social analytics tools.

The key metrics on Twitter are:

Followers (and follower growth over time)

The unique number of Twitter users who've "followed" your account, as

shown in Figure 8-23.

Figure 8-23. Twitter followers

Active followers

The number of followers who've logged into or used a feature of Twitter in the past 30 days (those who have not are likely inactive or nonhuman accounts). This data is challenging to get, and requires software that runs through your followers and determines which are actively using via the API. Some of the third-party tools discussed momentarily will show this information.

@ Replies

The number of tweets sent that begin with your account name. An example is shown in Figure 8-24. These are seen only by the mentioned account and anyone who follows both sender and receiver.

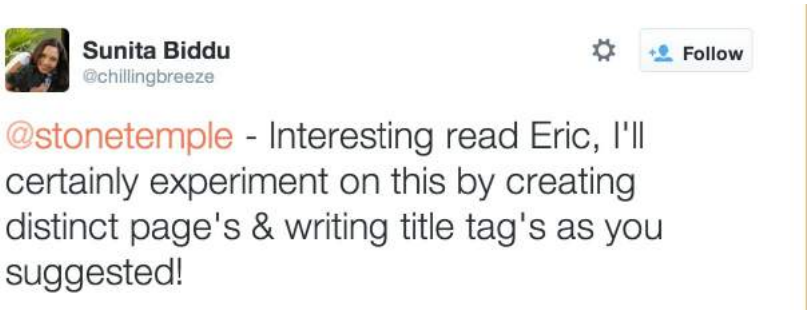


Figure 8-24. Tracking Twitter @Replies

@ Mentions

The number of tweets sent that include your account name (@ Replies are also @ Mentions). Figure 8-25 shows a tweet that includes Rand Fishkin's

account name, but uses it inside the tweet rather than at the beginning. This means others on Twitter can see the tweet by default, rather than only those who follow both accounts.

Figure 8-25. Tracking Twitter @Mentions

Brand mentions

Tweets that contain the brand/account name but don't use the @ symbol. An example is shown in Figure 8-26.

Figure 8-26. Brand mentions in Twitter

Domain/URL mentions

The number of tweets that include a link that contains your brand name/domain name. An example of this using Moz is shown in Figure 8-27. These now include, by default, any shortened URL that contains the brand/domain name, as Twitter automatically parses the final destination URL for matches to the query.

Figure 8-27. Tracking Twitter domain/URL mentions

Direct retweets

The number of retweets (using Twitter's native retweet button/functionality) of tweets you've posted that have appeared on the service, as shown in Figure 8-28.

Figure 8-28. Tracking Twitter retweets

RTs and vias

The number of tweets that contain an RT or via of your account (see the example in Figure 8-29). These are similar to direct retweets, but aren't necessarily counted by Twitter's automatic RT system because they modify the original message and appear to come from a unique source. Some users indicate a modified retweet with *MT*.

Figure 8-29. Modified RTs (MTs) in Twitter

Best performing content

The content you've shared on Twitter that has earned the most clicks, retweets, and shares. This is now available in Twitter's own analytics, but

some third-party tools give deeper analysis.

Direct traffic + non-Twitter.com drivers

Sources that have sent traffic to your site via Twitter's ecosystem, even if they come from desktop clients or other third-party software sources. See Figure 8-30 for an example. As of 2011, these show up (mostly) as coming from T.co (Twitter's link shortener), although some third-party link-shortening services, such as bit.ly, have their own analytics for clicks on links you create through them, even if Twitter publicly changes them to t.co links.[12]

Figure 8-30. Tracking Twitter traffic

It's often said that in analytics, nothing is worth tracking unless it can be used to take action and improve. For the preceding metrics, the primary actions you're tracking are ones that relate to you. The key to taking better actions is comparing successful interactions, tweets, and content against less successful ones to determine what has the best impact on growing your audience, bringing visits to your site, and eventually, driving conversion actions.

LinkedIn

LinkedIn functions like a hybrid of Twitter and Facebook. Connections require acceptance from both sides, but public entities (like company pages) and groups can be followed.

LinkedIn used to be known primarily as a sort of online address book of your business connections, and a place for job hunters to post their résumés. Recently,

though, it has made major strides to reinvent itself as a true social network. Users now get a Facebook-like news feed of content and conversations shared by people and companies they are linked to.

In early 2014, LinkedIn opened up its Publishing platform, formerly reserved to handpicked “influencers,” to all users. Publishing allows users to create bloglike content that gets pushed out to their connections and followers. Other users can now follow the Publisher content of anyone else without having to first gain a reciprocated connection with the author. These changes have made LinkedIn a much more significant player in the social media influence-building arena, especially for B2B concerns.

You can also leverage LinkedIn groups to develop your visibility and build new connections. You can use LinkedIn search to find relevant groups that show active participation. Review the group rules to make sure that some are at least in place, and being enforced. It’s best to find those that appear to have active discussions taking place.

You can also reach your target audience on LinkedIn using direct sponsored content, which can be a powerful supplement to your content marketing efforts.

Like Facebook, LinkedIn has some built-in analytics for businesses as well as individual profiles, and lots of data points that are useful to track, including:

Company page reach and engagement

You can track the number of times your company’s LinkedIn profile and its shared content have been viewed over time and the quantity and types of engagements earned. An example of this is shown in Figure 8-31.

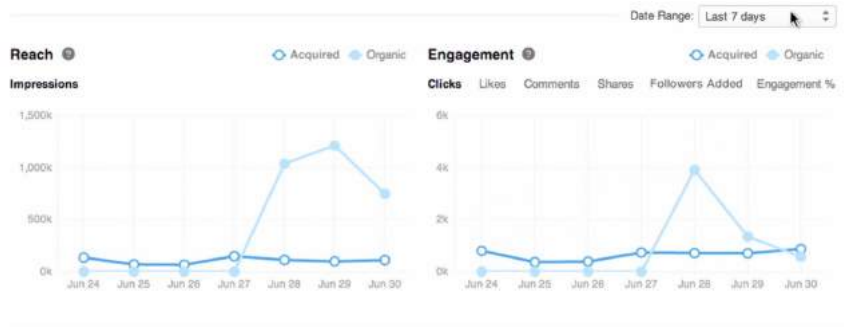


Figure 8-31. LinkedIn reach and engagement

Followers

As with Twitter, individuals can “follow” a brand account on LinkedIn and receive status updates in their “updates” stream. Figure 8-32 shows stats for followers of a LinkedIn account. The more followers you have, the greater your ability to reach more people on LinkedIn with the content you share there.

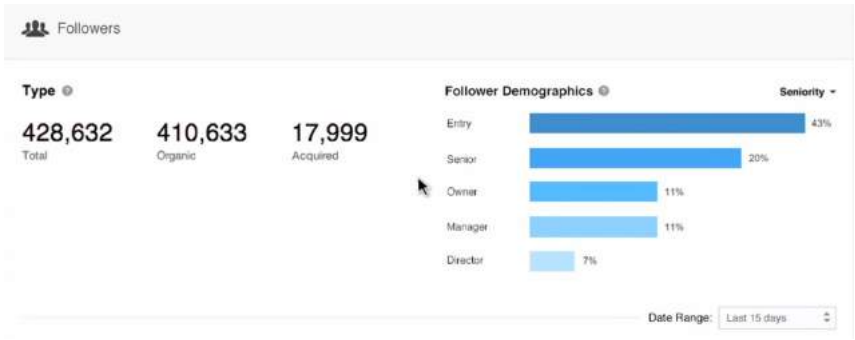


Figure 8-32. Followers in LinkedIn

Connections

The number of unique connections for an individual profile on LinkedIn is a worthwhile metric to track, but unfortunately, all that is available as of July 2014 is the raw, current count (as shown in Figure 8-33) and who those connections are.

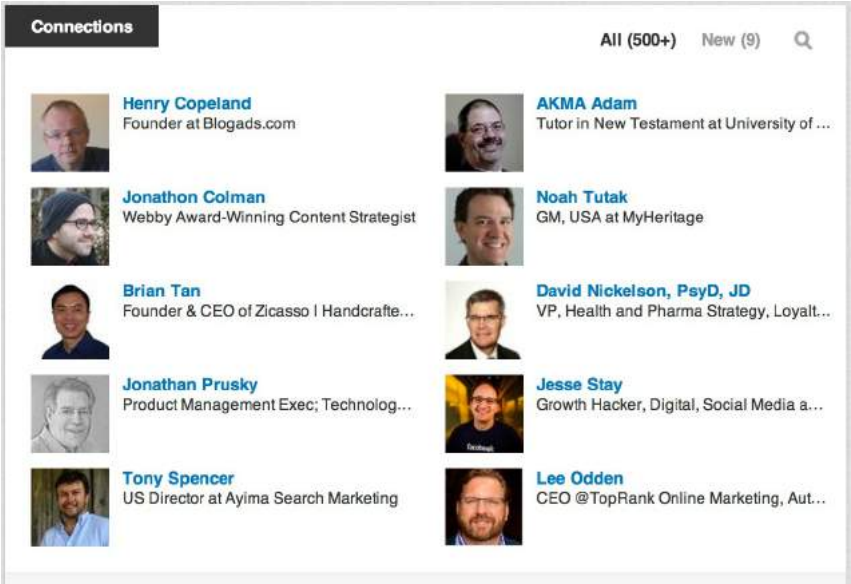


Figure 8-33. LinkedIn connections

Messages, invitations, and notifications

Figure 8-34 shows the status bar at the top of a LinkedIn page. The number on the envelope shows unread messages and unreviewed connection invitations, while the number on the flag represents unviewed news stream updates and other updates, such as people viewing your profile or commenting on one of your updates.

Figure 8-34. LinkedIn messages, invitations, and updates

Profile views

You can see how many people have looked at your profile over time, along with some data about who they are (only paid “Pro” LinkedIn accounts can see the identities of all profile viewers). Figure 8-35 shows an example.

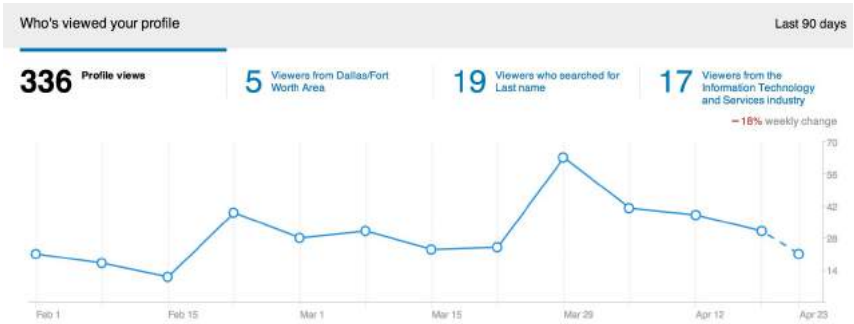


Figure 8-35. Profile views on LinkedIn

Top keywords and referral sources

This list shows you the top keywords that LinkedIn users searched for prior to discovering your profile, and where they came from to get to your profile, as shown in Figure 8-36. (Again, only a limited amount of data is shown unless you have a paid account.)

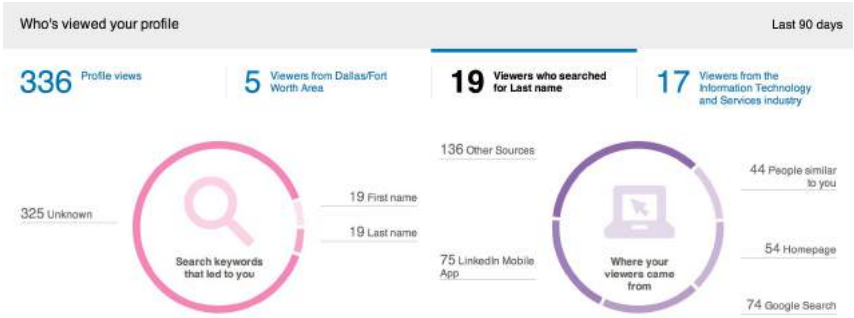


Figure 8-36. Keywords and referral sources on LinkedIn

Content shares

As of July 2014, there is no way to measure or record the number of regular status updates/shares you've sent out over LinkedIn, nor the number of "likes" you've received on the service.

LinkedIn Publishing

LinkedIn Publishing allows users to create original content with images, links, and formatting, which is then pushed out to their connections,

followers, and possibly beyond. Other users can follow someone's Publishing content without having to make a two-way connection. LinkedIn Publishers get some rudimentary metrics, such as number of views, likes, and comments for each post, as well as information on followers.

Traffic

LinkedIn isn't a huge traffic driver for most sites, but for certain B2B sites, it can be relatively substantial and the quality is often higher than other social sources. Figure 8-37 shows a screenshot from a Google Analytics account.

Social Network	Sessions
1. Google+	2,774 (56.83%)
2. Twitter	1,393 (28.54%)
3. Facebook	300 (6.15%)
4. reddit	114 (2.34%)
5. LinkedIn	70 (1.43%)

Figure 8-37. LinkedIn traffic

Few third-party tools exist to help with LinkedIn metrics, but over time, some of the tools for Facebook and Twitter may expand to LinkedIn.

Google+

Google's social network is still relatively young, but at least three factors make it worthy of marketers' attention:

- The network's rapid growth, especially in the global market (some estimates now make it the second largest social network, surpassing Twitter and second only to Facebook).[13]
- The integration of Google+ into some of Google's products and services—former Google Vice President for Social Vic Gundotra has called it "the next version of Google" and "the social layer of Google."[14] Google has since backed away from integrating Google+ into all of its services.
- The influence of Google+ on Google search, particularly in the area of personalized search.[15]

Google has promised analytics for pages and profiles, but as of this writing those still aren't available. However, several third-party tools have gone a long way in delivering the needed metrics. Among these are CircleCount, SteadyDemand, and NOD3X. The latter is a sophisticated tool for in-depth analysis of follower engagement, reach, and influence.

Significant metrics on Google+ include:

Number of followers

On Google+ "followers" are people or pages who have you or your page in at least one of their circles. You can see the number of followers on a profile or page, as shown in Figure 8-38 (if the profile or page owner has made followers visible in settings, that is).

Figure 8-38. Google+ followers and views

Views

View counts are new as of April 2014, and are displayed on the cover image of profiles and pages that have not opted out of their display in settings. The view count is an approximate total of all profile views, post impressions, and image impressions since October 2012. For more on what comprises views and what this metric might mean, see "New Google Plus Views Count: Important Metric or Vanity of Vanities?".

+Mentions

On Google+ one user links to another by typing a "+" and then the user's name. This creates a link to the mentioned user's profile, and also sends that

user a notification that she has been mentioned. As of April 2014, Google+ does not provide a count of these mentions, but some third-party tools do. You can view all your recent mentions by clicking on the Mentions tab. Figure 8-39 shows an example.

Figure 8-39. Name mentions on Google+

Brand mentions

As shown in Figure 8-40, you can find recent mentions of your brand through the Google+ search, using the Everything filter tab and changing the view from "Best of" to "Most recent."

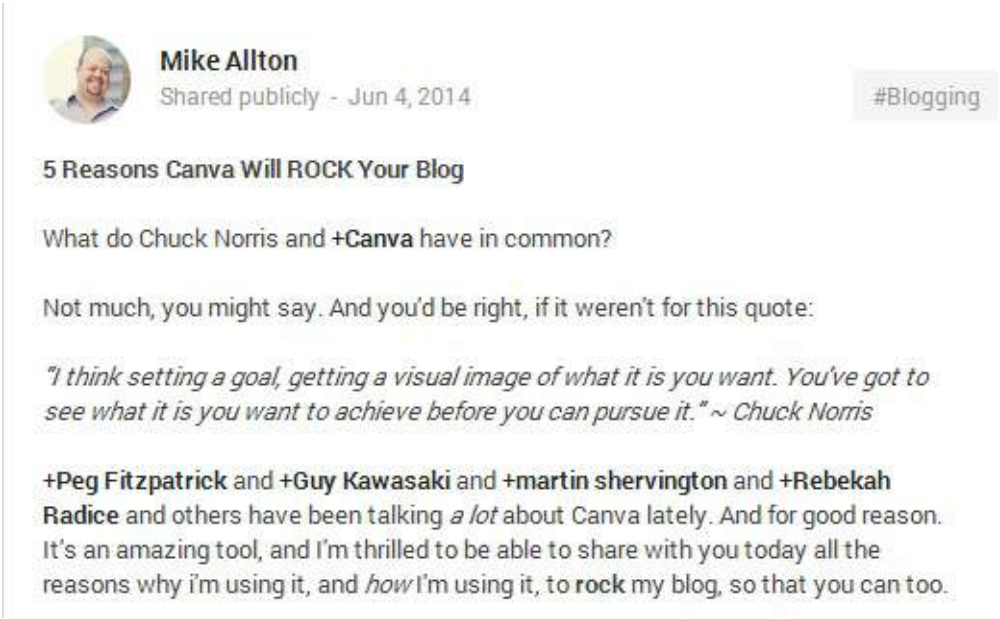


Figure 8-40. Brand mentions on Google+

Content shares, content +1s, link shares, link +1s

You can view recent shares of links to your own site on Google+ in Google Analytics under Acquisition > Social > Data Hub Activity, with Google+ as the selected network. Clicking the More button to the right of any Google+ page share allows you to filter for all shares of that link, and to view the Google+ Ripples for that link (Ripples are a graphic representation of all the public shares of the link on Google+). You can also view the number of posts, reshares, +1s, and comments for any link to your site that was shared on Google+ by navigating in Google Analytics to Acquistion > Social > Landing Pages, and then clicking the desired link. Next select "Google+" and then change the Primary Dimension to "Social Network and Action." See Figure 8-41 for an example.

Primary Dimension: Full Referrer Social Network and Action

Secondary dimension

Social Network and Action	Data Hub Activities	Data Hub Activities
1. Google+ : +1	35	54.69%
2. Google+ : comment	12	18.75%
3. Google+ : reshare	12	18.75%
4. Google+ : post	5	7.81%

Figure 8-41. Google+ site link activity in Google Analytics

Traffic

Google+ can become a major site traffic driver for brands that develop an active following there. Figure 8-42 is a screenshot of traffic from Google+ as shown in Google Analytics.

Social Network	Visits
1. Google+	4,260 (49.94%)
2. Twitter	2,345 (27.49%)
3. Facebook	679 (7.96%)
4. reddit	500 (5.86%)
5. LinkedIn	212 (2.49%)

Figure 8-42. Traffic from social networks in Google Analytics

Perhaps due to privacy issues, Google+ uses a single referring URL for all traffic, helping to consolidate the data in analytics reports, but making it frustrating to determine which shares/users/links sent the most, or most valuable, traffic.

Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

Depending on the quantity and value of the traffic that other social networks send, there may indeed be additional metrics worth tracking. Sites worthy of consideration are Instagram, Pinterest, StumbleUpon, SlideShare, reddit, and Quora. These are likely worth some investment on the metrics and effort front, and if small quantities of contribution/participation yield large returns, more investment is likely warranted.

Social news sites

The two most well known social news sites are reddit and Techmeme.

As we covered in Chapter 7, reddit is a social link aggregation and discussion site, self-organized into topical, independently moderated communities ("subreddits"). There are millions of visitors each day, and reddit has hosted "ask me anything" interviews with many celebrities and a few politicians, including Barack Obama.

Articles posted on reddit can potentially receive a lot of search traffic, so it can be tempting to try to influence this process to help achieve your goals. However, this is a bad idea, because registered users on reddit ("redditors") are very defensive of their platform, and quick to squash any attempts at manipulation or any commercial behavior.

Even trying to reach out to influencers on this platform will quickly draw a negative reaction, and can result in your being labeled as a spammer. Any attempt at solicitation runs a significant risk of your being downvoted, attacked, and/or blacklisted.

Still, there is a lot to be gained from reddit, but before you embark on such a campaign, spend some time at *http://www.reddit.com/wiki/reddiquette*, to learn the expected behavior on the platform. The best strategy for reddit is to try to participate normally and submit a lot of other people's great content to the

correct subreddits, and occasionally submit your own potentially viral content, without giving the appearance of self-promotion.

Techmeme is a site that provides a curated list of the hottest and most important tech industry news, and it does so on a single page. The site has a small team of editors who review what's going on and decide what to post. The only way to suggest content to this team is via Twitter.

Blogs and forums

The world of social media started out as one where discussion sites (forums, Q&A, bulletin boards, etc.) and the blogosphere reigned supreme. Eventually, the major social networks captured the hearts and minds of most users, but the reports of the death of blogs and forums have turned out to be greatly exaggerated.

Forums and blogs still hold rich opportunities for marketers. For example, Moz.com receives tens of thousands of visits each week from blogs and discussion sites of all sizes, and the site's participation/interaction with those sources often yields fantastic results in referral traffic, mindshare, and links. Many brands do likewise, hiring community managers or evangelists to engage in the sites where industry topics are discussed, building up strong, recognizable profiles that help bring awareness and produce traffic and links.

Thus, as responsible inbound marketers, it's our job to measure these channels and quantify their impact. Participation in blogs and forums can provide a lot of value, including:

Site/brand mentions

> Mentions of your site or brand name can help lead you to content and conversations worthy of engagement, and it's often useful to track the quantities of those mentions (and possibly the sentiment as well) over time. Google Alerts and Mention are both potentially worth looking at to help you monitor these mentions.

Links

> Direct links are nice because they appear either in link-tracking tools like Search Console, Open Site Explorer, Majestic SEO, Ahrefs,

LinkResearchTools, or directly in your web analytics (if they send any traffic).

Traffic

A must-have for any inbound channel, visit tracking is the simplest metric here.

For any inbound marketing channel (social or otherwise) you are considering, you should use a process similar to that shown in Figure 8-43.

Figure 8-43. Process for identifying high ROI social networks

Losing a few hours to channels that don't provide value is a small price to pay for the value of discovering and participating on those that do. You can see a presentation on a more in-depth process for evaluating social metrics in these two articles: "The Power of Inbound Marketing" and "Social Media ROI: How To Define a Strategic Plan".

NOTE

A special thanks to Mark Traphagen for his contributions to this portion of the chapter.

User Engagement as a Measure of Search Quality

Measuring user behavior on a website can provide strong signals about that site's quality. For example, if a visitor arrives at a site, visits 10 pages over the course of an hour, selects a product, adds it to her shopping cart, and then buys it, chances are pretty good that she found what she wanted there. Contrast that with the visitor who arrives at a web page and hits the back button of his browser in less than a second.

These are examples of user engagement signals, and search engines are beginning to use this kind of data in their algorithms. The specific signals they are using and how they're using them is not easily discerned, however. Search engines are secretive about the details of their algorithms because they are important trade secrets, and because it makes a spammer's job harder. However, we know that user engagement signals are useful in measuring search quality and may also be used as ranking signals. If a page is a poor result for a user's search queries, then the user's (probably very limited) interactions with that page will show it.

This section will examine the ways that the search engines can collect user engagement data, and the types of metrics they can collect. It should also be mentioned that in addition to the possible search result ranking impact of these signals, publishers should actively track user engagement metrics on their websites for a number of other reasons, including measuring the quality of the organic search traffic the site is receiving. Increased website engagement will result in increased conversion rates.

How Google and Bing Collect Engagement Metrics

Bing and Google both have a large number of data sources available to them. Some of the most important are:

Search results

User interactions with the search results are a key source of data. For example, if the user conducts a search and chooses not to click on the first or second result but does click on the third, that can be a signal that the third result may be the best result for that query, especially if this is a common occurrence. Search engines accumulate data like this in high volumes every day.

Browsers

Microsoft's Internet Explorer and Google's Chrome both have sizable market share and could be a rich data source for search engines (see "March 2014 Market Share" and "IE11 passes IE10 in market share, Firefox slips a bit, and Chrome gains back share"). Browsers are powerful data sources because they can monitor every action a user takes.

Contextual advertising

Google AdSense and Microsoft Content Ads are excellent sources of data as well. These offerings allow publishers to place ad units on their websites and earn revenue when users click on them, and data is passed back.

Toolbars

Surfers can install either the Google toolbar or the Bing toolbar in their browsers. These toolbars offer users a number of conveniences that improve their web surfing experience.

Phone operating systems

Both Google and Bing offer phone operating systems. Determining phone OS market share is complex for many reasons, but an April 2014 *TechRepublic* article pegged Google's Android OS as the #1 phone OS in the United States, with more than 75% market share,[16] and the Windows Phone OS is forecast to grow its share over time.[17]

Desktop and web-based applications

Microsoft Office is the dominant office productivity software suite worldwide. Links included in Microsoft Office applications can be tracked. Similarly, Google offers Google Docs as an office productivity suite, which operates entirely online.

Buy data from ISPs

Both major search engines have the ability to buy usage data from ISPs.

Google Analytics

Various studies show a range of market share statistics for Google Analytics. For example, one study showed 63% of the Fortune 500 companies use Google Analytics on their websites.[18]

Goo.gl

Since December 2009, Google has offered a URL shortener.

Internet services

Google offers free WiFi service to the city of Mountain View and offers Google Fiber to a growing number of U.S. cities. Google Fiber is already present or being installed in Portland (Oregon), San Jose, Salt Lake City, Provo, Phoenix, San Antonio, Austin, Kansas City, Nashville, Atlanta, Charlotte, and Raleigh-Durham.

Email services

Services like Gmail, Outlook, and Hotmail (now integrated with Outlook) provide yet another data source.

Payment services

Google Checkout, Google Wallet, and any mobile payment apps running on the Android OS are another potential source of data.

Potential User Engagement Signals

Each search engine has a rich array of data sources that enable it to measure a wide range of web surfer behavior. Here are some of the many signals search engines can extract from those sources:

Click-through rate (CTR)

The search engines can measure the click-through rate on links presented in search results, on web pages in URL shorteners, in their browsers, and more.

Clicks on other search results

Once a user completes a search and visits a link, a common behavior indicating a problem with a result is that she returns to the search results, often quite quickly, and then clicks on another result.

New search generation

Similarly, a user may look at a given search result, come back to the search engine, and modify his search query.

Bounce rate

Bounce rate is a measurement of the percentage of users who visit only one page on a website. Search engines extend that definition to take into account the user's interaction with the search results. For example, if a user clicks on a search result, then returns to the SERPs and clicks on another result, that could indicate that the first result was not a good response for that search query.

Time on page

Search engines can measure the amount of time spent on a given page using their browsers or toolbars. This is also referred to as *dwell time*, and more time on page might be considered signal of higher quality.

Time on site

Similarly, time on site could be considered a positive signal if the average user spends more time on your site than she does on the sites of your competitors. Of course, it could also mean that your site is difficult to

navigate or loads very slowly, so you'd need to look at this signal in conjunction with other signals.

Pages per visit

This metric can be easily measured by the browser, a toolbar, or Google Analytics. More page views implies greater user engagement with the site; however, some sites seek to increase page views by paginating their content to generate more ad impressions. As with time on site, a signal like this is hard to evaluate on a standalone basis.

Repeat visits

Do users return to the site? Return visits could also be viewed as a quality signal.

Pages printed

While many pages on the Web do not lend themselves to printing, certain classes of pages do, such as reference articles, recipes, maps, and similar content. If a user decides to print a page, that indicates a higher level of interest in the content.

Pages bookmarked in the user's browser

If a user bookmarks a page so he can return to it later, that is also a positive signal.

Scroll bar usage

Another relatively subtle indicator of engagement is whether the user scrolls down to see more of the content on a page.

Voting Mechanisms

Another set of signals that search engines can use involves voting mechanisms. These are methods by which users directly indicate their approval or disapproval of content. Here are some examples:

Chrome Blocklist extension

On February 14, 2011, Google released a Chrome extension that allowed users to block specific websites from their search results. This was not used as a ranking factor in the initial release of Google's Panda algorithm on February 24, 2011, but it was included as part of the second release of Panda on April 11, 2011.

Google+ +1s and Google+ shares of content

As discussed earlier, the evidence does not support these data points being used as ranking signals just yet (other than for personalized search), but in the future they could be.

Brand name searches

Another potential signal of prominence is a high number of brand name searches. For example, brands like Coca-Cola and Pepsi get searched on hundreds of thousands of times per month. This might make them better responses to generic search queries like *cola* than lesser-known brands.

Brand mentions

A related concept is that of brand mentions. People reference brand names on the Web all the time—on web pages, in tweets, and elsewhere—without actually implementing links to the page of the brand's website. While there is no documented evidence that search engines use this type of data, frequent brand references could be a strong signal.

Document Analysis

Document analysis is a somewhat different concept than user engagement, but it can be used to predict how users will perceive the quality of content on a site. Strong signals indicating a likelihood of poor-quality content could potentially be used as a ranking signal by the search engines.

Poor Editorial Quality

Poor spelling and bad grammar can indicate a lack of attention and care about the content. You must exercise care here, as user-generated content will not necessarily be checked for spelling or grammar. Publisher authored content that is not subject to quality editing, however, may indicate low-quality content.

Reading Level

The concept of reading level refers to the estimated grade level of the sophistication of the writing. One formula for this is known as the Flesch-Kincaid Reading Test, which measures the average words per sentence and the average syllables per word. It is actually not a measurement of the education level required to read the content; rather, you can think of it of as a measurement of the mental effort required to read a sentence.

If you want to check the reading level of a document, you can do so using the spelling and grammar checker in Microsoft Word. You can also configure Google's Advanced Search settings to show the reading level for documents in the search results. This modifies the results, as shown in Figure 8-44.

There is no single reading level that is optimal for each business; it varies depending on the nature of the content a site publishes. For example, if you have a site targeted to children, then a younger reading level is appropriate. Contrast that with content on the latest advances in artificial intelligence algorithms, where much higher reading level would be expected.

how to make french toast

Web Videos Images Shopping Books More Search tools

Any time Reading level Clear

How To **Make French Toast** - IncredibleEgg.org

Ad www.incredibleegg.org/**FrenchToast**

Look Up the American Egg Board's Recipes on Your Mobile Device!

Fast & Easy Recipes Health & Nutrition

How To Videos Hard Boiled Eggs

Results by reading level for **how to make french toast:**

Basic	88%
Intermediate	11%
Advanced	< 1%

Directions

- Beat together egg, milk, salt, desired spices and vanilla.
- Heat a lightly oiled griddle of skillet over medium-high flame.
- Dunk each slice of bread in egg mixture, soaking both sides. Place in pan, and cook on both sides until golden. Serve hot.

French Toast | Recipe - Allrecipes.com

allrecipes.com/recipe/**french-toast**-i/ Allrecipes.com

Feedback

How to Make French Toast: 9 Steps (with Pictures) - wikiHow

www.wikihow.com › ... › Breakfast › French Toast wikiHow

How to Make French Toast. Originally a way to use up stale bread, hot, steaming French toast is a perfect addition to a brunch, as well as making a delicious cold ...

How to make french toast - Instructables

www.instructables.com/id/how-to-**make-french-toast**-1/ Instructables

Basic reading level

My boyfriend makes fantastic **french toast**. Super fantastic! I've tried making **french toast** lots of times, and it never turns out like I want it to, so...

Figure 8-44. Reading level shown in Google Results

Keyword Stuffing/Lack of Synonyms

Keyword stuffing used to be the basic approach of spammers back in the days when Alta Vista was the leading search engine. There are still many people who believe that repeating the same keyword over and over on a web page will lead to higher rankings for that phrase. They create documents that include way too many instances of the same word, and often don't make much use of synonyms for the key phrase. This is an indicator of a focus on search engines and not users, and may correlate with lower-quality documents.

Ad Density and Offensive Ads

Search engines have learned that sites that have too many ads on them offer poor user experiences. These pages may result in high bounce rates, but in some instances it is also possible to measure this effect directly. For example, consider the concept of measuring the percentage of the portion of a page above the fold that is occupied by ads. Too high a density might be taken as a negative signal. Google also has a patent on detecting annoying ads and pages, which Bill Slawksi has written about in "How Google Rejects Annoying Advertisements And Pages".

Sameness

An exact match search on *"mortgage tips"* returns hundreds of thousands of results, as shown in Figure 8-45. It is likely that, of those hundreds of thousands of results, there is very little variation in the content.

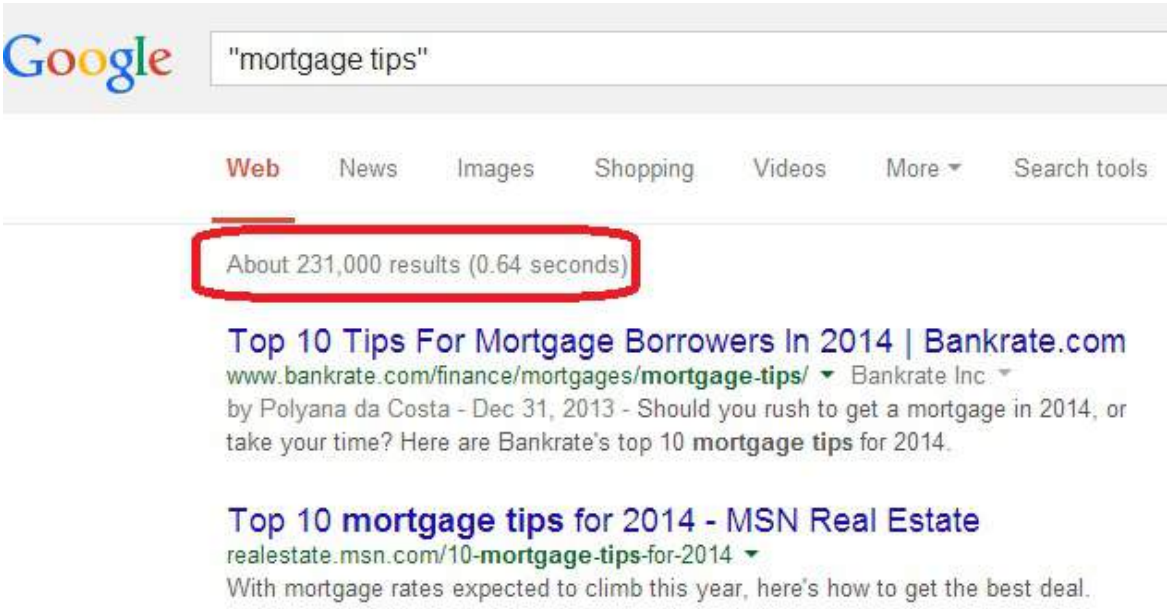


Figure 8-45. Search query returning hundreds of thousands of results

Broaden this concept to consider a site where the majority of articles are written about topics that many have written on before, and whose authors don't really bring any level of authority to the topic or offer a unique viewpoint. Such a site may be lowered in rankings, as users likely want to see authoritative opinions from those they trust.

Page Speed

In April 2010, Google announced that a website's page load time was now considered a ranking factor. However, Google has indicated that this affects only a small percentage of results (about 1%) and industry tests appear to confirm this, including the results of an investigation published by Moz.

The bottom line on page speed is that there are a lot of reasons why you should treat it as important, the biggest of which is the impact it has on conversion and bounce rates on your site. As an SEO factor, it is likely to have an impact on your results only if you have a particularly slow site. If that's the case, the drive to improve your conversion rate should already be compelling enough to address the issue.

Optimizing User Experience to Improve SEO

There are many reasons to address the user experience on your site, including conversion optimization and customer retention. As we have seen in the sections "User Engagement as a Measure of Search Quality" and "Document Analysis", SEO may also be impacted by the perceived user experience with a site. Measuring your site performance on these types of metrics can be a great start to improving the SEO performance of your website.

However, it is important to consider a broader question: What do your visitors want? If you are selling left-handed golf clubs, and you create a web page to rank on the search phrase *lefty ping putters*, what do you offer the visitors who come to that page from search? Reviews? Images? Videos? Putting tips? The ecommerce catalog front and center?

This is the most important part of optimizing a user experience. Creating a direct connection between the visitors arriving at your site, and what it is they are looking for. When you make this harder for them, metrics like bounce rate, time on site, and repeat visits will all suffer. When you make it easy, all of these metrics will improve.

One strategy is to survey your site visitors. If you don't have a lot of traffic, consider using a PPC campaign to bring in a number of visitors and then get them to take the survey. Next, we outline a three-step process to create a survey to find out from people what they are looking for on your site.

Step 1: Build a Survey

Start with a short series of questions asking the survey taker to imagine he's just performed the desired query. The basic structure is simple: request the top three content pieces your audience desires, then ask specifically about features that would make the page worthy of sharing (this is important, because it often differs substantively from what makes a page merely answer the user's query).

Finally, you can ask him to actually do a search on the target query (you don't want him to do it until the end, because what he finds might bias his responses) and report any results he liked, as these can provide additional insight.

Step 2: Send It to Your Customers/Potential Customers

You can find customers or potential customers virtually anywhere—among your friends, neighbors, coworkers, or social network connections, through existing SEO traffic or PPC campaigns, and so on. Anyone who fits your customer demographic or is creative enough to imagine herself as part of that demographic will work.

A link in the bottom of your email newsletter or a share on Facebook/LinkedIn/Twitter can often do the job, too. You might even try posting a link in a relevant industry forum or discussion group (so long as you're sure it won't be perceived as spammy).

Step 3: Record Responses and Leverage Them to Build What the People Want

Once you have your results, analyze them to see what you have. If you need things clarified, repeat the survey. It is better to do this than to act prematurely on incomplete or ambiguous data. Once you have built the site based on users' feedback, measure the results you get. Don't be afraid to go back and make more refinements, or even conduct another survey, to continue to educate yourself about what your visitors want.

Additional Social Media Resources

In this chapter, we have focused primarily on the interactions between social media and SEO. However, social media is itself a significant source of traffic. It deserves its own investigation and study. Here are some of the top resources available online to help guide you further into the world of social media:

- *http://www.socialmediaexaminer.com/*
- *http://blog.hubspot.com/*
- *http://socialmouths.com/blog/*
- *http://socialmediatoday.com/*
- *http://danzarrella.com/#*

- *http://www.convinceandconvert.com/blog/*
- *http://www.razorsocial.com/blog*
- *http://simplymeasured.com/blog/*
- *http://maximizesocialbusiness.com/*
- *http://www.briansolis.com/*
- *http://www.moz.com/blog*
- *http://marketingland.com*
- *http://allfacebook.com*
- *http://socialtriggers.com*
- *http://www.ducttapemarketing.com*
- *http://www.viralblog.com*
- *http://pushingsocial.com*
- *http://www.likeable.com/blog/*

Conclusion

While it does not appear that social media acts as a direct SEO ranking signal, except for personalized search and Google+, no one can debate that social media and user engagement have become significant factors for those practicing SEO. Links may still be the strongest source of signals to search engines, but these new factors are already an important part of the digital marketing landscape, and may gain additional weight over time as the search engines learn to use them better. Thus, a balanced SEO strategy needs to take into account social media, user data, and document analysis techniques.

In summary, brand publishers must seriously consider these four areas as a part

of their web promotional strategy:

- Participating in social media communities
- Providing an engaging user experience
- Offering unique and differentiated content
- Building a brand

Finding the right balance of these areas is already important for publishers to improve their conversion and SEO results, and it will only become more important over time.

[1] Danny Sullivan, "What Social Signals Do Google & Bing Really Count?", Search Engine Land, December 1, 2010, *http://searchengineland.com/what-social-signals-do-google-bing-really-count-55389*.

[2] Eric Enge, "Graph Search & Social Search With Bing's Stefan Weitz," Stone Temple Consulting, February 18, 2013, *https://www.stonetemple.com/graph-search-social-search-with-bings-stefan-weitz/*.

[3] Rand Fishkin, "Using Google+ to Appear in the Top Results Every Time - Whiteboard Friday," Moz Blogs, November 1, 2013, *http://bit.ly/using_google_plus*.

[4] For the full discussion, see "The Future of Search With Danny Sullivan".

[5] Ami Sedghi, "Facebook: 10 years of Social Networking, in Numbers," The Guardian's Technology Datablog, February 4, 2014, *http://bit.ly/fb_in_numbers*.

[6] See "Company | About".

[7] Kurt Wagner, "LinkedIn Hits 300 Million Users Amid Mobile Push," Mashable, April 18, 2014, *http://mashable.com/2014/04/18/linkedin-300-million-users/*.

[8] Claire Cain Miller, "The Plus in Google Plus? It's Mostly for Google," New York Times, February 14, 2014, *http://bit.ly/plus_for_google*.

[9] Piper Jaffray 27th Semi-Annual Taking Stock with Teens Survey, Spring 2014, *http://www.piperjaffray.com/private/pdf/TSWT%20Infographics.pdf*.

[10] Megan Ritter, "Market Your Brand to Pinterest's Highest User Base," Social

Media Week, August 14, 2014, *http://socialmediaweek.org/blog/2014/08/market-brand-pinterests-highest-user-base/*.

[11] Greg Jarboe, "Online Content Videos Break Record with 52 Billion Monthly Views," Search Engine Watch, January 13, 2014, *http://bit.ly/content_videos*.

[12] Tom Critchlow, "Twitter's t.co Link Shortening Service Is Game Changing - Here's Why," Distilled, August 18, 2011, *http://bit.ly/t_co_shortening*.

[13] Matt McGee, "Report: Google+ Now 2nd-Biggest Social Network Worldwide," Marketing Land, January 28, 2013, *http://bit.ly/google_plus_network*.

[14] Nick Bilton, Countering the Google Plus Image Problem," New York Times, March 6, 2012, *http://bit.ly/image_problem*.

[15] Mark Traphagen, "How Google Plus Profiles & Pages Gain Search Authority," Search Engine Land, November 21, 2013, *http://bit.ly/google_plus_search*; Eric Enge, "Google+ & SEO: How Google+ Impacts Search Results," Search Engine Watch, November 17, 2013, *http://bit.ly/plus_impact*.

[16] Conner Forrest, "Apple v. Google: The Goliath Deathmatch by the Numbers in 2014," TechRepublic, March 21, 2014, *http://bit.ly/apple_v_google_2014*.

[17] Michael Blair, "Windows OS Phones Should Increase Market Share in 2014 and Begin to Rival iOS," Seeking Alpha, March 6, 2014, *http://bit.ly/windows_mkt_share_2014*.

[18] Charles Farina, "Google Analytics Solidifies Lead in Fortune 500 Adoption in 2013," E-Nor, October 21, 2013, *http://bit.ly/google_analytics_2013*.

Panda, Penguin, and Penalties

Search engines use many techniques for measuring search quality, including manual penalties, algorithmic penalties, and ranking algorithms. Sometimes these algorithms and penalties can have a major impact on your organic traffic. Significant decreases in the search engine traffic to your website can be devastating to a business. As shown in Figure 9-1, sometimes these drops in traffic can be quite large.

Figure 9-1. Major loss in traffic

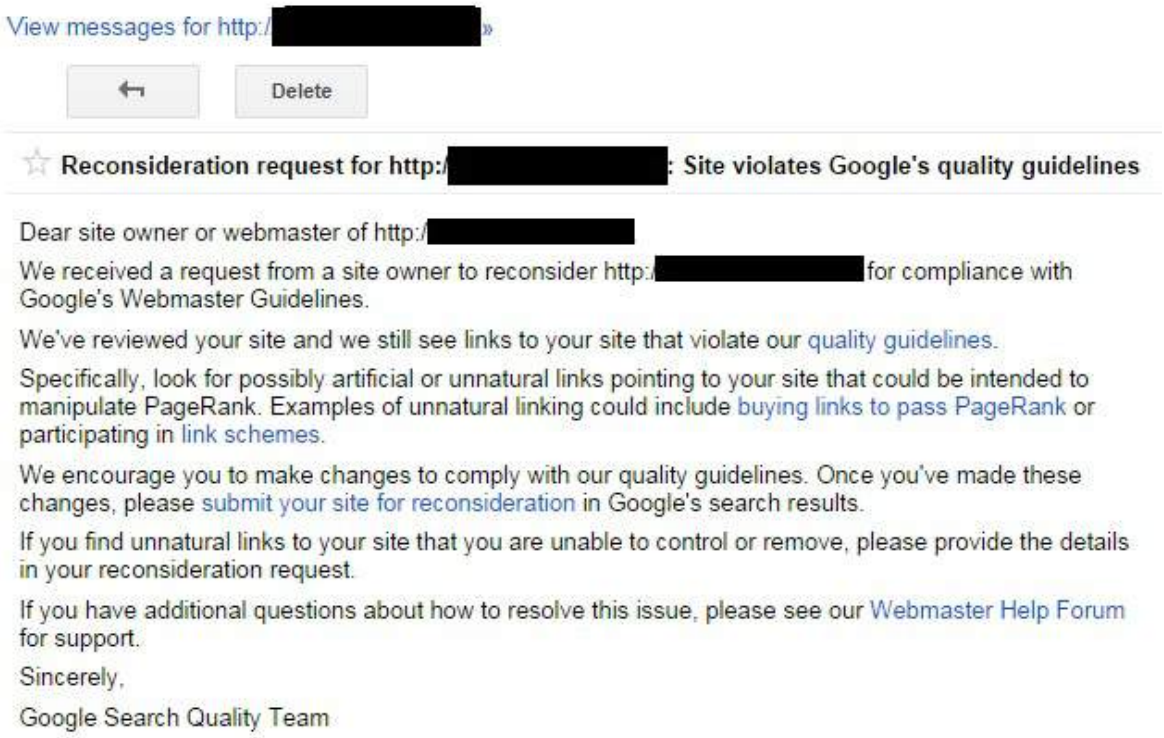
If the business shown in Figure 9-1 generates all of its revenue from organic search traffic, this would represent a devastating blow. This type of loss of revenue can mean laying off employees, or even closing the business.

For that reason, you need to have a working understanding of the potential causes of these traffic losses, and if you have already suffered such a loss, know how you can figure out what the cause is and what you need to do to recover.

Diagnosing the Cause of a Traffic Loss

The first step to identifying the cause of a traffic loss is to check your analytics data to see if the drop is in fact a loss of organic search engine traffic. If you have Google Analytics on your site, make sure you check your traffic sources, and then isolate just the Google traffic.

If you confirm that it is a drop in Google organic search traffic, then the next step is to check if you have received a message in Google Search Console indicating that you have been penalized by Google. Figure 9-2 shows an example of a manual penalty message from Google. This is what Google refers to as a *manual action*.


View messages for http:/[redacted]»

Delete

Reconsideration request for http:/[redacted]: Site violates Google's quality guidelines

Dear site owner or webmaster of http:/[redacted]

We received a request from a site owner to reconsider http:/[redacted] for compliance with Google's Webmaster Guidelines.

We've reviewed your site and we still see links to your site that violate our quality guidelines.

Specifically, look for possibly artificial or unnatural links pointing to your site that could be intended to manipulate PageRank. Examples of unnatural linking could include buying links to pass PageRank or participating in link schemes.

We encourage you to make changes to comply with our quality guidelines. Once you've made these changes, please submit your site for reconsideration in Google's search results.

If you find unnatural links to your site that you are unable to control or remove, please provide the details in your reconsideration request.

If you have additional questions about how to resolve this issue, please see our Webmaster Help Forum for support.

Sincerely,

Google Search Quality Team

Figure 9-2. Penalty message in Google Search Console

If you have received such a message, you now know what the problem is and you can get to work fixing it. It is not fun to have the problem, but knowing what you are dealing with is the first step in recovery.

If you don't have such a message, you will need to dig deeper to determine the source of your problem. The next step is to determine the exact date on which your traffic dropped. For example, if that date was April 24, 2012, take note of that. Then go to the Moz Google Algorithm Update Page and see if you can find the date of your traffic drop listed there.

Using our example of April 24, 2012, Figure 9-3 shows that sites that suffered traffic losses on this date were probably hit by Google's Penguin 1.0 algorithm

release.

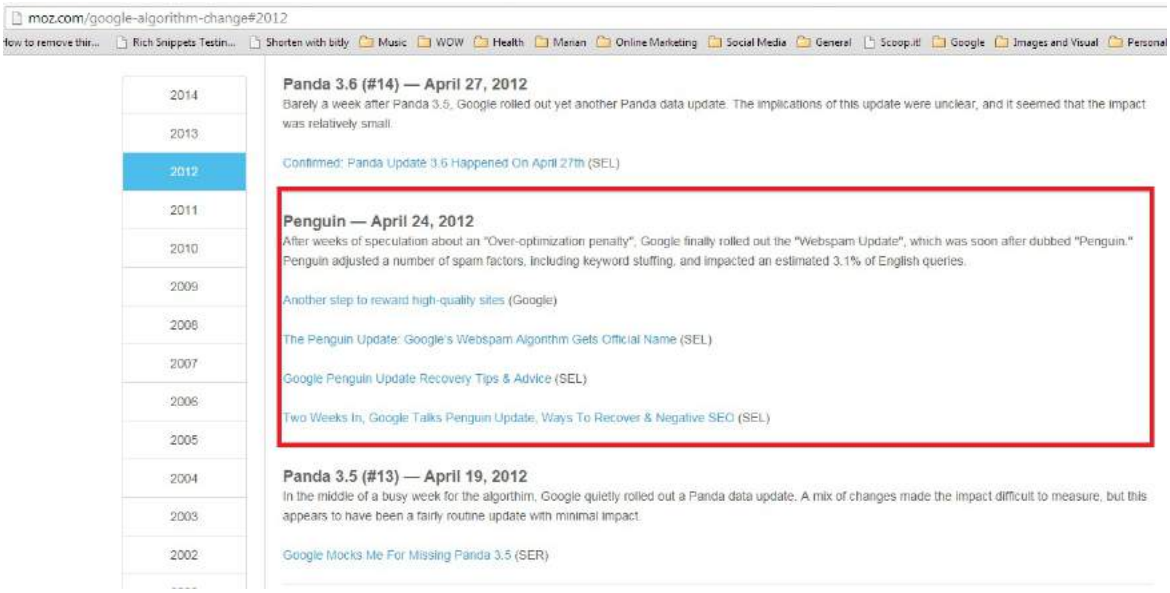


Figure 9-3. Moz Google algorithm update page showing Penguin 1.0

Barracuda Digital also offers a tool called the "Panguin Tool" that will overlay the Google algorithm updates on your Google Analytics data to make this date comparison analysis easier for you.

If you haven't gotten a message in Google Search Console, and the date of your traffic loss does not line up with a known Google algorithm update, the process of figuring out how to recover is much harder, as you don't know the reason for the drop.

Google does make smaller changes to its algorithms on a daily basis. From its perspective, these are smaller changes and not major updates. Such tweaks may be much harder to recover from as well. The best strategy is to focus on the best practices outlined in Chapter 6 and Chapter 7, or if you can afford SEO advice, bring in an expert to help you figure out what to do next.

Summary of Major Google Algorithms

The two most famous Google algorithms released since February 2011 are Panda and Penguin, but there are several others that Google has also released in that timeframe. Panda and Penguin are examined in detail later in this chapter. Here

is a summary of some of the other major algorithms released by Google:

Top Heavy Update

Released on January 19, 2012, this update impacts sites that have too many ads above the fold. This is true even if the ads are AdSense ads (from Google's own advertising programs). To avoid this penalty, avoid placing lots of ads near the top of your web pages. A banner ad is probably fine, as is something on the right rail of your site, but anything more than that starts to run into trouble with this algorithm. This algorithm was updated on October 9, 2012.

DMCA Penalty (a.k.a. Pirate)

Announced on August 10, 2012, this update applies penalties to sites that receive too many valid DMCA takedown requests. DMCA stands for Digital Millennium Copyright Act. You can file a DMCA takedown request if someone has stolen your content and published it as his own. Of course, to avoid this penalty, you should publish only your own original content. If you hire people to write for you, ensure that they are writing original content as well.

Exact-Match Domain (EMD)

With the release of this update on September 27, 2012, Google started to change the way it values exact-match domains. An example of an exact-match domain might be *http://www.blue-widgets.com* (note that this is not currently a real site, and is in fact a parked page), which historically would have had an advantage in ranking for the search phrase *blue widgets*. The purpose of this update was to reduce or eliminate that advantage.

Payday Loan (spammy sites)

Released on June 11, 2013, this algorithm targeted sites that it considered to be in particularly spammy market spaces. Google specifically mentioned payday loans, casinos, and pornography as targeted markets. This algorithm was updated on May 16, 2014.

Hummingbird

> Announced on September 26, 2013, Hummingbird was not really an algorithm update, but instead a rewrite of Google's search platform, with a design goal of making the overall search engine more flexible and adaptable for the future. Part of the design of this rewrite included more support for natural language queries and semantic search.

These are some of the major releases, but as we've noted, Google makes many smaller changes as well. In August 2014, the head of search quality for Google, Amit Singhal, indicated that it had made 890 changes in the past year.[1]

Panda

Google's Panda algorithm shook up the search landscape when it came out on February 24, 2012. In their announcement of the release, Google said the following:

> *Many of the changes we make are so subtle that very few people notice them. But in the last day or so we launched a pretty big algorithmic improvement to our ranking—a change that noticeably impacts 11.8% of our queries—and we wanted to let people know what's going on. This update is designed to reduce rankings for low-quality sites—sites which are low-value add for users, copy content from other websites, or sites that are just not very useful. At the same time, it will provide better rankings for high-quality sites—sites with original content and information such as research, in-depth reports, thoughtful analysis, and so on.*

Other information also came out pretty quickly about this release, which Danny Sullivan of Search Engine Land initially called the Farmer update.[2] Between this additional information and the initial announcement, several aspects of Panda quickly became clear:

- This was a very large change. Very few algorithm updates by Google impact more than 10% of all search queries.
- This algorithm is focused on analyzing content quality.

- Scraper sites were targeted.
- Sites lacking substantial unique information were also targeted. In particular, content farms were part of what Google was looking to address—a move that Eric Enge predicted would happen three weeks before this release.[3]
- Google clearly states a strong preference for new research, in-depth reports, and thoughtful analysis.
- The Panda algorithm does not use the link graph as a ranking factor.

The second release of Panda came out on April 11, 2011. What made this release particularly interesting is that it incorporated data gathers from Google's Chrome Blocklist Extension. This extension to Chrome allowed users to indicate that they wanted pages removed from the search results. Figure 9-4 illustrates what this looks like in action.

Figure 9-4. Chrome Blocklist Extension

This was the first time that Google ever publicly confirmed that it was using a form of direct user input as a ranking factor in any of its algorithms.

Initially, Panda was focused only on the United States, but was rolled out internationally on August 12, 2011, to most of the rest of the world except Japan, China, and Korea.

Google has done many releases for Panda since the initial one. Table 9-1 provides a full list of all known Panda releases.

Table 9-1. List of Panda releases

Panda release	Date	Notes
Panda 1.0	February 23, 2011	Impacted 11.8% of search results.
Panda 2.0	April 11, 2011	Added use of the Chrome Blocklist Extension.
Panda 2.1	May 9, 2011	
Panda 2.2	June 21, 2011	
Panda 2.3	July 23, 2011	
Panda 2.4	August 12, 2011	Rolled out internationally; impacted 6–9% of queries internationally.
Panda 2.5	September 28, 2011	
Panda “Flux”	October 5, 2011	Called “Flux” as a result of a Matt Cutts tweet; impacted 2% of queries.
Minor update	October 11, 2011	
Panda 3.1	November 18, 2011	Naming skipped an official 3.0 release.
Panda 3.2	January 18, 2012	
Panda 3.3	February 27, 2012	
Panda 3.4	March 23, 2012	Impacted 1.6% of results.
Panda 3.5	April 19, 2012	
Panda 3.6	April 27, 2012	

Panda 3.7	June 8, 2012	Impacted less than 1% of queries.
Panda 3.8	June 25, 2012	
Panda 3.9	July 24, 2012	Impacted about 1% of queries.
Panda #18	August 20, 2012	Industry agreed on nomenclature change to counting the release number.
Panda #19	September 18, 2012	
Panda #20	September 27, 2012	Impacted 2.4% of queries.
Panda #21	November 5, 2012	Impacted 1.1% of queries.
Panda #22	November 21, 2012	
Panda #23	December 21, 2012	Impacted 1.3% of queries.
Panda #24	January 22, 2012	Impacted 1.2% of queries.
Panda #25	March 14, 2013	Date estimated by Moz. Google announced Panda would be made part of the core algorithm. As of this date, rolling updates are no longer announced, but happen monthly and roll out over 10 days or so.
Panda #26	June 11, 2013	
Panda #27	July 18, 2013	Rumored to soften Panda's impact.
Panda 4.0	May 20, 2014	Major update confirmed by Google. Also appeared to soften Panda's impact on many sites, but new sites were also hit.

As of March 14, 2013, Google announcements of Panda updates became relatively rare, and Table 9-1 shows only the announced updates. However, these updates appear periodically as unannounced data refreshes and roll out over a period of more than a week.

Since that time, Google has confirmed only three Panda updates, with the most recent one being Panda 4.0 on May 20, 2014.

To track all Google updates over time, you can check the Google algorithm change history page on the Moz website.

Target Areas of Panda

Google has historically offered relatively vague information on how the Panda algorithm works to determine the quality of a site. For example, on May 6, 2011, Amit Singhal offered his advice on building high-quality sites. In it, he suggested a list of questions that you could use to determine if you were on such a site:

- Would you trust the information presented in this article?
- Is this article written by an expert or enthusiast who knows the topic well, or is it more shallow in nature?
- Does the site have duplicate, overlapping, or redundant articles on the same or similar topics with slightly different keyword variations?
- Would you be comfortable giving your credit card information to this site?
- Does this article have spelling, stylistic, or factual errors?
- Are the topics driven by genuine interests of readers of the site, or does the site generate content by attempting to guess what might rank well in search engines?
- Does the article provide original content or information, original reporting, original research, or original analysis?
- Does the page provide substantial value when compared to other pages in search results?
- How much quality control is done on content?
- Does the article describe both sides of a story?
- Is the site a recognized authority on its topic?
- Is the content mass-produced by or outsourced to a large number of creators, or spread across a large network of sites, so that individual pages or sites don’t get as much attention or care?

- Was the article edited well, or does it appear sloppy or hastily produced?
- For a health-related query, would you trust information from this site?
- Would you recognize this site as an authoritative source when mentioned by name?
- Does this article provide a complete or comprehensive description of the topic?
- Does this article contain insightful analysis or interesting information that is beyond obvious?
- Is this the sort of page you'd want to bookmark, share with a friend, or recommend?
- Does this article have an excessive amount of ads that distract from or interfere with the main content?
- Would you expect to see this article in a printed magazine, encyclopedia, or book?
- Are the articles short, unsubstantial, or otherwise lacking in helpful specifics?
- Are the pages produced with great care and attention to detail versus less attention to detail?
- Would users complain when they see pages from this site?

There are a few key points that can be extracted from this advice, and the industry has been able to determine and clarify a number of Panda's target areas. These include:

Thin content

As you might expect, this is defined as pages with very little content. Examples might be user profile pages on forum sites with very little information filled in, or an ecommerce site with millions of products, but very little information provided about each one.

Unoriginal content

These may be scraped pages, or pages that are only slightly rewritten, and Google can detect them relatively easily. Sites with even a small number of these types of pages can be impacted by Panda.

Nondifferentiated content

Even if you create all original articles, this may not be enough. If every page on your site covers topics that have been written about by others hundreds or thousands of times before, then you really have nothing new to add to the Web with your site. Consider, for example, the number of articles in the Google index about making French toast, as shown in Figure 9-5. There are over 30,000 pages on the Web that include the phrase *how to make french toast* in their title. From Google's perspective, it don't need another web page on that topic.

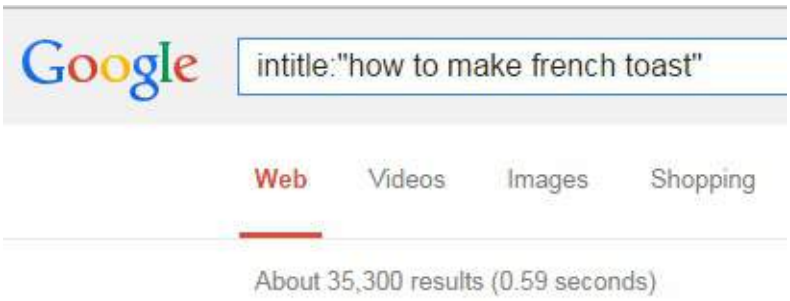


Figure 9-5. There are tens of thousands of pages on "how to make french toast"

Poor-quality content

This is content that is inaccurate or poorly assembled. In many cases, this may be hard to detect, but as mentioned in Amit Singhal's article, one indicator is content that includes poor grammar or a lot of spelling mistakes. Google could also potentially use fact checking as another way to determine poor-quality content.

Curated content

Sites that have large numbers of pages with lists of curated links do get hit by

Panda. Content curation is not inherently bad, but if you are going to do it, it's important to incorporate a significant amount of thoughtful commentary and analysis. Pages that simply include lots of links will not do well, nor will pages that include links and only a small amount of unique text. Content curation is explored in depth in "Content Curation & SEO: A Bad Match?".

Thin slicing

This was believed to be one of the original triggers for the Panda algorithm, as it was a popular tactic for *content farms*. Imagine you wanted to publish content on the topic of schools with nursing programs. Content farm sites would publish many articles on the same topic, with titles such as: "nursing schools," "nursing school," "nursing colleges," "nursing universities," "nursing education," and so forth. There is no need for all of those different articles, which prompted Google to target this practice with Panda.

Database-generated content

The practice of using a database to generate web pages is not inherently bad, but many companies were doing it to an excessive scale. This led to lots of thin-content pages or poor-quality pages, so many of these types of sites were hit by Panda.

Importance of Diversity in Rankings

Diversity is important to overall search quality for Google. One simple way to illustrate this is with the search query *Jaguar*. This word can refer to an animal, a car, a guitar, an operating system, or even an NFL team. Normal ranking signals might suggest the results shown in Figure 9-6.

Figure 9-6. Normal ranking signals may show these results for "Jaguar"

Note that the search results at the top all focus on the car, which may be what the basic ranking signals suggest the searcher is looking for. However, if the searcher is looking for information on the animal, those results are pushed down a bit. As a result, Google may use other signals to decide to alter the results to look more like those shown in Figure 9-7.

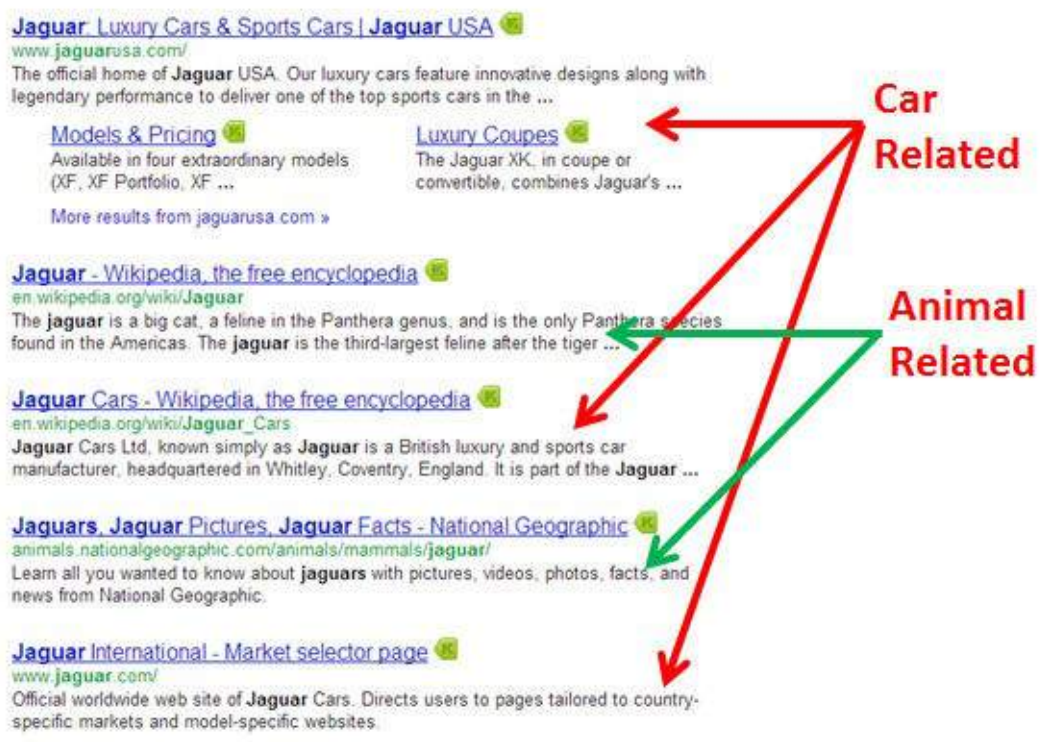


Figure 9-7. "Query Deserves Diversity" may alter the results for "Jaguar"

In this version of the results, one of the animal-related results has been pushed into the second position. Google makes these types of adjustments to the SERPs using a concept known as *Query Deserves Diversity*.

Google makes these adjustments by measuring user interaction with the search results to determine what ordering of the results provides the highest levels of user satisfaction. For example, even if traditional ranking signals would put another page for the car next, it might make sense for the next result to be about the animal.

Role of Authority in Rankings

Consider again the search query *how to make French toast* we showed in Figure 9-5. While Google has plenty of results on the topic, there are, of course, some sites that rank highly for this search query. How is their rank determined?

Very high-authority sites are likely to do fine when publishing content on a topic

that is already well covered on the Internet. There are a few possible reasons why this is the case:

- Reputation and authority is a big factor. For example, if the *New York Times* Lifestyle section posted a new article on how to make French toast, even though it is not particularly unique, readers might respond well to it anyway. User interaction signals with the search result for that content would probably be quite strong, simply because of the site's reputation.

- High-authority sites probably got to be that way because they don't engage in much of the behavior that Google does not like. Chances are that you won't find a lot of thin content, "me too" content, thin slicing, or any of the issues that are Panda triggers.

- Google may simply be applying looser criteria to a high-authority site than it does to other sites.

Exactly what factors allow higher-authority sites to have more leeway is not clear. Is it that Google is measuring user interaction with the content, the quality of the content itself, the authority of the publisher, or some combination of these factors? There are probably elements of all three in what Google does.

Impact of Any Weak Content on Rankings

Weak content on even one single section of a larger site can cause Panda to lower the rankings for the whole site. This is true even if the content in question makes up less than 20% of the pages for the site. When you are putting together a plan to recover from Panda, it is important to take this into account.

Path to Recovery

The road to recovery from a Panda penalty may be a long one. Oftentimes it requires a substantial reevaluation of your site's business model. You need to be prepared to look at your site with a highly critical eye, and this is often very hard to do with your own site. Thus, it's a good idea to consider bringing in an external perspective to evaluate your site. You need someone who is willing to look you in the eye and tell you that your baby is ugly.

Once you go through this reevaluation process, you may realize that even the basic premise of your site is broken, and that you need to substantially restructure it. Making these types of decisions is quite hard, but you need to be prepared to do it.

As you consider these tough choices, it can be helpful to look at your competition that did not get hit. Understand, however, that you may see instances of thin content, weak content, “me too” content, and other poor-quality pages on competitors’ sites that look just as bad as the content penalized on your site, and they may not appear to have been impacted by Panda. Don’t let this type of analysis deter you from making the hard choices. There are so many factors that Google uses in its ranking algorithms that you will never really know why your site was hit by Panda and your competitor’s site was not.

What you do know is that Google’s Panda algorithm does not like something about your site. This may include complex signals based on how users interact with your listings in the search results, which is data you don’t have access to.

To rebuild your traffic, it’s best to dig deep and take on hard questions about how you can build a site full of fantastic content that gets lots of user interaction and engagement. While it is not believed that social media engagement is a factor in Panda, there is likely a strong correlation between high numbers of social shares and what Google considers to be good content.

Highly differentiated content that people really want, enjoy, share, and link to is what you want to create on your site (the article “After Google Panda, Standout Content Is Essential” expands on this concept). There is a science to creating content that people will engage with. We know that picking engaging titles for the content is important, and that including compelling images matters too. Make a point of studying how to create engaging content that people will love, and apply those principles to every page you create. In addition, measure the engagement you get, test different methods, and improve your ability to produce great content over time.

Ways to address weak pages

As you examine your site, a big part of your focus should be addressing its weak pages. They may come in the form of an entire section of weak content, or a number of pages interspersed among the higher-quality content on your site.

Once you have identified those pages, there are a few different paths you can take to address the problems you find:

- Improve the content. This may involve rewriting the content on the page, and making it more compelling to users who visit.
- Add the `noindex` meta tag to the page (you can read about how to do this in Chapter 6). This will tell Google to not include these pages in its index, and thus will take them out of the Panda equation.
- Delete the pages altogether, and 301-redirect visitors to other pages on your site. Use this option only if there are quality pages that are relevant to the deleted ones.
- Delete the pages and return a 410 HTTP status code when someone tries to visit the deleted page. This tells the search engine that the pages have been removed from your site.
- Use the URL removal tool to take the page out of Google's index. This should be done with great care. You don't want to accidentally delete other quality pages from the Google index!

Expected timeline to recovery

Even though they are no longer announced, Panda releases come out roughly once per month. However, once you have made the necessary changes, you will still need to wait. Google has to recrawl your site to see what changes you have made. It may take Google several months before it has seen enough of the changed or deleted pages to tilt the balance in your favor.

What if you don't recover?

Sadly, if your results don't change, this usually means that you have not done enough to please the Panda algorithm. It may be that you were not strict enough in deleting poorer-quality content from your site. Or, it may mean that Google is not getting enough signals that people really care about what they find on your site.

Either way, this means that you have to keep working to make your site more interesting to users. Go beyond viewing this process as a way to deal with Panda,

and instead see it as a mission to make your site one of the best on the Web.

This requires substantial vision and creativity. Frankly, it's not something that everybody can accomplish without making significant investments of time and money. One thing is clear: you can't afford to cut corners when trying to address the impact of the Panda algorithm.

If you have invested a lot of time and made many improvements, but you still have content that you know is not so great, chances are pretty good that you haven't done enough. You will likely find yourself four months later wishing that you had kept at the recovery process.

In addition, the Panda algorithm is constantly evolving. Even if you have not been hit by this algorithm, the message from Google is clear: it is going to give the greatest rewards to sites that provide fantastic content and great user experiences. Thus, your best path forward is to be passionate about creating a site that offers both. This is how you maximize your chances of recovering from Panda, and from being impacted by future releases.

Successful Panda recovery example

In Figure 9-1, we showed an example of a site suffering a major traffic loss. The site in question, called The Riley Guide, is dedicated to providing information on jobs and careers. A thorough examination of the site showed that it had two major problems:

- The content on careers was too similar to the type of content you can find on a major government website, *http://www.bls.gov*. Even though The Riley Guide's content was not copied from that location, it did not offer any unique differentiation or value to the Web.

- The content on the rest of the site mainly consisted of links to resources for job information elsewhere on the Web. This content was structured like a librarian's index, with links and short commentaries on each. In some cases, the commentary was copied from the target website.

In short, the original site structure did not offer a lot of value. The publishers of The Riley Guide took the site through a complete overhaul. This was a lengthy process, as about 140 pages had to be rewritten from scratch, and another 60 or so pages had the `noindex` tag applied.

Figure 9-8 shows the results of this effort. Not only did the site recover from Panda, but it is now hitting new highs for traffic.

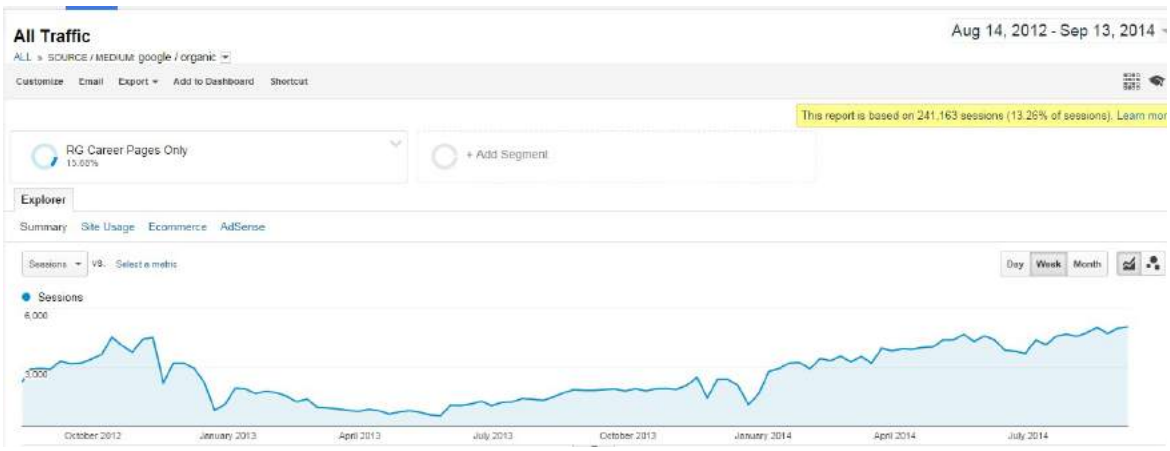


Figure 9-8. Panda traffic recovery

Penguin

Google's Penguin algorithm was first released to the world on April 24, 2012. It was the first major algorithm implemented by Google to address bad links. Like Panda before it, the Penguin release shook up the search landscape. Since the initial release of this algorithm, there have been several incremental releases, and an expansion of the scope of the types of links addressed.

Penguin algorithm hits are generally reasonably easy to detect, as there have only been five total releases of Penguin. These are the only dates on which you may have been hit by the alogrithm, and unfortunately, also appear to be the only dates on which you can recover from it. Table 9-2 lists all of the releases of Penguin as of June 2015.

Table 9-2. List of Penguin releases

Penguin release	Date	Notes
Penguin 3.0	October 17, 2014	Impacted 1.0% of queries
Penguin 2.1	October 4, 2013	
Penguin 2.0	May 22, 2013	Began targeting links more at the page level.
Penguin 1.2	October 5, 2012	Impacted about 0.3% of queries.

Penguin 1.1	May 25, 2012	
Penguin 1.0	April 24, 2012	Impacted 3.1% of U.S. search queries.

As you can see, the releases are quite rare, and this is one of the more jarring aspects of Penguin. Once you have been hit by it, you have to wait until the next release to have a chance to recover. That could be as long as one full year!

For purposes of diagnosis, if your traffic experiences a significant drop on one of the dates listed in Table 9-2, it's likely that you have been hit by Penguin.

Target Areas of Penguin

As you will see in the section "Links Google Does Not Like", there are many different types of links that are considered bad links. However, Penguin appears to focus on a more limited set of link types. These are:

Article directories

From the very first release of Penguin, Google targeted sites that obtained links from article directories. While these links may not always reflect manipulative intent by the publisher, Google found that people who leveraged article directories tended to operate lower-quality sites.

Cheap directories

These were also targeted in the very first release of Penguin. There are a few directories that are genuinely high quality, such as the Yahoo! Directory, DMOZ, Business.com, Best of the Web, and perhaps a few others specific to your industry vertical. Stay away from the rest.

Excessive rich anchor text

Excessive use of rich anchor text was also a part of the initial release of Penguin. Specifically, Penguin targeted too many instances of the same anchor text pointing to any of the URLs on your site. Google does not expect, or want, all links to say, "click here," but it sees it as a signal of spammy behavior when the exact same keyword-rich anchor text is used repeatedly.

Low-relevance international links

While it is not confirmed that this target area is a part of any Penguin release, anecdotal evidence suggests it might be. Consider any links from countries where you don't sell your products or services as a potential problem, unless they come from truly high-quality sites.

Comment spam

Excessively implementing links in comments on other people's blog posts and forums is also a problem for Penguin.

There may be more types of links that are a part of the Penguin algorithm or that could be included in future releases. A complete summary of problem links is discussed in "Links Google Does Not Like".

Path to Recovery

The first step on the road to recovery from a Penguin hit is to realize that Penguin releases are rare. As they happen roughly twice per year, you don't want to get cute and try to save some of your questionable links. Be aggressive and clean up every possible problem. Certainly, address the types of links listed in "Target Areas of Penguin", but you should also seriously consider dealing with all of the link problems listed in "Links Google Does Not Like".

Do not file a reconsideration request, as it is a waste of time. In fact, you can't file one unless you have a manual penalty on your site. Google does not have the mechanisms in place to adjust how the Penguin algorithm impacts you.

After you have addressed the bad links to your site, you must wait. Once the next Penguin release arrives, you will know whether you have done enough. If you are not successful in recovering during that release, you have no choice but to figure out which links you failed to clean up and try again for the next release. If you are successful, congratulations!

Penalties

There are two types of penalties, algorithmic and manual. Algorithmic penalties

do not involve any human component, whereas manual penalties do. While the details of what prompts Google to perform a manual review of a website are not always evident, there appear to be several ways that manual reviews can be triggered. Figure 9-9 illustrates how a manual review might be triggered in the case of a site that has problems with its link profile.

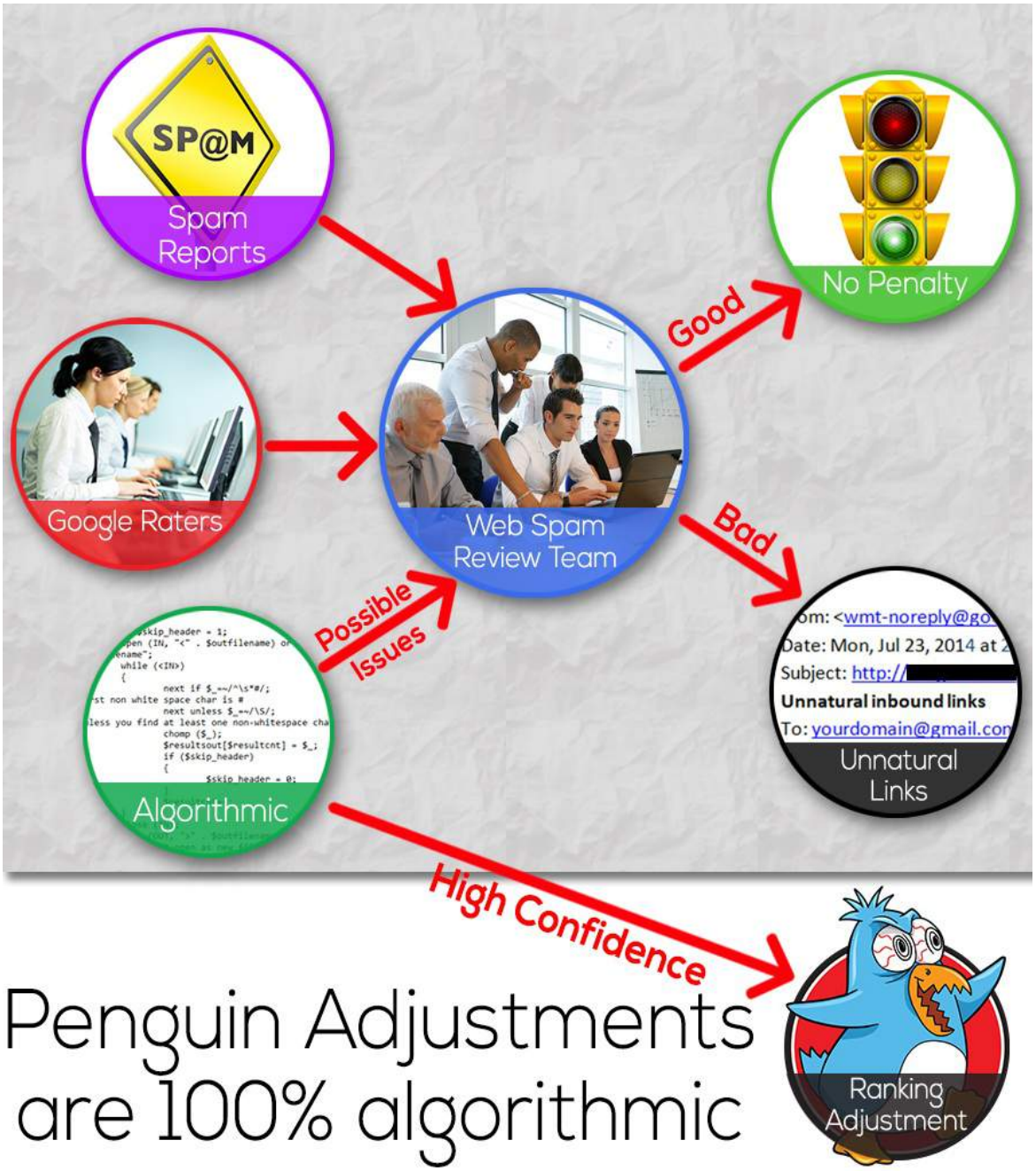


Figure 9-9. Possible ways that Google manual reviews may be triggered

Note that in some cases an algorithm may trigger an algorithmic ranking adjustment (as shown in Figure 9-9, algorithmic adjustments are made only when Google's confidence in the signals is very high; if the confidence level is not high but there are indications of a problem, a manual review might be initiated). Here is a summary of the major potential triggers:

Spam report

Any user (including your competitor) can file a spam report in Google. Google receives large volumes of these reports every day. Google evaluates each report, and if it finds one credible (it may run some type of algorithmic verifier to determine that), then it conducts a manual review.

Algorithmically triggered review

While this approach has never been verified by Google, it's likely that Google uses algorithms to trigger a manual review of a website. The premise is that Google uses algorithms like Panda, Penguin, and others that identify large quantities of sites whose behavior is bad, but not bad enough for Google to algorithmically penalize them, so these sites would be queued for manual review. Google could also implement custom algorithms designed to flag sites for review.

Regular search results reviews

Google maintains a large team of people who perform manual reviews of search results to evaluate their quality.[4] This effort is primarily intended to provide input to the search quality team at Google that they can use to help them improve their algorithms. However, it is quite possible that this process could also be used to identify individual sites for further scrutiny.

Once a review is triggered, the human reviewer uses a set of criteria to determine if a penalty is merited. Whatever the outcome of that review, it is likely that Google keeps the notes from the review in a database for later use. Google most likely keeps a rap sheet on all webmasters and their previous infractions, whether they result in a penalty or not.

Google Search Console Messages

As of April 2012, Google has maintained a policy of sending all publishers that receive a manual penalty a message in their Search Console describing the nature of the penalty. These messages describe the penalty in general terms, and it is up to the publisher to figure out how to resolve it. Generally, the only resource that Google provides to help with this is its Webmaster Guidelines. If you receive such a message, then the reconsideration request option in Google Search Console becomes available.

Types of Manual Penalties

Manual penalties come in many forms. The most well-known types of penalties are link related, but you can also get a variety of other penalties. Some of the most common types of manual penalties are discussed in the following sections.

Thin-content penalties

This penalty relates to pages that don't add enough value to users in Google's opinion. Figure 9-10 shows an example of a thin-content message from Google in Search Console.

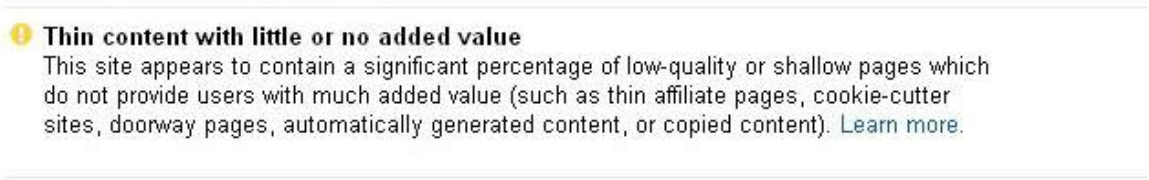


Figure 9-10. Thin-content penalty message

Unfortunately, when you receive this type of penalty, Google doesn't provide any guidance on what the cause might be. It does tell you that it is a thin-content penalty, but the rest is up to you. There are four primary triggers for thin-content penalties:

Pages with little useful content

As the name of the penalty suggests, pages with very little content are potential triggers for this penalty. This is especially true if there are a large number of these pages, or if there is a particular section on the site that has a significant percentage of its pages deemed thin.

Thin slicing

This happens to publishers who implement pages that are really designed to just garner search traffic. What these publishers often do is build pages for each potential search query a visitor might use, even if the variations in the content are quite small or insignificant. To use an earlier example, imagine a site with information on nursing schools with different pages with the following titles:

- Nursing schools
- Nursing school
- Nursing colleges
- Nursing universities
- Best nursing schools

Sometimes publishers do this unintentionally, by autogenerating content pages based on queries people enter when using the search function for the website. If you decide to do something like this, then it's critical to have a detailed review process for screening out these thin-slicing variants, pick one version of the page, and focus on it.

Doorway pages

These are pages that appear to be generated just for monetizing users arriving from search engines. One way to recognize these types of pages is that they are usually pretty much standalone pages with little follow-on information available, and/or they are pages that are largely written for search engines and not users.

The user arriving on these pages basically has two choices: buy now, or leave.

Poor integration into the overall site

Another issue to look for is whether parts of your site are not well integrated

into the rest of the site. Is there a simple way for users to get to these pages from the home page, from the main navigation of the site, or at least from a major section of the site?

If you have a section that appears to be isolated from the rest of your site, that could result in a thin-content penalty.

Once you believe you have resolved these issues, you need to submit a reconsideration request. You can read more about this in "Filing reconsideration requests". Once you have filed this request, you simply wait until Google provides a response. This process normally takes two to three weeks.

If you are successful, then you are in good shape and just need to make sure not to overstep your boundaries again in the future. Otherwise, it's back to the drawing board to see what you might have missed.

Partial link penalties

Another possible manual penalty is a partial link penalty. This is sometimes called an "impacts links" penalty, as that term is part of the message you get from Google (see Figure 9-11). These penalties indicate that one or a small number of your pages have been flagged for bad linking behavior.

Normally, only the rankings and traffic for those particular pages suffer as a consequence of this penalty.

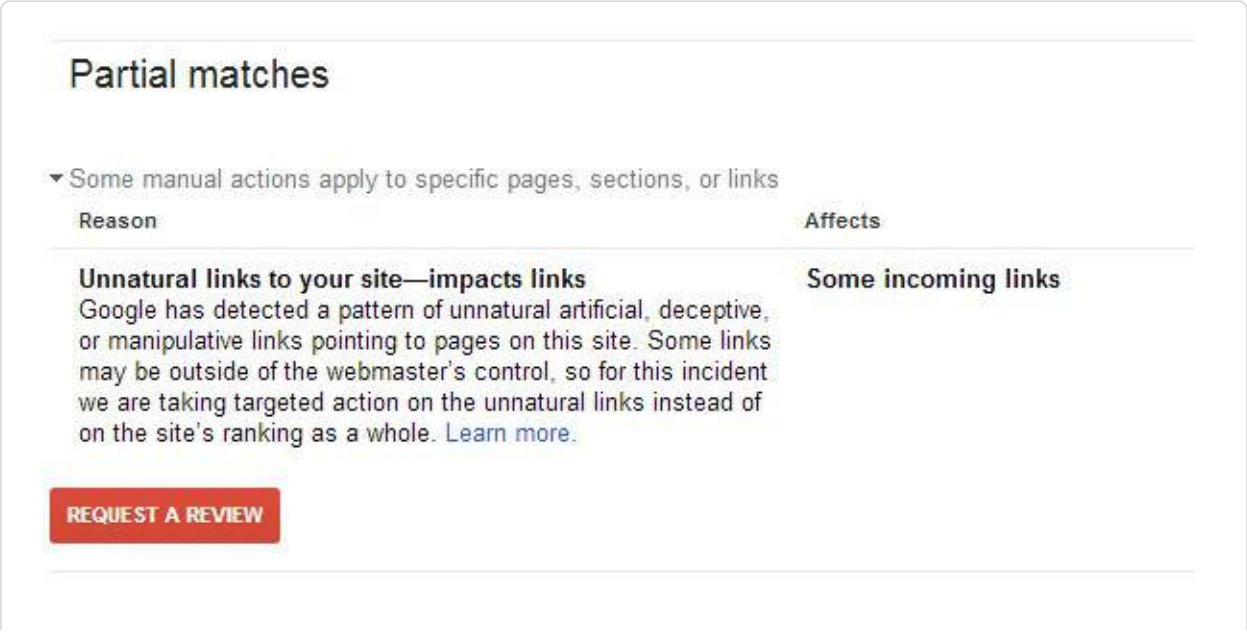


Figure 9-11. Partial link penalty message

Unfortunately, Google does not tell you which of your pages is receiving the penalty, so you have to determine that for yourself. This penalty is normally caused by too many questionable or bad links to pages other than your home page.

The cause is often a link-building campaign focused on bringing up the rankings and search traffic to specific money pages on your site. One of the more common problems is too many links with keyword-rich anchor text pointing to those pages, but other types of bad links can be involved as well. The steps to recover from this type of penalty are:

1. Pull together a complete set of your links as described in "Sources of Data".

2. Look for pages on your site, other than the home page, that have the most links.

3. Examine these pages for bad links as described in "Links Google Does Not Like".

4. Use the process described in "Link Cleanup Process" to deal with the bad links.

5. Submit a reconsideration request as described in "Filing reconsideration requests".

Once you have sent in the reconsideration request, the only thing you can do is wait. As noted previously, it normally takes two to three weeks before you get a response. Google will either let you know you have succeeded and confirm that it has removed the penalty, or it will tell you that you have failed, in which case you have to take a deeper look at your links and figure out what you missed in your previous attempt.

Sitewide link penalties

Manual link penalties can also be applied on a sitewide basis. This usually means more than a few pages are involved, and may well also involve the home

page of the site. With this type of penalty, rankings are lowered for the publisher on a sitewide basis.

As a consequence, the amount of lost traffic is normally far more than it is for a partial link penalty. Figure 9-12 shows an example of a sitewide link penalty message.

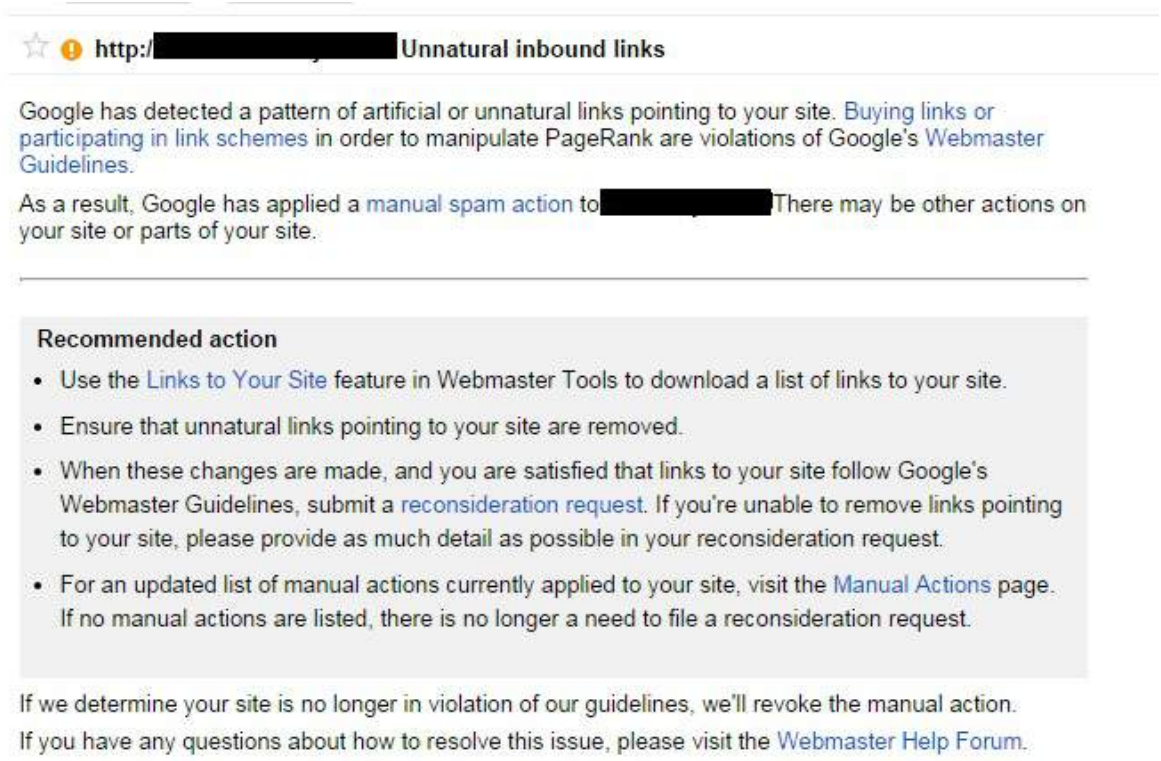


Figure 9-12. Sitewide link penalty message

The steps to recover from this type of penalty are:

1. Pull together a complete set of your links as described in "Sources of Data".

2. Examine this list for bad links as described in "Links Google Does Not Like".

3. Use the process described in "Link Cleanup Process" to deal with the bad links.

4. Submit a reconsideration request as described in “Filing reconsideration requests”.

Once you have sent in the reconsideration request, the only thing you can do is wait. It normally takes two to three weeks before you get a response. Google will either let you know you have succeeded and confirm that it has removed the penalty, or it will tell you that you have failed, in which case you have to take a deeper look at your links and figure out what you missed in your previous attempt.

Other types of manual penalties

Some of the other manual penalties include:

Cloaking and/or sneaky redirects

You can get this message if Google believes you are showing different versions of pages to Googlebot than you show to users. To diagnose this, use the “Fetch and Render as Google” tool in Search Console to retrieve the page. Use the tool to load the same page in another browser window and compare the two pages.

If you see differences, invest the time and effort to figure out how to remove the differing content. You should also check for URLs that redirect and send people to pages that are not in line with what they expected to see—for example, if they click on anchor text to read an article about a topic of interest but instead find themselves on a spammy page trying to sell them something.

Another potential source of this problem is *conditional redirects*, where users coming from Google search, or a specific range of IP addresses, are redirected to different pages than other users.

Hidden text and/or keyword stuffing

This message is generated if Google believes you are stuffing keywords into your pages for the purpose of manipulating search results—for example, if you put content on a page with a white background using a white font, so it’s invisible to users but search engines can still see it.

Another way to generate this message is to simply repeat your main keyword for a page over and over again in hopes of influencing search results.

User-generated spam

This type of penalty is applied to sites allowing user-generated content (UGC) that are perceived to not be doing a good job of quality control on that content. It's very common that sites with UGC become targets for spammers uploading low-quality content with links back to their own sites.

The short-term fix for this is to identify and remove the spammy pages. The longer-term fix is to implement a process for reviewing and screening out spammy content to prevent it from getting on your site in the first place.

Unnatural links from your site

This is an indication that Google believes you are selling links to third parties, or participating in link schemes, for the purposes of passing PageRank. The fix is simple: remove the links on your site that look like paid links, or add a `nofollow` attribute to those links.

Hacked site

Google will communicate this penalty by sending you a message in Search Console and/or by showing indications that your site has been hacked (and is dangerous to visit) in the search results. The most common cause for this penalty is failing to keep up with updates to your content management system (CMS).

Spammers take advantage of vulnerabilities in the CMS to modify your web pages, most often for the purpose of inserting links to their own sites, but sometimes for more nefarious purposes such as accessing credit card data or other personally identifiable information.

To resolve the problem, you will need to determine how your site has been hacked. If you don't have technical staff working for you, you may need to get help to detect and repair the problem. To minimize your exposure going forward, always keep your CMS on the very latest version possible.

Pure spam

Google will give you this message in Search Console if it believes that your site is using very aggressive spam techniques. This can include things such as automatically generated gibberish or other tactics that appear to have little to do with trying to add value for users.

If you get this message, there is a strong chance that you should simply shut down the site and start with a new one.

Spammy freehosts

Even if your site is clean as a whistle, if a large percentage of the sites using your hosting company are spamming, Google may take action against all of the sites hosted there. Take care to make sure you are working with a highly reputable hosting company!

For any of these problems, you need to address the source of the complaints. When you believe you have done so, follow the procedure outlined in "Filing reconsideration requests".

Links Google Does Not Like

To understand the types of links that Google does not like, we need only review Larry Page and Sergey Brin's original thesis, "The Anatomy of a Large-Scale Hypertextual Web Search Engine"). At the beginning of the thesis is this paragraph:

> *The citation (link) graph of the web is an important resource that has largely gone unused in existing web search engines. We have created maps containing as many as 518 million of these hyperlinks, a significant sample of the total. These maps allow rapid calculation of a web page's "PageRank," an objective measure of its citation importance that corresponds well with people's subjective idea of importance. Because of this correspondence, PageRank is an excellent way to prioritize the results of web keyword searches.*

The concept of a *citation* is critical. Consider the example of an academic research paper, which might include citations similar to those shown in Figure 9-

13.

Franklin, K. Anabaptists: A Complete History [Book]. New York: Siegfried Shaw. 2002 Aug - [Cited 2004 May 12]
James, R. Time and Space and Me. [Book]. San Franciso: Abington Krane. 1996 Sep - [Cited 2010 Aug 17]
Landry, P. Automotive Engineering III. [Book]. Boston: Helmsworth Claw. 1984 Jan - [Cited 2005 Feb 1]

Figure 9-13. Academic citations

The paper's author uses the citation list to acknowledge major sources he referenced as he wrote the paper. If you did a study of all the papers on a given topic area, you could fairly easily identify the most important ones, because they would have the most citations (votes) by other papers.

Generally speaking, you do not buy placement of a citation in someone else's research paper. Nor do you barter such placements ("I will mention you in my paper if you mention me in yours"), and you certainly would not implement some tactic to inject mentions of your work in someone else's research paper without the writer's knowledge.

You would also not publish dozens or hundreds of poorly written papers just so you could include more mentions of your work in them. In principle, you can't vote for yourself. Nor would you upload your paper to dozens or hundreds of sites created as repositories for such papers if you knew no one would ever see it there, or if such repositories contained a lot of illegitimate papers that you would not want to be associate with.

Of course, all of these examples have happened on the Web with links. All of these practices run counter to the way that search engines want to use links, as

they are counting on the links they find being ones that were earned by merit.

This means that search engines don't want you to purchase links for the purpose of influencing their rankings. You can buy ads, of course—there is nothing wrong with that—but search engines would prefer those ad links have the `nofollow` attribute so they know not to count them.

Additionally, pure barter links are valued less or ignored altogether. From 2000 to 2005, it was quite popular to send people emails that offered to link to them if they linked to you, on the premise that this helped with search engine rankings. Of course, these types of links are not real citations either.

Google will not place any value on the links from user-generated content sites, such as social media sites, either. Anywhere people can link to themselves is a place that search engines will simply discount, or even potentially punish if they detect patterns of abusive behavior.

Here is a list of some of the types of links that Google may consider less valuable, or not valuable at all:

Article directories

> These are sites that allow you to upload an article to them, usually with little or no editorial review. The articles can contain links back to your site, so the simple act of uploading the article results in a link. The problem is that this is a form of voting for yourself, and Google started punishing sites that actively obtained links from article directories with the Penguin 1.0 release on April 24, 2012.

Cheap directories

> Many directories have sprung up all over the Web that exist only to collect fees from as many sites as possible. These types of directories have little or no editorial review, and the owner's only concern is to collect as many listing fees as possible. There are a few quality directories, such as the Yahoo! directory, Business.com, and a few others. These are discussed in more detail in Chapter 7.

NOTE

These comments on directories are not meant to apply to local business directories, whose dynamics are quite different. These are discussed more in Chapter 10.

Links from countries where you don't do business

If your company does business only in Brazil, there is no reason you should have large numbers of links from Poland and Russia. There is not much you can do if people choose to give you links you did not ask for, but there is certainly no reason for you to proactively engage in activities that would result in your getting links from such countries.

Links from foreign sites with a link in a different language

Some aggressive SEO professionals actively pursue getting links from nearly anywhere. As shown in Figure 9-14, there is no reason to have a "Refinance Your Home Today" link on a page where the rest of the text is in Chinese.

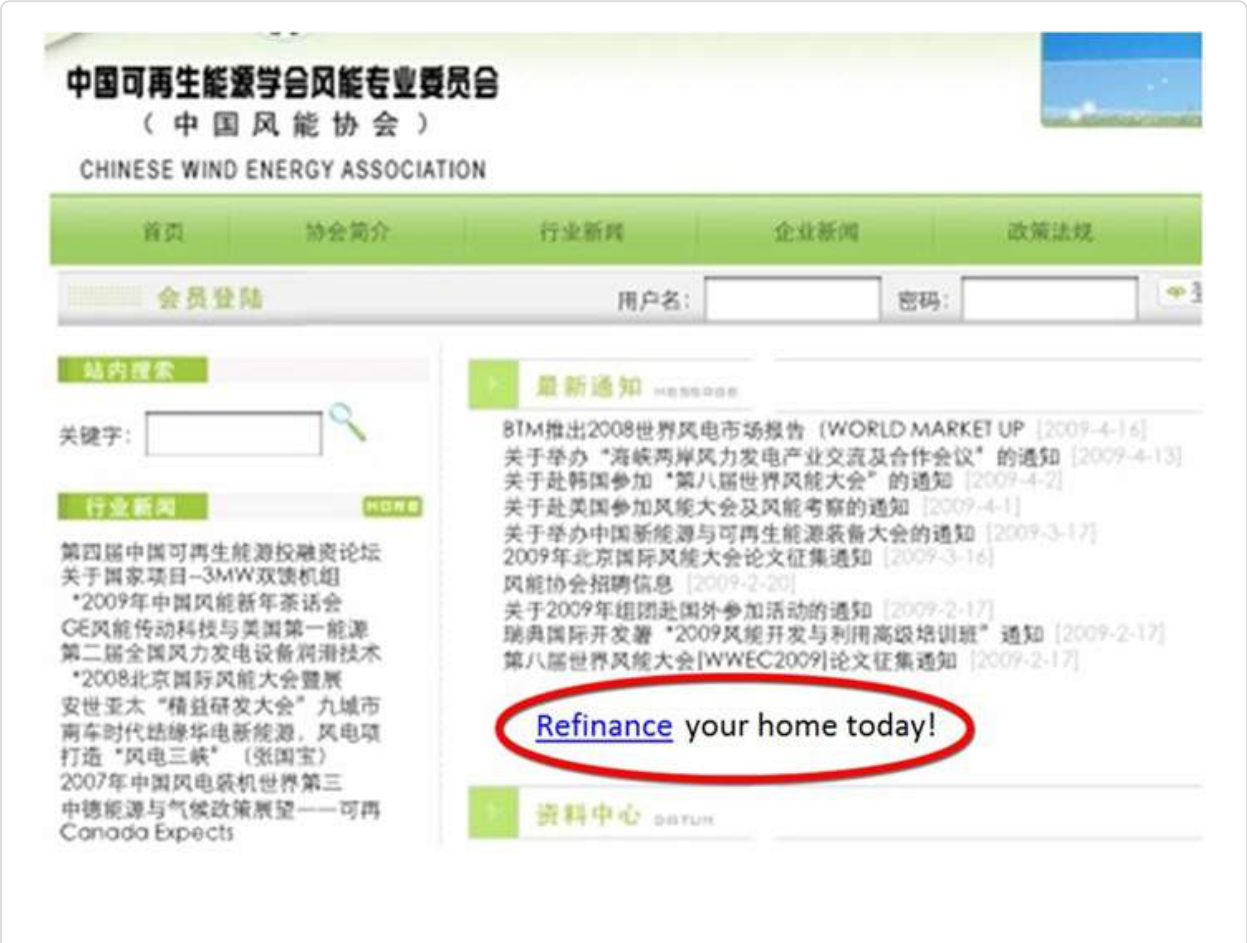


Figure 9-14. Foreign language mismatch

Comment spam

Another popular technique in the past was to drop links in comments on forums and blog posts. This practice became much less valuable ever since Google introduced the `nofollow` attribute, but aggressive spammers still pursue it. In fact, they make use of bots that drop comments on an automated basis on blog posts and forums all over the Web. They may post 1 million or more comments this way, and even if only .001% of 1 percent of those links are not `nofollow`ed, it still nets the spammers 1,000 links.

Guest post spam

These are generally poorly written guest posts that add little value for users and have been written just to get a link back to your own site. Consider the example in Figure 9-15, where the author was looking to get a link back to his site with the anchor text "buy cars." He could not even take the time to work that phrase into a single sentence!

St. Mary's

St. Mary's is a town just under an hour's drive from Bicheno, or an hour and a half's drive from Launceston. You can challenge yourself to a climb to the St. Patrick's Head, or the South Sister's Peak, or a relaxing drive in the Elephant Pass – all of which give spectacular views. In town, have a browse through all the craft galleries, bakeries and shops that this small town has to offer.

In terms of getting yourself a vehicle, scope out rental options, or if you're going for an extended holiday, you can consider looking at second hand cars to buy. Cars should be reliable, safe and road worthy of course, so don't buy something that's run down just because it's cheap.

Once you've got a set of wheels, it's time to hit the road!

Figure 9-15. Guest post spam

Guest posts not related to your site

This is a type of guest post spam where the article written does not really relate to your site. If you sell used cars, you should not expect Google to see any value in a guest post you write about lacrosse equipment that links back to your site. There is no relevance.

In-context guest post links

Another form of guest posting that Google frowns upon is posts that include links in the body of the article back to you, particularly if those links are keyword-rich, and if they don't add a lot of value to the post itself. Figure 9-16 shows a fictional example of what this might look like.

How to Get the Best Travel Deals

- Book early. People who wait until the last minute almost always pay higher prices.
- Use the Internet. Websites such as travel-o-matic can help you save hundreds of dollars on air travel and hotels.
- Try a vacation rental rather than a hotel.
- Look for travel package deals to help cut down the overall cost of both airfare and hotels combined.

This guest post was provided by Sarah Smith an avid traveler who also blogs for www.travel-o-matic.com.

Figure 9-16. Embedded keyword-rich anchor text links

Advertorials

This is a form of guest post that is written like it's an ad. Given the structure, it's highly likely that the site posting it was influenced to do so in some manner. If you are going to include guest posting as part of your strategy, focus on sites that don't permit these types of guest posts.

Widgets

One tactic that became quite popular is building useful or interesting tools (widgets), and allowing third-party websites to publish them on their own sites. These normally contained a link back to the widget creator's site. If the content is highly relevant, there is nothing wrong with this idea in principle, but the problem is that the tactic was abused by SEOs, resulting in Google wanting to discount many of these types of links.

Infographics

This is another area that could in theory be acceptable, but was greatly abused by SEOs. It is not clear what Google does with these links at this point, but you should create infographics only if they are highly relevant, highly valuable, and (of course) accurate.

Misleading anchor text

This is a more subtle issue. Imagine an example where the anchor text of a link says "information about golf courses," but the page receiving the link is about tennis rackets. This is not a good experience for users, and is not something that search engines will like.

Sites with malware

Of course, Google looks to discount these types of links. Sites containing malware are very bad for users, and hence any link from them is of no value, and potentially harmful.

Footer links

Once again, there is nothing inherently wrong with a link from the footer of someone's web page, but as these links are less likely to be clicked on or viewed by users, Google may discount their value. For more on this topic, you can read Bill Slawski's article Google's Reasonable Surfer: How the Value of a Link May Differ Based Upon Link and Document Features and User Data", which discusses Google's "reasonable surfer" patent.

Links in a list with unrelated links

This can be a sign of a purchased link. Imagine you find a link to your "Travel Australia" website mixed in a list of links with an online casino site, a mortgage lead generation site, and a lottery ticket site. This does not look good to Google.

Links from poor-quality sites

The links that have the most value are the ones that come from very high-quality sites that show substantial evidence of strong editorial control. Conversely, as quality drops, editorial control tends to as well, and Google may not count these links at all.

Press releases

It used to be quite popular to put out lots of press releases, complete with keyword-rich text links back to your site. Of course, this is a form of voting for yourself, and this is not the way that press releases should be used to promote your site. As shown in Figure 9-17, a much better way to use press releases is to get your news in front of media people and bloggers, and hope that it's interesting enough that they will write about you or share your news on social media.

Figure 9-17. The right way to use press releases

Bookmark sites

There are many quality sites for saving interesting links for your own benefit, such as Delicious, Evernote, Diigo, and many others. However, as these are user-generated content sites, their links are `nofollow`ed and thus have no value in ranking your site.

Not all of the types of links in the preceding list will necessarily result in your site bring penalized, but they are all examples of links that Google may not want to count.

Link Cleanup Process

The first part of the link cleanup process is to establish the right mindset. As you review your backlink profile, consider how Google looks at your links. Here are some rules of thumb to help you determine whether a link has real value:

- Would you want that link if Google and Bing did not exist?
- Would you proudly show it to a prospective customer right before she is ready to buy?
- Was the link given to you as a genuine endorsement?

As you review your backlinks, you may find yourself at times trying to justify a link's use. This is usually a good indicator that it's *not* a good link. High-quality links require no justification—it's obvious that they are good links.

Sources of Data

Google provides a list of external links in the Search Console account for your site. Figure 9-18 shows the steps to navigate to that report.

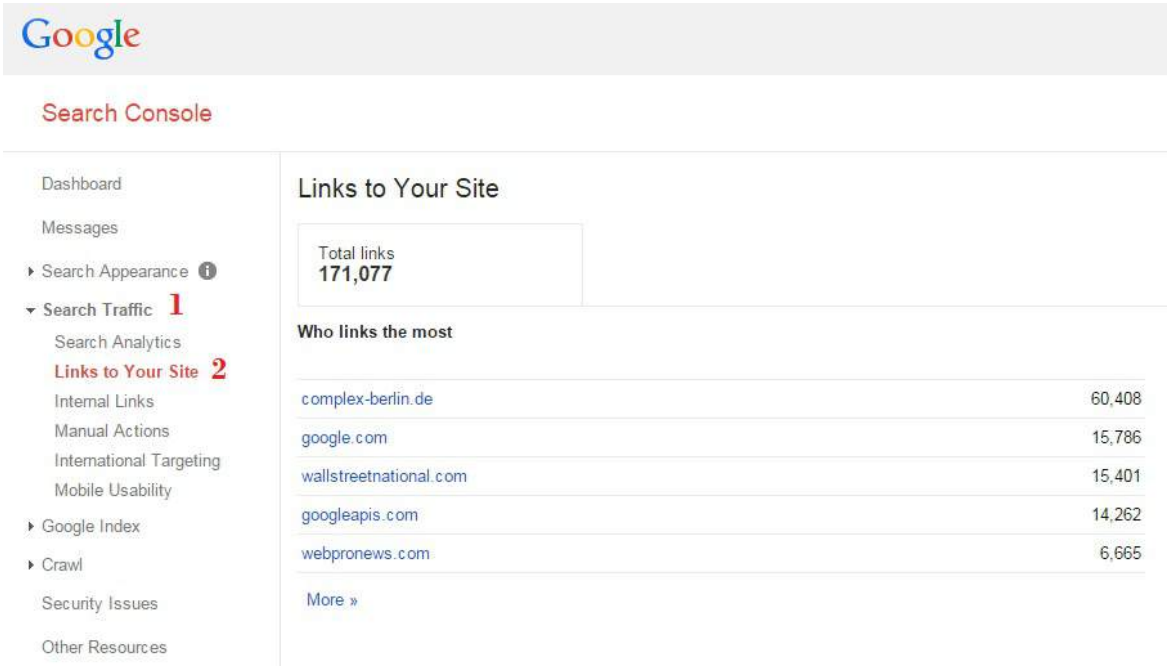


Figure 9-18. Steps to obtain the Search Console list of links

The problem with this list is that it tends to be incomplete, thus we recommend that you also pull links from several other sources. Some of the best additional sources include Open Site Explorer, Majestic SEO, Ahrefs, and LinkResearchTools.

The combination of data from all of these tools will show a more complete list of links. Of course, there will also be a lot of overlap in what they show, so make sure to deduplicate the list.

However, even the combination of all these sources is not comprehensive. Google shares only a portion of the links it is aware of in Search Console. The other link sources are reliant on the crawls of their individual companies, and crawling the entire Web is a big job that they simply do not have the resources for.

Using Tools

There are tools available to help speed up link removal by automating the process of identifying bad links. These include Remove'em and Link Detox. These tools can potentially help you identify some of your bad links. However, it is a good idea to not rely solely on these tools to do the job for you.

Each tool has its own algorithms for identifying problem links, and this can save you time in doing a full evaluation of all your links. However, keep in mind that Google has spent more than 15 years developing its algorithms for evaluating links and it's a core part of its business to evaluate them effectively, including detecting link spam. Third-party tools won't be as sophisticated as Google's algorithm. They can detect some of the bad links, but not necessarily all of the ones you will need to address. You should plan on evaluating all of the links—not only the sites labeled as toxic, but also any that are merely suspicious or even innocuous. Use your own judgment, and don't just rely on the tools to decide for you what is good or bad.

The Disavow Links tool

Google provides a tool to allow you to disavow links. The Disavow Links tool tells Google that you no longer wish to receive any PageRank (or other benefit) from certain links. This gives you a method for eliminating the negative impact of bad links pointing to your site. Figure 9-19 shows what the tool's opening screen looks like.

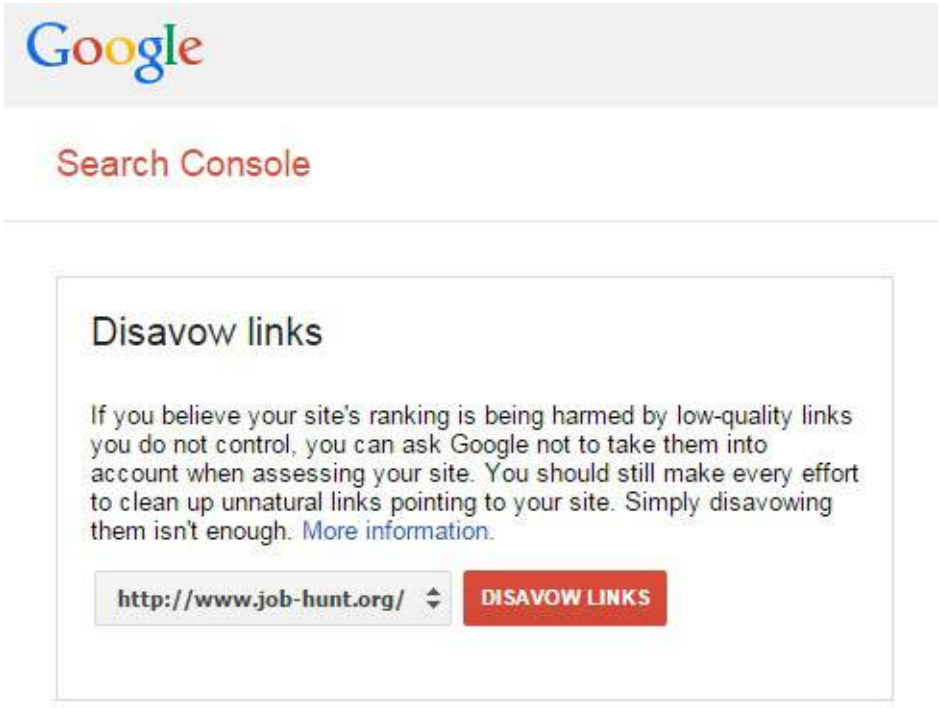


Figure 9-19. Opening screen of Google's Disavow Links tool

Note that Google includes the following text: "You should still make every effort to clean up unnatural links pointing to your site. Simply disavowing them isn't enough." This is good advice, as Google employees who review reconsideration requests like to see that you have invested time in getting the bad links to your site removed.

Once you select a site (blacked out in Figure 9-19) and click the Disavow Links button, you are taken to another screen that includes the following warning:

> *This is an advanced feature and should only be used with caution. If used incorrectly, this feature can potentially harm your site's performance in Google's search results. We recommend that you only disavow backlinks if you believe that there are a considerable number of spammy, artificial, or low-quality links pointing to your site, and if you are confident that the links are causing issues for you.*

On this screen, you need to click Disavow Links again, after which you'll be taken to a third and final screen, shown in Figure 9-20.

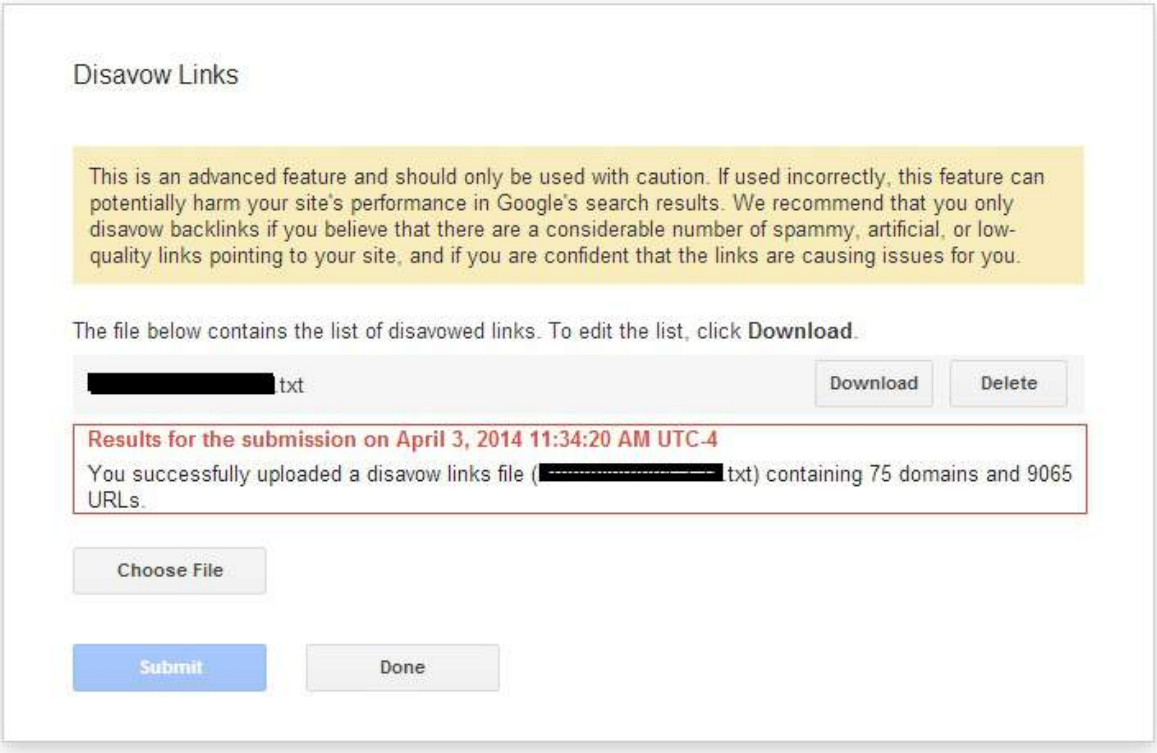


Figure 9-20. Final screen of Google's Disavow Links tool

The sample screen in Figure 9-20 shows that the current disavow file for this site (the name is blacked out) is disavowing 75 domains and 9,065 URLs.

One important tip: considering that the link data you have is incomplete (as described in "Sources of Data"), it is best practice to disavow entire domains. In other words, if you see one bad link coming to you from a domain, it is certainly possible that there are other bad links coming to you from that domain, and that some of these bad links are not in the data available to you.

An example would be a guest post that you want to disavow. Perhaps you have done only one guest post on that site, but the post will also appear in category pages and date-based archives on that blog. If you disavow only the post page, you still have many other bad links from that site that Google has found.

In the example shown in Figure 9-20, it is quite likely that this publisher has not solved his problem and that many (if not all) of the disavowed URLs should be made into full disavowed domains. That is usually the safest course of action.

Refer to the Google help page for more specifics on formatting the disavow file.

The Link Removal Process

The most important part of the link removal process is recognizing the need to be comprehensive. Losing a lot of your traffic is scary, and being impatient is natural. If there is a manual link penalty on your site you will be anxious to send in your reconsideration request, but as soon as you do, there's nothing you can do but wait.

If you don't do enough to remove bad links, Google will reject your reconsideration request, and you have to go through the whole process again. If you end up filing a few reconsideration requests without being successful, Google may send you a message telling you to pause for a while. Make a point of being very aggressive in removing and disavowing links, and don't try to save a lot of marginal ones. This almost always speeds up the process in the end. In addition, those somewhat questionable links that you're trying to save often are not helping you much anyway.

As we have discussed, the Penguin algorithm is updated only once or twice per year. If you have been hit by that algorithm and fail to recover during an update, that means waiting another 6 to 12 months before you have another chance.

NOTE

Google has indicated that it intends to speed up the Penguin update process in the future.

With all this in mind, you also want to be able to get through the process as quickly as possible. Figure 9-21 is a visual outline of the link removal process.

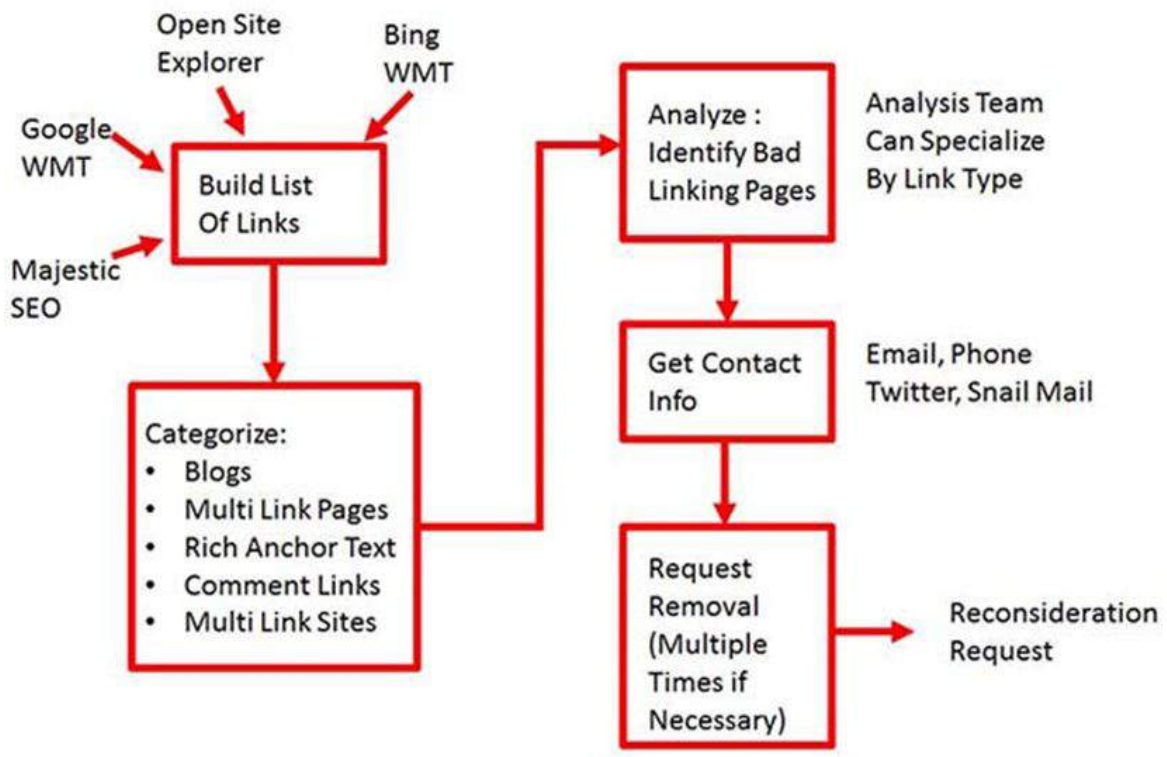


Figure 9-21. Process for removing links

Precategorizing the links is quite helpful in speeding up this process. For example, you can identify many of the blogs simply by using the Excel filter function and filtering on "blog." This will allow you to more rapidly review the links for problems. Tools such as Remove'em and Link Detox will do this for you as well.

This step is especially helpful if you know you have been running an aggressive guest posting campaign, or worse, paying for guest post placements. Some additional tips include:

- You do not need to worry about links that are marked as `nofollow`.
- Links from sites with very low PageRank for their home page probably are adding little to no value to your site.
- Links from very low-relevance pages are not likely to be adding much value either.

In addition, contacting sites directly and requesting that they remove links can be quite helpful. Google likes to see that you're putting in the effort to clean up those bad links.

Remember that reconsideration requests are reviewed by members of the webspam team at Google. This introduces a human element that you can't ignore. The members of this team make their living dealing with sites that have violated Google's guidelines, and you are one of them.

As we noted in "Sources of Data", even when you use all the available sources of link data, the information you have is incomplete. This means that it's likely that you will not have removed all the bad links when you file your reconsideration request, even if you are very aggressive in your review process, simply because you don't have all the data. Showing reviewers the good faith effort to remove some of the bad links is very helpful, and can impact their evaluation of the process.

However, there is no need to send link removal requests to everyone in sight. For example, don't send them to people where the link to you is marked with `nofollow`.

Once the process is complete, if you have received a manual penalty, you are ready to file a reconsideration request. If you have been hit by Penguin, all you can do is wait, as reconsideration requests won't help you there. The only way to recover from Penguin penalties is to wait for the next algorithm update.

Filing reconsideration requests

The first thing to realize about your reconsideration request is that a person will review it, and that person likely reviews large numbers of them every single day. Complaining about what has happened to your business, or getting aggressive with the reviewer, is not going to help your cause at all.

The best path is to be short and to the point:

1. Briefly define the nature of the problem. Include some statistics if possible.

2. Explain what went wrong. For example, if you were ignorant of the rules, just say so, and tell them that you now understand. Or, if you had a rogue SEO firm do bad work for you, say that.

3. Explain what you did to clean it up:
   - If you had a link penalty, let them know how many links you were able to get removed.
   - If you did something extraordinary, such as removing and/or disavowing all of your links from the past year, tell them that. Statement actions such as this can have a strong impact and improve your chances of success.
4. Clearly state that you intend to abide by the Webmaster Guidelines going forward.

As already noted, keep your reconsideration short. Briefly cover the main points and then submit it using the Search Console account associated with the site that received the penalty. In fact, you can't send it from an account without a manual penalty.

Expected reconsideration request timeline

Once you have filed the request, you now get to wait. The good news is that you generally get a response in two to three weeks. Hopefully, you will be successful! If not, you have to go back to the beginning of the process to figure out what you missed.

Conclusion

Traffic losses due to manual penalties or algorithmic updates can be devastating to your business. One defense against them is to reduce your overall dependence on Google traffic, but for nearly any business, traffic from Google is one of the largest opportunities available.

For that reason, it is critical that you conduct your SEO efforts in a way that is in line with Google's expectations of publishers—creating compelling websites and then promoting them effectively in accordance with Google's Webmaster Guidelines).

[1] Barry Schwartz, "Google Made 890 Improvements To Search Over The Past Year," Search Engine Land, August 19, 2014, *http://bit.ly/890_improvements*.

[2] Danny Sullivan, “Google Forecloses On Content Farms With ‘Panda’ Algorithm Update,” Search Engine Land, February 24, 2011, *http://bit.ly/panda_update*.

[3] Eric Enge, “The Rise And Fall Of Content Farms,” Search Engine Land, January 31, 2011, *http://bit.ly/content_farms*.

[4] Danny Goodwin, “Google’s Rating Guidelines Adds Page Quality to Human Reviews,” Search Engine Watch, September 6, 2012, *http://bit.ly/rating_guidelines*.

CHAPTER TEN

Mobile, Local, and Vertical SEO

Mobile, local, and vertical SEO have grown to become specializations in and of themselves, requiring a tremendous amount of dedicated effort, resources, and attention to fully maximize the opportunities they provide as the web ecosystem evolves. In this chapter, we will address these areas of search engine optimization at a high level.

The Mobile Landscape

If you have a web-based business, brand, or organization, mobile SEO will be absolutely vital to the success of your overall SEO strategy. But before we can examine the ins and outs of SEO for mobile, you need to first understand mobile search in the general context of mobile as a medium—one that is taking search market share and ad spend from desktop at a rapid pace.

As mentioned in Chapter 1, eMarketer published a study predicting that mobile ad spend would represent 72.2% of total U.S. digital ad spend by 2019 (Figure 10-1).

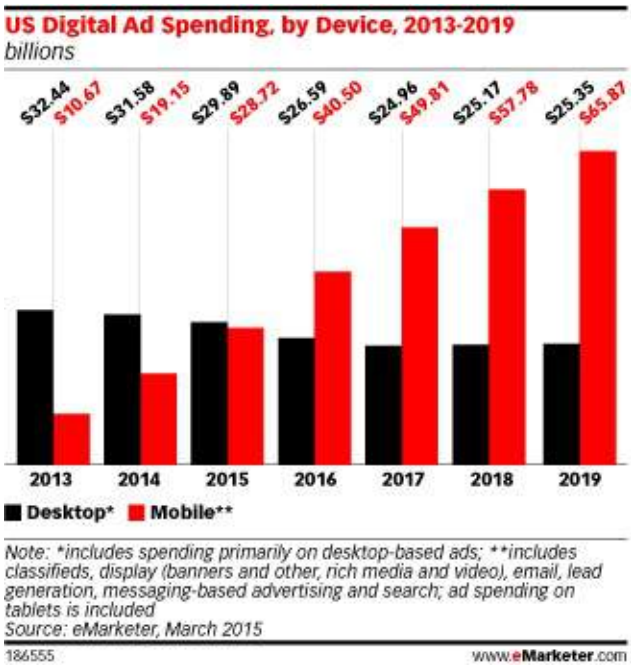


Figure 10-1. Mobile ad spend growth through 2019

The iPhone helped drive the growth of this market, because of its intuitive user interface and larger screen size (when compared to feature phones), and Android accelerated this trend with its increasing smartphone market dominance.[1].

The improved mobile web user interfaces and larger screen sizes on smartphones have been a big factor in the growth of mobile search. Having a specific mobile site experience for users is important, as mobile searchers are even more impatient and have a different intent than your typical desktop searcher. If you offer a desktop experience to a mobile user and she finds a competitor who offers a tailored mobile experience, you may have just lost a potential customer —and the search engines know this.

Rapid growth in mobile is expected to continue. On June 13, 2014, Matt Cutts told the audience at SMX Advanced in Seattle that before the end of 2014 Google would receive more search queries per day from mobile devices than it would from desktop devices. Google's mobilefriendly algorithm shift followed less than a year later. Today, a Google search from any mobile device shows us that the mobile landscape has evolved into an entirely unique, mobile-first search environment.

SEO for Mobile

The mobile tipping point in organic search has been reached. In April 2015, Google rolled out its mobilefriendly algorithm. Using a definitive, "yes or no" response ranking factor, this algorithm determines whether a site (or a specific page from a site) is mobile friendly, and then incorporates this information into its ranking decision for a mobile query. (Note that as of this writing, the mobilefriendly algorithm was not being applied to Google News content for mobile queries returning News results, or for Google News–specific searches).

Google's message with this algorithm update, which it warned would have more of an impact on its search results than either Panda or Penguin (learn more about Panda and Penguin in Chapter 9), was clear: if you want your site's content to show for mobile queries, you need to build your site for the mobile user experience. At the most basic level, this means that your mobile site must be:

Fast

Pages should ideally load in less than 1 second.

Functional

Page content should display and function properly in mobile browsers (no CSS, JavaScript, image, or other resource blocking).

Finger-friendly

Tap targets (e.g., buttons, links, form fields) should be large enough and properly spaced for small touchscreen use.

Free from redirects and errors

Mobile version URL requests should all return 200 (OK) status codes—no 301/302 redirects or errors, if possible.

There are various ways to implement a mobile version of your site that delivers user-friendly functionality to a mobile browser, with three being accepted by Google. Note that these fall into two general categories for SEO purposes—same URLs, and separate URLs:

Responsive web design, or RWD (same URLs)

Your site utilizes the same URLs for desktop and mobile. Your site's server always sends the same HTML to all devices, and cascading style sheets (CSS) are used to alter the rendering of the page on the device. Google can generally detect RWD as long as the site's assets (including CSS) are crawlable.

Dynamic serving, a.k.a. adaptive design (same URLs)

Your site utilizes the same URLs for desktop and mobile, but your site's server responds with different HTML and CSS depending on the user agent requesting the page.

Separate URLs

Your site utilizes different URLs for the desktop and mobile (and perhaps even tablet) versions of your site's content (utilizing a subdomain such as *m.site.com*, for example), with device-optimized content served at each URL version.

Same URL approaches: responsive web design and dynamic serving

As you can see, the first two options (responsive web design and dynamic serving) retain the same URLs across devices, while the third option (separate URLs) creates two entirely separate websites for the search engines to crawl, index, and rank. This is a major factor to consider before developing your mobile site.

A significant advantage offered by using the same URLs is that the mobile site may more effectively inherit the SEO characteristics of the desktop site. In other words, the link profile and social signals are automatically associated with the mobile version of the site (which is not the case with a mobile subdomain). These “same URL” approaches are considered by some to be the best option if you intend to support smartphone-type devices only, and do not plan to support feature phones (e.g., traditional flip phones).

The most cost-effective of these two approaches is usually RWD, as only one site “version” is being created and maintained. For this reason, RWD has seen rapid adoption in the mobile space.

In contrast to RWD, dynamic serving—whereby entirely different site versions are designed and coded with device-specific optimization and user experience in mind—renders different HTML for the same URL request, based on the user agent making the request. While this approach is more budget, development, and maintenance intensive, it often can allow for faster load times, and a more optimized user experience specifically tailored to device type.

Separate URL approach

You may wish to offer a mobile-specific URL for marketing and promotion purposes, or to allow access to the mobile version of your site regardless of the user agent making the request. In particular, if you plan to support both smartphones and feature phones, having a mobile subdomain allows you some additional flexibility in dealing with different screen sizes, or if your site is so large that you can’t provide a separate mobile experience for each page.

Some well-known examples of such URLs include *http://m.facebook.com*, *http://mobile.weather.gov*, *http://www.hotels.com/mobile/*, and *http://en.m.wikipedia.org*. The benefit of serving up a distinct mobile version

instead of serving your standard version with mobilefriendly CSS is that you can trim the file size of the HTML sent to the mobile device, eliminating code that the device won't be able to run, which improves page load time and makes the page look better to mobile bots. Google provides more information about separate URLs for mobile here: *http://bit.ly/separate_urls*.

Mobile user agents

Google employs different spiders for the mobile Web: Googlebot-Mobile for feature phones, and the standard Googlebot with an iPhone user agent for smartphones. Bing, on the other hand, uses the same crawler for both mobile content and desktop content, called Bingbot), and utilizes crawler variants in the following format:

```
Mozilla/5.0 + (Mobile Device) + Mobile Engine + Mobile Browser +
  bingbot/BingPreview/[version]
```

Both search engines use *user-agent detection* to determine if a searcher is on a mobile browser or a traditional desktop browser, and will show feature phone users a mobile experience. At the beginning of each browsing session, the user's hardware and browser combination will communicate a unique identifier to the website's web server, known as the *user agent*. This string identifies the nature of the hardware and software making the request.

The best practice for mobile site user experience, and hence for mobile SEO, is to detect the incoming user agent and show the user the version of your site specific to the user's device. If you are using the same URLs for both your desktop and mobile sites (as in responsive web design and dynamic serving), you simply show users the alternative version of your content. If you are using separate URLs, you will want to redirect mobile users to that mobile URL.

A useful tool for server-side user agent detection is MobileDetect.net, which provides an open source PHP class (`Mobile_Detect`) for user agent detection and is supported by third-party plugins for WordPress, Drupal, Joomla, Magento, Concrete5, and other content management systems, as listed in "3rd party modules".

General SEO guidelines for mobile

Some general tips for your SEO mobile strategy include the following:

- Review search engine recommendations for mobile SEO (and check for updates regularly!):

  — Google's Mobile SEO page from its Webmaster's Mobile Guide

  — Bing's "Meet Our Mobile Bots" page

  > **NOTE**
  >
  > — Bing generally does not provide as much webmaster/SEO support as Google.

- Ensure critical site files (CSS, JavaScript, images) are accessible to the search engines; many of these files signal to the search engines that the page is built to display and work well on a mobile browser (i.e., telling the search engine that the page is "mobilefriendly"). Check to make sure your site does not block access to any of these files via *robots.txt*.

- If your site is dynamically serving desktop and mobile content based on the user agent requesting the page, use the `Vary` HTTP header to signal your changes depending on the user agent (*http://bit.ly/dynamic_serving*).

- If your site is hosting desktop and mobile content on separate (unique) URLs, annotate the desktop page's HTML by adding the `rel="alternate"` tag pointing to the mobile URL, and annotate the mobile page's HTML by adding the `rel="canonical"` tag pointing to the desktop URL (*http://bit.ly/separate_urls*). This can also eliminate any potential duplicate content concerns between the two site versions.

- Publish only mobilefriendly content (Flash movies or Flash-based websites, for example, are not compatible with mobile!).

- Include a mobile sitemap according to updated search engine guidelines. As of this writing the most recent mobile sitemap format utilized by Google requires you to include a specific tag (`<mobile:mobile/>`) for mobile URLs to be crawled. For more information, refer to Google's documentation on

mobile sitemaps.

- Always test your pages in Google's MobileFriendly Test Tool. This will enable you to see if Google will treat your pages as mobile friendly in their search results.

Mobile site speed

Site speed is a crucial component of mobile SEO success, a point that has been driven home by Google for a number of years now. You can find suggestions for your specific site by using the following tools (the recommendations hold true regardless of whether you choose a same URLs or separate URLs approach):

- Google Page Speed Insights
- Mobi Test from Akamai
- WebPageTest
- W3C's Mobile OK Checker

Additional mobile SEO resources

Other resources you might find helpful for mobile SEO include:

- 50+ Essential Mobile SEO Resources
- "Mobile Site Audit Checklist"
- "SEO: Mobile Search"
- "Mobile SEO"

NOTE

A special thanks to Michael Martin for his contributions to portions of this section.

App SEO: Deep Linking and Indexing for Mobile Search

Based on the evolution of deep linking and indexation for mobile app content, significant opportunities have emerged for SEO exposure of mobile app content in the near future. App SEO for organic search exposure (different from "App Store SEO," which has generally referred to optimizing mobile apps for prominence within Apple's App Store and Google Play) provides a massive opportunity for content discovery via the indexation and ranking of deep content residing within mobile apps. As we have learned over the years with website content delivered via search, as deeper, more targeted content is delivered to a searcher, content (in this instance, app) engagement increases, time to conversion decreases, and search engine results improve.

Various technology companies are targeting the app-specific search landscape. URX has built a deeplink app search API, whereby developers can monetize apps or websites with deep links to contextually relevant in-app content and actions in other apps. In other words, URX has built a search engine that developers use to find a deep link into an app for a specific context. Quixey has built a search engine that relies upon deep linking to help users find the most relevant content within apps in the same way that search engines help users find the most relevant content within websites. Google is also actively including App content in its search results. Coauthor Eric Enge has interviewed Google's Mariya Moeva on this topic.

App Deep Linking

The indexation of deeplinked content within apps (in-app pages) and the delivery of this content to mobile searchers have opened a world of possibility for apps and the mobile users who use them. In its most basic definition, *app deep linking* enables a link to send a user (or a search engine crawler) to deep, targeted content within an app (such as a specific page, tab, or view) based on his intent, as opposed to sending him to the app's home page—something we take for granted on desktop when we share deeplinked content from various websites or when we click on a link in search results. In essence, deep links take the user directly to what he wants to do within the app, increasing app engagement, decreasing the time to conversion for targeted in-app behaviors,

and improving the relevance of results in mobile search. All things considered, the mobile search ecosystem is poised to support this trend of driving intent-based traffic to mobile apps.

Deep linking to in-app pages and functionality also allows users to share deep content across various platforms (email, social sharing, SMS), further increasing engagement levels. And from an SEO perspective, it is likely that we will eventually see external linking, social, and other engagement signals applied to deeplinked app URLs much in the same way that we see them applied to website URLs in SEO today.

Deep linking URL schemes

In order to provide deep links to your in-app content, you will need to create a unique URL scheme with routing parameters (path, query string, etc.) representing the custom actions to take within the app, register these with the operating system, and map the routes to content destinations or user actions within the app. An example of a deep linking scheme with a routing parameter is *twitter://timeline*. When entered into a mobile browser, this opens the Twitter app and links directly to the device user's timeline.

URX and MobileDeepLinking.org provide helpful tutorials on deep linking for both the novice and the app developer.

Deep linking standards

Deep linking standards attempt to define for app developers how to create in-app URL schemes so that the various platforms that publish these URLs can understand them. There are various cross-platform standards, as outlined by these providers:

- Facebook's App Links
- Twitter's App Card
- Google's App Indexing
- Quixey's AppURL
- MobileDeepLinking.org

Deep linking tracking and optimization opportunities

Increasingly, tools are being developed to provide deep linking capabilities as well as acquisition data "outside the app," filling a historical gap within traditional mobile app analytics solutions and enabling app marketers to utilize user intent metrics—which can provide greater insight into organic attribution and drive intent-based content delivery.

A sampling of these tools includes:

AppLinks

> Open, cross-platform solution for basic app-to-app linking (used by Facebook, Pinterest, Spotify, Venmo, and others)

Tapstream

> Marketing dashboard providing tools for deferred deep linking, onboarding links, and A/B testing

Yozio

> Platform offering solutions for deferred deep linking, dynamic linking, and A/B testing

A quick note on tools that provide third-party URLs for deep links to your app content, as opposed to using your own app link URLs: there may come a time in the future when it is smarter to develop your own deep link URLs for the same reason that you develop your own URLs for your site content—so that social shares and external links can be built over time and credited to your owned URL.

App Indexing

Google currently offers its App Indexing API to allow developers to notify Google about deep links in their native apps, which Google uses to index content that it then serves in mobile search results. Google also states that the App Indexing API allows the Google App to drive reengagement through Google Search query autocompletions.

Bing currently indexes apps for search on Windows and Windows Phone via the

Windows Phone Store URL.

Optimizing for Vertical Search

Vertical search refers to the search accessibility of various niches and formats of web content (referred to as *verticals*), including images, videos, events, news, travel, products, music, and people (this has also been referred to as "universal" search in Google). Google provides vertical search functionality, as do vertical-specific websites such as Expedia.com and Cars.com. The term "vertical search" is often used interchangeably with "blended search," although this is not quite correct. Blended search results include various types of vertical content, such as videos, images, and so on. SEO for vertical search refers specifically to efforts to optimize vertical content types so they appear prominently within vertical-specific search engines (Google Images search, for example) and within blended search results. At this stage, Google and Bing provide blended search results (in addition to their separate vertical search functions) to improve the user experience and increase the overall search value to users in ways that go beyond what traditional web search results have provided.

Generally speaking, blended search results can provide more relevant (and therefore more valuable) content within specific verticals. Vertical-specific search engines still exist, as Google and Bing provide numerous vertical-specific search functions; however, we are also seeing an increasing trend toward the search engines incorporating vertical results within their main web search results.

A June 2014 study by Searchmetrics found that the proportion of keyword searches in Google with Universal Search integrations (i.e., blended results) was 81% in 2013, up from 75% in 2012, reflecting the ongoing opportunity for increased exposure in blended search results and vertical search. SEO professionals should seriously consider what potential benefits vertical search optimization can provide to their, and their clients', websites. There are very specific ways in which vertical content can be optimized for success in these areas, as we will outline in this chapter.

Vertical search has been around for almost as long as the major search engines have existed. Some of the first vertical search engines were for image search,

newsgroup search, and news search, but many other vertical search properties have emerged since then, both from the major search engines and from third parties. This chapter will focus primarily on strategies for optimizing your website for the vertical search offerings from Google and Bing.

Universal Search = Blended Search

Google's web results search engine used to be a kind of vertical search engine itself: one focused specifically on web pages (and not images, videos, news, blogs, etc.). In May 2007, Google announced *Universal Search*—merging results from different types of digital media, such as images and videos, within its search results. Other search engines followed suit with similar search functionality, which is now referred to as "blended search" as a search engine–neutral term within the industry.

With the advent of Universal Search, Google changed the "web page" search engine into a search engine for any type of online content. Figure 10-2 shows Moz's "Mega-SERP" infographic, which reveals the wide array of search result types that can be pulled into a regular web search in Google, as it now incorporates various algorithmic and data components of its Knowledge Graph (or Carousel) into its main SERPs. Note that Figure 10-2 shows not an actual search results page, but rather the universe of vertical content types that can be pulled into general web search results.

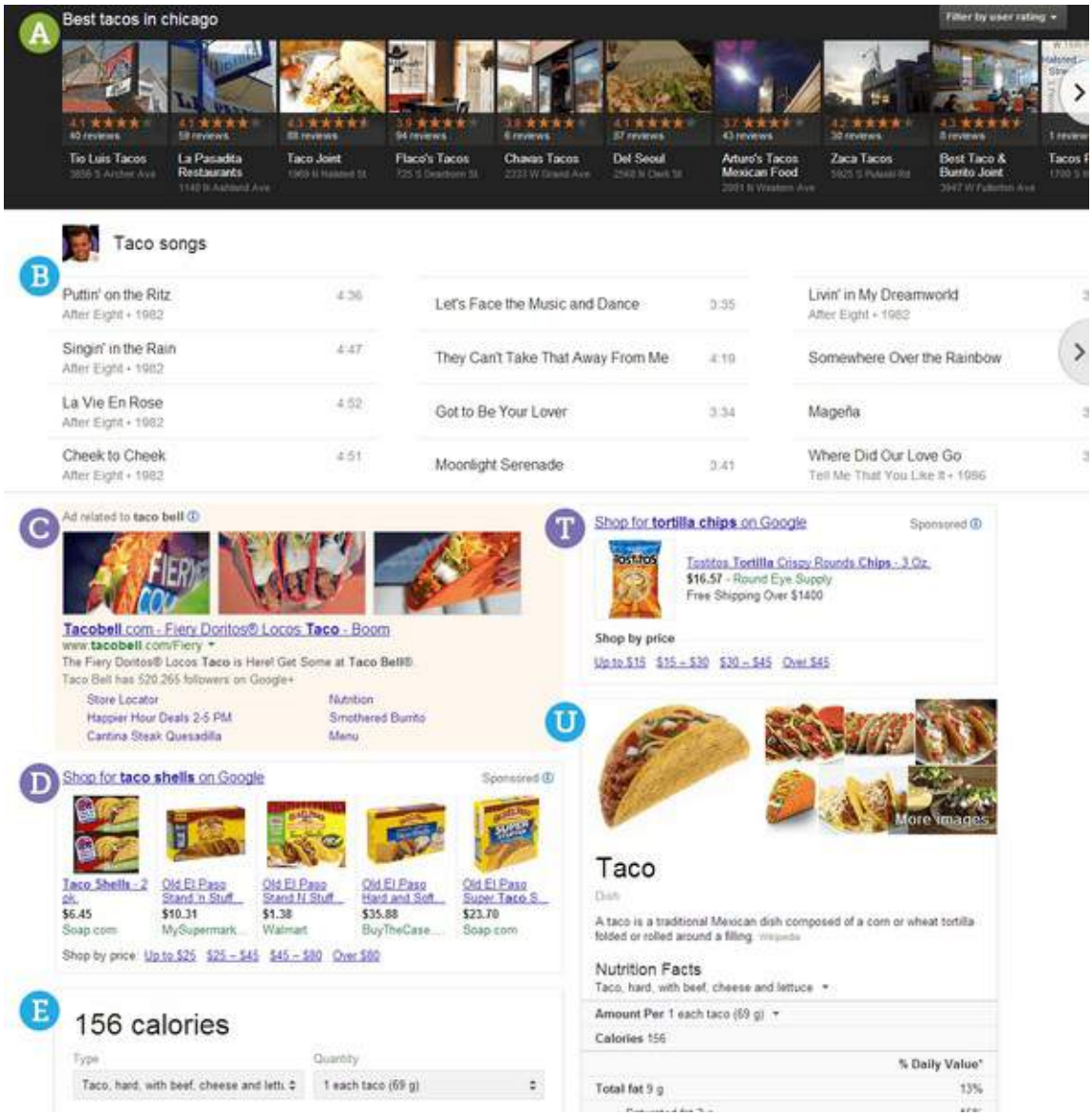


Figure 10-2. "Mega-SERP" infographic (Moz.com)

Figure 10-3 shows an example of how blended search plays out in a Google SERP for a search for *iphone*. Notice the news results (labeled "News for iphone"), the image results (labeled "Images for iphone"), and the Shopping results (labeled "Shop for iphone"—these are paid search results from advertisers). This SERP is an example of vertical search results being incorporated right into traditional web search results.

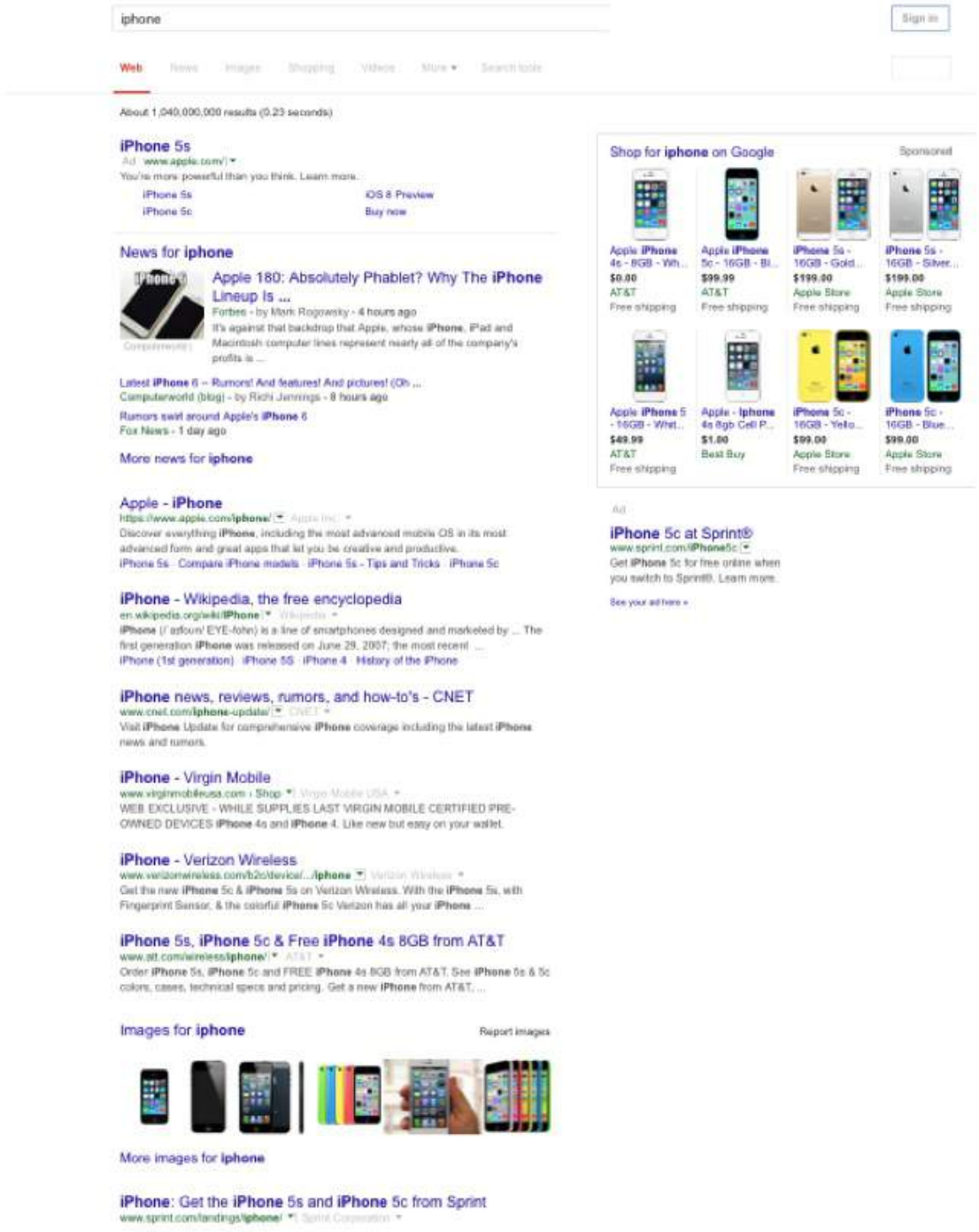


Figure 10-3. Blended search results for "iphone"

Figure 10-4 shows the results for a search on *Lebroning*.

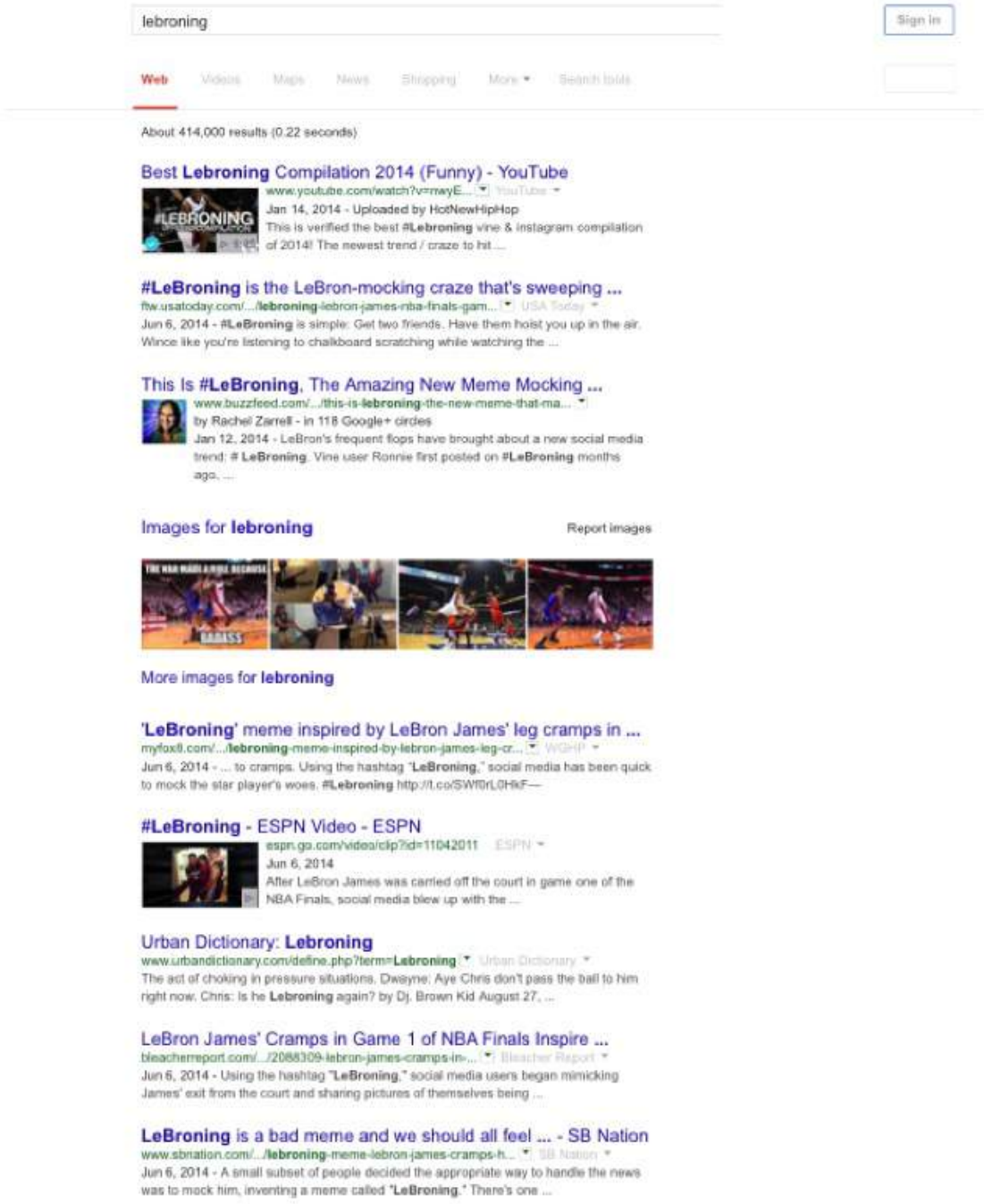


Figure 10-4. Blended search results for "Lebroning"

This shows images and videos in the search results. This enables users to click

on them and see for themselves exactly what “leboning” is.

The Opportunity Unleashed

As we noted at the beginning of this chapter, the opportunity in vertical search was significant *before* the advent of blended search. However, that opportunity was not fully realized because many (in fact, most) users were not even aware of the vertical search properties. With the expansion of blended search, however, the opportunities in optimization for vertical search have soared. By leveraging digital media assets such as images, videos, articles, events, recipes, TV and product reviews, and local business listings, businesses can increase their chances of having search result exposure that may directly or indirectly drive traffic to their sites.

Optimizing for Local Search

In 2014, Google published a research study called “Understanding Consumers’ Local Search Behavior”, which revealed numerous valuable insights into the state of local search. Figure 10-5 shows the executive summary from this study.

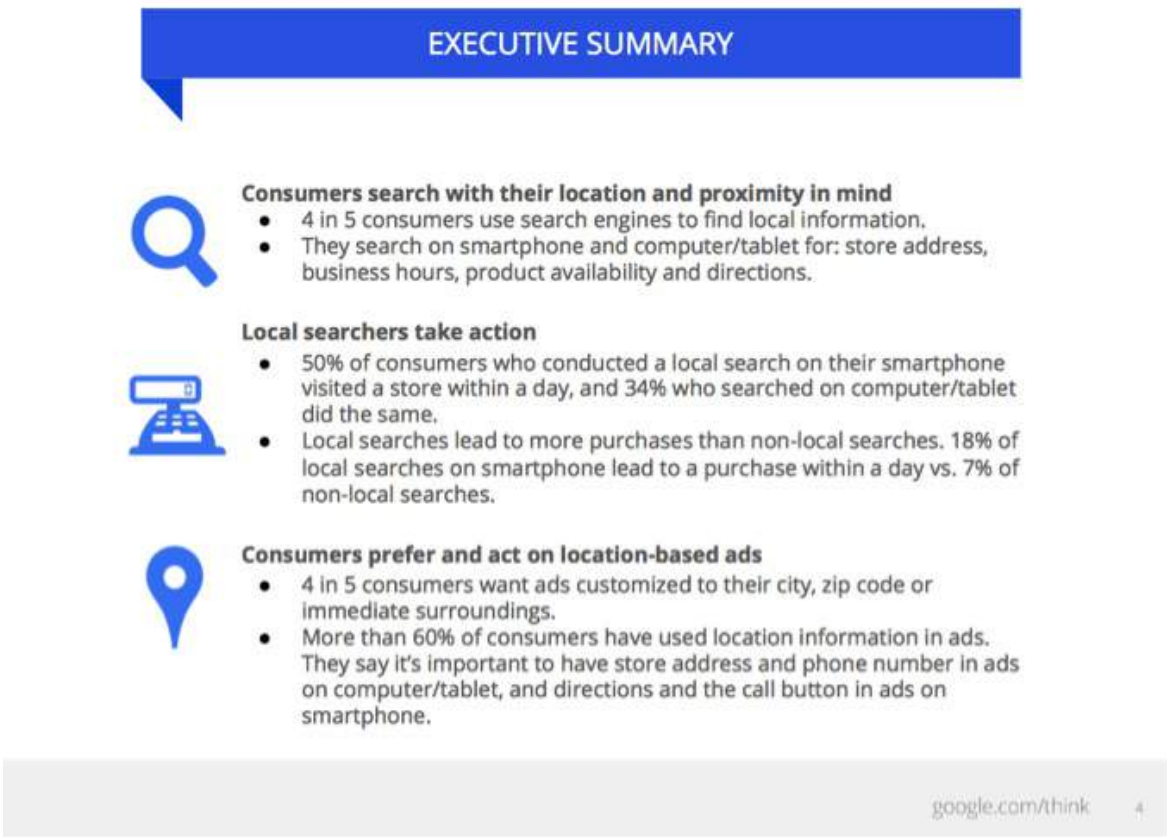


Figure 10-5. Google local search study: executive summary

Google's study also found that 56% of "on the go" searches have local intent, and 50% of smartphone users and 34% of desktop/tablet users who perform a local search visit a store within a day of their search. These valuable insights underscore the importance of developing a local search strategy.

The major search engines have sought to increase their advertiser bases by moving aggressively into providing local directory information. Applications such as Google's Google My Business and Bing Places for Business have introduced disruptive technology to local directory information by mashing up maps with directory listings, reviews and ratings, local citations, satellite images, and 3D modeling—all tied together with geotargeting and keyword search relevance. These innovations have excited users, and the mapping interfaces are growing in popularity as a result.

While there is no single dominant provider of local business information on the Internet, Google, Bing, and Yahoo! have ramped up their efforts to tackle this

market by providing local listings within their search results—and thus have become the go-to resources for local information. “Local SEO” has become a very complex and unique subset within the SEO practice, and has even developed its own unique vernacular, for which Moz has provided a helpful glossary. If you are hoping to gain traction in local SEO, it would be wise to familiarize yourself (and/or those working on your local SEO efforts) with the terminology in this glossary.

The opportunity for local search is huge. More than any other form of vertical search, local search results have a major presence in web search. For example, Figure 10-6 shows the results for a search on *minneapolis rental cars*.

Only three regular web search results are above the fold, and the following results are the “pack” of Google local search results (packs comprise three, five, or seven results).

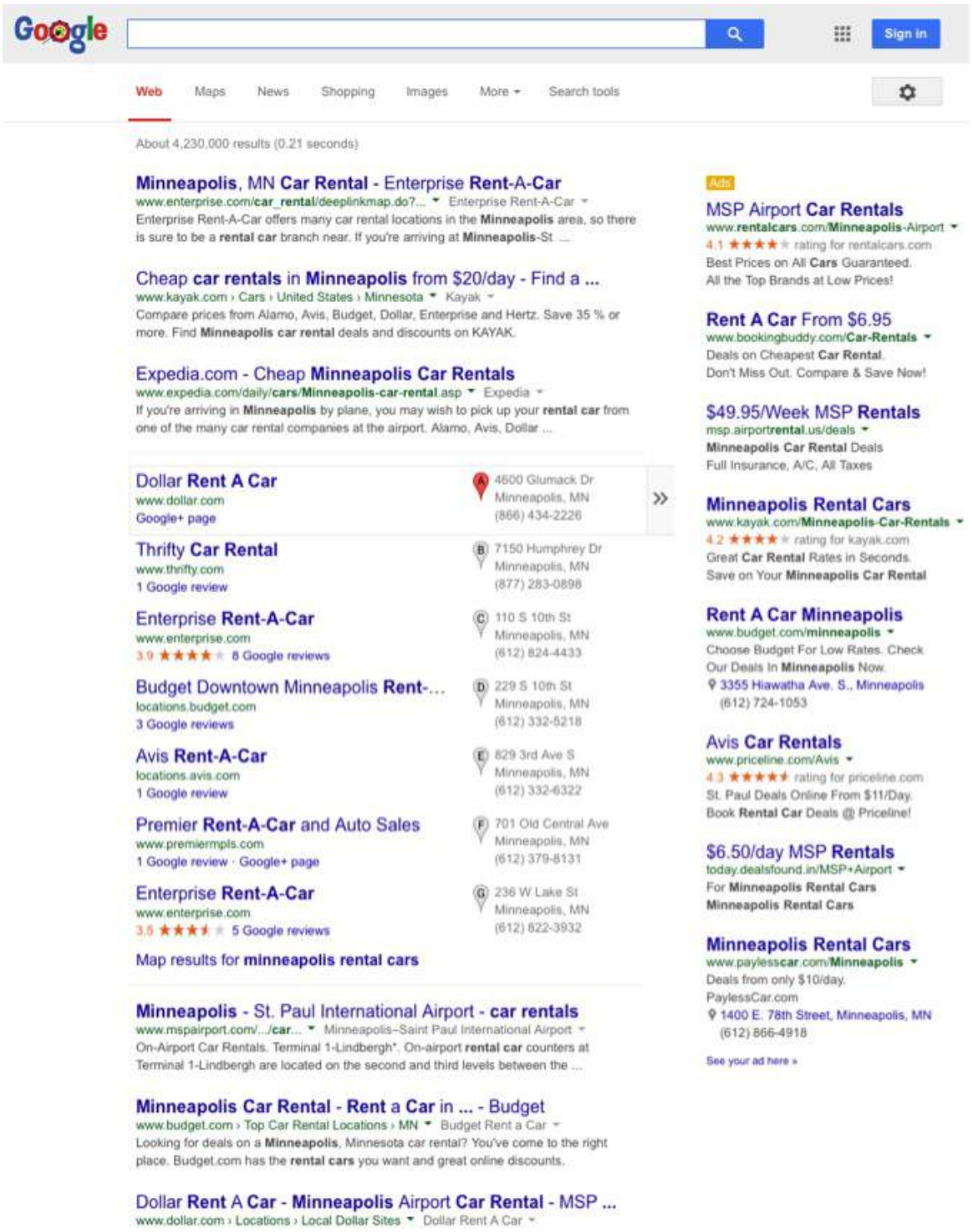


Figure 10-6. Local search results example for "minneapolis rental cars"

Local Listing Submissions

Enhancing your local listing profiles in Google, Bing, and Yahoo! is a matter of providing as much relevant data to the search engines as you can. The most important elements to focus on are your business name, address, phone number, website, and categories. These are the elements that you can influence that carry the most weight in the ranking algorithms. While adding a rich description can help with click-through rates and conversions, it has very little, if any, weight in the ranking algorithms.

Here are some guidelines for making choices for these data elements:

Business name

Make sure you use an accurate business name. Do not try to add keywords or locations—this is considered spam and may result in your listing being rejected. Be sure to use the exact same business name that you use on your website, and in any other online business profiles.

Address

Like the business name, your address is part of the "key" that the search engines use to identify your business. It is helpful to ensure that the address "maps accurately" in the search engines. Do this by performing a search in each engine on just your address. See if the engines show an exact match for your business address; if they do not, you may need to modify it (for example, there may be multiple ways a particular street is referred to, such as "Mc Andrew" versus "McAndrew," and you may have to try an alternate version), or you may need to report a mapping error to the search engine directly. To do so, click on the "Report a Problem" link, which can be found in several places on the map results—usually at the bottom of the map, or underneath the business address within the map display.

It is also useful to ensure that the USPS considers your address to be correct. You can check this with free tools available at *https://tools.usps.com/go/ZipLookupAction!input.action*.

Phone number

For the "main" phone number, be sure to use a unique local phone number. Avoid 800 numbers, and avoid tracking numbers if you do not plan to use

them on a long-term basis (i.e., for testing or other purposes). As with your business name and address, your main phone number is part of the “key” the engines uses to cross-reference information about your business around the Web. Using a nonlocal or frequently changing phone number can potentially confuse the search engines and cause them to have less confidence in your listing. In addition, it can lead to the creation of duplicate business profiles, which would compete against your core listing for search engine rankings.

Website

In many cases, for a business with a single location, your website is simply your home page. For businesses with multiple locations, it is ideal to have a separate page representing each location.

Be sure to provide the address and local phone number (as just described) on the page you link to. It can be helpful to encode the contact information on the web page using markup from Schema.org (*http://schema.org/LocalBusiness* or its industry-specific variants) or microformats (*http://microformats.org/wiki/hcard*). You can also use the Google Search Console Data Highlighter to “tag” the organization data on your pages in lieu of markup if you are not able to modify your web pages easily, but note that this will only serve to inform Google, not Bing or Yahoo!.

Having a website can be a strong ranking factor for the search engines, and it can be a strong factor in converting searchers to customers. It is advised that you provide at least a basic website associated with your business listing. Of course, if you do so, you will also want to follow much of the other advice in this book to optimize that website.

Description

Google now provides you with a description field that allows 4,000+ characters, which should be plenty of space to describe your business and promote your message. Make sure you provide a well-written, unique description of your business. While Google states that it no longer relies upon the business description for ranking purposes, it is still important to

accurately describe your service, product, and location specifics for users.

Categories

Category selection provides a very strong clue for the search engine to determine what kinds of searches your business is relevant to. The best advice here is to simply study the available categories, and be sure that the first one you choose is the “best” category. At least one category selection must be made from the search engines’ standard set of categories. Google and Bing allow for up to 10 categories, and Yahoo! allows for 5.

If you use your own categories, be sure to follow the search engine quality guidelines. Mainly, be sure to list categories that describe what the business is (as opposed to, for example, what products it sells or where it is located). Also, keep each category focused on a unique quality of the business. Do not attempt to jam keywords into your category choices, as the engines may consider this to be spam and reject your listing.

One good source for brainstorming categories is to examine the business listing pages of some competitors and see what they’ve chosen. It is helpful to be seen in the same categories as the majority of your competitors. This ensures that the engines consider your business as “the same type” as your main competitors, allowing you to be considered for display in the most important search results.

In addition to these key elements, your business profile can include information on your business hours and accepted payment methods, photos, videos, and other data items. Providing this extra information is worth the effort, as it enriches your business page and may help entice more users to contact you. It also opens up additional opportunities to provide keywords and context to help the search engines better understand how to rank your business.

NOTE

Your business’s name, address, and phone number (NAP) are perhaps the most critical components of your local listing data that contribute to local ranking, so it’s critical to keep this information consistent throughout your web presence and across your local listings. Most businesses go through normal transitions that require NAP information changes (business

location changes, contact phone number changes, etc.). Be sure to update your NAP information on your website, your local listings, and your third-party platform pages (YouTube, Facebook, etc.) after any of these transitions, and periodically monitor your NAP information for consistency.

Google My Business

In June 2014, Google announced the launch of Google My Business (GMB) in an effort to streamline and simplify the management of local business profiles—which had evolved into an increasingly complex and time-consuming array of business listing options and locations under the umbrella of Google properties. According to Google, businesses that have existing Google Places listings and Google+ Local Pages will automatically be upgraded to the GMB platform. At the launch of this service, Google added the ability to "View the Business on Google," which gives you a preview of your business info in Google Search, Maps, and Google My Business. The opening page for Google My Business is shown in Figure 10-7.

More information about Google My Business can be found on its FAQ page.

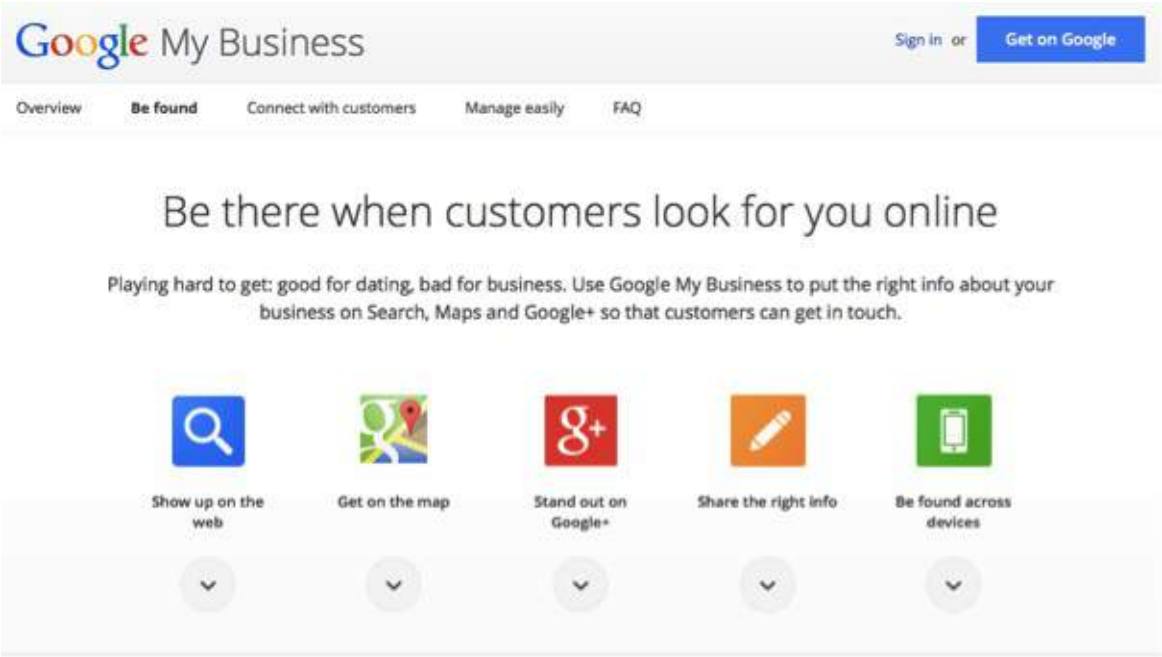


Figure 10-7. Google My Business splash page

Optimizing your GMB listing(s)

It is worth investing the time to learn how to optimize for GMB. One thing that may surprise you: if your business has been around for any length of time, Google probably already knows about it and has a listing in place. Google is very good at gathering and aggregating data from multiple sources to create a comprehensive representation of your business.

As of this writing, you will likely find this listing in a local SERP within the local pack with a link to a GMB page for your business. If Google has generated such a default listing for you, you should claim this GMB listing (rather than creating a new one). Do so by clicking on the "Manage This Page" link that appears on your business's GMB page (underneath "Is This Your Business?") and following the prompts, which will include creating a GMB account.

Claiming the listing will allow you to take control of the profile information for your business (the business name, address, and phone numbers(s), as well as a vast set of other data elements, including the website address, the business categories you should be associated with, photos, etc.). Claiming is relatively straightforward. Before granting you control, Google will have you prove you own the business being claimed. You will receive a secret code either through a postcard that's mailed to your business address or in a live phone call from a Google robot. You then enter this code into the portal, and Google grants you exclusive control of the listing. You may also be able to instantly verify your GMB listing through Google Search Console; for this option, the website listed on the GMB profile must be verified in Google Search Console with the same email address you are using to claim the GMB profile. You can also batch-upload listings if needed; see *http://bit.ly/bulk_mgmt*.

More information on creating, claiming, and verifying your Google My Business listing can be found here: *http://bit.ly/verify_local_business*.

As mentioned, it is almost always better to claim an existing listing, if there is one, than to create a new one. Local search engines struggle with duplication of listings, so it is best if you do your part to ensure that there is only a single listing representing your business. In fact, an advanced technique that you may want to explore is to search Google Maps using your business name, address, and phone number to see if there are any duplicates or near-duplicates of your business information, and if so, to claim those as well. Doing so will help

Google consolidate this information, and you will benefit from having several “weaker” listings consolidated into a single strong listing.

By taking control of your GMB listing, you’ve taken the first significant step toward improving your ranking in Google local search. Google will likely have more confidence in your listing data immediately, because you have proven ownership and asserted control. This confidence is a key ranking factor, as one very important function of local search engines is to cull the most accurate, up-to-date data from among the sea of conflicting information that defines the local business information landscape. You are now poised to take the next steps in optimizing your listing.

For more information from Google about increasing your business listing exposure, see *http://bit.ly/surfacing_info*.

Google Knowledge Graph Carousel

A recent eye tracking study by Mediative revealed some interesting data on click-through rates for search results showing a Google Knowledge Graph Carousel (see Figure 10-8).

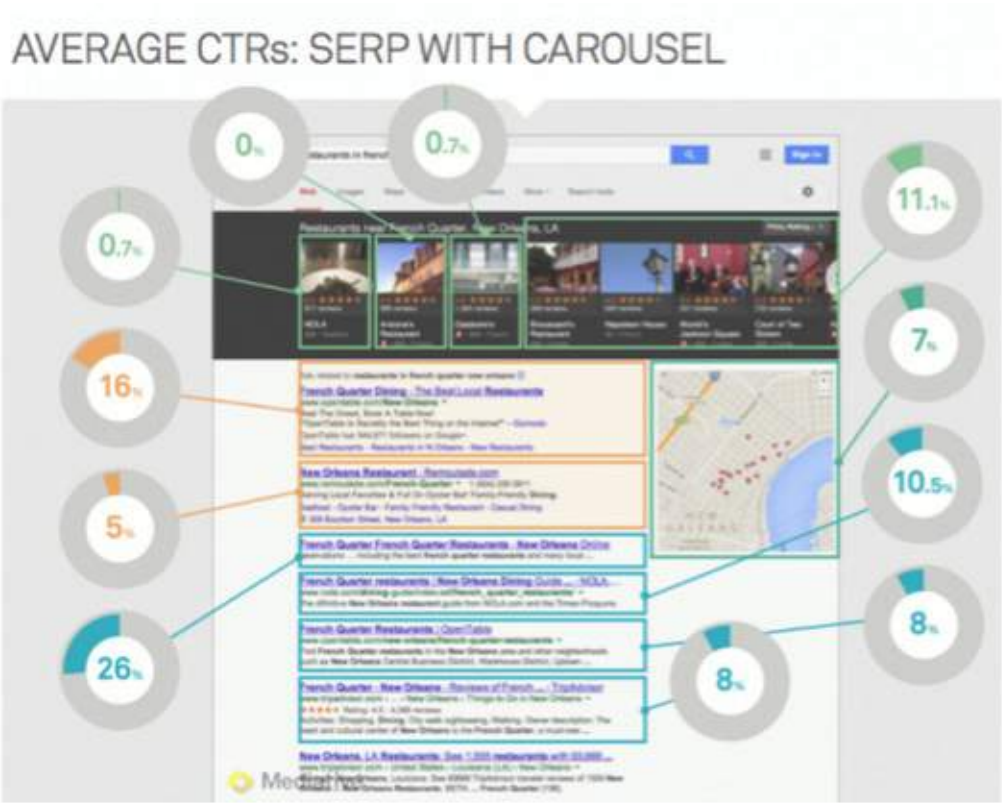


Figure 10-8. Local search result showing a Carousel

According to the study, the highest-performing Carousel results were the ones at the far right of the display, not those in the traditional 1, 2, or 3 positioning on the left. While this is just a single study, it shows the challenges that a new visual interface can create with regard to traditional user behavior.

As you can see in Figure 10-9, a search for *minneapolis pizza* is heavily biased toward local results, and Google's implementation of local Carousel results is displayed at the top. These Carousel results are sourced from the same set of data as regular local "pack" results; however, more results are displayed than the traditional pack of three, five, or seven. This graphic display, which appeared in mobile/tablet results in 2013, presents a new set of challenges and opportunities for business owners. It's important to note that Carousel has not affected all local results.

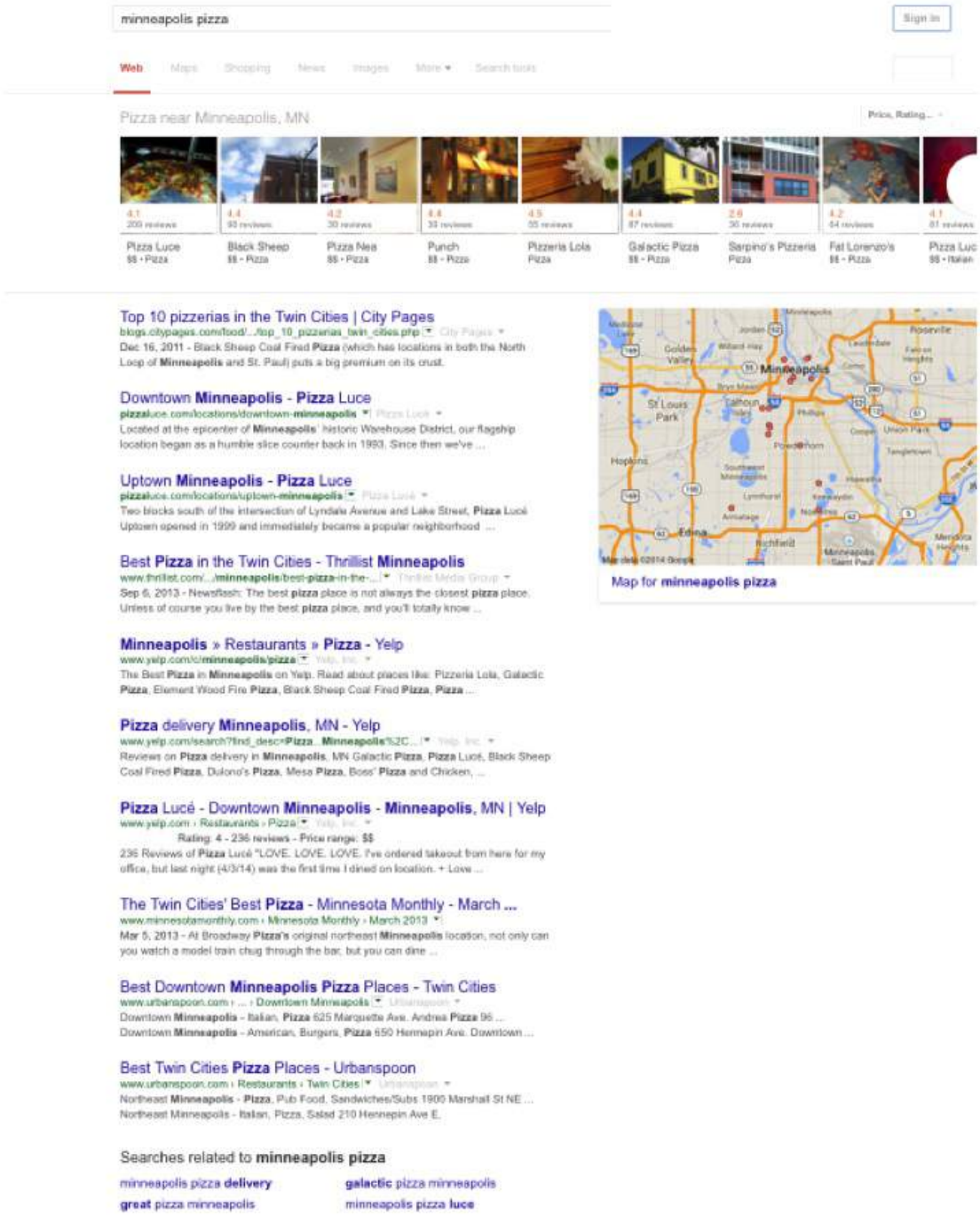


Figure 10-9. Local search results for "minneapolis pizza"

According to an October 2013 study by BrightEdge, Google Carousel impacted

14% of keywords across all industries, with Travel and Hospitality verticals being most affected, as shown in Figure 10-10.

NOTE

A fairly comprehensive list of keywords that trigger the Google Carousel can be found at *http://bit.ly/local_carousel*.

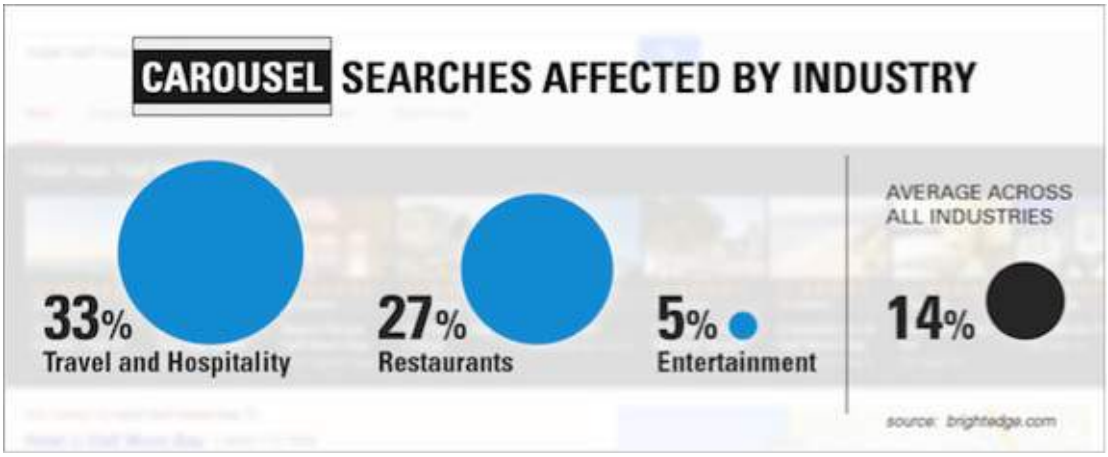


Figure 10-10. Google Carousel SERP impact across verticals

Inclusion within the Google Carousel (a.k.a. Knowledge Graph) is available for authoritative sources for organizations, events, movie reviews, and music/video play actions and requires the use of structured data markup, as outlined by Google's Developers Guide to Structured Data.

As mentioned earlier, Google also provides the Data Highlighter within Google Search Console, which enables webmasters and site owners to "tag" structured data on site pages so that it can potentially appear as "rich snippets" of content within Google search results. While webmasters have always been able to mark up structured data using various formats, with Data Highlighter now they can do so solely for Google's use without having to modify their site's pages on the server. Additional information about Google's handling of rich snippets and structured data can be found here: *http://bit.ly/structured_data_mrkup*.

Currently, Google supports rich snippets for the following content types:

- Reviews

- People
- Products
- Businesses and organizations
- Recipes
- Events
- Music

Chapter 13 provides more detail on using the Data Highlighter tool.

Bing Places for Business

Bing also enables businesses to create local listings for consideration in the organic search results with Bing Places for Business, which replaced the Bing Business Portal in April 2013. If you already have a Google My Business listing, Bing allows you to import that listing into its Places for Business interface to streamline the process (*http://bit.ly/manage_listing*). Bing also allows businesses to upload up to 10,000 business locations in one step. Figure 10-11 shows the splash screen for Bing Places for Business.

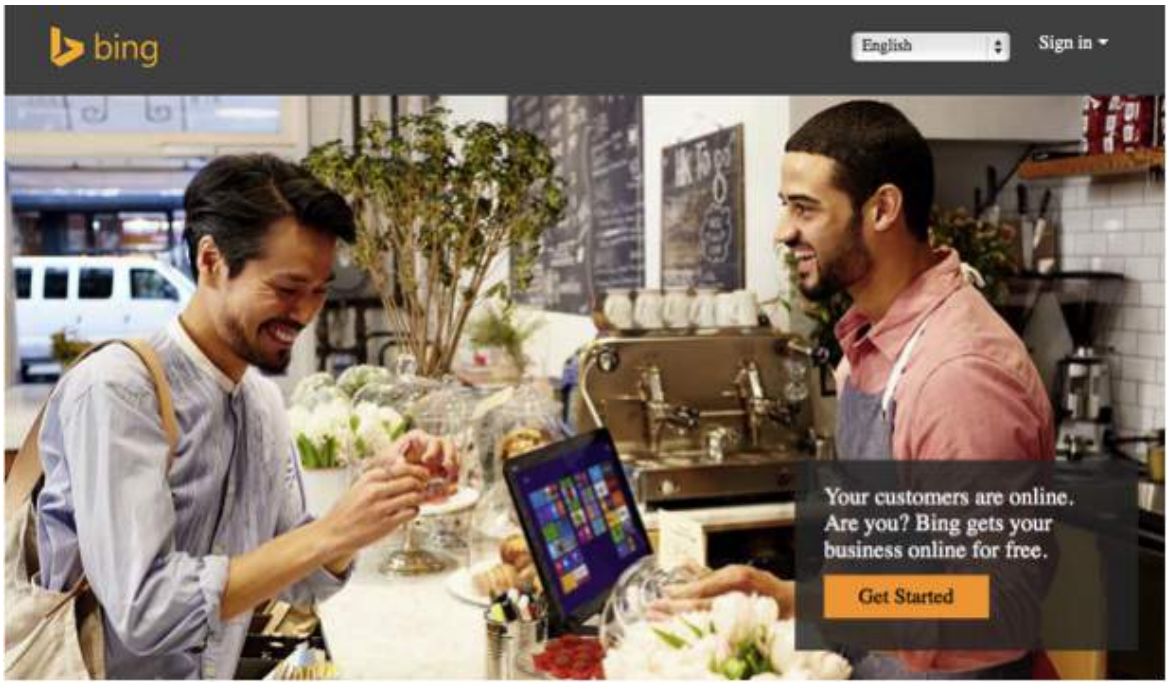


Figure 10-11. Bing Places for Business

Listing your business in Bing provides you with yet another angle from which to appear in Bing search results, according to Bing's Duane Forrester:

> *It can be a way to double dip in the SERP, as [a business's] Places listing data can appear, with [it] ranked up top in the organic stack.*

Bing also allows you to add links to your business's various social media profiles, so be sure to take advantage of this opportunity and list your Twitter and Facebook URLs.

Yahoo! Local

Yahoo! has been undergoing a steady evolution away from search as a core business function. However, it still provides value to businesses via its Small Business channel with its Local Marketing services, Localworks (local business listings in Yahoo! and 40+ directories), Mobile Offer Ads (mobile advertising), and Yahoo Gemini (its mobile search and native ad platform).

You can create a free Yahoo! Local Basic Listing, which includes your business's address, phone number, brief description, and URL; or for $9.95 per month you can create a Local Enhanced Listing, which enables you to add a

longer business description, multiple web links, and special offers/coupons.

Yahoo! also provides a mass-directory listing service called Yahoo! Localworks, which lists your business in 40+ directories including Yahoo! Local, Yelp, WhitePages, Bing, MapQuest, and Superpages; enables you to add videos and photos to your listing; and lets you track your listing performance with analytics.

Directories and aggregators

Directories can be built from the local phone company's database information, but no single phone company covers the entire country. For this reason, companies that host nationwide directories are primarily getting their content from *data aggregators* to form the foundation of their guides. Data aggregators build their content from a variety of sources, such as local-area print yellow pages, to have information that is as comprehensive as possible.

There are three top aggregators for U.S. business listings: Infogroup, Acxiom, and Neustar/Localeze. The first step in managing your local business's online presence is to check and update the business's listing information in each of these main aggregators. Ensure that the business name, address information, phone numbers, fax numbers, and any other contact information are correct. If you have multiple locations, make sure they're all listed.

In addition to checking for listing accuracy, it is important to address duplicate data issues at the data aggregator and local search publisher levels. Andrew Shotland, founder of Local SEO Guide, wrote an invaluable white paper for Yext on this topic entitled "The Definitive Guide to Duplicate Listings".

Business Website Optimization for Local Search

If you have been around for a while, your business probably is already included in the local search engines, as they compile data from the aggregators and other online directories. Once your business's listing is loaded into the local engines, you must determine how to get your business's website to rank higher when users search for your industry's keywords. On the next few pages, we will outline things you can do on your website to achieve better rankings in local search engines.

All of the basic SEO factors can come into play here and help to influence your rankings. These factors include having good, specific text in each page's `title`

tag, `<h1>` tags, meta description, page content, `alt` attributes, inbound links, and so forth. But some aspects are specific to local search, such as the following:

- If your company has multiple locations, it is not necessary to have a standalone website or subdomain (e.g., *loc1.example.com*, *loc2.example.com*, *loc3.example.com*) for each outlet. In fact, it is probably better if you don't, because each business location would likely produce similar website content. However, it probably would be helpful for you to create a separate profile web page on your site for each distinct location. Many sites with chain outlets list all outlets on one page—that is not optimal. It is better to have one page dedicated to each store location so that you can leverage all the on-page elements for that location. One caveat: you will want to have a different phone number listed for each location, preferably with a local area code (as opposed to a generic 800 number).

- Have your page title, `<h1>` tags, and content include the business name, the type of business, and the location—for example, "Acme Café: French Restaurant in Boston, MA." For multiple locations, make the title different on each location's page. Include the phone number, neighborhood, street address, area nicknames, and other location-distinguishing information.

- The home page and/or Contact Us page should have the main location's basic listing information (street address, city, state, zip code, phone numbers, etc.) displayed somewhere in the HTML text. You should also add the basic listing information in the Schema.org Organization or LocalBusiness format or the hCard microformat (these are methods for encoding address information on web pages; you can learn more about them in "Schema.org" and at *http://schema.org* or *http://microformats.org*). If you have multiple locations, display the basic information on each location's profile page. Google also accepts RDFa for this information (as well as tagging via Data Highlighter in Google Search Console). The Schema.org approach is the one we recommend, as this is a standard that has been agreed upon by Google, Bing, and Yahoo! since 2011.

- Place differentiating information on each store's pages, including items such as a map, hours of operation, brand names of products carried, product photos, "in business since <*year*>" information, menus with prices (for a

restaurant), ratings (if applicable), certifications, bonded status, and so on. As mentioned previously, be sure to include specifics about the physical location.

- In the second edition of this book, we discussed the concept of *centroid* (the location that the search engine defines as the center of the city). Many of the map search engines used to primarily display the businesses located closest to the centroid first for any particular category or keyword search. However, centroid is being replaced by business/searcher distance as a ranking factor, as the distance of a business from a searcher is more significant.

- Proximity works the same way for zip codes. If a user searches for businesses within a specific zip code, the businesses closest to the zip code area centroid will likely be displayed first in the list. Note, though, that zip code searches are not commonly performed in the United States (you can verify this with your favorite keyword research tool).

- Do not attempt to set up multiple user accounts to rate yourself well or reduce your competition's ratings! Likewise, do not pay customers for beneficial reviews—if such a practice were discovered, you could lose all of your beneficial ratings. Focus on working in ways that are allowed under all the search engines' rules: just ask people to rate you.

  — By asking enough happy clients to rate you, you might be able to drive up the positive reviews, which can bring rankings benefits. It is worth the effort to identify your happy customers and ask them to write reviews for you at one or more of the major directory sites. Never offer to pay for or otherwise incentivize positive reviews, however!

  — Consider doing post-sales surveys to find those happy customers, and then reach out to them to ask them to post reviews. You can offer to enter reviewers into a drawing to win a prize, but don't make their participation in the drawing contingent on a positive review and don't pay for each review. Monitor your online ratings in the various sites as part of good reputation management and try to respond to complaints/issues as quickly as possible. Consider using an online review monitoring tool such as Review Push and Review Trackers.

- Links from local directories and other locally relevant sites (local “citations”) are increasingly important ranking factors, and should be obtained when relevant.

- For businesses such as restaurants, it may be beneficial to add photos of celebrities who have visited the establishment. People perform searches on celebrity names all the time, so if you are lucky enough to have celebrities frequent your restaurant, uploading photos of them can help attract celebrity watchers to your business. However, be sure to get permission before using the celebrity’s photograph or “likeness” in your advertising or promotional materials, as without it you may be inviting a lawsuit. With permission, a celebrity photo endorsement can drive up inbound links and get users to linger longer on the site, perhaps bolstering the site’s quality scores.

- Community interaction can assist in online promotion. One great way to obtain local area links is to support charitable efforts in the area and to sponsor local events. Although it may seem crass to get link value out of charitable efforts, this is one of the best ways to support your local community while bolstering your business in the online marketplace. Don’t feel guilty for creating a win-win situation! Such opportunities could include:

  — Local charity races/walks, particularly if they list sponsors or sponsor logos on a page on their site. Request to have those linked to your local business site. Inbound links from these sites are very beneficial, and this is a valid way to get them!

  — Local college or high school sports teams and bands. You can request that sponsors are listed on the school website, if this isn’t done already.

  — Local fraternal organizations (e.g., the Freemasons, Lions Club, and Shriners).

  — Charitable events that you host yourself. Is anyone organizing the provision of food for the needy at Thanksgiving? If not, organize it yourself and host the information pages off your own website. This could draw many others to link to you as they help in promotion efforts.

NOTE

— A special thanks to John Biundo and David Mihm for their contributions to the "Optimizing for Local Search" section of this chapter in the second edition, and to Andrew Shotland and Dan Leibson from Local SEO Guide for their invaluable assistance with updates to this section in the third edition.

Optimizing for Image Search

According to the aforementioned Searchmetrics study, in 2013, images were displayed 45% of the time in Google SERPs, more frequently than any other blended search elements. Optimizing for image search can be a smart strategy for many search marketers. Even if you're working on a site that you don't feel truly lends itself to an image optimization strategy, you may be able to leverage images or photos on your website for SEO.

However, we should note that for some sites, there may not be a lot to gain here. Some etailers report poor conversion on image search traffic, and there is always the concern over third parties stealing their images. You need to weigh the benefit of image search capability against the costs and other opportunities for SEO on your site. Nonetheless, many sites do very well with image search, and image search optimization does not have to take much time.

A significant amount of traffic can come from image search, and the number of people competing effectively for that traffic is much lower than it is in general web search. Industries that don't immediately seem to provide compelling subjects for images may enjoy greater potential in this area, because the competition might never clue in to the advantages of integrating images into their sites and into an overall search marketing strategy.

There are a few different ways that image search optimization can help to improve traffic and conversions for your site:

Subtle reputation management

> Images of your products, services, or facility assist consumers during the research phase of their shopping, and lend an implicit message of openness/forthrightness to your business. Providing images can improve consumer confidence in your company, increasing the chances that they'll decide to do business with you.

Sales via image search results

Increasingly, consumers are searching for products via image search engines because they can rapidly find what they are seeking without having to dig through promotion-laden websites. If your products can be found in the image search engine, you have an improved chance of being found by those people. With no pictures, there's zero chance of your site being found in image search.

Increased chances of showing up in Universal Search/blended search results

Performing image search optimization improves your chances of showing up in additional positions on the main search results pages, as Universal Search pulls image search content into the main SERPs for some keyword search terms.

Site/business promotion opportunities

If you have a flexible enough organization and you hold the legal copyrights to your images, you can allow others to reuse the images in return for promotion of your site/business.

Image Optimization Tips

In comparison to indexing pages that are rich with text content, indexing images is much more difficult for search engines to perform. The image itself provides few clues to the content within it. Google is experimenting with artificial intelligence algorithms to improve detection of image content. For example, you can drag the image of a well-known person or place onto the Google Images search box, and Google will attempt to identify the contents and show you similar images.

Search engines are also experimenting with techniques such as optical character recognition (OCR) to read text content within images, but most images don't have any text to read. Search engines also use facial recognition software to determine when an image is of a face versus a body, or something else entirely.

However, although these types of technologies are very useful, they are limited in terms of what they can do, partly because they are computationally very

expensive and therefore not practical to use at the scale of the entire Web. For that reason, success in image search optimization depends on using all the signals available to you to increase the search engines' confidence in the content of your images.

This certainly includes the basic SEO techniques we have discussed in this book. The web page's `title` tag, the `<h1>` heading tag, the on-page content, and links to the page are all factors in image ranking. For example, if you have a picture of the Golden Gate Bridge as the single image on a page, and the title, headings, and content of the page all support that, the search engines' confidence in the content of the image increases. The same is true if you have 10 images on a page of many different bridges, and their content is reinforced by the title, headings, and page content. Consistency of content and theme is important in all of SEO, but it is especially critical in image SEO.

In terms of the page content, you should give particular emphasis to the text immediately preceding and following the image. This is what the user most closely associates with the image, and the search engine will view it the same way. A descriptive caption underneath the image is helpful.

You can do a number of things to further optimize your images. Here are the most important:

- Make sure the image filename or `img src` string contains your primary keyword. If it is a picture of Abraham Lincoln, name the file *abe-lincoln.jpg* and/or have the `src` URL string contain that keyword, as in *http://example.com/abe-lincoln/portrait.jpg*.

- Always use the `alt` attribute for images. This attribute helps the vision-impaired to understand your site, and search engines use it to better understand what your images are about. Our recent research indicates that this feature is still not used for lots of sites' images, and that many sites have tried to use it with invalid HTML. Make sure the `alt` parameter is valid, as in this example:

  ```
  <img alt="Abe Lincoln" src="http://example.com/abe-lincoln.jpg"/>
  ```

  Use the quotes if you have spaces in the text string of the `alt` content! Sites

that have invalid `<img>` tags frequently drop a few words without quotes into the `<img>` tag when they were intended for the `alt` attribute. Omitting the quotes is a common problem; without them, all terms after the first will be lost, if any are used at all.

- Avoid query strings for `img src` tags, just as you should for page URLs. Or, if you must use URLs that include query strings, use no more than two or three parameters. Consider rewriting the query strings in the URLs so that they do not contain an excessive number of parameters, which will cause spiders to refuse to crawl the links. Although Google claims to no longer have problems with these types of situations, it is better to be safe than sorry.

- Use good-quality pictures, which will read well when shown in thumbnail format. Good contrast is typically the key here. Lower-contrast images are visually harder to read, and it is common sense that if the thumbnail image doesn't look good, it will not invite a click.

- Do not save images as graphics files with embedded thumbnails—turn this feature off in Photoshop and other image editing software. Search engines may copy your image, reduce it in size, save it in compressed format, and deliver up a thumbnail of it for their results pages. An embedded thumbnail can wreak havoc with some compression software, and it increases your file size slightly, so just leave that feature disabled.

- Don't store the image in a sidebar column with your ads or inside the header/footer navigation elements; otherwise, the search engine algorithms will ignore the image as irrelevant, just as they ignore page decor and navigation graphics.

- Have a proper copyright license! You need to have a proper license to display the images found on your site so that you don't get sued. Be careful about trying to use images from Wikimedia Commons or other public stock photo sites, as you cannot be sure that those images really are in the public domain. For example, when you "purchase" an image from a stock photo site, you are not purchasing the copyright—you are purchasing the right to use the image. More often than not, these arrangements require that you link back to the

copyright holder.

- You need to ensure that your server configuration allows your site's images to be displayed when called from web pages on other domains. Some system administrators have disabled this setting to keep people from displaying their images on other sites, and this could cause problems if you want your images displayed in search engine image results pages. Likewise, make sure that your *robots.txt* file does not block the crawlers from accessing your image file directories.

- If it is a fit for your business, specify that others are free to use your images for online display as long as they link back to your website from a credit line below or adjacent to the image, where they display your copyright notice. Enabling others to use your photos invites more promotional attention when people wish to write about you in blogs or in news articles.

It is also a good idea to create an image XML sitemap to highlight images you consider particularly important, or to help the search engines find images that might otherwise be difficult for their crawlers to discover. Read more on image XML sitemaps at *http://bit.ly/adding_images* or in the section "Image sitemaps".

With the advent of blended search and the Web's overall movement toward more visually compelling, image-rich content, there are more opportunities to achieve organic exposure with images—and the area is also becoming more competitive. The following is a list of image optimization best practices to use when uploading your images to image sharing sites such as Flickr. These tips will help ensure your image files have the best chance of appearing for targeted, qualified searches in the SERPs:

- When you upload your photos, always add tags. The tags or keywords that you associate with your photo will ensure users can find it when they are searching and will lend keyword weight to the photo's page. Enter as many tags as possible that accurately describe your photo. Make sure you place any multiword tags within quotation marks (e.g., "pickup truck"). The Flickr Tag Cloud, Flickr's user-tag "folksonomy," generates a good link navigation system for both users and search engine spiders.

- This should be obvious, but have your photos publicly viewable, not

restricted to viewing by only your friends and family.

- Create a descriptive title for the image. This adds yet more keyword weight to the photo's page within Flickr.
- Enter a description under the photo, or write something about the picture.
- Consider adding a note or two directly onto the photo, particularly if it is humorous. Flickr allows you to select a rectangular area and associate some text with it that will appear as a tool tip when users mouse over it. Adding a humorous/interesting note or two may encourage users to participate on your photo's page, and the greater a page's level of participation/stickiness, the better quality score that page may attain.
- Ensure that photo geotagging is enabled (where applicable) for locally oriented images.
- Create thematic sets for your photos, and add each picture to the set(s) appropriate for it. This provides yet more contextual clues for search engines regarding the content of the photo's page, and it will allow a user arriving at the page to easily find similar pictures you've taken.
- Browse through Flickr's Groups for public "photo album" collections that are dedicated to pictures that could be related to your photo. Sometimes it helps to search for photos using keywords you have used in your tags, and then see what groups other people's photos belong to.
- Join those groups, and then add your photos apropos to each group's theme. The more links there are to your photo page, the more important your photo will be considered, so add it to a number of groups. Ideally, add it to groups that have a lot of members—the number of members indicates the popularity and traffic of the group.
- Link each of your Flickr photo pages to your website, or a related page on your site. You can add hyperlinks to the Description field below the photo. Use anchor text that has a call to action, or that tells the user what to expect if he clicks on the link (e.g., "We sell this product on our website"; "Enjoy this

view from the tables at our restaurant"; "This room is available at our bed & breakfast"). It is best to link to specific pages of related content as a richer indicator for link juice transfer.

- Post as many optimized pictures as possible. This is mostly a game of many small fractions adding up to large, cumulative results. The many pages of pictures linking back to your site will help build your overall authority. The more pages you have, the more likely it is that other Flickr users will find your content and link to it. This also increases your chances that a lucky picture or two might find their way onto a viral popularity wave that spurs many users to send links of your picture to their friends, or that a reporter might find one of your pictures ideal for her news story.

Many of these suggestions can generally be applied to images used on third-party social media platforms that heavily leverage images, such as Pinterest, which has become one of the leading referrers of ecommerce traffic on the Web. For these types of sites, be sure to use descriptive filenames and `alt` text, write keyword-rich descriptions, and make ample and appropriate use of hashtags.

Optimizing for Google Shopping Search

Google Product Search is now called Google Shopping. The aforementioned 2014 Searchmetrics study reported a 182% increase in 2013 in the total number of Shopping integrations in Google SERPs over the previous year. The biggest change within Google Shopping (besides the name) is that all product results now come from product listing ads—sponsored results for which Google collects a fee for every click. Google Shopping results have gained significant prominence within the SERPs, frequently appearing above the organic results and even pushing text ads down the page (see Figure 10-12). This makes it especially important for online retailers to optimize product visibility in Google Shopping.

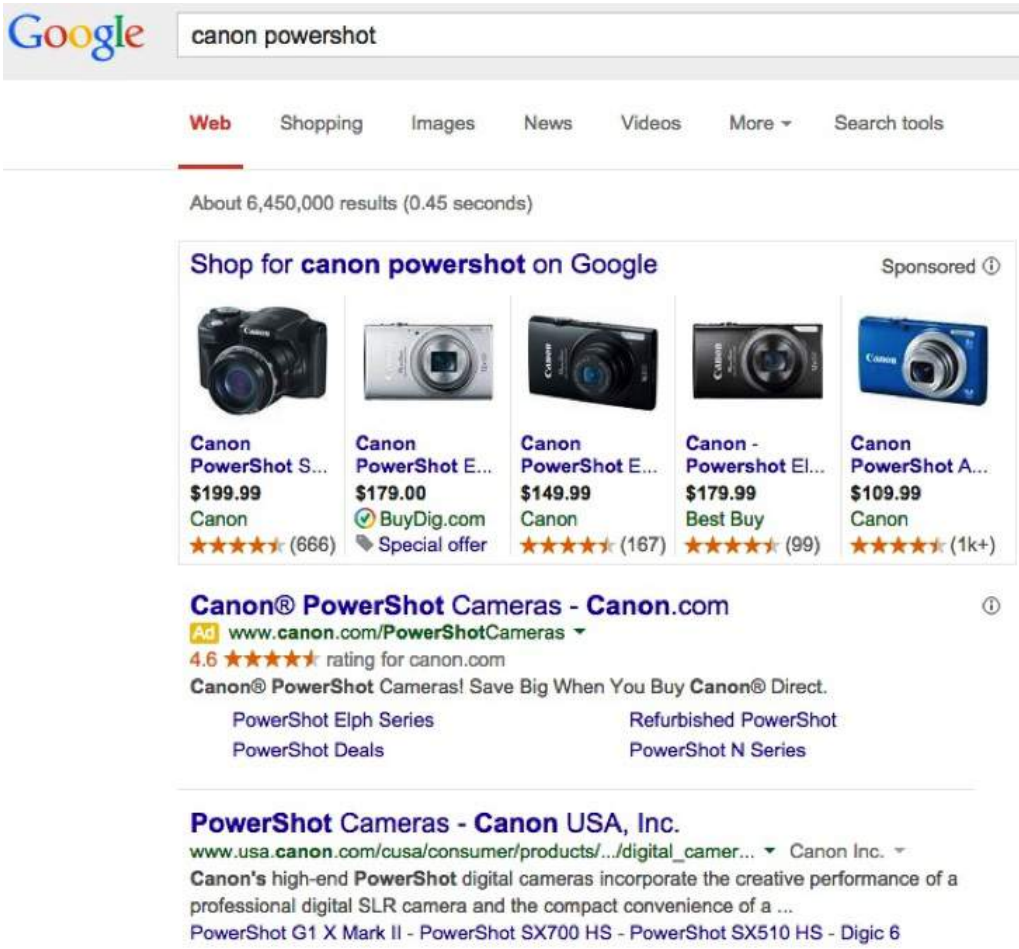


Figure 10-12. Google Shopping result for "canon powershot"

There are two main steps to list your products on Google Shopping:

1. Submit a product feed through the Google Merchant Center.
2. Create Product Listing Ads through Google AdWords.

Submitting a Product Feed

To be included in Google Shopping search, you need to upload a product feed that contains true physical/tangible products. The Google Merchant Center will accept other types of items, such as flights, hotels, car rentals, travel packages, and real estate, but such items will not get into Google Shopping search.

The latest technical specifications for submitting a product feed to the Google Merchant Center can be found online at *http://bit.ly/products_feed_specs*. It is

crucial to follow these requirements closely and to provide as much relevant data as possible to ensure Google can read what is in the feed and understand what products you're selling. The more data Google has about your products, the better it will be able to show your listings at the right time. In your feed, populate as many fields as possible with available product data (Brand, Category, Color, Price, Condition, etc.).

In addition to uploading accurate data, it is important to keep your feed data current, with the latest availability, pricing, and other specifications, to reduce the error rate and improve the chance that Google trusts your data (many merchants update their feeds daily). If the data in your feed does not match what Google finds on your landing pages, your feed quality will be lower, and it will be less likely that your products will appear in the SERPs for product queries. When in doubt, check your Google Merchant Center account to see how many errors Google finds in your feed, and either correct as many as possible or delete the products that are causing the feed errors.

Optimizing a Product Feed

Once your product feed is accepted, Google uses it to match your products to searches for those items. The extent to which Google's search engine correctly understands what products you sell is largely dependent on what you've put in your feed, and various optimization tactics can help you receive more impressions for relevant, product-related queries.

Optimize the title

Titles should contain your target keywords and be descriptive and concise. For example, for apparel you should include the brand, style, and gender. While you're allowed 150 characters, try to fit the most important words in the first 70 characters in case Google truncates it (see Figure 10-13). A strong title that will fit nicely in the limited space provided by ads might be "Red Wing Men's Oxford 133." Avoid overly descriptive or editorial language like "High Quality & Gorgeous North Face Down Jacket." Google prefers simple and product-focused language because that's what users prefer to see in the Shopping results.

Figure 10-13. Google Shopping result for "microwave ovens" showing product titles

Optimize the description

The optimal description length is between 500 and 1,000 characters, much shorter than the 5,000-character limit of the field. Like your product title, product descriptions should not be overly editorial or promotional, but rather include your keywords and specific attributes of the product—preferably near the start of the field in the event it gets truncated. The description should be grammatically correct and end with a punctuation mark. For products with variants (e.g., items that come in different colors), avoid listing the specifics of any of those variants, and instead state that the product comes in other colors. This can prevent inconsistencies between images and descriptions; for example, if you list a shirt and say it's available in green and red, Google may show the red shirt to a user searching for a green one.

Use correct product identifiers

Unique product identifiers (UPIs) are some of the most important attributes in your feed, yet they're also the ones most often submitted incorrectly. There are three types of UPIs—GTIN, MPN, and Brand—and most products require that two of the three be submitted using industry standard values.

The UPI that you submit for one of your products should be the same value that a competitor submits for that product, because these numbers are universal and Google uses them to match products to queries. If you're making up UPIs or submitting the "Identifier Exists" attribute for custom goods incorrectly, Google can't group your products correctly and you could be missing out on potential impressions.

The Identifier Exists attribute should be used for custom-made goods only; if that's not the case for your goods, you must get the proper UPI from the product manufacturer. Google will disapprove products that use incorrect UPIs.

Use the correct Google product category

Download the Google Product Taxonomy and make sure every product is categorized as specifically as possible. Using a more generic category won't get you more visibility. Google's goal is to show the most relevant results, so by giving your product the most specific categorization, you help Google better group it with similar products and map it to the most relevant queries. Google product categories are similar to UPIs in that you and your competitors should be submitting the same values for each product.

Google updates the product taxonomy every three months, so check back periodically to make sure that your products are still well categorized. If you find that your product is not showing for the right queries, experiment with the title, description, product identifiers, and categorization to improve your relevance for those missed queries.

Use appealing images

The image is the biggest visual element that users see in Shopping results, so you should submit a visually appealing, high-resolution image that clearly shows the product. You can submit up to 10 additional images in addition to the main image to show the product from different angles, in different settings, or with different packaging. If you find that your listings have a low click-through-rate (CTR), experiment with image rotation.

Submit a special offer with a promotion feed

When you're selling the same product as multiple competitors, the availability of a special offer is one of the few ways to differentiate your product listing. Use the Google Merchant Center to submit a separate feed with promotions such as discounts, free shipping, or free add-ons.

Become a Google Trusted Store

Google adds a special logo next to any merchant who is a Google Trusted Store, giving Trusted Store product listings more prominence at no additional cost (see

Figure 10-14).

Figure 10-14. Google Trusted Store "special offer" in Google Shopping

Promoting Products in AdWords

Because all shopping results are now sponsored, you will need to create product listing ads for all of the products you want to promote. To get started, create a new Shopping campaign in Google AdWords, and connect this campaign with your Merchant Center account. You can do a quick and simple setup with a single campaign, ad group, and product group that contains all products in the feed; more sophisticated setups take a bit more time and effort, but are well worth it for the additional level of control they provide over bidding and targeting.

To gain more control over how much you pay for a click and when your products appear in the SERPs, you can put product listing ads into different campaigns, ad groups, and product groups. This gives you greater control over settings like geotargeting, device bid modifiers, day parting, and negative keywords. Different bids can be set for each product group so that you're bidding the correct value for each type of product.

If you create multiple Shopping campaigns in an account, you can take advantage of an entirely new setting that's exclusive to Shopping campaigns: the ability to set priorities of high, medium, or low. In the event that multiple campaigns are eligible to show a shopping result, Google will show the product listing ad from the campaign with the highest priority. For example, you can use

this feature to set a higher priority for products that are on sale.

Reporting Results of Shopping Ads

AdWords has reports about all the key metrics such as clicks, impressions, cost, conversions, and conversion values driven by Shopping listings, helping you ensure your ads are successful. Reports also include benchmark data about the average CPC and CTR of competitors to help you see how your shopping ads perform in comparison.

Optimizing for Blog Search

News, video, and blog search provide tremendous opportunity for optimization. This has a bearing not only on obtaining traffic directly from the search engines, but also in promoting your business in whole new ways. Getting plugged in to news search, for example, can bring you plenty of quality traffic, but it can also result in your site being spotted by major media editors and writers, as well as bloggers who consume that media. This is the type of exposure that can lead to numerous links.

Blogs and RSS feeds offer a similar dynamic of getting your content in front of new readers through new channels. There is also a social aspect to blogging, due to the built-in mechanism for comments and the tendency for bloggers to interact heavily with each other.

Optimization for news and video search applies to a wide range of sites, and we'll come back to them later in the chapter. First, though, we'll look at the optimization for blog search.

Blogs are great publishing platforms for those who want to write articles on a regular basis. First, they make it easy to publish the content. Authors need only to log in and use a relatively simple set of menu choices to input what they want to publish, preview it, and then publish it. It is far easier than coding your own HTML pages by hand.

In fact, W3Techs has estimated that up to 60% of websites (from 24% of the Alexa Top 10 million sites) have been built using WordPress as the sole publishing platform. WordPress is the world's most popular blog platform, and

the search engines historically have handled WordPress-based sites very well.

Blogs are inherently social in nature, so they offer a host of social marketing benefits. Enabling comments allows for interaction with readers, and bloggers tend to have a significant level of interaction. Working this aspect of blogging as a social media platform is beyond the scope of this book. Nonetheless, be aware that a blog is an opportunity to establish yourself as an expert in a topic area, and to engage in a give-and-take activity that can dramatically change the visibility of your business.

In addition to these huge benefits, blogs can also bring you search engine and/or blog search engine traffic when they are properly optimized.

Structural Blog Optimizations

As we have discussed throughout this book, there are many key elements to successful SEO. These include things such as `title` tags, heading tags, good content, inbound links, and SEO-friendly architecture. Although the various blog publishing platforms are great, they can sometimes also require some tweaks to achieve optimal SEO results:

- Blogs usually offer the ability to categorize each post. Make sure the tag name is used in the title of that tag page.

- Override default `title` tags with custom ones. You can do this using one of the best WordPress SEO plugins available, WordPress SEO by Yoast. Along with many other SEO features, this plug-in allows you to supply a custom `title` tag, defined through a custom field in a post or a page.

- Rewrite your URL to contain keywords, and to use hyphens (preferred over underscores) as word separators. Do not let the blog platform include the date in the URL.

- Make sure you 301-redirect from *http://yourblog.com* to *http://www.yourblog.com* (or vice versa). Note that if you have a site at *http://www.yourdomain.com* and a blog at *http://www.yourdomain.com/blog*, you may need to implement a separate redirect just for the blog. This has to be handled not just for the home page, but also for all internal pages (e.g.,

permalink pages). Each URL must redirect to the corresponding URL on the *www* version.

- If you change from one blog platform to another one, the URL structure of your blog will likely change. If so, make sure you maintain legacy URLs by 301-redirecting from each page's old location to the new one.

Optimizing Your Anchor Text

Anchor text is just as important in blogging as it is in general SEO, and should also be treated with an equal amount of caution (i.e., not overoptimized). Here are some guidelines:

- Make the post's title a link to the permalink page. You do not want your only link to the post to say "Permalink."

- Use a tool such as Open Site Explorer or Majestic to see who is linking to your site. Using these tools or tools like them, you can see who is linking to you and what anchor text they have used. Look for opportunities to request revisions to anchor text on inbound links, but before making such a request make sure you are confident that your relationship with the linker will not result in her simply removing your link instead of changing it.

- Internally link back to old, relevant posts within the body of a blog post. Don't use *here* or *previously* or similar words as the anchor text; use something more descriptive for the user instead.

Sticky Posts

Sticky posts are a way to add content that always shows up first on a page of your blog—for example, a post that serves as the introduction/overview for all the content on one of your category pages. Using this technique, you can add keyword-rich introductory copy to a category page or tag page. WPMUDev.org, a WordPress developer and resource site, offers a tutorial on how to create and style sticky posts.

Author Profile Pages

If you have a multiauthor blog, another smart tactic is to create author profile pages. This has a multitude of benefits. First, many of your readers will have a stronger interest in one writer than in the others. Better still, offer RSS feeds on a per-author basis. In addition, the authors are likely to link back to their author pages on your blog from their own sites.

Figure 10-15 is an example of an author profile page.

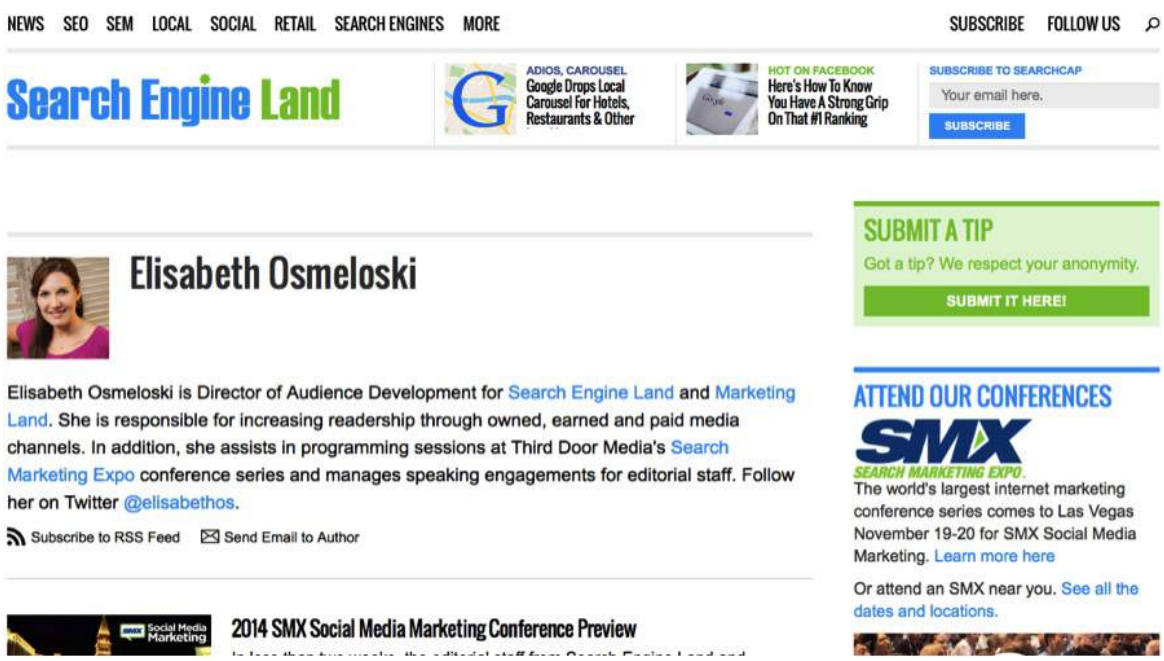


Figure 10-15. An author profile page

Links

Obtaining links and managing your link juice remain critical activities. Blog platforms provide limited ability to manage your internal link juice, so this may require some customization to accomplish. Fortunately, in the WordPress environment, some really good plugins are available to help you with this.

For starters, you can also create cross-links between related posts using a plug-in such as WordPress Related Posts. This is a great way to get people who just finished reading one post on your blog to consider reading another one. Finally, develop relationships with other blog and article authors, and get them interested in your content; hopefully, they will follow your work and offer relevant citations when the opportunity arises.

Optimizing for News Search: Google News

As a major traffic driver to news-based web content, Google News should be at the top of your list for news search optimization. Google News offers publishers visibility beyond just the Google News site, as Publisher articles will also show up in regular Google searches. Certain queries around headline news and current events will trigger Google to show a fresher set of results. Some of these articles are presented in a cluster and labeled “In the news,” which gives publishers the opportunity for their most recent articles to appear at or near the top of the results page.

There are many types of queries that will trigger a “news box.” Examples are disasters (tsunamis, earthquakes, storms), attacks (school shootings, terrorists), elections, holidays, significant sporting events (Olympics, World Cup, Superbowl), large brand names, and deaths of notable persons. Figure 10-16 shows a news box inserted at the top of the SERP for the query *ebola*.

Until very recently, news clusters in search results contained only articles from approved publishers accepted into Google News. In October 2014, Google began including other sources in the news box as well, such as reddit, Twitter, or very small niche sites.

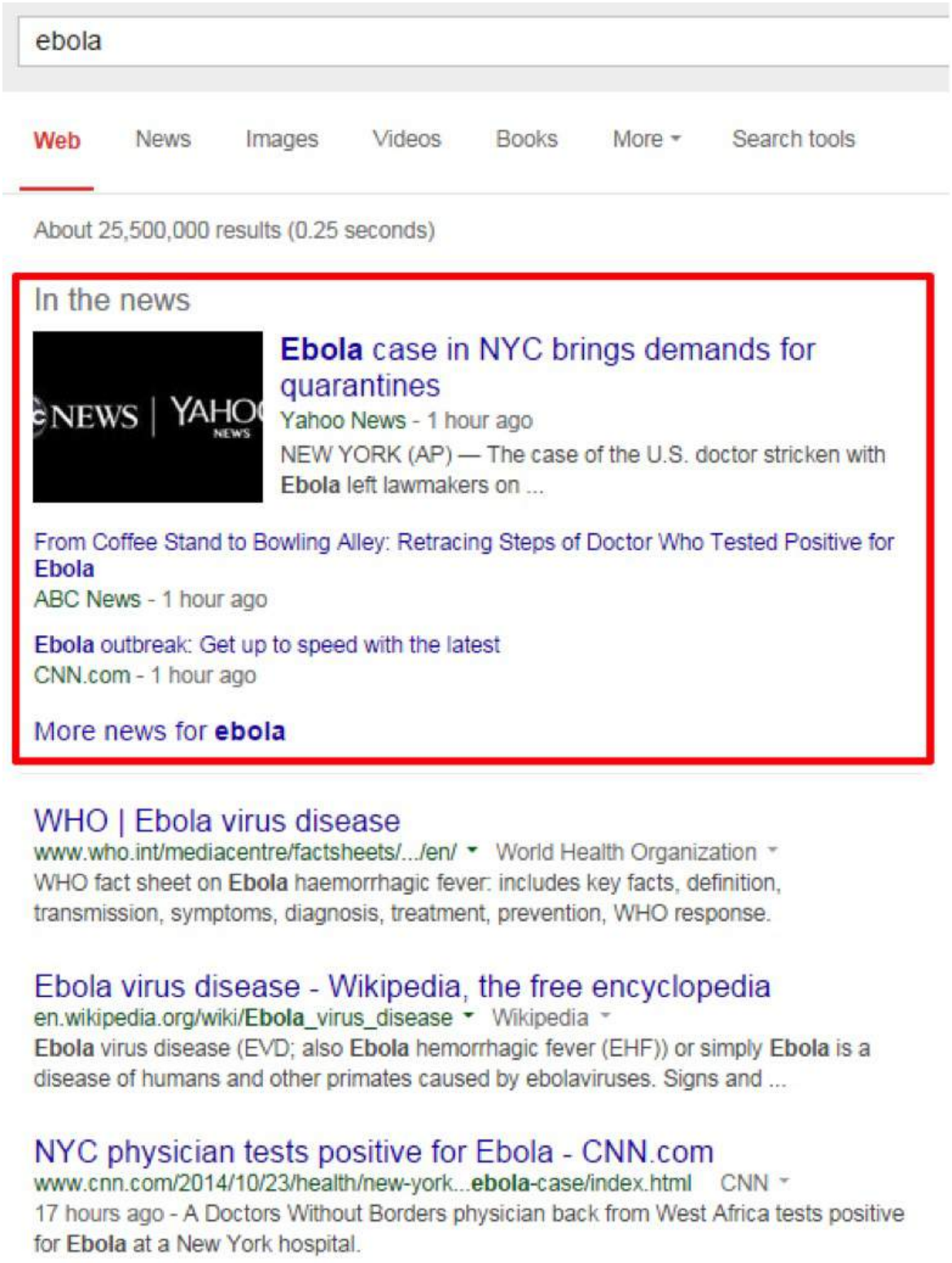


Figure 10-16. A news box in a Google SERP for a search on "ebola"

Acceptance Criteria

Publishers must meet very strict quality standards to be accepted into Google News, and these standards are strongly enforced (both for getting in and for staying in). In Google's News Publisher Help section, Google outlines its News requirements in detail. You should read through these guidelines very carefully before applying.

Here are some of the general requirements:

- Your site must contain clearly accessible contact information, including a physical address, phone numbers, and email.
- If your site has a common publication name, there is likely already a publisher in Google News with the same name. You will need to submit your site with a unique name. For example, the *Star Tribune* submitted its publication name as *Minneapolis Star Tribune*. Be sure to omit extraneous articles (such as *the*) and descriptive clauses.
- If your site is available in multiple languages, you must submit separate requests for each version.

Application Process

Once you are certain that your site complies with all the requirements, submit your application to Google News at *https://support.google.com/news/publisher/answer/40787#contact=1&ts=317919*

Be thorough and precise when filling out the application. It can take a couple of weeks for Google to review an inclusion request. If the application is denied, no reason will be provided and you cannot reapply for 60 days, so be sure that your submission is as accurate and complete as possible.

Paywalls and Subscription Sites

Sites that require registration to view articles or have a "paywall," whereby article access requires a subscription, are labeled as such in Google News (see Figure 10-17).

Figure 10-17. Google News content behind a paywall

Stories labeled as "subscription" tend to receive fewer clicks than articles that are free. This consequently reduces the click-through-rate for that publication. Because CTR is a strong signal in the ranking algorithm, subscription sites will tend to get less visibility in Google News than sites that are free to users.

As an alternative to labeling a site as a subscription, Google offers a program called First Click Free, or FCF (*http://bit.ly/reg_sub*). With FCF, publishers must allow users referred from Google to bypass the paywall for that particular article. For any additional clicks during that session, the publisher can present the user with the paywall.

FCF requires that visitors from Google be permitted five free clicks each day, regardless of whether the paywall is *hard* (no free article views) or *metered* (views allowed up to a certain number of articles within a certain timeframe, typically one month). For example, the *New York Times* limits nonsubscribers to viewing 10 articles per month. It is important to understand that FCF *overrides* the publisher's metered allotment. Once the user has viewed her monthly allotment of free articles, she must subscribe (or register) to view additional articles. However, sites that implement FCF must still allow this user to access the first article free if she is referred from Google, and they must allow this five times each day.

Google News Publisher Center

Google News is different from Google web search in that it is a static environment when it comes to content discovery. The crawler goes no more than one level deep from a set of pages that are entered into its database when a publisher is accepted into Google News. If your site launches a new section page

or subsection page, Google News will not find that page just from crawling a new link in your navigation. You must notify the Google News publisher help team, and they will manually update their system to add a new source page to your crawl.

For dynamic, growing publisher sites, this creates a tedious and often uncertain cycle of requesting frequent changes to your registration, which is a manual process. Google hates manual processes, hence the new Google News Publisher Center. This is a tool designed to facilitate the addition, removal, and "labeling" of new site pages by allowing publishers some control over updating section pages and their labels.

How to add, edit, or delete a section URL

Figure 10-18 and Figure 10-19 show the screens for managing sections in Google News.

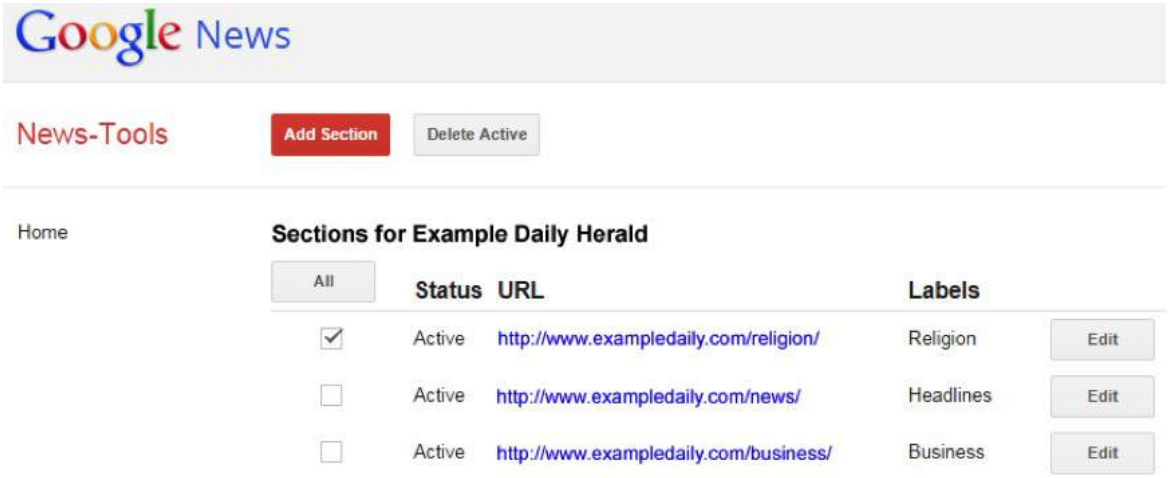


Figure 10-18. Add/deleting a section URL in Google News

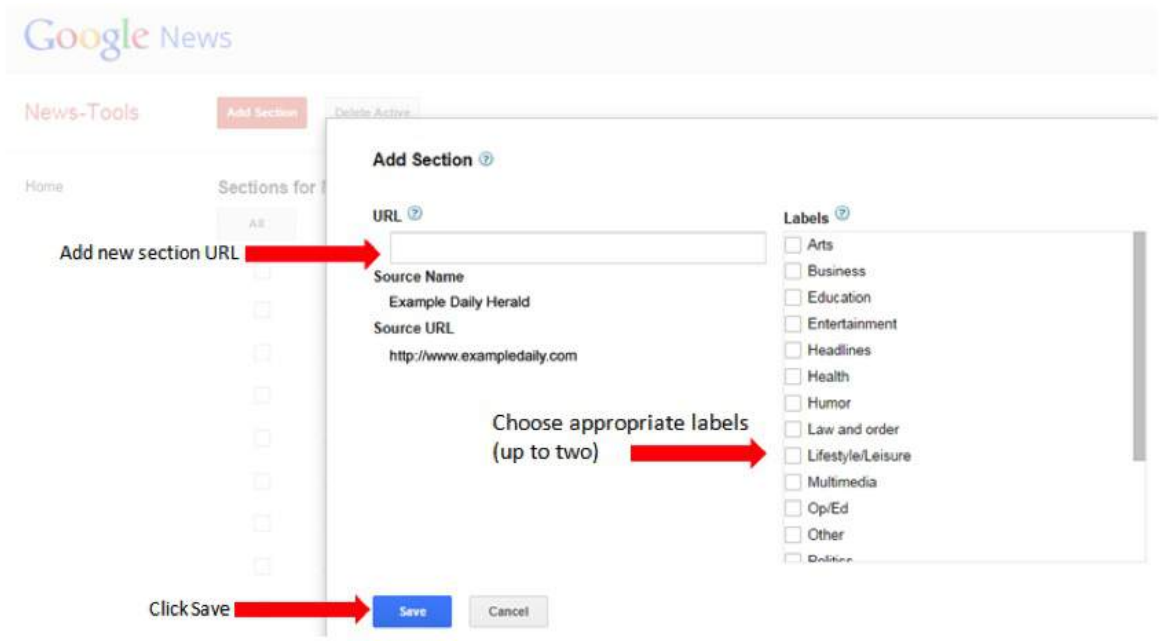


Figure 10-19. Adding and labeling a section URL in Google News

How to update source details

Figure 10-20 shows the screen for managing source details in Google News.

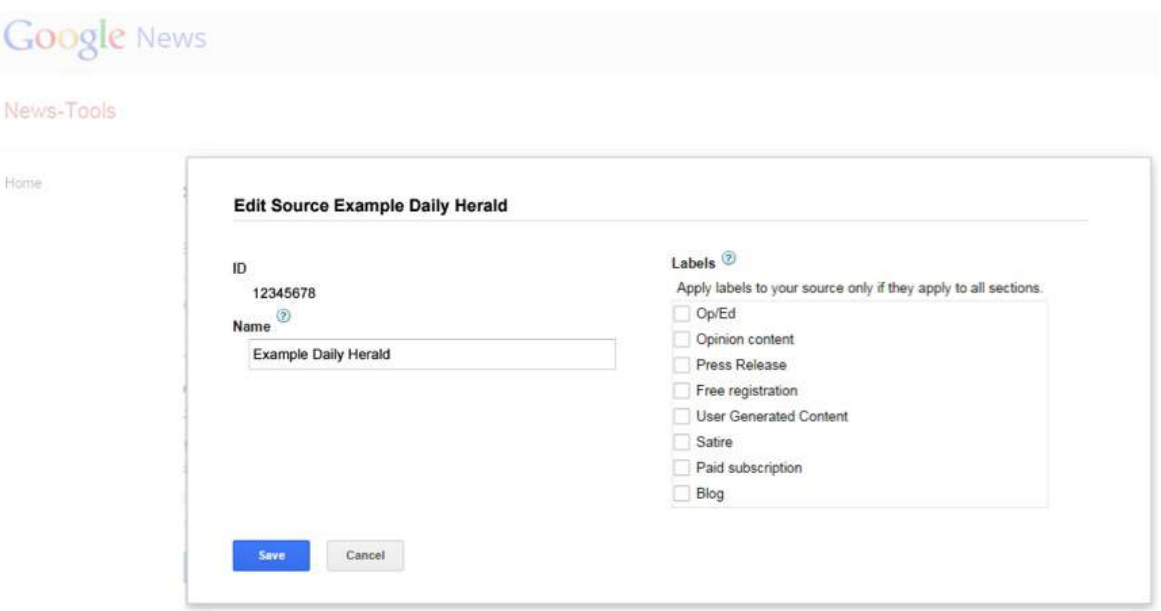


Figure 10-20. Editing source details in Google News

Technical Requirements

Google News has some unique technical content requirements, many of which are more particular than the web search algorithm. For example:

- The text that makes up an article's date information and author byline should be the only text or code between the article's headline and the article's first sentence. It should be placed in a separate line of HTML between the article's headline and the article's body text.

- The URL for each article must contain a unique number consisting of at least three digits (*http://bit.ly/article_urls*); however, this rule is waived for URLs submitted via News sitemaps.

- Articles must contain a minimum of 80 words. This means that short news briefs or pages with a video and a single brief paragraph are unlikely to be included in the news index.

- The HTML source page of a section or article must be less than 256 KB in size. This can be an issue for pages that use infinite scroll or have a substantial number of user comments.

- Pages that display multiple articles at the same URL will not be included.

- Image links or links embedded in JavaScript cannot be crawled.

- Section and article pages must be on the same domain (or subdomain) as your main site (exceptions are videos on YouTube.com and RSS feeds on Feedburner.com).

- Sites encoded in UTF-8 are optimal, and article pages must have an HTTP `content-type` of `text/html`, `text/plain`, or `application/xhtml+xml`.

- If your publication mixes news with other types of content such as advice columns, how-to articles, weather forecasts, stock data, classified ads, or paid advertorials, Google News expects you to prevent this content from getting crawled. There are three methods you can use to accomplish this:

  — Use a page-specific meta tag:

```
<meta name="Googlebot-News" content="noindex, nofollow">
```

— Place all of your non-news content in different subdirectories (or subdomains), such as *http://example.com/jobs* or *http://jobs.example.com* and block Googlebot-News access to these sections in the *robots.txt* file.

— Create a Google News sitemap for your news articles only and inform Google News (*http://bit.ly/report_issue*) that you prefer to have your articles crawled exclusively from there.

Headlines

The technical specifications for headlines are as follows:

- Headlines must contain at least 10 characters and be between 2 and 22 words in length.
- Do not link article headlines (linked headlines are a default setting in some site templates).
- Wrap the article headline in an `<h1>` tag. Ideally this will be the only `<h1>` tag on that page.
- The article headline should also be in the `<title>` tag, placed before any additional information such as the publisher name or the section.

  — Like this:

  ```
  <title>This is the headline - Lifestyle - Daily Publisher</title>
  ```

  — But not like this:

  ```
  <title>Daily Publisher - This is the headline</title>
  ```

  — And not like this:

  ```
  <title>Lifestyle: This Is the Headline - Daily Publisher</title>
  ```

- If the headline in the `<title>` tag is significantly different from the headline

in the `<h1>` tag, this can prevent the story from being included in Google News.

Meta tags

Google News has certain requirements regarding the use of meta tags as well.

canonical

If you publish the same article on multiple pages within your site, use the `rel="canonical"` link element to specify which URL to rank for a story.

Meta description

Google News does *not* use the meta `description` tag for snippets. Instead, the crawler looks at article body text near the headline (*http://bit.ly/incorrect_snippets*).

There are two meta tags that are specific to Google News and do not apply to web search: `news_keyword` and `standout`.

news_keyword

This tag is the only on-page keyword tag that is used for ranking by the major search engines.

The `news_keyword` tag is placed within the page's `<head>` tags. It can be used to specify keywords that are highly relevant to the article but might not be in the headline or the first paragraph of the story. For example, an article with the headline "USA vs. Belgium Breaks Another Ratings Record" could use the tag like this:

```
<meta name="news_keywords" content=" FIFA, soccer, football, World Cup 2014">
```

Commas are the only punctuation allowed in this field and should be used to separate each keyword or phrase. You can use up to 10 terms for a given article and all keywords are given equal value, so the first keyword is not considered a stronger signal than the tenth keyword. Words that are contained in the headline or the first paragraph do not need to be repeated in the `news_keywords` tag.

standout

When a publisher breaks a news story, it is important to Google News to attribute that article as the original source, but algorithmically this has been a challenge. In late 2010, Google announced a new meta tag named `standout` that publishers can use to designate a particular article as the source of a story or as an exceptional piece of journalism. Google provides the following criteria to determine when the use of this tag is appropriate:

- The article is an original source for the story.
- Your organization has invested significant resources in reporting or producing the article.
- The article deserves special recognition.
- The `standout` tag has not been used on your own articles more than seven times in the past calendar week.

When another publisher breaks a story and your organization writes about it, refers to it, or otherwise draws from it, Google strongly suggests using this tag in your article to credit that publisher by citing its URL in your `standout` tag. If your article draws on more than one piece of original or exceptional journalism, you can use multiple `standout` tag citations within your article. Citing `standout` articles from other publishers does not count against your limit of seven self-citations per week.

The `standout` tag cannot be added after the article has been published. This tag is placed in the page `<head>` and can be coded as a meta tag or as a link tag:

```
<head>
...
<meta name="standout"
content="http://www.example.com/breaking_exclusive_story_2314">
<head>
<head>
...
<link rel="standout"
href="http://www.example.com/breaking_exclusive_story_2314">
<head>
```

When the `href` URL points to itself, Google interprets the tag as a self-citation. When the `href` URL points to another publisher's page, the tag is considered to be an out-citation.

Article expiration

This tag is treated as a removal request and is used to specify that an article should be removed from the Google index at a certain time. To function properly, the tag must be included in the article at the time that it is first crawled. It will take about a day after the removal date passes for the page to disappear from the search results. The date and time must be specified in the RFC 850 format as follows:

```
<meta name="googlebot" content="unavailable_after: 25-Aug-2011 15:00:00 EST">
```

Date

This tag can be used in addition to the date and time specified in the article byline. Use W3C format:

```
<meta name="DC.date.issued" content="YYYY-MM-DD">
```

Thumbnail Images in Google News

Google News includes a thumbnail image in almost every cluster displayed in either web search results or on the Google News home and section pages (see Figure 10-21). Having your article's image used as the thumbnail for a cluster can increase the amount of clicks that story receives.

Figure 10-21. Google News article thumbnail image

Images must be hosted on the publisher's domain—not an external domain or content delivery network (CDN)—and must use a standard filename extension like *.jpg*.

Use the following guidelines to increase the likelihood of getting your article images included:

- Place images near the article headline.
- Place images inline within the article content.
- Use well-written captions.
- Use images that are fairly large in size (at least 60 pixels by 60 pixels).

You can find additional image information for Google News at *http://bit.ly/missing_images*.

Recrawling

Google News does recrawl articles, but most recrawls occur only within the first few hours after discovery. Google's system is generally better at detecting and displaying bigger updates, such as a substantial change to a headline or the lead paragraph. If an article is edited after being indexed in Google News, there is no

guarantee that the article will get recrawled and updated.

Google News Sitemaps

Google News uses the standard Sitemaps protocol, but with additional news-specific tags such as related keywords, article images, content genres, and relevant stock tickers. Even though a standard sitemap file can include multiple types of content, it is better to create a separate News sitemap for your news content. A News sitemap allows your content to be discovered more efficiently and comprehensively because you can feed all of your news articles to Google News even if they are not linked from a section page in the News database. If your site contains a mix of news and non-news content, you can use the sitemap to exert more granular control over which content is submitted. The News sitemap will accept URLs that do not include the three numbers required for discovery from crawling.

A News sitemap is particularly recommended for:

- Sites that have been recently added to Google News
- Sites that often add new section pages or change existing section URLs
- Sites that require users to follow several links to reach news articles

---

NOTE

The Google Sitemap Generator *cannot* be used to create a News sitemap. There are a number of third-party tools recommended by Google that can help you to generate a Google News sitemap (*http://bit.ly/sitemap_generators*).

News sitemap guidelines

Here are the guidelines for creating a News sitemap:

- Upload the sitemap to your root directory.
- Update the sitemap immediately after publishing a new article.
- Do not include articles older than 48 hours.

- For the publication date, refer to the W3C format, using either the “complete date” format (`YYYY-MM-DD`) or the “complete date plus hours, minutes, and seconds” format with time zone designator (`YYYY-MM-DDThh:mm:ssTZD`) format. Use the original date and time at which the article was published on your site, not the time the article was added to your sitemap.

- A News sitemap with more than 1,000 URLs must be broken into multiple sitemaps with a sitemap index file.

News sitemap fields

The following subsections describe the guidelines and requirements for the News sitemap fields.

Publication name

Make sure that your publication name in the sitemap exactly matches the name you used when you submitted your application for Google News. If your publication name is *Daily Herald*, but you submitted it as *Gotham Daily Herald* because Google News already contained a publication named *Daily Herald*, your sitemap publication name must also use *Gotham Daily Herald*.

Access (paywalls and registrations)

The `<access>` tag takes one of the following values:

`Subscription`

An article that requires users to pay to view content

`Registration`

An article that requires users to sign up for an unpaid account to view content

For publishers that include a mix of free and paid articles, the only way to designate this on an article-by-article basis is by using the `<access>` tag within a News sitemap feed.

Genres (content types)

Values for the `<genres>` tag are required when applicable and restricted to the

following five options. The tag can contain more than one value, separated by commas:

`Satire`

An article that ridicules its subject for didactic purposes

`OpEd`

An opinion-based article that comes specifically from the Op-Ed section of your site

`Opinion`

Any other opinion-based article not appearing on an Op-Ed page

`PressRelease`

An official press release

`Blog`

Any article published on a blog, or in a blog format

Keywords

The `keywords` field contains a comma-separated list of keywords describing the topic of the article. Keywords may be drawn from, but are not limited to, the list of existing Google News keywords.

Images

The most effective way to get your images indexed is to include them in your News sitemap feed. Here is an example of a sitemap entry for a story with an associated image:

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
xmlns:news="http://www.google.com/schemas/sitemap-news/0.9">
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1">

<url>
```

```
<loc>http://www.example.org/sample.html</loc>
<news:news>
<news:publication>
<news:name>Example Times</news:name>
<news:language>en</news:language>
</news:publication>
<news:publication_date>2014-10-23</news:publication_date>
<news:title>President Announces New Education Initiative</news:title>
</news:news>

<image:image>
<image:loc>http://img.example.com/story-image.jpg</image:loc>
</image:image>
</url>
```

Stock tickers

For business articles, you can include a comma-separated list of up to five stock tickers of companies that are the main subject of the article. Each ticker must be prefixed by the name of its stock exchange, and must match its entry in Google Finance (*http://bit.ly/missing_images*).

Example Google News sitemap

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
 xmlns:news="http://www.google.com/schemas/sitemap-news/0.9">
xmlns:image="http://www.google.com/schemas/sitemap-image/1.1">

<url>
<loc>http://www.example.org/business/article55.html</loc>
<news:news>
<news:publication>
<news:name>Example Times</news:name>
<news:language>en</news:language>
</news:publication>
<news:publication_date>2012-12-23</news:publication_date>
<news:title>Companies A, B in Merger Talks</news:title>
<news:access>Subscription</news:access>
<news:genres>Opinion,Blog</news:genres>
<news:keywords>business, merger, acquisition, Company A, Company B</news:keywords>
<news:stock_tickers>NASDAQ:A, NASDAQ:B</news:stock_tickers>
</news:news>
<image:image>
<image:loc>http://img.example.com/story-image.jpg</image:loc>
```

```
</image:image>
</url>
```

Google News sitemap validation

The following XML schemas define the elements and attributes that can appear in a News sitemap file. A News sitemap can contain both News-specific elements and core sitemap elements:

- For News-specific elements, schemas can be downloaded from *http://bit.ly/news-specific_schemas*.
- For core Sitemaps elements, schemas can be downloaded from *http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd*.

There are a number of tools available to validate the structure of your sitemap based on these schemas. You can find a list of XML-related tools at the following locations:

- *http://www.w3.org/XML/Schema#Tools*
- *http://www.xml.com/pub/a/2000/12/13/schematools.html*
- *http://www.xml.com/pub/a/2000/12/13/schematools.html*

To validate your News sitemap file against a schema, you'll need to provide additional headers in the XML file as shown here:

```
<?xml version='1.0' encoding='UTF-8'?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
     xmlns:news="http://www.google.com/schemas/sitemap-news/0.9"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
     xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
       http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd
       http://www.google.com/schemas/sitemap-news/0.9
       http://www.google.com/schemas/sitemap-news/0.9/sitemap-news.xsd">
<url>
    ...
</url>
</urlset>
```

Google News sitemap submission

There are a couple of different ways to submit a sitemap to Google News.

Via robots.txt

You can add the location of your News sitemap file to the *robots.txt* file.

```
Sitemap: http://example.com/sitemap_location.xml
```

This directive is independent of the `user-agent` line, so it does not matter where you place it in your file. If you have a sitemap index file, you only need to include the location of the index file rather than each individual sitemap.

Via Google Search Console

Even if you specify the location in the *robots.txt* file, you should also upload the sitemap through your Search Console account. Doing this will provide you with access to detailed information should your sitemap contain errors.

Here's how to submit a sitemap in Search Console:

1. Log in to Google Search Console or set up an account if you don't have one.
2. On the Search Console home page, click the site for which you want to manage sitemaps.
3. On the lefthand menu, click Crawl to expand the Crawl menu, and then click Sitemaps from the Crawl menu list.
4. Click the Add/Test Sitemap button in the top-right corner of your screen.
5. In the text box that pops up, add the complete URL path to your sitemap or the sitemap index file.
6. Click Submit.
7. Refresh your browser to see your new sitemap in the sitemaps list.

Here's how to test a sitemap:

1. Log in to Google Search Console or set up an account if you don't have

one.

2. On the Search Console home page, click the site for which you want to manage sitemaps.

3. On the lefthand menu, click Crawl to expand the Crawl menu, and then click Sitemaps from the Crawl menu list.

4. Click on a sitemap in the list to open the Sitemaps Details page and click Test Sitemap in the top-right corner. Alternatively, if you don't see the sitemap you're looking for, click Add/Test Sitemap on the main Sitemaps page, enter in a URL path in the text box of the dialog that appears, and click Test.

5. Once the test is completed, click Open Test Results to see your test results.

6. Once you correct any issues identified by the test, you can click Resubmit to alert Google of any changes.

And finally, here's how to view the Sitemap Details page:

1. Log in to Google Search Console or set up an account if you don't have one.

2. On the Search Console home page, click the site for which you want to manage sitemaps.

3. On the lefthand menu, click Crawl to expand the Crawl menu, and then click Sitemaps from the Crawl menu list.

4. In the list of sitemaps, click the sitemap you want to examine to open the Sitemaps Details page. From this page, you can view errors, indexing statistics, and more.

News crawl errors

Publishers in Google News can view news-specific crawl error reports (*http://bit.ly/news_crawl_errors*) in their Google Search Console account:

1. From the Dashboard, click Crawl > Crawl Errors.

2. Click on the News tab to see crawl errors specific to your news content.

3. Crawl errors are organized into categories, such as “Article extraction” or “Title error.” Clicking on one of these categories will display a list of affected URLs and the crawl errors they’re generating.

Videos in Google News

Only videos from publishers accepted in Google News will be included. Videos can be embedded in an article with the YouTube player, or publishers can create a YouTube channel and share it with Google News. To be included in Google News, YouTube channels and embedded videos must adhere to the following guidelines:

- Videos should be timely, reporting on current news events. No how-to videos, promotions, movie trailers, or music videos are allowed.

- Videos should be uploaded as quickly as possible. This will help them to be grouped with the most recent articles.

- Videos containing a single story rather than multiple segments are preferred.

- Video titles should be descriptive and specific (not generic like “Breaking News”).

- The video description field should contain a lot of detail. Convey the who, what, when, where, and why of the story.

- The video `keyword` tags field is very helpful. Include the names of the people, places, companies, and events mentioned in the video.

- Each video must be assigned to a category when it is uploaded. For general news coverage, select the YouTube category “News & Politics” (which is `youtube_category_id 25`).

- Audio must be clear and easy-to-understand. Images should be in focus.

- All content must be original and unique. Music, images, and text within

videos must belong to your organization or be used with legal permission.

- Videos must be embeddable on other sites to be displayed in Google News.
- Video that are geoblocked in certain regions will not appear in Google News.

Additional YouTube channel guidelines are as follows:

- Only YouTube channels from publishers already in Google News will be included.
- Regularly updated channels are more likely to be accepted.
- A YouTube channel can be customized to reflect your organization's branding. In the channel profile section, include information about your organization and links to your site.
- If you have multiple channels, clearly identify each category (politics, business, entertainment, sports, etc.).

Once the YouTube channel is ready, inform Google by using its "Adding a New Section" form:

1. Add the URL of your YouTube channel in the field under "New section URLs."
2. In the pull-down menu for Category, select "YouTube channel," which is near the end of the list.
3. Make sure you also include your publication name and URL at the start of the form.

You can find additional information about videos in Google News at these sites:

- *http://bit.ly/submitting_video*
- *http://bit.ly/best_practices_news*

Editor's Picks

The Editor's Picks feature allows publishers to showcase original, innovative news content with their publication logo. Editor's Picks are displayed in a sidebar module in the right column on the Google News home page and certain section pages. Editor's Picks can include content that Google News might not index otherwise, such as long-form narrative articles, slideshows, interactive graphics, or video stories. Publishers can provide up to five links that represent the organization's best original journalistic work at any given moment.

Publishers can participate in Editor's Picks by creating and submitting a custom RSS or Atom feed. Up to three different feeds can be submitted—one for the home page, and one each for the Technology and Business sections.

The following are a handful of important guidelines about submitting and maintaining this feed:

- Include only news content (no links to subscription offers, how-to articles, stock quotes, classified ads, weather forecasts, etc.).
- Individual items in the feed can be older than 48 hours, as long as at least one item in the feed has been recently updated. Ideally, the feed should be updated once or twice per day. At least one additional item must be added to the feed every 48 hours. Feeds not updated at all within 48 hours will not be shown to users.
- Only feeds containing at least three articles are displayed. If there are more than five articles, only the top five articles will be shown.
- It is recommended to include an author name for each article. Without this information, the byline will be presented as "<*Publisher*> Staff."
- Headlines longer than 75 characters get truncated.

Once your feed is ready, submit it at *https://support.google.com/news/publisher/contact/editors_picks*.

Additional guidelines and instructions on Editor's Picks are available at: *http://support.google.com/news/publisher/bin/answer.py?answer=1407682*.

NOTE

A special thanks to Alex Bennert for her contribution to the Google News section of this chapter.

Optimizing for Video/Multimedia Search

Video search offers a tremendous opportunity for organic search traffic via your video content. The first component of video search optimization involves improving the ranking of your video content in Google's "universal" search results. The second involves improving the ranking of your videos in YouTube (the second largest search engine).

According to data from comScore qSearch for May 2014, in the United States, Google had 12.5 billion explicit core searches that month, YouTube had 4.0 billion search queries, Microsoft had 3.5 billion searches, and Yahoo! had 1.8 billion. So, even if most SEOs know YouTube is the second largest search, they still focus on Google.

An analysis of Google Universal Search results by Searchmetrics (June 18, 2014) found that videos appear in 65% of Google searches in the United States. However, 54% of these video results were from YouTube, about 5% were from Vimeo, about 5% were from Dailymotion, and no other video provider got more than about 1%. In addition, the average first video integration from YouTube was about two positions ahead of the average first position of a competitor's video.

So, although it may seem counterintuitive, your best strategy for improving the ranking of your video content in Google Universal Search results often involves improving the ranking of your videos in YouTube first.

Video SEO for YouTube

Today, more than 100 hours of video are uploaded to YouTube every minute. That means 144,000 hours of video are uploaded to YouTube every day. So you've got to optimize your videos, playlists, and channel to give them the best potential for success on the platform.

Although YouTube's algorithm doesn't change over 500 times a year like Google's algorithm does, the tips outlined here should come with the advisory

best if used before the next algorithm update. For example, YouTube officially replaced “view count” with “watch time” as a ranking factor in October 2012. This adjusted the ranking of videos in YouTube search to reward engaging videos that keep viewers watching and benefit channels if their videos drive *more viewing time across YouTube*.

On the other hand, “relevance” has remained a ranking factor since the early days—that is, before Google acquired YouTube for $1.65 billion in October 2006. So, recognizing that some things will change while others will remain the same, here are key strategies, best practices, and important tips for video SEO (freshness date: July 2014).

Conduct keyword research

When conducting keyword research for video SEO, one of the first tools that you should consider using is Google Trends. Google Trends enables you to take popular search queries and explore traffic patterns over time and geography. In March 2013, Google Trends added YouTube search data going back to 2008, making it a great tool to look at video trends. Just visit Google Trends and enter up to five terms or phrases in the search box, and the default setting will display web search interest. Then, on the upper right, click on “Web Search” and choose “YouTube Search” for YouTube search interest. You can slice by region, time, or category as well. You will quickly see—and be able to show others—the similarities and/or differences in Google and YouTube search trends. For example, web search interest in “Chevy” and “Chevrolet” was the same in June 2014. But, if you compared the YouTube search interest for these two terms, you would see that “Chevy” was almost twice as popular as “Chevrolet” that month.

Another trending tool to check out is YouTube Trends, which shows you the latest trending videos and topics on YouTube.

Optimize your metadata

YouTube uses metadata—your video’s title, tags, and description—to index your video correctly. To maximize your presence in search, suggested videos, and promotion, make sure your metadata is well optimized. Your title can be up to 100 characters long. Your description can include up to 5,000 characters. Your tags can be up to 120 characters. Use every one of these limits for each of your videos.

Title

Make your title compelling—this is your video's headline. If it showed up in a search, would you click on it? Other guidelines include:

- Always represent your content accurately.
- Offer keywords first, branding at the end.
- For serial content, add the episode number to the end of the title.
- Update video titles so they continue to grab views.

NOTE

Avoid titles that trick viewers into clicking on the video. This will cause drop-offs in the first few seconds of your video and will negatively impact your video's watch time.

Description

Only the first few sentences of your description will appear in search results or above the fold on a watch page, so make them count! Follow a template for all of your video descriptions to create consistency.

A video description should:

- Accurately describe your video's actual content (transcript excerpts are ideal) in one or two sentences.
- Describe your channel and link to your channel page.
- Drive viewers to subscribe (so include a subscribe link).
- Link to other episodes or related videos and playlists.

NOTE

While you may want to link offsite in your videos, keep in mind that this could affect your video's watch time. Videos with lower watch times appear lower in search results. YouTube's new "metadata defaults" feature allows you to create templates for your metadata and ensure important text or links are always included when you upload a video.

A video description may:

- Include your channel's release schedule.
- Include links to time-codes in the video for long-form content.
- Include a recurring keyword tagline. The keyword tagline is a group of sentences that describe your channel. They should include several search-driven keywords. Repeating this tagline in episode descriptions will inform first-time viewers about your channel.

NOTE

Remember that it is a violation of YouTube's Terms of Service to use misleading metadata on your videos.

Tags

Tags are descriptive keywords that will help people find your videos. Create a set of standard tags for your channel that can be applied to any video you publish (e.g., filmmaking, animation, comedy, "Funny Videos," "Pet Videos," etc.):

- Include a mix of both general and specific tags.
- Use enough tags to thoroughly and accurately describe the video.
- Update catalogue videos' tags when new search trends emerge.
- Properly format tags to ensure appropriate indexing of your video.
- Include keywords from your title in your video's tags.
- List tags in order of relevance to the video and try to use the whole 120-character limit.

NOTE

Update the metadata on your older videos if their titles, tags, or descriptions are not optimized. This can increase views even if a video's been public for a long time.

Create custom thumbnails

Thumbnails show up in different sizes and formats all across the platform and beyond. Make sure you've got a strong, vibrant image that pops no matter what size it is.

General guidelines:

- When shooting a video, take shots that will make great thumbnails.
- Always upload custom thumbnails with the video file.
- Make sure the thumbnail is not racy.
- Consider the legibility of your thumbnail at multiple sizes. Thumbnails change size depending on the YouTube placement and device.

Visual guidelines:

- Use visual cues (colors, images, shapes, personalities) that are consistent with your brand.
- Thumbnails should be clear, in focus, and high-resolution (640px × 360px minimum, 16:9 aspect ratio).
- Use bright, high-contrast images.
- Include close-ups of faces.
- Use visually compelling imagery.
- Make sure your thumbnail is well framed with good composition.
- Ensure the foreground stands out from background.
- Check that the thumbnail looks great at both small and large sizes.

- Be sure the image accurately represents the content.

> **NOTE**
>
> Upload high-resolution thumbnails so they appear crisp and clear wherever viewers happen to see them.

Optimize your annotations

Annotations are clickable overlays that you can add to your YouTube videos. You can use annotations in a huge variety of ways; brands and partners are always thinking of new and innovative uses.

Annotations are a great way to encourage your viewers to engage with your video and take meaningful actions as a result. The key is to create an annotations experience that is conversational and interactive.

Use annotation in these cases:

- Subscribe direct link. Annotate the “subscription confirmation” page to make it easy for viewers to subscribe to your channel.
- Navigation. Highlight your newest video, create a table of contents for long videos, or prompt users to enter a sequential playlist viewing experience (e.g., “Click here for the next video in this series”).
- Calls to action. Drive engagement by inviting viewers to like, comment, or respond to questions. Complement scripted calls to action in the video with textual annotations.

Annotation best practices:

- Avoid annotations along the very top of the frame. This is where your title will show if it’s embedded.
- Don’t obstruct the actual content.
- Don’t bombard the viewer. This can feel “spammy.”
- When appropriate, set annotations to open a new window when clicked. Be

careful! Don’t take viewers away from a video too soon.

- Annotations at the end of a video should open in the same window.

NOTE

Measure the effects of your annotations in YouTube Analytics with the “Annotations” report.

Spotlight annotations

Most annotation types are self-explanatory. Spotlight annotations stand out because they allow creators to subtly create clickable areas within a video. The text appears only when a viewer mouses over it; a light outline shows otherwise. This is a great way to include unobtrusive but clickable annotations.

InVideo Programming annotations

Unlike regular annotations, InVideo Programming allows you to promote both your channel and any one of your videos on YouTube across all of your uploads. Here’s how InVideo Programming works:

- When promoting videos, it pulls in the thumbnail as the annotation. Videos with optimized thumbnails perform better.
- It allows you to upload a custom, transparent, square image to promote your channel. Unsubscribed viewers who mouse over the channel image can subscribe to your channel without interrupting the video.
- InVideo Programming annotations are viewable on mobile devices.

Optimize your captions

Providing captions makes your video content accessible to a wider audience. It also acts as additional metadata, which helps your video show up in more places on the site. If your video is captioned for multiple languages, it will also be searchable in those languages.

Create and optimize playlists

Creating playlists allows you to collect, organize, and publish multiple videos

together. This increases watch time, encourages user engagement with your channel, and creates another asset that will appear in search results and in Suggested Videos. You can create playlists using your own videos, other videos, or a combination of both.

Use playlists as follows:

- Group a set of videos that you want viewers to enjoy in a single session or in a particular order.
- Organize videos around a theme or a tent-pole event.
- Separate multiple shows into playlists and feature them on your channel.
- Combine your most-viewed videos with new uploads.
- Curate good brand-advocating videos (reviews, testimonials, hygiene, etc.) created by your community.

Playlist best practices:

- Choose a strong thumbnail for your playlist. Make it pop!
- If a playlist needs context, upload a short, snappy intro video or interstitial videos with a host. Create a hosted playlist.
- Make your metadata work for you. A strong title, tags, and description will help people find your playlist.
- Use Playlist Notes to write conversational asides about individual videos.
- Use invideo messaging, annotations, end cards, and links to send viewers to a playlist.
- Feature your playlist on your channel page by creating a new “section.”

NOTE

To link to a video in Autoplay playlist mode, click the Share button and then copy and paste that URL. The video link will launch the whole playlist.

Optimize your channel

Your channel is the face of your brand on YouTube. It allows you to collect and organize all your videos in one place. Several customizable channel features will help ensure that you're delivering the best experience for your viewers while making your brand more discoverable across YouTube.

A number of channel features will follow your videos across the YouTube site and on devices—make sure that they effectively represent your brand personality:

Channel name

Pick a short, memorable channel name in line with your brand identity. This will appear widely across YouTube, so make sure it's the best representation of your brand.

Channel icon

Upload a square, high-resolution image to be your channel's icon across YouTube. The image will appear alongside all your videos on the watch page. (In most cases, you can use your brand's logo.)

Channel description

- The first few words of your channel description appear most frequently across the site, so highlight your most important branding up front.
- Include your upload schedule, especially if you host multiple content types or series.

Channel art

- Channel art is your channel's primary branding across all devices. Create customized, visually compelling channel art. Busy images don't scale well, so keep the image simple.
- Add website and social media links to the About tab. Include these links in your channel art to help tie your YouTube presence to the rest of your

online brand.

- Check how your channel looks in search, related channels, and the channel browse page. Do your channel icon, channel name, and channel art do a good job representing your brand to potential fans?

NOTE

Associate your official website with your YouTube channel.

Optimizing for subscribed fans

In the Browse view of the channel page, subscribed and unsubscribed viewers see different versions of your channel.

What to watch next

- Subscribed viewers see personalized recommendations of “What to Watch Next” based on their viewing history.
- Promote a video with InVideo Programming to make it appear first in the “What to Watch Next” recommendations.

Recent activity

- Your most recent feed posts will appear on your channel page in the recent activity feed and on the activity tab.
- Keep your feed active with uploads, likes, and channel posts. This will give your fans another reason to come back regularly.

NOTE

Your top section will show up most frequently across devices. Make sure it will draw in new viewers.

Optimizing for unsubscribed viewers

The unsubscribed view is your first opportunity to convince potential fans to subscribe.

Channel trailer

- Enable the channel trailer for unsubscribed viewers. This video will autoplay, so tell new visitors what your channel's all about and why they should subscribe.
- Keep your trailer short, but make sure you ask the fans to subscribe!
- Show, don't tell. Give potential fans a taste of your best content.

Sections

- Sections organize videos, playlists, and channels on your channel's Browse page. Subscribed and unsubscribed viewers will see them.
- Sections can divide your content by genre, theme, show, or any other criteria. Effective organization will help your audience find the content most relevant to them.

Gain subscribers by promoting your channel

Once you've optimized your channel for new viewers, it's time to use YouTube tools to bring them there.

InVideo Programming – Channel Promotion

- Enable Channel Promotion in InVideo Programming to place your channel icon across your entire video library.

Annotations and calls to action

- Use end cards or other clickable annotations to push viewers to your channel page.
- Vocal calls to action along with annotations or video description links can help increase click-through rates.

Linking with Google+

- Link your YouTube channel to your Google+ page to amplify the social reach of your videos, increase discoverability of your content, and offer a more streamlined Google experience. Doing this also enables new features for easier management of your YouTube channel.

Related Channels

- Related Channels are promoted channels populated by YouTube that appear on your channel page. Recommendations are based on channels that are similar to yours.

- Be sure to enable Related Channels. Disabling the feature will prevent your channel from being promoted on other channels.

Featured Channels

- Featured Channels are channels you choose to promote on your channel page. For instance, highlight other channels in your brand's portfolio.

- If you are promoting a large number of channels, rotate through the list using the "shuffle" feature. This ensures all your channels get visibility.

NOTE

Make sure your icon doesn't distract from the videos or overlap with existing annotations.

Optimize watch time

As mentioned earlier in this chapter, watch time is an important metric for video SEO on YouTube. The algorithm for suggesting videos prioritizes those that lead to a longer overall viewing session over those that receive more clicks. Viewers benefit from more enjoyable content being suggested to them, and creators benefit from more focused, engaged audiences.

If you're making videos that people are watching well beyond the first click, those videos will be suggested more often. Here are some guidelines:

- Be an effective editor. Create a compelling opening to your videos and then

use programming, branding, and packaging techniques to maintain and build interest throughout the video.

- Build your subscriber base. Subscribers are your most loyal fans and will be notified of new videos and playlists to watch.
- Involve your audience in your videos, encourage comments, and interact with your viewers.
- Build long watch-time sessions for your content by organizing and featuring content on your channel.
- Create a regular release schedule for your videos when uploading to encourage viewers to watch sets of videos over single videos.

Check YouTube Analytics

Measurement is key, both for defining success and optimizing toward it. YouTube Analytics enables you to track a number of your most important metrics. For example, the Traffic Sources Report shows you which search terms people are using to find your video. The Views Report identifies which videos have the greatest view times and view-through rates. And the Audience Retention Report indicates which videos are successful at keeping viewers watching.

The key is to decide what success looks like and then translate that into metrics that are relevant for *your* brand. Most brands find it useful to select one metric for each of the three buckets that measure engagement with your content:

Audience

Are you reaching the right audience? How well?

Expression

Is your target audience engaging with your content? How much?

Participation

Is your audience endorsing and sharing your content? How much?

NOTE

The YouTube Analytics API can be configured to schedule regular reports on specific metrics.

Leverage paid advertising

While this book is focused on SEO, video SEO can leverage paid advertising in a way that website SEO can't. In fact, YouTube enables you to ignite earned media with paid media.

Given the abundance of videos on the Web, it's risky to assume that your content will be organically discovered by a large audience. That is why it's vital to design a solid plan to promote your content and ensure it's viewed by your target audience.

"Going viral" plays a key role in building your audience on YouTube, but unless you already have millions of subscribers, you'll need to seed your content when it launches. Indeed, when they don't have an existing subscriber base or engaged social following, many brands use paid advertising on YouTube to ignite sharing and accelerate audience building.

The YouTube platform allows you to optimize your investment in paid media by activating the virtuous circle that links *owned media* (your video content), *paid media* (paid video advertising), and *earned media* ("free" views obtained when people share the video ads).

This provides you with a major benefit from advertising on YouTube that you don't get from Google: earned impact. According to YouTube, more than 6,000 campaigns generated at least one earned view as a result of every two paid views in the first half of 2014.

And the YouTube audience isn't shy about sharing brand content: two-thirds of YouTube users agree that "if there is a brand I love, I tend to tell everyone about it." Across the globe, they are twice as likely to be early adopters, agreeing that "I am among the first of my friends and colleagues to try new products." They are 1.8 times more likely to be influencers, agreeing that "people often come to me for advice before making a purchase."

Indeed, hundreds of campaigns get more than two earned views per paid view. Although you don't get that benefit with Google AdWords, you can get that

benefit with Google AdWords for video, which powers YouTube TrueView video ads.

Video SEO for Google

Once upon a time, one of the best ways to improve your site's appearance in video search results was to make sure that Google knew about all of your rich video content. So it made sense to submit a sitemap to Google that included video information in a supported format, making the included video URLs searchable on Google Video.

Google Video was launched on January 25, 2005, after which Google bought former competitor YouTube on October 9, 2006. On June 13, 2007, Google announced that Google Video search results would begin to include videos discovered by its web crawlers on YouTube, on other hosting services, and from user uploads. On January 14, 2009, the Official Google Video Blog announced, "In a few months, we will discontinue support for uploads to Google Video." On April 15, 2011, Google announced via email that after April 29 it would no longer allow playback of content hosted on Google Video. However, it reversed that decision one week later to provide users with greater support for migration to YouTube. The service, which was renamed Google Videos, was shut down on August 20, 2012.

The remaining Google Videos content was automatically moved to YouTube, so for all intents and purposes, Google Video has been replaced by YouTube.

Embed YouTube videos and playlists

Today, one of the best ways to improve the ranking of your video content in Google Universal Search results is to upload it to YouTube—though there is still debate in the industry as to whether uploading your video to your own site first, for origination credit, provides longer-term site benefits.

Whichever approach you decide upon, these are instructions for uploading to YouTube:

Embed a video:

1. Click the Share link under the video.

2. Click the Embed link.
3. Copy the code provided in the expanded box.
4. Paste the code into your blog or website.

Embed a playlist:

1. Click Playlists on the left side of your YouTube page.
2. Click the playlist title you'd like to embed.
3. Click the Share link.
4. Click the Embed link.
5. Copy the embed code that appears in the menu below.
6. Paste the code into your website or blog.

Use a video distribution service

Searchmetrics' analysis of Google Universal Search results, cited earlier in this chapter, indicates there's another way to improve the ranking of your video content in Google Universal Search results: upload it to Vimeo or Dailymotion. Uploading your video content to YouTube makes it 10 to 11 times more likely that it will appear in Google Universal Search results—about two positions ahead of the average first position of a video from Vimeo or Dailymotion. However, it makes much more sense if you upload your video content to YouTube as well as Vimeo, Dailymotion, and other top video and social networking sites. To do this efficiently and effectively, consider using a video distribution service, such as OneLoad. OneLoad is powered by its parent company, TubeMogul, a media buying platform for video advertising founded in 2006. You can leverage its Destination feature to easily distribute to custom sites as well as encode and create RSS feeds to syndicate your video anywhere. And videos distributed through OneLoad are tracked by TubeMogul's powerful analytics. OneLoad is the video distribution tool of choice for some of the world's largest brands, agencies, media companies, government institutions, and nonprofits. These organizations use this specialized service to get their videos

watched by more people in more places.

Use a Google video sitemap

One of the last video SEO strategies to consider using is a Google video sitemap. Creating a video sitemap is a way to make sure that Google knows about all the video content on your site, especially content that it might not otherwise discover via its usual crawling mechanisms.

However, as the analysis by Searchmetrics indicates, about 1% of the videos in Google Universal Search results come from websites like MUZU TV, Howcast, Artistdirect, and eHow. In other words, creating a video sitemap may let Google know about your rich video content, but your videos still need to get a significant number of views or watch time to appear in Google Universal Search results.

When a user finds your video through Google, she will be linked to your hosted environments for the full playback. Search results will contain a thumbnail image (provided by you or autogenerated by Google) of your video content, as well as information (such as title) contained in your sitemap. It's also worth noting that Google can't (or won't) predict or guarantee when or if your videos will be added to its index.

You can create a separate sitemap listing your video content, or you can add information about your video content to an existing sitemap—whichever is more convenient for you.

The Google video extension of the Sitemaps protocol enables you to give Google descriptive information—such as a video's title, description, duration, and so on—that makes it easier for users to find a particular piece of content. Google may use text available on your video's page rather than the text you supply in the sitemap's video content, if this differs.

Adding video information to a sitemap

You can create a video sitemap based on the Sitemaps protocol. You can also use an existing mRSS feed as a sitemap. Or, if you're the belt-and-suspenders type, you can use both.

Your sitemap will need to include the following minimum information for each video: title, description, playpage URL, thumbnail URL, and the raw video URL

or URL to Flash video player. Without these five pieces of information, Google cannot surface your videos in its search results.

Once you've created your sitemap, you can submit it using Search Console.

Use recognized file types

Google can crawl Flash SWF objects and the following raw video file types: *.mpg*, *.mpeg*, *.mp4*, *.mov*, *.wmv*, *.asf*, *.avi*, *.ra*, *.ram*, *.rm*, and *.flv*. All files must be accessible via HTTP. Metafiles that require a download of the source via streaming protocols are not supported at this time.

Conclusion

As we've outlined in this chapter, optimizing your site for mobile search and your vertical content for Universal Search, engaging in local SEO efforts, and if you're a news publisher, participating in Google News can provide significant exposure within organic search for your business. Not only can you gain visibility in the increasingly diverse search ecosystems, but you can also increase the likelihood that your content assets are working for you in all areas of the Web to promote increased user sharing and engagement—which provides a positive feedback loop for all of your SEO efforts.

[1] comScore, "comScore Reports May 2014 U.S. Smartphone Subscriber Market Share," July 3, 2014, *http://bit.ly/comscore_may_2014*

CHAPTER ELEVEN

Tracking Results and Measuring Success

Because the nature of what SEO practitioners do is not clear to many people, the expert practitioner knows how important it is to establish quantifiable goals and demonstrate results. Expectation management is key, and effective communication of timetables for deliverables and improvements is crucial to the success of any SEO project.

Web analytics and other data-mining tools can provide a rich array of data that is invaluable to the SEO process. These include tools for discovering new links, finding long-tail search terms, determining the best links your competitors' sites have, and more. All in all, a strong understanding of these tools will make you a better SEO practitioner.

However, thousands of tools are available, and it can be difficult to figure out which ones you need to use, which ones are merely nice to have, and which ones are not worth your time. It is a good idea to start simple and focus on establishing a baseline measurement of where you are before you begin your SEO efforts so you can better measure your results. A proper baseline should account for as many variables as are applicable to your business and industry—organic sales, indexing, crawl rates, conversion rate analysis, and so on.

In other words, you need to measure whether, and to what extent, you have been successful. The first thing you need to do is define the goal(s) of your site. Here are some possibilities:

- Generating sales
- Obtaining leads
- Generating ad revenue
- Getting people to read specific articles
- Obtaining donations
- Developing an audience (which can lead to the above)

These are all valid objectives that relate directly to the business. Missing from

this list are practices such as measuring rankings, measuring the total number of links, and measuring the average cost per acquired link (organically obtained links still have a cost: the cost of the marketing campaign to get them). These other measurements may be interesting, but they are not the end goal.

Closely connected with the site goals are direct drivers of results. These include parameters such as:

- Total site traffic (all visitors)
- New versus returning visitors
- Traffic to your most popular pages
- Pages receiving the search traffic
- Pages resulting in conversions down the line
- Traffic per page and conversion per page

Measuring results in terms of the main business goal should always be the first set of metrics you put in place. From there you can diverge and look at other metrics that might help you diagnose problems with the site or give you additional insight into how to set up a campaign.

Why Measuring Success Is Essential to the SEO Process

Although quantifying deliverables and measuring progress are important for external reporting purposes, it is just as important for SEO practitioners to measure the efficacy of their own efforts to make timely adjustments as necessary. As you will see in this chapter, numerous tools are available to aid you in this process.

At the beginning of any SEO project, it is wise to establish baseline data points for the website. This includes the following:

- Quantifying organic search traffic by search engine

- Quantifying a breakout of what site content areas are getting the current organic search traffic by search engine
- Quantifying data on conversions broken down by search engine
- Identifying poorly performing pages
- Identifying best-performing pages (in terms of traffic and conversions)
- Tracking search engine crawler activity on the site
- Determining the number of indexed pages
- Determining whether the indexed pages are getting search traffic
- Determining whether best-selling product pages are indexed and getting search traffic
- Identifying 404 error pages and external sites linking to these pages, if any
- Determining whether you've been hit by a search engine penalty or algorithm update

You cannot methodically improve what you cannot measure, or as Peter Drucker put it: "What gets measured gets managed." Defining and mapping the path toward concrete goals are crucial aspects of the SEO process—and over time, these goals may change. Thus, it is also important to make sure the data you capture helps you understand your progress toward these goals.

In the world of web analytics, this is referred to as picking *actionable key performance indicators* (or actionable KPIs). The best data measurements are those that potentially result in an action being taken in response. Think of this as data-driven decision making.

It is important to recognize that SEO requires operating effectively within two distinct worlds. In the first, SEO-related metrics such as rank, visits, and backlinks rule the day; in the second—the one likely occupied by the rest of the organization—business metrics such as sales, costs, returns, and contract renewals are the focus. As an SEO practitioner, you should address both spaces

by clearly tying SEO metrics to business goals.

The Tracking Cycle: Produce, Launch, Measure, Refine

In summary, the basic tracking and measurement process usually looks something like this:

1. **Define an SEO strategy, determine an implementation schedule, and establish a clear understanding of goals.** What are you going to accomplish, and what is the strategy for accomplishing it? How will you measure progress?

2. **Discuss your strategy.** The marketing and business development teams are your allies here—you want to ensure that your SEO objectives are based on the overall business and site objectives, both long-and short-term.

3. **Establish a baseline.** Now that you are about to start and you have decided how you are going to measure progress, establish a baseline by recording the current stats prior to beginning work. Make sure you don't get a false baseline due to seasonal factors or some other unusual event. Comparing year-over-year data will usually eliminate fluctuation due to seasonality. However, you must also consider how changes in the market, new competition, elimination of competition, industry consolidation, changes in your business strategy, and changes in the search engines themselves may have affected that year-over-year data.

4. **Proceed with your project.** Implement the SEO strategy—new URLs, technical site changes, new content development, and so on.

5. **Collect data.** Collect the newest data for each metric you decided to focus on. SEO can take days, weeks, or months to show results, so make sure you wait long enough for your efforts to have an impact. Many factors could influence the length of time you should wait, including:

   - If your site is brand new, it may take longer for your changes to take effect.

   - If the scope of the change is drastic (such as a complete redesign

incorporating new information architecture, new URLs, etc., versus a simple visual reskinning), it will probably take longer to see results.

- Sites that get crawled at great depth and frequency will probably yield visible results faster.
- Sites seen as authoritative may also show faster results.

6. **Compare the baseline data to the new data.** The new data has little meaning unless it is compared to your baseline. This is the time when you can really assess your progress.
7. **Refine your campaign.** Now that you have compared your old data with your new data, you can make some decisions. Is the campaign a bust? If so, abandon it and move on to the next one. The old business axiom "Fail quickly" applies here. The faster you diagnose a failure and move on to the next thing, the better.

If you are achieving great results—that is, a noticeable increase in traffic and/or conversions—then you've got a great process in place; apply it to more pages on your site, or look for ways to further improve conversions through content development and design. If you don't see any change in your traffic or conversion rate (or if the changes are measurably negative) within a few weeks, don't give up just yet. Perhaps you've overlooked something important, or there was an error in your implementation. Or perhaps you are recovering from a larger site overhaul. In this case, things may get worse before they get better. Be patient, and at this point of the process, focus on action-based goals instead of results-based goals. For very damaged or very large sites, the wait for improvement could be six months to a year.

Establishing a Proper Baseline

It might sound easy to record a set of SEO metrics before the start of the project and then compare the same set of metrics after 30, 60, or 90 days to measure the progress. But what if you don't make any changes and the metrics reflect an improvement? What if the improvements are because of external business factors? How can your SEO project get attribution for its business impact?

Depending on the type of business, and assuming there have not been drastic strategic shifts from the previous year, it is possible to account for the seasonality of the business if you compare each period to the corresponding period of the previous year. Then adjust for overall year increase (or decrease). You can measure by comparing a couple of months year to year where you have not made any SEO changes. The assumption is that everything else being equal, a performance increase in search should be attributed to your SEO efforts.

It is important to make sure that there haven't been any major strategic shifts in the business. Sometimes companies completely change the way they do business, and this dramatically impacts your ability to measure results. In cases like this, you can fall back to the less reliable approach of tracking the increase in pages getting search clicks and overall increase in average ranking across the board.

One other issue that may significantly impact your ability to establish a baseline for measurement is a seemingly spontaneous drop in traffic. When analyzing the organic traffic to a site, if you notice a large drop it's vital to determine its cause before proceeding with creating a set of baselines and goals. Large traffic drops can be caused by a number of factors, including a large-scale site redesign or rebuild, a shift in the nature of the business as just noted, seasonal factors (which you can usually determine by looking at several years of data at a time), or organic search algorithm updates.

An extremely useful tool you can use in determining if a large traffic shift might be caused by an algorithm update is the Panguin (a concatenation of the words *panda* and *penguin*) tool created by Barracuda Digital. This tool allows you to overlay your Google Analytics organic traffic with Moz's history of algorithm updates to see if traffic shifts coincide with the updates. It's not always possible to definitively determine if you have been impacted by an algorithm update, as occasionally several updates happen over a very short period of time or you may have a traffic drop caused by another factor that happens to coincide roughly with the algorithm update. However, before you embark on an SEO project, it's important to be aware of the possibility that an algorithm update may impact the potential success of the project.

It's also important to monitor rank in Google to understand what is working in your strategy. For example, a jump from 99 to 15 on a targeted keyword might

not bring you a substantial increase in traffic; however, it is a huge indicator that something you are doing is working. AuthorityLabs, a subscription SEO ranking tracker, offers a one-stop page for tracking all of the keywords your business is targeting as well as their shifts over time.

Using Analytics as a Business Case for SEO

You can use a properly structured plan as the business case for an SEO project. The way to do this is to express the target results of an SEO project in terms of financial impact. You could include a variety of metrics in a business case, such as:

- Revenue
- Lead generation
- Margin
- Branding value
- Reach
- Other action triggers (newsletter sign-ups, contact requests, demo requests, free-trial acceptance, viewing and/or sharing a specific piece of content, etc.)

Measuring such factors requires that you tie organic search engine visits to the revenue and other conversions that result.

Measuring Search Traffic

Classic web analytics data is an incredible asset to SEO. Here are three examples of ways to utilize this data for SEO purposes:

- Look at your daily referral reports to detect newly received inbound links (the great majority of webmasters click on a link after implementing it, to make sure it works).
- Measure the results of your campaigns by tracking the increase in

conversions you are driving over time.

- Try to find missed opportunities—search terms and phrases that people used that did not result in a click, visit, or conversion.

Web analytics are a must-have for any web publisher.

Basic Overview

Your hosting company most likely provides a free web analytics solution, such as AWStats, Webalizer, or something similar. Although these tools provide valuable data, they are very limited in scope, and other tools out there provide significantly more data. Here are some of the best-known ones:

- Google Analytics
- Woopra
- Clicky
- Piwik
- Crazy Egg
- Adobe Analytics (enterprise-level solution)
- IBM Digital Analytics (enterprise-level solution)
- Webtrends (enterprise-level solution)

Web analytics platforms track your site's traffic in two major ways. The older of the two methodologies is to use software that analyzes your web server logfiles after traffic activity. Setting up this type of solution generally involves installing the software on an internal server that is able to gain access to the logfiles.

The newer methodology involves implementing JavaScript tags on all of the pages of your website. Provided that you have a basic template for your pages, this generally is a fairly straightforward process. JavaScript then tracks the activity on your web pages as it happens and builds a data profile reflecting that

activity.

Selecting the Right Analytics Package

Logfile tracking and JavaScript tracking are equally valid methods, and each has its own strengths and weaknesses. The biggest advantage of the logfile method is that you can track search engine crawler activity on your site. This is something you cannot do in JavaScript implementations, because search engine crawlers do not execute the JavaScript.

The second big advantage of a logfile-based solution is that you run the software in-house, so no third party has a copy of a logfile with your proprietary traffic data on it. This distinction can be a big security issue for some organizations.

Another key advantage to logfile analysis is the ability to track:

- Users who don’t have JavaScript enabled (or are using privacy or ad filters)
- Media types beyond HTML
- Partial requests (incomplete page loads and load errors)

Ultimately, though, most companies opt for JavaScript tracking because JavaScript offers a much greater level of flexibility than logfiles. You can tweak the JavaScript to do custom conversion tracking, or gather pages into logical groupings in a manner that cannot be done as easily in logfile-based applications.

Other key advantages of JavaScript tracking include the ability to:

- Track outgoing link requests
- Track events that don’t involve HTML requests, such as playing a video
- Record visitors’ screen resolution

Some analytics packages, such as Webtrends, offer both options or a combined solution. This kind of approach can bring you the flexibility and power of JavaScript, but still get you your search engine robot crawling data as well.

Making this decision is only the first step in picking an analytics package. We

listed some of the more popular vendors earlier, but there are many more. Each package has different strengths and weaknesses. Not only that, but they all do their counting a little bit differently. These differences stem from how the analytics packages conduct visitor tracking. None of them is right or wrong; they just differ in the exact thing they are measuring. For instance, one package might count a “visitor session” within a certain timeframe.

The more important factor in this decision is whether the functionality of the web analytics software fits your needs. You often do not know what your requirements will be until you have used analytics for a while. As you engage with analytics, you will continually uncover more things to investigate, and develop new requirements.

For many companies, one of the best solutions is to start with a free analytics package such as Google Analytics, then look to buy a higher-end solution once you’ve exhausted the capabilities of the free package. By pushing the limits of these free analytics products first, you will end up developing a set of requirements you can use in deciding where to go next.

None of this is meant to imply that you should not brainstorm your requirements in detail before selecting an analytics package. You should (and must). Just expect that you will develop new requirements along the way. Web analytics is a journey that unfolds over time.

Based on the requirements you establish in your upfront brainstorming, you may find that you require a set of features that the free packages do not provide. Use that knowledge to select the right package to start with.

Extracting Valuable SEO Data in Web Analytics

You can extract all kinds of data from web analytics. Next, we’ll look at a few of the more interesting types of information you may want to extract.

Traffic by search engine

One of the first things you may want to know is the breakdown of traffic by search engine. Figure 11-1 provides an example of such a report in Google Analytics.

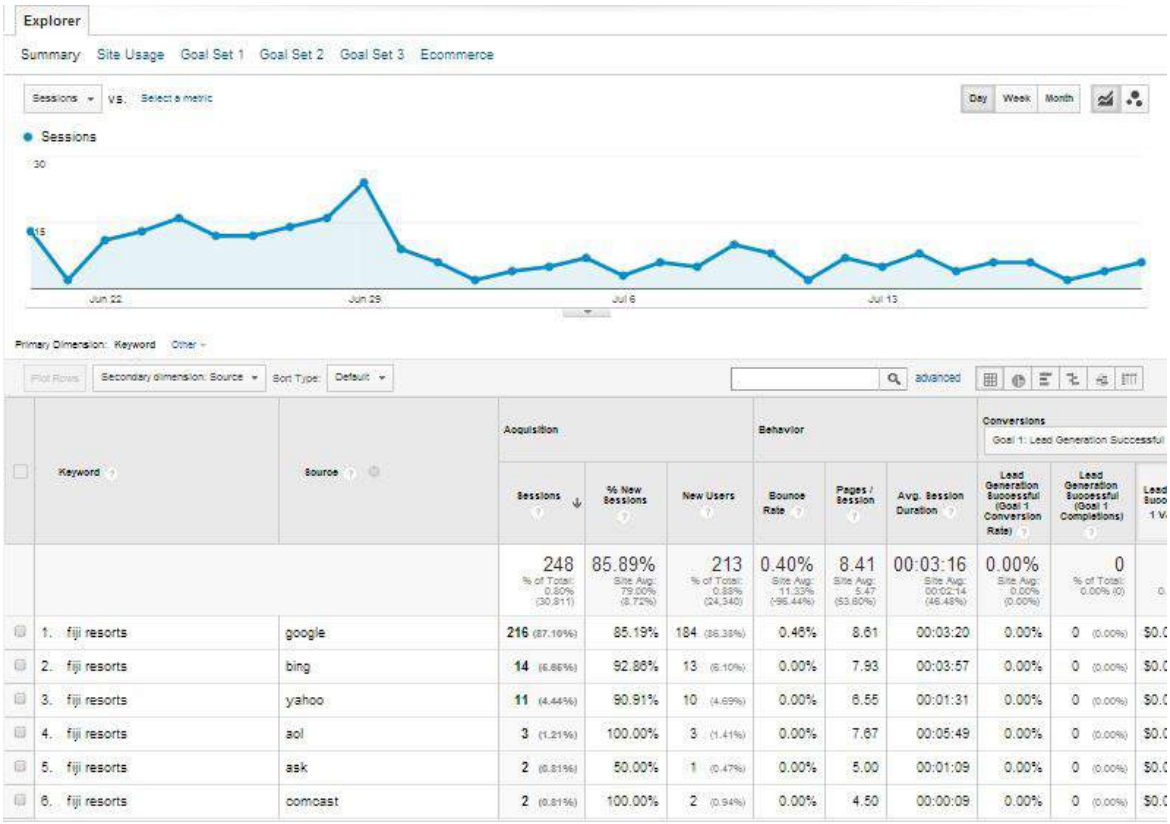
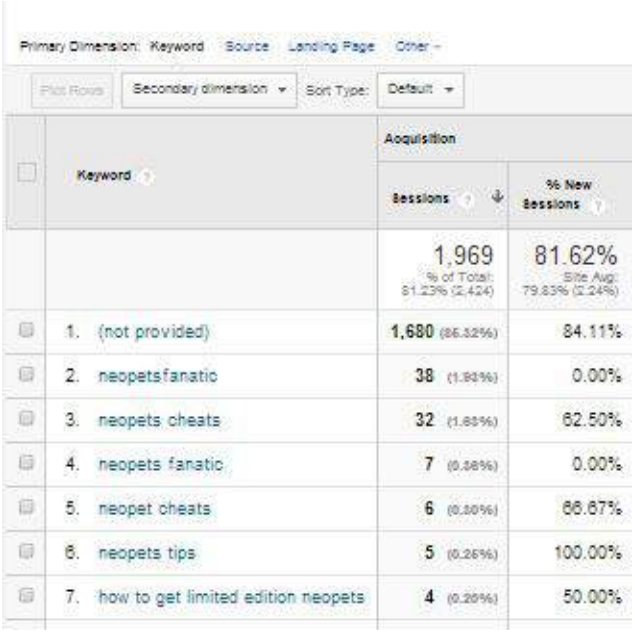


Figure 11-1. Traffic by search engine

Notice how small the traffic is from Bing and Yahoo! compared to Google. This may be indicative of a problem with how the site is indexed by Bing (which supplies the search results to Yahoo!); the webmaster, designer, or developer might want to explore why the combined traffic from Bing and Yahoo! is so low. Or, this result could just stem from the fact that Google by far has a dominant market share—67.3%, according to a comScore report from August 2014. However, in your specific niche, that percentage may be far higher or lower, which is why browser choice can be such an important demographic data set.

Traffic by keyword

Beginning in October 2011, Google began “securing” its search traffic, hiding most keyword data for organic search. As of 2014, most search terms reported by Google Analytics appear as “not provided” (see Figure 11-2).



Primary Dimension: Keyword Source Landing Page Other

Plot Rows Secondary dimension Sort Type: Default

	Keyword	Acquisition	
		Sessions	% New Sessions
		1,969 % of Total: 81.23% (2,424)	81.62% Site Avg: 79.83% (2.24%)
1.	(not provided)	1,680 (86.32%)	84.11%
2.	neopetsfanatic	38 (1.93%)	0.00%
3.	neopets cheats	32 (1.63%)	62.50%
4.	neopets fanatic	7 (0.36%)	0.00%
5.	neopet cheats	6 (0.30%)	66.67%
6.	neopets tips	5 (0.26%)	100.00%
7.	how to get limited edition neopets	4 (0.20%)	50.00%

Figure 11-2. Keyword not provided

There are several ways to attempt to extract more detailed keyword data, though none of them are as detailed as when Google used to provide you all of the data. One solution is to use keyword data from your Google Search Console (see Figure 11-3). You can now even access your Google Search Console data directly from Google Analytics (*http://bit.ly/search_console_data*). Google provides you with some estimates of the number of impressions, number of clicks, your click-through rate, and your average ranking position for a variety of your keywords. Note, however, with Search Console, Google will not provide you with all of your search queries (though it will provide several thousand) and the data is stored for only 90 days (so get in the habit of downloading an export file every 90 days to maintain your archive). This information is found in your Google Search Console interface under Search Traffic – Search Queries. In addition to getting traffic estimates for web searches, you can segment the information by web/mobile/image, by geographical location, or for only specific queries.

Note that by default, filters are different on Google Search Console and Google Analytics. If you're noticing differences in query results, try syncing up the filters.[1]

Query	Impressions	Clicks ▲	CTR	Avg. position
costa rica weather	12,946	1,544	12%	3.7
costa rica map	28,014	895	3%	4.3
guanacaste costa rica	13,319	821	6%	1.7
costa rica rainy season	1,058	673	64%	1.1
limon costa rica	6,884	393	6%	1.9
rainy season costa rica	464	346	75%	1.1
weather in costa rica	3,711	284	8%	4.7
map of costa rica	15,943	280	2%	6.0
rainy season in costa rica	342	250	73%	1.0
tourist attractions in costa rica	1,547	246	16%	1.4

Figure 11-3. Google Search Console keyword data

If you participate in paid Google AdWords advertising, you can also look at your Matched Search Queries report in Google Analytics to see the exact terms users searched on before clicking on one of your ads (*http://bit.ly/adwords_search*). This may help give you some idea of the types of terms people are using to get to your website. In addition to Google Analytics, Bing Search Console also provides some keyword data.

Although it is a paid tool, SEMrush can also provide some insights into what keywords your site ranks for and approximately how much traffic you might be receiving from each of them.

Another paid solution for obtaining insight into keyword data is Searchmetrics' Traffic Insight solution. Traffic Insight allows you to reconstruct "not provided" keywords and traffic data on a page-by-page basis by incorporating various data sources and running them through a proprietary algorithm to allow you to "reverse engineer" much of your "not provided" data.

One of the data points of interest for an SEO practitioner is what search terms are bringing traffic to the website. This provides a quick way to see where the SEO campaign is going well and where it is not. You can also use this data to spot opportunities where a key search term is providing some traffic, but not as much as you would expect if you were ranking highly for that term. Remember that search terms are no longer simply considered in isolation. If you spot an

opportunity to increase traffic on a key search term, you should really think about what related terms might also drive traffic to the same page as the keyword you are considering.

You can then look to see where you are ranking for that term. Perhaps you are in a lower position on the first page, or on the second page of the SERPs. If so, it might make sense to focus some attention on this term. With a little effort, such as launching a focused promotional campaign for the page in question, and improving and expanding upon the information you provide related to that term, you may be able to move up several positions and obtain a traffic boost. By improving the content on the page to better satisfy the user's query, you may also begin ranking on related terms.

A traffic-by-keyword report can also show you the long tail of search as it relates to your current site. Figure 11-4 depicts a snippet from the organic search phrases report from Google Search Console showing some of the lower-volume terms for ArtofSEObook.com.

seo book	2,158	4
the art of seo book	33	3
the art of seo 2nd edition	24	3
best seo books	159	2
seo art	40	2
art of seo book	57	1
the art of seo free download	10	1
art of	1	1
seo books	1,449	0
seobook	320	0
best seo book	170	0
middle east search engines	78	0
seo book.com	65	0

Figure 11-4. Long-tail keywords

Notice that a lot of variations of “seo book” are showing up in the list, in addition to the rather interesting *Middle East search engines*. Looking into this query data can present you with opportunities for delivering new content that addresses the intent behind these queries.

Branded versus nonbranded traffic

Branded search traffic refers to queries that involve your company name, domain name, or product brand name. *Nonbranded search traffic* can be much more valuable because it often represents new visitors to your site. Branded traffic will convert at a higher rate on average, but nonbranded search traffic is often from people who did not previously know your brand, or at least they were not yet committed to it to the point where they searched specifically for a variation on your brand name or domain name.

In the past, you could see the data on keywords driving traffic to your website in great detail, as the search engines provided that data and you could readily examine it in your analytics software. Now it’s common that 90% or more of the keywords for your organic search traffic will be reported as “not provided,” meaning that it is difficult to get a strong measurement of what keywords are driving traffic to your site, or even the split between branded and nonbranded traffic.

Number of pages getting search traffic

An indirect way of measuring effective indexation is to keep an eye on how many pages are getting search traffic. This number should represent a subset of the total pages indexed, but it is more valuable because these pages are not just indexed, they also ranked highly and were interesting enough that visitors decided to click on the listing.

This is an important metric to track as you work on addressing site architecture issues like duplicate content and bot traps. As you remove search engine spider roadblocks, the bots will find and index more pages. Those pages will rank and get search traffic, and this number should increase. The same happens as you consolidate duplicate content and the indexation of the site improves.

There are several ways in which someone knowledgeable in Google Analytics can configure your account to give you at least some insight into which of your pages are receiving Google organic traffic where the keyword is “not provided” (which, as noted earlier, will be the great majority of them). From this data, you can extrapolate which groups of related keywords those landing pages are optimized for and which keywords they actually rank for, with the assumption that the “not provided” traffic is coming from those keywords. For some examples of these solutions, see Dan Parker’s “How to steal some ‘not provided’ data back from Google” or Moz’s “Easing the Pain of Keyword Not Provided: 5 Tactics for Reclaiming Your Data”.

Segmenting Search Traffic

Next, you can consider putting these metrics together. Even the free tools provide substantial capability for building out custom reports. Google Analytics has options that show the search phrases just for Google, and the pages of the site that Google sent users to when they clicked on your link.

This is a key improvement to your SEO research for two reasons:

- If you are looking into what terms can bring fast traffic benefits as a result of some additional optimization, you’ll want to know which search engines you are ranking in.
- If you are going to optimize a page to rank higher, you’ll need to make sure you’re optimizing the right page!

Referring Sites

Referring site reports are useful for a number of reasons, but one of their more interesting SEO applications is to spot when you receive new links. You can often see those new links in these reports first, even before the search engines report them. Figure 11-5 shows a sample portion of the referring sites report from Google Analytics.

If you notice a site for the first time in your list of referrers, it can be a leading indicator that you have received a new link. This report can help you detect new links that result from your linkbuilding campaigns, and therefore help you measure which of your link development campaigns are yielding the best results.

Source	Acquisition	
	Sessions	
	1,069 % of Total: 59.89% (1,785)	
101. biteclubeats.com	1	(0.09%)
102. blog-search.com	1	(0.09%)
103. blog.landerapp.com	1	(0.09%)
104. blog.myrss.jp	1	(0.09%)
105. bloggerbabes.com	1	(0.09%)
106. blogher.com	1	(0.09%)
107. bowhuntamerica.com	1	(0.09%)
108. boxingnewsonline.net	1	(0.09%)
109. bravoacademy.com	1	(0.09%)
110. businessblogconsulting.com	1	(0.09%)

Figure 11-5. Referring sites report

Using Custom Analytics Dashboards

In analytics terms, a *dashboard* is a single-page view that contains your most critical metrics all in one place. Of course, your most critical metrics are different from those of the next publisher, because the needs of different sites vary greatly.

In addition, multiple dashboards may be required in any given organization. For example, the CEO of a large public company probably wants to see different data (and a lot less of it) than a senior business analyst.

Each analytics package provides methods for implementing a custom dashboard. Figure 11-6 is an example of one from IBM Digital Analytics.

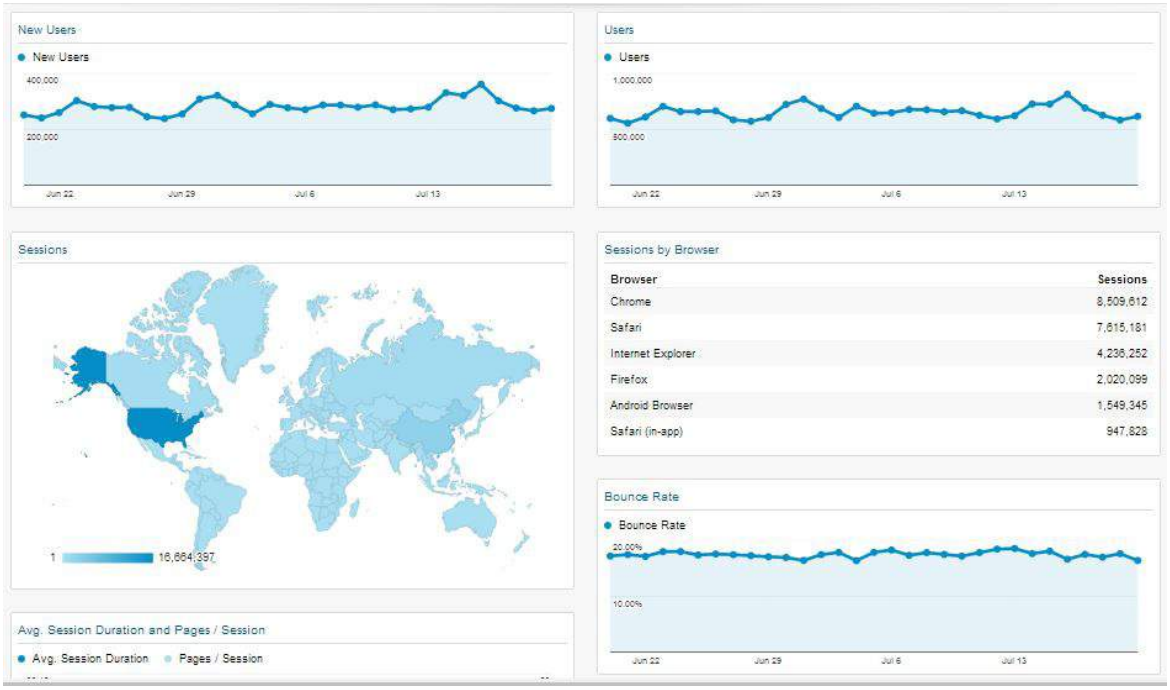


Figure 11-6. Custom dashboard

As you can see from Figure 11-6, a dashboard can be quite visual. What is most important, though, is that it provides the data that is most important to the person for whom it was designed. As an SEO practitioner, you can implement a dashboard to show your progress toward the goals you've set for your SEO

campaign.

Providing this type of visibility has two important benefits:

- The person viewing the report will appreciate the fact that she does not have to work hard to do a quick health check on the progress of the SEO efforts. As we suggested at the beginning of this chapter, the fact that you agree to measurable goals will be a great comfort to management.

- You will know what data your managers are looking at. When something goes wrong (or right), or when management wants to discuss some aspect of the business, they will have started from the dashboard you set up for them.

NOTE

A key consideration in selecting analytics and dashboard solutions is whether your web analytics solution has prebuilt integrations with your preferred SEO solution, or whether your SEO solution is part of a larger cross-media dashboard platform. Otherwise, you may end up with siloed information or quite different conversion metrics in one dashboard than in another.

Taking a Deeper Look at Action Tracking

Action tracking goes one step deeper than basic analytics. Rather than simply observing what pages are visited and how many unique sessions are logged, action tracking allows you to narrow down groups of visitors based on the actions they take on your site.

In most instances, action tracking requires setting up a code in your analytics program and attaching that code to a button, page load, image rollover, or other JavaScript-trackable task (a click or hover).

You can see from Figure 11-7 that:

- Moz has action tracking in place where the users click a particular button. Note that it is often the best choice to have the action tracking on the action confirmation page rather than the sign-up form itself, as tracking on the sign-up form can result in incomplete form fills being included as part of your action total.

- Based on the people who take this action, you can predict which search terms

will be better at converting visitors into applicants.

- The Event Value and Avg. Value columns show 0, but if you were tracking ecommerce buyers, you could assign values to each event and track high-volume buyers. It is also a good idea to assign a value to different actions such as “contact us” requests, subscribers to free services, or any other user actions that could later lead to a conversion.

- Expanding on this idea, you could also track users by time of day, the search engine they used, their geographic location, and so on.

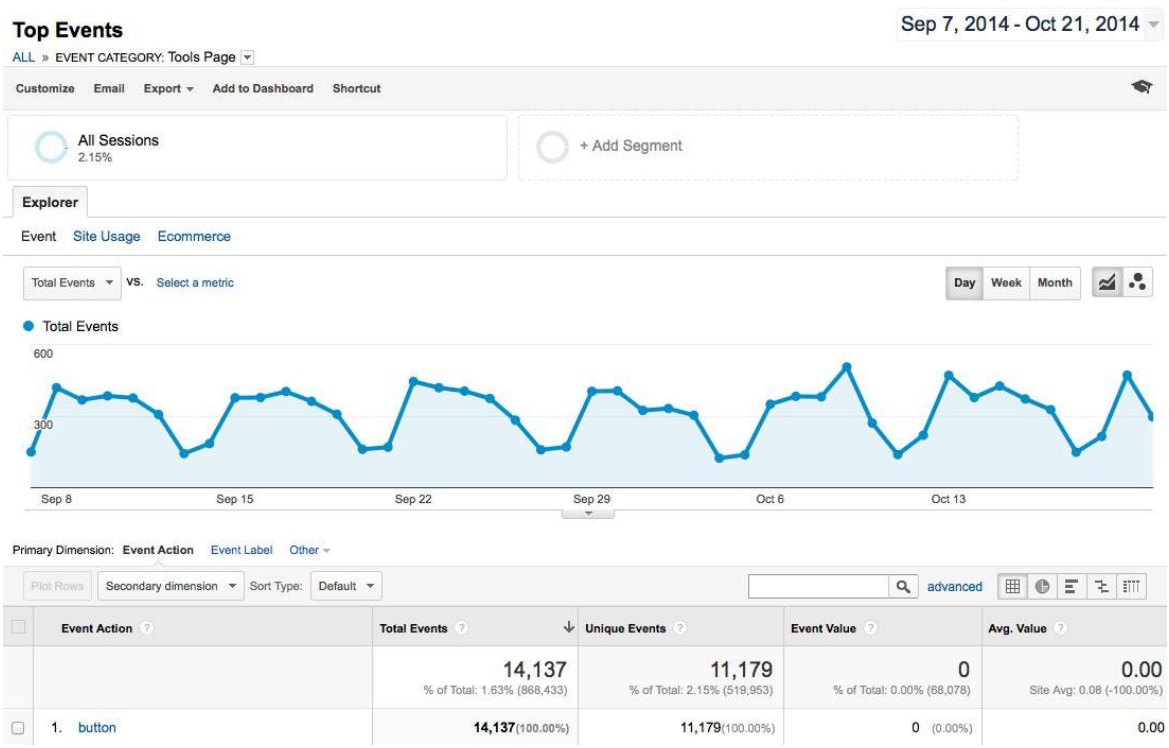


Figure 11-7. Action tracking in analytics

So, what types of actions should you track on your site? The answer varies depending on your business and site structure. Here are some suggestions segmented by site type.

Ecommerce sites:

Add to Cart button

Studies have shown us that users who “add to cart,” even if they do not

complete the checkout process, are more likely to return to make a purchase. This is also a good way to calculate shopping cart abandonment and make changes to refine and improve the process.

Complete checkout

An obvious one, this action will show you what percentage of each user group is converting into sales. It is also useful to measure what percentage of people start the checkout process but do not complete it.

Save to wish list

Ecommerce sites offering wish lists are still in the minority, but wish lists are a great way to track interest that isn't quite a purchase.

Send this to a friend

Many sites offer a "share this page" function, and it is a great action to be aware of. If folks are sending out your link, you know you have a hit.

B2B sites:

Subscribe to newsletter

A subscription is a tacit endorsement of your brand and a desire to stay in contact. It may not be a conversion, but for B2B, it may be the next best thing.

Contact form submission

A contact form submission can be even more valuable than a newsletter subscription, in some cases. Though some of these forms will report support issues, many may contain users' questions about your products/services and will indicate their desire to open a sales conversation.

Email link

As with contact forms, direct email links have the possibility of becoming sales contacts. The best thing you can do is clearly label sales emails and track them separately from support or business issues.

Blogs:

Subscribe to RSS feed

An RSS feed subscriber is a blog's equivalent of a conversion; tracking these is imperative. While the number of RSS subscribers is declining, this is still a useful metric to track.

Add comment

Anyone who is contributing content to the blog or otherwise participating merits your attention (as do those channels that earn user engagement).

Social bookmark/share

Anyone submitting your content to Facebook, Twitter, Delicious, StumbleUpon, reddit, and other social media and news aggregation/discussion sites deserves to be recognized (and sought after).

Forums or other sites based on user-generated content (UGC):

Sign up for an account

These users are active contributors; you need to know where they come from.

Contribute content

When a user publishes, discovering his path is important (especially if it is not from a bookmark/type-in).

Add comment

As in the preceding list item, comments are a great predictor of engagement.

Vote/rate

Even low levels of participation, such as a rating or a vote, are worth tracking when every piece of participation counts.

Social shares

Facebook friends and likes, Google +1s, Twitter followers are all important data points. You can read more about this in “Tracking Social Media” in Chapter 8.

You can get very creative with many of these and track all sorts of actions. If you offer a tool or calculator, track its usage. If you want to follow who clicks a particular link (rather than just a page), add an action to it. You can even see which users hover on an image.

For many marketers, action tracking is the secret sauce. Once you install and activate this data, it is easy to adjust SEO (or PPC) marketing efforts on particular terms, phrases, or pages that will make a huge difference to the campaign’s efficacy.

For SEO practitioners, action tracking is a “cannot live without it” tool. The recommendations made based on this kind of tracking are backed up by real data. High-volume action-tracking studies almost always make a huge impact on a site’s bottom line. Action tracking also allows you to compare the performance of organic search against other channels such as paid search, email marketing, and social media.

Separating the Analytics Wheat from the Chaff

One of the big problems with web analytics is that there is so much data. There is seemingly an endless supply of the stuff—so much that you can get lost in the weeds and waste a lot of time. As Dennis Mortensen of Visual Revenue likes to say, do not “report-surf.”

By this he means you shouldn’t get lost in the data, looking at one interesting fact after another, without understanding what business objective you are attempting to accomplish. To avoid this trap, you need to have a plan.

The basis of this plan is the definition of actionable key performance indicators (KPIs). To summarize this concept very briefly, do not view a report unless there is a chance that it will provide you with enough insight to take action (such as reallocating resources to fix a problem or pursue a new opportunity you discovered).

Common analytics mistakes

Because of the complexity of web analytics, there are plenty of opportunities to make mistakes. Even if you are laser-focused on actionable KPIs, you can still draw wrong conclusions from the data. Here are a couple of examples:

Making a decision on too small a sample size

A simplistic example would be if you had a keyword that brought you two visitors, and both of them bought something. What does that mean? Your conversion rate will always be 100%? Of course not.

You need to have a test run for a length of time that is statistically significant. According to Jonathan Mendez, long-time evangelist for Offermatica (which is now Adobe Target), the "golden rule" for test length is 100 conversions per branch (or tested element). For an A/B test, ideally you would have 200 conversions before you looked at the results. However, if after 50 conversions you have 40 on one branch and 10 on the other, your test is over. If your variance in results is lower, you need more data to make sure what you are seeing is accurate.

Not recognizing when you have enough data

You can make this mistake in the opposite direction too. When you have enough data, you need to be willing to act on it.

Working with biased data

You can easily collect biased data and make the wrong decision. For example, an ecommerce site basing its SEO plans on December's traffic numbers is probably being misled. And temporal factors aren't the only cause—visitor locations, referral sources, and in-house campaigns to existing lists can all skew information.

Tying SEO to Conversion and ROI

As we discussed previously in this chapter, it is important to tie your SEO efforts to the results they bring to the business. A fundamental piece of that is measuring the conversions driven by organic SEO traffic. Here are some of the most common types of conversions:

Sales/sales revenue

This is the one that everyone assumes is part of conversions. Sales and sales revenue (or better still, margin) conversions can be the simplest things to track, except when you are selling many different products at different price points and in different quantities. In this case, the process would need to be a bit more sophisticated.

If your site is advertising-driven, you need to look at the impact of organic search traffic on advertising revenue. If you have no financial goals for your site, you need to look at some of the other types of conversions and determine their value or worth.

Email/blog/newsletter subscriptions

Anytime a user signs up to receive regular communications from you, it is a win. Even though there are not direct financial consequences to this, it is still a conversion. Someone who has subscribed to something you offer is more likely to become a customer than a first-time visitor to your site, so you need to credit this type of conversion.

Sign-ups

Closely related to the notion of subscriptions are other types of sign-ups. Perhaps you offer a service such as a tool that people need to sign up for to use. Even if the tool is free, you should track this as a conversion.

You most likely received the person's email address in the process, and even if she indicates that she does not want to receive commercial communications from you, you should be building loyalty with the tool you provided her access to (or whatever she signed up for); otherwise, why would you be providing her with that service?

Downloads

Many sites offer free downloads, such as white papers or tools. Even if you do not require a sign-up of any type, you should still count this as a conversion. You are getting your message out there with the downloads you

offer.

Contact forms and phone calls

When someone contacts you or requests that you contact him, via a “contact us” form or a phone call, it might count as a conversion. Phone calls can be tracked through unique numbers per type of action, landing page, and traffic source.

Visitors who share

This conversion happens when a visitor shares the information she found on your site with someone else. This includes Facebook shares, Google+ shares, and links on Twitter. In addition, if your site has a “share with a friend” or an “email to friend” feature, you could keep track of conversions by noting each time a user used that feature.

Visitors who link

Users who find your site’s content useful, entertaining, or otherwise compelling may link to it from their own sites. You should count such linking as a conversion.

Visitors who publicize

Visitors can publicize your site by mentioning it in forums, blog comments, or social media sites such as Twitter and Facebook, or by writing about it on their own sites. These actions should count as conversions.

You should place a dollar value on every type of conversion you receive. Even a white paper download has some value to it; perhaps it may increase the potential of a future sale to that user. For example, if 100 people download your white paper and two of them become customers and buy a $100,000 consulting contract, you can estimate the value per download at $2,000.

SEO software packages such as Conductor Searchlight and Raven SEO Tools allow SEO professionals to view search ranking data together with traffic and revenue data. This enables them to tie natural search traffic to conversions, and therefore measure return on investment (ROI). You can see sample screens from

Conductor Searchlight in Figure 11-8 and Raven SEO Tools in Figure 11-9.

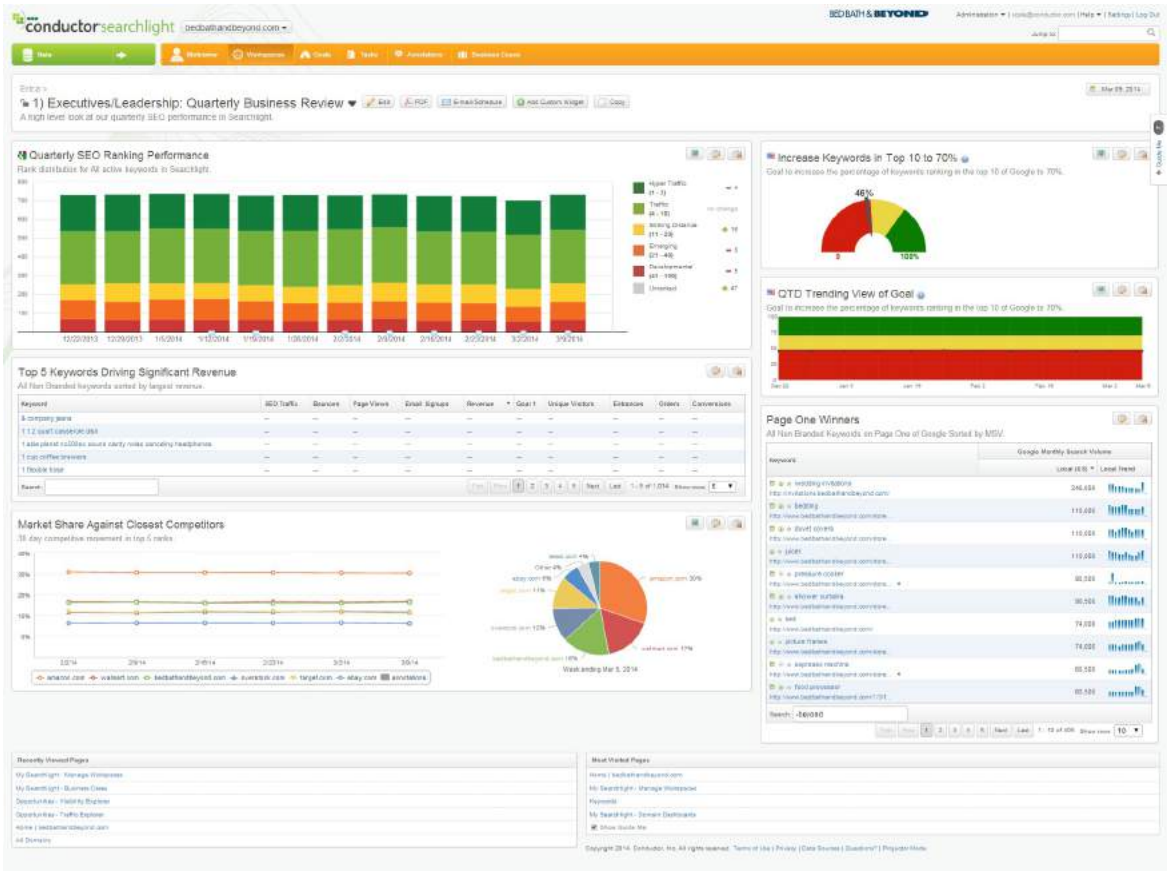


Figure 11-8. Conductor Searchlight ranking data overlaid with traffic and revenue metrics

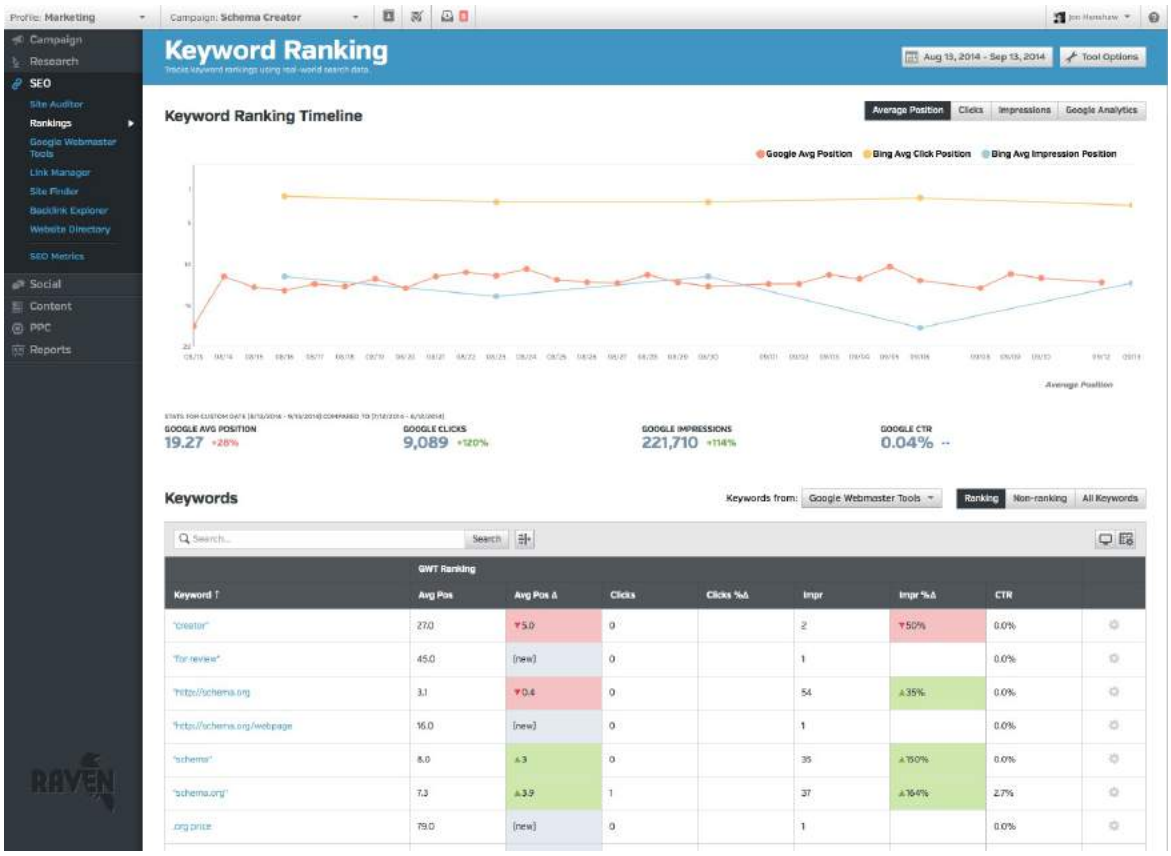


Figure 11-9. Raven SEO Tools ranking data and conversions

Managing Attribution

Another issue to be aware of is attribution—identifying the user actions that led to a specific outcome. The challenge is that there is a tremendous amount of interaction between types of media. Here are some examples:

- A user does a search, clicks on an organic search result, and then reads a few things and leaves. The next day, she remembers what she read, does another search, and this time clicks on a paid search ad for the site and buys a product (organic search should receive some credit).

- A user does a search, clicks on an organic search result, and leaves. A few days later, because of what she learned, she goes into a store and buys a product (organic search should receive some credit).

- It goes in the other direction too: a user sees a TV ad, then does an organic

search to find the website, and buys the product (the TV ad should get some credit for the sale!).

Such cross-channel interactions are becoming increasingly common. Data published on Search Engine Watch supports this.[2] A ScanAlert study of a large number of ecommerce sites showed that 30% of transactions took place one to three days after the customer's initial visit. So, the problem of attribution management is significant!

It is not always easy to track multivisit conversions in detail. Certainly, a direct interaction between TV/offline stores and a website is very difficult to track. But even looking at multiple site visits with an interaction between PPC and organic search can be difficult, partly because the tracking systems rely on cookies, which users have a tendency to delete.

So, attribution is something you should think about, and ideally you should attempt to track the impact across multiple site visits. However, the methods for attributing sales across multiple site visits are imperfect. One thing you can do to mitigate that is to count different types of conversions, such as filling out contact forms, downloading white papers, visiting key product pages, and so on. Tracking a larger number of events that indicate progress toward a sale (or whatever your end goal may be) can help you identify at which stage in the research and conversion process organic search comes into play for users.

Setting Up Analytics Software to Track Conversions

One of the great things about tracking conversions is that you can use the resulting data to focus your efforts. If a given page or referrer is showing much higher conversion rates than another source, you may want to focus more of your SEO efforts on those rather than on something else that does not convert as well.

Capturing conversion data takes some setup. Your analytics software does not know what you consider a conversion unless you define it. For example, in Google Analytics, you need to go into your Profile Settings to configure conversion goals and funnels (see Figure 11-10).

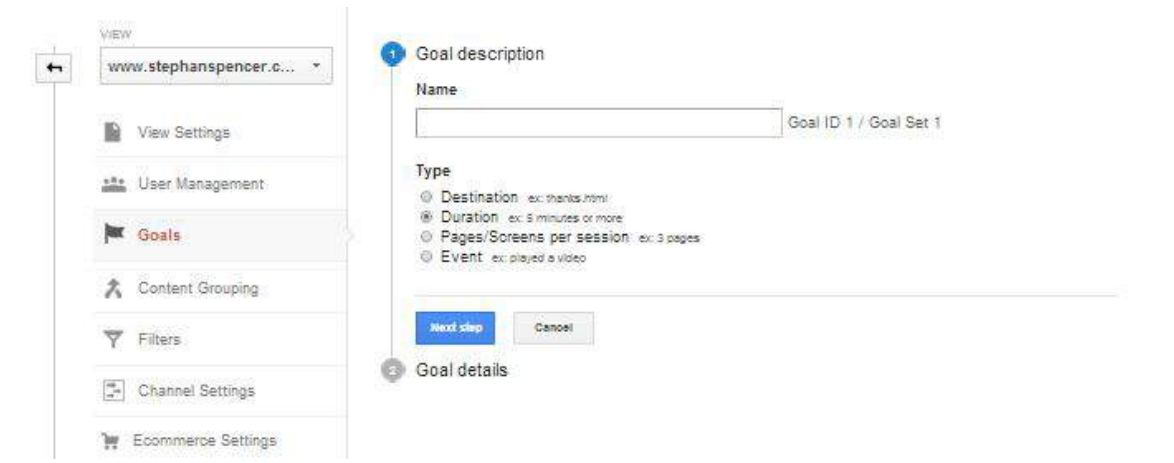


Figure 11-10. Google Analytics goals

Once you click Edit (on the right) for one of the goals, you come to a screen that allows you to define a goal (see Figure 11-11).

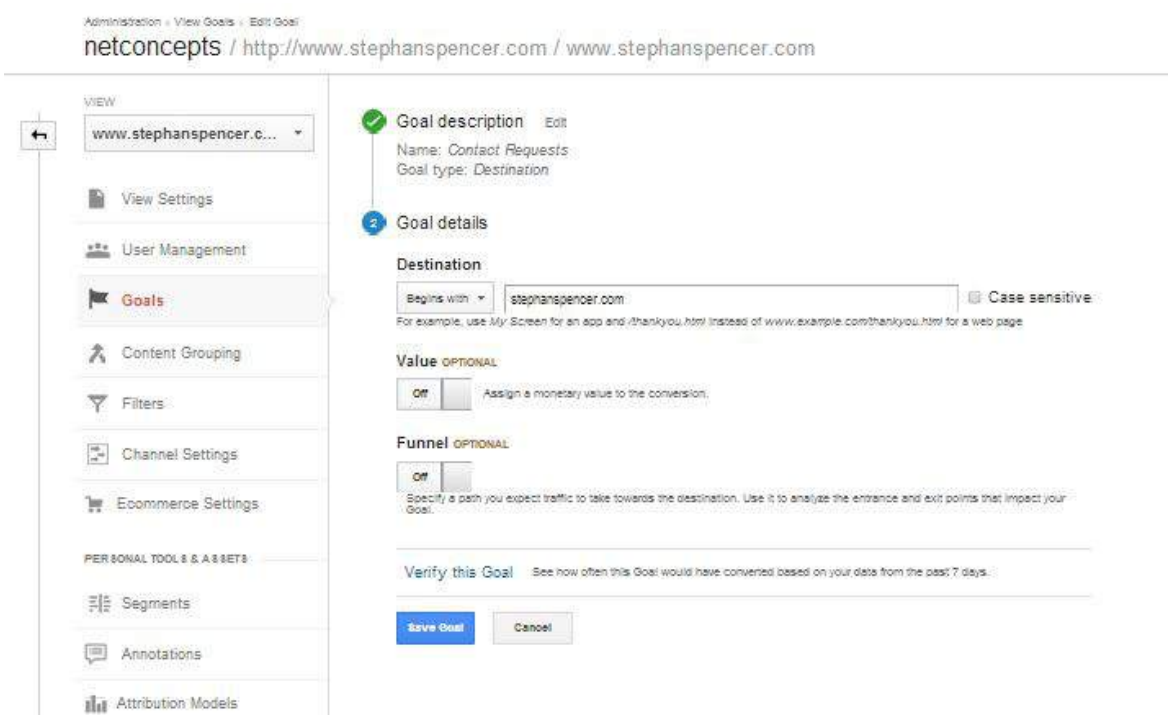


Figure 11-11. Google Analytics goals setup

In Figure 11-11, the primary data being tracked is when someone arrives at a given web page. These types of conversions are easy to set up.

Event tracking, such as someone clicking on a link, downloading a white paper, clicking on Flash or AJAX elements, or playing a video, is slightly more complicated. Because you cannot count on a page load to tell you when a link has been clicked on, you need to add a bit of tracking code to the link itself. Here is what the tracking code looks like for Google Analytics:

```
<a href="#"
  onClick="gaq.push(['trackEvent', 'Video',
    'Play', 'Art of SEO Book Launch']);">
  Play Video</a>
```

The key code is `onClick="gaq.push(['trackEvent', 'Video', 'Play', 'Art of SEO Book Launch'])`. This tells Google Analytics to associate this event to the user's action. Event configuration is best handled by an experienced developer. More information is available through the Google Analytics Event Tracker Developer Guide).

Google has also introduced a new analytics package called Universal Analytics, which allows you to track users across multiple devices and sessions. It also offers you the flexibility to implement tracking (including event tracking) through both JavaScript on your website and the Google Analytics SDK for mobile, and there's even a protocol to measure engagement across other digital devices, such as game consoles. Implementing Universal Analytics requires several steps that are not intuitive, so we recommend engaging an experienced analytics professional to help you through the process.

Google Analytics provides the particularly useful capability of remarketing to previous site visitors in Google AdWords.[3] You can segment your users based on a specific behavior, such as visiting a particular product page, and make a list of these users in Google Analytics. If the list is longer than 100 users, it is automatically sent over to Google AdWords. Then your ads will be marketed to solely those users across the Web.

Form submissions are another metric you may want to track. This category could include when someone has completed a request to sign up for a newsletter, requested or started a white paper download, or simply used a form to request that you contact him. These are usually quite easy to set up, provided that the result of the form submission is some sort of "thank you" page. All you need to

do in this case is tell your analytics software that the loading of that “thank you” page is considered a conversion, and then give it a value.

Analytics is particularly useful if a portion of your customers tend to log in to a company’s features with a social media login, such as a “Login via Facebook” mechanism. In that case, you have the ability to track users across several devices, for even more advanced and thorough customer data.

Facebook Insights gives you yet another valuable data set for tracking conversion. Domain Insights shows all of the traffic from Facebook to your domain, including traffic from your Facebook business page, from links on status updates, from social plug-ins, and more.

Life gets more complicated when you want to track actual product sales on sites where products vary in price. Generally speaking, this will require putting some custom JavaScript on your order confirmation page.

In addition, you will need to have your web application populate certain data in the JavaScript before executing it, such as the total amount spent. Oftentimes, publishers choose to include a lot more data, such as listing all the products sold, the price per product, how many units, shipping costs, and so on.

Although the preceding examples used Google Analytics, any of the major analytics packages can do this type of conversion tracking.

Conversion tracking strategy

An important step in conversion tracking is deciding what you want to call a conversion. As we outlined earlier in this chapter, it is important to consider all types of conversions, such as sign-ups, contact requests, downloads, and so forth. There is one other decision you need to make: how to value each conversion.

What is a contact request worth? What about a newsletter sign-up? One way to back into this is to look at how many of your newsletter sign-ups become customers. For example, if 100 people sign up for the newsletter, 10 of them become customers, and the total dollar value of their orders is $200, the average value per sign-up is $2.

When you first set up conversion tracking, you may need to estimate some of these things. But as your data improves, you can improve your calculation and

refine the value you place on each conversion.

Segmenting Campaigns and SEO Efforts by Conversion Rate

Now that you have conversion tracking set up, how do you use it to focus your SEO efforts? One of the things you will want to do is to track conversion data in different ways, such as looking at:

Conversions by keyword

What keywords are bringing the best results? Note that this data has become far more obscured since the “not provided” update in 2013 but can still be useful to measure. Even linking Search Console to Google Analytics does not enable conversion tracking for those terms.

Conversions by referrer

Which traffic source is offering the best conversion?

Conversions by web page

Which pages on your site result in the highest number of conversions?

Conversions by initial entry page

Which initial entry pages ultimately lead to the highest number of conversions?

For example, Moz actively tracks its conversions using Google Analytics. Tracking by keyword can be a useful way to do this, as shown in Figure 11-12.

	Keyword	Acquisition			Conversions Goal 14: Moz Subscription Signup		
		Sessions	% New Sessions	New Users	Moz Subscription Signup (Goal 14 Conversion Rate)	Moz Subscription Signup (Goal 14 Completions)	Moz Subscription Signup (Goal 14 Value)
		28,475 % of Total: 0.70% (4,072,000)	38.48% Site Avg: 51.13% (-24.74%)	10,957 % of Total: 0.53% (2,082,039)	1.81% Site Avg: 0.17% (967.09%)	516 % of Total: 7.46% (6,915)	$0.00 % of Total: 0.00% ($0.00)
1.	moz	12,694 (44.58%)	34.37%	4,363 (39.82%)	1.61%	204 (39.53%)	$0.00 (0.00%)
2.	seomoz	4,609 (16.19%)	36.99%	1,705 (15.56%)	1.76%	81 (15.70%)	$0.00 (0.00%)
3.	seo moz	1,858 (6.53%)	35.04%	651 (5.94%)	2.58%	48 (9.30%)	$0.00 (0.00%)
4.	moz com	1,289 (4.53%)	30.88%	398 (3.63%)	1.24%	16 (3.10%)	$0.00 (0.00%)
5.	seo tools	763 (2.68%)	58.58%	447 (4.08%)	2.23%	17 (3.29%)	$0.00 (0.00%)
6.	seo mox	731 (2.57%)	39.12%	286 (2.61%)	4.24%	31 (6.01%)	$0.00 (0.00%)
7.	moz pro	372 (1.31%)	20.97%	78 (0.71%)	2.15%	8 (1.55%)	$0.00 (0.00%)
8.	moz analytics	317 (1.11%)	28.71%	91 (0.83%)	0.95%	3 (0.58%)	$0.00 (0.00%)
9.	seo website analyzer	299 (1.05%)	52.84%	158 (1.44%)	2.01%	6 (1.16%)	$0.00 (0.00%)
10.	website seo analysis	284 (1.00%)	57.04%	162 (1.48%)	0.35%	1 (0.19%)	$0.00 (0.00%)

Figure 11-12. Google Analytics tracking conversions

You can see that the conversion rate on the keyword *seo tools* is high. Perhaps an SEO campaign to improve the ranking of Moz on these keywords is in order. Not only do you see high conversion, but you also see reasonable material volume.

Implementing an SEO campaign that doubles related traffic for this set of keywords will likely have a significant impact on the bottom line. Google's move to obscure keyword data makes this form of measurement less valuable than it used to be, but it can still be a useful directional indicator.

Increasing Conversion

As an SEO practitioner, you should strive to become an expert at conversion optimization, because higher conversion rates mean higher impact for your SEO campaigns. Coauthor Stephan Spencer offers a viewpoint on this topic in a MultiChannel Merchant article "Be a Converter".

Nonetheless, the average site conversion rate for a typical merchant today is only about 2.4% combined for new and returning visitors (source: Monetate ECommerce Quarterly Report EQ1 2013).

So, how do you help your conversion rate by managing your content and refining your site design? Most web resources that give content advice are

focused on blogs—specifically, what and how often you should publish. Those of us who use websites strictly for sales and marketing are equally concerned with good content, but from a totally different angle—one that traditional content advice doesn't properly address.

What's more, much of the design advice out there seems contradictory. The key to enhancing your conversions by adjusting your design and content is not in thinking strictly about the technology or the mechanics of site layout, but in understanding the needs of your visitors. In that, there are many different elements to consider. A book about SEO can't cover this topic in depth, but you may want to pick up one of the following excellent books that do: *You Should Test That!* by Chris Goward (Sybex), *A/B Testing* by Dan Siroker and Pete Koomen (Wiley), *Always Be Testing* by Bryan Eisenberg and John Quarto-von Tivadar (Sybex), *Landing Page Optimization* by Tim Ash (Sybex), and *Web Design for ROI* by Lance Loveday and Sandra Niehaus (New Riders Press).

The content marketing bump

A specific example of how to use action tracking by referral source is to track a *content marketing bump*. The idea is that after successful content draws thousands of visitors to your site, a small percentage of those visitors will stick around and continue to visit on a semiregular basis (this is particularly true if you're regularly delivering fresh, valuable content).

Action tracking by referral source

Figure 11-13 shows a sample report from Google Analytics.

In Figure 11-13, you can see a list of domains that sent Moz traffic over a few weeks' time, along with the number of resulting Moz subscription sign-ups. You can see some clear patterns, including the following:

- WebStatChecker, Feedly, Google Plus, Stack Overflow, and reddit are sending traffic, but less than .001% of those visitors sign up for membership on the site.

- The domains that point primarily to the Moz tools (such as Backlinko) are responsible for a high number of sign-ups, but this is to be expected, as a visitor has to register to see the full data sets available in Moz's tools such as Open Site Explorer.

- SearchEngineJournal, Facebook, and QuickSprout are amazingly on-topic sources of traffic, which is directly borne out in the higher percentages of their referrals who become members.

	Source	Acquisition	Conversions Goal 14: Moz Subscription Signup		
		Sessions	Moz Subscription Signup (Goal 14 Conversion Rate)	Moz Subscription Signup (Goal 14 Completions)	Moz Subscription Signup (Goal 14 Value)
		974,998 % of Total: 23.94% (4,072,006)	0.09% Site Avg: 0.17% (-47.39%)	871 % of Total: 12.60% (6,915)	$0.00 % of Total: 0.00% ($0.00)
1.	analytics.moz.com	70,681 (7.25%)	0.06%	45 (5.17%)	$0.00 (0.00%)
2.	moz.com	52,470 (5.38%)	0.00%	0 (0.00%)	$0.00 (0.00%)
3.	t.co	30,139 (3.09%)	0.10%	29 (3.33%)	$0.00 (0.00%)
4.	webstatschecker.net	21,658 (2.22%)	0.00%	0 (0.00%)	$0.00 (0.00%)
5.	example.com	16,738 (1.72%)	0.00%	0 (0.00%)	$0.00 (0.00%)
6.	feedly.com	14,028 (1.44%)	0.04%	6 (0.69%)	$0.00 (0.00%)
7.	opensiteexplorer.org	13,783 (1.41%)	0.04%	6 (0.69%)	$0.00 (0.00%)
8.	facebook.com	10,115 (1.04%)	0.28%	28 (3.21%)	$0.00 (0.00%)
9.	plus.url.google.com	10,067 (1.03%)	0.04%	4 (0.46%)	$0.00 (0.00%)
10.	seocheki.net	9,985 (1.02%)	0.01%	1 (0.11%)	$0.00 (0.00%)
11.	searchenginejournal.com	9,347 (0.96%)	0.26%	24 (2.76%)	$0.00 (0.00%)
12.	l.facebook.com	8,106 (0.83%)	0.32%	26 (2.99%)	$0.00 (0.00%)
13.	inbound.org	7,679 (0.79%)	0.09%	7 (0.80%)	$0.00 (0.00%)
14.	reddit.com	7,015 (0.72%)	0.04%	3 (0.34%)	$0.00 (0.00%)
15.	backlinko.com	6,668 (0.68%)	0.19%	13 (1.49%)	$0.00 (0.00%)
16.	stackoverflow.com	5,578 (0.57%)	0.02%	1 (0.11%)	$0.00 (0.00%)
17.	searchengineland.com	5,492 (0.56%)	0.11%	6 (0.69%)	$0.00 (0.00%)
18.	quicksprout.com	4,793 (0.49%)	0.15%	7 (0.80%)	$0.00 (0.00%)

Figure 11-13. Action tracking by referral source

Action tracking is one of the most valuable things you can do on any type of site. Most people limit its application to ecommerce, but you can see from the data in Figure 11-13 how valuable it can be, even for a simple blog. Even if you do not have a membership sign-up, you could attach the action tracking to your “add comment” link in a blog or to a newsletter sign-up on a content site. Segmenting that data by incoming referrals (search engines, direct type-ins, and links) gives

you an even better picture of who's sending you truly valuable traffic. You can also get an indication of traffic quality by looking at bounce rate, time on site, and number of pages viewed for each referrer.

In Figure 11-14, you can see that around late May, Moz published some successful content that helped bump up the site visitor count at least through mid-June. This traffic "bump" after the spike is another great benefit of creating viral-worthy content; this is a terrific way to grow the popularity of a site.

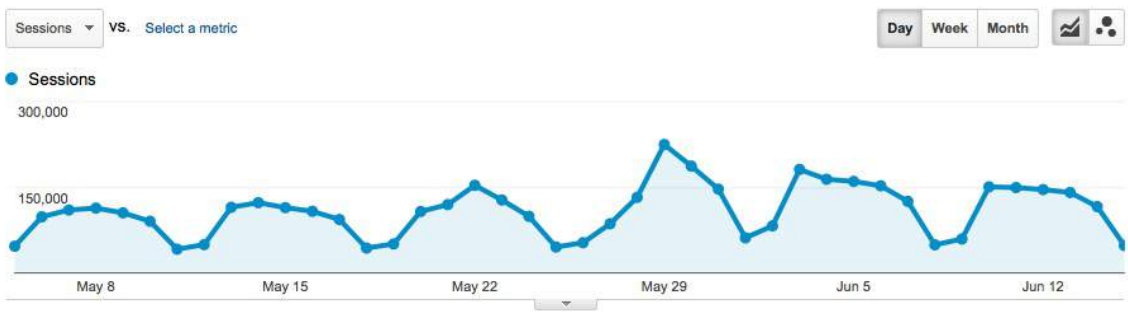


Figure 11-14. Traffic report

The tricky part is to continue to deliver great content after a viral piece and maintain the creative energy necessary to prove to visitors that you're worth coming back to again and again. It is not an easy sell—people's time is valuable. You have to make your content so targeted, interesting, and worthy of their attention that they cannot look away. It is challenging and time-consuming, but it is a great way to build a successful site.

Determining Project ROI

An effective SEO process is one that continuously works toward a positive return on investment. A basic SEO ROI business case would include these elements:

Number of people searching for your keywords

This can be hard to estimate, because you cannot completely map out the long tail. One rough estimation strategy is to multiply the search volume for the top terms for your business by 3.3 (i.e., assume that the head terms are about 30% of the available volume).

Expected click-through rate from organic search

This relates to the position you think you can achieve. The #1 position can net you approximately 19% of the available clicks.

Average conversion rate

Once you have captured the visitor, how successful are you at completing a conversion? Typical conversion rates for a website might be between 2% and 5%. It should be easy to get this data from your analytics. You should already know what your conversion rate is!

Average transaction amount

Last but not least, factor in the average transaction value. Again, this is data you already have.

Ultimately, the formulas look like this:

- SEO revenue = people searching for your keywords *click-through rate* average conversion rate * average transaction amount

For example: 10,000 per day * 10% *5%* $100 = $3,000 per day.

- SEO ROI = SEO revenue / SEO cost (use total $ spent for salaries and consulting, or number of hours spent)

For example: $3,000 per day / $500 per day = an ROI of 6×.

An alternate theory and calculation

Predicting an SEO project's ROI based on rankings and click-through rate can, in some cases, be problematic because you have very little control over the variables. You end up relying on numbers that you have a limited ability to influence. As an alternative approach, you can measure and track SEO ROI based on an increase in search visibility.

To do this, begin by determining two things:

- How many pages are getting search clicks for the month

- How many clicks per page your site is getting

Next, record these supporting metrics:

- Average ranking across the whole keyword spectrum
- Average click-through rate

Now, by making it easier for search bots to find more pages, consolidating duplicate content, improving the depth and breadth of content, changing your pages to better match your users' needs, and improving page titles, metadata, and microdata, you should see an increase in the number of pages getting search clicks and/or the number of clicks per page. The combination should result in more traffic when compared year-over-year.

Modifying the previous formula to account for this alternate approach, we get:

- SEO revenue = increase in (pages getting search clicks *search clicks per page)* average conversion rate * average transaction value

Note that CTR in the long tail is far more variable. Not as many searchers are going to go past the first couple of results, so this formula won't always accurately show your return. ROI is always going to be a "best estimate," but this type of ROI calculation will be more accurate for the short tail.

Competitive and Diagnostic Search Metrics

Thus far in this chapter we have focused on the basics—the dollars and cents of determining whether you have succeeded in your SEO campaign. As we noted at the beginning of the chapter, these should be your first priority in your analytics efforts. In this section, we will start looking at metrics that you can use to diagnose specific SEO issues. An example would be finding out whether a major section of your site is not indexed. Another example is seeing how your traffic growth compares to that of your competitors (this helps you decide whether you have set the right objectives for your efforts).

Search Engine and Competitive Metrics

Numerous tools allow you to monitor your site and those of your competitors,

providing insight into your SEO progress. You can also use these tools to figure out what your competitors are doing from an SEO perspective. This type of intelligence can provide you with new ideas on how to adapt your strategy to get better results.

As with all such tools, it is always important to understand the context of the tools and to have an idea of what you are looking for. Better knowledge of your competitors' strategy is certainly one valuable goal. Detecting a problem in how your website is crawled is another.

By selecting specific and actionable goals, you can set yourself up for the highest possible return.

Site Indexing Data

It is valuable to know how many pages in your site are in a search engine's index. This is important for two reasons:

- If you discover that key parts of the site are not in the index, you can embark on an effort to determine why.
- Looking at competitors' indexed pages can help you learn about those competitors' sites and strategies.

You can get basic information on the number of indexed pages for a site using the *site:<yourdomain.com>*, as in a Google, Yahoo!, or Bing search. Figure 11-15 uses the results for Stephan Spencer's blog.

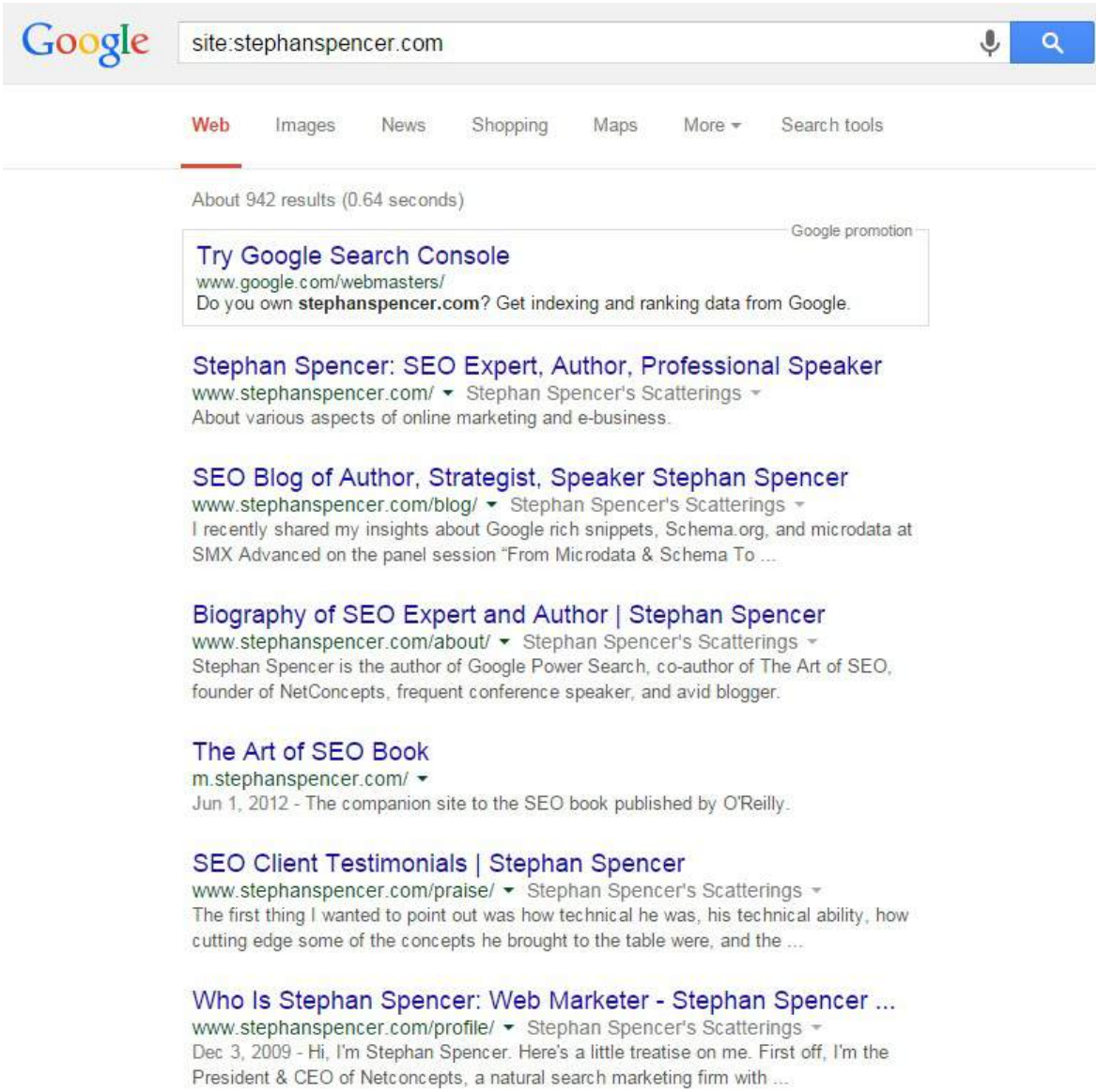


Figure 11-15. Site indexing data

As shown in Figure 11-15, Google reports 787 pages in the index. Is that all of the site's pages? Many site owners will have a pretty good idea of the number of pages on their site and therefore can get a gut feeling as to whether they have a problem. One important thing to note is that the results that the *site:* search operator return are fast estimates, and aren't as accurate as the reports generated by Google Search Console based on your XML sitemap. You can use them to give you a general impression, but that is all.

You can improve this data by performing your site query, switching to 100

results per page, and going to the last page of the Google results to see if the number of results has changed. This will give you more accurate totals for the number of pages indexed.

If you think many pages are missing from the index, you can begin to explore a bit deeper. You can start by narrowing down sections of the site. Figure 11-16 shows the 2014 blog post archives on Stephan's site.

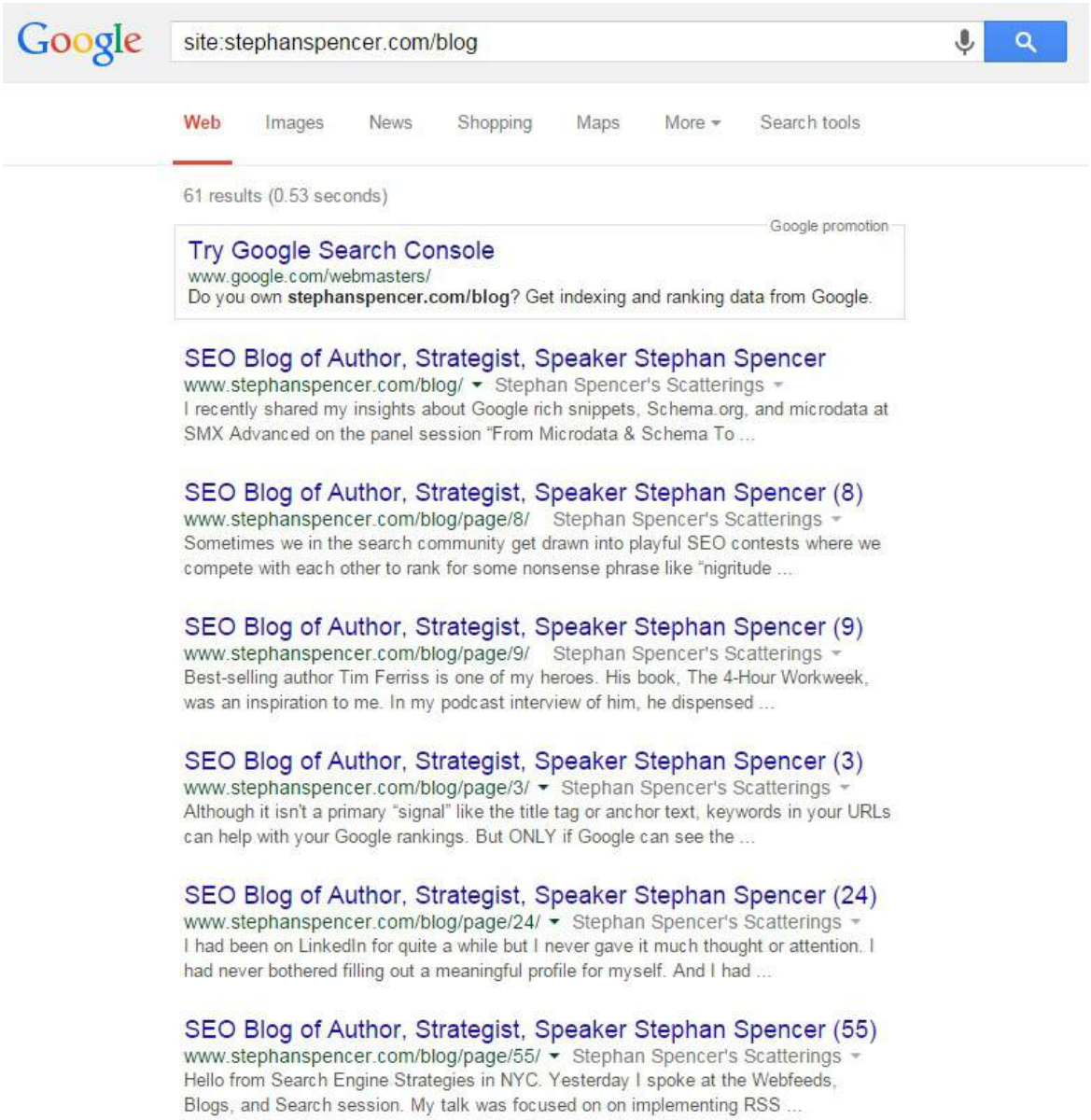


Figure 11-16. Indexing data on a website section

As shown in Figure 11-16, this reports 23 pages in the Google index. By using

these different search options, you can isolate what portions of the site appear to be less well indexed than others. You can also combine *site:* with other operators, as we outlined in Chapter 2. For example, consider combining it with the *inurl:* operator, as in Figure 11-17.

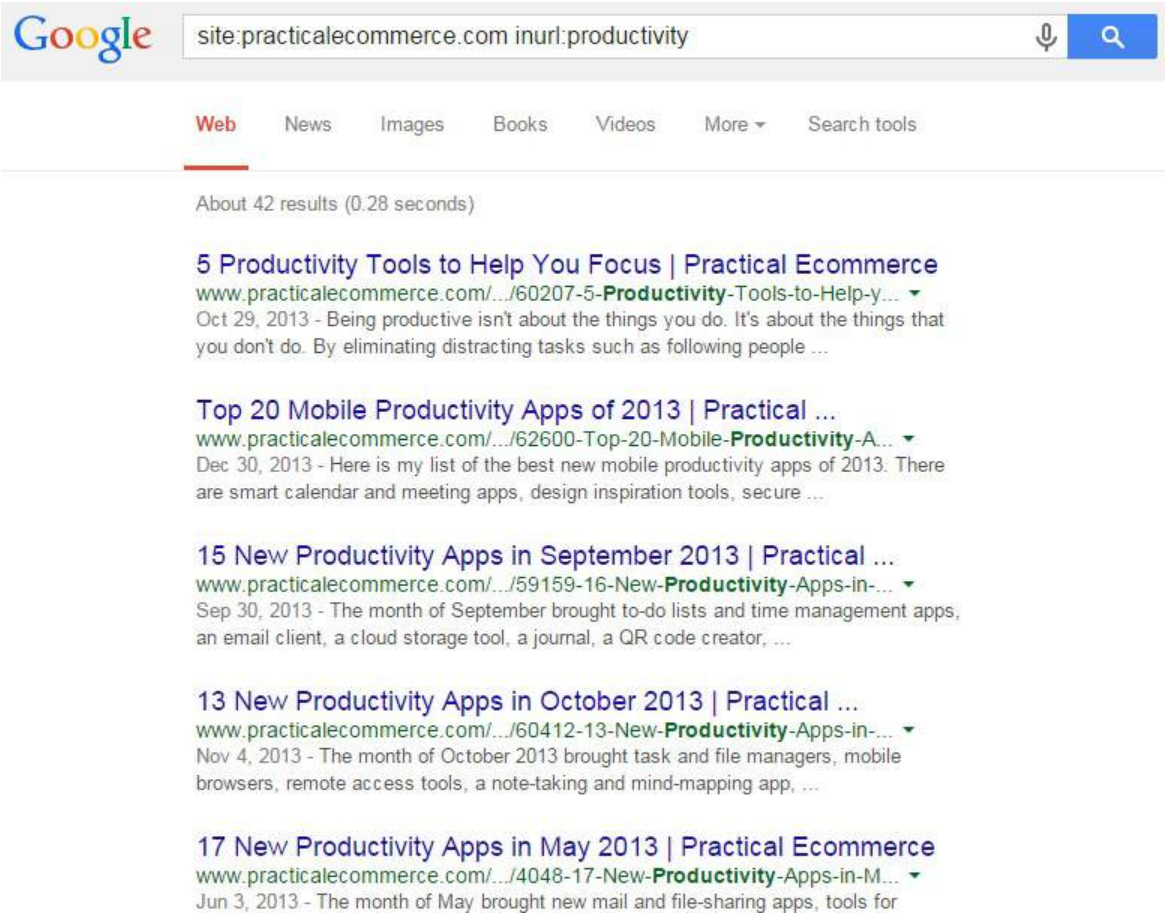


Figure 11-17. Combining the site: operator with inurl:

Figure 11-17 shows that *http://www.practicalecommerce.com* has 42 pages in the index that have the word *productivity* in the URL.

So far, we've shown website indexing tools that you can use on any website, which is great for competitive research. However, you can pull more accurate statistics for your website using Bing Webmaster Tools. You can also pull site indexing data from Google Search Console.

Both of these tools are very useful, and you should use them on every site you touch. By simply signing up and verifying site ownership, you will get a

tremendous amount of valuable data at no cost. For example, Figure 11-18 shows the opening screen for Bing Webmaster Tools for *http://stonetemple.com*.

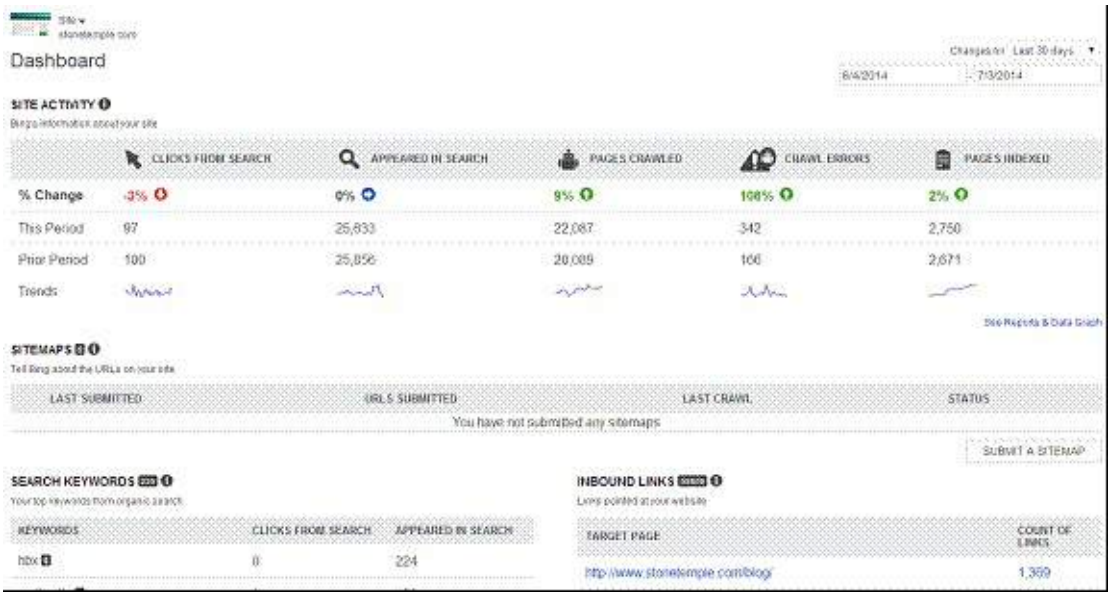


Figure 11-18. Bing Webmaster Tools

In Figure 11-18, it looks like Bing has 2,750 pages from the site indexed. If you search for *site:stonetemple.com* in Bing itself, it shows something rather different (see Figure 11-19).

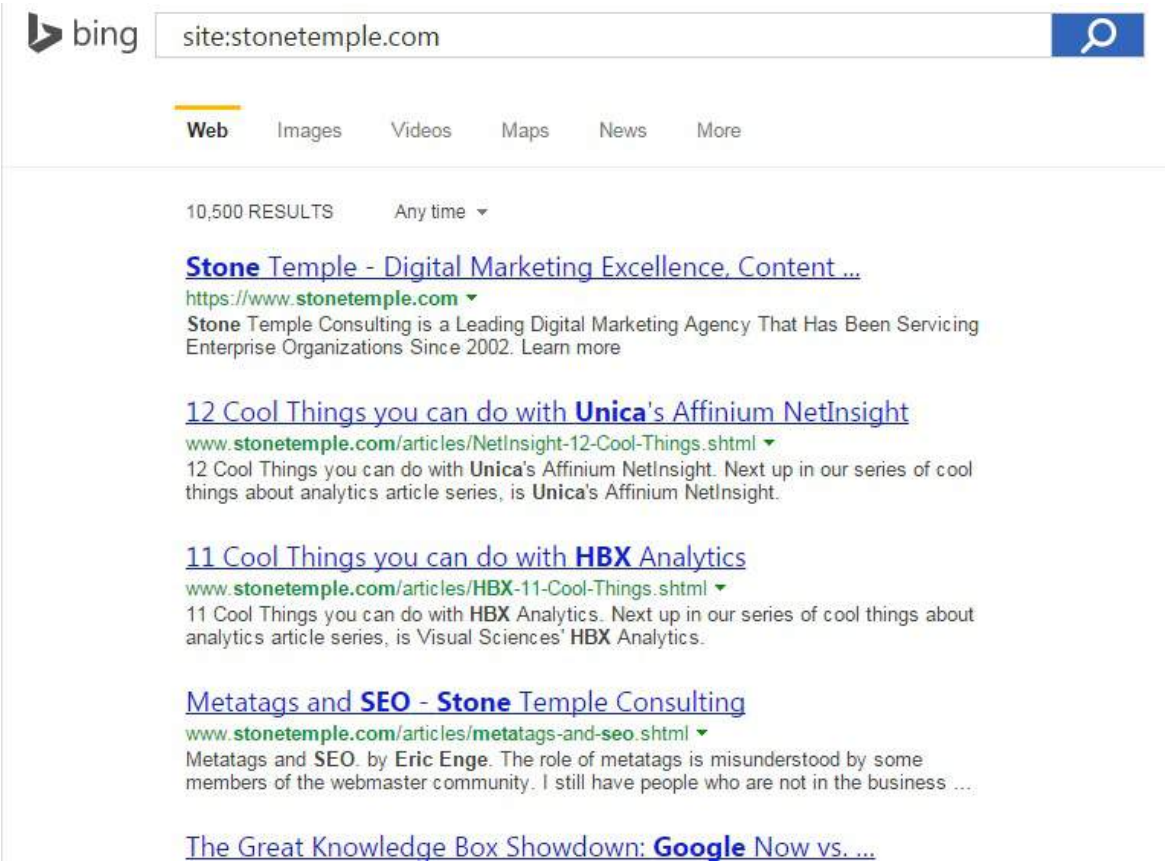


Figure 11-19. Results of the site:stonetemple.com search on Bing

This search shows 10,500 indexed pages. Of the two, you should consider the report from Bing Webmaster Tools to have more accurate data than the results from a site search at Bing.

Bear in mind that Google Search Console and Bing Webmaster Tools do not keep track of historical data for you. If you want to maintain a history of the data, you should plan on downloading and saving the reports.

LinkBased Tracking of Content Marketing

Measurement of content marketing campaigns is another must for the expert SEO practitioner. This can help you refine your campaigns to bring better results, compare link growth to your search rankings, and perform analysis to see how your campaigns are faring compared to the efforts of your competitors.

NOTE

Some analytics tools have built-in functions for comparative anchor text analysis.

Search engine–supplied tools

There are three basic tools for checking backlinks:

- Google Search Console is a powerful start. With Google Search Console, publishers can easily download a spreadsheet of all the links Google has in its database (note that Google might not include some links that it does not consider significant). With this tool, publishers can see only the links to their own site. Figure 11-20 shows the report.

- Bing Webmaster Tools is also a great asset. It offers a similar capability for downloading a spreadsheet of the links that Bing has in its database for a site. Once again, publishers can use this tool only to see the links to their own site. Figure 11-21 shows what the report looks like.

Overview » **All domains**
Top 1,000 domains that have links to pages on your site.

Download this table | Download more sample links | Download latest links | Show 25 rows | 1-25 of 1,000

Domains	Links	Linked pages
aggregage.com	5,250	4
vkistudios.com	1,762	1
wordpress.com	1,409	12
businessblogconsulting.com	1,373	2
nicklaskoski.fi	1,070	1
seobook.com	789	2
seokicks.de	611	3
problogger.net	319	2
thebillionairegirl.com	299	1
aboutonlinematters.com	285	2
blogspot.com	281	17
57883.net	217	1

Figure 11-20. Google report on external links

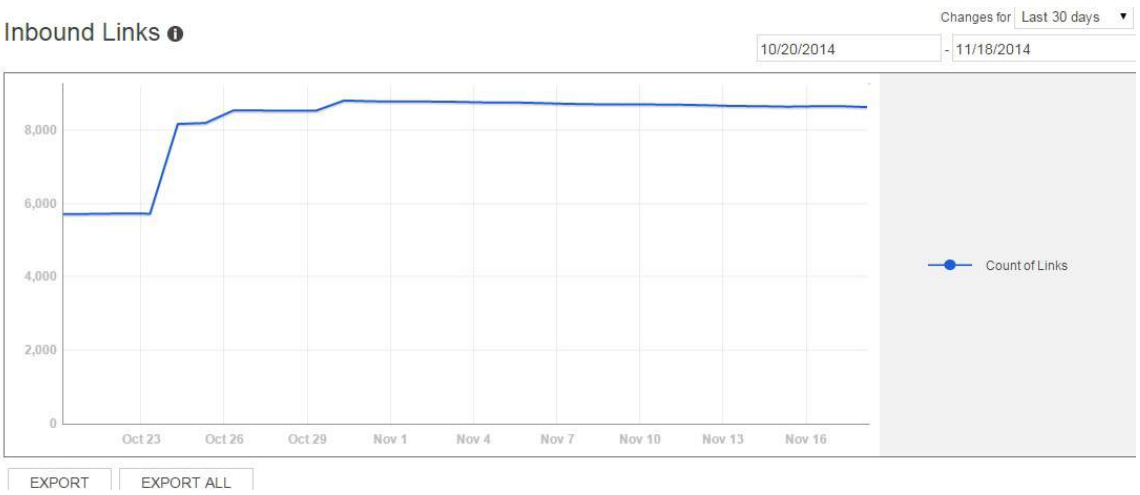


Figure 11-21. Bing report on external links

For quick-and-dirty link totals, you can use a Firefox plug-in known as SearchStatus, which provides basic link data on the fly with just a couple of mouse clicks. Figure 11-22 shows the backlinks menu you'll see.

Figure 11-22. SearchStatus plug-in

Notice also in Figure 11-22 that the SearchStatus plug-in offers an option for highlighting `nofollow` links, as well as many other capabilities. It is a great tool to help you pull these kinds of numbers much more quickly than would otherwise be possible.

An option for quickly referencing link totals in Google Chrome is Majestic's Backlink Analyzer browser extension, shown in Figure 11-23.

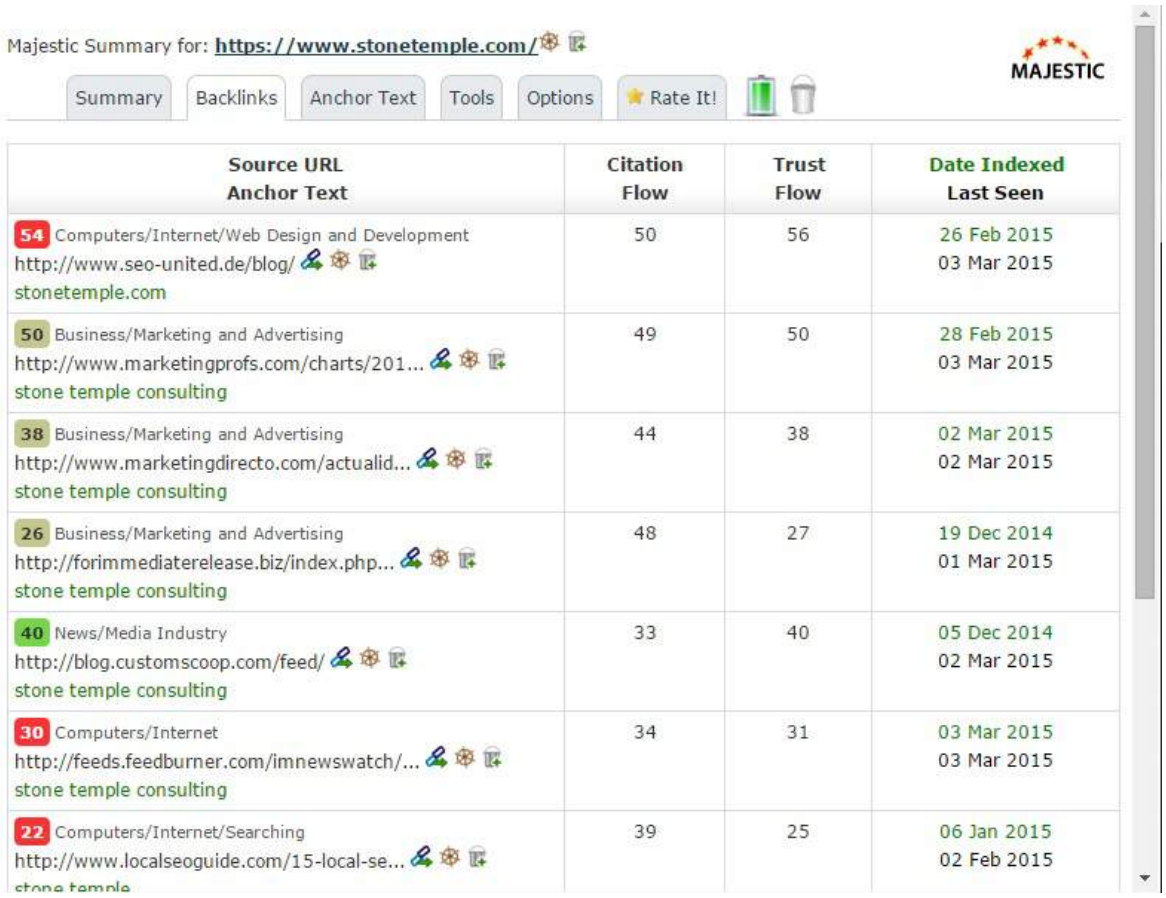


Figure 11-23. Majestic Chrome extension

The Majestic plug-in gives you rankings for Citation Flow (authority) and Trust Flow (trustworthiness) as well as the number of external backlinks and referring domains for the current URL, the subdomain that URL resides on, and the root domain for the current site.

Another useful tool for seeing link data at a glance is the MozBar Chrome and Firefox extension is shown in Figure 11-24.

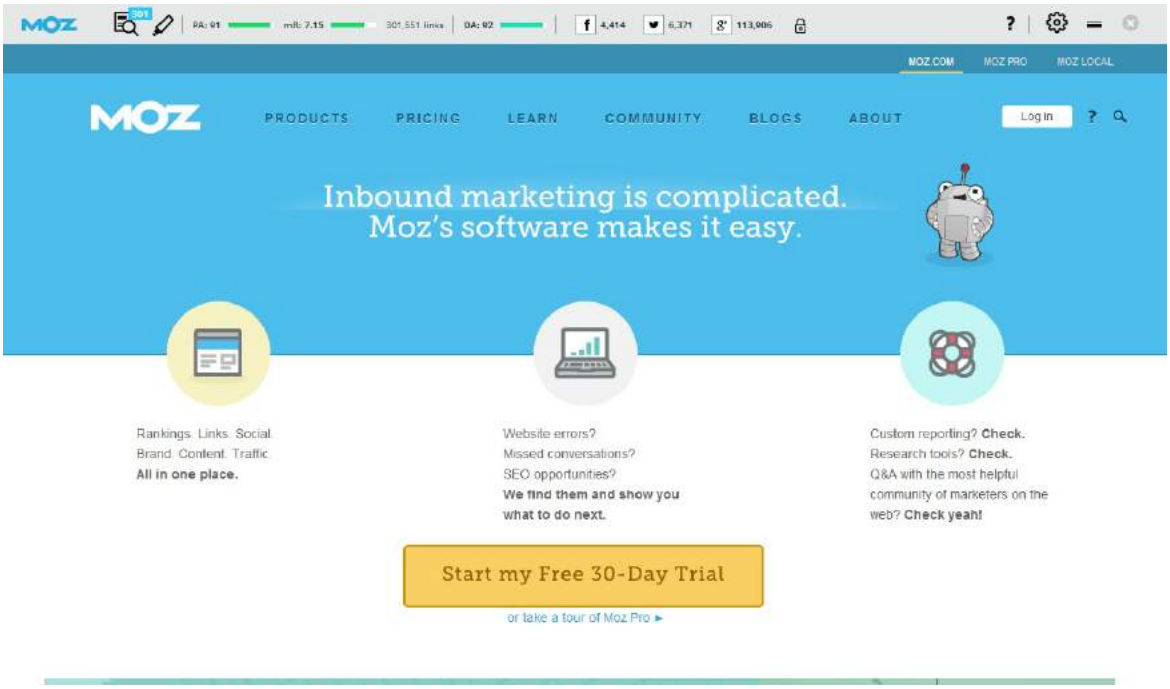


Figure 11-24. MozBar Chrome extension

The MozBar shows a great deal of data that is useful for SEO. For the purposes of quick backlink analysis, it shows Page Authority, MozRank (Moz's answer to PageRank), total number of links to the current URL, and Domain Authority. Figure 11-24 shows the data you can get for free. If you have a paid Moz membership, the toolbar will display additional link metrics such as the number of unique linking domains and data for the subdomain.

Third-party link-measuring tools

Next, we'll look at some of the better-known advanced third-party tools for gathering link data.

Open Site Explorer

Open Site Explorer was developed based on crawl data obtained by Moz, plus a variety of parties engaged by Moz. When you enter a competitor site's URL in Open Site Explorer, you can scope out where it's getting links, discover which pages are the most influential, and compare its link profile to yours to see how your performance stacks up. This tool lets you get as many links as it can find and then extract them all into a spreadsheet. Figure 11-25 shows a sample report.

Inbound Links

Gauge a site's influence. See inbound links to the page, subdomain, or root domain you've entered and analyze the linking pages.

Target: this page | Link Source: only external | Link Type: all links | Group links by subdomain

1 - 50 Inbound Links | Request CSV | Prev | Next

Title and URL of Linking Page	Link Anchor Text	PA	DA
Gizmodo - Tech By Design gizmodo.com/	io9 io9.com/	94	92
Lifehacker - Tips and downloads for getting things done lifehacker.com/	io9 io9.com/	93	92
Gawker - Today's gossip is tomorrow's news gawker.com/	io9 io9.com/	93	91
Advertising Age - Advertising Agency & Marketing Industry News - Advertisi... adage.com/	io9 io9.com/	92	91
Kotaku - The Gamer's Guide kotaku.com/	io9 io9.com/	91	89
Jezebel - Celebrity, Sex, Fashion for Women. Without Airbrushing. jezebel.com/	io9 io9.com/	91	89
Deadspin - Sports News without Access, Favor, or Discretion deadspin.com/	io9 io9.com/	90	88
Jalopnik - Drive Free or Die jalopnik.com/	io9 io9.com/	89	86
Valleywag - Sorry to disrupt. valleywag.gawker.com/	io9 io9.com/	89	91
This Is Apple's Next iPhone gizmodo.com/5520164/this-is-apples-next-iphone	io9 io9.com/	86	92

Figure 11-25. Open Site Explorer report

Included with this report are Page Authority and Domain Authority scores for each URL listed and each domain. Be aware that these metrics are useful in assessing the value of a domain or a page, but what the search engines actually measure and use may be substantially different.

The tool also flags redirects so that you can see them. The beauty of Open Site Explorer is that it allows SEO practitioners and publishers to collect competitive intelligence on other people's domains. All the data is extractable in a spreadsheet to allow for easy offline processing.

Majestic

Majestic offers comprehensive backlink analysis for your domains (see Figure 11-26). Majestic's flagship tool is the Majestic Site Explorer, which is

particularly useful for auditing the links to a website. You can quickly analyze any website's backlinks, new links, lost links, and anchor text—all in one place.

Majestic has two unique metrics: Trust Flow and Citation Flow. These are scored between 0 and 100. Trust Flow is a good quality metric; it allows you to easily analyze how trustworthy a website or backlink is. Citation Flow can give you an idea of how influential a site or page is—its "link juice"—based on the number of links coming into it and the authority of those links. Looking at the ratio between the Trust Flow and Citation Flow scores can also help you determine if a backlink might be toxic. A page with a high Citation score but low Trust may be one you want to avoid getting a link from, as there are many links pointing to the page but overall they're untrusted.

Topical Trust Flow, a newly available feature of Trust Flow, is also worth noting. It breaks down the links of your site by industry to identify what comprises your Trust Flow. Then, you can check out the Topics tab for a more in-depth look at the exact number of external backlinks and linking domains from that industry. This can be great for scoping out PR opportunities; you can filter the chart by industry to get more specific topics. For example, filtering by "computers" gives you topics about web design, Internet resources, open source, domain names, shareware, and dozens more.

The Clique Hunter is another clever tool that you can use to see where your competitors may have the edge on their backlinks. Input up to 10 different domains or competitors to see where they are getting their backlinks and where you have missed out within your industry. Needless to say, this is another little gem from Majestic that can be used for link prospecting.

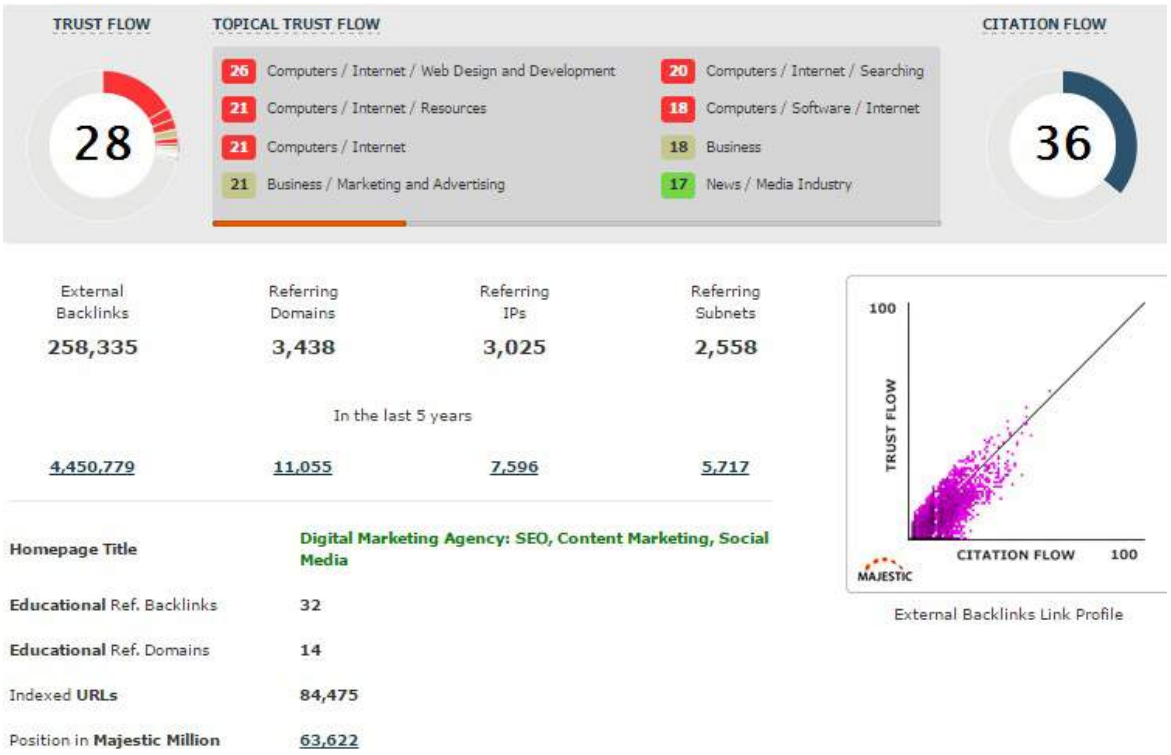


Figure 11-26. Majestic backlink report

LinkResearchTools.com

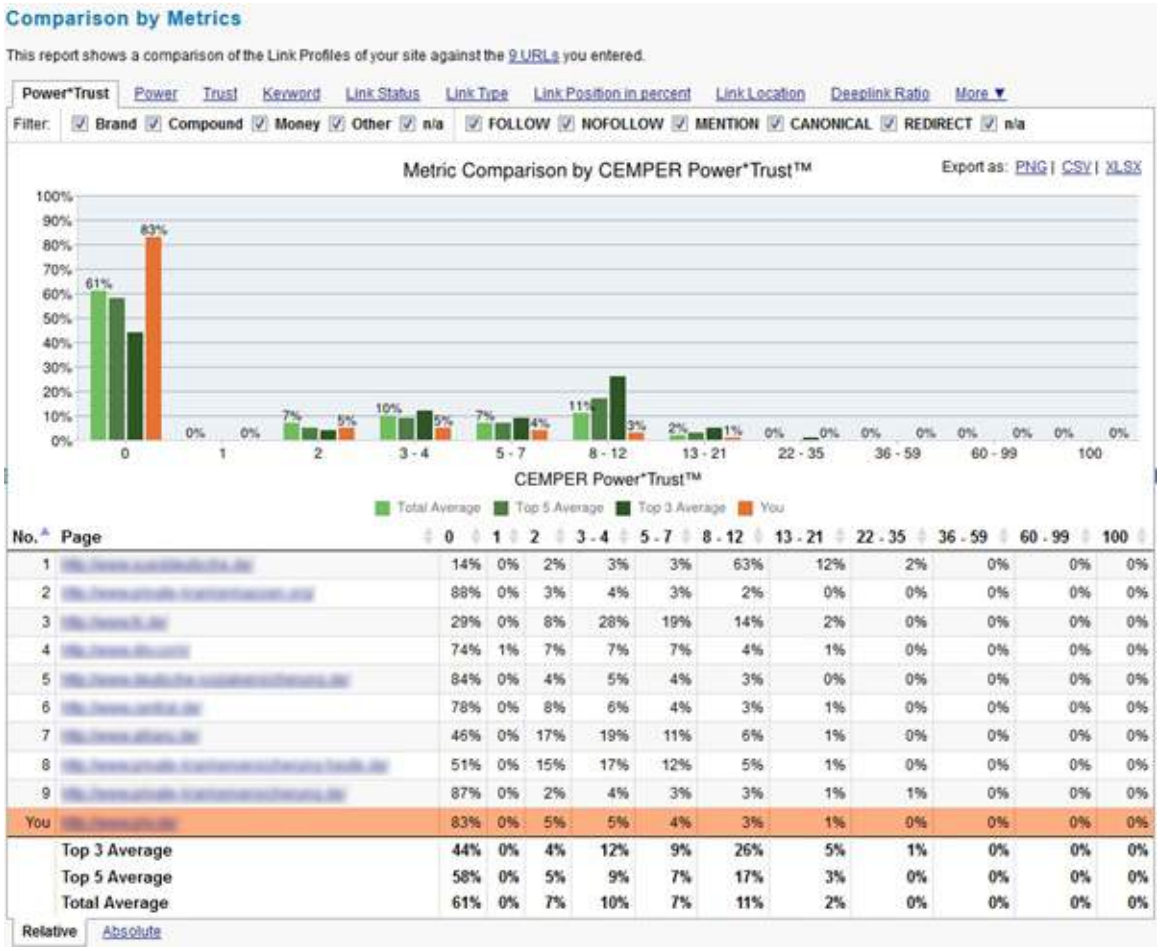
European-based Cemper.com offers a link building toolkit called LinkResearchTools. The toolkit is a report-oriented system that provides qualified link data in real time from 23 sources. Some basic SEO metrics include the number of links to the page and domain, linking root domains (domain popularity), theme of the site, and CEMPER Power* Trust™ metrics. In addition, LinkResearchTools reports can be customized to international search engine, language, and country settings. When analyzing links from a domain in such great detail, you can find potential backlinks, identify link networks, analyze competition, and discover toxic and suspicious links.

LinkResearchTools includes a number of tools for different use cases. The Quick Backlinks and Quick Domain Compare tools provide the top links of a domain/page and enable you to compare SEO metrics of other domains within a minute. The Backlink Profiler and the Juice Tool allow you to analyze link profiles for any domain. The Link Juice Thief, Missing Links Tool, and SERP Research Tool help you identify new link prospects. Link Alerts monitor link

growth for you or your competitors. The Common Backlinks Tool is a hub finder that helps you find sites that are powering your competitors. The Strongest Subpages Tool identifies a domain's strongest subpages. The Link Juice Recovery Tool helps a webmaster identify missing/broken pages with external links. The Contact Finder finds contact information for domains by searching the site for email addresses, social media profiles, or contact forms.

The Link Detox tool is primarily used to identify bad links easily and create the necessary file to submit to Google's Disavow Links tool. The data-driven organic Link Detox Genesis algorithm helps locate unnatural links by analyzing a website link profile and automatically classifying each link as toxic, suspicious, or healthy. Furthermore, the Link Detox Risk Metric assists with link risk management and provides safe search engine optimization. Webmasters should allocate the time and effort to clean toxic and suspicious links from their sites' link profiles.

Figure 11-27 shows a screenshot from the Competitive Landscape Analyzer. For competitive research, LinkResearchTools includes the Competitive Landscape Analyzer, Competitive Link Detox, Competitive Link Velocity, and Competitive Keyword Analyzer. All relevant information and SEO metrics are crawled and displayed in clear graphics. The overview gives you a qualitative statement of how your link profile appears compared to the top competitors in your niche/industry. In this figure, nine competitors are analyzed by their Power*Trust score, which measures both the Power and Trust of a site.



No.	Page	0	1	2	3 - 4	5 - 7	8 - 12	13 - 21	22 - 35	36 - 59	60 - 99	100
1		14%	0%	2%	3%	3%	63%	12%	2%	0%	0%	0%
2		88%	0%	3%	4%	3%	2%	0%	0%	0%	0%	0%
3		29%	0%	8%	28%	19%	14%	2%	0%	0%	0%	0%
4		74%	1%	7%	7%	7%	4%	1%	0%	0%	0%	0%
5		84%	0%	4%	5%	4%	3%	0%	0%	0%	0%	0%
6		78%	0%	8%	6%	4%	3%	1%	0%	0%	0%	0%
7		46%	0%	17%	19%	11%	6%	1%	0%	0%	0%	0%
8		51%	0%	15%	17%	12%	5%	1%	0%	0%	0%	0%
9		87%	0%	2%	4%	3%	3%	1%	1%	0%	0%	0%
You		83%	0%	5%	5%	4%	3%	1%	0%	0%	0%	0%
	Top 3 Average	44%	0%	4%	12%	9%	26%	5%	1%	0%	0%	0%
	Top 5 Average	58%	0%	5%	9%	7%	17%	3%	0%	0%	0%	0%
	Total Average	61%	0%	7%	10%	7%	11%	2%	0%	0%	0%	0%

Figure 11-27. LinkResearchTools report

Ahrefs

Ahrefs offers a comprehensive website analysis toolset with a strong focus on backlink analysis. Ahrefs builds its own index and crawler, which can crawl billions of pages daily and update databases with fresh backlinks found on the Web every 15 minutes. With this capability, Ahrefs is well known as having the largest index of live backlinks and covering universal results (all top-level domains included) to support users worldwide. Ahrefs also designs its own 100-scale AhrefsRank, which has a strong positive correlation value with the PageRank model for you to easily measure and compare your domain (or link) quality to your competitors.

Ahrefs Site Explorer is a link analysis tool that helps you perform granular analysis on link profiles for your or your competitor’s sites. With this tool, you

can identify where the site links are coming from, detect whether any problematic backlinks (or unnatural links) are linked to the site, determine how many backlinks are newly acquired or lost in 60 days, and explore new linkbuilding opportunities by reverse-engineering your competitors' backlinks. The SEO report allows you to rapidly evaluate on-page SEO performance of the site. This often involves a thorough examination to assess the site's accessibility to the search engines, and the SEO-friendliness of the site information architecture.

Ahrefs tools are straightforward, with a neat UI and easy-to-read data visualization (see Figure 11-28). Social metrics including Facebook shares/likes, tweets, and Google +1s are available at both the domain and link level. Filtering options are available for almost every metric, which makes the tools highly customizable. Data is downloadable in CSV format for further data sorting or uploading to third-party research tools, while the backlinks report is available in PDF format.

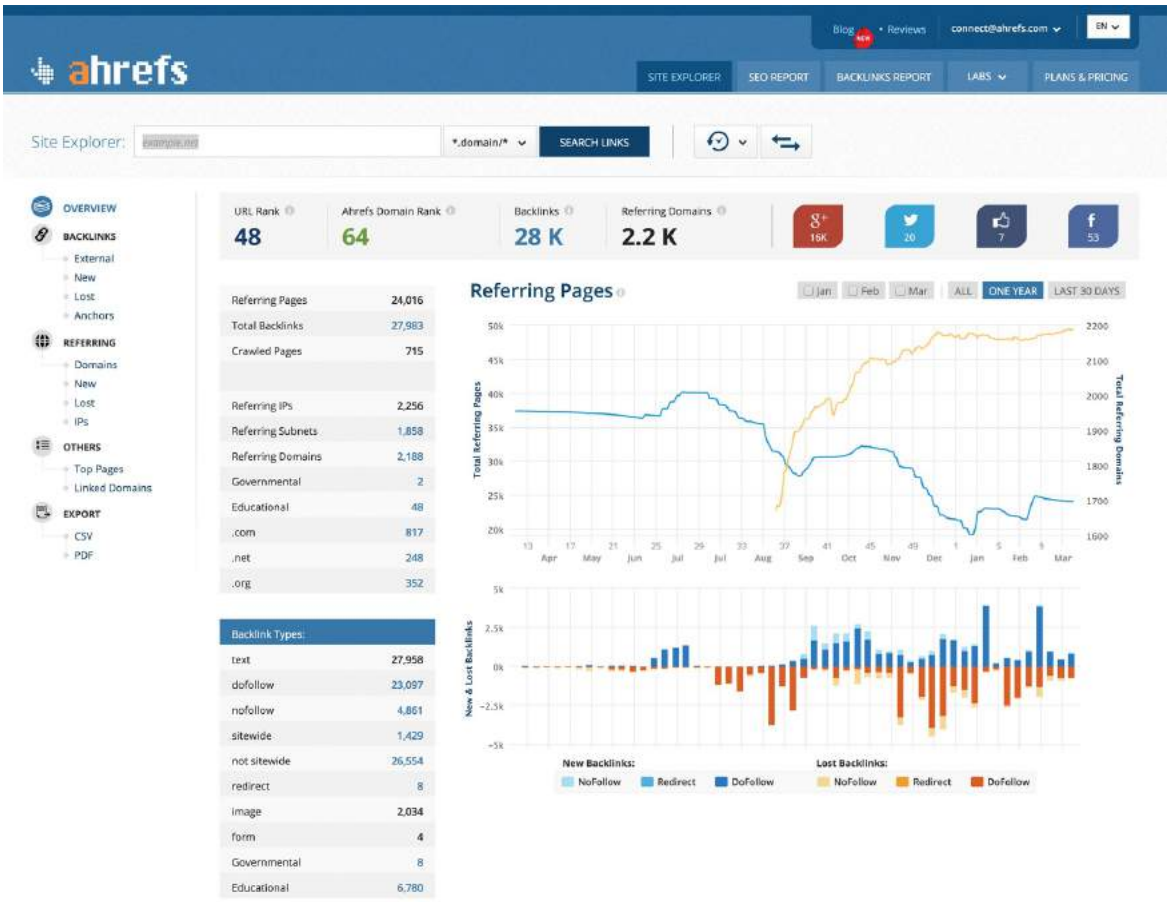


Figure 11-28. The Overview report under ahrefs

Raven Tools

Raven provides another comprehensive toolset for link development. One of Raven's tools is the Link Manager, which tracks links that you plan on acquiring, have requested, or have successfully built (see Figure 11-29). The Link Manager also includes conversion and ROI tracking. It will go out and automatically check the identified pages to see whether you have acquired a link on them, and return the results—thus automating the process of identifying targets and tracking the progress of your campaign on an ongoing basis.

Raven also offers Research Central, which mashes up data from Majestic, Open Site Explorer, and SEMRush. Research Central generates a screenshot of the site you're researching and displays results for the domain's Authority/Ranking, Social Metrics, Site Metrics, and Keywords Metrics. Raven also offers a tool

called Site Finder that allows you to enter a keyword phrase, and then the tool will automatically review the top 10 Google results for that phrase and evaluate the value of a link from each of those pages.

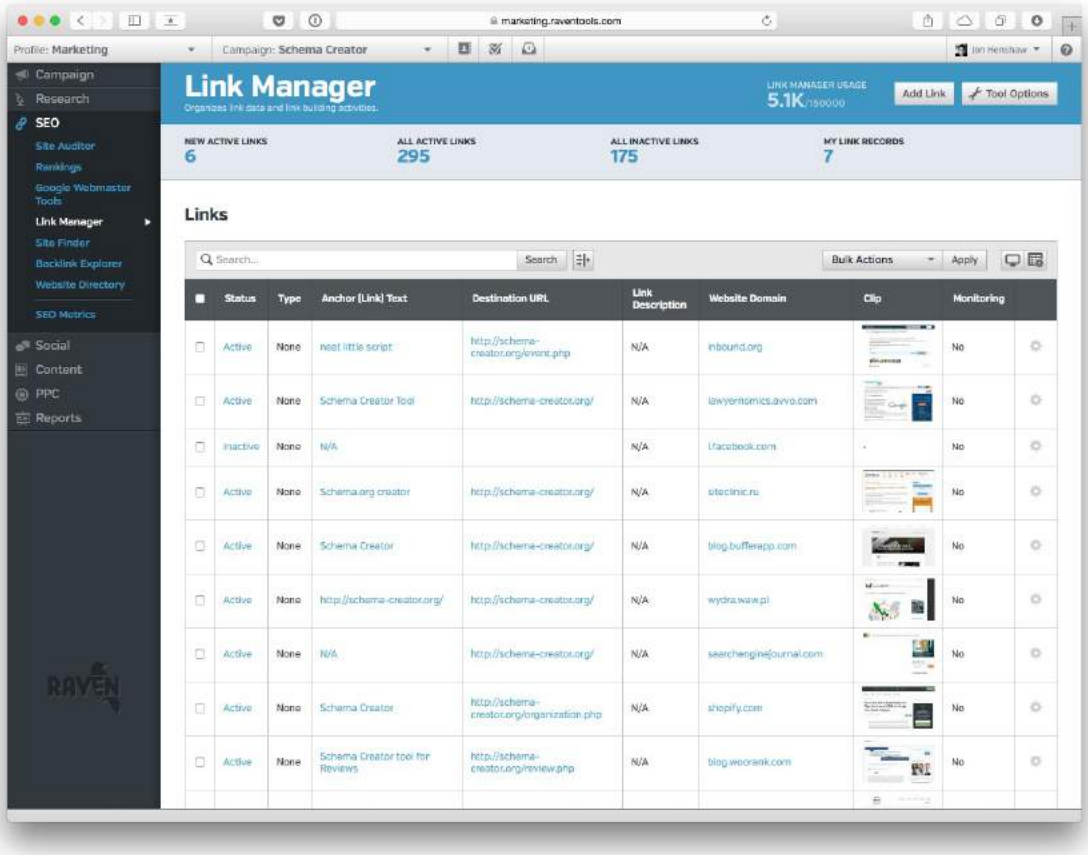


Figure 11-29. Raven Tools Link Manager

LinkBuilder and LinkGalaxy

One commercial alternative for link analytics is Rio's LinkBuilder™ solution powered by its LinkGalaxy link collection vehicle, which is geared more toward enterprise-class clients.

Link quality is much more important than link quantity, and the LinkBuilder feature provides multiple unique views to analyze and prioritize which link hubs are the most critical to obtain, in order of importance.

Figure 11-30 shows a LinkHub analysis report from the Rio LinkBuilder tool.

Figure 11-30. Rio LinkBuilder LinkHub Analysis for "toothpaste"

Conductor Searchlight

The Conductor Searchlight platform transforms a brand's web presence in "unpaid" channels (e.g., organic search, content, and social) into a powerful acquisition method resulting in meaningful traffic and revenue increases. The technology gathers data from the Web to deliver insight into where a brand's content is getting discovered, how its competition is being found, and what is needed to win the "unpaid" media war.

Conductor Searchlight makes use of the Open Site Explorer data set from Moz. A sample report is shown in Figure 11-31.

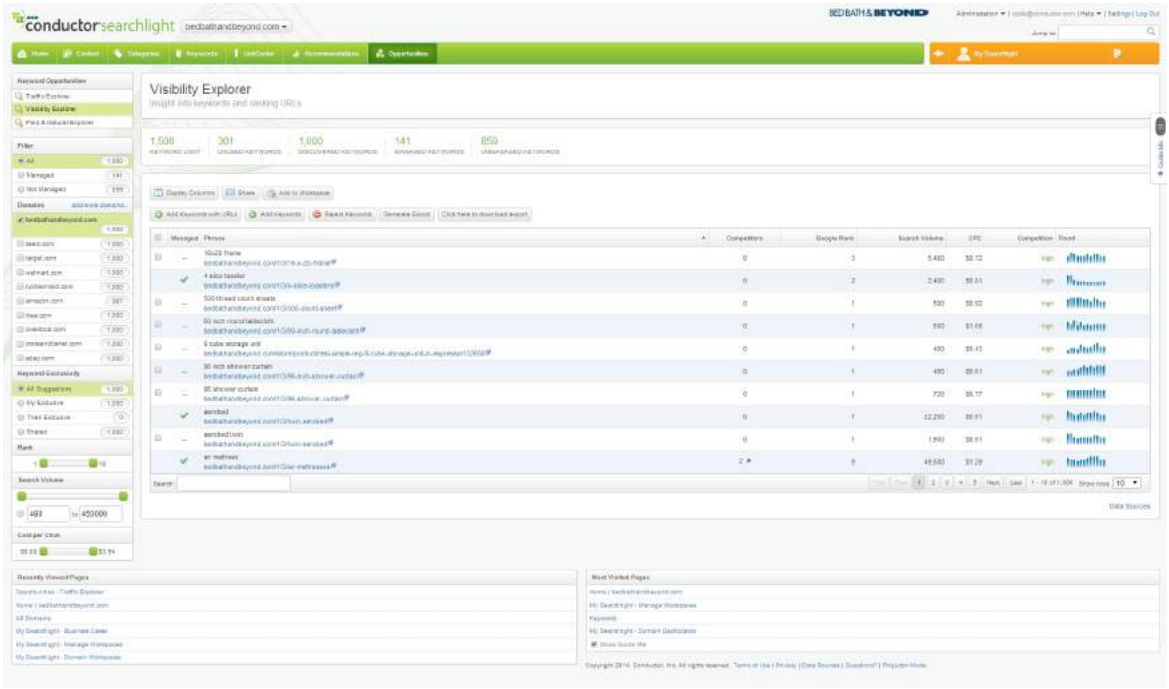


Figure 11-31. Conductor Searchlight

Exalead

Exalead has developed a European search engine that offers the capability to use a *link:* operator. What makes Exalead particularly interesting are the advanced filtering capabilities this operator provides, as you can see in Figure 11-32.

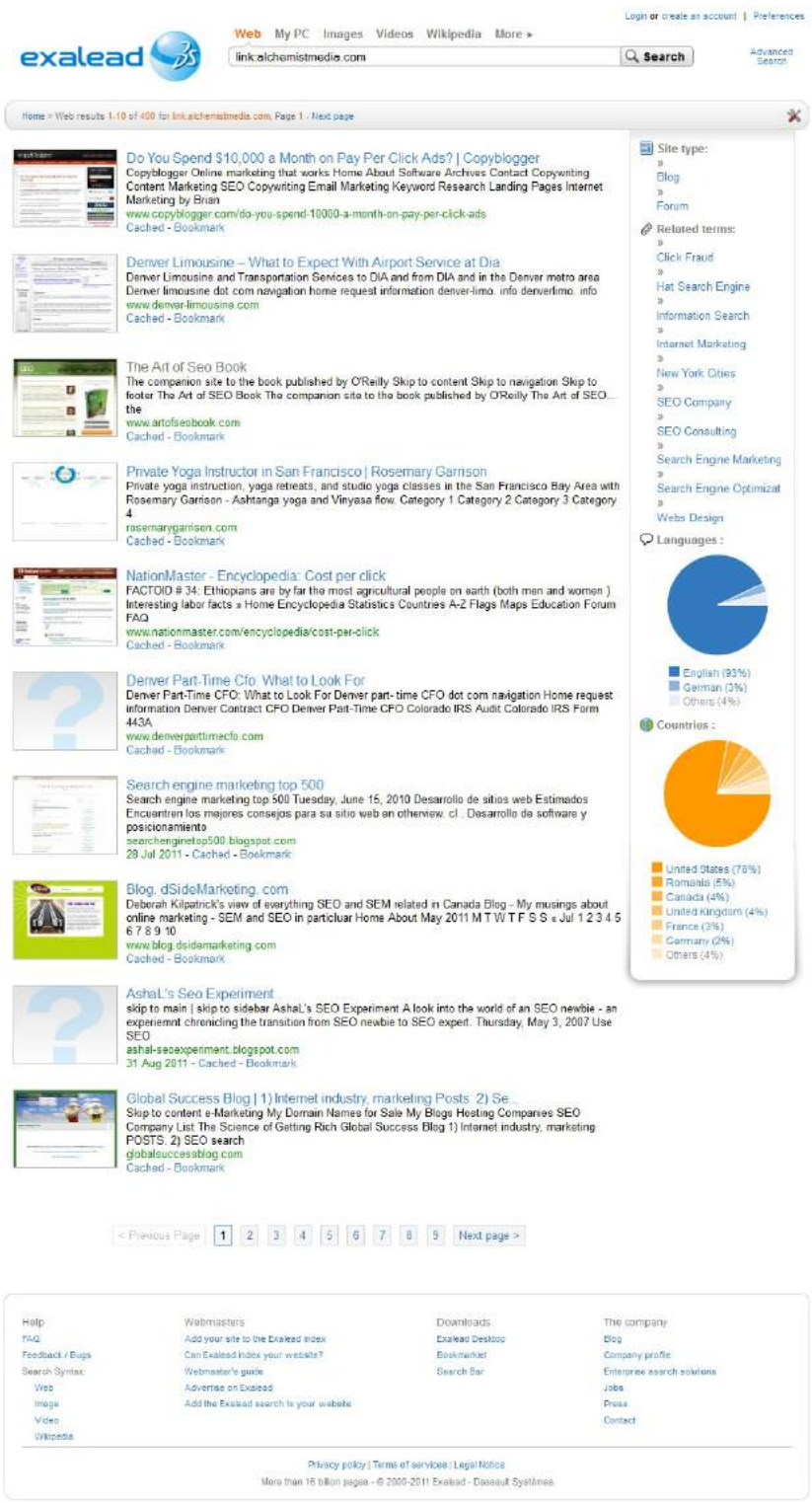


Figure 11-32. Exalead link: data

Note the ability to specify options such as language and site type, which could

be useful for segmenting link data. These filters are shown in a red box.

Measuring the value of a link

People often ask what the value of a particular inbound link is. There is no simple way to answer that question, but there are some metrics that can give you a sense of a link's importance.

These are the primary indicators of a link's value:

Where does the linking page rank for the term/phrase you want to rank for?

> If the page is ranking #1 at Google for *sliced bread* and you want to be #1 at Google for *sliced bread*, guess what? That's the #1 most valuable link you can get. Keep going down the list to about positions 25 to 30, and you're still getting solid gold in link value.

Where does the linking page rank for one to two important, competitive terms in its `<title>` *tag?*

> This will give you a very solid idea about how much overall link juice and respect the search engines are giving the page. It is also a good way to identify the global link value that could be provided.

Where does content on the linking domain generally rank for competitive terms in its pages' respective `<title>` *tags?*

> As in the preceding list item, we're trying to identify how positively the engines view pages on the domain. If the pages generally rank in the top 20 results, you can rest assured that search engines think the domain's value is pretty high, and that links from that domain will pass significant value.

How many keyword phrases do the linking domain and page rank for in the top 20 results?

> Sites that rank for a very small set of keywords may be overoptimized and thus have lower-value links, whereas sites that rank for a larger set of relevant keywords provide your site with more relevant links.

Does the linking site carry any brokered sets of links?

Sites that sell links may lose their ability to pass link authority. This really applies to any type of low-quality, manipulative linking. If you can see it, chances are that Google might see it someday too. In addition, Google may penalize a site retroactively for repeated bad behavior, even if that behavior has been stopped.

What is the relevance of the linking page/site to your target page?

Answering this question requires you to think critically about the visitors to both the potential linking page and the domain. If the relevance of the subject matter to your site is high, the link will provide more semantic and topic-specific value.

When was the last time Google crawled the page?

A fresh timestamp is a great indicator of the importance of the page. The older the page is, the less likely it is to be relevant to a modern search query. You can check a page's timestamp by looking at its cache in Google.

Inlinks to page

Next, look at the links to the specific page you want to get a link from (or perhaps that already links to you). You want to know whether the domain links into this individual page heavily, or whether it is practically an orphan. Also, the more traffic that site gets, the more likely it is that the traffic will pass on to your page. See whether it is a page that other sites reference, which can help illuminate its potential value.

The Hilltop algorithm, adopted by Google in 2003, essentially says that if your site is listed along with other relevant and quality sites within a niche on a site, the site that lists it is considered a topical authority, and your site is given more value. An example might be "10 Startups to Watch in 2014" by *Entrepreneur Magazine*. If your site is listed as one of the "startups to watch" along with other high-quality startup sites, you would benefit from the authority conferred by *Entrepreneur Magazine* and your site would be valued more highly.

These are some secondary indicators of link value:

Links to high-ranking competitors

Although this isn't always an indication of direct value, it can be a good signal. Your competitors are obviously ranking partly based on the strength of their links, so researching those sources can provide insight into where they derive that value.

PageRank of the domain

Though Google's PageRank of a domain is not a value you can get at directly, you can approximate it by looking at the PageRank of the website's home page. The home page's PageRank should give you an idea of whether the site has been penalized and of its overall link authority (i.e., the potential value of its links). You can also assess a domain's link strength using Moz's Open Site Explorer tool, which we discussed earlier in this chapter.

A PageRank 6 (or higher) home page clearly has some link love and respect, while PageRank 2 obviously has less; a gray bar can be a red flag; and a PageRank 0 tells you that the domain is either new or completely invisible. However, as linkbuilding expert Eric Ward points out, avoiding lower PageRank domains simply because they are PageRank 3 or 4 is not necessarily a good thing, as you could be missing out on very relevant links that, in volume, contribute to your relevance.

NOTE

PageRank may decline as a link quality signal over time. Google's John Mueller noted in an October 2014 Google Webmaster Central Hangout that the company had not updated the PageRank ratings available to third parties (Toolbar PageRank) for over a year—that is, since October 2013—and may well not update them again in the future.

Inlinks to domain

If a site has been picking up many new links over the past few months, it might be a much better candidate than the PageRank or other link data might indicate. This is a pretty indirect measurement, but it's not completely useless. However, the number often takes into account lots of links from a

single domain, so it can be misleading.

Number of external links on the page

Pages generally pass some of their link authority to each page they link to. Therefore, if you have a link from a page with 10 outbound links, it is far better than having a link from an equivalent page with 100 outbound links, so this is a parameter worth knowing about.

Domain Authority/Page Authority

Domain Authority and Page Authority are two of the main metrics Moz uses to estimate the link authority of the site. Domain Authority is a combination of over 40 signals intended to track the overall "strength of a website" as compared to other sites. Page Authority measures similar metrics but is intended to rank the overall strength of individual pages as opposed to the domain overall. Google has stopped updating Toolbar PageRank, so many in the industry use Domain Authority and Page Authority metrics instead of PageRank.

Citation Flow and Trust Flow

Citation Flow and Trust Flow are metrics from Majestic. Essentially, these flow metrics look not only at links to a given page, but also to the links that point at *those* links, and so on up a "chain" of links. They calculate a rating that reflects not only the number of links flowing through to a URL, but also their power and trustworthiness. Citation Flow looks at the "strength" of the chain of links leading to a URL, while Trust Flow looks at the chain's trustworthiness.

*LRT Power*Trust*

As another alternative to Google's PageRank, Power*Trust (or CEMPER Power*Trust) is a metric calculated by CEMPER's LinkResearchTools that combines the power (or authority) and trustworthiness rankings from many of the aforementioned tools into a single metric. Power*Trust is a useful metric to use to quickly gauge the overall influence of a set of links pointing to a URL. It combines data from many sources—including Google, Moz, and

Majestic—in real time.

Ranking

Many people get obsessed with checking their rankings. While rankings do have some correlation with traffic, you shouldn't spend too much time checking them. For one thing, many tools that are available today for checking rankings scrape Google results, which is against Google's Webmaster Guidelines. Google results vary by user, and the scraped results do not represent what everyone is seeing—because there is no such thing as a search result that "everyone" sees at this stage.

Rankings data can be helpful in a few scenarios, however:

- Where a #1 (or a very high) ranking for a trophy search term provides brand value
- When there is a sudden significant change (such as a move from position #5 to #50, or a similar change in the other direction)
- Where a long-term trend can be established, such as a steady improvement in position, or a steady decline
- In aggregate, where you can measure a site's overall activity and popularity

While evaluating this data, bear in mind that the search engines are constantly being fine-tuned. Google is believed to adjust and test its algorithms every day. This can result in a lot of movement in rankings on a regular basis, and this type of movement is usually meaningless from an SEO perspective.

What follows are some examples of tools that allow you to obtain rankings data. First, you can get the data directly from Google Search Console (see Figure 11-33).

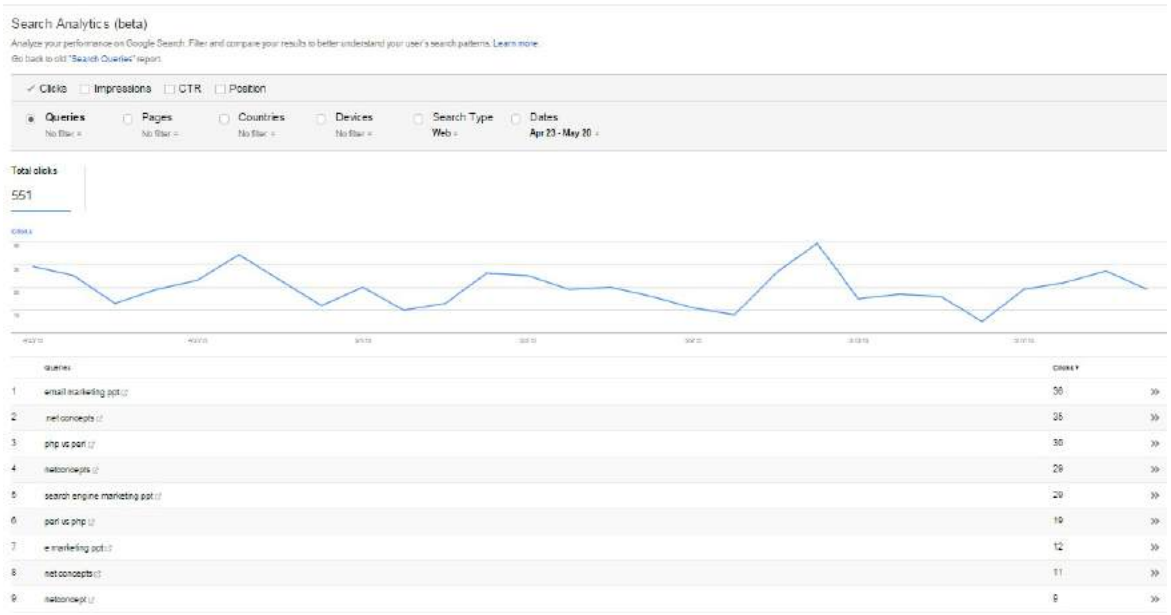


Figure 11-33. Top search queries report

Google handily shows the top queries for which your site came up, as well as the queries that were clicked on the most. This is nice data to give you a snapshot of where things stand, and you can download it into a spreadsheet.

AuthorityLabs

AuthorityLabs provides an array of reporting tools that help you track rankings across the major search engines. One reporting tool of note is its new Now Provided tool (see Figure 11-34), which is positioned as an answer to the difficulties brought on by the “not provided” update of 2013. Now Provided uses a mixture of Google Analytics data, competition data, and a few volume and ranking metrics to seek out your most important keyword opportunities. This helps you understand how much traffic a keyword is getting you. Another feature of the tool allows you to see the number of pages sending traffic to your site on any given day.

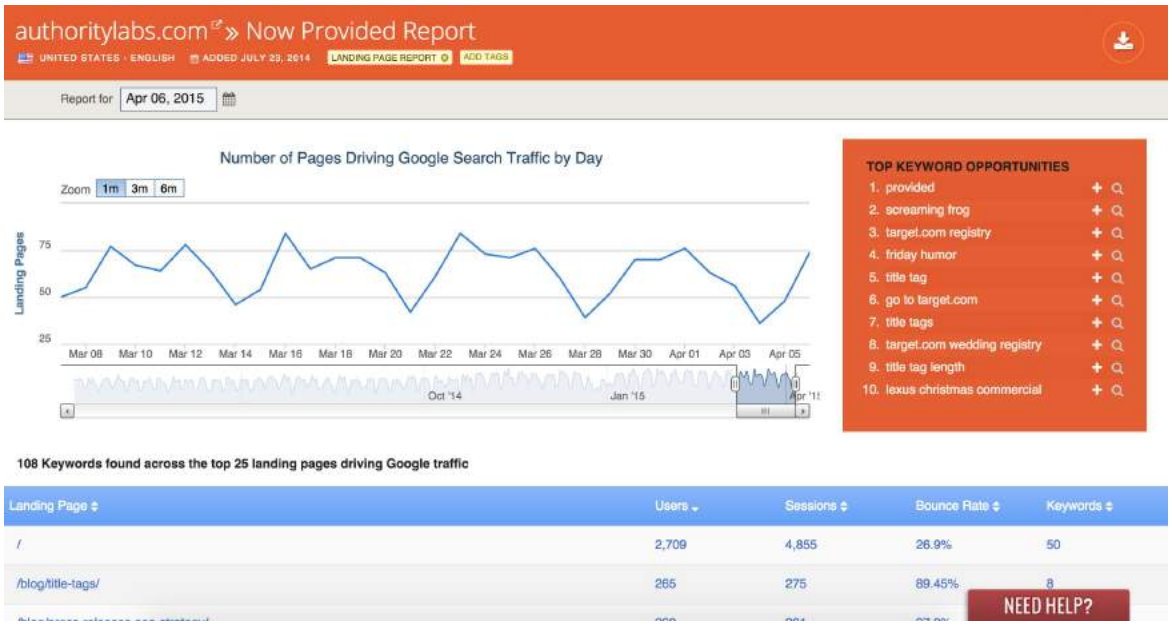


Figure 11-34. AuthorityLabs Now Provided report

Shelf space

Shelf space is an SEO metric that assigns a fraction of the total SERP to each listing—for example, 33% for the first listing, 10% for the second listing, 0.0000001% for the 123rd listing, and so on. Each point of presence yields more shelf space. Links to your site, images or videos in the results, and anything else that appears on a search result page all count toward shelf space. The sum of the shelf space for all of the listings for a particular query is 100%.

Shelf space summarizes the value of several listings on the same SERP. It allows you to easily examine the trend of your shelf space over time and compare it to that of your competitors. You can also compute your average shelf space across 1,000 keywords. If it's 10%, you know that you control 10% of those SERPs.

Here are some tips:

- When looking at SERP shelf space (or rank), treat your social media and user-generated content listings (Facebook, YouTube, etc.) as your own, including in "universal search" results.

- In reporting, draw a distinction among the shelf space controlled by your social media pages, your preferred landing pages, your other web pages, your

competitors' pages, and independent (Wikipedia, government, university, etc.) pages.

- If you care about more than a few keywords, or if you want to get a lot of long-tail traffic to your site, categorize your keywords and queries. Typical categories include branded versus nonbranded, product, subject or topic, and the geographical location mentioned in the query.

- Benchmark your SERP shelf space changes against your competitors. Then you can tell whether a change is due to something you did or due to a change in Google's algorithms.

- YouTube and Facebook are search engines, too, so you should care about your rankings there just as much as you do on Bing.

- When you're dealing with paid/organic interaction, the following simple formula estimates the cost per incremental click (CPIC; a.k.a. cost per incremental conversion) for paid search above and beyond what organic would drive to your site in the absence of paid:

  CPIC = CPC / (1 – shelf space)

  You can estimate CPIC at the query level or at the category level, as just described. Paid search marketers can optimize their campaigns to minimize their measured CPIC, truly maximizing their return, rather than choosing the ads that are best at shifting clicks from their own organic to paid listings.

SEO Platforms

Visibility into your and your competitors search rankings over time is an important indicator of your search landscape and share of search. Here are two examples of companies that offer SEO platforms:

- Conductor offers Conductor Searchlight, which can crawl your website, enable visibility into your competitive search landscape, and bring on-and off-page data (such as backlink and on-page factor analysis) together in one place.

- Rio offers SEO Search Analytics, which uses the company's patented SEO

Audit Score methodology to measure the SEO health of your websites.

Crawl Errors

Complex websites will ultimately develop errors. For example, a publisher might move a page and forget to redirect the old page to the new page, or make a mistake in the *robots.txt* file that prevents the crawling of key parts of her site. Or someone might implement a link to the publisher's site incorrectly by linking to a page that does not exist (possibly due to a typographical error).

Tools from the search engines

All of these errors are easy to diagnose in either Google Search Console or Bing Webmaster Tools. Figure 11-35 shows the Crawl Information page in Bing Webmaster Tools.

Figure 11-35. Bing Crawl Information page

You can see that Bing will provide data on 400–499 errors, 500–599 errors, 301 redirects, 302 redirects, *robots.txt* exclusion, DNS failures, and connection timeouts in this report.

Google provides similar data, with some notable differences (see Figure 11-36).

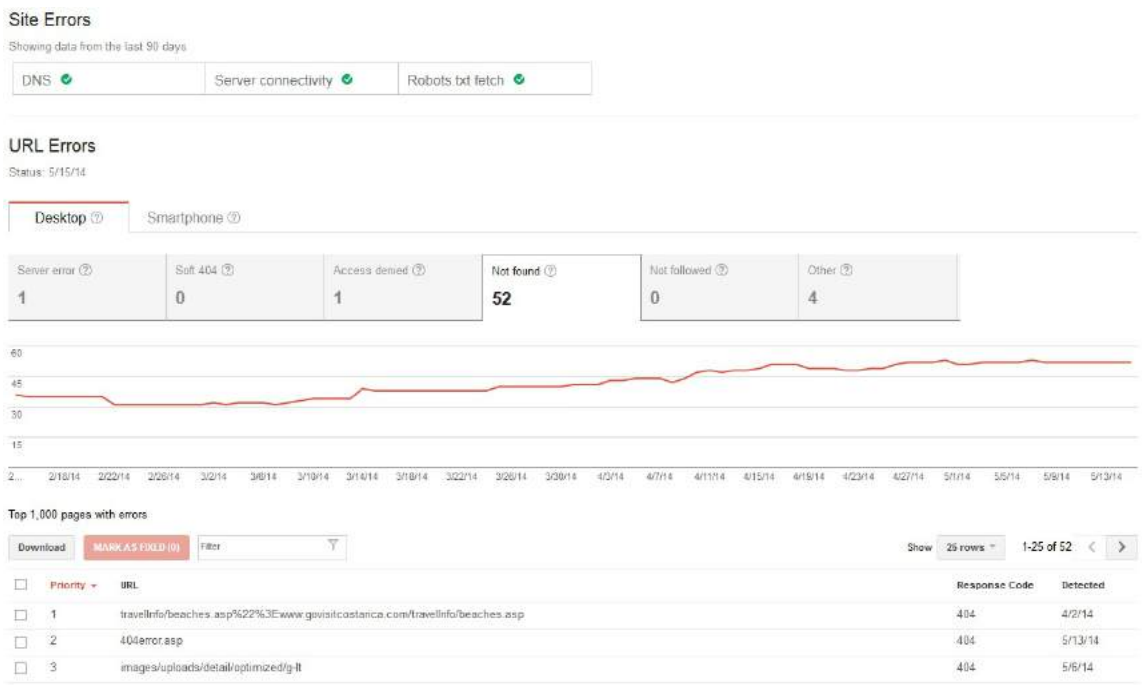


Figure 11-36. Google crawl errors

Notice the additional types of errors provided by Google. In addition, if you look at the 404 report in Google Search Console, you have the option of seeing what pages link to the page generating the 404 error, as shown in Figure 11-37.

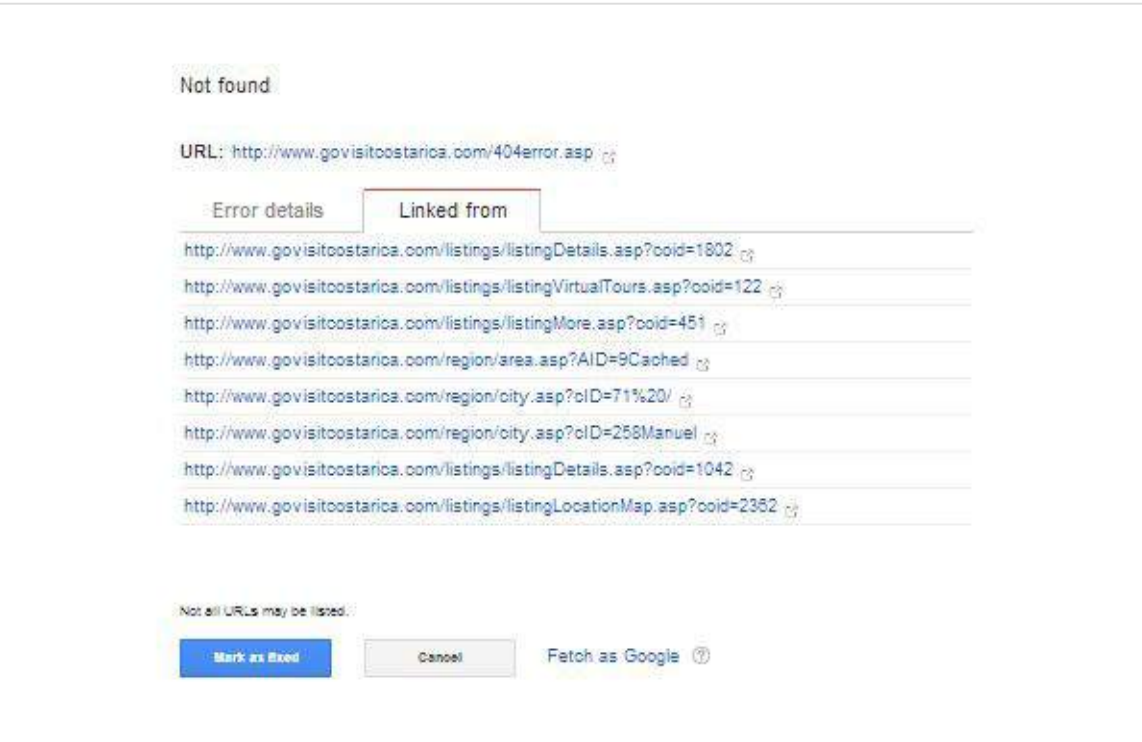


Figure 11-37. Google 404 errors report

The great thing about the 404 reports in both tools is that they can help you spot pages that have been implemented with the wrong URL on your site, or instances when someone links to one of your pages using the wrong URL. In the latter case, as we suggested in Chapter 7, you can 301-redirect the page that received the bad link to the correct page, and pick up an easy link. In short, 404 errors provide great opportunities to request updated links from sites linking to old URLs.

Another problem occurs when a page no longer exists but the web server erroneously returns a 200 status code. This is sometimes referred to as a *soft 404*, and it can occur when you've created a custom 404 landing page. In these instances, a "Page not found" error will redirect to a valid page, which results in an HTTP 200 "OK" result. This can hide the presence of broken links. A soft 404 can limit a site's crawl coverage because the landing page URL may be indexed instead of the intended (but missing) unique content pages.

So in addition to traditional HTTP 404 errors, examine your reports for HTTP 200 errors and give them the same weight as a standard 404. Google Search Console provides 404 and 200 error codes in its reports, and can be helpful for identifying broken links that need to be 301-redirected.

Third-party tools to check for crawl errors

There are some excellent third-party tools available to help you see how a search engine crawler will look at your site. One of these is SEO-Browser. This tool can be useful during a technical SEO evaluation of a site, or when you are debugging site indexation problems. Figure 11-38 shows the basic view from SEO-Browser for the TripAdvisor home page.

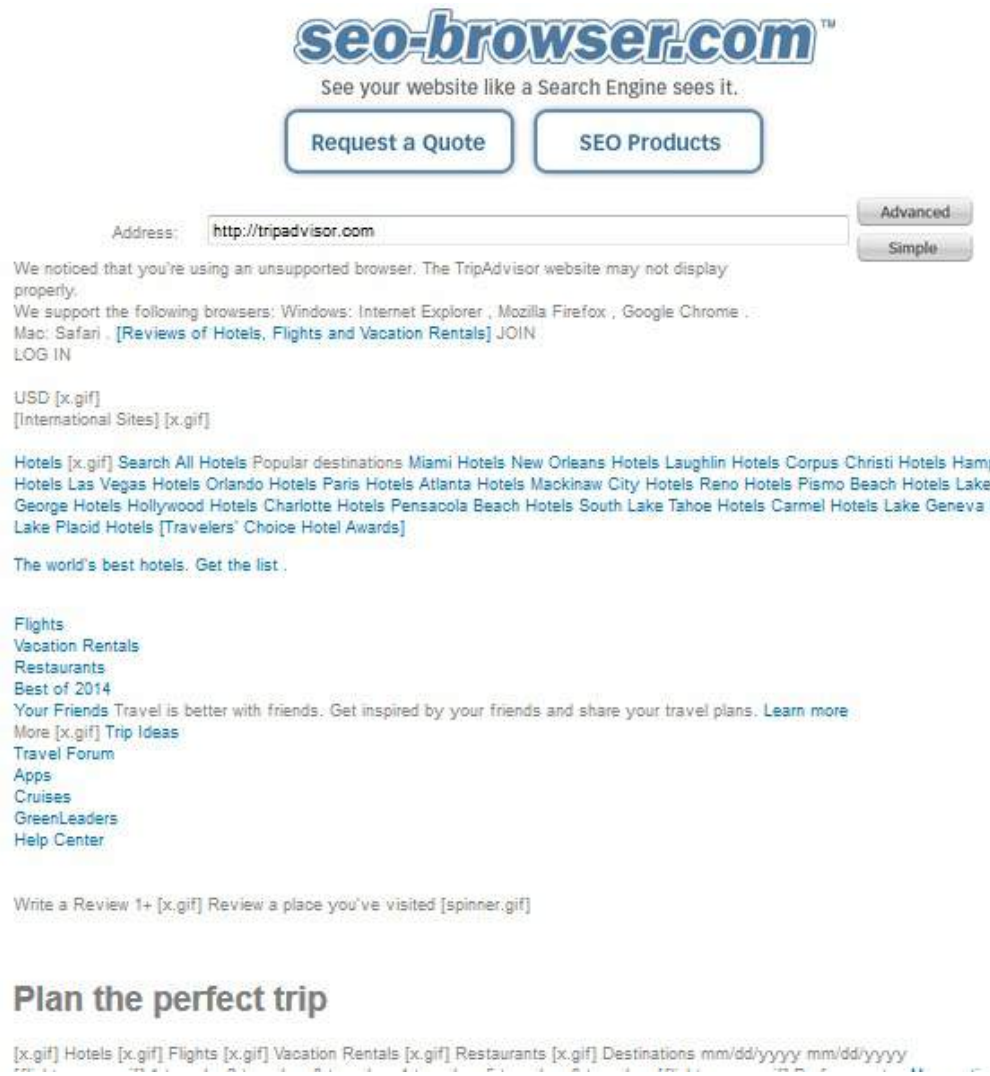


Figure 11-38. Trip Advisor home page in SEO-Browser

Interestingly, this report indicates that the website shows that the surfer used an unknown browser. It might be interesting to determine why that is happening. If you look at the Amazon home page using the Advanced mode, you see some interesting data there as well (see Figure 11-39).

Figure 11-39. Amazon home page in SEO-Browser Advanced mode

The Advanced view provides a wide range of stats above the view of the web page itself.

Another excellent tool is the Moz Crawl Test tool, which provides an extensive set of statistics on any website you choose to check. Figure 11-40 shows a small sample of the data provided.

Crawl Tes http://pro.seomoz.org/tools/crawl-test
www.staples.com
Report created: 10 Sep 04:13

URL	Time	Craw	Title Tag	Meta Desc	HTTP Stat	Referrer	Link Coun	Content-T	4XX (Clier	5XX (Serv	Title Missi	Duplicate	URLs with	Duplicate	URLs with	Long URL	Overly-Dy	301 (Perm	302 (Temp	301/302 Ti	Meta Refr	Meta Refr	Title Ele
www.stap	0		Office Sup	Shop Stap	200	www.stap	528	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1-1/2" Nor	Staples&r	200	www.stap	194	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1-1/2" Sta	Staples&r	200	www.stap	170	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1-1/2" Sta	Staples&r	200	www.stap	183	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1-1/2" Sta	Staples&r	200	www.stap	171	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1-1/2" Sta	Staples&r	200	www.stap	166	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1/2" Non-	Staples&r	200	www.stap	140	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1" Avery&	Staples&r	200	www.stap	193	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1-Hole Pa	Staples&r	200	www.stap	188	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1" Non-Vi	Staples&r	200	www.stap	189	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		1" View Bi	Staples&r	200	www.stap	228	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		100% Rec	Staples&r	200	www.stap	135	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		100% Rec	Staples&r	200	www.stap	144	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		100% Rec	Staples&r	200	www.stap	153	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		100% Rec	Staples&r	200	www.stap	129	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		Recycled I	StaplesÃ®	200	www.stap	141	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		100% Rec	Staples&r	200	www.stap	154	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		12" Chrom	Staples&r	200	www.stap	129	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		13 Watt Cl	Staples&r	200	www.stap	156	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		14" x 4" x	Staples&r	200	www.stap	144	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		15 Watt In	Staples&r	200	www.stap	136	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		16 oz. Rub	Staples&r	200	www.stap	129	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		16" x 4" x	Staples&r	200	www.stap	144	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		2" Cardina	Staples&r	200	www.stap	163	text/html	No	No	No	No		No	No		No	No	No		No		No
www.stap	0		2" Cardina	Staples&r	200	www.stap	167	text/html	No	No	No	No		No	No		No	No	No		No		No

Figure 11-40. Moz Crawl Test tool

The tool also provides a detailed look at some of the more important pages on the site, including the indexing status in the major search engines, the most important keywords, metadata, and the number of internal links to the page.

Tools such as these that provide a crawler's-eye view of your web page or website can be effective in helping you to rapidly diagnose problems with your site. Other companies that offer proprietary crawl analysis tools are:

- Rio
- Screaming Frog, shown in Figure 11-41
- DeepCrawl

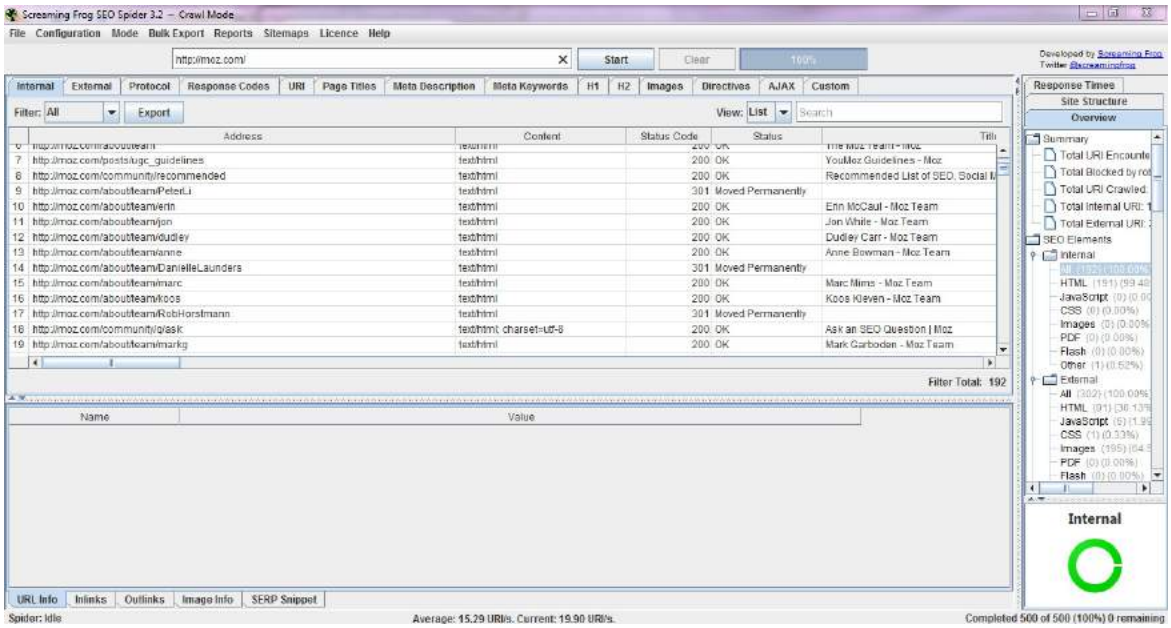


Figure 11-41. Screaming Frog tool

There are free tools that offer a lot of functionality as well. These include:

- Xenu Link Sleuth
- GSite Crawler

Tracking the Blogosphere

The blogosphere is a very dynamic part of the Web—always fresh and constantly updated. News tends to break in the blogosphere, and in Twitter, before it hits mainstream media or the Web at large. Tracking the blogosphere will help you stay current on both your industry and the SEO industry. Major influencers tend to hang out in the blogosphere as well as have their own blogs. Identifying these influencers is the first step to reaching them.

Here are four major uses for tracking blogs:

Reputation monitoring

We talked earlier in this chapter about the importance of tracking your domain and brand mentions across the Web. Reputation monitoring takes this idea one step further, where you look to identify problems and risks to your reputation as they materialize on the Web. For example, knowing when

someone starts to write negatively about your brand somewhere on the Web is important. Tracking down those issues early and addressing them quickly is a wise thing to do.

Tracking buzz and public relations campaigns

A closely related activity is tracking buzz and the response to your PR campaigns. When you make a major press push of some sort or you succeed in getting a write-up about your website on an influential blog, you should monitor the ripple effect across the rest of the Web.

Using the type of brand tracking we've already discussed can be a great way to do that. For example, did the number of mentions you received on the Web this week increase significantly over the number of mentions the week before?

Identifying potential influencers

Identifying influencers is a key part of link building. As we've discussed throughout the book, establishing yourself as a recognized expert is an important part of your content strategy. Part of that is creating world-class content, but that won't mean much unless you get the word out somehow. A great way to do that is to reach out and develop relationships with the key influencers.

Competitive analysis

If you can do this type of research for your own site, why not do it for your competitors' sites? You can see reputation problems that are developing for them, or see the impact of their media campaigns as they roll them out.

Tracking Your Blog(s)

Blogs offer numerous benefits to the online marketer. They can offer a great way to position yourself or the team behind your site as experts. For that reason, you can reach major influencers as well as a broad audience of people at the same time. This can become a nice source of links as a part of your ongoing content and link development efforts.

Blog links

There are a few ways to extract data on the number of links to your posts. You can also use either Google Search Console or Bing Webmaster Tools (see Figure 11-42) to track the number of links to your blog.

As we outlined in the discussion on link tracking, you can use these tools to get link data on your website. You can also take a more granular look at your data. For example, if you have a blog at *http://www.example.com/blog*, you can pull a report from Bing Webmaster Tools (or your tool of choice) and sort your spreadsheet on the URLs of the pages receiving the links. When you are done with this, your data might look something like Table 11-1.

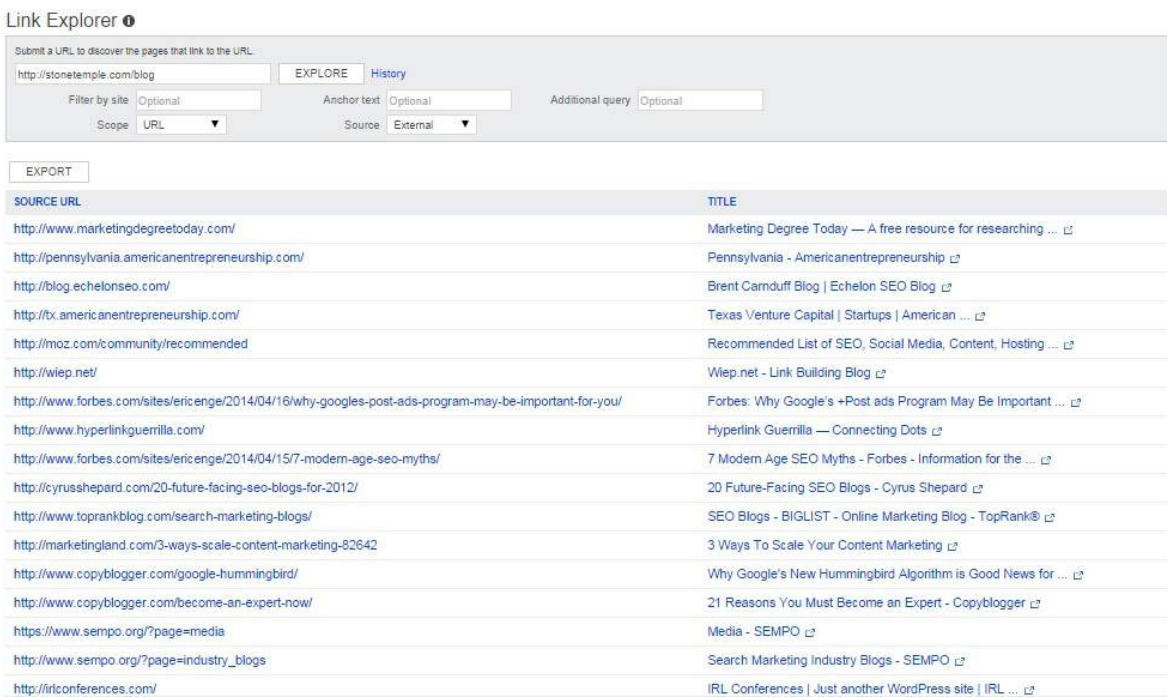


Figure 11-42. Inbound blog links crawled from Bing Webmaster Tools

Table 11-1. Sorted link report

8433	*http://www.stonetemple.com/articles/video-*	*http://www.*	7-Jul-11
8434	*http://www.stonetemple.com/articles/video-*	*http://www.*	12-Jul-11
8435	*http://www.stonetemple.com/articles/video-*	*http://www.*	4-Jul-11
8436	*http://www.stonetemple.com/articles/video-*	*http://www.*	3-Jul-11
8437	*http://www.stonetemple.com/blog*	*http://www.*	9-Jul-11

8438	*http://www.stonetemple.com/blog*	*http://www.*	21-Jul-11
8439	*http://www.stonetemple.com/blog*	*http://www.*	2-Jul-11
8440	*http://www.stonetemple.com/blog*	*http://www.*	1-Jul-11
8441	*http://www.stonetemple.com/blog*	*http://www.*	14-Jul-11
8442	*http://www.stonetemple.com/blog*	*http://www.*	6-Jul-11
8443	*http://www.stonetemple.com/blog*	*http://www.*	10-Jul-11
8444	*http://www.stonetemple.com/blog*	*http://www.*	23-Jul-11
8445	*http://www.stonetemple.com/blog*	*http://www.*	22-Jul-11
8446	*http://www.stonetemple.com/blog*	*http://www.*	25-Jul-11
8447	*http://www.stonetemple.com/blog*	*http://www.*	28-Jul-11
8448	*http://www.stonetemple.com/blog*	*http://www.*	24-Jul-11
8449	*http://www.stonetemple.com/blog*	*http://www.*	20-Jul-11
8450	*http://www.stonetemple.com/blog*	*http://www.*	17-Jul-11
8451	*http://www.stonetemple.com/blog*	*http://www.*	26-Jul-11
8452	*http://www.stonetemple.com/blog*	*http://www.*	25-Jul-11
8453	*http://www.stonetemple.com/blog*	*http://www.*	24-Jul-11

You can then find the end of the list of links that go to the blog and see how many total links you have. So, for example, if the last link is found on line 12,367 of the spreadsheet, the total number of links to the blog is 12,367 – 8,437 = 3,930. To go further, you can start analyzing what parts of the blog have the most links as well.

Google Alerts

A good way to monitor search engines' awareness of your company, brand, and products is to set up alert notifications. Use Google Alerts to get automatic emails when there are new results for the keywords you specify. Note that sometimes this "news" can be outdated and redundant.

In addition to monitoring your own web presence and reputation, it pays to keep tabs on your competitors' keywords, rankings, links, traffic, and news. Consider setting up alerts for your competitors as well as your own brands. You may even want to sign up for your competitors' email marketing messages so that you can monitor the news and offers they are sending to their customers and prospects.

Search Engine Robot Traffic Analysis

Understanding how robots are spidering their sites is another thing that expert SEO practitioners should know how to do. For one thing, spidering frequency is a clue as to which pages on your site have the highest PageRank and trust, because Google generally crawls the Web in roughly reverse PageRank order. It can also help you detect spidering problems on your site.

You can use the tools we discuss in this section to help you find potential spidering issues, and analyze how important the search engines consider your content to be. You should be looking for clues of SEO problems, such as *robots.txt* blocking the crawlers, architectural problems, or even signs of a penalty (as might be signified by a big drop in crawling frequency).

However, it is important to know that this data will not tell you everything. For example, the fact that a web page was crawled does not mean it will be placed in the index. The page will still need to pass some additional tests to achieve such placement (such as having unique content and enough links to justify its inclusion).

For pages that are indexed, you can look at how often the spiders visit/crawl your pages versus how often the engine actually shows a new version of your page in the index. To do this, look at the last cached date the search engine reported, and compare it with your crawling data.

Tools that perform log analysis include Webtrends, IBM Unica NetInsight, and Lyris HQ Web Analytics.

Other logfile analysis programs include:

- AWStats
- Sawmill
- Webalizer
- W3Perl
- Piwik

What is crawl efficiency and why is it important?

Analyzing search bot behavior involves three main goals:

- Ensuring all of your important pages are found by the bots
- Preventing PageRank from being diluted by duplicate content problems
- Getting time-sensitive content (ads/promotions, breaking news) indexed as soon as possible

Crawl efficiency is about helping search bots reach these goals. To do this, you must analyze their behavior on your site, understand the obstacles they face, and remove any roadblocks that prevent successful indexation.

Manually removing duplicate content and providing an XML sitemap are an excellent ways to address crawl efficiency problems. If the XML sitemap is comprehensive, and there is enough PageRank flowing through the site, the most important pages will be found. If the XML sitemap includes additional metadata, like how often specific pages are changed, and this metadata is accurate, then your new and time-sensitive content can be discovered quickly.

Addressing duplicate content properly is extremely important, from the perspective of both increasing crawl efficiency and improving the PageRank of canonical pages. As the PageRank consolidates, canonical page rankings will increase, and the result will be more clicks and more sales.

Google Search Console

You can also get detailed information about spidering activity on your own website using Google Search Console. Figure 11-43 shows the statistics for one site.

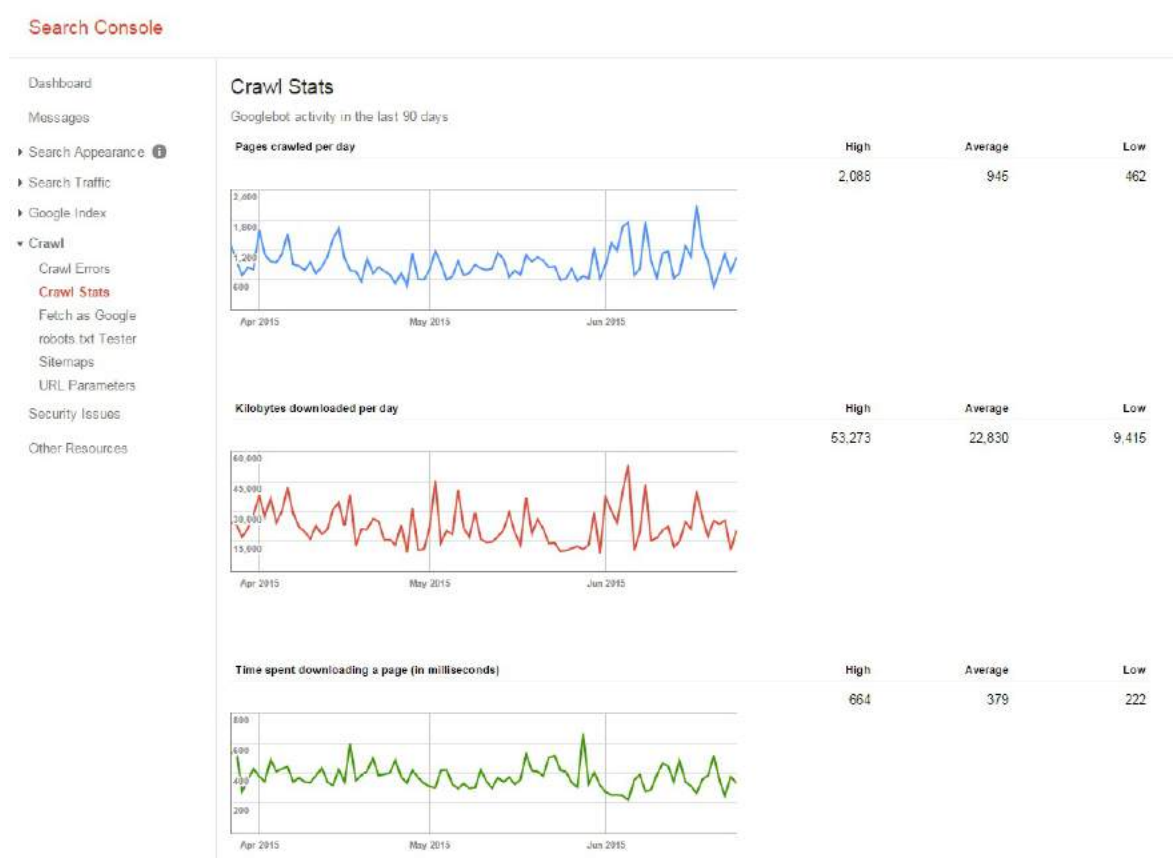


Figure 11-43. Google Search Console spider activity report

This report provides a great visual snapshot. It is most valuable when you see major changes in the data. For example, if you change your hosting provider, and the page download time subsequently jumps upward, this may be an indication of a problem with the new hosting company. When you see these types of changes, it can be a flag that something happened with your website, and you should investigate it.

Web Traffic Comparison

When you set your goals for growing your online business, you may, for example, target a sales increase of 50% from search referral traffic. However, in setting those goals you may also want to take into account where your competitor is relative to you. If your competitor has twice your traffic, for example, you may not be happy with a goal to increase yours by 50%.

In addition, when you measure the results of a completed SEO campaign, it is

useful to see how much your competitor grew during the same timeframe, because you can use this data to set your goals for your next campaign. For example, if you grew your traffic by 50% and your competitor grew by 100%, you may want to adjust accordingly.

Several tools offer ways to compare the traffic of different websites. However, each tool has accuracy issues and significant bias.

For example, Alexa relies on data obtained from users who have the Alexa toolbar on their systems. This is a limited number of users, and of course, only certain types of people are willing to have the toolbar on their system. This introduces a bias to the data, which means it is not completely reliable. In addition, you can get software to artificially inflate your Alexa rankings by spoofing the Alexa toolbar. This is not something you should do, but you should understand the issues with these types of measurement systems.

Other companies that offer traffic measurement services, such as Compete and Quantcast, suffer from similar problems with bias and limited sample sites. Note also that all of these data sets are limited to the United States. However, these tools still have value when you are comparing the traffic of one site to another in the same market space (perhaps you and your competitor) because the bias should affect both sites equally.

Alexa

Alexa provides you with a quick and easy look at how the traffic for two sites compares, as shown in Figure 11-44.

Alexa
The Web Information Company

Register Login Help

Home | Site Tools | Top Sites | Site Info | What's Hot | Toolbar

Search for more Search

searchenginewatch.com

Search Engine Watch

Is this your site? Run an Alexa Site Audit.

About Search Engine Watch (searchenginewatch.com): Search Engine Watch is the authoritative guide to searching at Internet search engines and search engine registration and ranking issues. Learn to submit URLs, use HTML meta tags and boost placement.

Statistics Summary for **searchenginewatch.com**

Searchenginewatch.com is ranked #1,872 in the world according to the three-month Alexa traffic rankings. The site is in the "Searching" category of websites. The site has been online for more than...

Show More

Alexa Traffic Rank		Reputation	
1,872 Global Rank	1,398 Rank in US	16,443 Sites Linking In	5 Reviews

Traffic Stats | Search Analytics | Audience | Contact Info | Reviews | Related Links | Clickstream

Traffic Rank | Reach | Pageviews | Pageviews/User | Bounce % | Time on Site | Search %

Daily Reach (percent)
searchenginewatch.com
searchengineland.com

Estimated percentage of global internet users who visit searchenginewatch.com:

	Reach	Change
Yesterday	0.10900	-4%
7 day	0.08500	+0.9%
1 month	0.08540	-9%
3 month	0.09220	+18.7%

Compare **searchenginewatch.com** to:

searchengineland.com | Compare

Learn more about Alexa Traffic Stats.

Share this: facebook Like Tweet +1

Average Load Time for Searchenginewatch.com

Very Slow (4.513 Seconds). 94% of sites are faster.

Note: Slow sites may be penalized by search engines.

Searchenginewatch.com's Regional Traffic Ranks

Country	Rank
India	656
United Kingdom	719
Pakistan	892
Philippines	1,007
Australia	1,229

More

Where Visitors Go on Searchenginewatch.com

Subdomain	Percent of Site Traffic
searchenginewatch.com	86.61%
forums.searchenginewatch.com	11.52%
newsletters.searchenginewatch.com	0.59%
jobs.searchenginewatch.com	0.53%
cms.searchenginewatch.com	0.38%

More

Audience Snapshot

Based on internet averages, searchenginewatch.com is visited more frequently by users who browse this site from work.
Get complete site demographics.

View the full Audience Profile

Top Search Queries for Searchenginewatch.com

	Query	Percent of Search Traffic
1	meta tags	2.21%
2	search engine watch	1.80%
3	google panda update	1.77%
4	google panda	1.53%
5	kudlie ball	0.50%
6	link building strategies	0.46%
7	google guitar	0.36%

View the complete Search Analytics

Featured Alexa Site Tools

Site Audit
If you want to improve your SEO techniques, increase your organic search traffic, and ensure complete site analytics tag coverage, this tool is for you.

Toolbar Creator
Generate your own custom toolbar and promote it to your visitors to increase engagement and generate more frequent visits to your site.

Edit Your Listing
Update your listing on Alexa, including your site description, contact information, and more.

View All
Access all of Alexa's Site Tools.

Alexa Toolbar Creator
Create a Toolbar!

GROW YOUR BUSINESS ONLINE
Degrees in:
Internet Marketing
Mobile Development
Web Design
FULL SAIL UNIVERSITY. LEARN MORE
Advertise with Alexa

High Impact Search Queries for Searchenginewatch.com

	Query	Impact
1	meta tags	High
2	search engine watch	Medium
3	google panda update	Medium
4	google panda	Medium
5	search engine	Medium
6	link building	Medium
7	panda update	Medium

View the complete Search Analytics

Top contact lens brands on Sale!
Acuvue Oasys $22.99(USD) BUY NOW
Air Optix Aqua $29.89(USD) BUY NOW
Air Optix Night & Day Aqua $49.89(USD) BUY NOW
Biofinity $29.49(USD) BUY NOW
DAILIES AquaComfort Plus 30
CoastalContacts.com

Top Search Queries for Searchenginewatch.com

	Query	Percent of Search Traffic
1	meta tags	2.21%
2	search engine watch	1.80%
3	google panda update	1.77%
4	google panda	1.53%
5	kudlie ball	0.50%
6	link building strategies	0.46%
7	google guitar	0.36%

View the complete Search Analytics

Wayback Machine
See how Searchenginewatch.com looked in the past

Like searchenginewatch.com? Download the Alexa toolbar and access exclusive analytics content.

您现在可以在Alexa中文官方网站cn.alexa.com上获取searchenginewatch.com的网站流量信息。

NETFLIX Instantly watch TV shows & movies FREE TRIAL Click here >

Company | Help | Alexa Tools for Site Owners | Related Services

Download the Alexa Toolbar
Privacy Policy | Terms of Use
© Alexa Internet, Inc.
An amazon.com company

Figure 11-44. Alexa traffic comparison

Underneath the chart you will notice a lot of additional data on SearchEngineWatch.com (the site listed first in the search box), such as the most popular content and where the readers are coming from.

Compete

Compete offers both a free and a paid service. As with Alexa, you can get basic traffic data on a site or compare traffic between sites (see Figure 11-45).

Figure 11-45. Compete traffic comparison

Compete also offers additional data for a fee. This includes data on the major keywords driving traffic to a site (see Figure 11-46).

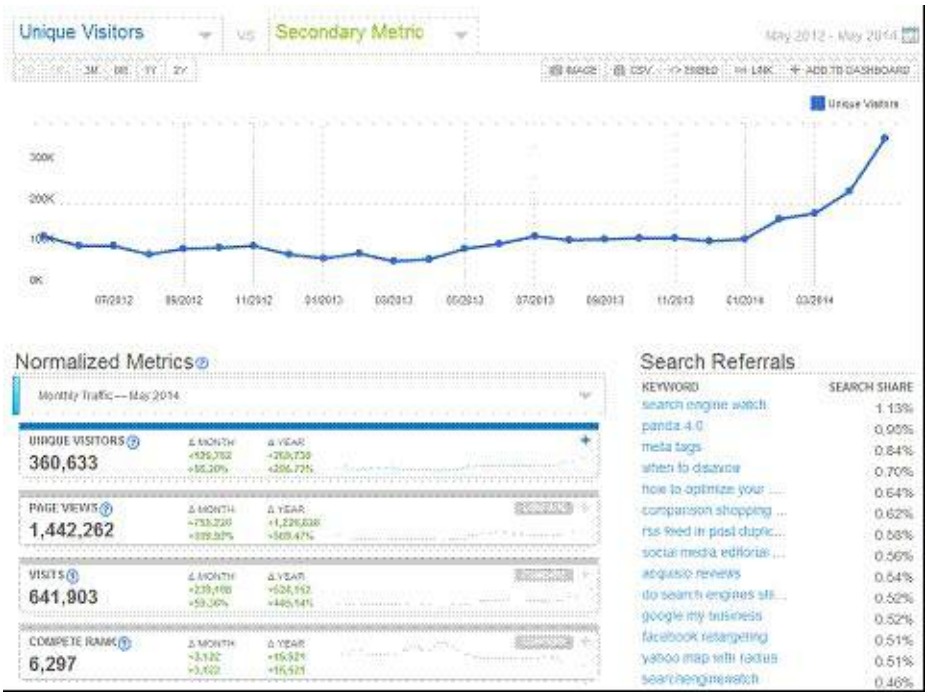


Figure 11-46. Compete keyword data

Compete leverages multiple sources, including ISP, panel, and toolbar data, and has a total panel of about 2 million users. Compete's diverse data sources help in reducing bias in the data.

Quantcast

Quantcast is a competitor of the other services we've listed. With Quantcast, you can get detailed data about individual sites.

The demographic data is interesting, as well as the other sites that users visit. Quantcast collects "pixel data" from scripts running on websites that are part of the program. It claims that these sites represent about 50 million visitors per month. Quantcast supplements this data with data from ISPs. Figure 11-47 shows a sample report for USA.gov.

Quantcast also offers publishers an opportunity to participate in its Quantcast Advertise for Publishers program. This program enables publishers to make public much more accurate data through Quantcast and requires a tracking tag to be placed on their sites. Publishers who sell advertising on a cost-per-impression basis can use this as third-party validation of their traffic levels.

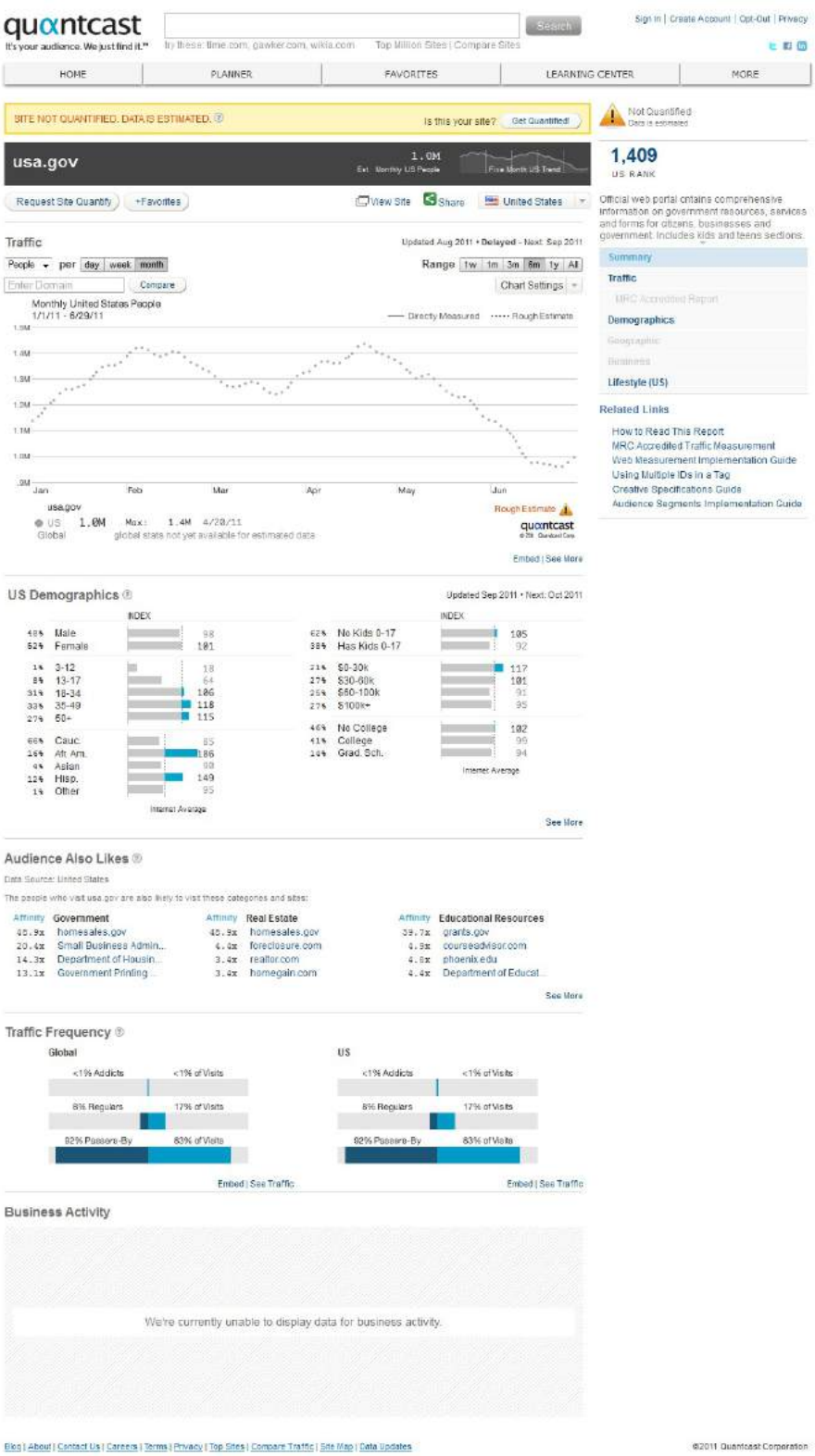


Figure 11-47. Quantcast site data

Temporal Link Growth Measurements

Since Google released its patent application on historical information in March 2005 (read more at *http://moz.com/blog/google-patent-on-historical-data*), search marketers have recognized that trends in temporal link and content analysis do have a real impact on rankings.

The engines are trying to measure patterns—they're looking for indications of increasing or decreasing relevance and authority that temporal trends provide. They want to identify several specific items:

Content growth patterns

How often does a particular site tend to add new pages?

Content update patterns

How often are documents edited and updated?

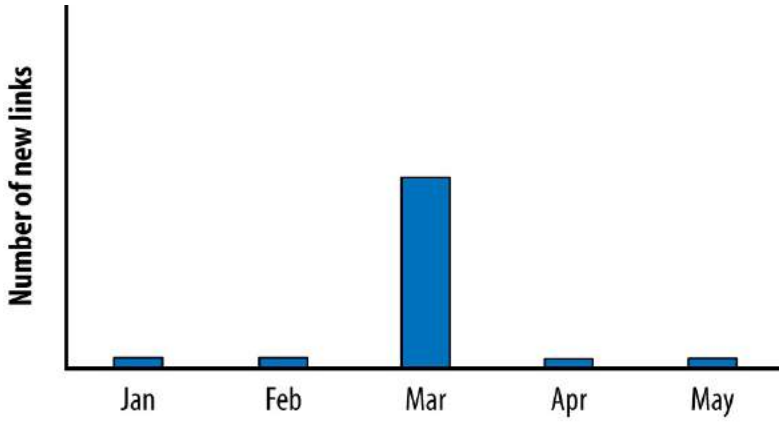
Link growth patterns

How often do new links appear pointing to the site?

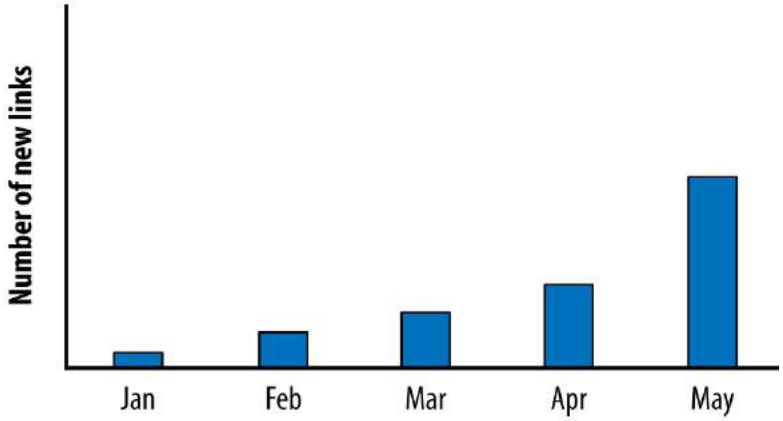
Link stagnation patterns

Does the number of links to the site stagnate or decrease?

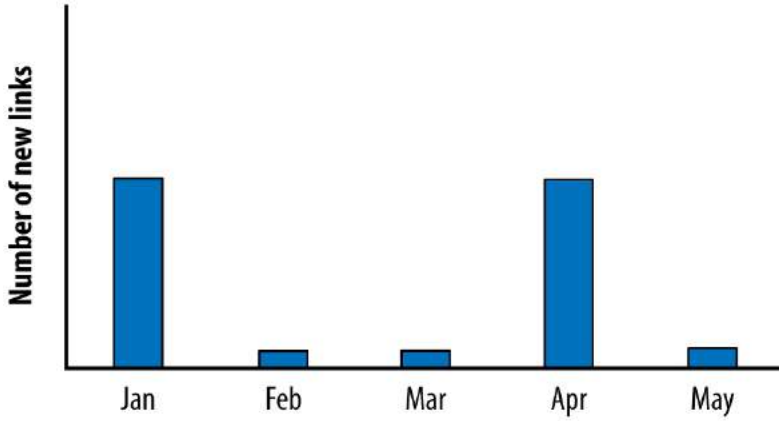
The engines aren't interested only in how many links pointed to the site today versus yesterday (or how many pages were added); they are also fundamentally interested in tracking patterns over time. Figure 11-48 and Figure 11-49 depict some example graphs showing the rate of new external links (and in the last two instances, pages) created over time, with some speculation on what the trends might indicate.



Linkbait? A big news story? A large link purchase? A sensation in the blogosphere?

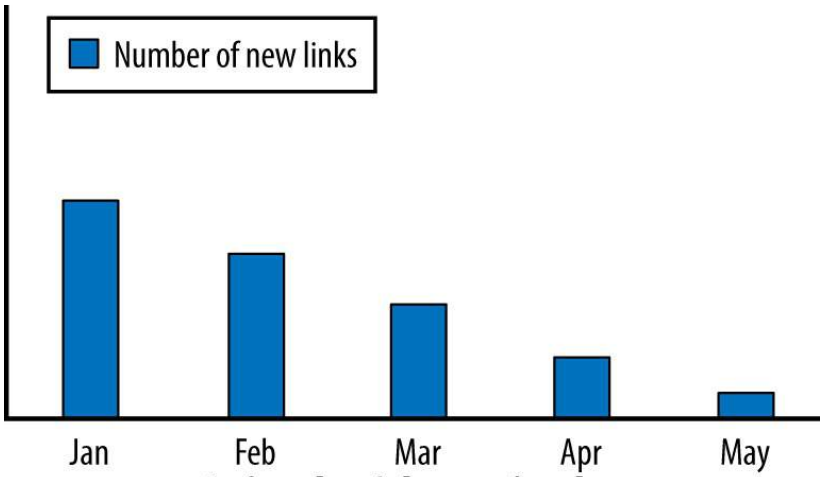


A site that's getting more popular over time and growing in reach?

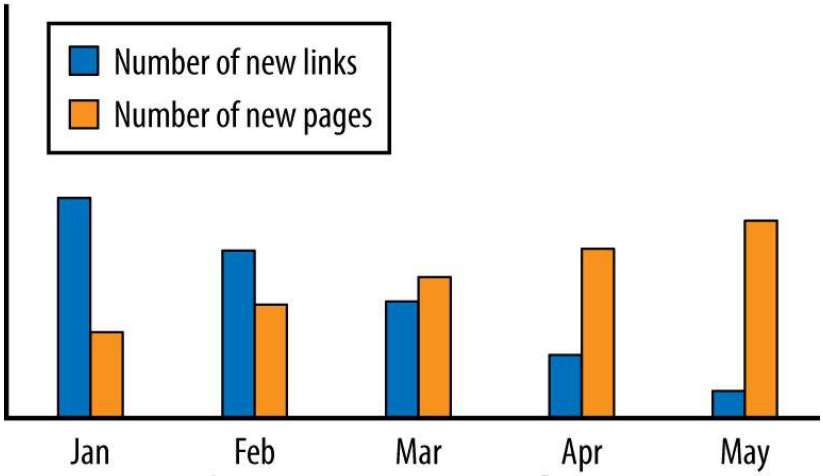


Inconsistent link trends? Cyclical relevance?

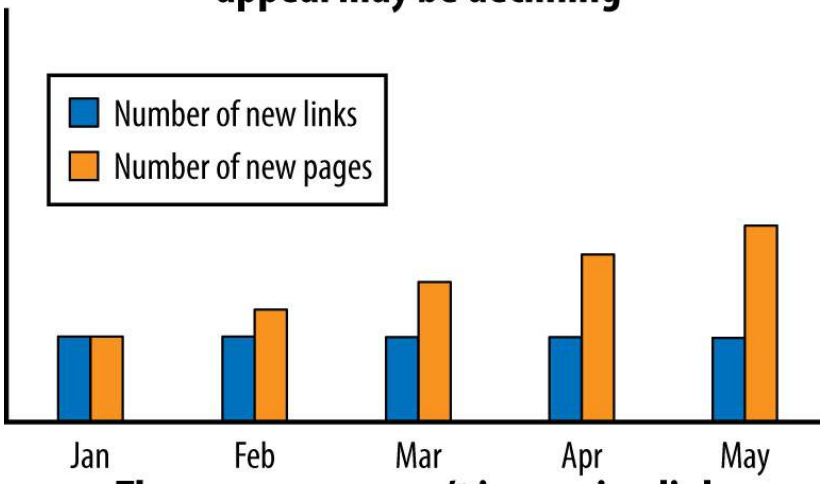
Figure 11-48. Interpreting new external link data



A site that's becoming less relevant, interesting, or worthwhile?



Despite more content, relevance or appeal may be declining



The new pages aren't increasing link numbers, a potential sign of staleness

Figure 11-49. More link data speculation

These assumptions do not necessarily hold true for every site or instance, but the graphs make it easy to see how the engines can use temporal link and content growth information to make guesses about the relevance or worthiness of a particular site. Figure 11-50 shows data of a few real sites and how these trends have affected them.

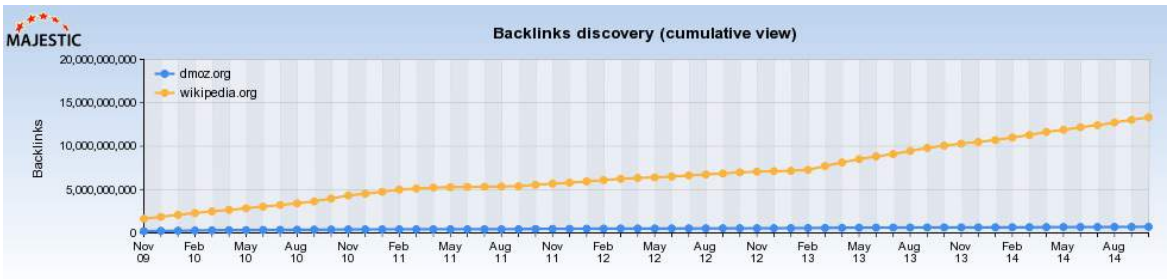


Figure 11-50. DMOZ versus Wikipedia backlink discovery

As you can see in Figure 11-50, Wikipedia has had tremendous growth in both pages and links from 2009 through 2014. This success manifests itself in the search engines, which reward Wikipedia's massive link authority with high rankings for much of its content.

DMOZ was once a default reference link for many sites, but its relative influence has waned. Although it continues to grow, Wikipedia simply trumps it immensely. Figure 11-51 shows backlink domain discovery indicating that Wikipedia not only has breadth of links, but also of linking domains.

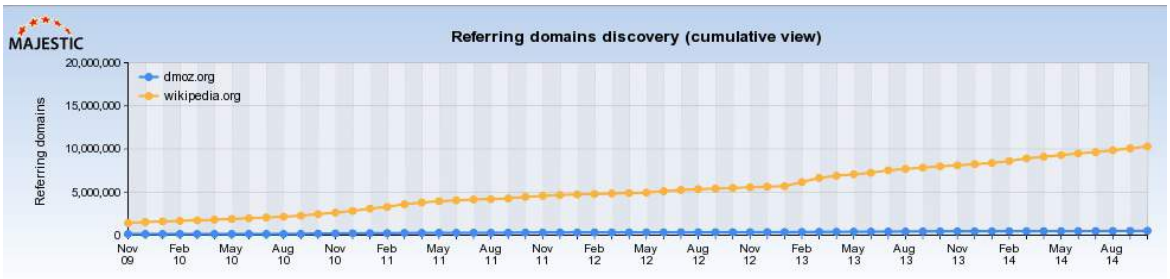


Figure 11-51. DMOZ versus Wikipedia referring domains

Many forms of spam and manipulative link building are likely to stand out like a

sore thumb when put under the temporal microscope. When there's a large gain in links relative to a site's sphere, influence, and historical link growth, the engines can take a closer look at the source of the links or even trigger a manual review. Common sense would dictate that a small-time local real estate site doesn't usually attract a few thousand new links in a week unless it has done something newsworthy or link-worthy.

There are few limits to what the engines can do with such data, and there is no reason they shouldn't be analyzing it (it is readily available, after all). Do consider how the link and content growth patterns for your sites may affect the engines' perspectives on your rankings and trustworthiness.

Key Performance Indicators for LongTail SEO

As we have discussed throughout the book, the long tail is an important part of SEO. Here are some available metrics for diagnosing the health of your long-tail search traffic:

Branded-to-nonbranded ratio

This is the percentage of your natural search traffic that comes from branded keywords versus nonbranded keywords.

If the ratio is high and most of your traffic is coming from searches for your brand, this signals that your SEO is fundamentally broken. The lower the ratio, the more of the long tail of natural search you likely are capturing. This metric is an excellent gauge of the success of your optimization initiatives. What constitutes a high ratio is going to vary depending on your brand's level of recognition and how web savvy your users are. Certainly anything over 50% brand searches warrants closer examination.

Unique crawled URLs

This is the number of unique (nonduplicate) web pages crawled by search engine spiders such as Googlebot and BingBot. Your website is like your virtual sales force, bringing in prospects from the search engines. Think of each unique page as one of your virtual salespeople. The more unique pages you have, the more opportunities you have to sell through the search engines.

Search visitors per contributing page

This is the percentage of unique pages that yield search-delivered traffic in a given month.

This ratio is a key driver of the length of your long tail of natural search. The more pages yielding traffic from search engines, the healthier your SEO program. If only a small portion of your website is delivering searchers to your door, most of your pages—your virtual salespeople—are warming the bench instead of working hard for you. You can think of these nonperforming pages as “freeloaders.”

Keywords per page

This is the average number of keywords driving traffic to each page (minus the freeloaders) in a given month. Put another way, it is the ratio of the total number of keywords driving search traffic to pages receiving search traffic.

The higher your keywords per page, the more of the long tail of natural search your site will capture. In other words, the more keywords each page attracts or targets, the longer your tail. So, an average of eight search terms per page indicates pages with much broader appeal to the engines than, say, three search terms per page. The average online retailer in a Netconcepts study on the long tail of natural search) had 2.4 keywords per page. A larger number of keywords per page should also indicate that your content is more varied in its keyword usage and that it is better meeting the needs of a wide variety of users (which the search engines will recognize).

Search visitors per keyword

This is the ratio of search engine–delivered visitors to search terms. This metric indicates how much traffic each keyword drives and is a function of your rankings in the SERPs. Put another way, this metric determines the height or thickness of your long tail.

The average merchant in the aforementioned Netconcepts study obtained 1.9 visitors per keyword.

Index-to-crawl ratio

This is the ratio of pages indexed to unique crawled pages. If a page gets crawled by Googlebot, that doesn't guarantee it will show up in Google's index. A low ratio can mean your site doesn't carry much weight in Google's eyes.

Search visitors per crawled page

Calculated for each search engine separately, this is how much traffic the engine delivers for every page it crawls.

Each search engine has a different audience size. This metric helps you fairly compare the referral traffic you get from each. The Netconcepts study found that Bing and Yahoo! tended to crawl significantly more pages, but the yield per crawled page from Google was typically significantly higher.

As you optimize your site through multiple iterations, watch the aforementioned KPIs to ensure that you're heading in the right direction. Those who are not privy to these metrics will have a much harder time capturing the long tail of SEO.

Duplicate Content

Duplicate content—one or more pages with multiple unique URLs—is, in effect, making your own long tail unnecessarily longer. You only want one canonical URL for each page of content, and you don't want that content to be repeated verbatim on other pages or sites.

Once you find duplicate content, figure out which page and URL you want to be the canonical resource (typically the ones with the highest PageRank), and 301-redirect the others to it.

Some analytics packages and independent SEO tools can track duplicate content, such as SEObook.com's Duplicate Content Checker.

Other Third-Party Tools

Throughout this chapter, several third-party tools have been mentioned or

recommended for analysis, research, and testing. This section outlines a few more worth noting.

MozBar

As previously mentioned in this chapter, the MozBar provides at-a-glance link and social sharing metrics, but it also offers much more. Through a series of available drop-down menus (see Figure 11-52) the MozBar can provide information on on-page elements, general attributes, in-depth link metrics, semantic markup, and HTTP status of the page. Some of the on-page elements it details are the URL, `<title tag>`, meta description, meta keywords, heading tags, bolded text, italicized text, and `alt` text for images on the page. Some of the "general attributes" it details are the content of the meta `robots` tag, the `rel="canonical"` link element, the page load time, a link to the URL for Google's cache of the page, the IP address, and the country the site is perceived to be located in. It offers a wide variety of link metrics for the current URL, the current subdomain, and the current root domain including Page Authority, Domain Authority, External Followed Links, Linking Root Domains, MozRank (mR), and MozTrust (mT). The tool also shows whether various types of markup are present on the page and provides links to further analyze that markup. Some of the types of markup it can detect are: Schema.org, Open Graph Protocol, Twitter cards, microformats, and authorship. Also offered are links to more of Moz's SEO tools.

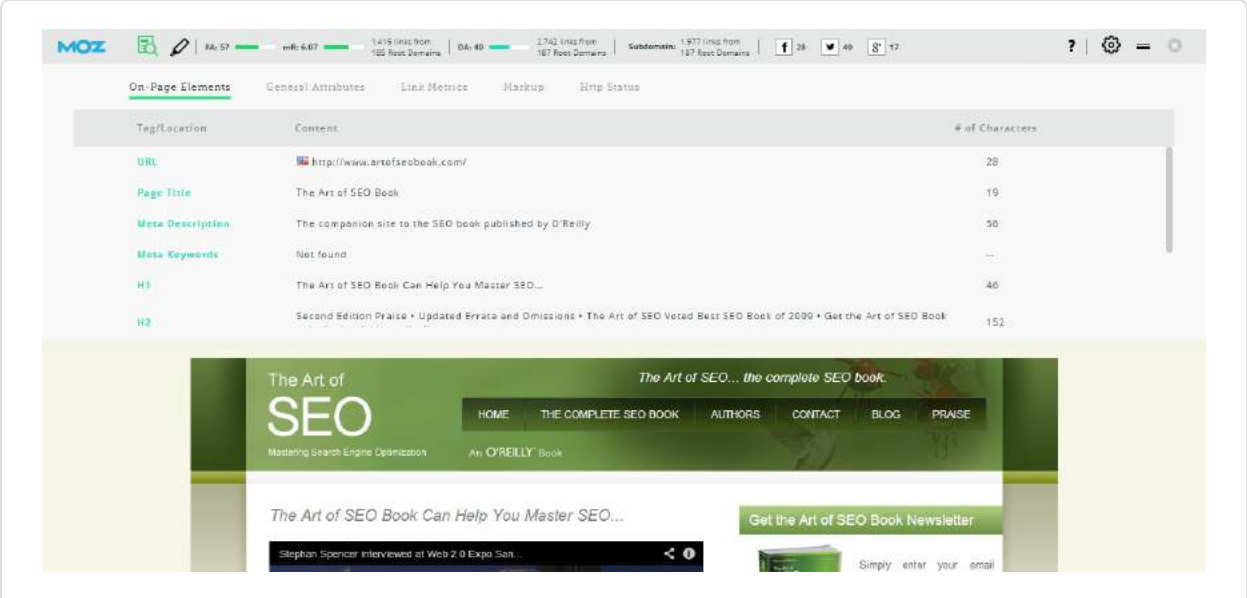


Figure 11-52. MozBar drop-down menus

SEO Quake

SEO Quake offers both a Firefox plug-in and a Chrome extension. SEO Quake adds a toolbar to your browser for at-a-glance statistics and offers links to more in-depth page information (see Figure 11-53). Some of the metrics this add-on allows you to view are: PageRank, number of pages indexed in Google, SEMrush links to both the page and domain, number of pages indexed in Bing, Alexa rank, domain age, social shares, links to whois information, the page's source code, and the SEMrush domain rank.

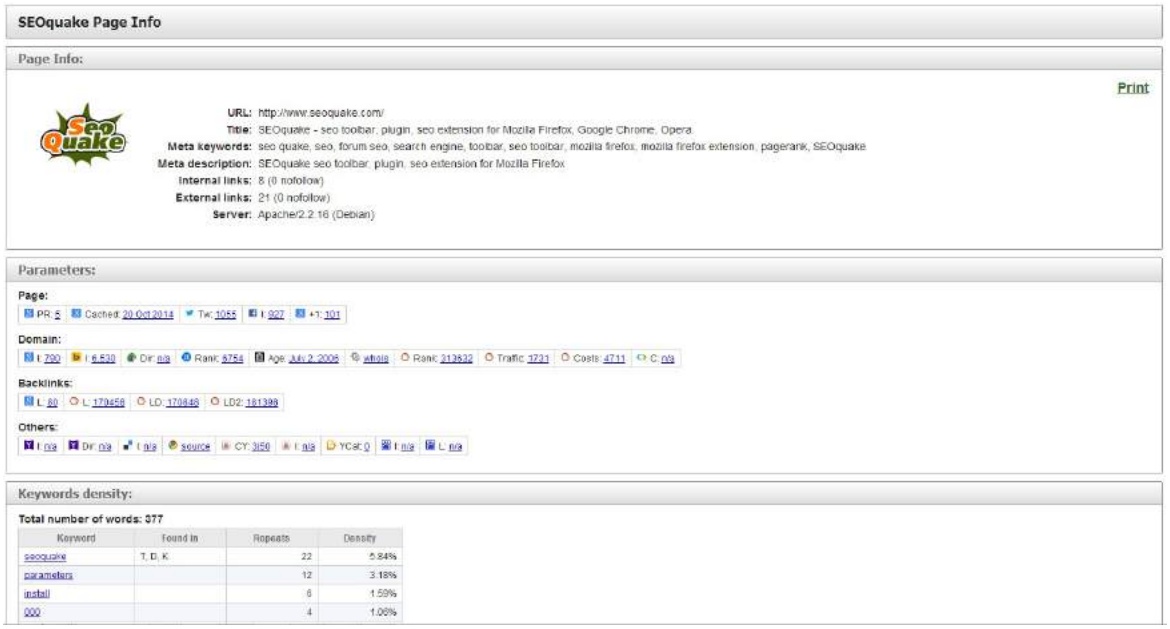


Figure 11-53. SEO Quake page info report

SEO for Firefox

Aaron Wall of SEO Book created a useful data enhancement tool with the SEO for Firefox extension. When turned on, this tool provides extra data that can be populated automatically or upon request from within the regular search results of Google, Yahoo!, and Bing. Some of the most popular data points this add-on provides for the URL or domain of each search result include: Google PageRank, site age, various backlink metrics, *.edu* and *.gov* link counts,

bookmarking and social metrics (like number of YouTube views), DMOZ listings, and additional reporting from other third-party tools.

SpyFu

SpyFu offers valuable keyword data for SEO professionals as well as resources for PPC advertisers. There is both a free version that provides valuable but limited data, as well as a paid version that provides additional data. For SEO, SpyFu offers the Domain Research, Keyword Research, and Backlink Discovery tools.

Domain Research allows you to review a site's organic keywords by their latest changes, by their page, and by clicks (see Figure 11-54). This allows you to monitor for rank, scout for potential algorithm penalties and manual penalties, and even compare these rankings to competitors.

Keyword Research, as expected, focuses on the stats of an individual keyword. The tool provides a history of domains that have ranked for a specific keyword, the URLs of pages ranking for a keyword, and keyword mapping ideas. These results can be incredibly helpful for developing a content strategy.

The Backlink Discovery tool doesn't work like other backlinking tools, which show current backlinks to the site. Instead, it suggests potential backlinking opportunities, such as hubs that are linking to various competitors or other sites within the niche. Getting deeper into it, you can even organize results by keyword, and single out backlinks that would be helpful to rank for a specific keyword.

For PPC, SpyFu can provide information on a site's estimated daily ad budget, total ad clicks per day, average cost per click, average ad clicks per day, average ad position, average ad competitors, average ad percentile, paid and organic keywords, top paid and organic competitors, and other domains that share the same owner.

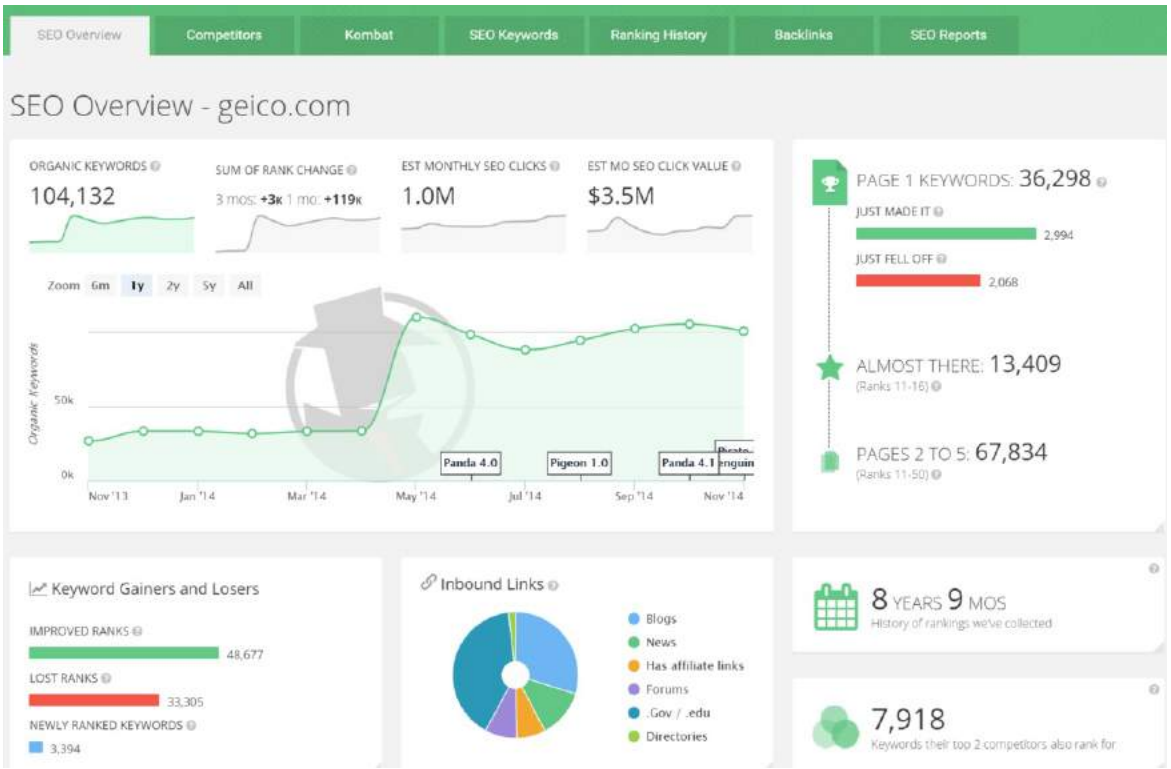


Figure 11-54. SpyFu Domain Research for SEO

SEMrush

SEMrush is another tool that offers useful keyword information to SEO and paid search professionals. Figure 11-55 shows a paid search report that is interesting from an SEO perspective because it shows the target landing page for the keyword, as well as the estimated CPC. This can provide insight into the value your competitors are placing on various keywords and landing pages on their site.

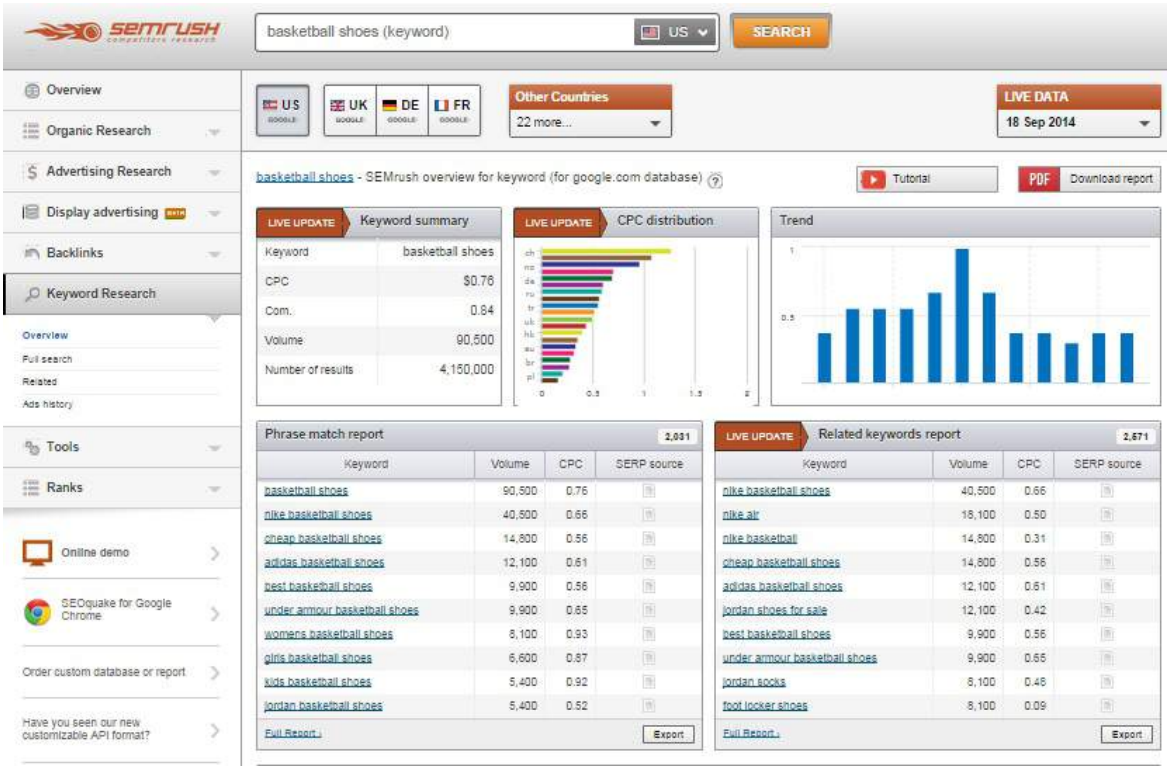


Figure 11-55. SEMrush keyword report

Rio SEO Search Analytics

Rio SEO Search Analytics is an enterprise class SEO platform that includes many more features than simple rank checking, crawl error alerts, and basic recommendations. Some of its features include:

- Keyword research solutions
- Long-tail analysis and recommendations
- Workflow tracking and automation
- Portfolio analysis for multiple sites, microsites, and landing pages
- Prioritization mechanisms
- Competitive analysis features
- Advanced link analytics

- Ad hoc analysis capabilities
- Advanced reporting capabilities
- Alerting capabilities
- Universal search metrics (image, video, etc.)
- Social signal inclusion and analytics

Rio SEO Website Optimizer

In some enterprise environments there are fundamental challenges with the content management system (CMS) or the ecommerce platform that make it very difficult to implement desired SEO changes.

There are only a few solutions in this space, and one of the best known is Rio's SEO Website Optimizer, formerly Covario's Organic Search Optimizer, originally developed by Stephan Spencer (one of the coauthors of this book) at Netconcepts and previously known as GravityStream.

Rio's Website Optimizer takes a complementary approach to whatever system is currently in place and helps SEOs get around IT to dynamically fix the SEO errors. Website Optimizer enables you to modify the pages without having to directly deal with the CMS or ecommerce platform.

Searchmetrics Essentials

Searchmetrics Essentials allows you to rapidly see how your competition is doing in organic search, paid search, news search, image search video search, and more (see Figure 11-56). The Keyword Research tab gives you search volume, cost per click, competition for keywords, and the average ad spend on a per-keyword basis. This can be more helpful and informative than simply using the Google Keyword Planner. You can actually see what people are spending to target specific keywords.

Searchmetrics Essentials also lets you peer deep into your competition's social media campaign. You can see which pages are the most liked, commented on, and shared throughout all of the popular networks. The platform offers many

other features as well.

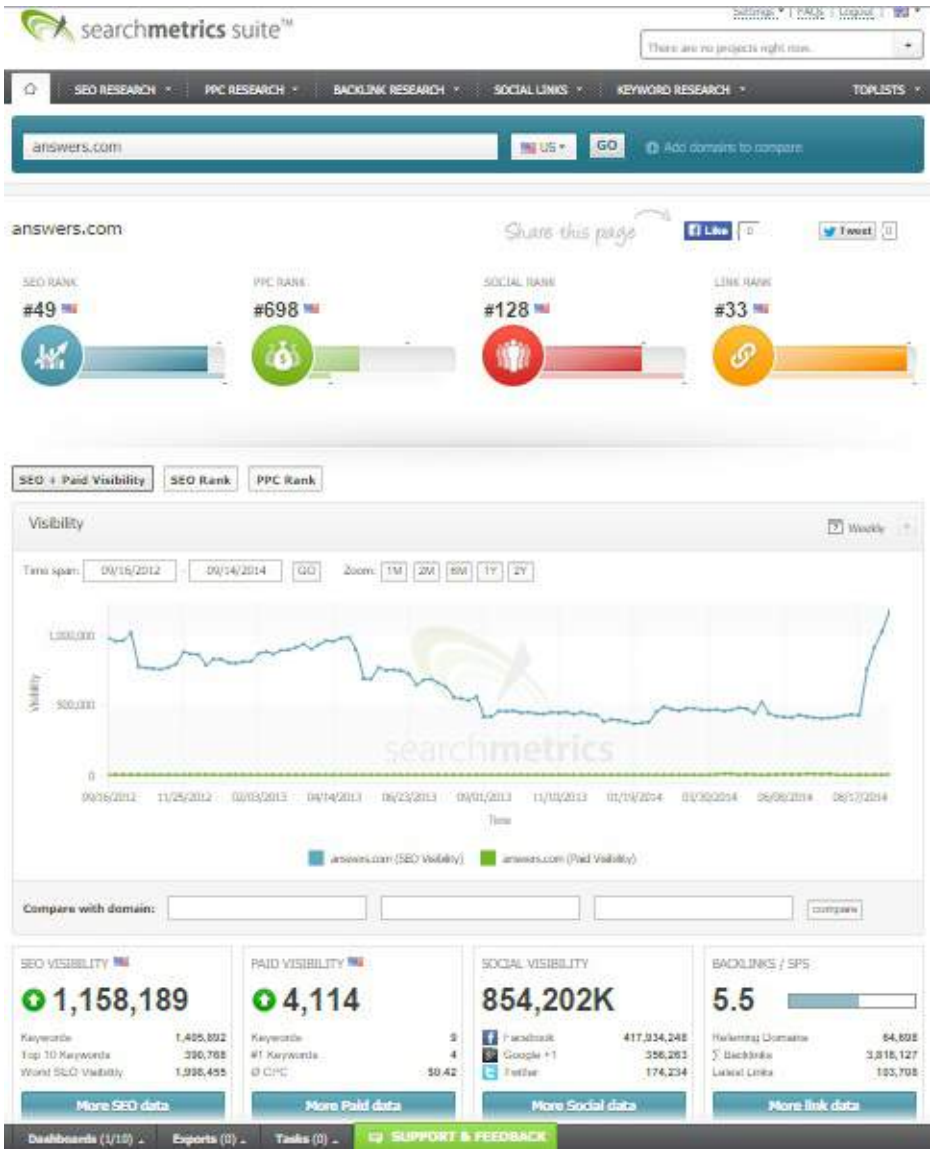


Figure 11-56. Searchmetrics Essentials

Conclusion

One of the biggest challenges for an SEO professional is the relatively long delay between implementation and results. Paid search campaigns can begin to show results in 24 hours, yet for SEO the process can take months. This makes effective measurement of results and progress critical. Putting the right analytics tools in place can provide the accountability required by budget managers. For

the SEO practitioner, it can also provide valuable insight that can be used to fine-tune SEO efforts for the best possible results.

[1] Christine Churchill, “Getting Query Data from Google Webmaster Tools to Correlate with SEO Query Data in Google Analytics,” SEM Clubhouse, March 11, 2013, *http://bit.ly/correlate_data*.

[2] Eric Enge, “The Impact of Comparison Shopping Sites on ECommerce Sales,” Search Engine Watch, August 1, 2007, *http://bit.ly/comparison_shopping*

[3] Justin Cutroni, “Getting Started with Google Analytics Remarketing,” Anaytics Talk, November 20, 2012, *http://cutroni.com/blog/2012/11/20/getting-started-with-google-analytics-remarketing/*.

CHAPTER TWELVE

Domain Changes, Redesigns, and Troubleshooting

Whenever you make structural changes to your website, there is a risk that you will confuse the search engines and harm your search rankings and organic traffic. The types of site changes that can influence your site's organic search performance include changing your domain name, changing your content management system (CMS), redesigning your site, adding ecommerce functionality, changing your blog platform, and many others—basically, anything that fundamentally alters your site's frontend and/or backend visual or functional elements can potentially influence your organic search performance.

In this chapter, we will review the various scenarios from the top down. Be sure to refer back to "Content Delivery and Search Spider Control" to review the technical specifics of options for moving content from one location to another.

Google groups site moves into one of two categories (*http://bit.ly/site_moves*): moves "with" URL changes, and moves "without" URL changes. It provides specific guidelines for handling moves within each category.

The Basics of Moving Content

"Moving content" refers to a situation in which content that used to be located and accessed at one URL (e.g., *http://www.yourdomain.com/pageA*) is moved to another URL (e.g., *http://www.yourdomain.com/products/pageA*). One of your goals when you move content is to make sure users and search engine crawlers that attempt to visit the old URL (*/pageA*) are presented with the content from the new location (*/products/pageA*). In addition, when you move content from one URL to another, the links to the old URL will stop providing value to your rankings in the search engines for that content unless you properly implement a 301 (permanently moved) redirect. There are very important reasons to move content properly—both of which can be easily overlooked by inexperienced or hurried webmasters and development teams.

In "Duplicate Content Issues", we covered the technical specifics of how to do this in detail, including the golden rule of moving content: the search engine needs to see a 301 HTTP status code whenever you redirect the content's URL to a new location.

The 301 HTTP status code causes the search engine to pass most of the value of any links for the original page over to the new page, and should result in the rapid deindexation of the old URL. Because link authority is a precious asset, you want to make sure you use a 301 redirect every time.

Large-Scale Content Moves

Setting up the redirects can become difficult when changes result in movement of large quantities of content. For example, when you change your domain name, every single piece of content on your site will move to a new URL, even if the site architecture is identical (*http://www.olddomain.com/*... moves to *http://www.newdomain.com/*...).

This is challenging because you might have to set up individual 301 redirects for every single page on the site, as in this example:

- *http://www.olddomain.com/page1* to *http://www.newdomain.com/page1*
- *http://www.olddomain.com/page2* to *http://www.newdomain.com/page2*
- *http://www.olddomain.com/page3* to *http://www.newdomain.com/page3*
- ...
- ...
- *http://www.olddomain.com/page1000* to *http://www.newdomain.com/page1000*

Unfortunately, some systems still require that these redirects be set up one at a time, so this could be quite a painful process. Imagine a site with 1 million pages!

Fortunately, publishers who use an Apache web server (Unix and Linux servers) can take advantage of the power of Apache's mod_rewrite module, which can perform the redirect of every URL on the old domain to the same URL on the new domain in two lines of code:

```
RewriteCond %{HTTP_HOST} ^olddomain\.com [NC]
RewriteRule ^/(.*) http://www.newdomain.com/$1 [R=301,L]
```

The preceding code presumes that you prefer the “www” version as the canonical URL. You can also use two similar lines of code to specify the “non-www” version as the canonical URL (see “URL Rewrites & Redirects: The Gory Details (Part 1 of 2)” for examples without “www” and other alternative approaches).

Another highly popular web server is Microsoft’s IIS. In many installations of IIS, you will find yourself in a situation where you have to implement a separate redirect instruction for each page, one at a time. Fortunately, you can utilize an ISAPI plug-in such as ISAPI_Rewrite, which enables you to perform large, scalable rewrites in a language similar to that used by Apache’s mod_rewrite. You can learn more about mod_rewrite, ISAPI_Rewrite, and regular expressions in Chapter 6.

Mapping Content Moves

Sometimes a site redesign is simply a “reskinning” of the visual elements of the old site with a new look and feel, retaining the same technical elements of information architecture, URL file and directory names, and user navigation. Other times, a redesign changes both the visual design and the technical elements. And sometimes it’s a combination of the two approaches. For sites changing both design and function, the first stage of planning is to figure out which content will be moved where and which content will be removed altogether. You will need this information to tell you which URLs you will need to redirect and to which new locations.

The best way to start this process is by getting a complete map of your information architecture with full URLs. For many websites this is not as simple as it sounds. Fortunately, tools are available to make the job easier. Here are some ways to tackle this problem:

- Extract a list of URLs from your web server’s logfiles and site architecture documentation.
- Pull the list from your XML sitemap file, provided you believe it is reasonably complete.
- Use a free crawling tool, such as Screaming Frog SEO Spider Tool.

- Use tools such as Moz.com's Open Site Explorer and Google Search Console to pull a list of the external links to your site, and make sure all pages that have received links on your site are included.
- Check Bing Webmaster Tools' Index Explorer to find all of the crawlable URLs that you may not know still exist on the site.

These tools should help you assemble a decent list of all your URLs. After determining which URLs have content that will remain on the site, you must then map out the pages that you want to redirect the "migrating" content to. Additionally, for content that is being "retired," you need to determine whether to redirect them at all (a definite yes if the URLs for these pages have many internal and external links), and if so, what new URLs to redirect them to.

One way to do this is to lay it out in a spreadsheet, which might end up looking like Table 12-1.

Table 12-1. Planning your content moves in advance

Old URL	New URL
http://www.olddomain.com/page1	*http://www.newdomain.com/page1*
http://www.olddomain.com/page2	*http://www.newdomain.com/page2*
http://www.olddomain.com/page3	*http://www.newdomain.com/page3*
http://www.olddomain.com/page4	*http://www.newdomain.com/page4*
http://www.olddomain.com/page5	*http://www.newdomain.com/page5*
http://www.olddomain.com/page6	*http://www.newdomain.com/page6*
http://www.olddomain.com/page7	*http://www.newdomain.com/page7*
http://www.olddomain.com/page8	*http://www.newdomain.com/page8*
http://www.olddomain.com/page9	*http://www.newdomain.com/page9*
http://www.olddomain.com/page10	*http://www.newdomain.com/page10*

If you are redirecting a massive number of URLs, you should look for ways to simplify this process, such as writing rules that communicate what you need to know. For example, you could abbreviate the list in Table 12-1 to the short list in Table 12-2.

Table 12-2. Simplifying content move planning with

wildcards

Old URLs	New URLs
*http://www.olddomain.com/page**	*http://www.newdomain.com/page**

Then you can save the individual lines for the more complicated moves, so your resulting spreadsheet would look like Table 12-3.

Table 12-3. Mapping all your content moves completely

Individual page moves	
Old URL	**New URL**
http://www.olddomain.com/about-us	*http://www.newdomain.com/about-us*
http://www.olddomain.com/contact-us	*http://www.newdomain.com/contact-us*
http://www.olddomain.com/press-relations	*http://www.newdomain.com/press*

Large-scale page moves	
Old URL	**New URL**
*http://www.olddomain.com/content/**	*http://www.newdomain.com/content/**
*http://www.olddomain.com/page**	*http://www.newdomain.com/page**

The purpose of this is to efficiently give your developers a map for how the content movement should take place. Note that the spreadsheet should contain a map of all changed URLs, which may include downloadable content such as PDF files, PowerPoint presentations, Flash files, multimedia, or any other types of content being moved.

You can also note retiring content via additional entries in the left column, with the entries in the right column indicating where users looking for that old content should be sent. Now your spreadsheet might look like Table 12-4.

Table 12-4. Identifying pages that have been removed

Individual page moves	
Old URL	**New URL**
http://www.olddomain.com/about-us	*http://www.newdomain.com/about-us*
http://www.olddomain.com/contact-us	*http://www.newdomain.com/contact-us*

http://www.olddomain.com/press-relations	*http://www.newdomain.com/press*

Large-scale page moves	
Old URL	**New URL**
*http://www.olddomain.com/content/**	*http://www.newdomain.com/content/**
*http://www.olddomain.com/page**	*http://www.newdomain.com/page**

Eliminated pages	
Old URL	**Redirect to**
http://www.olddomain.com/widgets/azure	*http://www.newdomain.com/widgets/blue*
http://www.olddomain.com/widgets/teal	*http://www.newdomain.com/widgets/green*
http://www.olddomain.com/widgets/puce	*http://www.newdomain.com/widgets/*

The new entries show what should happen to retired pages. The first two retired pages may represent products that you no longer carry, so you would likely want to redirect them to the closest existing products you have. The third retired page represents a URL where there is no sufficient replacement, so you may choose to redirect that one to the parent page for that topic area.

As you can see, a major SEO objective during content migration is to preserve as much link authority and traffic from the old URLs as possible, while providing the best possible user experience for people who arrive at the old URLs.

Expectations for Content Moves

The big downside to content migration is that the search engines won't necessarily adapt to the URL changes immediately. Many sites temporarily lose rankings after making a large-scale content move, then recover after a period of time. So naturally, the question is, how long will it take to get your organic rankings and traffic back?

The reality is that a number of factors are involved, depending on your particular situation. Some examples of these factors might include:

The size and complexity of your site

Bigger, more complex sites may take longer to process.

The complexity of the move

If the site has been fundamentally restructured, it is likely to take more time for the search engines to adapt to the new structure.

The perceived authority of the site

Sites that have a higher (search engine) perceived authority may be processed faster. Related to this is the rate at which the site is typically crawled.

The addition of new links to the new pages

Obtaining new links to the new URLs, or changing old links that used to point to the old URLs so that they point to the new URLs, can help speed up the process.

If you are moving to an entirely new domain, you can aid the process in Google by using the Change of Address tool inside Google Search Console. Before using this tool, make sure that both your old domain and your new domain are verified in Search Console. Then, on the Search Console home page, click on the old domain. Under Site Configuration, click "Change of Address," and then select the new site. You can also use the Bing Change of Address tool to let Bing know about your site move.

When all is said and done, a reasonable estimate is that a significant traffic dip from the search engines after a move should rarely last longer than 60 to 90 days, and many sites recover in a shorter timespan.

Another approach to content moves (especially when you're updating and redirecting an entire site's URLs) is to perform the URL migration in a phased manner as opposed to "wiping out" and redirecting all of a site's URLs at once. You'd do this for a few reasons, such as wanting to test the search engines' handling of such a migration on your site before committing to the sitewide change and its resulting impact. Another reason is to mitigate potential organic traffic dips that will occur during the updating period; it is often easier to tolerate 10% traffic loss at a time, versus a 30%–40% traffic loss all at once. This can be especially true for websites that rely upon traffic-based advertising revenue. Google's Matt Cutts describes the benefits of this phased approach in a quote we

cite in “Pre-Move Preparations”.

Maintaining Search Engine Visibility During and After a Site Redesign

Companies may decide to launch a site redesign as part of a rebranding of their business, a shift in their product lines, a marketing makeover, or for a variety of other business reasons. As discussed, any number of things may change during a site redesign. For example:

- Content may move to new URLs.
- Content might be eliminated.
- Content may be changed.
- Content could be moved behind a login or paywall.
- New site sections may be added.
- New site functionality may be added.
- Navigation/internal linking structure may be changed significantly.

Of course, the move may involve moving everything to a new domain as well, but we will cover that in the next section, “Maintaining Search Engine Visibility During and After Domain Name Changes”. Here are some best practices for handling a site redesign that involves these technical elements:

- Create 301 redirects for all URLs from the original version of the site pointing to the new URLs on the redesigned site. This should cover scenarios such as any remapping of locations of content and any content that has been eliminated. Use a spreadsheet similar to the ones we outlined at the beginning of this chapter to map out the moves to make sure you cover all of them.
- Review your analytics for the top 100 or so domains sending traffic to the moved and/or eliminated pages and contact as many of these webmasters as possible about changing their links. This can help the search engines

understand the new layout of your site more rapidly and provides both better branding and a better user experience.

- Review a backlink report (using your favorite backlink analysis tool) for your site and repeat the process in the preceding bulleted item with the top 200 to 300 or so results returned. Consider using more advanced tools, such as Open Site Explorer or Majestic SEO, which allow you to filter your links to more easily identify the most important ones. Ahrefs' Site Explorer and Cemper's LinkResearchTools are good resources too.

- Make sure you update your sitemap and submit it to Google Search Console and Bing Webmaster Tools. Consider using multiple sitemaps, one for each content type and/or content area, to submit and monitor the indexing of your new site URLs.

- Monitor your rankings for the content, comparing old to new over time—if the rankings fall, post in the Google Webmaster Central Help Forum detailing what you did, what happened, and any information that might help someone help you. Google employees do monitor these forums and sometimes comment in situations where they think they can help. Don't use the forums to complain; state what has happened and ask for help, as this gives you the best chance of getting feedback.

- Monitor your Search Console account and your analytics for 404 errors and to see how well Google is handling your 301s. When you see 404 errors occurring, make sure you have a properly implemented 301 redirect in place. Don't limit this checking just to 404 errors. Also be on the lookout for HTTP status codes such as 500 and 302. Maintain the XML sitemap of old URLs until search engines discover the 301s.

Maintaining Search Engine Visibility During and After Domain Name Changes

There may come a time when you have a strong business need—such as a rebranding, renaming, or merger/acquisition—to change your site's domain name. This section will cover some of the considerations and challenges

involved in a domain name change.

Unique Challenges of Domain Name Changes

One of the more challenging aspects of a domain name change is potentially losing the trust the search engines have associated with your old domain. Another issue is that if there were business-specific keywords present in your old domain name that are not in your new domain name, you may see a decline in organic traffic, even if you maintain or recover placement after migration. This decline is a result of "domain bias"—the propensity for searchers to click on domains in search results that include keywords they used in their search query. You may also see a slightly negative impact in organic search terms related to the keywords in your previous domain; although Google is cracking down on exact-match-domain (EMD) websites with low-quality content that were ranking well in search, it still places weight on the words in a domain.

Another unique challenge is the "youth" of the new domain, especially if it was recently purchased and/or has no historical backlink profile. Because of its age, the new domain may be slow to rank. Although the site's relevance and inbound link profile (including the links 301-redirected from the old domain) may suggest a high ranking for some search queries, because the new domain is not yet trusted, the rankings are suppressed and traffic is much lower than it would otherwise be. Domain youth is another reason why updating valuable third-party links to reflect your new domain is important.

If the prospect of taking a "young domain" hit is too unappealing, another tactic you can try is to make the move to a different domain that has a backlink history associated with it—just make sure that history is a positive one! You don't want to move to an old domain that had any historical spam, manual reviews, or other negative associations, so be sure to perform a thorough backlink audit with your preferred link auditing tools. And while you're at it, see if you can get Google Search Console access to research whether there were any manual spam actions reported against the domain.

Pre-Move Preparations

Unfortunately, lost traffic is common when you make domain name changes, though the traffic loss is usually temporary. If you do things properly, you can

and should recover from any negative impact, and hopefully quickly—but you should be prepared for the potential traffic impact of a domain switch.

If you are planning a domain migration, buy the new domain as early as you can, get some initial content on it, and acquire some links. The purpose of this exercise is to get the domain indexed and recognized by the engines ahead of time.

Then, register the new domain with Google Search Console and Bing Webmaster Tools. This is just another part of making sure Google and Bing know about your new domain as early as possible and in as many ways as possible.

Once you've done this, follow these best practices for handling a domain name change:

- Create 301 redirects for all URLs from the old site pointing to the proper URLs on the new site. Hopefully you will be able to use mod_rewrite or ISAPI_Rewrite to handle the bulk of the work. Use individual rewrite rules to cover any exceptions. Have this in place at launch.

- Review your analytics for the top 100 or so domains sending traffic to the old pages, and contact as many of these webmasters as possible about changing their links.

- Make sure that both the old site and the new site have been verified and have sitemaps submitted at Google Search Console and Bing Webmaster Tools.

- Launch with a media and online marketing blitz—your goals are to get as many new inbound links as possible pointing to the new site as quickly as possible, and to attract a high number of branded searches for the redesigned site.

- Monitor your rankings for the content, comparing old to new over time.

- Monitor your Search Console account for 404 errors and to see how well Google is handling your 301s. When you see some 404s pop up, make sure you have a properly implemented 301 redirect in place. If not, fix it.

- Monitor the search engine spidering activity on the new domain. This can provide a crude measurement of search engine trust. Search engines spend more time crawling sites they trust. When the crawl level at the new site starts to get close to where it was with the old site, you are probably most of the way there.

- Watch your search traffic referrals as well. This should provide you some guidance as to how far along in the process you have come.

- You can also check your server logs for 404 and 500 errors. These will sometimes flag problems that your other checks have not revealed.

As mentioned earlier, Google's Matt Cutts discussed the benefits of a phased approach at a PubCon event:

> *So here's the extra step. Don't just move the entire domain from the old domain to the new domain. Start out and then move a subdirectory or a subdomain. Move that first; if you've got a forum, move one part of your forum. Move that over to the new domain, and make sure that the rankings for that one part of your site don't crash. Sometimes it takes a week or so for them to sort of equalize out, because we have to crawl that page to see that it's moved. So if you move a part of your site first, and it goes fine, then you know that you're pretty safe. So instead of doing one huge move, if you can break it down into smaller chunks and start out by moving a small part of your site first, you'll know that you'll be gold.*

The value of this approach is that it reduces the risk associated with the move by breaking the migration process down into more manageable chunks. Even if you use this approach, however, you should still follow the guidelines outlined in this section to implement the move of each site section, and check on its progress.

Changing Servers

You might decide you want to move servers without changing your domain name or any of your URLs. A common reason for this change is that the growth of your traffic requires you to upgrade your hosting environment to a faster server. If you are using third-party hosting, perhaps you are changing your

hosting company, or if you have your own data center, you may need to move or expand your facilities, resulting in a change in the IP addresses of your servers.

This is normally a straightforward process, as you can simply go to the registrar where you registered the domain name and update the domain name system (DNS) records to point to the new server location. You can also temporarily decrease the site's DNS Time to Live (TTL) to five minutes (or something similar) to make the move take place faster. This is really the bulk of what you need to do, though you should follow the monitoring recommendations we will outline shortly.

Even if you follow this process, certain types of problems can arise. Here are the most common:

- You may have content that can't function on the new platform—for example, if you use Perl in implementing your site and Perl is not installed on the new server. This can happen for various other reasons as well, and the result can be pages that return 404 or 500 errors instead of the content you intended.

- Unfortunately, publishers commonly forget to move key content or files over, such as *robots.txt*, analytics files, *sitemaps.xml*, or the *.htaccess* file. It is imperative that these important files are migrated to your new server.

- Server configuration differences can also lead to mishandling of certain types of requests. For example, even if both your old server and your new server are running IIS, it is possible that the new server is configured in such a way that it will transform any 301 redirects you have in place into 302 redirects. Be sure to double-and triple-check that all server directives are properly migrated from the old server to the new.

The best advice for dealing with these concerns is to make a list of special files and configuration requirements and verify that everything is in place prior to flipping the switch on any server moves.

In addition, you should test the new site in its new location before making the move. You will need to access the content on the new site using its physical IP address. So, the page at *http://www.yourdomain.com/pageA* will be found at an address similar to *http://206.130.117.215/pageA*. To access the site, add that IP address to your test machine's `hosts` file (this assumes you are running

Windows) with a corresponding hostname of *http://www.yourdomain.com*, which will allow you to surf the site at the new IP address seamlessly. This advance testing should allow you to check for any unexpected errors. Note that the location of the `hosts` file varies across different versions of Windows, so you may need to search online to get information on where to find it on your machine.

Monitoring After Your Server Move

As with our other scenarios, post-launch monitoring is important. Here are the basic monitoring steps you should take:

- Monitor your Google Search Console and Bing Webmaster Tools accounts for 404 errors and to see how well the search engines are handling your 301s. When you see 404 errors, make sure you have a properly implemented 301 redirect in place.
- Monitor the spidering activity on the new domain to make sure no unexpected drops occur.
- Watch your search traffic referrals for unexpected changes.
- You can also check your server logs for 404 and 500 errors, which will sometimes expose problems that your other checks have not revealed.

Hidden Content

In "Content Delivery and Search Spider Control", we discussed ways that you can hide content from the search engines when you want to. However, at times this is done unintentionally—that is, sometimes publishers produce great content and then, for one reason or another, fail to expose that content to search engines.

Valuable content can be inadvertently hidden from the search engines, and occasionally, the engines can find hidden content and construe it as spam, whether that was your intent or not.

Identifying Content That Search Engines Don't See

How do you determine when you have unintended hidden content? Sometimes the situation is readily apparent; for example, if you have a site that receives high traffic volume and then your developer accidentally `noindex`es every page on the site, you will see a catastrophic drop in traffic. Most likely this will set off a panicked investigation, during which you'll quickly identify the `noindex` issue as the culprit.

Does this really happen? Unfortunately, it does. Here is an example scenario. Suppose you work on site updates on a staging server. Because you don't want the search engines to discover this duplicate version of your site, you keep the pages on the staging server `noindex`ed. Normally, when you move the site from the staging server to the live server, you remove the `noindex` tags, but unfortunately, many site owners forget to do this.

This type of problem can also emerge in another scenario. Some webmasters implement a *robots.txt* file that prohibits the crawling of their staging server website. If this file gets copied over when the site on the staging server is switched to the live server, the consequences will be just as bad as in the `noindex` scenario just outlined. The best way to prevent this type of situation is to implement a series of safety checks on the site that take place immediately after any update of the live server.

There are potential problems, however, that are much more difficult to detect. First, with a new site launch, you won't have any preexisting traffic, so there will be no drop in traffic levels to alert you that something is wrong. In another scenario, you may have an established site where you accidentally do something to hide only a portion of the site from the engines, so the issue is less obvious.

Regardless of your situation, web analytics can help you in the detection process. Use your analytics software to find pages on your site that get page views but no referring search traffic. By itself, this is not conclusive, but it provides a good clue as to where to start. Note that the converse of this is interesting for another situation—if you see content that is getting search referrals even though you don't want or expect it to, you may want to hide that content.

Another data point you can examine is the number of pages the search engines report as indexed for your site. In a new site scenario, you can look at this to see whether the search engines appear to be picking up your content. For example, if you have a site with 1,000 pages with a good inbound link profile, and after

three months only 10 pages are indexed, that could be a clue that there is a technical problem. Using multiple sitemap files, one for each site content area covering a specific segment of URLs, can be helpful in diagnosing such problems.

You do need to be careful not to overreact to the count of indexed pages, because the numbers that the search engines report will naturally fluctuate quite a bit. But if you are aware of the types of numbers typically reported for your site, and they drop to an unusually low level and stay there (or keep dropping), you probably have a problem.

Identifying the Cause of Non-Spidering

Once you realize you have a spidering problem, there are a series of checks you can perform to identify the potential causes.

Blocked by robots.txt

Check your *robots.txt* file (*http://www.yourdomain.com/robots.txt*) to see whether you are preventing the crawlers from accessing parts of the site that you actually want them to see. This mistake is quite common.

Both Google Search Console (see Figure 12-1 and Figure 12-2) and Bing Webmaster Tools provide simple ways for you to see whether they are aware of content that *robots.txt* is blocking them from crawling.

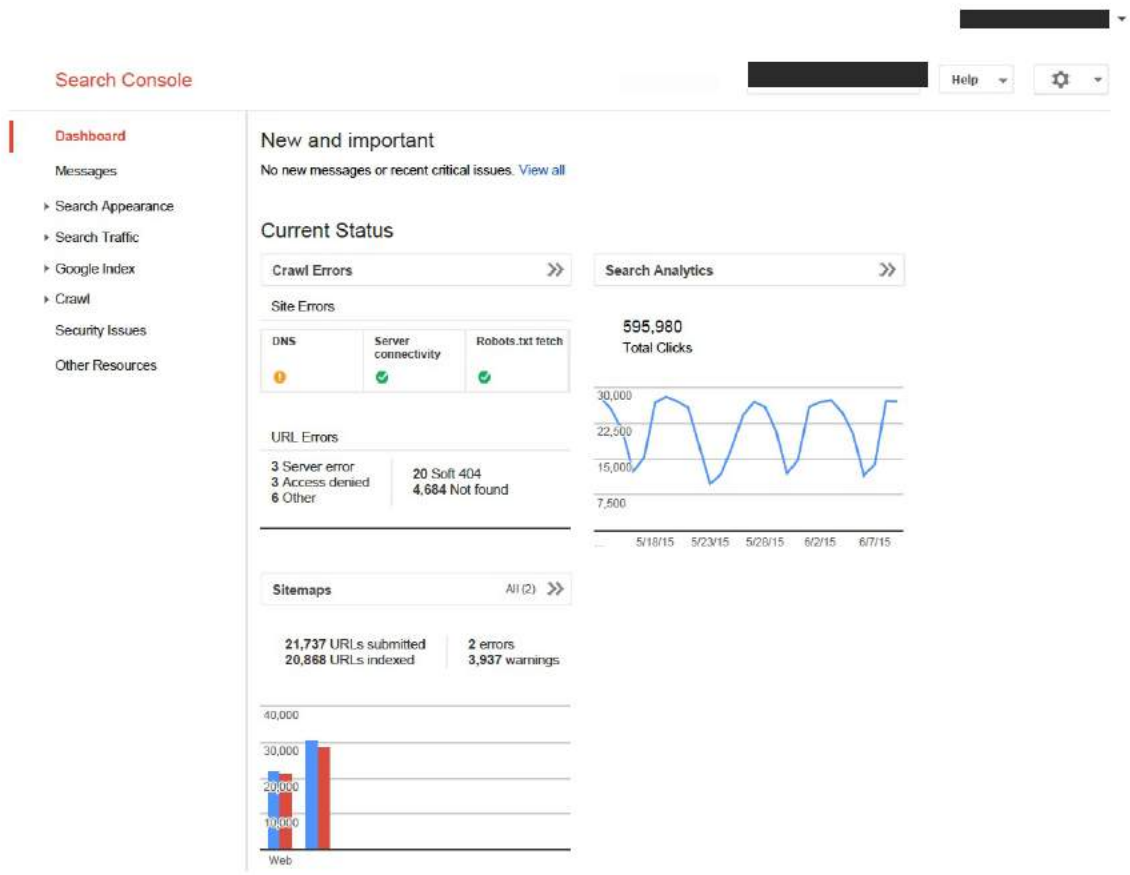


Figure 12-1. Google Search Console: restricted by robots.txt

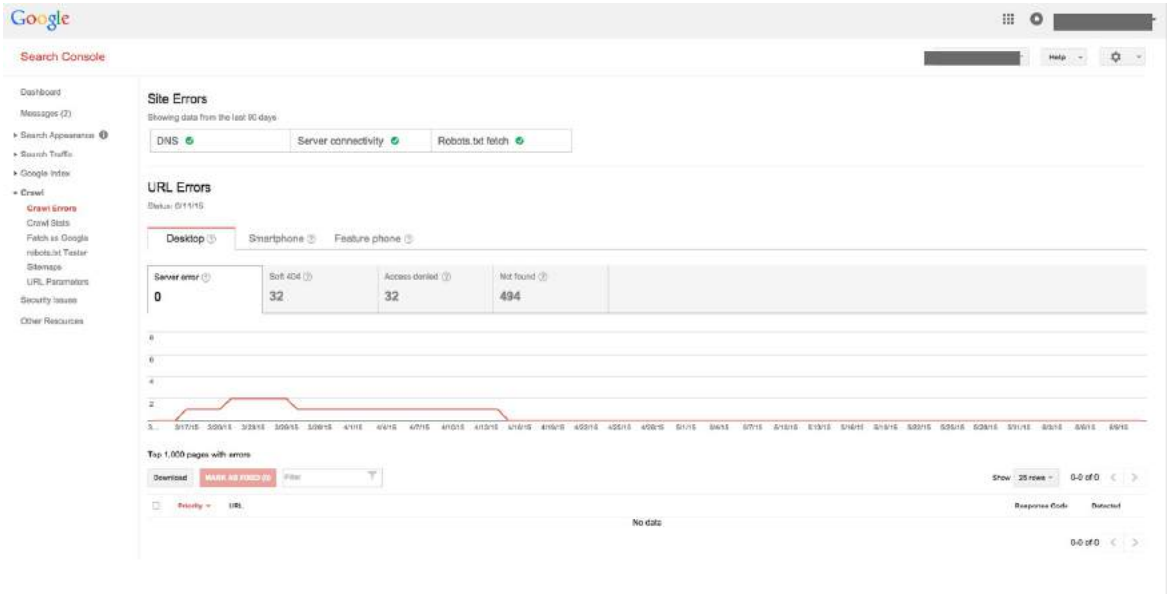


Figure 12-2. Google Search Console: crawl errors

These reports are helpful when you have content on the site that has links to it (either internal or external) but that the search engines don't crawl because they are excluded from it in *robots.txt*.

The solution is simple: figure out what line in your *robots.txt* file is blocking the search engines and remove it, or update it so that the content is no longer being blocked.

Blocked by the robots meta tag

The `robots` meta tag in a page's header might look something like this:

```
<meta name="robots" content="noindex, nofollow">
```

As we discussed in "Content Delivery and Search Spider Control", a setting of `noindex` will tell the search engine that it is not allowed to include the page in its index. Clearly, you should check to see whether you have made this error if you find that the engines are not crawling certain pages that you want crawled.

In addition, the `nofollow` tag will tell search engines not to pass any link authority to the pages linked to on that page. If all the links on your site to a particular piece of content are `nofollow`ed, you are passing no link authority to the page. This tells the search engines that you don't value the page, and as a result they won't treat the links as endorsements for it.

Solving this problem requires locating the places where these `robots` meta tags are on your site and removing them. Note that the default setting for the `robots` meta tag is `"index,follow"`, so there is no need to implement the tag if that is your desired setting. Just make sure you don't have `robots` meta tags in place that change the default in places where that is not the desired behavior.

No direct links

You may find that a particular piece of content has no links to it, or that all of the links to that content are `nofollow`ed. You can also make links invisible to the search engines (possibly unintentionally) by showing them only to visitors that support cookies, or by encrypting the links to the content in some fashion. The solution here is to make sure you implement plain-text (or image) links to the

content. Better still, get some third-party websites to link to the content as well.

Form submission requirement

Requiring a login or some other type of form submission to see content is another common cause of non-spidering. Search engines will not attempt to fill out forms to see what is behind them. The simplest solution is often to remove the requirement for the form if you want the search engines to index this content.

However, some sites sell content on a subscription basis (also referred to as being behind a "paywall"), and they will not want to offer their content for free. In October 2008, Google announced the First Click Free program (*http://bit.ly/reg_sub_sites*), which allows subscription-based sites to have their content crawled and indexed by Google, but still allows the publisher to require human visitors to subscribe to access the content. As of 2014, Google requires that any user coming from a Google domain must be able to see a minimum of five free articles per day, which it refers to as "metering." You can read more about First Click Free in "Content Delivery and Search Spider Control".

Session IDs

Session IDs confuse search engine crawlers: every time the engines come to your site, they see a different page. For example, they may see *http://www.yourdomain.com?SessID=2143789* one time and *http://www.yourdomain.com?SessID=2145394* the next. Even though your intent is to track the session of a particular user, and you think of these URLs as the same page, the search engine does not. You can read more about session IDs in "Controlling Content with Cookies and Session IDs".

Not enough link authority to remain in main indexes

Sometimes the non-spidering problem has nothing to do with the issues we just discussed. The search engines may see the page just fine, but there may not be enough link juice going to it to merit inclusion in their main indexes. This is more common than people think, and it happens because the search engines do not attempt to index all the world's web pages.

For example, content that Google perceives to be of low importance (i.e., content that doesn't have enough link authority, or is perceived to be duplicate content) will be excluded from the main index. In previous years, this content may have

been relegated to what Google called its "supplemental index," but in 2014 Google's John Mueller confirmed that for the purposes of treating pages differently, Google no longer has a supplemental index.

Google wants to emphasize the more important pages on the Web, and doesn't want the rate at which it delivers search results to be slowed down by pages that most people probably don't want to see.

Identifying Hidden Content That May Be Viewed as Spam

Hidden text is one of the challenges that webmasters and search engines still face. Spammers continue to use hidden text to stuff keywords into their pages, for the purposes of artificially boosting their rankings. Search engines seek to figure out when spammers are doing this and then take appropriate action. There are many ways to create hidden text unintentionally, though, and no one wants to be penalized for something they did not intend to do. Google's Webmaster Guidelines for hidden text state the following:

> *Hiding text or links in your content to manipulate Google's search rankings can be seen as deceptive and is a violation of Google's Webmaster Guidelines.*

In a Google Groups thread (*http://groups.google.com/group/Google_Webmaster_Help-Indexing/browse_thread/thread/928aa76a1226cf89/32b089e3248cef78?pli=1*), Googler Susan Moskwa had this to say:

> *Of course, as with many techniques, there are shades of gray between "this is clearly deceptive and wrong" and "this is perfectly acceptable." Matt [Cutts, former head of Google's webspam team] did say that hiding text moves you a step further towards the gray area. But if you're running a perfectly legitimate site, you don't need to worry about it. If, on the other hand, your site already exhibits a bunch of other semi-shady techniques, hidden text starts to look like one more item on that list. It is like how 1 grain of sand isn't noticeable, but many grains together start to look like a beach.*

Related to this is a posting by Matt Cutts on Threadwatch (*http://www.threadwatch.org/node/4313#comment-26883*):

> *If you're straight-out using CSS to hide text, don't be surprised if that is called spam. I'm not saying that mouseovers or DHTML text or have-a-logo-but-*

also-have-text is spam; I answered that last one at a conference when I said, "imagine how it would look to a visitor, a competitor, or someone checking out a spam report. If you show your company's name and it is Expo Markers instead of an Expo Markers logo, you should be fine. If the text you decide to show is "Expo Markers cheap online discount buy online Expo Markers sale..." then I would be more cautious, because that can look bad.

Obviously, this is a fate you want to avoid. Note the use of the word *perceived* in the Google Webmaster Guidelines snippet. It doesn't sound like a simple black-and-white problem, does it? In fact, it is not, as there are many ways to create hidden text.

Unintentionally creating hidden text

There are a few ways to create hidden text without intending to do so. One of the most common ways is via your CMS, which has some CSS-based methods built into it. For example, many content management systems use the `display:none` technique to implement drop-down menus or other widgets that "expand" to display more text when clicked. Tab folders are a great example of this. Sometimes the `display:none` technique is used in user-generated content systems where the page normally shows the number of comments on a post, but suppresses the text "0 Comments" in the event that no comments have been made.

People can also commonly create hidden text when they provide enhancements for the visually impaired. For example, you may have a Flash object on your web page and want to provide users with a text description of the content. You may not want to place the text on the page, as it might make the page look cluttered to a user with normal vision. The solution some people use to serve both audiences is to hide the text from the sighted users.

Many of these scenarios have no SEO value, even when manipulated by spammers. These types of techniques generally do not carry a risk of being penalized, because there is no reason to suspect negative intent.

Spam Filtering and Penalties

Over time, it has become a lot more difficult to "game" the search engines and a

lot easier to fall victim to a search engine penalty or outright ban. It is hard to recover from these.

Consequences can include ranking penalties, removal of the site's "voting" power (i.e., ability to pass PageRank), incomplete indexation (i.e., a partial site ban), or, worst of all, a total site ban.

Not even the largest corporations spending big dollars on Google AdWords are immune. In 2013, both BBC News and the popular music lyrics website Rap Genius were hit with Google penalties. In the case of BBC News, Google Webmaster Tools (what Search Console was called at the time) sent an "unnatural link" notification (which represents 1%–2% of Google Webmaster Tools/Search Console notifications.[1] While the case turned out to be targeted to only one article, it caused a stir.[2]

The Rap Genius case, however, was a much more serious matter that involved the company's invitation to bloggers to add links to Rap Genius lyric URLs in exchange for Rap Genius subsequently tweeting the bloggers' posts to its Twitter followers. This link scheme came to the attention of Google's Matt Cutts, who vowed Google would investigate, and it did—resulting in Rap Genius incurring an official Google penalty for link schemes.[3]

Search engines rely primarily on automated means for detecting spam, with some auxiliary assistance from paid evaluators, spam vigilantes, and even your competitors. Search engineers at Google and Microsoft write sophisticated algorithms to look for abnormalities in inbound and outbound linking, in sentence structure, in HTML coding, and so on.

As far as the search engines are concerned, SEO has both an acceptable side and an unacceptable side—and in general terms, many efforts intended to boost a site's search engine ranking without improving the measurable value of a page can be considered spamming.

Each search engine has different published guidelines. Here is where you can find them:

- Google's Webmaster Guidelines
- Bing Webmaster Guidelines

The search engines have varying degrees of tolerance for SEO tactics. Anything

that violates these guidelines, pollutes the search results with irrelevant or useless information, or would embarrass you if your Google AdWords or Bing rep discovered it is unsustainable, and should generally be avoided.

There is a difference between "search engine friendly," and crossing the line into spam territory. Search engine friendly can mean, for example, that the site is easily accessible to spiders, even if it is database-driven; that HTML code is streamlined to minimize the amount of superfluous code; that important headings, such as product names, are set apart from the rest of the text (e.g., with `<h1>` tags) and contain relevant keywords; or that link text is contextual, instead of comprising just "click here" or "more info" references.

Contrast these basic SEO practices with the following manipulative search engine spam tactics:

- Serving pages to the search engines that are useless, incomprehensible, unsuitable for human viewing, or otherwise devoid of valuable content—such as *doorway pages*, which SEO vendors may refer to by more innocuous names, including *gateway pages*, *bridge pages*, *jump pages*, *attraction pages*, *advertising pages*, *channel pages*, *directory information pages*, *search engine entry pages*, *satellite sites*, *mini sites*, *magnet sites*, or *shadow domains*. Whatever you call them, by definition they are created for the sole purpose of boosting search engine rankings.

- Creating sites with low-quality content. There are many techniques for doing this, including:

  — Duplicating pages with minimal or no changes and exposing them to the same search engines under new URLs or domains.

  — Machine-generating content to chosen keyword densities (e.g., using a technique such as Markov chains, which are not recommended).

  — Incorporating keyword-rich but nonsensical gibberish (also known as *spamglish*) into site content.

  — Creating a low-value site solely for affiliate marketing purposes (see "Duplicate Content" for a more complete definition of *thin affiliate*).

- Repeating the same keyword phrase in the `<title>` tag, the `<h1>` tag, the first `alt` attribute on the page, the meta description, the first sentence of body copy, and the anchor text in links pointing to the page.
- Targeting topically irrelevant keywords.
- Concealing or obscuring keyword-rich text or links within the HTML of a page so that it is not visible to or accessible by human users (i.e., by placing it within comment tags, `<noscript>` tags, or `<noframe>` tags; or by using colored text on a similarly colored background, tiny font sizes, layers, or links that don't show as links to users because they are not highlighted in some manner, such as with an underline).
- Hijacking or stealing content from other sites and using it as content fodder for search engines. This is a practice normally implemented through scrapers.
- Purchasing links for the sole purpose of influencing search rankings.
- Participating in *link farms* (which can be distinguished from directories in that they are less organized and have more links per page) or *reciprocal linking schemes* (link exchanges) with irrelevant sites for the purpose of artificially boosting your site's importance.
- Peppering websites' guest books, blogs, or forums in bulk with keyword-rich text links for the purpose of artificially boosting your site's importance.
- Conducting *sneaky redirects* (immediately redirecting searchers entering your site from a keyword-rich page that ranks in the search engine to some other page that would not rank as well).
- *Cloaking*, or detecting search engine spiders when they visit and modifying the page content specifically for the spiders to improve rankings.
- *Negative SEO*, or submitting your competitors to link farms and/or otherwise creating third-party associations on the competitors' behalf so that they will be penalized.

These tactics are questionable in terms of effectiveness and dubious in the eyes of the search engines, often resulting in the offending site being penalized by or banned from the search engines—a risk that's only going to increase as the engines become more aggressive and sophisticated at identifying and removing offenders from their indexes. We do not advocate implementing these tactics to those interested in achieving the long-term benefits of SEO.

The search engines detect these tactics not just through sophisticated automated spam-catching algorithms, but also through spam reports submitted by searchers—and yes, by your competitors.

Oftentimes search marketers don't even know they're in the wrong and running afoul of search engine guidelines. For example, search engines place more scrutiny on pages that show signs of potential deception, such as `<no-archive>` tags, `<noscript>` tags, `<noframe>` tags, and cloaking, even though all of these can be, and often are, used ethically.

Seeing SEO strictly as a chess game between yourself and the search engines is a shortsighted view, as the goal of the search engines is to provide relevant search results to their users, and your goal as a search marketer is to provide people with valuable content and a positive website experience. Trying to fool the search engines and take unfair advantage by exploiting how they work is simply not a sustainable approach.

Chapter 9 addresses the topic of Google's manually applied penalties, as well as its Panda and Penguin algorithms, in detail.

Low-Quality Domains and Spam Sites

You can spot a poor-quality website in many ways, not the least of which is the "common sense" check. Would you hire a company with a website named *www.seoseosanfranciscoseo.com* or *www.sonysonyelectronicssonyforsale.com*? The domain, of course, is only one signal, and search engines rely on a wide range of signals as indicators of quality. Some of the most obvious signals are site owners who are actively spamming the search engines with their offsite activities—for example, if the site is actively buying links, or text-link-spamming blog posts, forums, and article comments.

However, there are also less obvious signals. Many such signals mean nothing

by themselves and gain significance only when they are combined with a variety of other signals. When a number of these factors appear in combination on a site, the likelihood of it being seen as a low-quality or spam site increases.

Here is a long list of some of these types of signals:

- Short registration period
- High ratio of ad blocks to content
- JavaScript redirects from initial landing pages
- Use of common, high-commercial-value spam keywords such as *mortgage*, *poker*, *texas hold 'em*, *porn*, *student credit cards*, and related terms
- Many links to other low-quality spam sites
- Few links to high-quality, trusted sites
- High keyword frequencies and keyword densities
- Zero, or very little, unique content
- Very few direct visits
- Registered to people/entities previously associated with untrusted sites
- Not registered with services such as Google Search Console or Bing Webmaster Tools
- Rarely have short, high-value domain names
- Often contain many keyword-stuffed subdomains
- More likely to have longer domain names (as above)
- More likely to contain multiple hyphens in the domain name
- Less likely to have links from trusted sources

- Less likely to have SSL security certificates
- Less likely to be in high-quality directories such as DMOZ, Yahoo!, and Best of the Web
- Unlikely to have any significant quantity of branded searches
- Unlikely to be bookmarked in services such as Delicious
- Unlikely to get featured in social voting sites such as Digg, reddit, StumbleUpon, and so forth
- Unlikely to have channels on YouTube, communities on Google+ or Facebook, or links from Wikipedia
- Unlikely to be mentioned on major news sites (either with or without link attribution)
- Unlikely to be registered with Google and Bing local listings
- Unlikely to have a legitimate physical address/phone number on the website
- Likely to have the domain associated with emails on blacklists
- Often contain a large number of snippets of “duplicate” content found elsewhere on the Web
- Frequently feature commercially focused content
- Many levels of links away from highly trusted websites
- Rarely contain privacy policy and copyright notice pages
- Rarely listed in the Better Business Bureau’s Online Directory
- Rarely contain high-grade-level text content (as measured by metrics such as the Flesch-Kincaid Reading Level)
- Rarely have small snippets of text quoted on other websites and pages

- Commonly employ cloaking based on user agent or IP address
- Rarely have online or offline marketing campaigns
- Rarely have affiliate link programs pointing to them
- May have links to a significant portion of the sites and pages that link to them
- Extremely unlikely to be mentioned or linked to in scientific research papers
- Unlikely to use expensive web technologies (Microsoft Server and coding products that require a licensing fee)
- More likely to contain malware, viruses, or spyware (or any automated downloads)
- Likely to have privacy protection on the whois information for their domain

It is important to note that while many of these signals can be viewed negatively in aggregate, having one, such as a private domain registration, in and of itself is not going to be interpreted as a spam signal. Many legitimate sites will have one or more of these signals associated with them. For example, there are many good sites with a *.info* TLD.

There are also some signals that require data from a web analytics tool (which Google may be able to obtain from the Google Analytics account of the site in question):

- Rarely receive high quantities of monthly visits
- Rarely have visits lasting longer than 30 seconds
- Rarely have visitors bookmarking their domains in the browser
- Unlikely to buy significant quantities of PPC ad traffic
- Rarely have banner ad media buys
- Unlikely to attract significant return traffic

There can be legitimate reasons for the occurrence of many (possibly even most) of the aforementioned signals. For instance:

- Businesses outside the United States will not be in the Better Business Bureau directory.
- The site may not be relevant to scientific research papers.
- The publisher may not be aware of Google Search Console or Bing Webmaster Tools.

As these three examples illustrate, it's important to put all of these signals into proper context before making an assessment of site quality.

Spam Reports

Search engines supplement their internal spam fighting efforts by allowing users—including your competitors—to submit spam reports. For example, Google provides a form for reporting spam at *https://www.google.com/webmasters/tools/spamreport*.

In addition, the search engines can and do make use of human reviewers who conduct quality reviews.

In late 2012, Google published "Search Quality Rating Guidelines", a "CliffsNotes" version of the criteria it uses for search quality rating.

Duplicate Content

As we discuss in "Content Management System Issues", there are many ways to unintentionally create duplicate content. For this reason, search engines' normal response to duplicate content is to filter it out, but otherwise not penalize the publisher for it. They filter it out because they don't want to show multiple copies of the same piece of content in their search results, as this does not really bring any value to users. They don't punish the publisher because the great majority of these situations are unintentional.

In fact, in late 2013, Matt Cutts declared that 25%–30% of content on the Web is duplicative and that unless the content duplication was egregious and malicious, you don't need to worry too much about it.[4]

However, there are three notable exceptions that can result in your site being impacted by Google's Panda algorithm (*http://bit.ly/google_panda_update*):

Copyright violations

In this scenario, there is an actual copyright violation where a publisher is showing a copy of another publisher's content, usually via scraping, without permission.

In 2014, Google announced its new Scraper Report), with Matt Cutts tweeting: "If you see a scraper URL outranking the original source of content in Google, please tell us about it."

Thin affiliate sites

In this scenario, the publisher has permission from another publisher (the "Content Owner") to display its content, but the Content Owner also shares that same content with many other sites. The common scenario is a Content Owner running an affiliate network and generating leads or sales for its products or clients, largely by offering an affiliate program to other publishers.

All of the affiliate sites publish the same (or very similar) content. The problem the search engines have with these sites is that they offer very little value, as there is nothing truly unique about their content.

The affiliate site may also create hundreds or thousands of pages to target vertical search terms with little change in content. The classic example of this is creating hundreds of web pages that are identical except for the city name referred to on each page ("Phoenix Oil Changes," "Austin Oil Changes," "Orlando Oil Changes," etc.).

Massive duplication

This scenario applies to a site that is not a thin affiliate but has a very large amount of duplicate content (even if it has duplicated that content with permission). The threshold is unknown, and it probably changes over time, but our experience suggests that sites in which 70% or more of the pages are duplicates of other pages on the Web are likely to be subject to a penalty.

There may be other spammy forms of duplicate content not identified here. It is likely to be spam if it is implemented intentionally (e.g., the thin affiliate site example qualifies here), if it is intended to manipulate search results, and if the content adds no value to the Web. On the other hand, syndicated content (content published on the original source's website and then copied with express permission on third-party websites), when published properly, presents no duplicate content issues. Coauthor Eric Enge's Search Engline Land article "Syndicated Content: Why, When, and How" provides guidelines for how to properly and safely syndicate your site content.

Basic Rules for Spam-Free SEO

Especially if you are new to SEO, the first and most important rule is to be familiar with the guidelines of the search engines (see the beginning of this section for their locations).

Second, it is essential that you learn to apply a basic personal filter to your SEO activities. Generally speaking, if you engage in an activity for the sole purpose of influencing search rankings that has no benefit to users or search engine crawlers, you are putting your site's rankings at risk.

For example, if you start buying keyword-rich text links from a bunch of sites across the Web and you do not expect to get significant traffic from these links (just enough to justify placing the advertisements), you are headed for trouble and likely will eventually be hit by Google's Penguin algorithm.

Of course, there are more ways you can get into trouble than through purchasing or aggressively spamming other websites with keyword-rich text links. Earlier in this chapter, we listed a number of spammer techniques. Most publishers/SEO practitioners won't run into the majority of these, as they involve extremely manipulative behavior, as we outlined previously in this section. However, newcomers to SEO do tend to make certain mistakes. Here are some of the more common ones:

- Stuffing keywords into your web page so it is unnaturally rich in those words.
- Overoptimizing internal links (links on your website to other pages on your website). Generally speaking, you do this by overstuffing keywords into the anchor text of those links.

- Cloaking, or showing different content to the search engine crawlers than you show to users.
- Creating websites with lots of very thin content pages, such as the thin affiliate sites we discussed previously.
- Implementing pages that have search engine–readable text that is invisible to users (a.k.a. hidden text).
- Participating in link schemes, link farms, or aggressive guest blog posting, or using other tactics to artificially boost link popularity.

As we discussed previously, there are many other ways to end up in the crosshairs of the search engines, but most of those fall into the category of highly manipulative SEO tactics (generally referred to in the industry as *black hat SEO*).

Ultimately, you want to look at your intent in pursuing a particular SEO practice. Is it something you would have done if the search engines did not exist? Would it have been part of a publishing and promotional plan for your website in such a world?

The notion of intent is something that the search engines look at very closely when evaluating a site to see whether it is engaging in spam tactics. In fact, search engine representatives speak about *intent* (why you did it) and *extent* (how much you did it) as being key things they evaluate when conducting a manual review of a website.

A perceived high degree of intent and pursuing something to a significant extent can lead to more severe penalties. But even if your intent is pure and you don't pursue a proscribed practice to a significant degree, the search engines will want to discount the possible benefits of such behavior from the rankings equation. For example, if you buy a few dozen links and the search engines figure out that the links were paid for, they will discount those links. This behavior may not be egregious enough for them to penalize you, but they don't want you to benefit from it either.

Search Engine Penalties and Reconsideration Requests

The reality is that penalties are imposed on websites. It is often hard to know what has caused a penalty; although the search engines might describe the general reason for it (such as “unnatural links”), they will not identify the behaviors specifically. You will need to review your site to see where you have deviated from best practices and address all such issues. For a thorough discussion on how to identify search engine penalties and how to recover from them, see Chapter 9.

Content Theft

The Web is sometimes a little bit of a Wild West environment, and one of the bad things that can occur is that another publisher (a spammer) can copy your content and republish it on one of his own sites. This is very easy for infringers to do. They simply “View Source” in their browsers and take whatever they like without regard to copyright, or send a spider to collect all of your pages in one fell swoop.

You can locate copyright infringers pretty easily using various online tools, including Plagium and Copyscape, if they’ve lifted some of your page copy. It’s much more difficult if they’ve limited their sticky fingers to just your design.

There are six major options for dealing with an infringer:

- File a DMCA (Digital Millennium Copyright Act) infringement notification.
- Contact the infringing company’s CEO, head of marketing, legal counsel, or whomever you feel is most likely to respond.
- Notify the site’s hosting company, if it is not self-hosted.
- Report the activity on your blog, or if your blog is not that popular, get someone with a popular blog to report it.
- Have your lawyer send the infringer a cease and desist letter.

Doing nothing is not an option if you’re serious about protecting your intellectual property rights. At the very least, you want to show you have made an effort to protect your IP in the event that ownership comes into question in

other scenarios.

A phone call to the infringing company is inexpensive, and it can work. Sometimes publishers who use contract writers are not even aware that they are infringing. But this is often not a viable option if you're dealing with an entity that has masked its domain contact information and is hosting its site with a company that will not respond to your requests, and where a copyright complaint will not be effective (i.e., in countries without copyright laws).

In the United States, the DMCA provides a useful hammer to beat on copyright infringers without the need to contact them yourself, in the form of a DMCA infringement notification. In DMCA legal speak, this is also known as a "takedown notice." Other countries with strong copyright laws have similar processes.

As a content producer, you have the right to enforce your copyright. When your content gets "repurposed" on others' websites without your permission, you can simply file a DMCA infringement notification with the infringer's web hosting provider to get that infringer's website shut down.

You can also get the infringing content delisted from the search engines by filing DMCA notices with Google and the other major engines. It is not a daunting procedure. It might take an hour of your time, and it is well worth it. The process is as follows:

1. First, look up the web host and the domain registrar of the offending site using lookup tools such as DomainTools. You can usually ascertain who the web host is from the name servers and/or the netblock owner.

2. Check the official directory of designated DMCA agents for the host and the registrar. (Hopefully, they're listed.)

3. Prepare a letter to send to the designated agent of the web host. The notice you write should include your contact information, a description of and a link to the content that was copied, the web address of the copied content, a statement that you have a good faith belief that the duplication of the material is not legal, a statement that under penalty of perjury you are the copyright holder, and your signature. Some web hosts will allow you to email your notice to them, making this process even easier.

4. Send a similar notification to the search engines. That will cut off the site's air supply (its traffic) in case the content doesn't get taken down right away. Instructions and contact details for each engine are available at the following URLs:
   - Google: *http://bit.ly/removing_content*
   - Bing: *https://www.microsoft.com/info/cpyrtInfrg.aspx*
5. If the web host doesn't take the site down promptly, submit a DMCA notice to the infringer's domain registrar. It might be worth sending a notice to the data center that the web host uses before you try the registrar.

NOTE

DMCA letter templates specific to each search engine can be found at *http://www.mcanerin.com/EN/articles/copyright-03.asp*.

Be aware that the DMCA takedown procedure is a double-edged sword. Like any other tool, a DMCA notice can be used for good or evil. Specifically, someone could use one unfairly against you! It happens—sites do sometimes use the DMCA to silence their competitors.

Therefore, you as a website owner need to protect yourself from unwarranted (or at least unwelcome) prosecution. If the potential exists for you to inadvertently host infringing material on your website—for example, if you are hosting online forums, group blogs, blog comments, or other types of content that can be submitted by others—here are some actions you can take to help protect yourself:

- It's helpful if you can qualify as a service provider that can be covered under the Safe Harbor provision. For example, you may qualify if you offer a search engine or a bulletin board system.

  — If so, notify your customers of your policies regarding copyright infringement and the consequences of repeated infringing activity. One way is to make this part of your Terms of Use.

- Publish a page on your website with DMCA filing instructions and state that, if and when you get a DMCA notification, you will act on it.

- Most importantly, check the directory of designated agents; if your company isn't listed there, complete the required form for inclusion in the directory and file it with the Copyright Office (*http://www.copyright.gov/onlinesp/list/a_agents.html*).

Regardless of the path you choose, it is helpful to be able to simply and rapidly prove that you are the original author of the content. One tool that can help with that is the Internet Archive's Wayback Machine. This site keeps copies of websites over time. You enter a website name, click Search, and get a screen that may look a bit like Figure 12-3.

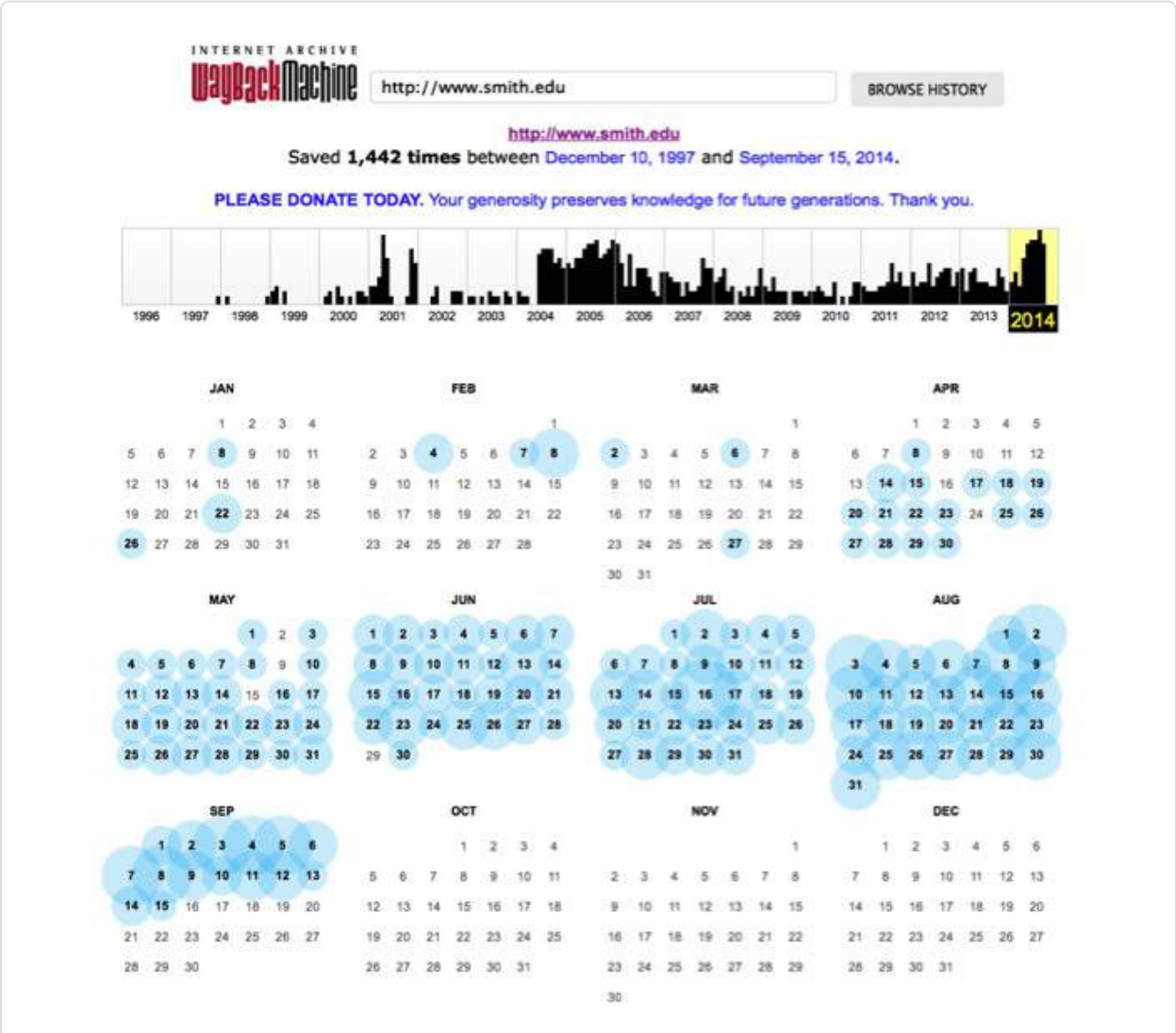


Figure 12-3. Sample archived website data from the Wayback Machine

You can then click on a date to see the state of the website at that time. The Wayback Machine does not always keep complete site copies, but a lot of data is available there, and it can often be used to provide clear proof that you were the first person to publish a given piece of content.

Changing SEO Vendors or Staff Members

Another major transition that can happen is when the publisher changes its SEO staff. This can happen both with in-house SEO staff and when you use an external SEO agency (e.g., the publisher can decide to switch agencies, or the agency can have internal changes). Companies make changes in their SEO staffs for many reasons, and issues can arise when this happens.

Potential Problems with SEO Staff Changes

One potential issue is a loss of momentum. The prior SEO team probably had some programs under way, and the new team will inherit those programs and have to sustain them, or there will be a delay before any results are seen from the new team's SEO efforts.

Even if the new team does its best to sustain an existing SEO strategy, it will take them some time to become familiar with it, and the specific tactics the old team used may not fit the new team's skill set that well. It is best to be prepared for this lost momentum when you're making significant changes in your SEO team. One way to try to deal with this is to hire the new team and have them work for a while in parallel with the old team. This can work in some situations, but it will not be an option in others.

Another situation that can occur is that the new team may have a different philosophy than the old team (this may be why the publisher made the switch). This will invariably mean a restart of efforts, and it can potentially take many weeks to regain the lost momentum.

SEO Documentation for Actions and Progress

One way to limit exposure to changes in the SEO team is through detailed

documentation. Keeping a detailed record of what has been done, when it was done, and why it was done can be invaluable in these scenarios. Ideally, this should be combined with your website analytics data.

Traditional web analytic tools make it easy to go back in time and see traffic data over time, but tools such as Google Search Console and Bing Webmaster Tools don't; nor do the search engines provide historical data on metrics such as indexed pages, link counts, and the like. You should also keep a record of this data over time. If your site meets the data inclusion criteria for services like SEMrush, you may be able to obtain historical ranking data going back up to two years for your site, which can help you understand past organic trends. Archive Google Search Console data every 90 days.

Make sure you have ownership of all relevant website accounts, usernames, and passwords. You don't want any transition to leave you in a situation where you can't access your critical data.

SEO Documentation for Rapid Training

Make sure that you have documented company processes regarding SEO, and that people associated with the SEO team and/or the management of the SEO team are all familiar with those processes. This can help speed up transitions as well, as the new team can be trained on what those processes are. Part of this includes having a detailed outline of the tasks and opportunities that are the first priorities for the new team.

With this information in hand, you can get things running smoothly much faster after a team reorganization. You can start the new team with one or more training sessions to get them up to speed on campaign history, login information, and what needs to be done going forward.

The better prepared you are to get new SEO staff up to speed and moving forward, the better off you will be. Bear in mind that you have hired the new team for their expertise; while the initial training sessions may result in the creation of a modified plan, you will get there much faster if you bring all these issues to the table on day one.

Cleanup and Auditing

Once your migration to new SEO staff (or a new SEO vendor) is complete, you should revisit and remove the old team's access to your analytics and search engine tools systems. This is one good reason to have universal, or "anonymous" business user and email accounts associated with important analytics and SEO logins. For example, tie your Google Analytics to a Google account with an email address like *mybusinessanalytics@gmail.com* or *analytics@mywebsite.com*, as opposed to the personal email account of someone who may or may not be with your company a year or two from now. Each user can be verified and removed individually. Audit this list on a regular basis.

Conclusion

Large-scale website changes to already established site properties are a fact of life, and they should be handled with as much foresight and planning as possible. In SEO, as in life, the only constant is change—and at the end of the day, we are grateful for this, as it means there is always opportunity for increased search exposure, and for new businesses to enter the playing field and succeed.

[1] Barry Schwartz, "Google Sends Hundreds of Thousands of Webmaster Notifications Each Month & 90% Are Black Hat Related," Search Engine Land, February 14, 2013, *http://bit.ly/black_hat_notifications*.

[2] Barry Schwartz, "Google Penalized One Article On BBC's Web Site," Search Engine Land, March 18, 2013, *http://bit.ly/bbc_penalized*.

[3] Barry Schwartz, "Google Has Officially Penalized Rap Genius for Link Schemes," Search Engine Land, December 25, 2013, *http://bit.ly/rap_genius_penalized*.

[4] For more information, check out "How Does Google Handle Duplicate Content?" on the Google Webmasters YouTube channel.

CHAPTER THIRTEEN

SEO Education and Research

Learning everything you need to know about SEO in order to be effective today, and in the months and years to come, is a difficult process that is made even more challenging and complex by the constant evolution of web technologies, search engines, searcher behavior, and the landscape of the Web.

Fortunately, many resources that make the job easier are available online. Leading SEO pros regularly post their thoughts on social media, blogs, and forums, and many speak at various industry conferences. These platforms create numerous opportunities to interact with folks doing SEO at a high level, to learn from their experience, and to keep your optimization skills sharp and up to date.

In this chapter, we will talk about how you can leverage the many resources available to be more successful in your SEO efforts today and in the future.

SEO Research and Search Performance Analysis

It should be fairly evident to you at this stage that SEO is always evolving. Search engines are constantly changing their algorithms (Penguin, Panda, Hummingbird, oh my!), and new media and technologies are being introduced to the Web on a regular basis. Staying current requires an ongoing commitment to research, study, and participate in the process of SEO.

SEO Resources

One of the easiest ways to research what is happening in the world of SEO is to study the websites and periodicals that cover SEO in detail. Simultaneously, ongoing testing of your (and others') SEO hypotheses should also play a major role in your day-to-day work if you hope to be, and remain, competitive.

Websites

A large number of online sites cover the search marketing space. Here is a short list of some of the better-known ones:

- Search Engine Land, owned and operated by Third Door Media

- Search Engine Journal, owned and operated by Alpha Brand Media
- Search Engine Watch, owned and operated by Incisive Media
- Moz, owned and operated by SEOMoz, Inc.

Each of these sites publishes columns on a daily basis, usually with multiple posts per day. The columns are typically written by industry experts who have been chosen for their ability to communicate information of value to their reader bases. Moz.com also provides a wide range of tools and resources for SEO practitioners.

Commentary from search engine employees

Search engine representatives sometimes actively participate in forums, or publish blog posts and/or videos designed for webmasters. The main blogs for the two major search engines at the time of this writing are:

- Google Webmaster Central Blog
- Bing Search Blog

The search engines use these blogs to communicate official policy, announce new products or services, and provide webmasters with useful tips. You can reach Google personnel via the Google Webmaster Help group in Google Groups. Members of the Google webspam team are active in this group, answering questions and even starting their own new threads from time to time. Google's YouTube channel also holds weekly Hangouts Office Hours, where people around the world can ask questions directly to Googlers. You will also find helpful advice via the Google Webmasters YouTube channel.

You can also interact with search engine representatives in various forums, such as WebmasterWorld, and in social media, often via their individual Twitter accounts. We will discuss the value of forums in more detail in "The SEO Industry on the Web".

Interpreting commentary

Search engine representatives are "managed" by their corporate communications departments, and some aren't even allowed to go on the record with statements

about how their search engines function. A rare few search engine reps have free reign (e.g., Google's Gary Illyes). Often they can't be very specific or they can't answer questions at all. The algorithms the search engines use are highly proprietary, and they need to be guarded as extremely valuable business assets.

This means there are certain types of questions they won't answer, such as "What do I have to do to move from position 3 to position 1 on a particular search?" or "How come this spammy site ranks so much higher than mine?"

In addition, the search engines have their own motives and goals. They will want to reduce the amount of spam in their search indexes and on the Web overall (which is a good thing), but this may lead them to take positions on certain topics based on those goals.

As an example, Google does not outline its specific methods for detecting paid links, but it has made clear that its ability to detect them has increased tremendously—and the updates to the Penguin algorithm over the past couple of years certainly prove it. Taking this position is, in itself, a spam-fighting tactic, as it may discourage people from buying links who otherwise might have chosen to do so.

In spite of these limitations, you can gather a lot of useful data from interacting with search engine representatives.

SEO Testing

SEO is both an art and a science. As with any scientific discipline, it requires rigorous testing of hypotheses. The results need to be reproducible, and you have to take an experimental approach so as not to modify too many variables at once. Otherwise, you will not be able to tell which changes were responsible for specific results.

And although you can gain a tremendous amount of knowledge of SEO best practices, latest trends, and tactics online, it is hard to separate the wheat from the chaff and to know with any degree of certainty that an SEO-related tactic will provide benefit. That's where the testing of your SEO efforts comes in: to prove what works and what doesn't.

Unlike multivariate testing for optimizing conversion rates, where many experiments can be run in parallel, SEO testing requires a serial approach.

Everything must filter through the search engines before the impact can be gauged. This is made more difficult by the fact that there's a lag between when you make the changes and when the revised pages are spidered, and another lag before the spidered content makes it into the index and onto the search engine results pages (SERPs). On top of that, the results delivered depend on various user-specific pieces of information, including the user's search history and location, the Google data center accessed, and other variables that you cannot hold constant.

Sample experimental approach

Let's imagine you have a product page with high SERP exposure in Google for a specific search term, and you want to improve the positioning and resultant traffic even further. Rather than applying a number of different SEO tactics at once, start testing specific tactics one at a time:

1. Modify just the HTML `<title>` tag and see what happens. Depending on your site's crawl frequency, you will need a few days to a few weeks to give Google and Bing enough time to recognize what you have done, and respond.

2. Continue making further revisions to the HTML `<title>` tag in multiple iterations until your search engine results show that the tag truly is optimal.

3. Move on to your heading tag, tweaking it and nothing else.

4. Watch what happens. Optimize it in multiple iterations.

5. Move on to the intro copy, then the breadcrumb navigation, and so on.

You can test many different elements in this scenario, such as:

- HTML `<title>` tag

- Headline tags (`<h1>`, `<h2>`, `<h3>`, ...)

- Placement of body copy in the HTML

- Presence of keywords in the body copy

- Keyword prominence
- Keyword repetition
- Anchor text of internal links to that page
- Anchor text of inbound links to that page from sites over which you have influence

Testing should be iterative and ongoing, not just a “one-off” in which you give it your best shot and never revisit the issue. If you’re testing `<title>` tags, continue trying different things to see what works best. Shorten them; lengthen them; move words around; substitute words with synonyms. If all else fails, you can always put a `<title>` tag back the way it was.

When doing iterative testing, do what you can to speed up the spidering and indexation so that you don’t have to wait so long between iterations to see the impact. For example, you can flow more link authority to the pages you want to test. That means linking to them from higher in the site tree (e.g., from the home page). Be sure to do this for a while before forming your baseline, though, because you will want the impact of changing the internal links to show in the search engines before you initiate your test (to prevent the two changes from interacting).

Alternatively, you can use the XML Sitemaps protocol to set a priority for each page, from 0.0 to 1.0. Dial up the priority to 1.0 to increase the frequency with which your test pages will be spidered.

NOTE

Don’t make the mistake of setting *all* your pages to 1.0; if you do, none of your pages will be differentiated from each other in priority, and thus none will get preferential treatment from Googlebot.

Throughout your SEO efforts, always remember that geolocation and personalization factors mean that not every searcher is seeing the same search results; therefore, you shouldn’t rely solely on specific rankings as the

bellwether for the impact of your SEO tactics.

Other useful SEO metrics

As we discussed in Chapter 9, there are many other meaningful SEO metrics, including:

- Organic traffic to the page
- Search engine spider activity
- Search terms driving traffic per page (to the extent you can determine keyword referrer data)
- Number and percentage of URLs yielding search traffic
- Searchers delivered per search term (again, Google's blocking of keyword referrer data can make this difficult)
- Ratio of branded to nonbranded search terms
- Unique pages spidered
- Unique pages indexed
- Ratio of pages spidered to pages indexed
- Conversion rates

An effective testing regimen also requires a platform that is conducive to your performing rapid-fire iterative tests, in which each test can be associated with reporting based on these new metrics. Such a platform comes in very handy with experiments that are difficult to conduct under normal circumstances.

Testing a category name revision applied sitewide is harder than, say, testing a `<title>` tag revision applied to a single page. Consider a scenario in which you're asked to make a business case for changing the category name "kitchen electrics" to the more targeted, search engine–optimal "kitchen small appliances" or "small kitchen appliances." Conducting the test to quantify the value would require applying the change to every occurrence of "kitchen

electrics” across the website—a tall order indeed, unless you can conduct the test as a simple search-and-replace operation, which you can do by applying it through a proxy server platform.

By acting as a middleman between the web server and the spider, a proxy server can facilitate useful tests that normally would be invasive on the ecommerce platform and time-intensive for the IT team to implement.

NOTE

During the proxying process, you can replace not only words, but also HTML, site navigation elements, Flash, JavaScript, iframes, and even HTTP headers—almost anything, in fact. You also can do some worthwhile side-by-side comparison tests: a champion/challenger sort of model that compares the proxy site to the native website.

Start with a hypothesis

A sound experiment always starts with a hypothesis. For example, if a page isn’t performing well in the SERPs and it’s an important product category for you, you might hypothesize that it’s underperforming because it’s not well linked-to from within your site. Or you may conclude that the page isn’t ranking well because it is targeting unpopular keywords, or because it doesn’t have enough copy.

In the case of the first hypothesis, you could try these steps to test your theory:

1. Add a link to that page on your site’s home page.
2. Measure the effect, waiting at least a few weeks for the impact of the test to be reflected in the rankings.
3. If the rankings don’t improve, formulate another hypothesis and conduct another test.

Granted, this can be a slow process if you have to wait a month for the impact of each test to be revealed, but in SEO, patience is a virtue. Reacting too soon to changes you see (or don’t see) in the SERPs can lead you to false conclusions. You also need to remember that the search engines may be making changes in their algorithms at the same time, and your competitors may also be making

SEO changes simultaneously. In other words, be aware that the testing environment is more dynamic than one using a strict scientific method with controls.

Analysis of TopRanking Sites and Pages

There are many reasons to analyze topranking sites, and particularly those that rank at the top in your market space. They may be your competitors' sites—which is reason enough to explore them—but even if they are not, it can be very helpful to understand the types of things these sites are doing and how those things may have helped them get their top rankings. With this information in hand, you will be better informed as you decide how to put together the strategy for your site.

Let's start by reviewing a number of metrics of interest and how to get them:

1. Start with a simple business analysis to see how a particular company's business overlaps with yours and with other topranking businesses in your market space. It is good to know who is competing directly and who is competing only indirectly.

2. Find out when the website was launched. This can be helpful in evaluating the site's momentum. Determining the domain age is easy; you can do it by checking the domain's whois records. Obtaining the age of the site, however, can be trickier. You can use the Internet Archive's Wayback Machine to get an idea of when a site was launched (or at least when it had enough exposure that the Internet Archive started tracking it).

3. Determine the number of Google results for a search for the site's domain name (including the extension) for the past six months, excluding the domain itself. To get this information, search for *<theirdomain.com>* -site: *<theirdomain.com>* in Google. Then append *&as_qdr=m6* to the end of the results page URL and reload the page (note this only works with Google Instant).

4. Find out from Google Blog Search how many posts have appeared about the site in the past month. To do this, search for the domain in Google Blog Search, then append *&as_qdr=m1* to the end of the results page URL and

reload the page.

5. Obtain the PageRank of the domain's home page as reported by the Google toolbar or a third-party tool.

6. Use an industrial-strength tool such as Moz's Open Site Explorer, Majestic SEO, or LinkResearchTools to analyze backlink profiles. These tools provide a rich set of link data based on their own crawl of the Web, including additional critical details such as the anchor text of the links.

7. If you are able to access a paid service such as Experian's Hitwise or comScore, you can pull a rich set of additional data, breaking out the site's traffic by source (e.g., organic versus paid versus direct traffic versus other referrers). You can also pull information on their highest-volume search terms for both paid and organic search.

8. Determine the number of indexed pages in each of the two major search engines, using *site:<theirdomain.com>*.

9. Search on the company brand name at Google, restricted to the past six months (by appending *&as_qdr=m6* to the results page URL, as outlined earlier).

10. Repeat the preceding step, but for only the past three months (using *&as_qdr=m3*).

11. Perform a Google Blog Search for the brand name using the default settings (no time frame).

12. Repeat the preceding step, but limit it to blog posts from the past month (using *&as_qdr=m1*).

Of course, this is a pretty extensive analysis to perform, but it's certainly worthwhile for the few sites that are the most important ones in your space. You might want to pick a subset of other related sites as well.

NOTE

> As valuable as website metrics are, brand names can sometimes provide even more insight. After all, not everyone is going to use the domain name when talking about a particular brand, nor will they all link. Thus, looking at brand mentions over the past few months can provide valuable data.

Analysis of Algorithmic Differentiation Across Engines and Search Types

Each search engine makes use of its own proprietary algorithms to crawl and index the Web. Although many of the basic elements are the same (such as links being used as votes), there are significant differences among the different engines. Here are some examples of elements that can vary in on-page SEO analysis:

- Role of user engagement measurements
- Strategies for measuring content quality
- Strength of social media signals
- Weight of `<title>` tags
- Weight of heading tags
- Weight placed on synonyms
- Value of internal link anchor text
- How internal links are weighted as votes for a page
- Duplicate content filtering methods

Similarly, there are many different ways a search engine can tune its algorithm for evaluating links, including:

- Percentage of a page's link authority that it can use to vote for other pages
- Weight of anchor text
- Weight of text near the anchor text

- Weight of overall linking page relevance
- Weight of overall relevance of the site with the linking page
- Factoring in placement of the link on the page
- Precise treatment of `nofollow`
- Other reasons for discounting a link (paid, not relevant, etc.)

A detailed understanding of the specifics of a search engine's ranking system, at least to the point of exact and sustained manipulation, is not possible. However, with determination you can uncover various aspects of how the search engines differ. One tactic for researching search engine differences is to conduct some comparative searches across the engines. For example, when we searched on *folding glass doors* in Google and Bing, we obtained the results outlined in Table 13-1. (Note that these search results differ daily by search engine, so the results you may see today will likely differ from what we saw at the time we conducted this search; the results also vary by searcher location.)

Table 13-1. Comparison of top five results for "folding glass doors" in Google and Bing

Google	Bing
http://*www.nanawall.com/...*	http://*www.solarinnovations.com/...*
http://*www.jeld-wen.com/...*	http://*www.lacantinadoors.com/...*
http://www.milgard.com/...	http://*www.jeld-wen.com/...*
https://www.marvin.com/...	http://*www.andersenwindows.com/...*
http://www.solarinnovations.com/...	http://*www.panorammicdoors.com/...*

There are some pretty significant differences here. For example, the search engines have only two sites in common (*http://www.jeld-wen.com/* and *http://www.solarinnovations.com/*). You can conduct some detailed analysis to try to identify possible factors contributing to *http://www.lacantinadoors.com* showing up in Bing, but not in Google.

At the most basic level, your analysis of competing websites may indicate that

Google appears to be weighting HTML `<title>` tags more heavily, whereas Bing seems to place greater value on keywords in the domain name. These are, of course, hypothetical examples of how you can begin to analyze SERPs across different engines for the same query to find differentiators across the search algorithms. However, there is a very large number of variables, so while you will be able to form theories, there is no way to ever truly be certain! Herein lies one of the persistent challenges of SEO.

The Importance of Experience

There are some commonly perceived differences among the search engines. For example, Google is still believed to place greater weight on link analysis than Bing (note that link analysis is very important to both engines). As we discussed in the second edition of this book, Google had fallen prey to much text-link forum/blog/article comment spam, with sites exploiting this tactic for higher organic positioning, and we expressed our hope that Google would improve its policing on that front. By the end of 2013, with various updates to its Penguin algorithm, Google had done just that by aggressively penalizing sites with low-quality, mostly paid, and often keyword-stuffed text links.

There are also institutional biases to consider. Given its huge search market share, Google has the richest array of actual search data. The nature of the data that the different engines have available to them can be influencing factors in how they make their decisions.

Over time and with experience, you can develop a sixth sense for the SERPs so that when you look at a set of search results you will have a good grasp of the key factors in play without having to analyze dozens of them. A seasoned SEO pro, when approaching a new project, will often be able to see important elements and factors that someone who is just starting out will not be able to pick up on. This kind of expertise is similar to that of a seasoned auto mechanic who can put an ear to an engine and tell you what’s wrong—that kind of knowledge doesn’t come from a blog or a book, but from years of trial and error working on many different engines. So while this book is a great resource that can help you get started, remember that actually *doing* SEO is the only way to build your abilities.

Competitive Analysis

Everything we discussed previously regarding analyzing top sites applies to analyzing competitors as well. In this section, we'll cover some additional analysis methods that can help you gain a thorough understanding of how your competitors in search are implementing their SEO strategies.

Content Analysis

When examining a competing website, ask yourself the following questions:

- What types of content are currently on the site, and how much content is there? Are there articles? Videos? Images? Music? News feeds? Answering these questions will tell you a lot of things, including how your competitors view their customers and how they use content to get links. For example, they may have a series of how-to articles related to their products, or a blog, or some nifty free tools that users may like. If they do have a blog, develop a sense of what they write about. Also, see whether the content they are developing is noncommercial or simply a thinly disguised ad.

- How rapidly is that content changing? Publishers with rapidly changing websites are actively investing in their websites, whereas those who are not adding new articles or updating content may not be.

- Are they collecting user-generated content? Sites that gather a meaningful amount of user-generated content tend to have an engaged user audience.

- Are they trying to generate sign-ups or conversions in a direct way with their content? Or is the content more editorial in its tone and structure? This can give you more insight into the way their marketing and SEO strategies are put together.

Internal Link Structure and Site Architecture

The organization and internal linking structure of your competitors' sites can indicate their priorities. Content linked to from the home page is typically important. For example, the great majority of websites have a hierarchy in which the major subsections of the site are linked to from the home page, and perhaps

also from global navigation that appears on all or most of the pages on the site.

But what else is linked to from the home page? If the competitor is SEO-savvy, this could be a clue to something that she is focusing on. Alternatively, she may have discovered that traffic to a given piece of content has a high conversion rate. Either way, these types of information can be helpful in understanding a competitor's strengths and weaknesses.

External Link Attraction Analysis

You can extract a tremendous amount of information via link analysis. A thorough review involves looking at both anchor text and link authority, and requires an advanced tool such as Open Site Explorer, Majestic SEO, and LinkResearchTools. Blekko also provides some useful SEO tools, as do SEMrush and Searchmetrics.

Conducting an external link analysis of your competitors allows you to do many things:

Dig deeper

What pages on the site are attracting the most links? These are likely the most important pages on the site, particularly if those pages also rank well for competitive search terms.

Determine where their content is getting its links

This can help you develop your own content and link development strategies.

Analyze anchor text

Do they have an unusually large number of people linking to them using highly optimized anchor text? If so, this could be a clue that they are engaged in broadscale link spam, such as paid article/blog commenting strictly for text-rich keyword linking.

If your analysis shows that they are *not* buying links, take a deeper look to see how they are getting optimized anchor text.

Determine whether they build manual links

Manual links come in many forms, so look to see whether your competitors appear in lots of directories, or whether they are exchanging a lot of links.

Determine whether they are using a direct or indirect approach

Try to see whether they are using indirect approaches such as PR or social media campaigns to build links. Some sites do well building their businesses just through basic PR. If there are no signs that your competitors are doing a significant amount of PR or social media–based link building, it could mean they are aggressively reaching out to potential linkers by contacting them directly.

If they are using social media campaigns, try to figure out what content is working for them and what content is not.

Determine whether they are engaging in incentive-based link building

Are they offering incentives in return for links, such as award programs or membership badges? Do those programs appear to be successful?

Determine their overall link-building focus

Break down the data some more and see whether you can determine where they are focusing their link-building efforts. Are they promoting viral videos or implementing a news feed and reaching influencers through Google News and Bing News?

What Is Their SEO Strategy?

Wrap up your competitive analysis by figuring out what your competitors' SEO strategies are. First, do they appear to be SEO-savvy? If they're not well versed and up to date with SEO, you may identify the following site characteristics:

- Poorly optimized anchor text (if the majority of the site's in-content links say "Read more" or "Click here," it is possible there is no extensive SEO expertise behind the scenes; research further).

- Content buried many levels deep. As you know from Chapter 6, flat site architecture is a must in SEO.

- Critical content made inaccessible to spiders, such as content behind forms or content that can be reached only through JavaScript.

In general, if you see sites with obvious errors that do not follow best or even good practices, the business is either not very SEO-savvy or is continuing to employ developers and use platforms that do not enable it to implement effective, SEO-friendly site development. These competitors may still be dangerous if they have a very strong brand, so you can't ignore them. In addition, they can fix a lack of SEO expertise by simply hiring a smart SEO expert and getting key stakeholder buy-in to make needed changes.

If, however, you are looking at a competitor who is very strong on the SEO front, you will want to pay even closer attention to their efforts.

Competitive Analysis Summary

If your competitors are implementing a strategy and it is working for them, consider using it as an effective model for your own efforts. Or, simply choose to focus on an area that appears to be a weak spot in their strategy, and/or one that leverages your current resources most effectively.

For example, if your competitor is not focusing on a segment of the market that has the potential to provide quality links through content development efforts, you can stake out that portion of the market first to gain a competitive edge.

In addition, if your competitors have previously bought or are still buying links, reporting them to the search engines is an option—but this is a decision you will have to make in consideration of your organization's business ethics. Their spam tactics will eventually be discovered, so it is just as well to focus your energies on building your own business.

Ultimately, you need to make some decisions about your own SEO strategy, and having a detailed understanding of what your competitors are doing can be invaluable in that process.

Using Competitive Link Analysis Tools

It can be frustrating to see competitors shoot ahead in the rankings and have no idea how they've achieved their success. The link analysis tools we have

mentioned can help you reverse-engineer their tactics by letting you quickly sort through the links pointing to a given site/page.

For example, if you were curious about how GoAnimate, a site providing business video creation tools, had earned its links, you could peek inside an advanced Open Site Explorer report and see that, according to Moz, there are 14,624 links to the domain from 2,332 domains, as shown in Figure 13-1.

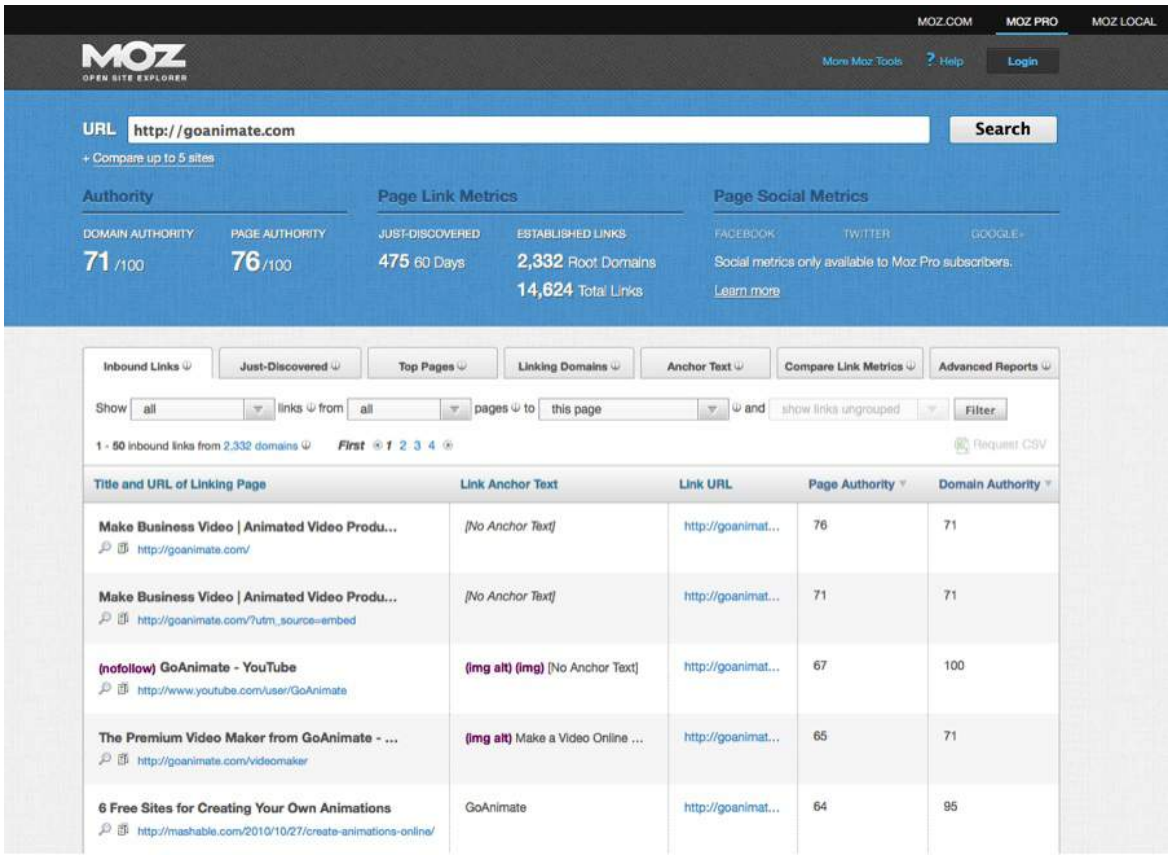


Figure 13-1. Exploring your competitor's inbound links

Looking at the data in Figure 13-1, you can see links that appear to come from content on the GoAnimate YouTube channel, as well as from tech article websites like Mashable.com.

This type of information will help your organization understand what the competition is up to. If your competitors have great viral content, you can attempt to mimic or outdo their efforts.

And if they've simply got a small, natural backlink profile, you can be more

aggressive with content strategies and direct link requests to overcome their lead. It's always an excellent idea to be prepared, and the ability to sort and filter out `nofollows` and internal links and see where good anchor text and link authority flow from makes this process much more accessible than it is with other, less granular or expansive tools.

Competitive Analysis for Those with a Big Budget

In Chapter 5, we introduced you to Experian's Hitwise, an "online competitive intelligence service" with a price tag that runs into the tens of thousands of dollars. If you have a big budget, there are four more expensive packages you might want to consider: RioSEO, Conductor, Searchmetrics, and seoClarity. These are powerful enterprise-grade SEO assessment and competitive analysis solutions that grade your website in a number of areas and compare it side-by-side to your competitors and their scores.

Using Search Engine–Supplied SEO Tools

Google and Bing both make an active effort to communicate with webmasters and publishers and provide some very useful tools for SEO professionals, and it is imperative that you verify your site(s) with these tools to take advantage of them.

Search Engine Tools for Webmasters

Although we have referred to the tools provided for webmasters by Google and Bing throughout the book, it is worth going into these in greater depth. Using search engine–provided tools is a great way to see how the search engines perceive your site.

Setting up and using a Google Search Console or Bing Webmaster Tools account provides no new information about your site to the search engines, with the exception of any information you submit to them via the tools, and the basic fact that you, the site owner, have an interest in the very SEO-specific data and functionality they provide.

You can create an account with either Google or Bing to access these tools quite

easily. An important part of creating these accounts is verifying your ownership of the site. Google provides you with the following options to verify your site:

- Add a meta tag to your home page (proving that you have access to the source files). To use this method, you must be able to edit the HTML code of your site's pages.

- Upload an HTML file with the name Google specifies to your server. To use this method, you must be able to upload new files to your server.

- Verify via your domain name provider. To use this method, you must be able to sign in to your domain name provider (for example, GoDaddy or Network Solutions) or hosting provider and add a new DNS record.

- Add the Google Analytics code you use to track your site. To use this option, you must be an administrator on the Google Analytics account, and the tracking code must use the asynchronous snippet.

Bing provides the following options to verify your site with its Webmaster Tools (*http://bit.ly/verify_ownership*):

XML file authentication

Click `BingSiteAuth.xml` to save the custom XML file, which contains your customized ownership verification code, to your computer, and then upload the file to the root directory of the registered site.

Meta tag authentication

Copy the displayed `<meta>` tag with your custom ownership verification code to the clipboard. Then open your registered site's default page in your web development environment editor and paste the code at the end of the `<head>` section. Make sure the `<head>` section is followed by a `<body>` tag. Lastly, upload the revised default page file containing the new `<meta>` tag to your site.

CNAME record authentication

This option requires access to your domain hosting account. Inside that

account you would edit the CNAME record to hold the verification code (series of numbers and letters) Bing has provided you. When complete, you can verify your ownership of the site.

The intent of these tools is to provide publishers with data on how the search engines view their sites. This is incredibly valuable data that publishers can use to diagnose site problems. We recommend that all publishers leverage both of these tools on all of their websites.

In the following sections, we will take a look at both of these products in more detail.

Google Search Console

Figure 13-2 shows the type of data you get just by looking at the opening screen once you log in to Google Search Console.

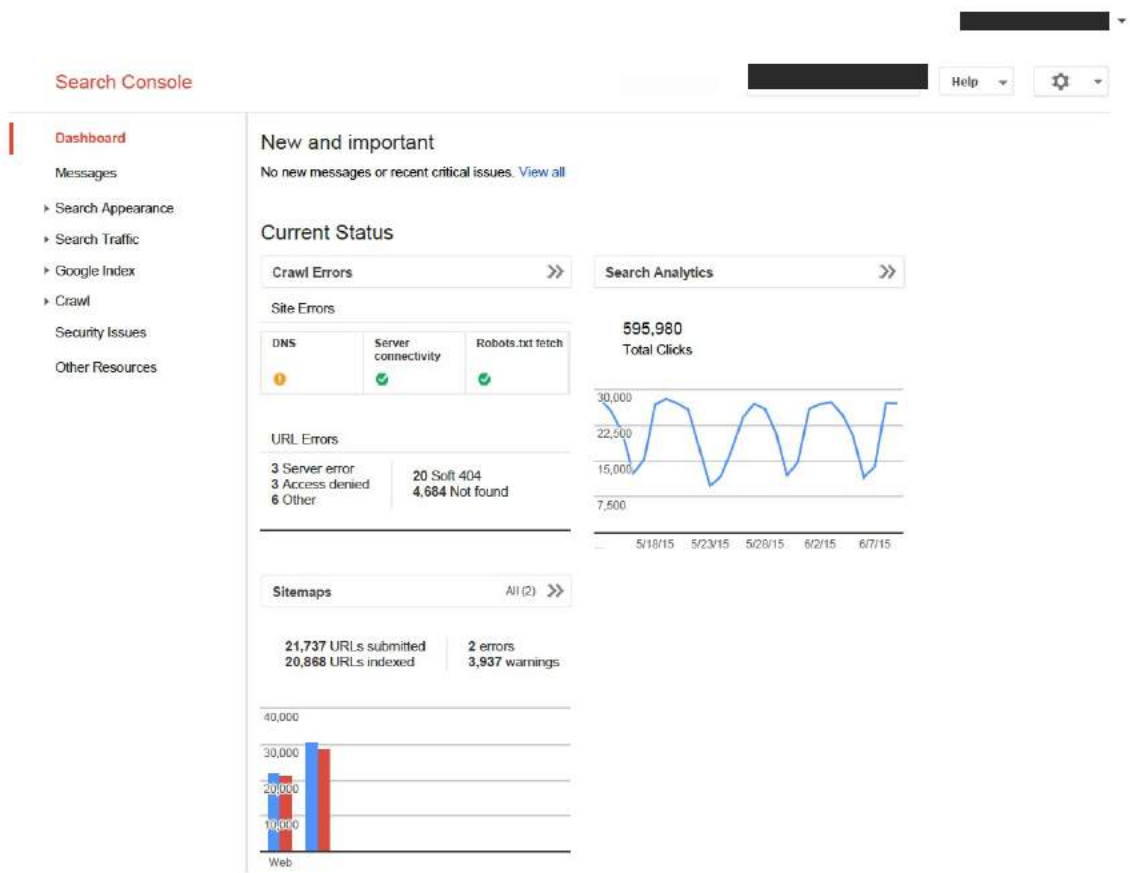


Figure 13-2. Google Search Console opening screen

Google updated its Search Console offering to provide additional information it maintains about your site, including Search Appearance. Figure 13-3 shows Google's Search Appearance Overview (which you can reach by clicking on the *i* in the grey circle next to Search Appearance), which brings users to an interactive overlay with links to information and instructions for specific components relevant to this section.

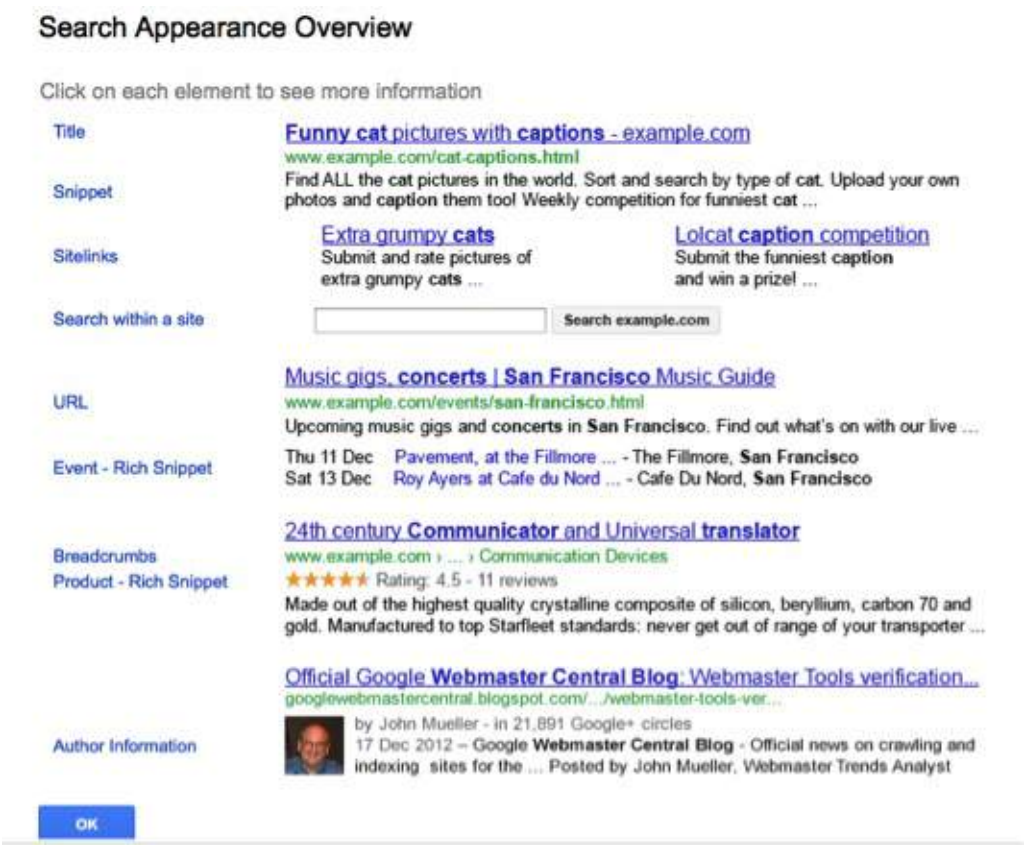


Figure 13-3. Google Search Appearance Overview

Search Appearance also gives you an inside look at potential problems with your meta description tags and your `<title>` tags, as shown in Figure 13-4.

You should investigate all meta description and `<title>` tag issues to see whether there are problems that can be resolved. They may be indicating duplicate content, or pages that have different content but the same `<title>` tag. Because the HTML `<title>` tag is a strong SEO signal, this is something that you should address by implementing a tag that more uniquely describes the

content of the page.

You can find valuable data in each of the report sections.

Another valuable section, called “Search Traffic,” shows Google search query, impression, and click data for your website (Figure 13-5).

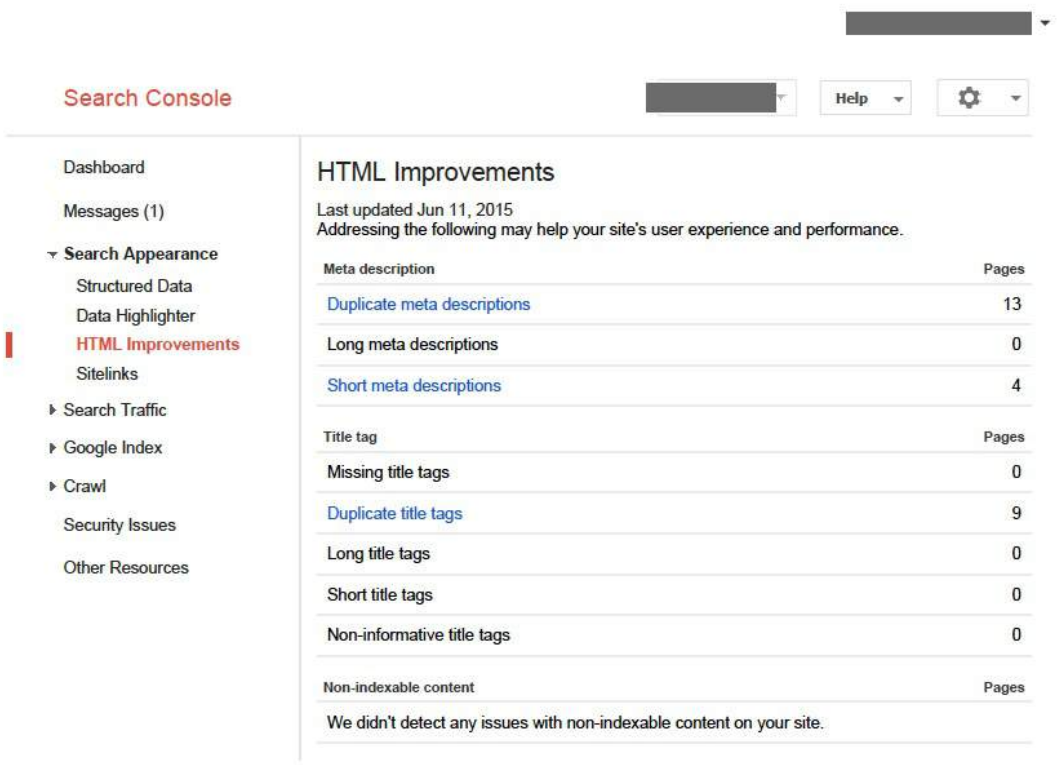


Figure 13-4. Google Search Console HTML Improvements report

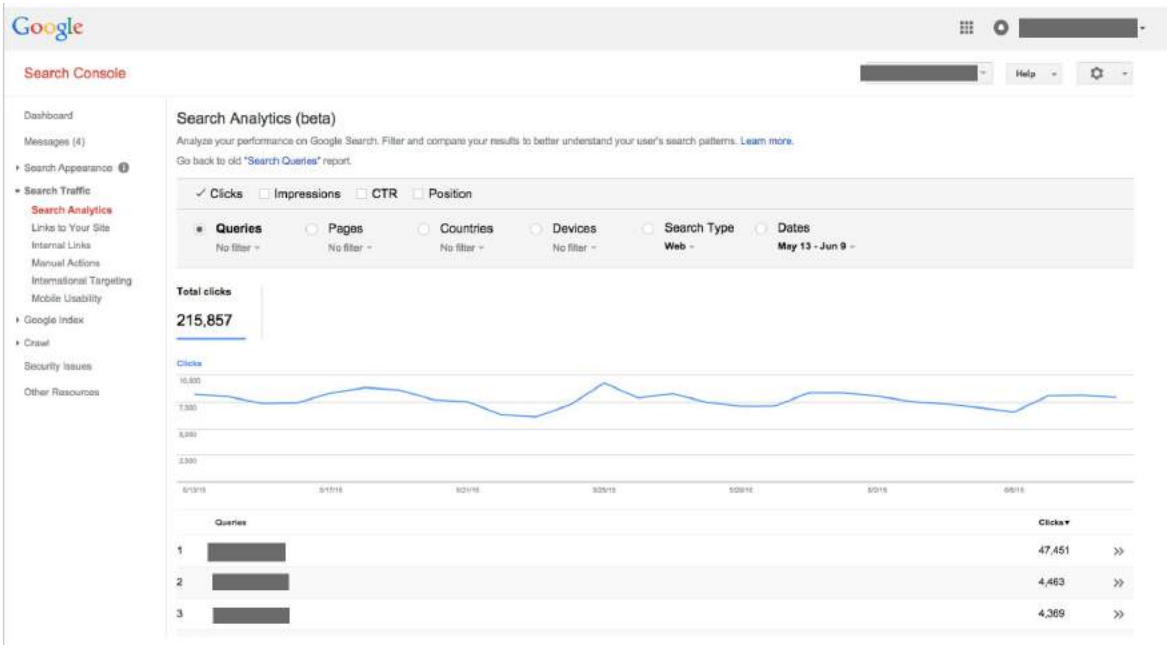


Figure 13-5. Google Search Console top search queries report

Figure 13-5 shows Google's view of which search queries the site is showing up for most often in the SERPs, how many impressions the site is receiving for each query, and how many clicks come from these impressions. You can also see data from Google regarding what the site's average position in the SERPs was during the selected period for each of the search terms listed. Note that this data is pretty limited, and most publishers will be able to get better data on their search queries from web analytics software.

Another valuable reporting area in Search Console is the Crawl section (Figure 13-6), which enables webmasters to look at Google's findings based on its crawling activities of the site's URLs and *robots.txt* file.

The charts in Figure 13-6 look normal and healthy for the site. However, if you see a sudden dip that sustains itself, it could be a flag that there is a problem.

The Crawl Errors report (Figure 13-7) is an incredibly valuable tool to identify and diagnose various site issues, from DNS resolution and server down time (Site Errors) to *robots.txt* and URL accessibility (URL Errors).

Figure 13-8 shows another diagnostic data point available in the Crawl section of Google Search Console: Blocked URLs.

It is common to find blocked URLs in your Search Console account. They can

occur when sites mistakenly restrict access to URLs in their *robots.txt* file; when that happens, this report can be a godsend. The report flags pages it finds that are referenced on the Web but that Googlebot is not allowed to crawl. Other times, blocking specific URLs is intentional, in which case there is no problem. But when your intention was not to block crawler access to your pages, this report can alert you to the problem and the need to fix it.

Another important diagnostic tactic is to look at Google's handling of your sitemap file from within the Crawl section, as shown in Figure 13-9.

Figure 13-6. Google Search Console Crawl Stats report

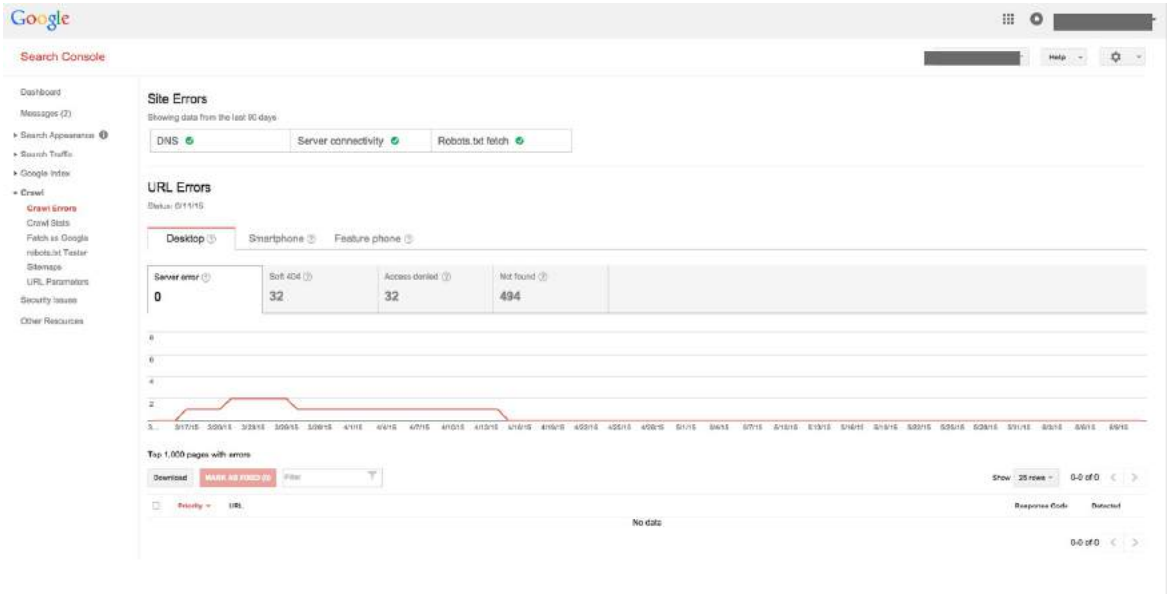


Figure 13-7. Google Search Console Crawl Errors report

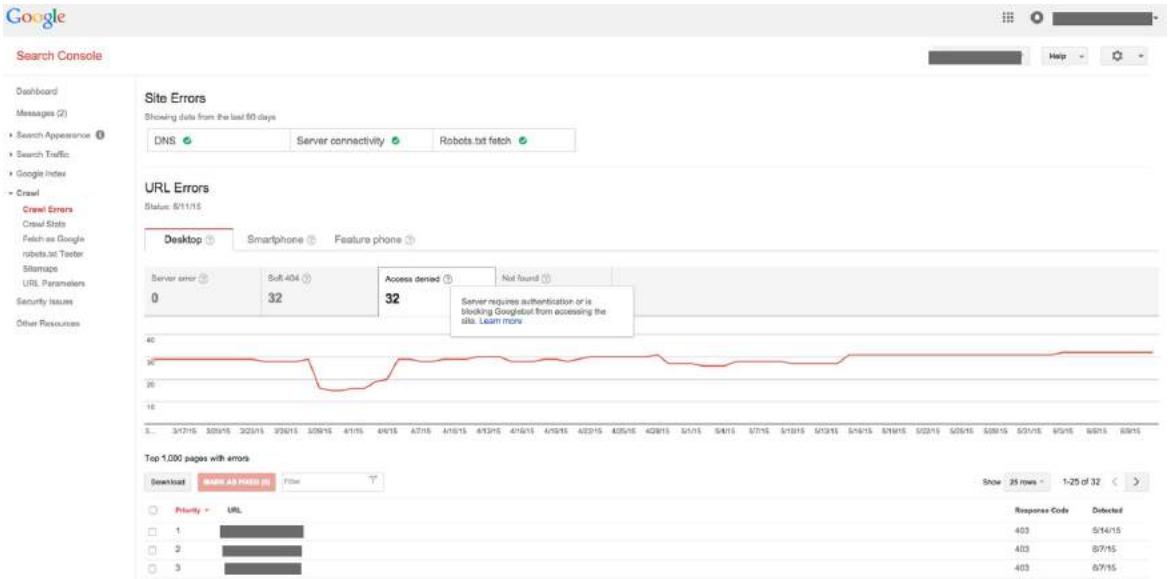


Figure 13-8. Google Search Console Blocked URLs report

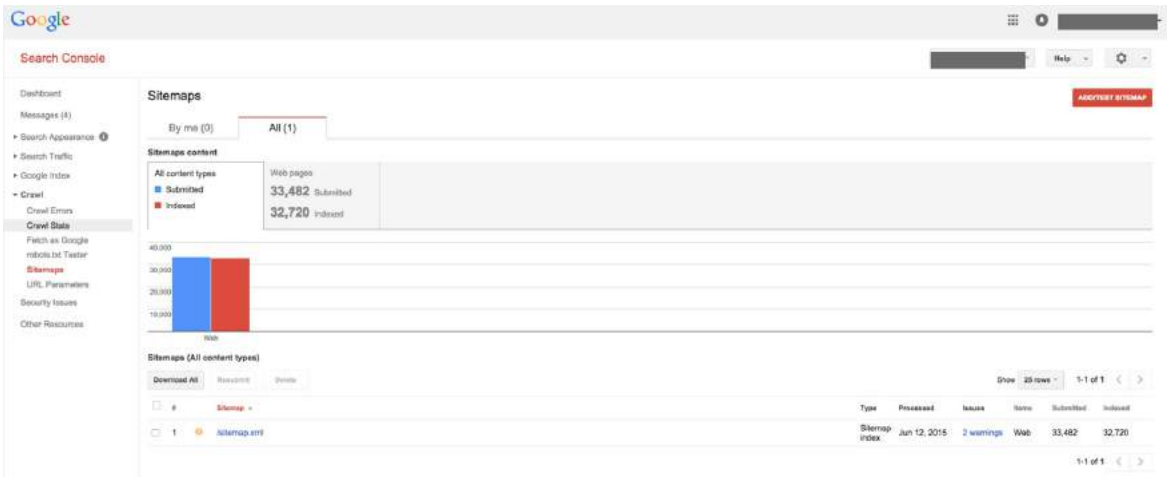


Figure 13-9. Google Search Console Sitemaps report

With this data, you can analyze the nature of any sitemap problems identified and resolve them, as you don't want to have any broken links in your sitemap file. In the example shown in Figure 13-9, we can see the "caution" icon to the left of the filename (*sitemap.xml*), and underneath the Issues column we see the linked text "1 warnings." Clicking on this brings us to detailed information about the warning, which in this instance appears to show that there is an unreachable URL within the sitemap (see Figure 13-10).

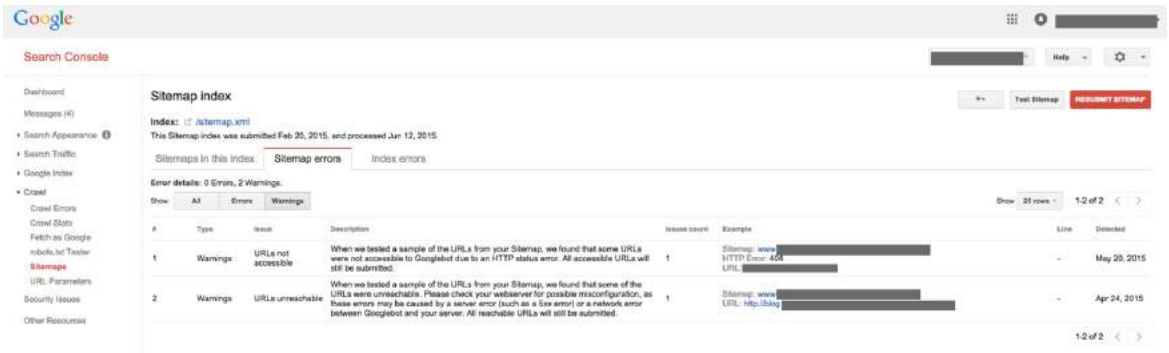


Figure 13-10. Google sitemaps error details

As Google continues to update its Search Console offerings, it is important to familiarize yourself with the options available to you as you perform SEO and manage your site(s). This information can be found in the Settings section of Google Search Console, and in the following list:

Geographic target

If a given site targets users in a particular country, webmasters can provide Google with that information, and Google may emphasize that site more in queries from that country and less in others.

Preferred domain

The preferred domain is the domain the webmaster wants to be used to index the site's pages. If a webmaster specifies a preferred domain as *http://www.example.com* and Google finds a link to that site that is formatted as *http://example.com*, Google will treat that link as though it were pointing at *http://www.example.com*.

Crawl rate

The crawl rate affects the speed of Googlebot's requests during the crawl process. It has no effect on how often Googlebot crawls a given site.

robots.txt test tool

The *robots.txt* tool (accessible at Crawl → Crawl Errors → Robots.txt fetch) is extremely valuable as well.

In addition to the services just outlined, Google Search Console has added functionality for site owners to increase their organic exposure by streamlining various processes that otherwise would be time-consuming to implement on the client side.

The Data Highlighter), for example, enables webmasters and site owners to "tag" structured data on site pages so that it can potentially appear as "rich snippets" of content within Google search results. While webmasters have always been able to mark up structured data using various formats, with Data Highlighter they can now "tag" structured data for Google's use without having to modify the site's pages on the server. More information about Google's handling of rich snippets and structured data can be found here: *http://bit.ly/structured_data_mrkup*.

Currently, Google supports rich snippets for the following content types:

- Reviews

- People
- Products
- Businesses and organizations
- Recipes
- Events
- Music

Figure 13-11 shows the main Data Highlighter page within your Google Search Console account.

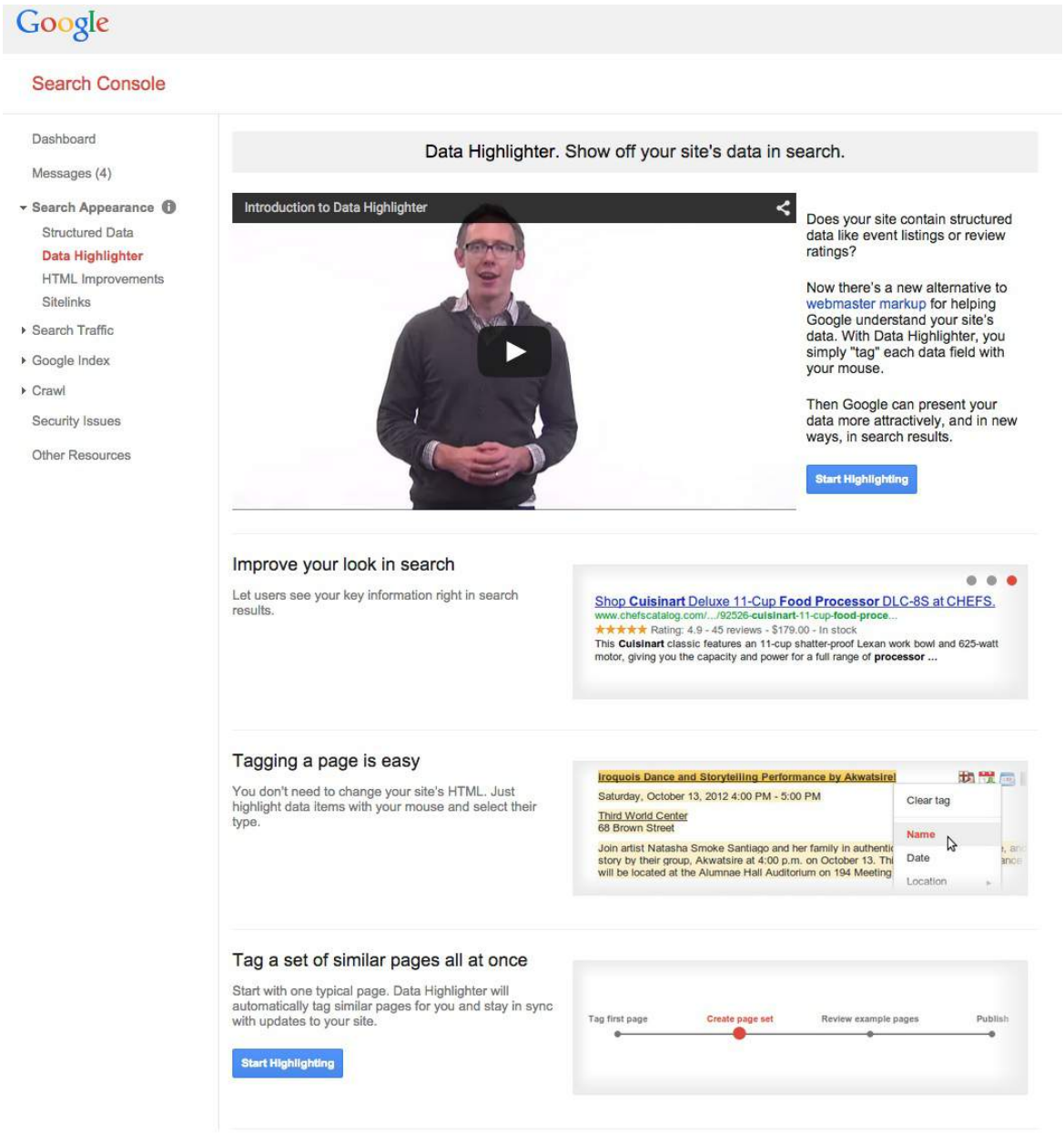


Figure 13-11. Google Data Highlighter

Clicking on Start Highlighting brings you to the entry fields to select the appropriate options for your site (Figure 13-12).

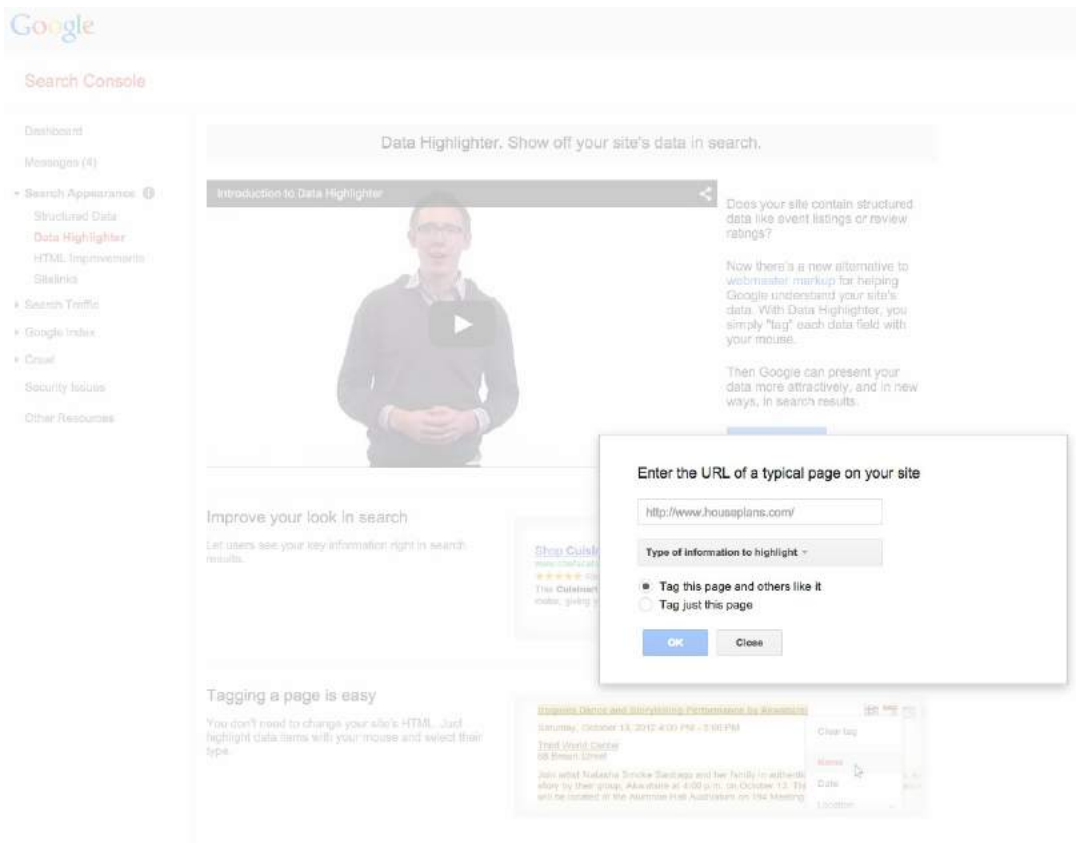


Figure 13-12. Google Data Highlighter options

In Figure 13-12, you would select the data type to highlight from the drop-down list, and then choose whether this data type is unique to this page. For example, if you have only one page with this type of data, but other pages have different data needing markup, you'd choose the option "Tag just this page"; conversely, if your site has thousands of product pages with the same structured data format, you can streamline the process and instruct Google to "Tag this page and others like it."

Manual spam actions

Google will also notify webmasters via the Search Console interface (and via the email notification settings specified within the account) of any manual spam actions it has taken against a site. Figure 13-13 shows the type of message you would receive from Google if it has identified "a pattern of unnatural, artificial, deceptive, or manipulative links" pointing to pages on your site (*http://bit.ly/unnatural_links_impacts*).

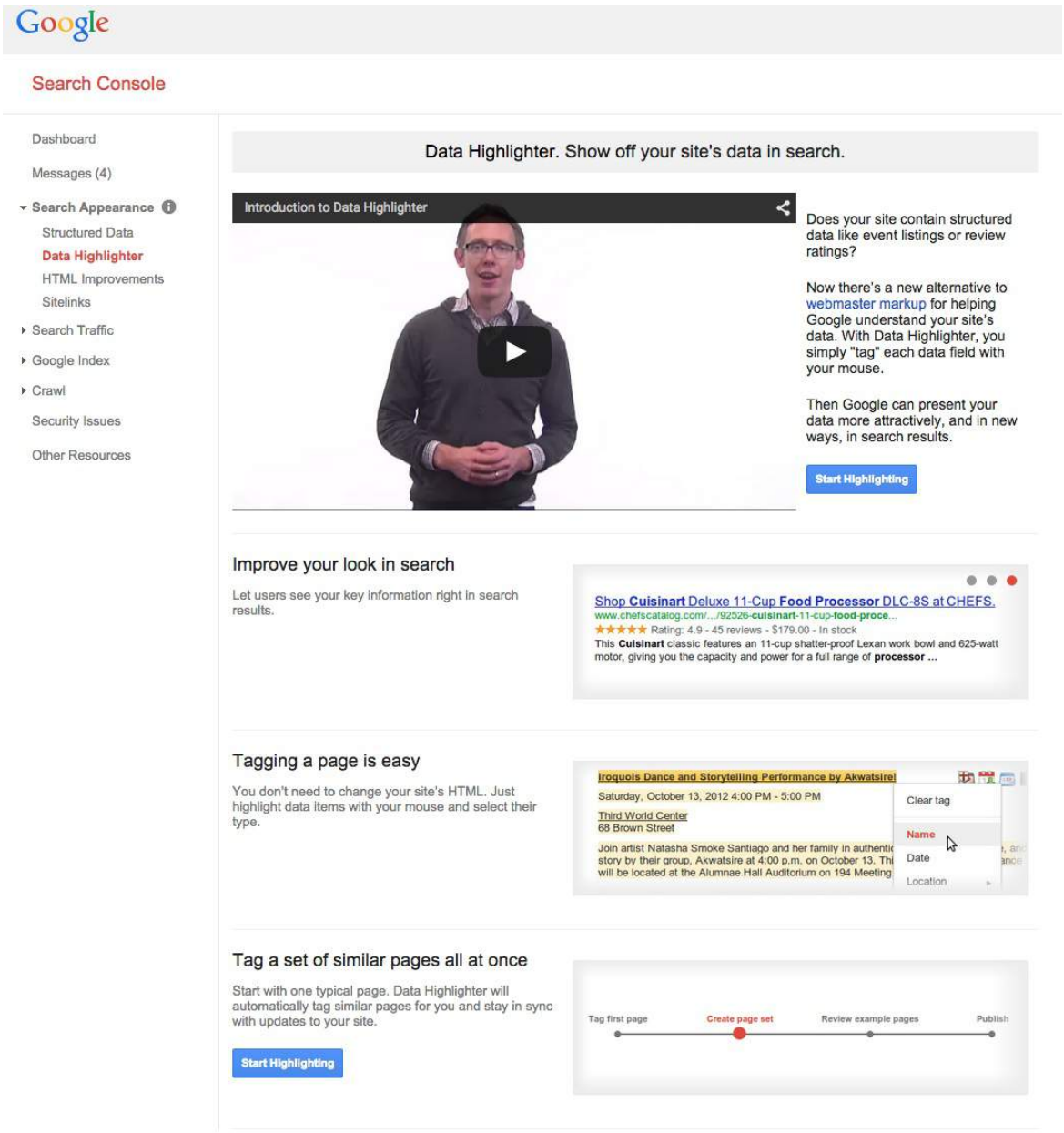


Figure 13-13. Google Search Console manual penalty notification

Bing Webmaster Tools

Microsoft also offers a product for webmasters with some great features. With Bing Webmaster Tools you can get a rich set of data on many of the same metrics that Google Search Console offers, but you are getting the feedback from the Bing crawler instead. Being able to see the viewpoint from a different search engine is a tremendous asset, because different search engines may see different

things. Figure 13-14 shows the opening screen for Bing WMT.

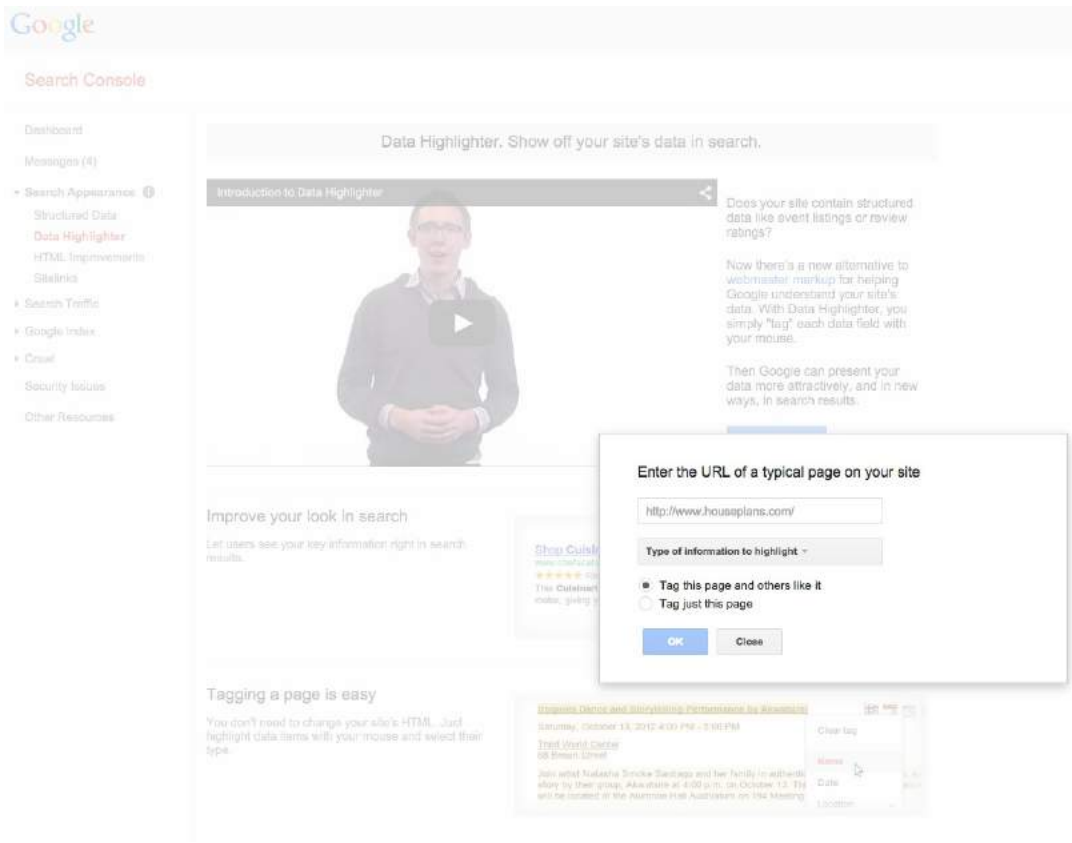


Figure 13-14. Bing Webmaster Tools opening screen

Already you can see some of the great data points that are available. At the top, you can see Clicks from Search, Appeared in Search (Impressions), Pages Crawled, Crawl Errors, and Pages Indexed.

To see Crawl Errors identified by Bingbot (Bing's search crawler), you would click on Reports & Data→Crawl Information, as shown in Figure 13-15.

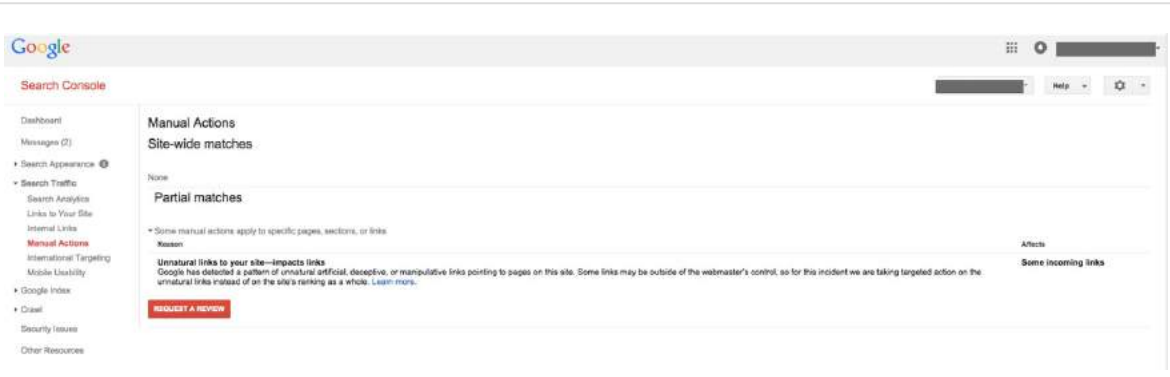


Figure 13-15. Bing Webmaster Tools Crawl Information report

You can pick from several options:

- 400–499(Request errors)
- 500–599 (Internal server errors)
- 301 (Pages moved permanently)
- 302 (Pages moved temporarily)
- *Robots.txt* Exclusion
- DNS Failures
- Connection Timeouts

Bing Webmaster Tools also includes an incredibly valuable tool called Index Explorer, accessible from Reports & Data→Index Explorer, and from the upper-right corner of the Crawl Errors report in Figure 13-15. This tool provides tremendous insight into your site's internal architecture as viewed by a crawler, often showing subdirectory folders that you, and your webmaster, did not know were still "live" and crawlable (see Figure 13-16).

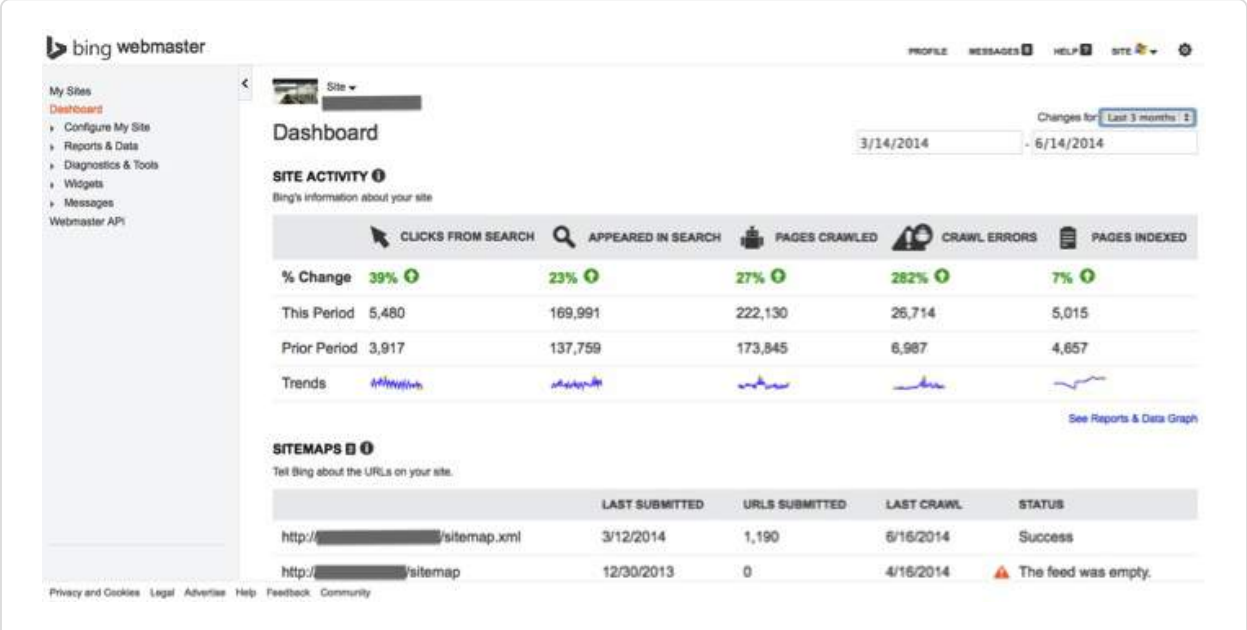


Figure 13-16. Bing Webmaster Tools Index Explorer

Bing Webmaster Tools also provides you with a section called SEO Reports, as shown in Figure 13-17.

Figure 13-17. Bing Webmaster Tools SEO Reports

This section provides data similar to Google's HTML Improvements report, and gives suggestions for where to improve SEO-relevant areas of your site's pages for better performance in organic search; focus on the items with High severity first, as these are likely to be the most impactful in the near term.

Bing also provides a new section called Diagnostics & Tools that is a mini-toolset in and of itself, with a host of valuable SEO functions (Figure 13-18).

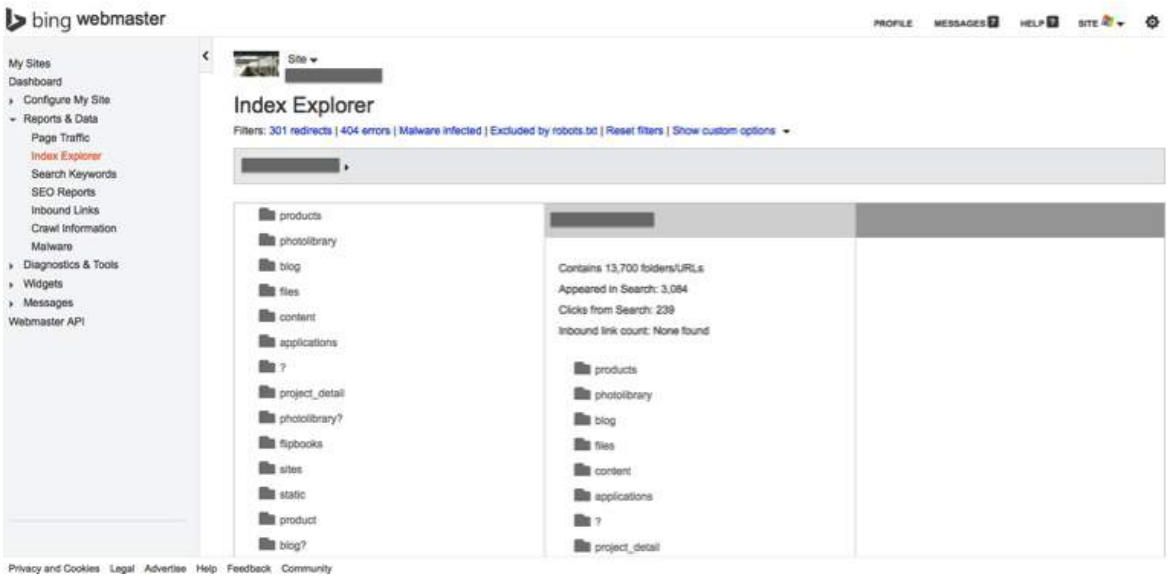


Figure 13-18. Bing Webmaster Tools Diagnostics & Tools

The first tool, Keyword Research, enables you to gather estimated query volume for keywords and phrases based on the Bing search engine's historical query database. Note that if you want only query data for the *exact* keyword or phrase you enter, be sure to select the Yes checkbox for Strict, just underneath and to the right of the keyword entry field. Checking Yes means that if you enter the phrase "feng shui santa barbara" you will see only query volumes specific to that exact phrase, not for phrases that include it, such as "feng shui consultant santa barbara."

NOTE

In the summer of 2015, Bing announced the move to HTTPS—ultimately blocking search query data in the referrer path, resulting in the inability to track the keyword searches performed for search traffic sent from Bing. It has announced that it will continue to provide "limited query data" from the Search Query Terms Report in the Bing Ads UI or API and in Bing Webmaster Tools reports (*http://bit.ly/bing_to_encrypt*).

There is much more here than we have been able to show in just a few pages, and the possibilities are limitless for how to utilize these valuable tools from both Google and Bing to your SEO advantage.

The SEO Industry on the Web

SEO-specific social media communities, forums, and blogs are almost always abuzz with new information and insights on the topic of SEO. Following this online community closely can help you increase your SEO knowledge and keep your skills sharp.

Blogs

A significant number of people and companies in the SEO industry regularly publish high-quality content about SEO and web marketing. This provides a great resource for newcomers who want to learn, and even for experienced SEO practitioners, who can pick up tips from others. However, note that not all top-notch SEO pros monetize their tactics by publishing them on social media or other publishing platforms, preferring instead to keep much of their expertise close to the chest.

The reality is that SEO is a deep and complex topic, and the opportunities for learning never cease. That is one of the most rewarding aspects of the profession, and fortunately, the large number of available blogs about SEO and web marketing make it easy to research the profession on a regular basis.

Be aware, however, that anyone can create a blog and start writing about SEO with self-appointed authority. There is a big difference between publishing an opinion and publishing content based on empirical knowledge. Many people have been able to create consulting businesses out of “flash in the pan” service offerings rooted in opportunism, as opposed to experience and love of the mystery, dynamism, and challenge of search. At all stages of your SEO research process and lifelong learning curve, it is important to know how to separate out those who know what they are talking about from those who do not.

One way to do that is to see which bloggers have the respect of other senior people in the industry. This is not unlike the process search engines go through to identify trusted authority sites and see who they recommend and trust. Proximity to clearly authoritative figures is a big plus. (This is why we always recommend that you run a backlink profile check before hiring a firm or consultant. If the majority of links to them are low-quality text link spam, you have a red flag.)

Another way to develop a voice is to become a contributor to these communities. As you accumulate SEO experiences, share the ones you can. Comment in blogs, forums, and social media and become known to the community. If you contribute, the community will reward you with help when you need it.

The search technology publications we discussed at the beginning of this chapter are a great place to start. However, a number of other SEO blogs provide a great deal of useful information. Here are some of the best, listed alphabetically.

First up are blogs by the authors:

- Eric Enge: Digital Marketing Excellence
- Stephan Spencer: Scatterings
- Jessie Stricchiola: Alchemist Media Blog

And here is a list of blogs by the search engines:

- Bing Search Blog
- Google Webmaster Central Blog

Here are some other highly respected blogs and website resources:

Search Engine Roundtable

Barry Schwartz's roundup of all things search-related

Rae Hoffman

Affiliate marketer and SEO pro Rae Hoffman's writings

Annie Cushing

Analytics expert Annie Cushing's writings on all things analytics, with SEO-specific information

Dave Naylor

Dave Naylor's coverage of a wide range of online marketing issues, with a focus on SEO

Small Business SEM

Matt McGee's blog focused on search marketing for small businesses

SEO Book

Aaron Wall's accompaniment to his excellent book on SEO

TechnicalSEO.info

Merkle/RKG's newly forming resource for technical SEO issues

You can also get a more comprehensive list by visiting Lee Odden's BIGLIST of SEO and SEM Blogs.

SEO News Outlets, Communities, and Forums

Much information can be gleaned from news sites, communities, and forums. Reading them does require a critical eye, but you can pick up a tremendous amount of knowledge by reading not only articles and posts, but also the differing opinions on various topics that follow those articles and posts. Once you develop your own feeling for SEO, interpreting what you read online becomes increasingly easier.

SEO news outlets and communities

- Search Engine Land
- Search Engine Watch
- Search Engine Journal
- Moz.com

Popular SEO forums

Search Engine Roundtable Forums

These are the forums associated with Barry Schwartz's Search Engine Roundtable blog.

SEOChat and WebmasterWorld Forums

Founded by Brett Tabke, WebmasterWorld is now owned and operated by Jim Boykin. These forums have been among the most popular places for SEO pros to discuss issues since 1996—nearly two decades!

WebProWorld Forums

This is a large web development and SEO forum with a very diverse group of posters.

High Rankings Forums

Jill Whalen, a well-known and respected SEO practitioner, is the founder of High Rankings. As of November 2014, Jill was still on her hiatus from SEO consulting, which she announced in 2013; however, there is still an active forum community.

NOTE

In Memoriam

In the second edition of this book, we sang the praises of longtime WebmasterWorld moderator Ted Ulle, lovingly known in the search industry as "Tedster." Sadly, Ted passed away in 2013, and the industry mourned (*http://searchengineland.com/mourning-ted-ulle-165466*). We miss you, Ted, and we thank you for your incredible contributions to our industry over the years.

Communities in Social Networks

A significant number of SEO professionals are very active on social media, while others are not—it is a matter of personal and professional preference (and available time!). You can find this book's authors on Facebook:

- Eric Enge: *https://www.facebook.com/ericenge.stc*
- Stephan Spencer: *https://www.facebook.com/stephanspencerseo*
- Jessie Stricchiola: *https://www.facebook.com/jessiestricchiola*

on Twitter:

- Eric Enge: *https://twitter.com/stonetemple*
- Stephan Spencer: *https://twitter.com/sspencer*
- Jessie Stricchiola: *https://twitter.com/itstricchi*

on Google+:

- Eric Enge: *https://plus.google.com/+EricEnge*
- Stephan Spencer: *https://plus.google.com/+StephanSpencer*
- Jessie Stricchiola: *https://plus.google.com/+JessieStricchiola*

on Pinterest:

- Stephan Spencer: *https://www.pinterest.com/stephanspencer/*

and on LinkedIn:

- Eric Enge: *https://www.linkedin.com/in/ericenge*
- Stephan Spencer: *https://www.linkedin.com/in/stephanspencer*
- Jessie Stricchiola: *http://www.linkedin.com/in/jessiestricchiola*

Social media engagement creates opportunities to interact with and learn from other people doing SEO. And, as in the forums, you can encounter search engine employees on these sites as well.

Participation in Conferences and Organizations

Conferences are another great way to meet leaders in the search space and network with them. The various panels may cover topics you're looking to learn more about, and this by itself makes them a great resource. Even experienced SEO practitioners can learn new things from the panel sessions.

Make sure you target the panels with the following objectives in mind:

- Attend panels from which you can pick up some information, tips, and/or tricks that you need for your immediate SEO projects or that broaden your

SEO knowledge.

- Attend panels where someone that you want to get to know is speaking.
- Engage during panels and presentations by tweeting your questions during the session (as available!).

NOTE

If you want to meet someone who is speaking on a panel, sit in the front row, and at the end of the session introduce yourself and/or ask a question. This is an opportunity to get some free and meaningful advice from experts. Get the speaker's business card and follow up with her after the event. Some panels also have search engine representatives speaking on them, so this tactic can be useful in meeting them as well.

Of course, you may encounter the people you want to meet in the hallways in and about the conference, and if so, take a moment to introduce yourself. You will find that most attendees are very open and will be happy to share information and tips.

In addition to the day's sessions, each conference has a series of networking events during the day and sometimes in the evening. These events are often held in the bar of the hotel where (or near where) the conference is being held.

Each conference usually has a published party circuit, with different companies sponsoring parties every night (sometimes more than one per night). In addition, some conferences are starting to incorporate more "healthy" ways to network and connect with industry peers during the events. Third Door Media's Search Marketing Expo (SMX) spearheaded this effort at SMX San Jose in 2014 by offering "Social Matworking"—yoga sessions for speakers and attendees taught by coauthor Jessie Stricchiola. Consider going to these types of industry networking events as a way to meet people and build new relationships.

Here is an alphabetical list of some recommended conferences for search-specific content:

Ad-Tech

This is an interactive marketing conference that has begun to address many of the issues of SEO and search marketing.

PubCon

Two to three times per year, Brett Tabke, founder and former owner of WebmasterWorld, runs conferences that attract a large group of SEO practitioners.

ClickZ Live

Formerly Search Engine Strategies (SES), then SES Conference & Expo, and rebranded in 2014 to ClickZ Live, this group runs 8 to 12 conferences worldwide. San Francisco is the biggest event, but New York is also quite substantial.

SEMPO

The Search Engine Marketing Professionals Organization, the largest nonprofit trade organization in the world serving the search and digital marketing industry and the professionals engaged in it, focuses on education, networking, and research. Founded in 2002 by coauthor Jessie Stricchiola and search industry leaders Brett Tabke, Christine Churchill, Kevin Lee, Dana Todd, Noel McMichael, Barbara Coll, and Fredrick Marckini, SEMPO has local chapters that host events, hosts frequent webinars and Hangouts, and is an ongoing presence with networking events at industry conferences.

Third Door Media Events

Third Door Media runs more than a dozen conferences worldwide, focusing primarily on search marketing (Search Marketing Expo, or SMX) but also covering other online marketing topics, such as social media (SMX SMM) and marketing technology (MarTech). Third Door Media also offers a conference specifically for advanced-level search marketers and SEO professionals called SMX Advanced, held annually in Seattle.

MozCon

Operated by Moz (formerly SEOMoz), these conferences are focused on what Moz terms “inbound marketing” and cover many areas of digital marketing.

More in-depth workshops are offered by:

The Direct Marketing Association

The DMA is the world's largest trade association for direct marketers. It provides quality instruction, including certification programs and 90-minute online courses.

Instant ETraining

Instant ETraining provides training on Internet marketing topics such as SEO, social media, paid search, and web usability. Its training videos are bite-sized, ranging anywhere from 20 to 50 minutes in length. Instant ETraining also offers live workshops for SEO and social media multiple times per year. Instructors include experts such as Eric Enge, Shari Thurow, Christine Churchill, Carolyn Shelby, and Bob Tripathi.

Market Motive

Market Motive offers web-based Internet marketing training and certification. Market Motive was one of the first companies to offer a credible certification program for SEO, and the faculty includes leading industry figures such as Avinash Kaushik, John Marshall, Greg Jarboe, and Todd Malicoat.

Conclusion

Mastering SEO requires a great deal of study and effort. Even when you achieve a basic level of effectiveness in one vertical or with one website, you have to continue your research to stay current, and you still have many other variables and environments to consider before you will be able to point your expertise at any site and be effective. Even then, you will always be working against some business-specific variables that are out of your control, and will need to adapt and modify accordingly. Thankfully, industry resources abound in the form of news sites, blogs, forums, conferences, industry organizations, workshops, training materials, competitive analysis tools, and search engine–supplied toolsets and knowledge centers. Take advantage of these rich sources of information and, if you are so inclined, participate in the vibrant online

community to stay in front of the SEO learning curve and ahead of your competition.

CHAPTER FOURTEEN

SEO Support: InHouse, External, or Both?

The Business of SEO

SEO has its roots in the web development and technology communities. Early search engine optimizers were web developers, small-business owners, and forward-thinking consultants looking to make their websites visible in the search engines and focused on tweaking HTML to make their sites more search engine–friendly.

As search has matured over the years and taken its place at the marketers' table as a legitimate sales and brand development channel, SEO pros have increasingly been asked to focus on the business side of the practice—to evolve beyond the code and Microsoft Excel jockeys of the early days of the practice (though let's be clear: Excel still serves some very valuable functions in SEO work!).

The benefits of focusing on SEO as a business are:

- It legitimizes the practice of SEO. Talking about `canonical` tags is great for chatting at an SEO conference, but talking about organic search share will go over better in the C-suite. In many companies, winning the C-suite battle to obtain budget and functional support is the only way to really succeed long-term in SEO.

- It bridges the gap between the practice of SEO and the impact on the business. Focusing on the business aspects of SEO will enable you to be clear about the ROI you are getting from organic search. Translating SEO-centric metrics such as rankings and traffic to revenue and dollars earned enables you to tie a tangible business impact to your efforts.

Understand Your Market Opportunity

Most SEO consultants probably cannot say with certainty what ranking in the top search positions for their keywords is (or would be) worth to them in terms of traffic and revenue. However, you should never set out to develop a new product or launch a new service without first understanding the business

case/market opportunity, and organic search should be no different.

Understanding the market opportunity prior to editing even the first `<h1>` tag is important, because it will help you to:

- **Understand what SEO is potentially worth to your organization.** It will give you an idea of what organic search visibility is worth to your business and answer questions like "How much organic traffic could we potentially procure from higher organic positioning?" Answers to these questions will guide the SEO process, including budget and infrastructure investments, and will help you make the SEO business case to the rest of the organization.

- **Track progress as you go.** Building a market opportunity benchmark prior to embarking on SEO will enable you to measure progress as you go and demonstrate the ROI of SEO in real, quantifiable metrics that will resonate with the company stakeholders.

A market opportunity assessment in organic search shows the current visitors captured for your keywords, versus the potential opportunity were you to rank in a top position. There are various tools available on the Internet that can help simplify the process.

To conduct a market opportunity assessment:

1. **Isolate your top group of nonbranded keywords.** Determine your target keywords based on their perceived conversion potential. The number of keywords may vary based on factors such as your SEO maturity or vertical market space, and can range from around 50 for those just starting out up to several thousand keywords for those further along in organic search.

2. **Gather search volumes and rankings for your keywords.** Find out what your current search volumes and rankings are for your targeted keywords.

3. **Use a click-through rate (CTR) curve to identify the potential for new visitors.** A CTR curve shows the expected percentage of clicks based on the search position of the keyword. Plugging your current search volume and rank into a CTR curve can give you a general idea of the number of new visitors you can expect to get by moving up in the search rankings.

4. **Scale that across your total traffic.** The first three steps of the analysis give you a sense of the traffic gain you can get for your top keywords, but it is likely that you have many other keywords driving traffic to your site. If you have an opportunity to double the traffic on your top keywords, you likely also have the opportunity to double your overall nonbranded organic search traffic volume.

5. **Check the CPC values for your keywords.** Using a tool like SEMrush or Searchmetrics, or the Google Keyword Planner from within your own Google AdWords account, run some numbers on the current market value of your keyword targets.

Get Buy-In Across the Organization

Now that you've done an opportunity assessment and know what you will be working toward in organic search, it's time to internally publicize the opportunity and obtain buy-in across the organization. For SEO efforts to be successful in any organization, active participation from many parts of the business—from C-level executives approving budget, to the engineering department implementing web and infrastructure changes, to content writers using SEO best practices—is an absolute necessity. Organizational buy-in is as critical to successful SEO as information architecture and `<h1>` tags.

To obtain organizational buy-in:

1. **Get people excited about SEO—then scare the heck out of them.** Show people the opportunity in organic search that has come out of your opportunity assessment. Make them aware of what the business is missing out on by not performing well in organic search. Then, make them aware of where your company stands relative to the competition, and identify competitors who are currently cashing in on the opportunities you are missing out on.

2. **Keep the SEO opportunity foremost in people's minds.** Make posters outlining the market opportunity and hang them on the walls. Speak about it often, and generally socialize it in the organization. The more you can keep the opportunity front and center in people's minds and keep them excited about it, the better.

3. **Create a service-level agreement (SLA) with stakeholders.** To formalize the working relationship with organizational departments and increase their buy-in to SEO, sign an SLA with them that defines how quickly they will respond to requests and the metrics that will be used to measure success.

Lay the Groundwork

Once you've obtained buy-in from the organization, it's time to position yourself to succeed. We've heard of many instances of chief marketing officers (CMOs) who have dipped a toe in the water with SEO but labeled it "black magic" or ineffective because the company's website failed to skyrocket to the top of the search rankings for the high-volume, highly competitive keywords they've targeted in 90 days or less. To avoid problems like this and to lay the groundwork for SEO success in your organization, it is important to do two things:

- **Set the appropriate expectations.** SEO is a long-term strategy where you are likely to see incremental improvements that may start slowly at first. Focus expectations on gains in nonbranded search traffic and conversions from that traffic, not on rankings. Many marketers' only exposure to search marketing is the world of pay per click, so it is important to educate them about SEO with its unique characteristics and differentiators.

- **Show early wins.** You may be tempted to go after high-volume, highly visible keywords right out of the gate, but ranking for competitive keywords can—and does—take time. It is better to initially target less competitive keywords that can result in quick wins (what we call "low-hanging fruit"), while simultaneously developing a long-term strategy for more competitive keywords. This will help keep people engaged and believing in the opportunity in organic search.

To help set reasonable expectations and report on your organic search progress in the most effective way possible:

- **Make a distinction between milestones and success metrics.** The fruits of your labor can take some time to materialize in organic search results. To maintain organizational buy-in and gauge your progress before traditional success metrics such as conversions kick in, report on key milestones such as

number of indexed backlinks, changes made to the website, and recommendations made.

- **Report on SEO metrics as business metrics.** Report on SEO like you would any sales effort, by continuously reporting on your search pipeline and the competitive landscape. Create a mission around SEO: you might strike a gong when certain milestones are hit, or install TVs where you can show regularly updated traffic and conversion stats.

- **Be a businessperson.** Understanding when to use technical SEO tactics is important, but speaking the language of business the C-suite understands is critical to the ongoing success of SEO in most organizations. Expand your horizons beyond Excel and analytics and learn to speak the language of dollars and cents—this is what executives will notice and understand.

Motivate Resources That Don’t Share Your Goals to Help You

As we’ve mentioned, succeeding in SEO requires support and buy-in across your organization. Many SEO efforts are dependent on non-SEO team members in the company to get things done, often relying on resources across numerous departments who may not be trained on, or oriented toward, thinking about SEO. Finding ways to work with others to get things done is key to SEO success.

To most effectively work with resources organization-wide on SEO tasks:

- **Show how it will benefit them.** Trying to get content creators to use targeted keywords in content can be a losing proposition. If you demonstrate conclusively, however, how it will benefit them—that they will reach a wider audience with their content—they are much more likely to buy in, implement changes to their content development process, and commit to the effort over time.

- **Recognize contributions to organic search visibility.** Recognize when team members do something that contributes to organic search success, such as optimizing a press release or tweaking an XML sitemap. At one Fortune 100 company, the director of SEO handed out an award every quarter to the individual who had contributed most significantly to organic search visibility.

- **Become an SEO evangelist.** Evangelism is just one more on the long list of job requirements for the SEO pro. One of the more important tasks is walking the halls, spreading the gospel of SEO. Buy the IT person breakfast, the PR director lunch, and the content writer dinner. Get people excited about organic search and then be sure to keep them that way!

Progress Through the Stages of SEO Maturity

Now that you have organizational buy-in and appropriate resources to execute, how do you go about scaling your keywords and continuing to drive organic search traffic to your website?

Our observations over the years have shown that organizations often move through five distinct stages when it comes to SEO. Each stage is characterized by investments in budget, strategy, and metrics appropriate to the stage's maturity level. Organizations should continuously be striving to move into the next maturity stage until they are at the "Compete" level and are actively vying to outperform the top firms in the search rankings. The stages are:

1. **Foundational.** In the first stage of SEO maturity, the organization dips its toe in the water and focuses on low-hanging fruit, optimizing for a small set of initial keywords. In the Foundational phase, the majority of the budget is spent on personnel, with a portion allocated to engineering/production. Efforts begin to move their top target keywords on deeper pages of the SERPs (page 7 and below) into striking-distance (page 4 or below), so they are primed to move into top-visibility positions in later phases. One employee and/or independent consultant is often assigned to SEO part-time.

2. **Investment.** In this stage, organizations begin to expand their coverage, moving into more competitive keywords with higher search volumes. Average ranks begin to improve as keywords optimized in the previous stage begin to move into page 2, 3, or 4 of the search rankings. Organizations now begin to invest in some link-building activities and content creation and start to expand their investment in SEO technology. One or two employees are assigned full-time to SEO.

3. **Measurement.** In this stage, to maintain investment in SEO, it becomes

critical for organizations to begin showing ROI and tying SEO metrics such as rankings and traffic to business metrics such as conversions and revenue. In the “Measure” stage, rankings continue to increase as the organization becomes increasingly SEO-savvy, beginning to invest in continuing education. Several full-time employees are assigned to SEO, along with specialized outside agencies and/or consultants.

4. **Scaling.** In this stage the organization organizes and prioritizes, beginning to look at organic search from a holistic perspective. An organic search pipeline develops as SEO starts to be treated as a sales campaign organizationally. Many more keywords move into top-visibility positions as investments in link building and content creation are increased. Several employees remain assigned full-time to SEO along with outside agencies and consulting specialists.

Building an SEO team

At this point in the SEO maturity cycle, you are ready to catapult yourself into organic search success by moving from relying on internal “generalist” SEO professionals with outside consultant specialists to building internal teams of specialized professionals. From content writers who write search engine–friendly content, to link developers who focus on leveraging that content to acquire strategically important links, to keyword jockeys adept at discovering hidden opportunities through spreadsheet analysis, to the project manager who pulls it all together, your organization is now ready to become a lean, mean SEO machine and drive to compete with the best the SERPs have to offer. This is the final stage:

5. **Compete.** In this stage the organization makes additional gains in organic search visibility such that it begins to look at its share of search relative to the competition. Budgets stabilize, while mature SEO technology is in use company-wide. In the “Compete” phase you are building internal teams of SEO professionals.

NOTE

A special thanks to Seth Besmertnik for his contribution to this portion of the chapter.

Using InHouse SEO Talent Versus Outsourced Support: Dynamics and Challenges

Once your organization has embraced the need for SEO, you must decide whether to hire someone inhouse or work with an outside consultant or agency. The same decision must be made when the time comes to expand an existing SEO team. Here are the factors you should consider:

- Is there enough work to justify a full-time, part-time, or contract employee?
- Is inhouse knowledge on SEO strategically necessary for the business?
- Are people with the right SEO skill set for your website available for hire within your target salary range? (What each company needs will vary.)
- Do you want to access and leverage experienced talent from an external agency to supplement an internal team?
- Is your organization so complex that it would benefit from the presence of an inhouse SEO pro?

The Value of InHouse SEO

One of the significant advantages of inhouse SEO is the ability to access search expertise whenever you need it. Some additional advantages are:

- A greater likelihood that SEO best practices will be integrated into the company culture.
- Greater accountability.
- Faster ramp-up time (familiarity with the core business).
- Faster implementation time (greater leverage in the development process).
- The ability to ensure that SEO requirements remain in scope (and raise the flag when they are descoped).

- The SEO expert can be tapped anytime and her work reprioritized immediately for quick deliverable turnarounds.
- The ability to provide cross-departmental communications around complex SEO issues, and support project management of ongoing SEO implementation.

The Value of Outsourced SEO Support

Although it may seem that inhouse SEO is the ideal solution, there are also plenty of reasons to use an outsourced team instead of, or in conjunction with, an inhouse team. Here is a summary of some of the best reasons to leverage outsourced talent:

- Finding expert-level (15+ years of experience) SEO talent to hire is difficult, so it may be smarter to contract with such talent.
- It allows you to focus on your core business and outsource your noncore activities.
- Experts generally have a broader reach in the marketplace via industry contacts, and as a result of working with many clients.
- SEO is multidisciplinary, so using external people makes it easier to pull in components when needed.
- Outsourced teams don’t have the myopic focus that sometimes occurs in long-time inhouse personnel.
- Outsourcing brings a fresh perspective and new ideas, and outsourced talent can often identify what inhouse staff may have missed.

Many SEO consulting firms, such as those of the authors, are willing to take on contracts that include training inhouse personnel. This approach can give you the best of both worlds: the ability to immediately inject a high level of expertise into your organization (albeit on a temporary basis), and a long-term plan to bring the capability inhouse, or at least build up the expertise with ongoing or occasional augmentation via outside consulting.

The Case for Working with an Outside Expert

You might want to work with an outside expert if any of the following scenarios apply to you:

- You don't have any inside SEO resources, are not in a position to hire them, and need someone to work directly with your development and marketing teams to get your SEO campaign going.

- You don't have any inside SEO resources, and while you want to build a team, you need to jumpstart the process in the meantime. You may also want to use your outsourced SEO practitioner to help train your team.

- You have an inside team but it is missing a key piece of expertise, or you just want to sanity-check the conclusions your team is drawing or see whether they are missing anything. In addition, because SEO at its best involves lots of brainstorming, an injection of fresh ideas can be just what the doctor ordered.

- You have an inside team, but they are busy working on other projects.

Regardless of the scenario, an outsourced SEO professional can bring two assets to the table that will be difficult for your internal team to provide:

The power of external authority

If you hire an industry-acknowledged expert, that person can often speak with a voice of authority that is not readily available to the inside team. This is not necessarily because the inside team are less senior; it's just because they are, well, inside. The view of an unbiased third party with industry credentials can be invaluable.

From management's point of view, the opinion of a recognized authority can increase their comfort that the right path is being pursued. It is very common for inhouse SEO professionals to engage an outside expert to help their cause. Engagements range from a presentation or training to site audits and quick consulting calls with web development and engineering teams.

The value of experience on multiple sites/campaigns

Outside experts will have the benefit of having worked on hundreds, if not thousands, of different sites and SEO strategies. Each site's SEO strategy is like a lab experiment in which the SEO expert can benefit from many different types of tests to learn what works well. In addition, if your outside expert is active in the SEO community, he may also benefit from discussing different experiences with other SEO experts.

How to Best Leverage Outside Help

As we've just discussed, there are several scenarios in which you might bring in an outside SEO expert. In all of these scenarios, you will benefit most by working with someone who is good with people, can effectively and diplomatically negotiate, and can teach. Bringing in someone you can trust is also a must.

To leverage these skills, you need to put your SEO expert in a situation where he can succeed. A good candidate will tell you what will be needed from you to be successful. Take this input seriously. If you can't provide the things you are being asked for, your ability to succeed will be limited. It is vital that management paves the way for successful SEO plan implementation. Although many managers may believe they are giving SEO projects what they need, this is not always the case, and as a result, SEO suffers. The key is for management to understand how to give the right type of SEO support.

For example, the SEO expert may need time from your development team to implement the suggestions he makes. If the development time is not available, his suggestions will be useless to you. Prearrange time with developers to talk to the SEO consultant, and work with project management to schedule time to work on the recommendations.

On the content side of things, the SEO expert may define a content plan to support ongoing content development and marketing efforts. Before you commit to a new hire, make sure you have a rough idea of the scope of the content development and marketing efforts the expert thinks might be needed to support this.

How to Implement Your Expert's Recommendations

Take your expert's input seriously, because he's telling you what he thinks he needs to be successful. If he doesn't get that support and ends up thinking he can't win, what are the chances he will succeed? The thing to remember about SEO is that implementation is crucial; you cannot implement a partial strategy and expect to see results. Unfortunately, most companies implement only a fraction of what is needed and wonder why their traffic isn't increasing, which is truly a shame—and why buy-in is so crucial. Remember, SEO is *extremely competitive*. There's a reason Google makes so much money with AdWords: because not everyone, and not every business, has what it takes to make high volumes of organic traffic happen. To be competitive in SEO isn't the same as finishing a 5K race; it's more like running a marathon and finishing in the top 0.001%. Do you think any of the top marathoners implement only 50% of their training, diet, and sleep regimens? Not a chance.

Ultimately, the cost of hiring an outside SEO expert goes beyond the direct cost for the expert's services. The SEO expert will need a certain amount of support to have a chance for success. Before you decide to hire an SEO professional, ensure that you have the internal support (and are willing to provide it) to make the recommended changes. You don't want to be one of those companies that pays for SEO advice that sits on the shelf and collects dust—we have seen this more times than we care to count!

You also need to have a plan to manage your outsourced SEO help, as high-quality SEO firms will also be placing demands on your organization. Quality SEO cannot happen in a vacuum. This includes content marketing, which requires high-quality content that represents your brand.

The types of internal resources you need to be able to provide include:

- Engineering and web development resources to implement the SEO firm's recommended changes to the website, on both a near-term and long-term basis
- Marketing resources to review plans and any content created to support those campaigns
- Content development resources to create any required content, whether it be adding copy to the website or creating content to support content acquisition strategies and link earning

- One or more project managers to oversee the entire process

How to Integrate SEO Knowledge in the Organization

Although it makes sense to involve an experienced SEO professional (employee or consultant) in your optimization efforts, it is important to remember that for SEO to be successful in the long term, you need to integrate SEO expertise into every aspect of your online business.

A basic level of SEO knowledge within the software and website development departments can speed up the production process, as a well-educated development team already knows a lot about how SEO affects what they do every day and can modify their efforts accordingly, on an ongoing basis. They will also be more likely to make the right decisions when the person in charge of SEO is not immediately available. Ultimately, this can save a large organization hundreds of thousands, maybe even millions, of dollars—because it won't be necessary to do a large amount of extra work to make the site search engine–friendly.

SEO is no longer merely about `<title>` tags and meta descriptions. Although this wasn't the case a few years ago, user engagement signals and website usability play increasingly large roles in search engine visibility today. It is vital that website designers think about the experience of the user who comes to the site from a search engine results page. Likewise, social media strategies need to incorporate keyword themes and links that search engines are looking for as signals in their ranking formulas, as well as the overall element of social engagement that leads to natural link development and increased brand awareness.

COLLABORATION WITH INTERNAL TEAMS

If you are in a larger organization, you must persuade your various teams/departments to have a stake in your site's SEO success because you will rely on them to implement many of the optimization efforts. For example, the engineering department can be your biggest ally, or a thorn in your side. They need to weave SEO into every facet of what they do, and that happens only when they want to work with SEO.

If conflicting priorities and insufficient internal resources stymie you with regard to any specific department, you must work hard to resolve the issues—or resign yourself to having to find alternative solutions. In some cases you can still accomplish what you need to with your site, but it will require a bit of out-of-the-box thinking and some technology.

Making sure your web development team has some basic SEO knowledge can dramatically reduce the work required of the SEO expert, so that when he does review your website for search engine friendliness and to create an SEO strategy, instead of spending time talking to you about Flash, JavaScript, header tags, and `alt` text, he can focus on more advanced solutions and strategies. The goal is for the organization as a whole to address 80% of the SEO workload, and the inhouse SEO guru or outside consultant to address the more complex 20% of the SEO workload.

Similarly, basic SEO knowledge is an incredibly powerful tool in both online and offline marketing. Many companies have failed to integrate their offline marketing efforts (such as television or radio commercials promoting a new product) with their online efforts, and SEO planning in this area can often make (or break) a product or website launch.

SEO knowledge at the management level is essential for similar reasons. Decisions made by management, which can affect the entire business, can often have disastrous effects on SEO efforts if they are done in a vacuum—one common example of this is contractual agreements with partners for content syndication. It is a common problem that even after the development teams, designers, and copywriters are trained on SEO, managers end up becoming the obstacle because they haven't been briefed on, or haven't entirely bought into, the aspects of SEO that executives need to know about.

When all of the members of an organization have and act on basic SEO knowledge, the likelihood of success is exponentially greater. The SEO specialist or team can focus most of their energy on new initiatives and continuing to develop the process, while the rest of your team makes good business decisions and taps them for on-demand answers to questions.

NOTE

A special thanks to Jessica Bowman for her contributions to this portion of the chapter.

The Impact of Site Complexity on SEO Workload

The amount of time that is needed to perform effective SEO depends on the size and complexity of the website, as well as the competitiveness of the market space the organization is pursuing. An organization's size and vertical scope also have an effect on the overall complexity of the SEO process. Here are some of

the ways that SEO considerations can affect the complexity of a site:

Keyword research

More pages mean more keyword research. Solid keyword research is needed to help drive the site architecture and content plans.

`<title>` *tags*

`<title>` tags (your page titles) are still an important ranking factor, and an important piece of the search results page itself. This means that you need to take the time to write the best possible copy for generating the highest click-through rates. For very large sites, you may have to design an algorithmic method to choose these for you.

Page content

The content must be unique and substantial enough for a search engine to understand what the page is about. Google's Matt Cutts has mentioned that two to three sentences of unique content can suffice to achieve ranking.[1] If you sell products and receive content from manufacturers, you need to invest the resources to write unique product descriptions; otherwise, you risk being omitted from the search engine's indexes or ranking lower because your content is a duplication of other sites' content.

Meta descriptions

Meta descriptions are important because search engines often use an excerpt from your meta description in the SERPs, and the description they provide for a page can influence its click-through rate. While you cannot directly control the description used in the SERPs, you can influence it. For example, by including the keywords that you are targeting with a page within its meta description text, you can make the relevance of your page clearer to the searcher. The larger the site is, the more writing you will have to do, because search engines value unique meta descriptions for each page on the site. For very large sites, you may have to design an algorithmic method to choose these for you.

Link development efforts

As sites scale, the complexity and need for links grow. You need to research the competitiveness of your targeted keywords, and make plans for link development so that execution neither grossly exceeds nor falls short of the necessary effort. The more websites/domains your company owns, the more link development is required. Likewise, the less authoritative your website is, the more link development work is required.

Web-based partnerships

Websites of all sizes engage in partnerships and relationships with other entities (charities, businesses, consultants, clients, distributors, agents, etc.). SEO professionals know that all of these partnerships represent opportunities for acquiring links and link relationships, and that when they are properly leveraged they can result in massive value-adds for the organization. For larger organizations, these partnerships can be more complicated in nature. The more complex the organization is, the longer it will take to leverage these opportunities for SEO.

PR team/agency

Your PR team is your friend when it comes to link development and content distribution, so be sure to integrate your SEO efforts with your PR efforts. Truly advanced organizations task a link metric to the PR firm they use so that the teams are asking for links with every story they pitch. You should too.

Development platforms and content management systems

The development platforms and content management systems used on larger sites can often create a number of limitations regarding SEO implementation, and frequently require extensive, costly, and time-consuming workarounds before optimization efforts can be implemented. If you have a non-search-friendly CMS (most are not search engine–friendly), you will have to do more customization work. It is recommended that you work with an expert to understand what is needed so that you develop it right the first time (because you may need to recode things every time you upgrade the CMS).

Solutions for Small Organizations

Some organizations are not equipped—either structurally or financially—to have an entire SEO team to handle the SEO workload we have been discussing. In fact, only one person may be knowledgeable about SEO, and that person may be only a part-time employee. Or there may not be anyone within the organization with the time or skills necessary to optimize the site, so outsourcing may be required. This section will give you direction on how to handle SEO in a small organization.

Developing the InHouse SEO Specialist

Building SEO knowledge inhouse can be challenging in a small organization where most of the employees are already performing multiple tasks. It is often good to have an SEO consultant on call to answer questions and validate solutions.

If you need to assign SEO to existing talent because you don't have the budget to hire an agency, consultant, or contractor, you should consider engaging an SEO professional to evaluate the aptitude of the person(s) you have in mind for the role. The cost of this assessment from an independent consultant can run anywhere from several hundred to a few thousand dollars, depending on a variety of factors. Countless companies have underestimated the specialized skills and investment required for successful SEO. It's a role unlike any your organization has seen before—this person needs the technical skills, marketing panache, and political savvy to work her way into your website development processes, and the innovative, outside-the-box thinking that will generate creative solutions to search engine crawler needs. Choosing the wrong person can be a costly mistake in terms of time wasted and missed opportunities.

As with larger organizations, it is important to develop a basic level of SEO knowledge throughout the organization. SEO still touches on management, marketing, and development, and it is important that all of these departments have a basic understanding of the issues.

We will outline some of the ways to effectively and rapidly learn about SEO and build up inhouse knowledge next.

Making the Most of Limited Resources or Budgets

Learning SEO and doing it on your own can be a challenging task, for two major reasons:

- The demanding, ever-changing landscape of search algorithm behavior is often unpredictable and nonintuitive.
- There are literally thousands of details and tactics to learn, some of which may have little impact on their own but, when used in various combinations with other components, can have a powerful influence on rankings. Herein lies the "art" aspect of mastering SEO.

Fortunately, many SEO training tools and materials are available via paid subscription at Moz, SEO Book, Instant ETraining, Market Motive, ClickZ Academy, and others. If you don't have the budget for a subscription, you can try public blogs and resources such as Moz.com and SearchEngineLand.com. Also consider the many resources for learning that we discussed in Chapter 11.

You can do numerous things at a fairly low cost to improve your site's overall optimization, including the following:

Use the free search engine tools

Use the free tools provided by the three major search engines. Create accounts in Google Search Console and Bing Webmaster Tools and verify yourself as a site owner on both. This will provide you with access to diagnostic tools, such as *robots.txt* validators, as well as reports on backlinks, spidering activity, server errors, top search queries, anchor text, and more.

Find the best keywords to target

Use the Google Keyword Planner to find keywords with high search volumes. Then use the Moz Keyword Difficulty and SERP Analysis Tool to get an estimate of how hard it would be to rank for the terms you have identified.

Check out your competitors

Assess your site and those of your competitors for SEO success factors such

as keyword-rich URLs, `<title>` tags and `<h1>` tags, keyword prominence, and so on. To survey your and your competitors' `<title>` tags across a large number of pages, use the search engines' *site:* operators and set (in the preferences) the number of results returned per page to 100.

Optimize your `<title>` *tags*

You want each `<title>` tag across your site to be unique and focused on a relevant keyword theme. Make each `<title>` tag count, because of all the elements on the page, it's what the search engines give the most weight; it also heavily influences the searcher's click decision from among the search results.

Optimize other critical elements

Analyze the text, HTML, inbound links, internal links, anchor text, and so on to determine your ideal configuration for success. Include a dose of your own critical thinking.

Measure, test, measure, and refine

Test your assumptions and the assertions of others—particularly SEO bloggers (not every piece of advice you find will be accurate or applicable). Measure key performance indicators (KPIs) and treat SEO like a series of experiments. Make iterative improvements to your URLs, title tags, `<h1>` tags, internal linking structure, anchor text, page copy, link acquisition efforts, and so on.

What sorts of KPIs should you measure and aim to improve? At a minimum, consider checking rankings, traffic, and conversion metrics. However, you can also check other metrics, such as the number of different search terms used to find your site (for this, use Google Search Console, SearchMetrics, or SEMrush), the number of different landing pages where search visitors arrive, the growth of inbound links and the addition of any notable inbound links, and so forth.

NOTE

We recommend these books to help you understand using analytics: *Web Analytics 2.0: The Art of Online Accountability and Science of Customer Centricity* by Avinash Kaushik (Sybex) and *Advanced Web Metrics with Google Analytics* by Brian Clifton (Sybex).

Test different ideas

Get one *great* idea for a tool, service, resource, or page, and bounce it off some folks in social media, participants in SEO forums, or privately through email to an SEO expert whom you trust. Hire a developer who can help you build it; consider leveraging your online relationships via LinkedIn, Facebook, and Twitter to mine recommended developer talent.

Many inhouse SEOs have had success finding copywriters via sites like Craigslist or Textbroker to write articles for nominal fees ranging from $10 to $50 per article (few companies and inhouse resources can compete with these rates), though remember the old adage that you get what you pay for, and this applies online as well. Again, consider leveraging your online contacts to find highly regarded talent within your budget.

Leverage low-cost tools

Consider using one of the following tools:

WordPress or Drupal

These tools are popular with dynamic and static website/application developers, respectively. Most of the time they are used to build web pages, but they also offer a range of reporting tools.

Xenu's Link Sleuth

This is a simple link-based crawler. Web developers use Xenu to check for broken links on a regular basis, but for SEO purposes the best value comes in the form of simple internal link analysis. By ordering the Xenu Sitemap based on "links in" and "page level," you can easily detect possible internal linking abnormalities that may interrupt PageRank flow or decrease anchor text value; and of course, you can save all this

information as a report. Xenu gathers loads of information, and it is a very useful tool, even for in-depth SEO purposes.

Screaming Frog SEO Spider

This is a small desktop program you can install on your PC or Mac that spiders websites' links, images, CSS, scripts, and apps from an SEO perspective. It fetches key on-site page elements for SEO, presents them in tabs by type, and allows you to filter for common SEO issues (or slice and dice the data how you see fit by exporting it into Excel). You can view, analyze, and filter the crawl data as it's gathered and updated continuously in the program's user interface.

Microsoft Word

Although it may seem to be an unconventional tool for a web developer/SEO practitioner, Microsoft Word is undeniably one of the best copywriting and publishing tools, and practically all users are familiar with it. It has several built-in features that help you to produce high-quality content, analyze existing content, fix and locate essential grammar errata, and above all, easily automate and synchronize all features and changes with other users and publishing tools. For more tech-savvy folks, there is always the scripting option for fine-tuning.

As with most SEO tools, the beauty is in the eye of the beholder. If you use the preceding tools properly, they can be very helpful, but if you lack experience or try to use them for the wrong kind of task, they can cause pain and misery.

Making proper on-page optimization decisions usually takes days. Even for a relatively small site, it is possible to cut that down to fewer than two hours by using the tools and methods we just described. Of course, there is a difference in the quality of work and the documentation you'll get with these tools compared to what you'd get with more conventionally priced SEO services, but they do have their place.

SEO webinars

There are numerous free SEO webinars available that will expand your knowledge. Great sources include Instant ETraining, Moz, Digital Marketing Depot, and SEMPO.

Limited cash options

If you're on a tight budget, create a blog with great content and promote it via social media (Twitter, Facebook, and Google+). Execute the SEO action items identified in this book, and leverage the free tools and guides we've mentioned. Attend free webinars. Do what you can for SEO. Then, invest in an hour or two with a consultant to do an ad hoc site review. You can also check with SEO firms to see which ones are offering special introductory packages, as some firms offer special pricing just to get started with a client.

These are just examples of things you can do, and countless other options are also available. Whatever you do, don't short-circuit the process of developing inhouse expertise, even if you hire outside help. That inhouse knowledge will be invaluable over time, and if you have hired a quality outsourcing partner, she will be delighted because it will make her job easier. She may even help train you or your team.

Solutions for Large Organizations

The challenges of performing SEO for a large organization are a bit different from those for a small organization. This is true regardless of whether you are working from an inhouse position or are an outsourced SEO consultant. Some of the challenges can stem from the size of the site, which can range from 10,000 to tens of millions of pages, potentially spread across multiple domains, countries, and languages.

Large organizations are usually complex entities, and many of these organizations may make decisions by committee, or review all decisions with a committee before finalizing them.

Many important projects (SEO or otherwise) can be delayed or even canceled in a large organization due to a lack of understanding by a single key player. In the world of SEO, efforts are often delayed by someone in IT/development,

marketing, sales, or management—and it can be challenging to get a meeting with the people you need to persuade, let alone accomplish the task of persuading them.

Patience and persistence are essential, and adhering to the following large-organization SEO guidelines is recommended:

- Get buy-in, if you can, from the head of IT, the head of marketing, the head of sales, and senior management. Those people will likely control your fate. If you can get two or three department heads and senior managers on board, you should be in good shape.

- Always make sure you talk about opportunity cost. Make sure all people involved understand that they're potentially leaving *X* visitors on the table every day, and that at the current conversion rate that means *N* potential leads and *Y* potential dollars. Include the opportunity gap—where the clients are versus where they could be—in every report.

- Insist on a sound web analytics plan. Successful SEO projects depend highly on quality analytics information. If your company currently cannot implement a quality analytics solution, yet you have support for implementing SEO, look at other KPIs, such as rankings combined with search volume and estimated clicks per position in the search results. You can also rely on the data from Google Search Console and Bing Webmaster Tools, which will provide some quality information.

- Provide detailed reports even when you're not asked to. As the disciplined people around you are likely to be unfamiliar with the basics of SEO, SEO practitioners need to overcommunicate.

- Educate, educate, educate. Always explain why you are making a given recommendation. Just make sure you are speaking in the language of the recipient (i.e., talk technically to the developer and in business terms to the marketer). This is where many SEO efforts fizzle out. When you conduct training properly, you can create an almost overwhelming amount of buy-in and interest in SEO. Companies often go wrong when they think one training session is enough. Unfortunately, interest will wane, and if you don't offer continuing education you can be back to square one in 6 to 12 months.

- Be confident. Cover your bases before each meeting. It's important to speak confidently on what you know, but it's also important not to speak about what you do not know. There is nothing wrong with saying, "That's a good question. I don't have an answer for you right now. Let me do some research and get back to you." Most development teams will respect that answer—if you can do your research and get back to them quickly. Be mindful of the Ugly Baby Conundrum: the website is the baby of the programmers, designers, and business sponsors; when you point out a site's SEO needs and flaws, it's like you're saying, "Your baby is ugly." When you look at a site for SEO, you will probably find many issues. Be sure to mention what has been done well in terms of SEO, and be unbiased when presenting the issues you've found.

- Don't make technical presentations to the executive team—they want metrics and action plans showing progress (or lack of progress). Although they might understand what you're saying, their brains are tied up with 99 other things. Just get to the point: what's gone well, what hasn't, and what you need from the team to fix it.

Contracting for Specialist Knowledge and Experience

Even if you have a solid inhouse SEO team, sometimes it makes sense to get help from the outside. Why would you want to bring in external expertise when you have a strong team already?

The answer lies in the complexity of SEO itself. SEO has dozens of subdisciplines, including video optimization, local search, image optimization, competition for search traffic in other countries, link development, usability, and strategies for social media properties. If you need help in one or more of these areas, don't hesitate to seek assistance from outside your organization. It's also possible that your SEO team has mid-level, but not expert-level, knowledge of SEO—and contracting out for expert-level guidance of your team is often a wise investment.

Another reason to bring in an SEO consultant is to conduct an audit of your site. This should help identify the things you've missed, confirm the things you've done right, and highlight new opportunities. Here are the ideal times to engage a

consultant to perform an SEO audit:

- When you need to understand the amount of work that will be involved in optimizing your site
- When you have done what you know about SEO and need advice on what to do next (or as part of an annual training, to help develop better expertise inhouse)
- When others in your organization need justification of your SEO recommendations
- Before a site redesign (or at any time you make major changes to a site), so that you can learn what needs to change and what you can do to make the next website more search engine–friendly

Applying SEO Recommendations Intelligently

One of the biggest challenges in SEO is the gap between strategy and implementation. The SEO artist knows how to explain his recommendations in simple terms and in the language of his audience. However, simple expression of a concept does not always mean seamless implementation.

When you look at a site from an SEO perspective, you must always keep the bigger picture in mind: to do what is best for the business by balancing both short-and long-term goals while managing available resources effectively. Changes that minimize implementation cost and hassle but have a big impact have the best ROI. Before you recommend throwing out that old CMS and reinvesting millions of dollars in a brand-new website, ensure that the ROI to the organization will justify the expense, and back it up with data.

This does not mean you should never recommend drastic changes, but rather that you should do so at the right time and place, and break out the implementation of the overall SEO strategy into prioritized SEO "projects." Be mindful of the potential benefit for each SEO recommendation. Sometimes this is difficult to quantify, and you may need to build a table similar to Table 14-1.

Table 14-1. Rating potential SEO projects

Project	Value	Effort
SEO Project 1	High	High
SEO Project 2	Low	Medium
SEO Project 3	Medium	Low
SEO Project 4	High	Low

This type of analysis is incredibly valuable in keeping an organization moving forward with an SEO program and increasing search traffic at the most rapid rate possible. Of course, sometimes the executives in your company may want to see dollar numbers instead of High, Medium, and Low ratings for the Value, so you may need to be prepared to give an estimated range of dollar impact. Just make sure to be clear that you are providing a very rough estimate, as this is usually difficult to predict precisely.

Hiring SEO Talent

If you have decided that you want to add to your inhouse SEO team (or start one), you need to think seriously about the type of person you want to hire. Do you want someone with prior experience, or do you want the person to learn from other team members along the way? Do you know which facets of SEO are crucial to the success of your specific website and organization, and how to hire for these skill sets?

Hiring top people can be prohibitively expensive, because there are a lot of income opportunities for them to pursue. It is hard to match the earning potential of a top SEO pro because, at the very least, he can usually make more from building sites and operating them in an affiliate or lead-generation model. That does not mean you cannot hire real talent; it just means you will pay a premium. An alternative would be to hire someone less seasoned and experienced, and pair her up with a consultant or agency to oversee the work.

Selecting the Right SEO Person

Although experience costs more, under normal circumstances it will bring faster results. On the other hand, a less experienced person is a lot like a rookie in baseball: if you hire a future star, you can potentially accomplish a great deal at a

much lower cost.

When sifting through résumés, examine candidates' specific skill sets. If you are looking for someone to develop social media campaigns, you should probably be a bit more focused on that than on whether the candidate is an expert at keyword research. No matter what his skill level is, he should "get" the Internet. Does he blog? Tweet? Have a professional and well-connected LinkedIn profile? If he doesn't, and there isn't a specific personal or professional reason why (privacy, confidentiality, etc.), he's possibly under-experienced for the position you are looking to fill.

Also, when evaluating a person's skills, don't forget to assess her political finesse. What makes a successful SEO practitioner at an agency or in a freelance position is different from what makes a successful candidate inhouse. Often, companies focus strictly on results and types of projects; however, if your company is plagued with red tape and political minefields, you need someone who can remain diplomatic, sell SEO to the entire organization, and integrate it into the organization in the right places so that it is implemented successfully. There is an 80/20 rule for inhouse SEO: 80% of the time you spend selling and 20% of the time you spend doing SEO.

Pitching the Person

As we've noted, finding good SEO talent can be challenging, as top SEO experts are in high demand and are likely to have more than one job option to pursue; for these SEO pros, a high six-figure salary is the norm.

Many companies overlook this. The "jobs offered" sections of forums are often full of offers, but few employers have made any real effort to market their openings. It really is a basic marketing problem that many of those vacancies show. Many don't even put a location in the title. Far worse is that almost all of the posts are company-centric: all about what the company wants from applicants, and no thought for what the prospective applicant will want from an employer. SEO is not like other job functions. A good SEO guru is in high demand; you have to make her think, "I want to work there!"

Fail to pitch the prospect, and you will get to choose from a diminishing pool of folks who can't make it alone, aren't able or willing to make a good living from affiliate programs, and don't prefer to be with a more forward-thinking and

proactive company that looks to develop its staff, rather than merely recruit them. Many SEO pros are looking for companies that have more support and resources; they are frequently stuck as isolated one-person departments, with few tools and “toys” to help them advance.

SAMPLE JOB OPENING: AN OLDIE BUT GOODIE

The following job opening appeared in the job postings at Search Engine Watch nearly a decade ago:

I’m totally serious about this unique opportunity for a person of the right calibre to come and work with Propellernet UK, a Brighton-based Search Engine Marketing company, as a trainee Consultant.

This is not your average trainee position, and a high level of skills are already expected. We do not have the time to spend years on this training, and we are looking to take on an experienced SEO, who already understands the special aspects of Marketing as it applies to the web. Then I will personally train and coach you to take that further still.

The ideal candidate will have a background in sales or marketing, as well as a broad understanding of SEO.

(Clue: If you don’t know who I am, your knowledge of SEO is not going to be broad enough.)

You must be a great communicator, as an essential part of this role is to bridge the communication gap that often exists between the Marketing and Technical Departments of Client Companies.

You need to be able to talk to Marketing Directors on their level, and then turn around and talk with Heads of IT Departments on their level without a pause.

You need to have equal passions for both Teaching and for Learning, as you will be expected to do both things each week as a matter of course.

You need high initiative, and the ability to think on your feet. You need to be proactive yet still have fast reactions too.

You must be in possession of a full, clean, UK Driving License, as considerable travel may be involved in the consulting work.

What you get in return:

A salary based upon ability and experience that will be reviewed regularly as your skills are proven. The salary will also include performance-related elements, and on-target earnings could easily surpass £70,000 as soon as your skills are up to the demand.

You get to work right on the cutting edge of SEO consultancy, where the work is never dull nor repetitive and we seek unique solutions for every client.

You get to be trained to be a world-class SEO and Web Marketing Consultant, able to stand alongside the very best that our industry has to offer.

I am looking to train an equal, not a junior, so you’ll be learning everything that I can impart to you just as quickly as you can absorb it.

This is a completely sincere offer. Your time will not be wasted—I don’t expect mine to be either.

CVs should be sent to Propellernet Ltd.—you can find the address for yourself if you’re a qualified

candidate.

Sincerely,

Ammon Johns (aka Black_Knight)

Marketing Director, Propellernet Ltd.

Clearly, this type of job posting is not what every company would want to put out there, but that is part of what makes it effective. There is a certain amount of attitude to it, and that speaks volumes to a potential applicant: either this is what you are looking for, or it isn't.

The job requirements also spell out many of the personal and professional attributes an applicant would need to have to get the job. The compensation is well detailed too, so the applicant knows what is in it for him. Good SEO professionals are well compensated and hard to attract. They know the value of their time and will not waste it on jobs they do not feel will give them valuable experience and just compensation.

In addition, the posting makes it clear that Ammon Johns was looking for someone who could be a true peer. This and many other points in the job outline make it evident that he is looking for a senior person. One thing that might make this type of listing a bit more effective would be the addition of some examples of search problems that a qualified applicant should be able to solve prior to applying (the "find the address for yourself" requirement at the bottom is not a challenging problem).

Making the Offer

Move quickly! Once you know the candidate is the right fit, make an offer. Good SEO professionals are always in demand, and waiting two or three weeks to make an offer could cost you the superstar hire. It is not uncommon for top-notch pros to land a new job in 3 to 4 weeks, nor is it uncommon for a company to search for a candidate for 6 to 12 months. In the SEO recruiting space, if you snooze, you lose. Prioritize time for the interviews, and act fast. If the process just seems too daunting, consider paying for a few hours of a top SEO pro's time to help you in the hiring process.

Selecting an SEO Firm/Consultant

Once you have decided to hire an outside firm, you need to go through some sort of selection process. This can be daunting if you are not already familiar with the space and if you don't already have some understanding of SEO yourself. There are many solid and reputable companies out there, but unfortunately there are a large number of bad ones as well.

In addition, there are many different types of projects you may want them to

work on. For example, if you have a local SEO–focused project in mind, you will need to find a team with the right type of experience. Similarly, you may have specific contract requirements that will impact which firm is right for you. Consider looking at resources such as the Moz vendor directory.

Getting the Process Started

Just as a bad hire can be a very expensive proposition, so can engaging the wrong consultant or firm, either through a haphazard selection process, or no selection process at all. Here is an outline of the steps you should take for a more methodical, disciplined search for outsourced SEO expertise:

1. Start with your goals. Connect with your team and make sure you have a good idea of what you want to accomplish by working with an SEO expert. Sometimes it's as simple as an audit to make sure you're following best practices; other times it's training for your marketers and content builders to get them up to speed on how to actively promote the site. If you have a solid list of things you want completed at the end of an engagement, you'll be better able to find the consultant or firm that best suits your needs—and you'll be better able to judge the eventual results.

   If you don't really know what you want, that's OK, but it does make the SEO practitioner's task a little less structured, and it means you need an exceptional level of trust in that person. Sometimes an SEO expert can help you define your goals—to get the best of both worlds, you can ask her what she thinks your goals should be and combine those with your own list. It is still fairly common to hear companies tell us, "We don't know what we need to do for SEO, but we need to do something; can you tell us?"

2. Connect with your social network. The people who can best assist you in finding a good match are often those who know you well. Talk to friends, fellow business owners, SEO bloggers, and people you know and trust.

   When you communicate with your network, make sure you share a little bit about your site, your business, and your goals with regard to SEO (to the extent that you know them). The more information you can share, the better folks are able to assist. At a minimum, if you have geographic

considerations or want the culture of a large, multi-offering agency versus that of a smaller, SEO-focused consultant/consultancy, this is good to include in the request.

3. Get advice from SEO-savvy people you trust. Familiarize yourself with the Twitter handles and Facebook pages for the people you trust, and engage. Tap your Google+ circles and LinkedIn connections to request advice and guidance. As a rule, SEO experts are incredibly friendly about referring business to good companies and good people, and if your network has already given you a few names, running these by the professionals can get you valuable insight on potential winners versus lemons. Even the authors of this book are willing to provide recommendations, so feel free to reach out to us.

4. Make sure the person is available. This is a fairly obvious step, but many people do not put this at the beginning of the communication process. Ask the person whether he is available at the start, and then continue the conversation if applicable (there is no need to waste energy on folks who can't help, though you should always ask for a referral). Your time (and theirs) is precious, so don't waste it.

5. Define what type of expertise you need. If you have an enterprise-level, highly complex website with legacy systems and a complex development life cycle, you need an SEO consultant who knows how to work within these constraints. If local SEO is your priority, you need someone well versed in local SEO. Few companies do extremely well in all facets of SEO, so you must identify which aspect of SEO you need most and find a company that excels in that area.

Preparing a Request for Proposal

A well-written request for proposal (RFP) can significantly improve the overall process. If your RFP is not written well, it hinders the SEO firm's ability to understand and define your needs and to scope and price your project. This in turn leads to a disconnect in expectations for both parties. A lousy RFP can discourage a busy SEO firm from even responding—which is a very unfortunate outcome, as it takes the best firms out of the running. That being said, there are

some SEO companies and consultants who do not participate in the formal RFP process, and with good reason: some businesses use the bait of an RFP to gain valuable SEO strategy tips, and even specific tactics, for free and at the expense of the SEO vendor. So while it can seem like a good idea to ask for a lot of great information in the RFP, this can place an undue burden on the RFP respondents and may result in some of the best agencies and consultants not responding, simply because they have other opportunities they can pursue without having to participate in a lengthy RFP process.

Many companies intuitively "know" what they want but find it difficult to ask for it in a way that is clear, succinct, informative, and constructive. If written properly, an RFP can facilitate the sales process and ensure that everyone involved on both sides understands the purpose, requirements, scope, and structure of the intended engagement. Following a few key steps in the beginning of the RFP process will enable you to rest easy, knowing that you are more likely to get what your company wants and needs.

Step 1: Nominate a "point person" for the engagement

One of the most challenging concepts for any large company with multiple working parts is to determine the end goal of the engagement. Often, marketing departments may voice different wants and needs than IT departments; even when they are asking for the same thing semantically, they may not be using the right language to communicate what they are looking for.

Every successful project needs a project manager who is invested in the project's success and can pull together the disparate groups that have a stake in the outcome or a role to play. Without that person at the helm, the project will struggle. By nominating that person as the SEO firm's "point person" even before you send out your RFP, you will ensure a steady flow of information throughout the process so that internal and external expectations are met.

Step 2: Define "needs" and "wants" using a decision matrix

RFP recipients will understand that you aren't the SEO expert, and therefore you aren't going to be able to adequately define the scope of your desired SEO engagement. As the saying goes, "You don't know what you don't know." Just try to be reasonable when formulating your needs and wants, and recognize that the RFP recipient may have a better idea of what you need than you do.

In addition to disclosing to the SEO firm the basics, such as your objectives, your site's conversion types (e.g., online purchases, newsletter subscriptions, white paper downloads, etc.), target audience, constraints, and so forth, in all fairness you will also need to clearly spell out the criteria by which you will evaluate the SEO firm's suitability. This means you will need to define these criteria in advance.

SEO firms know the selection process is usually governed by gut feel, so any RFP recipient who does not have a preexisting relationship with you is automatically inclined not to respond because he knows the deck is likely stacked against him. This is further exacerbated when you send the RFP to multiple SEO firms. You can allay this concern by candidly sharing with the firms your biases and the clearly defined criteria by which you will be evaluating them.

You should map this out in a *decision matrix*, which is simply a chart listing the attributes you are looking for in the SEO firm, a weighting factor for each attribute, a score from 0 to 10 for each, and the weighted score (e.g., the score multiplied by the weighting factor). The weighted scores are added to arrive at a total score. A highly simplified example of a decision matrix is shown in Figure 14-1.

Candidate SEO Firm Scoring Matrix

	Trustworthy	Trust Weight	Experience	Experience Weight	Availability	Availability Weight	Total
SEO Firm 1	8	0.4	6	0.3	9	0.3	7.7
SEO Firm 2	4	0.4	8	0.3	10	0.3	7
SEO Firm 3	7	0.4	7	0.3	8	0.3	7.3

Figure 14-1. Decision matrix for evaluating competing SEO firms

Judging criteria that are both quantitative and qualitative brings objectivity into a subjective process, which will aid you in managing expectations internally. And by sharing your list of criteria and the weighting factors with the RFP recipients, you increase the likelihood—as well as caliber—of responses.

We recommend that you place a high premium on trust in hiring an SEO firm. Unfortunately, there are lots of firms that claim a high level of expertise that either don't really have it or don't put in the effort to drive results (this is much

the same in many consulting disciplines). This is one factor on which we recommend you place a lot of weight.

Step 3: Define your success metrics

Now that you've identified what you are looking for in an SEO firm, it's important to let the firms know how you will measure the success of the engagement. Some firms can report on SEO health metrics well beyond just rankings (e.g., page yield, keyword yield), thus facilitating troubleshooting and reviews of program performance. Others will rely heavily on your own analytics package to track the program's success.

For example, you might pose this question internally: "What are your KPIs?" SEO firms often use KPIs to prove the value of the services they provide. If you aren't sure which KPIs your company would like to use, think about baking that request into the RFP, stating something such as, "Do you offer monthly program performance reviews indicating the program's growth? If so, please elaborate on the deliverables of such reviews."

You can assist this process by defining your business metrics clearly, whether they be sales, leads, views or shares of a piece of content, or something else. Make sure the SEO firm is under a nondisclosure agreement (NDA), and provide them with a baseline of traffic and conversion metrics and ask them to show what strategies they are going to use to grow your business from that baseline. Don't get distracted by artificial metrics such as rankings or number of links. While these may be of some interest, keep your eye on your bottom line and make sure that your prospective SEO firm knows that they will need to do that too.

Step 4: Prepare to disclose all known influencing factors

An SEO firm can easily examine your existing website, evaluating the on-page factors (`<title>` tags, navigation, HTML) to gauge project scope. But other factors will influence your organic exposure during the engagement, some of which won't be immediately obvious or known to the RFP recipient without prior disclosure. Important disclosures include:

- Has your site received a manual action notification from Google?
- Has your site seen traffic drops resulting from the Panda, Penguin, Mobile, or

other updates?

- Does your company have other domain names, subdomains, or microsites?
- Is a massive redesign of your website in the works, or are you about to migrate to a new CMS?
- Do you employ a third party for your site's internal search, and if so, who is it?
- Have you ever engaged in text link purchasing, article/blog spam, or link networks?

By disclosing as much information as possible, you spare the SEO firm the time and expense of discovering these things on their own and increase your chances of getting a good ROI in them.

Step 5: Provide an estimated timeline and budget for project completion

One of the biggest deterrents in any RFP is confusion over when a project should be completed and how much it might cost. A company hiring an SEO firm may not know how much time it takes to complete an SEO audit, but it may have pressing internal matters that require the project be done by a specific deadline.

Retail sites may want to schedule new launches around a particular theme or season; for example, if you want your site launched in time for back-to-school, be sure to write that in your proposal. Not only will an expected deadline save your company time in looking for an appropriate firm, but it also serves as a professional courtesy to the SEO firms you are querying.

Budgets often coincide with a project deadline, depending on how your company conducts its business. Consider assessing a budget based on a range of services; that is, instead of saying the project absolutely has to cost $*X*, say you are willing to spend within a range of $*X* to $*Y*. Budgets may or may not be included as part of the RFP, and there are benefits (and drawbacks) to both approaches. Even if you do not provide an SEO firm with your budget in the RFP, we recommend determining a budget beforehand because knowing how much money you are willing to spend will help you determine a target ROI.

Be aware that high-quality work at high speed is clearly going to cost more. Consulting with people you trust that have used SEO firms can help you get a sense of what the costs are likely to be. Once you have an idea of the market costs for a project of your complexity, you can then decide to spend more (to try to get faster results) or less (to manage budgeting concerns).

A sample RFP document outline

Now that you've gathered the preceding information for your RFP, you're ready to sit down and write it. You can structure an RFP in several ways. Here is a brief outline:

Section 1: Summary and overview

This section is where you will introduce the challenge you are having and provide the SEO firm with an overview of the rest of the RFP. We also recommend outlining how you would like SEO firms to respond to your RFP. Think of this section as an executive summary, where you will provide the highlights of the RFP without the technical details.

Section 2: Technical summary

Often, SEO firms need the gritty details to determine how they can best help you. The technical summary is where you will provide key pieces of information relevant to your project, such as technical requirements, a description of technical issues with the project, your site's current platform, and so on. The technical summary might originate from your IT department, as this section is often for an SEO firm's programmers and delivery team.

Section 3: Administration and management

By describing who will be involved in the project on your end and what the timeline is for completion, you are defining the framework of the project.

Section 4: Project expectations and delivery

If you've done your homework, this should be the easiest section to write. From outlining your evaluation criteria to specifying what monthly deliverables or training you would like to receive to assessing your KPIs, this section helps an SEO firm determine its cost to complete your project, as

well as its suitability.

When you write an RFP, keep in mind that the reason you're doing so is to hire an expert SEO firm that will propose its recommended actions for your company's website(s) to achieve the greatest chance for success. Sometimes an SEO firm's feedback highlights other potential issues that you may not have considered, which may change the scope of your project entirely—and this is often the case!

Communicating with Candidate SEO Firms

You should also plan to talk to candidate SEO firms. This is your opportunity to dig a little bit deeper and move beyond the RFP. If you are a smaller company and have bypassed doing an RFP, you will still need to be in direct contact with your prospective SEO firm. Here are some things for you to consider during this stage of the process:

Ask for a list of past success stories (not just clients)

Let the company tell you where it has helped someone be successful. You can ask for a client list, but be aware that a failed client engagement does not necessarily mean the SEO firm failed. In addition, many SEO service contracts prohibit the SEO firm from disclosing their clients. Many times clients don't implement the SEO practitioner's recommendations, or they try to, but do it badly.

You can use this to your benefit. Ask the firm for a case study of a client failure, and then ask them why they think it failed. This can provide some great insight into what they require from their clients, because the failure they disclose to you will inevitably be due to the client not providing them with what they needed.

Talk on the phone or (if possible) meet in person

Email is a great initial communication medium, but a phone call or in-person meeting gives you a real sense of the team you'll work with. Make sure you're not talking to a salesperson (if you go with a big company), but rather to the person who will be interacting with you throughout the contract

execution process. There's nothing worse than getting sold by a charismatic, knowledgeable SEO leader and then being passed on to a junior team for management.

In the phone call/meeting, be sure to establish rapport. If you don't feel comfortable on a personal level, don't dismiss it—brilliant SEOs are great, but if you can't work well together, the project is unlikely to succeed. Everyone works better with people they like and get along with. There's also the incentive of not letting down someone you've developed a working relationship with, which extends subconsciously and consciously into every part of the work you do together.

Present a few initial issues over email

Some SEO consultants won't engage in any work before a contract is signed, but you should at least get them to talk about how a problem can be approached, whether it's an issue, and what strategies they might recommend to fix it (even from a broad perspective). If you are an SEO practitioner, put a few examples of issues you've already identified on the table. Seeing how the consultant responds builds trust in her knowledge, and gives you insight into how she solves problems and what your work relationship and the advice you receive will be like down the road.

Ask for references

Asking for references is a good idea, but make sure you qualify the reference as part of the process. You want to protect yourself from people who might get some friend of theirs, instead of a real client, to act as a reference. It is helpful if the reference the candidate provides works within a well-known business. Make sure the relationship between the SEO firm and the reference provider is clear.

Once you get the reference on the phone, you should try to learn what the SEO consultant was asked to do, and what he did. Get specific details on this. If you want someone to provide link development strategy for you but he did not do link development for the other client, the value of that reference is decreased. Then see whether the reference will give you some

input on the benefits they received from the work of the SEO consultant (e.g., increases in traffic, conversions, or other metrics of value). Then verify what the reference says about the SEO firm by checking their site metrics with Compete or Quantcast.

Also, think about doing your own research. Find out what site the consultant worked on for the reference. Then use the Wayback Machine to see prior versions of the site if they're available.

Contact other past clients

Look at the SEO firm's site and see if it provides client lists. Call ex-clients the candidate firm didn't give as references. In addition, ask about good and bad experiences with the company in forums that cover SEO.

Making the Decision

Regardless of whether you have done an RFP or used a more informal process, make sure you understand the process you will use to make a decision. Here are some tips:

Get a written response from two or three vendors

Don't get an RFP response or quote from just one vendor, even if you're sure you're going to use them. Multiple quotes aren't just a best practice, they're a good way to learn about pricing and scope discrepancies. The quote from the vendor you don't like might include some pieces that your preferred vendor overlooked. Pricing is harder because it fluctuates so wildly in the search marketing world.

However, don't choose an SEO expert on price alone (or even make price the biggest part of your decision), unless your budget is a real problem. SEO is usually an incredibly high-ROI activity. Companies that spend larger amounts on SEO services tend to make it back in a matter of weeks or months from traffic and conversion increases (remember that more targeted traffic means higher conversion rates too).

Thus, going with an SEO expert who costs a lot less might seem like a good

idea, but if you don't work as well together and you think she might not do as good a job, you're hurting yourself in the long term. Of course, don't get fleeced by an exorbitantly overpriced firm (watch out; some of the biggest SEO companies have some of the most ruthless pricing models because they know that Fortune 1,000 companies will work only with consultants who've done work for other Fortune 1,000 companies), but be aware that the prospect of saving a few thousand dollars is not the best reason to choose a different provider.

Have smart, sensible people review the contract

Lawyers are great, but sometimes legal folks get overly concerned with liability and risk management details and overlook big-picture business items. Make sure your savvy business/ROI-focused personnel get a good look too. If there are legal issues that cause a rift, it can even be wise to get your C-level executives involved. Sometimes the folks from legal won't budge on an issue that's going to be a deal breaker, but if the CEO says do it anyway, you can get around the more problematic demands. Also be aware that different things are important to different firms. For example, some SEO companies will walk away from a deal if the client insists on a no-publicity clause, if you insist on them assuming millions of dollars in liability for a small engagement, or if it's got a "work for hire" clause, which the company has if they want to share but not lose ownership.

Go with your gut

When you finally make your choice, go with the team you feel will bring results. Making a matrix of price versus service versus reputation versus estimated productivity is fine if you're into that, but consider going with your initial gut feeling. If you don't feel like there is a "right" company, go back to the table, get more information, and even try to get another bid or two. When you do find the best choice, things just "feel" right. How do you know?

- The way they talk about search engines makes sense and the advice they've given lines up with the best practices you've seen expressed on SEO sites you trust.

- They are familiar with recent trends at the engines and the links they send or news they cite is timely, relevant, and logical.
- They use the engines like pros, rapidly combing through site queries, link information, and analytics data to unveil the underlying factors that are hurting your rankings or helping the competition to succeed.
- Their interactions with your team "fit." The team members like the consultants and are eager to implement their advice.

Choose someone you are comfortable dropping your guard around. This is your expert and ideally your confidant, someone you can reach out to and talk to about challenges and knowledge gaps in a way that you cannot with inhouse colleagues.

Remember, an SEO business relationship is a relationship, and if the two parties involved simply speak different languages (professionally and/or culturally) within their organizations, no amount of SEO expertise—and no amount of SEO budget—will make the relationship work. It is much smarter to work with people you like than to pay begrudgingly for a working relationship that simply doesn't work. Cultural fit is essential, so be sure you can identify your own business culture and then look for a compatible, resonant culture within the SEO company/consultancy you hire.

Mixing InHouse SEO with Outside SEO Help

When you are combining the resources of an external team with those of your inhouse SEO team, you need to determine who does what, who decides what is to be done, and so on. It is best to have this all clearly defined up front, or else there will be conflicts and confusion, which will slow down the process of increasing your organic search traffic. Be sure to talk to the website program or project manager, not only upper management. In the end, it will be up to the program or project manager(s) to fit the changes into the schedule. Get their buy-in from the beginning. You may want to even involve them in the vendor selection process, because they will have to work with the vendor as well.

If we look first at the process of defining who does what, there are a few

considerations, such as the following:

- What areas of expertise does your inhouse team have? How does the outsourced resource complement or reinforce that?
- Is the outsourced resource bringing a unique new area of expertise to your organization, or is it a supplemental resource to an existing area of expertise?
- Is the outsourced resource going to be working on an area where you consider it essential to build expertise inhouse?
- Can you leverage the outsourced resource with your internal resources? Perhaps you can use this to keep the outside help working on higher-value tasks that provide you a bigger return on your consulting dollar.

All of these considerations factor into the overall division of labor between your outsourced team and your internal team.

You also need to be clear about who is making the decisions and who is communicating those decisions to the outside SEO consultant. He will need guidance and direction on what you are looking to accomplish. Although you hopefully defined this in detail up front, it is likely to change over time, so updating that guidance on a regular basis is important. And as the outside consultant may not sit in your offices all day long, more effort may be required to communicate changes in plans to him than to your internal team.

In addition, the outsourced SEO consultant will have questions and will make recommendations, and he needs to know who to pass these on to and who will give him any related updates and let him know about decisions that are made. A clear communication channel is critical to success.

Building a Culture of SEO into Your Organization

While SEO is at its core a marketing channel, it is unlike any other marketing channel or company function you have seen, because it spans so many disciplines.

SEO involves a deep understanding of and integration with nearly all web-related disciplines, from website development and user experience (UX) design

to site analytics and tracking. Engaging in SEO efforts is ultimately a process of mobilizing all aspects of your organization to come together to determine target audiences, develop messaging, identify content resources and develop robust content development strategies, build out a user-focused yet search-friendly technology platform, and leverage the business model for external engagement and link acquisition.

SEO is also a bit of a strange beast because much of it is nonintuitive—although some of it, for many SEO pros, is very intuitive and based on “gut feel” and two decades’ worth of experience. As an SEO practitioner, you can often explain things to people in great detail, and they will look at you like you have three heads and say, “Huh?”

Building an SEO culture requires that you get the organization as a whole to accept a handful of important facts:

- Organic search engine traffic is essential to the growth and success of the organization.
- Search engines impose some specific constraints on how you can structure your website.
- Search engine success requires specialized marketing programs.
- You can gain a strategic advantage over your competition by embracing these limitations and requirements.

These concepts are pretty high-level and very simple by design. If you can get all parts of the organization to buy into them, though, it will have a very powerful impact on your chances for success. Once everyone agrees that SEO is needed, you will have taken the first step in building an enduring SEO culture into your organization.

The value of this is that you will have made all team members aware, at least at a basic level, of how SEO affects them. Before the senior manager makes a sweeping policy decision about something, she may pause to consider the SEO impact and get the advice she needs on the topic. Before the developer implements something a particular way, he will likely make sure he is doing the right thing from an SEO point of view. In this environment, your inhouse SEO

practitioner is invited to the right meetings, pulled into impromptu discussions, and called for quick answers as people are executing project tasks.

To build an SEO culture at your organization, you need to institutionalize SEO by integrating it into everyday business activities and making it part of everyone's job description. Successful SEO will not be the result of a solo effort by your inhouse SEO practitioner(s); it will come from collaboration with everyone who influences and touches the website.

Conclusion

As we have outlined throughout this chapter, SEO touches engineering, web development, marketing, public relations, UX, design, copywriting and editorial, content strategy, social media, sales, and management. Getting every one of those groups to understand the basics of SEO is essential to long-term success. SEO is not the be-all and end-all, but it is a significant component of all these disciplines.

SEO culture is focused on growing relevant, converting traffic to the website and the business overall, and building a brand that resonates with users. Educating your organization in these simple concepts will ensure that everyone is working toward the same objective and will bring increased efficiency to your overall operations, which will ultimately help you achieve SEO success.

[1] Matt Cutts and Eric Talk About What Makes a Quality Site," Stone Temple Consulting, July 9, 2012, *http://www.stonetemple.com/matt-cutts-and-eric-talk-about-what-makes-a-quality-site/*

CHAPTER FIFTEEN

An Evolving Art Form: The Future of SEO

As we have noted throughout this book, SEO is about leveraging your company's assets through search engine–friendly content creation and website development (on-site) and targeted content promotion (offsite) in order to increase exposure and earn targeted traffic from organic search results. Therefore, the ultimate objective of the SEO professional is to make best use of organic search traffic as determined by various business goals of a website by guiding organizations through SEO strategy development, implementation, and ongoing measurement of SEO efforts.

This role will change as technology evolves, but the fundamental objectives will remain the same as long as "search engines" with nonpaid search results exist. The complexity of search will continue to grow, as all search engines seek to locate and index all of the world's digital information. As a result, we expect various levels of expansion and evolution in search within the following areas:

- Continued rapid expansion of mobile search, and with it, voice search, as the world continues to increase its demand for this capability.
- Large-scale investments in improved understanding of entities and relationships (semantic search) to allow search to directly answer more and more questions.
- Growth in social search as users begin to leverage social networks to discover new and interesting content and solutions to problems from their friends.
- Indexation of multimedia content, such as images, video and audio, including a better understanding of the content of these types of files.
- Indexation of data behind forms (something that Google already does in some cases—for example, if you use First Click Free).
- Continued improvements in extraction and analysis of JavaScript and AJAX-based content.
- Increased localization of search.

- Expanded personalization capabilities.

Mobile search is already driving an increasing demand for linguistic user interfaces, including voice recognition–based search. Voice search greatly improves the ease of use and accessibility of search on mobile devices, and this technology will continue to evolve and improve. In October 2011, Apple released the iPhone 4s with Siri. While most of its capabilities were already present in Google Voice Actions, Siri introduced a more conversational interface, and also showed some personality.

Business deals also regularly change the landscape. For example, on July 29, 2009, Microsoft and Yahoo! signed a far-reaching deal that resulted in Yahoo! retiring the search technology that powers Yahoo! Search and replacing it with Microsoft's Bing technology.[1] Bing also came to an agreement with Baidu to provide the English-language results for the Chinese search engine.[2]

Many contended that this deal would result in a more substantive competitor for Google. With Microsoft's deep pockets as well as a projected combined Bing/Yahoo! U.S. market share of just over 30%, Bing could potentially make a formidable competitor. However, six years later, it is not clear that there has been a significant shift in the search landscape as a result.

Since that time, it's been rumored that Yahoo! has undertaken projects (codenamed Fast-Break and Curveball) aimed at getting it back into the search game, but its lack of adequate search technology makes it seem unlikely any major changes will be made soon. But Yahoo! is not quitting: on November 20, 2014, it was announced that Yahoo! had reached a deal with Mozilla to replace Google as the default search engine for the Firefox web browser.[3]

Perhaps the bigger shift may come from the continuing growth of Facebook, which reports having 1.44 billion monthly active users as of April 2015 (*http://bit.ly/fb_passes_1_44b*), including about half the population of the United States and Canada.[4] Facebook apps boast significant numbers of users as well; among the most popular are WhatsApp (500 million), Instagram (200 million), and Messenger (another 200 million).

As we noted in Chapter 8, Bing's Stefan Weitz suggests that 90% of people use their friends to help them make one or more decisions every day, and 80% of people use their friends to help them make purchasing decisions.[5] In addition, as

we showed in Chapter 8, Google+ has a material effect on Google's personalized search results. If Google can succeed in growing Google+, the scope of this impact could grow significantly over the next few years.

Third Door Media's Danny Sullivan supports the notion that some amount of search traffic may shift to social environments. In a July 2011 interview with Eric Enge, he said: "I think search has his cousin called discovery, which is showing you things that you didn't necessarily know you wanted or needed, but you are happy to have come across. I think social is very strong at providing that."

Neither Stefan Weitz nor Danny Sullivan believes traditional web search is going away, but it seems likely that there will be some shifts as people discover other ways to get the information they want on the Web.

In fact, Stefan Weitz believes that search will get embedded more and more into devices and apps, and is focusing much of Bing's strategy in this direction, partly because the search engine recognizes that it won't win head to head with Google in what Weitz calls the "pure search space."[6]

These developments and many more will impact the role SEO plays within an organization. This chapter will explore some of the ways in which the world of technology, the nature of search, and the role of the SEO practitioner will evolve.

The Ongoing Evolution of Search

Search has come a long way, and will continue to progress at an increasingly rapid pace. Keeping up with these changes, the competitive environment, and the impact of new technology provides both a challenge and an opportunity.

The Growth of Search Complexity

Search has been evolving rapidly over the past decade and a half. At the WSDM conference in February 2009, Google Fellow Jeff Dean provided some interesting metrics that tell part of the story:

- Google search volume had grown 1,000 times since 1999.
- Google has more than 1,000 times the machines it had in 1999.

- Latency dropped from less than 1,000 ms to less than 200 ms.
- Index update latency improved by about 10,000 times. Whereas updates took Google months in 1999, by 2009 Google was detecting and indexing changes on web pages in just a few minutes.

These are staggering changes in Google's performance power, and these are from six years ago. And of course this is just part of the changing search environment. Some of the early commercial search engines, such as Web Crawler, InfoSeek, and Alta Vista, launched in the mid-1990s. At that time, web search engines' relevancy and ranking algorithms were largely based on keyword analysis. This was a simple model to execute and initially provided pretty decent results.

However, there was (and is) too much money in search for such a simple model to stand. Spammers began abusing the weakness of the keyword algorithms by stuffing their pages with keywords, and using tactics to make them invisible to protect the user experience. This led to a situation in which the people who ranked first in search engines were not those who deserved it most, but were in fact those who understood (and could manipulate) the search algorithms the best.

By 1999, Google launched, and the next generation of search was born. Google was the search engine that most effectively implemented the concept of citation analysis (or link analysis). As we outlined earlier in the book, link analysis counted links to a website as a vote for its value. More votes represent more value, with some votes being worth more than others (pages with greater overall link authority have more "juice" to vote).

This created a situation that initially made the spammers' job more difficult, but the spammers began to catch up with this advance by purchasing links. With millions of websites out there, many of them with little or no revenue, it was relatively easy for the spammer to approach a site and offer it a nominal amount of money to get a link. Additionally, spammers could implement bots that surfed the Web, finding guest books, blogs, and forums, and leaving behind comments with links in them back to the bot owner's site.

The major search engines responded to this challenge as well. They took two major steps, one of which was to build teams of people who worked on ways to detect spamming and either discount it (by not attributing value) or penalize it.

The other step was to implement an analysis of link quality that goes deeper than just the notion of PageRank. Factors such as anchor text, relevance, and trust became important as well. These factors also helped the search engines in their war against spam.

But the effort to improve search quality as well as fight spammers continued. Historical search result metrics, such as how many clicks a particular listing got and whether the user was apparently satisfied with the result she clicked on, have made their way into search algorithms. In 2008, then–Yahoo! Chief Scientist Jan O. Pederson wrote a position paper that advocated use of this type of data as follows:

> *Search engine query logs only reflect a small slice of user behavior—actions taken on the search results page. A more complete picture would include the entire click stream; search result page clicks as well as offsite follow-on actions.*
>
> *This sort of data is available from a subset of toolbar users—those that opt into having their click stream tracked. Yahoo! has just begun to collect this sort of data, although competing search engines have collected it for some time.*
>
> *We expect to derive much better indicators of user satisfaction by considering the actions post click. For example, if the user exits the clicked-through page rapidly [then] one can infer that the information need was not satisfied by that page.*

Throughout the last few years, Google, Bing, and Facebook have invested in making use of social signals as well. As we discussed in Chapter 8, links and mentions on social media sites could potentially be used to corroborate to the search engines that your content is worthy of ranking, though this does not appear to be in use today (other than the impact of Google+ on personalized search).

Consider also the notion of author authority. While Google killed the `rel="author"` tag, it can still create associations between a website and an author. If the author begins to establish a high level of authority, this could become a strong signal to the benefit of the site, even if that authority is established elsewhere.

In May 2007, Google made a big splash with the announcement of Universal

Search (delivering what we have referred to in this book as “blended” search results). This was the beginning of the integration of all types of web-based data into a single set of search results, with data from video, images, news, blogs, and shopping search engines all being integrated into a single search experience.

This was only the very beginning of improved indexing for multimedia content. Google already has the ability to recognize the content of images in many cases. As we discussed in Chapter 10, Google Images already allows you to drag an image into the search box and it will try to recognize that image’s content.

Search engines also can make use of other data sources, such as domain registry data to see who owns a particular website. In addition, they have access to analytics data, data from their browsers, data from their ad networks, data from their web search toolbars, and data from free WiFi and Internet access distribution to track actual web usage on various websites. Other data sources will come into play too, such as data from wearables (e.g., Google Glass), embedded devices, and even your thermostat.[7] Although no one knows how, or how much, the search engines use this type of data, these are additional information sources at their disposal.

Search engines continue to look for more ways to improve the quality of search results, and of the overall user experience within search. Google’s efforts toward personalization allow it to look at a user’s search history to get a better idea of what results will best satisfy that user. In 2008, Danny Sullivan summarized this entire evolution into four phases:[8]

- Search 1.0: keywords and text
- Search 2.0: link analysis
- Search 3.0: integration of vertical results
- Search 4.0: personalization

So, what makes up Search 5.0? Increased use of social media data appears to be one major possibility. The “wisdom of the crowds” may be becoming a factor in ranking, as discussed in Chapter 8. Mike Grehan talks about this in his paper, “New Signals to Search Engines”. He summarizes the state of web search as follows:

We're essentially trying to force elephants into browsers that do not want them. The browser that Sir Tim Berners-Lee invented, along with HTML and the HTTP protocol, was intended to render text and graphics on a page delivered to your computer via a dial-up modem, not to watch movies like we do today. Search engine crawlers were developed to capture text from HTML pages and analyze links between pages, but with so much information outside the crawl, is it the right method for an always-on, ever-demanding audience of self producers?

Universal search was a step that acknowledged part of this problem by making all types of data available through web search. But many of these data types do not provide the traditional text-based signals that search engines rely on. Here is more from Mike Grehan's paper:

Signals from end users who previously couldn't vote for content via links from web pages are now able to vote for content with their engagement with that content. You can expect that these types of signals will become a significant factor in the future.

User engagement and interaction with content provides the search engines with information about the value of content beyond links and on an array of data types (such as images and video), and this gives the engines more tools to improve search quality. This type of data was first confirmed as a factor in the rankings of videos on sites such as YouTube, which has become the third largest search engine for the United States (according to comScore, Hitwise, and Nielsen Online, and behind the combined Bing/Yahoo! results), and was confirmed as a factor in general web search in 2011. YouTube's ascent in search volume is particularly interesting because it is not a general web search engine, but one that focuses on a specific vertical—videos. This speaks to demand shifts taking place among the consumers of search results.

At the end of the day, the best results are likely to be provided by the best sites (there are exceptions; for example, for some search queries, the best results may be "instant answers"). The technology that has driven today's engines was based on two ranking signals: good keyword targeting and good links, but user experience and engagement are now significant factors too.

More data collection means more opportunity to win, even if your site doesn't conform flawlessly to these signals, and a better chance that if there is only one

indicator you're winning, you could be in big trouble. Keywords and links will likely remain the primary ranking factors for the next few years, but the trend toward search engines using the new signals is steadily gaining momentum and strength.

Following these advances, what will be next? Artificial intelligence (AI) already plays a huge role in search. As Google's director of research Peter Norvig indicated to Eric Enge in an October 2011 interview:

> *If you define AI as providing a course of action in the face of uncertainty and ambiguity, based on learning from examples, that's what our search algorithm is all about.*

Further evidence of Google's fascination with AI can be seen in Google cofounder Larry Page's March 2014 TED presentation, "Where's Google Going Next?". In that presentation, he discussed how Google was applying machine learning to allow a computer to do things like play video games or, on their own, extrapolate the concept and nature of an entity like a cat. Google's goal is to apply that kind of learning to enable machines to understand the nature of information and how to present it to searchers.

One example of this is how Google Translate works. It examines and analyzes millions of real-world translations to learn how one language translates into another. It learns by example, rather than trying to learn an artificial set of rules. Learning a language based on grammar does not work, because language is far too dynamic and changing. But learning from real-world usage does work. Using this technology, Google can offer instant translation across 58 languages.

Voice search works much the same way. Historically, speech recognition solutions did not work very well and required users to train the system to their voice. Google Voice uses a different approach, as noted by Peter Norvig in that same interview:

> *[For] Voice Search, where you speak your search to Google, we train this model on around 230 billion words from real-world search queries.*

Google's Dominance

Thousands of posts, news articles, and analysis pieces have covered the central topic of battling Google's dominance in web search, but few have discussed the

most telling example of the search giant's superiority. Many believe that the key to Google's success, and more importantly, a key component in its corporate culture, is its willingness and desire to return information relevant to a searcher's query as soon as possible.

Some also believe that Google's biggest impact in the search engine market is its advertising platform, which is the world's largest. By expanding its search capabilities, it is able to create a more enticing advertising platform through AdWords, AdSense, and its embeddable Google search box. Its incredible infrastructure allows for site speed and crawling depth that is unmatched by any other engine.

However, it goes a bit deeper than that. In late 2008, tests were performed in which users were asked which search engine's results they preferred for a wide variety of queries—long-tail searches, top-of-mind searches, topics about which their emotions ranged from great passion to total agnosticism. They were shown two sets of search results and asked which they preferred (see Figure 15-1).

Numerous tests like this have been performed, with a wide variety of results. In some, the brands are removed so that users see only the links. Testers do this to get an idea of whether they can win from a pure "quality" standpoint. In others, the brands remain to get an unvarnished and more "real-world" view. And in one particular experiment—performed many times by many different organizations—the results are swapped across the brands to test whether brand loyalty and brand preference are stronger than qualitative analysis in consumers.

It is this last test that has the most potentially intriguing results. Because in virtually every instance where qualitative differences weren't glaringly obvious, Google was picked as the best "search engine" without regard for the results themselves (see Figure 15-2).

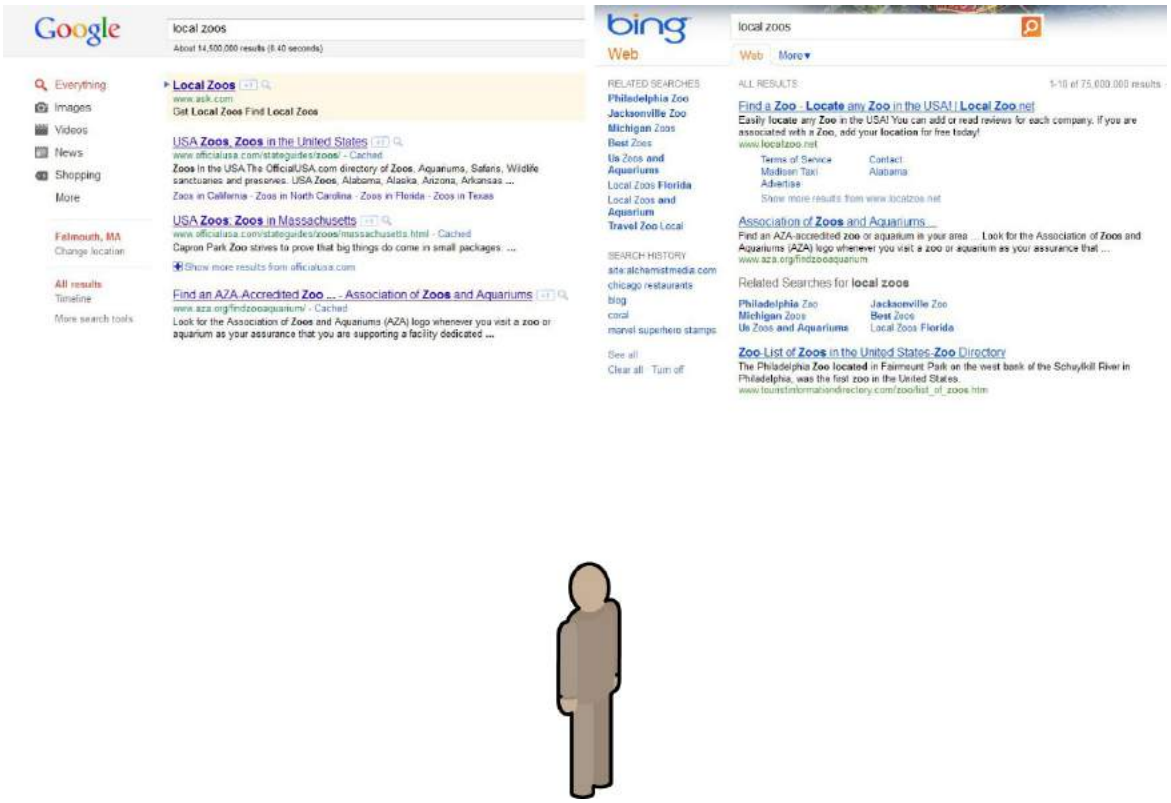


Figure 15-1. Comparing Google and Bing results

Figure 15-2. Results indicating that users may have a strong emotional preference for Google

Fundamentally, testers find (again and again) that the brand preference for Google outweighs the logical consideration of the search results quality.

Search engines that plan to take market share from Google are going to have to think differently. If Microsoft or a start-up search engine wants to take market share, it's going to have to think less like a technology company trying to build a better mousetrap and more like a brand trying to win mindshare from a beloved competitor. How did Pepsi take share away from Coke? Or Toyota from Ford? That is beyond the scope of this book, but it is a process that can take more than a great idea or great technology. It requires a massive psychological shift in the way people around the world perceive the Google brand relative to its competition.

One strategy that Bing has embarked on is to leverage Facebook data in the Bing results; this data is shown in the right rail of the results page as a separate set of results. On October 13, 2010, the two companies signed a deal that provides Bing with substantial access to Facebook's data on shares, likes, and more.[9]

In addition, as noted earlier, Bing's Stefan Weitz talked about Bing's focus on

embedding search technology into devices and apps:

> *The question is, where is search really going? It's unlikely we're going to take share in [the pure search] space, but in machine learning, natural language search...and how we can make search more part of living. For us, it's less about Bing.com, though that's still important. It's really about how we can instead weave the tech into things you're already doing.*

Google's social media network, Google+, isn't the runaway success Google executives hoped it would be. Google purportedly continues to invest heavily in Google+[10], yet in February 2014, the *New York Times* likened Google+ to a ghost town, citing Google stats of 540 million "monthly active users," but noting that almost half don't visit the site. A significant redirection with Google+ is likely inevitable

Also consider the official Google mission statement: "Google's mission is to organize the world's information and make it universally accessible and useful." It is already moving beyond that mission. For example, Google and NASA are working on new networking protocols that can work with the long latency times and low bandwidth in space.

Google is also pursuing alternative energy initiatives, which clearly goes beyond its mission statement. One example is its investments in self-driving cars. In addition, Google has ventures in office productivity software with Google Docs. These two initiatives have little to do with SEO, but they do speak to how Google is trying to expand its reach.

Another recent, now cancelled, venture for Google was its exploration into wearable technology, primarily in the form of Google Glass. Though Google Glass fell out of favor with much of the technology community, Glass was likely intended only as a first foray into the realm of wearable tech.[11] Despite the fact that Glass itself may have been released too soon, and in too much of a "beta" format, it dominated the early wearable search technology market. As wearable technology becomes more pervasive, expect Google to reintroduce another iteration of Google Glass in a bid to take a leading role in that market just as it has in mobile technology with Android.

Another potential future involves Google becoming a more general-purpose pattern-matching and searching engine. The concept of performing pattern matching on text (e.g., the current Google on the current Web) is only the first

stage of an evolving process. Imagine the impact if Google turns its attention to the human genome and creates a pattern-matching engine that revolutionizes the way in which new medicines are developed. And the plethora of potential applications of pattern matching in the “real world” (such as disaster preparedness and logistics) is astounding; Google is uniquely poised to capitalize on this global opportunity with its understanding of our physical world from the ground (e.g., by Google’s self-driving cars), from in our homes (e.g., Google’s Nest acquisition), and from the sky via satellites and drones (e.g., Google’s acquisitions of Skybox Imaging and Titan Aerospace).

More Searchable Content and Content Types

The emphasis throughout this book has been on providing the crawlers with textual content semantically marked up using HTML. However, the less accessible document types—such as multimedia, content behind forms, and scanned historical documents—are increasingly being integrated into the search engine results pages (SERPs) as search algorithms evolve in the ways that the data is collected, parsed, and interpreted. Greater demand, availability, and usage also fuel the trend.

Engines Will Make Crawling Improvements

The search engines are breaking down some of the traditional limitations on crawling. Content types that search engines could not previously crawl or interpret are being addressed.

In May 2014, Google announced that it had substantially improved the crawling and indexing of CSS and JavaScript content. Google can now render a large number of web pages with the JavaScript turned on so that its crawlers see it much more like the average user would. In October 2014, Google updated its Webmaster Guidelines to specifically advise that you do not block crawling of JavaScript and CSS files.

Despite these improvements, there are still many challenges to Google fully understanding all of the content within JavaScript or CSS, particularly if the crawlers are blocked from your JavaScript or CSS files, if your code is too complex for Google to understand, or if the code actually removes content from

the page rather than adding it. Google still recommends that you build your site to "degrade gracefully", which essentially means to build the site such that all of your content is available whether users have JavaScript turned on or off.

Another major historical limitation of search engines is dealing with forms. The classic example is a search query box on a publisher's website. There is little point in the search engine punching in random search queries to see what results the site returns. However, there are other cases in which a much simpler form is in use, such as one that a user may fill out to get access to a downloadable article.

Search engines could potentially try to fill out such forms, perhaps according to a protocol where the rules are predefined, to gain access to this content in a form where they can index it and include it in their search results. A lot of valuable content is currently isolated behind such simple forms, and defining such a protocol is certainly within the realm of possibility (though it is no easy task, to be sure). This is an area addressed by Google in a November 2011 announcement. In more and more scenarios you can expect Google to fill out forms to see the content that exists behind them.

Engines Are Getting New Content Sources

As we noted earlier, Google's stated mission is "to organize the world's information and make it universally accessible and useful." This is a powerful statement, particularly in light of the fact that so much information has not yet made its way online.

As part of its efforts to move more data to the Web, in 2004 Google launched an initiative to scan in books so that they could be incorporated into a Book Search search engine. This became the subject of a lawsuit by authors and libraries, but a settlement was reached in late 2008. In addition to books, other historical documents are worth scanning. To aid in that Google acquired reCAPTCHA (e.g., see *http://www.google.com/recaptcha*), and in December 2014, Google announced a major enhancement to how reCAPTCHA works, with the goal of making it much more user-friendly.

Similarly, content owners retain various other forms of proprietary information that is not generally available to the public. Some of this information is locked up behind logins for subscription-based content. To provide such content owners

an incentive to make that content searchable, Google came up with its First Click Free program (discussed earlier in this book), which allows Google to crawl subscription-based content.

Another example of new sources is metadata, in the form of markup such as Schema.org, microformats, and RDFa. This type of data, which is discussed in "CSS and Semantic Markup", is a way for search engines to collect data directly from the publisher of a website. Schema.org was launched as a joint initiative of Google, Bing, and Yahoo! to collect publisher-supplied data, and the number of formats supported can be expected to grow over time.[12]

Another approach to this would be allow media sites and bloggers to submit content to the search engines via RSS feeds. This could potentially speed indexing time and reduce crawl burden at the same time. One reason why search engines may not be too quick to do this, though, is that website publishers are prone to making mistakes, and having procedures in place to protect against those mistakes might obviate the benefits.

However, a lot of other content out there is not on the Web at all, and this is information that the search engines want to index. To access it, they can approach the content owners and work on proprietary content deals, and this is also an activity that the search engines all pursue.

Another direction they can go with this is to find more ways to collect information directly from the public. Google Image Labeler was a program designed to do just this. It allowed users to label images through a game where they would work in pairs and try to come up with the same tags for the image as the person they were paired with. Unfortunately, this particular program was discontinued, but other approaches like it may be attempted in the future.

Multimedia Is Becoming Indexable

Content in images, audio, and video is currently not easily indexed by the search engines, but its metadata (tags, captioning, descriptions, geotagging data) and the anchor text of inbound links and surrounding content make it visible in search results. Google has made some great strides in this area. In an interview with Eric Enge, Google's director of research Peter Norvig discussed how Google allows searchers to drag an image from their desktop into the Google Images search box and Google can recognize the content of the image.

Or consider *http://www.google.com/recaptcha*. This site was originally used by Google to complete the digitization of books from the Internet Archive and old editions of the *New York Times*. These have been partially digitized using scanning and OCR software. OCR is not a perfect technology, and there are many cases where the software cannot determine a word with 100% confidence. However, reCAPTCHA is assisting by using humans to figure out what these words are and feeding them back into the database of digitized documents.

First, reCAPTCHA takes the unresolved words and puts them into a database. These words are then fed to blogs that use the site's CAPTCHA solution for security purposes. These are the boxes you see on blogs and account sign-up screens where you need to enter the characters you see, such as the one shown in Figure 15-3.

Figure 15-3. ReCAPTCHA screen

In this example, the user is expected to type in *morning*. However, in this case, Recaptcha.net is using the human input in these CAPTCHA screens to translate text that wasn't recognized by OCR. It makes use of this CAPTCHA information to improve the quality of this digitized book.

Google used this methodology for years, but has since changed reCAPTCHA to focus more on images instead.[13] This new approach is image based, and is intended to help Google with its computer vision projects.

Similarly, speech-to-text solutions can be applied to audio and video files to extract more data from them. This is a relatively processing-intensive technology, and it has historically had trouble with many accents and dialects, so it has not yet been universally applied in search. Apple's Siri and Google Voice are leading the charge in addressing this issue. In addition, if you upload a video to YouTube, you can provide a search-indexable caption file for it, or request that Google use its voice recognition technology to attempt to autocaption it.

The business problem the search engines face is that the demand for information and content in these challenging-to-index formats is increasing exponentially. Search results that do not accurately include this type of data will begin to be deemed irrelevant or wrong, resulting in lost market share and declining ad revenues.

The dominance of YouTube is a powerful signpost of user interest. Users want engaging, multimedia content—and they want a lot of it. For this reason, developing improved techniques for indexing such alternative content types is an urgent priority for the search engines.

Interactive content is also growing on the Web, with technologies such as AJAX. In spite of the indexing challenges these technologies bring to search engines, their use is continuing because of the experience they offer broadband users. The search engines are hard at work on solutions to better understand the content wrapped up in these technologies as well.

Over time, our view of what is "interactive" will likely change dramatically. Two-or three-dimensional first-person shooter games and movies will continue to morph and become increasingly interactive. Further in the future, these may become full immersion experiences, similar to the holodeck on *Star Trek*. You can also expect to see interactive movies where the audience influences the plot with both virtual and human actors performing live. These types of advances are not the immediate concern of today's SEO practitioner, but staying in touch with where things are headed over time can provide a valuable perspective.

More Personalized, Localized, and User-Influenced Search

Personalization efforts have been under way at the search engines for some time. As we discussed earlier in this book, the most basic form of personalization is to perform an IP location lookup to determine where the searcher is located, and tweak the results based on the searcher's location. However, the search engines continue to explore additional ways to expand on this simple concept to deliver better results for each user. It is not yet clear whether personalization has given the engines that have heavily invested in it better results overall or greater user satisfaction, but their continued use of the technology suggests that, at the least,

their internal user satisfaction tests have been positive.

Indeed, Google has continued to expand the factors that can influence a users' personalized search. For example, one major signal it uses is the user's personal search history. Google can track sites a user has visited, groups of related sites the user has visited, whether a user has shared a given site over social media, and what keywords the user has searched for in the past. All of these factors may influence the given personalized search engine results page.

User Intent

As just mentioned, Google personalized results are tapping into user intent based on previous search history, and serving up a mix not just of personalized "blue links" but of many content types, including maps, blog posts, videos, and local results. The major search engines already provide maps for appropriate location searches and the ability to list blog results based on recency as well as relevancy. It is not just about presenting the results, but about presenting them in the format that maps to the searcher's intent.

User Interactions

One area that will see great exploration is how users interact with search engines. As the sheer amount of information in its many formats expands, users will continue to look to search engines to be not just a search destination, but also a source of information aggregation whereby the search engine acts as a portal, pulling and updating news and other content based on the user's preferences.

Marissa Mayer, then Google's VP of Location and Local Services (now CEO of Yahoo!), made a particularly interesting comment that furthers the sense that search engines will continue their evolution beyond search:

> *I think that people will be annotating search results pages and web pages a lot. They're going to be rating them, they're going to be reviewing them. They're going to be marking them up...*

Indeed, Google already offers users the ability to block certain results. Mayer's mention of "web pages" may be another reason why the release of Google Chrome was so important. Tapping into the web browser might lead to that

ability to annotate and rate those pages and further help Google identify what content interests the user. As of February 2014, StatCounter showed that Chrome's market share had risen to an impressive 44%.

Chris Sherman, executive editor of Search Engine Land, offered up an interesting approach that the search engines might pursue as a way to allow users to interact with them and help bring about better results:

> *[F]ind a way to let us search by example—submitting a page of content and analyzing the full text of that page and then tying that in conjunction with our past behavior...*

New Search Patterns

This is all part of increasing the focus on the users, tying into their intent and interests at the time of search. Personalization will make site stickiness ever more important. Securing a position in users' history and becoming an authoritative go-to source for information will be more critical than ever. Winning in the SERPs will require much more than just optimizing for position, moving toward an increased focus on engagement.

Over time, smart marketers will recognize that the attention of a potential customer is a scarce and limited quantity. As the quantity of information available to us grows, the amount of time we have available for each piece of information declines, creating an attention deficit. How people search, and how advertisers interact with them, may change dramatically as a result.

In 2008, *The Atlantic* published an article titled "Is Google Making Us Stupid?". The thrust of this article was that Google was so powerful in its capabilities that humans need to do less (and less!). Google has made huge advances since this article, and this is a trend that will continue. After all, who needs memory when you have your "lifestream" captured 24/7 with instant retrieval via something akin to Google desktop search or when you have instant perfect recall of all of human history?

These types of changes, if and when they occur, could transform what today we call SEO into something else, where the SEO of tomorrow is responsible for helping publishers gain access to potential customers through a vast array of new mechanisms that currently do not exist.

Growing Reliance on the Cloud

Cloud computing is transforming how the Internet-connected population uses computers. Oracle founder Larry Ellison's vision of thin-client computing may yet come to pass, but in the form of a pervasive Google operating system and its associated, extensive suite of applications. Widespread adoption by users of cloud-based (rather than desktop) software and seemingly limitless data storage, all supplied for free by Google, will usher in a new era of personalized advertising within these apps.

Google is actively advancing the mass migration of desktop computing to the cloud, with initiatives such as Google Docs, Gmail, Google Calendar, Google App Engine, and Google Drive. These types of services encourage users to entrust their valuable data to the Google cloud. This brings them many benefits (but also concerns around privacy, security, uptime, and data integrity). In May 2011 Apple also made a move in this direction when it announced iCloud, which is seamlessly integrated into Apple devices.

One simple example of a basic application for cloud computing in the notion of backing up all your data. Most users don't do a good job of backing up their data, making them susceptible to data loss from hard drive crashes and virus infections. Companies investing in cloud computing will seek to get you to store the master copy of your data in the cloud, and keep backup copies locally on your devices (or not at all). With this approach you can more easily access that information from multiple computers (e.g., at work and home).

Google (and Apple) benefits by having a repository of user data available for analysis—which is very helpful in Google's quest to deliver ever more relevant ads and search results. It also provides multiple additional platforms within which to serve advertising. Furthermore, regular users of a service such as Google Docs are more likely to be logged in a greater percentage of the time when they are on their computers.

The inevitable advance of cloud computing will offer more and more services with unrivaled convenience and cost benefits, compelling users to turn to the cloud for their data and their apps.

Increasing Importance of Local, Mobile, and

Voice Search

New forms of vertical search are becoming increasingly important. Search engines have already embraced local search and mobile search, and voice-based search is an area in which all the major engines are actively investing.

Local Search

Local search was an active business before the advent of the Internet, when Yellow Pages from your phone company were the tool of choice. As the World Wide Web gained prominence in our lives, the Yellow Pages migrated online. These types of local search sites have themselves evolved to leverage more of the unique nature of the Web. Some of the major players are CitySearch, Local.com, YellowPages.com, and Superpages.

As the major search engines evolved, they integrated sophisticated mapping technology into their systems, and then began to map in local business data as well. You can find these local search engines at *http://maps.google.com/*, *http://local.yahoo.com*, and *http://www.bing.com/maps*. These search engines map in some of their business data from the same types of data sources that drive the Yellow Pages websites, such as Acxiom, InfoUSA, and Localeze, but they also supplement that data by crawling the Web.

One of the big challenges facing the local search engines is *map spam*, the proliferation of spam results in local search. Here are two blog posts that provide examples:

- "Illusory Laptop Repair—A Most Elegant Google Places Hack"
- "Google not Good Enough in Controlling Local Map Spam"

Resolving these issues is obviously critical for the search engines and is a major area of investment for them. We believe that local search is an area whose importance will continue to grow. With a very large number of searches having local intent, this is potentially a major frontier for developments in search.

Over time, you can expect that these problems will be worked out. Local search is already extremely powerful, but you can anticipate that the search engines will control the flow of customers in cyberspace and the real world, because

customers will be finding merchants through search. You can also expect that more and more of these local searches will take place on mobile devices. Success in local search will be what makes or breaks many businesses.

Consumers are becoming increasingly reliant on local search and its auxiliary services—street maps, directions, satellite views, street views, 3D visualizations (Google Earth), enhanced listings, user reviews, and ratings. Through Google Mobile's image recognition (formerly Google Goggles), Google even has the capability of searching based on photos you take with your mobile device. Augmented reality, where we see metadata juxtaposed on top of the physical world via a "looking glass," is around the corner, and Google Glass is just a hint of what is to come.

With the help of social apps such as Yelp, FourSquare, and Facebook, mobile users can locate nearby friends, and search for special offers and discounts when they "check in" at certain stores or restaurants. Such tools guide consumers' movements and their decision-making processes; this means the search engines and other search services are in a powerful position to manage the flow of customers from the Web to brick-and-mortar businesses. So, in a way, the search engines are in the logistics business, building the connection between online and offline.

Mobile Search

According to the June 2015 Ericsson Mobility Report, there were 7.1 billion worldwide mobile subscribers in 2014, with a forecast of 9.2 billion subscribers by 2020. (*http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf*).[14] As a result, the opportunity for mobile search has begun to grow rapidly. A Google study showed that search engines were the most visited sites on mobile devices, with 77% of those surveyed reporting their use.[15] The same study showed that 9 out of 10 searchers took action as a result of conducting a search.

In addition, as of October 2014, Google began experimenting with using the mobile friendliness of a site as a ranking factor, and in May 2015 Google made Mobile Friendliness a significant ranking factor impacting results for users searching from Smartphones.[16] This provides strong incentive to have a mobile search strategy in place (see "The Mobile Landscape" for more on this)!

The convenience of being able to get the information you need while on the go is just way too compelling. Why be tied to a desk if you do not have to be? Further, in many countries, freedom from a desktop is a necessity because users have much less personal space than their U.S. counterparts.

The ever-expanding versatility and power of tablet and mobile devices—from indispensable utility apps, to immersive multimedia players, to massively multiplayer online games, to paradigm-shifting hardware advances such as the iPhone's multitouch display, proximity sensor, GPS, and gyroscopes—will fuel this growth. The *network effect*, whereby the value of the network grows by the square of the size of the network (Metcalf's Law), gives further incentive for users to migrate to their mobile devices as more and more apps allow them to interact with their peers in increasingly more interesting ways.

The small keyboard/typing surface is currently a limitation, but Apple's and Google's voice-based solutions have already made great strides in replacing the keyboard as the input device of choice. The advent of the Linguistic User Interface (LUI) will continue to revolutionize mobile search, and with the LUI comes a whole new set of skills that will be required of the SEO practitioner.

In addition, wearable devices such as the Apple Watch are gaining traction. Voice technology is improving to the point where the absence of a keyboard will not be a significant limitation.

Voice Recognition Search

When users are mobile they must deal with the limitations of their mobile device, specifically the small screen and small keyboard. These make web surfing and mobile searching more challenging than they are in the PC environment.

Voice search is a great way to improve the mobile search experience. It eliminates the need for the keyboard, and provides users with a simple and elegant interface. Speech recognition technology has been around for a long time, and the main challenge has always been that it requires a lot of computing power.

Several examples already exist:

Google Voice Search

This is a free service from Google that enables you to perform a Web search based on text-to-speech voice queries instead of traditional text-typed queries.

Google Mobile App for iPhone

Included in this application is voice searching capability. You can speak your query into the application and the results are displayed on-screen.

Apple's Siri

What made Siri a significant step forward after its launch was its use of more natural human speech (Google offers similar capabilities now). This was the start of a movement toward the device learning the user, rather than vice versa.

Microsoft Cortana

In April 2014, Micorosft unveiled its personal assistant offering, known as Cortana. Cortana is being promoted as offering more capabilities for learning your personal preferences than the competing options. Currently, it is designed to work only on Windows Phones.

Processing power continues to increase, even on mobile devices, and the feasibility of this type of technology is growing. This should be another major area of change in the mobile search landscape.

In addition, it is reasonable to expect that voice recognition technology will be applied to the actual recorded content of audio and video files, to determine content and aid in ranking these media types within search results—something that will likely lead to video "script" optimization as an added component of video SEO.

Increased Market Saturation and Competition

One thing you can count on with the Web is continued growth. Despite its constantly growing index, a lot of the pages in Google may be low-quality or duplicate-content-type pages that will never see the light of day. The Web is a

big place, but one where the signal-to-noise ratio is very low.

One major trend emerges from an analysis of Internet usage statistics. According to Miniwatts Marketing Group, 84.9% of the North American population uses the Internet, so there is not much room for growth there. In contrast, Asia, which already has the most Internet users (1.4 billion) has a penetration rate of only 34.7%. Other regions with a great deal of opportunity to grow are Africa, the Middle East, and Latin America.

This data tells us that in terms of the number of users, North America is already approaching saturation. Europe has some room to grow, but not that much. However, in Asia, you could see two times that growth, or 2 to 3 billion users! The bottom line is that a lot of Internet growth in the coming decade will be outside North America, and that will provide unique new business opportunities for those who are ready to capitalize on that growth.

With this growth has come an increasing awareness of what needs to be done to obtain traffic. The search engines are the dominant suppliers of traffic for many publishers, and will continue to be for some time to come. For that reason, awareness of SEO will continue to increase over time. Here are some reasons why this growth has continued:

The Web outperforms other sales channels

When organizations look at the paths leading to sales and income (a critical analysis whenever budgets are under scrutiny), the Web almost always comes out with one of two assessments. Either it is a leading sales channel (especially from an ROI perspective), or it is the area with the greatest opportunity for growth. In both scenarios, digital marketing (and, in correlation, SEO) take center stage.

It is the right time to retool

Established companies frequently use down cycles as a chance to focus attention inward and analyze themselves. Consequently, there’s a spike in website redesigns and SEO along with it.

Paid search drives interest in SEO

Paid search spending is still reaching all-time highs, and when companies

evaluate the cost and value, there's a nagging little voice saying, "75%+ of the clicks do not even happen in the ads; use SEO."

SEO is losing its stigma

Google is releasing SEO guides, Microsoft and Yahoo! have in-house SEO departments, and the "SEO is BS" crowd have lost a little of their swagger and a lot of their arguments. No surprise—solid evidence trumps wishful thinking, especially when times are tough.

Marketing departments are in a brainstorming cycle

A high percentage of companies are asking the big questions: "How do we get new customers?" and "What avenues still offer opportunity?" Whenever that happens, SEO is bound to show up near the top of the "to be investigated" pile.

Search traffic will be relatively unscathed by the market

Sales might drop and conversion rates might falter a bit, but raw search traffic isn't going anywhere. A recession doesn't mean people stop searching the Web, and with broadband adoption rates, Internet penetration, and searches per user consistently rising, search is no fad. It is here for the long haul.

Web budgets are being reassessed

We've all seen the news about display advertising falling considerably; that can happen only when managers meet to discuss how to address budget concerns. Get 10 Internet marketing managers into rooms with their teams and at least 4 or 5 are bound to discuss SEO and how they can grab that "free" traffic.

Someone finally looked at the web analytics

It is sad, but true. When a downturn arrives or panic sets in, someone, maybe the first someone in a long time, checks the web analytics to see where revenue is still coming in. Not surprisingly, search engine referrals with their

exceptional targeting and intent matching are ranking high on the list.

Although more and more people are becoming aware of these advantages of SEO, there still remains an imbalance between paid search and SEO. The SEMPO Annual State of Search Survey (*http://bit.ly/2015_state_of_search*; membership is required to access the report) includes information suggesting that as much as 90% of the money invested in search-related marketing is spent on PPC campaigns and only about 10% goes into SEO.

This suggests that either SEO could see some growth to align budgets with potential opportunity, or firms that focus solely on SEO services had better diversify. SEO budgets continue to expand as more and more businesses better understand the mechanics of the Web. In the short term, PPC is easier for many businesses to understand, because it has more in common with traditional forms of marketing. Ultimately, though, SEO is where the most money can be found, and the dollars will follow once people understand that.

SEO as an Enduring Art Form

Today, SEO can be fairly easily categorized as having five major objectives:

- Make content accessible to search engine crawlers.
- Find the keywords that searchers employ (understand your target audience) and make your site speak their language.
- Build content that users will find useful, valuable, and worthy of sharing. Ensure that they'll have a good experience on your site to improve the likelihood that you'll earn links and references.
- Earn votes for your content in the form of editorial links and social media mentions from good sources by building inviting, shareable content and applying classic marketing techniques to the online world.
- Create web pages that allow users to find what they want extremely quickly, ideally in the blink of an eye.

Note, though, that the tactics an SEO practitioner might use to get links from

editorial sources have been subject to rapid evolution. We now turn to content marketing instead of link building. In addition, a strong understanding of how the search engines measure and weight social engagement signals is increasingly important to SEO professionals.

One thing that you can be sure about in the world of search is change, as forces from all over the Web are impacting search in a dramatic way. To be an artist, the SEO practitioner needs to see the landscape of possibilities for an individual website, and pick the best possible path to success. This currently includes social media optimization expertise, local search expertise, video optimization expertise, an understanding of what is coming in mobile, and more. That's a far cry from the backroom geek of the late 1990s.

No one can predict what the future will bring and what will be needed to successfully market businesses and other organizations on the Web in 2 years, let alone 5 or 10. However, you can be certain that websites are here to stay for a long time, and that websites are never finished and need continuous optimization just like any other direct marketing channel. SEO expertise will be needed for a long time—and no one is better suited to map the changing environment and lead companies to success in this new, ever-evolving landscape than today's SEO practitioner.

The Future of Semantic Search and the Knowledge Graph

In Chapter 6, we explored the state of semantic search and the Knowledge Graph as we know it today. All the search engines are continuing to investigate these types of technologies in many different ways, though Google is clearly in the lead. The Knowledge Vault is just one of many initiatives that Google is pursuing to make progress in this area.

Part of the objective is to develop a machine intelligence that can fully understand how people evaluate the world, yet even this is not sufficient. The real goal is to understand how *each* human being evaluates the world, so that the results can be fully personalized to meet each individual's needs.

Not only do search engines want to give you the perfect answer to your questions, they also want to provide you with opportunities for exploration.

Humans like to conduct research and learn new things. Providing all of these capabilities will require a special type of machine intelligence, and we are a long way from reaching those goals.

There are many components that go into developing this type of intelligence. In the near future, efforts focus largely on solving specific problems. For example, one such problem is maintaining the context of an ongoing conversation. Consider the following set of queries, starting with *where is the empire state building?* in Figure 15-4.

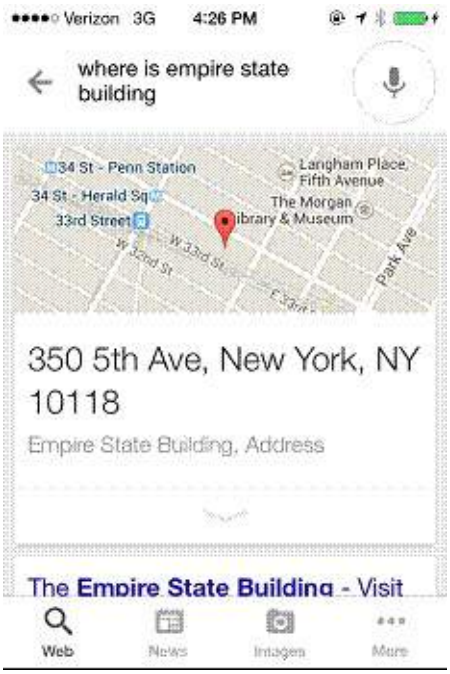


Figure 15-4. Response to the query "where is the empire state building?"

Notice how the word *the* was dropped in the query display. Figure 15-5 shows what happens when you follow this query with the one on *pictures*.

Figure 15-5. Response to the query "pictures"

Notice again how the query was modified to *empire state building pictures*. Google has remembered that the prior query was specific to the Empire State Building, and did not require us to restate that. This query sequence can continue for quite some time. Figure 15-6 shows the result when we now ask *who built it?*

Figure 15-6. Response to the query "who built it?"

Once again, the query was dynamically modified on the fly, and Google has remembered the context of the conversation. Figure 15-7 shows what happens when we now try the query *restaurants*.

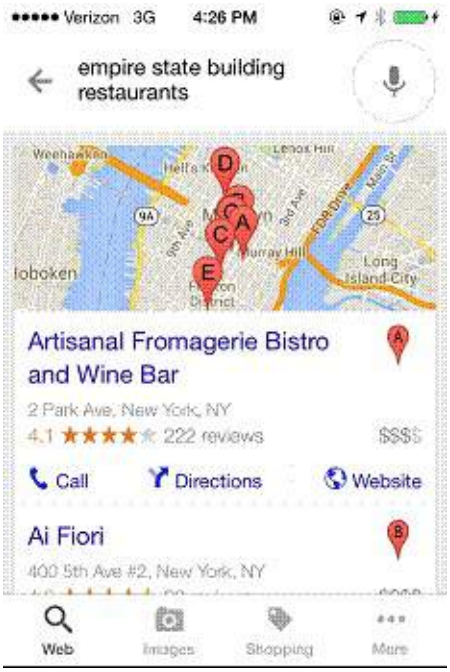


Figure 15-7. Response to the query "restaurants"

Finally, we can follow this query with the more complex query *give me directions to the third one*, as shown in Figure 15-8.

Figure 15-8. Response to the query "give me directions to the third one"

This entire sequence of queries is quite complicated, capped off by Google's understanding of the concept of the "third one" in the final query. Even though this is very sophisticated, it is nonetheless an example of a point solution to a specific problem.

To truly model human thought, search engines will need to build machines that can reason like humans, are able to perceive the world around them, understand how to define objectives and make plans to meet them, and can independently work to expand their knowledge.

Many disciplines are involved in developing artificial intelligence, such as computer science, neuroscience, psychology, philosophy, and linguistics. Even just understanding linguistics is a major challenge, as there are thousands of different languages in the world, and this by itself multiplies the complexity of the task.

The computing power to take on these challenges does not yet exist, so developing expanding capabilities is a major piece of the puzzle. For example, Google is pursuing efforts to build a quantum computer.[17]

In the near term, we can expect changes in search results to come in the form of

more point solutions to specific problems. As the understanding of how to model human intelligence expands, and as processing power grows with it, we may see much more significant changes, perhaps 5 to 10 years down the road.

Conclusion

SEO is both art and science. The artistic aspect of SEO requires dynamic creativity and intuition; the search engine algorithms are too complex to reverse-engineer every aspect of them. The scientific aspect involves challenging assumptions, analyzing data, testing hypotheses, making observations, drawing conclusions, and achieving reproducible results. These two ways of thinking will remain a requirement as SEO evolves into the future.

In this chapter, we conveyed some sense of what is coming in the world of technology, and in particular, search. Although the previous decade has seen an enormous amount of change, the reality is that it has simply been the tip of the iceberg. There's a lot more change to come, and at an ever-increasing (exponential) rate. If the Law of Accelerating Returns holds, we're in for a wild ride.

In this fast-moving industry, the successful SEO professional has to play the role of early adopter. The early adopter is always trying new things—tools, tactics, approaches, processes, technologies—to keep pace with the ever-evolving search engines, ever-increasing content types, and the ongoing evolution of online user engagement.

It is not enough to adapt to change. You will need to embrace it and evangelize it. Many in your (or your client's) organization may fear change, and steering them through these turbulent waters will require strong leadership. Thus, the successful SEO professional also has to play the role of change agent.

The need for organizations to capture search mindshare, find new customers, and promote their messaging will not diminish anytime soon, and neither will the need for searchable, web-based, and instantaneous access to information, products, and services. This ability—to generate traction and facilitate growth by connecting the seeker and provider—is perhaps the most valuable skill set on the Web today. And although there is the possibility that the search engines could eventually be called "decision," "dilemma," or even "desire" engines, the

absolute need for the understanding of and interactions between both the psychological and the technological natures of search will ensure that SEO as a discipline, and SEO professionals, are here to stay.

[1] Steve Lohr, "Microsoft and Yahoo Are Linked Up. Now What?", New York Times, July 29, 2009, *http://bit.ly/ms_yahoo_linked_up*

[2] Michael Bonfils, "Bada Bing! It's Baidu Bing – English Search Marketing in China," Search Engine Watch, August 3, 2011, *http://bit.ly/baidu_bing*

[3] Alexei Oreskovic, "Yahoo Usurps Google in Firefox Search Deal," Reuters, November 20, 2014, *http://bit.ly/yahoo_usurps_google*

[4] Jim Edwards, "'Facebook Inc.' Actually Has 2.2 Billion Users Now — Roughly One Third Of The Entire Population Of Earth," Business Insider, July 24, 2014, *http://www.businessinsider.com/facebook-inc-has-22-billion-users-2014-7*

[5] Emil Protalinski, "Bing Adds More Facebook Features to Social Search," ZDNet, May 16, 2011, *http://www.zdnet.com/blog/facebook/bing-adds-more-facebook-features-to-social-search/1483*

[6] Brid-Aine Parnell, "Microsoft's Bing Hopes to Bag Market Share with ... Search Apps," The Register, November 4, 2014, *http://bit.ly/bing_search_apps*

[7] Marcus Wohlsen, "What Google Really Gets Out of Buying Nest for $3.2 Billion," Wired, January 14, 2014, *http://www.wired.com/2014/01/googles-3-billion-nest-buy-finally-make-internet-things-real-us/*

[8] Danny Sullivan, "Search 4.0: Social Search Engines & Putting Humans Back in Search," Search Engine Land, May 28, 2008, *http://searchengineland.com/search-40-putting-humans-back-in-search-14086*.

[9] Bing Blogs, "Bing Gets More Social with Facebook," October 13, 2010, *http://blogs.bing.com/search/2010/10/13/bing-gets-more-social-with-facebook/*

[10] Kurt Wagner, "New Google+ Head David Besbris: We're Here for the Long Haul (Q&A)," Re/code, October 7, 2014, *http://bit.ly/david_besbris*

[11] Rachel Metz, "Google Glass Is Dead; Long Live Smart Glasses," MIT Technology Review, November 26, 2014, *http://bit.ly/google_glass_dead*

[12] Google Official Blog, "Introducing Schema.org: Search Engines Come Together for a Richer Web," June 2, 2011, *http://bit.ly/intro_schema_org*.

[13] Frederic Lardinois, “Google’s reCAPTCHA (Mostly) Does Away With Those Annoying CAPTCHAs,” TechCrunch, December 3, 2014, *http://bit.ly/googles_recaptcha*.

[14] More People Around the World Have Cell Phones Than Ever Had Land-Lines,” Quartz, February 25, 2014, *http://qz.com/179897/more-people-around-the-world-have-cell-phones-than-ever-had-land-lines/*

[15] Julie Batten, “Newest Stats on Mobile Search,” ClickZ, May 23, 2011, *http://bit.ly/newest_mobile_stats*

[16] Barry Schwartz, “Google May Add Mobile User Experience To Its Ranking Algorithm,” Search Engine Land, October 8, 2014, *http://searchengineland.com/google-may-add-mobile-user-experience-ranking-algorithm-205382*

[17] Tom Simonite, “Google Launches Effort to Build Its Own Quantum Computer,” MIT Technology Review, September 3, 2014, *http://www.technologyreview.com/news/530516/google-launches-effort-to-build-its-own-quantum-computer/*.

Glossary

301 status code

An HTTP status code returned by a web server indicating that content has been moved from one location to another. A 301 status code indicates a permanent server redirect, which is a signal that the original page no longer exists and that users and search engines should view the new page as the canonical version of that content.

302 status code

The 302 status code means that the document requested is “found” but temporarily resides under a different URL. Because a permanent redirect has not been used, the client should continue to use the original requested URL for future requests.

400 status code

The 400 status code means a “bad request,” indicating that the server is not able to understand the document request due to malformed syntax.

401 status code

The 401 status code means “unauthorized.” The server is requesting user authentication prior to fulfilling the document request.

403 status code

The 403 status code means “forbidden.” The server understood the request, but is refusing to fulfill it. The webmaster may wish to alert the user as to why her request has been denied. If the organization does not wish to provide a reason, a “404 (Not Found)” status can be displayed instead.

404 status code

The 404 error message represents a document “not found.” This means that the client was able to communicate with the server, but the server could not find the requested document. Alternatively, the server could be configured to not fulfill the request and not provide a reason why.

410 status code

Similar to a “404 (Not Found)” error message, the 410 status code indicates that the requested document is intentionally “gone” (i.e., is no longer available), and there is no forwarding address.

500 status code

The 410 status code is usually used for limited-display documents such as promotional information. It is up to the webmaster to determine at what point to remove the 410 status message.

500 status code

The 500 error message states that there was an “internal server error” that has prevented the document request from being fulfilled.

501 status code

The 501 error message is displayed when the server does not recognize the document request method. The server is not capable of fulfilling this request and states that the request was “not implemented.”

Absolute link

A link that specifies the exact location on the Web of a particular page. An example of this would be *http://www.netconcepts.com/learning-center*, which specifies all aspects of where the content is to be found.

This is in contrast to a relative link, which specifies the location of a web page relative to your current location. For example, if you are on *http://www.stonetemple.com* and you specify a link as */about-eric-enge*, that

will be interpreted as *http://www.stonetemple.com/about-eric-enge*.

Accessibility

The practice of making websites usable by disabled people (especially vision-impaired users).

adCenter

Microsoft's cost-per-click (CPC) advertising program. These are ads that you see in Bing's search results. Bing uses a variety of factors to determine the order in which the paid search results are shown, including historical click-through rate and the bid price.

AdSense

According to Google, "AdSense is a fast and easy way for website publishers of all sizes to display relevant Google ads on their website's content pages and earn money. Because the ads are related to what your visitors are looking for on your site—or matched to the characteristics and interests of the visitors your content attracts—you'll finally have a way to both monetize and enhance your content pages."

AdWords

Google's cost-per-click (CPC) advertising program. These are the ads that you see in Google's search results. AdWords takes click-through rate into consideration, in addition to the advertiser's bid and other factors, to determine the ad's relative position within the paid search results.

Affiliate site

An affiliate site markets products or services on behalf of another website. It sends visitors to another website to complete the sale, in exchange for fees or commissions.

AJAX (Asynchronous JavaScript and XML)

Allows you to create a more user-friendly web application by working

behind the scenes (inside a web browser) to make web pages feel more responsive. Using AJAX, you can create applications on your pages that change the content of the page without having to reload the entire web page.

alt

Designed to provide an alternative text description (a text equivalent) for images. People often mistakenly refer to these as “alt tags.”

Analytics tool

A program that assists in gathering and analyzing data about website usage. Google Analytics is a feature-rich, popular, free analytics program.

Anchor text

The actual text part of a link (usually underlined). Search engines use anchor text as an important ranking factor because it helps them determine the relevancy of the referring site and the link to the content on the landing page. Google pays particular attention to the text used in a hyperlink and associates the keywords contained in the anchor text to the page being linked to.

API

Programming tools that provide a gateway to data from an application or website without having to visit that website to obtain the desired data. For example, Google used to provide Google Suggest autocomplete keyword suggestions to third-party websites via an API but closed off access to that API in August 2015.

Article directory

A website that hosts syndicated articles submitted by users. EzineArticles.com and The Free Library are examples of article directories.

Article syndication

The process of creating content intended for publication on other websites (and normally not on your own site). This is an SEO strategy used to

increase the number of backlinks pointing to your site.

Authority site

A website that has many incoming links from other related authoritative expert/hub sites.

Automated submitting

The use of automated software to submit your web pages to the search engines. The search engines frown upon this practice.

Backlinks

Inbound links pointing to a web page or site that originates from another page or site.

Ban

When a search engine blocks your site from appearing in its search results.

Beacon

A line of code placed in an ad or on a web page that helps track the visitors' actions, such as registrations or purchases. A web beacon is often invisible because it's only 1×1 pixel in size and has no color. Also known as a *web bug*, *1-by-1 GIF*, *invisible GIF*, or *tracker GIF*.

Bing

Search engine owned by Microsoft, formerly known as Live Search and before that MSN. Yahoo! sources its search results information from Bing.

Bingbot

The name of the program that Bing uses to crawl the Web.

Black hat SEO

The opposite of white hat SEO. Sometimes called *spamdexing*, black hat

SEO is the practice of using optimization tactics that cause a site to rank more highly than its content would otherwise justify, or making changes specifically for search engines that don't improve the user's experience of the site. In other words, black hat SEO is optimization that goes against search engine guidelines. If you step too far over the mark, your site may be penalized or even removed from the index.

Blacklists

Lists that either search engines or vigilante users compile of search engine spammers, which may be used to ban those spammers from search engines or to boycott them.

Blog (a.k.a. weblog)

An online diary or journal with entries made on a regular if not daily basis. Blog authors choose whether to blog openly or anonymously; some blogs are maintained by anonymous authors who use nicknames or handles instead of their real names.

Weblog entries are made regularly and chronologically but are displayed in reverse chronological order. The range of topics covered is endless. Some blogs focus on a particular subject, like travel, fashion, or astrology, while others are personal online diaries.

Body copy

Popular blogging platforms include WordPress, Blogger, TypePad, and Moveable Type.

Body copy

The "meaty" textual content of a web page. Body copy refers to text visible to users and does not include graphical content, navigation, or information hidden in the HTML source code.

Bot

Short for "robot" and often also referred to as a *spider* or *crawler*, a bot is a

program that performs a task in an automated fashion. Search engines use bots to find web pages and add them to their search indexes. Spammers often use bots to “scrape” content for the purpose of plagiarizing it for exploitation.

Bounce rate

The percentage of users who enter a site and then leave it without viewing any other pages (i.e., who visit only one page).

Breadcrumbs

Website navigation in a horizontal bar above the main content that helps the users to understand where they are on the site and how to get back to the root areas.

Cache

A collection of copies of web pages stored locally on an Internet user’s hard drive or within a search engine’s database. The cache is what allows web pages to load so quickly when a user hits the back button in the web browser; the page is stored, so it does not have to be downloaded again. Google is unusual among search engines in that it allows Internet users to view the cached versions of web pages in its index. Simply click on the word “Cache” next to the search result of interest, and you will be taken to a copy of the page as Googlebot discovered and indexed it. This feature of Google makes it easy to spot cloaking.

Cascading Style Sheets (CSS)

A mechanism that allows publishers to control the design and layout of website through a stylesheet without having to embed the details in the site’s HTML code.

Click-through

The action of clicking a link, such as one found in an ad element or a search result, causing the user to be sent to another web page.

Click-through rate (CTR)

A measure of the number of clicks received from the number of impressions delivered. If your page or ad is viewed 100 times and this generates 8 clicks, the CTR is 8%.

Cloaking

The practice of using user agent or IP address detection to recognize incoming visitors as either search engine spiders or users, and then delivering content to the spiders that differs from the content seen by human users.

Comment spam

Blog or forum comments posted solely for the purpose of generating an inlink to another site. This is the reason most forums and blogs nofollow comment links and/or use third-party comment management systems such as Disqus and IntenseDebate that utilize iframes or JavaScript to render comments.

Content management system (CMS)

A publishing platform intended to make the task of publishing and maintaining a website easier.

Content marketing

The practice of publishing high-quality content that helps build a company's reputation and visibility, but at the same time results in people linking to the company's website.

Contextual advertisement

Advertising that is related to the content.

Conversion

The act of converting a website visitor into a customer, or at least taking that visitor a step closer to customer acquisition (e.g., by convincing him to sign

up for your email newsletter, download a white paper or free tool, etc.).

Conversion rate

The rate at which visitors get converted to customers or are moved a step closer to customer acquisition.

Cookie

Information placed on a visitor's computer by a web server. While the website is being accessed, data in the visitor's cookie file can be stored or retrieved. Cookies are primarily used as unique identifiers (i.e., user IDs or session IDs) to isolate a visitor's movements from others' during that visit and subsequent visits. Other data that may get stored in a cookie includes an order number, email address, or referring advertiser.

Cost per action (CPA)

The cost incurred or price paid for a specific action taken, such as signing up for an email newsletter, entering a contest, registering on the site, completing a survey, downloading trial software, or printing a coupon.

Cost per click (CPC)

The cost incurred or price paid for a click-through to your landing page. This cost normally relates to a scenario where you have placed an advertisement on another site, such as the search engines' sites (through Google AdWords or Microsoft adCenter) or a third-party website, or through an affiliate program.

Cost per thousand (CPM)

The price paid for a thousand impressions.

Crawler

Another name for a search engine's program for traversing the Web by following hyperlinks. This program can also be referred to as a *bot*, *robot*, or *spider*.

Database-driven

A database-driven website is connected to a database, and web page content is based in part on information extracted from that database.

Directory

This is a site where human editors group websites into categories and provide site descriptions (or edit the descriptions that are submitted to them). With a directory, picking the right category and composing a description rich in key phrases will ensure maximum visibility. Contrast this with a search engine, which is not human edited.

Doorway

Also known as a *bridge page* or a *gateway page*, a doorway page is a web page full of keyword-rich copy that doesn't deliver any useful information other than a link into a site, and whose sole purpose is to capture traffic from search engines and then send that traffic to a highly commercial web page. Microsites that don't add much new information and that focus on sending traffic to a main site can also be considered doorways.

Duplicate content

Content on one web page that is similar or identical to that found on another website or page.

Ecommerce site

A website devoted to retail sales.

Editorial link

A link that is published only because the webmaster considers it to add value for users of that site.

FFA (Free For All) site

Sometimes called *link farms*, these are sites (or pages) with many outgoing

links to unrelated websites, containing little if any unique content. Link farms are intended only for spiders and have little or no value to human users; when discovered, they are ignored or penalized by the search engines.

Flash

A technology developed by Macromedia Inc. (now owned by Adobe) that allows a web designer to embed interactive multimedia into web pages.

Forum (a.k.a. discussion forum)

A virtual community used for information exchange. Users can post messages in different forums, either to the group at large or to certain users. However, all postings can be seen by anyone else who has access to that forum, so save sensitive materials for private email! Forums are also threaded, which means a reply to a particular post becomes part of that post's "thread" and can be followed to provide a cohesive progression through a particular topic.

Frames

A method for combining separate web pages into one, each potentially with its own scroll bar. You know you're on a framed website when part of the page scrolls while the rest of the page stays in place. Frames can frustrate people because when they try to bookmark a specific page it often doesn't work as intended but instead bookmarks the "frameset" page, which is typically the home page.

Geotargeting

Pinpointing the user's location for the purpose of serving targeted advertisements or web pages. For example, Google AdWords and Microsoft adCenter allow for targeting of countries, states, cities, and suburbs.

Google bowling

A black hat SEO technique used to knock competitors down in, or out of, the search engine results. This form of SEO sabotage is conducted by pointing

hundreds of questionable links from low-quality sites at a competitor's site so that site ends up being banned or penalized by Google.

Generally, newer sites are more susceptible to Google bowling, as older sites are better established with a range of existing high-quality links.

Google supplemental index

A secondary Google database containing supplemental results pages that are deemed to be of lesser importance by Google's algorithm or are less trusted. These are pages that are less likely to show up in the search results.

Googlebot

The name of the program that Google uses to crawl the Web.

Grey hat SEO

SEO using both black hat and white hat techniques. This can also refer to the use of techniques that do not strictly follow the search engines' guidelines but that are only slightly over the line.

Heading tag

An HTML tag that is often used to denote a page or section heading on a web page. Search engines pay special attention to text that is marked with a heading tag, as such text is set off from the rest of the page content as being more important.

Hidden text

Contextual HTML text that is concealed from human visitors to a web page but is made available to search engines spidering the site. Using hidden text is considered an SEO spam tactic.

Hijacking

Hijacking of websites is a practice that makes search engines believe that a specific website resides at another URL. It is a form of cloaking and an

illegal spam tactic. Spammers may do this to increase a page's ranking in the search engine results pages.

HTML (HyperText Markup Language)

The main markup language for the creation of web pages, used to mark up web content and display it in a formatted manner. HTML defines how data is structured and informs the web browser how the page is to be displayed. It's up to the web browser software (e.g., Microsoft Internet Explorer or Mozilla Firefox) to render the HTML source.

Some of the page elements that can be coded with HTML include page titles, text (paragraphs, lines, and phrases), lists (unordered, ordered, and definition lists), tables, forms, basic HTML data types (character data, colors, lengths, content types, etc.), and much more.

The source HTML code of any web page can be viewed in a web browser such as Firefox (right-click and select "View Page Source") or Internet Explorer (select View → Source).

HTML is not a programming language and therefore is quite static in nature. It is considered to be a subset of SGML (Standard Generalized Markup Language).

Tim Berners-Lee first described HTML in 1991, in a publicly available document called "HTML Tags." The Internet Engineering Task Force (IETF) published the first draft proposal in 1993. HTML became an international standard (ISO/IEC 15445:2000), and its specifications are maintained by the World Wide Web Consortium (W3C).

HTTP

Short for HyperText Transfer Protocol. This is one of the basic methods used by browsers and web servers to transfer data information (such as web pages) across the Web. One of its weaknesses is that the information is sent as simple text, and is easily intercepted, read, and modified, by third parties.

HTTPS

Short for HyperText Transfer Protocol Secure. This is one of the basic methods used by browsers and web servers to transfer data information (such as web pages) across the Web. Unlike with HTTP, the data sent is encrypted, making it quite difficult for third parties to intercept and read or modify the data.

Hub

Trusted pages or sites that link to the authority websites in a given market space. For example, in the online music space, a hub would be a page or site that links to most of the authoritative sites for online music (and does not link to any, or extremely few, poor-quality sites).

Impression

Any event where a user views a web page or ad. Each view is considered one impression.

Inbound link

Any link that points to your site from a site other than your own. Inbound links are an important factor in the search engines' algorithms.

Index (n.)

A search engine's database, in which it stores textual content and other information about web pages that have the potential to show up in its search results.

Index (v.)

The act of adding a web page to a search engine's index.

Indexed pages

The pages on a site that have been incorporated into a search engine's database.

Information retrieval (IR)

The part of computer science dealing with the retrieval of information (not data) from a collection of written documents. The retrieved documents attempt to satisfy a user information need, usually expressed in natural language.

Interlinking

The linking structure of various web pages within a site that helps users and spiders navigate its content.

Internal link

A hypertext link that points to another page within the same website. Internal links can be used as a form of navigation, directing visitors to pages within the website. Links assist with creating good information architecture within a site.

Search engines also use internal text links to crawl pages within a website. The way internal links are structured will impact the way in which search engine bots spider and subsequently index a site's pages.

Internet service provider (ISP)

An ISP provides a range of Internet-related services to customers, including Internet connectivity, email, website hosting, and domain name registration.

Inventory

Usually provided services for a monthly fee, an ISP can be a commercial business, a university, a government organization, a school, or any other entity that provides access to the Internet to members or subscribers.

ISP

An acronym for Internet service provider.

Inventory

Advertising space available for purchase on a website. Based on projections, inventory may be specified as number of available impressions or as a share of voice.

Invisible Web

A term that refers to the vast amount of information on the Web that is not indexed by the search engines. Coined in 1994 by Dr. Jill Ellsworth.

IP (Internet protocol)

The communications protocol used for sending datagrams (packets of information) across the Internet.

IP address

Each server or device connected to the Internet is assigned a unique permanent (static) or temporary (dynamic) IP address. It is expressed as a four-part series of numbers separated by periods (.) that identifies every sender and receiver of network data. The numbers, each of which has a range of 0–255, represent the domain, the network, the subnetwork, and the host computer (for example, 127.0.0.10). The IP address sometimes translates into a specific domain name.

ISAPI_Rewrite

A module or plug-in for Microsoft IIS web servers that can be used to rewrite requested URLs on the fly. Modeled after mod_rewrite, it supports an unlimited number of rules and an unlimited number of attached rule conditions for each rule to provide a flexible and powerful URL-manipulation mechanism that can be used to offer search engine–friendly URLs, thus increasing indexing chances for a dynamic database-driven website.

Java applets

Small programs written in the Java programming language that can be embedded into web pages. Applet programs run on the Internet user’s

computer rather than the web server's computer.

JavaScript

A programming language that runs on the Internet user's computer rather than the web server's computer. In the past, search engines did not run JavaScript, but that is changing.

Key performance indicators (KPIs)

Metrics selected by an organization to help it achieve organizational goals through the definition and measurement of progress. The KPIs are agreed upon by an organization and are indicators that can be measured that reflect success factors. The KPIs selected must reflect the organization's goals, they must be key to its success, and they must be measurable.

Keyword

A word or phrase that a search engine user might use to find relevant web pages. If a keyword doesn't appear anywhere in the text of your web page, it's less likely that your page will appear in the search results for that term (unless of course you have bid on that keyword in a pay-per-click search engine).

Keyword cannibalization

The excessive reuse of the same keyword on too many web pages within the same site. This practice makes it difficult for the users and the search engines to determine which page is most relevant for the keyword.

Keyword density

The number of occurrences of a given keyword on a web page. It used to be that the more times a given word appeared on your page, the more likely it was that the search engine would return that page when users searched for that keyword. As search algorithms have gotten much more sophisticated, this is no longer the case, although it is helpful to have the target word or phrase show up a small number of times on the page.

Keyword popularity

The number of searches Internet users have done for a given keyword during a period of time.

Keyword prominence

A measure of the location (i.e., placement) of a given keyword in the HTML source code of a web page. The higher up in the page a particular word is, the more prominent it is and thus the more weight that word is assigned by the search engine when that word matches a keyword search done by a search engine user. Consequently, it's best to have your first paragraph contain important keywords rather than superfluous marketing speak. This concept also applies to the location of important keywords within individual HTML tags, such as heading tags, `<title>` tags, or hyperlink text.

Keyword research

The process of determining the words and phrases that people use to find something, then compiling them into a list for use on web pages, websites, and more.

Keyword stuffing

Placing excessive amounts of keywords into the page copy and the HTML in such a way that it detracts from the readability and usability of a given page, for the purpose of boosting the page's rankings in the search engines. This may include techniques such as hiding keywords on the page by making the text the same color as the background, hiding keywords in comment tags, or overfilling alt tags with long strings of keywords. Keyword stuffing is just another shady way of gaming the search engines, and as such, it is strongly discouraged.

KPIs

An acronym for key performance indicators.

Latent semantic indexing (LSI)

An algebraic model of document retrieval based on a singular value decomposition of the vectorial space of index terms. In plain English, LSI is a method for extracting the words on a web page to determine the topics that it is likely to be relevant for.

Link bait

A web page with useful or entertaining content intended to attract incoming links, often mostly via social media. This term is not used as much anymore because it sounds manipulative.

Link building

A marketing discipline whose purpose is to obtain inbound links from webmasters of other sites with the aim of increasing “link popularity” and/or PageRank. There are many methods for link building, as described in Chapter 7.

Link exchange

The practice of exchanging links with another site (a.k.a. reciprocal linking).

Link farm

A group of highly interlinked websites formed for the purpose of inflating link popularity (or PageRank). This is a form of link spam.

Link juice

A generic term used to refer to the value of a link. This term is much broader than the original term coined by Google, PageRank. Unlike PageRank, link juice is meant to consider the notions of the trust and authority of a link.

Link popularity

A measure of the value of a site based upon the number and quality of sites that link to it.

Link spam

Links between pages that exist for the sole purpose of inflating search engine rankings. Frequently these are links that do not provide end user value.

Link text (a.k.a. anchor text)

The user-visible text of a link. Search engines use anchor text to determine the relevancy of the referring site and the link to the content on the landing page. Ideally all three will share some keywords in common.

Linkerati

Internet users who own websites that have a tendency to link to other quality content. Bloggers and news media are part of the linkerati, as are other site owners who link to high-quality content.

Logfile

Most web servers maintain a log of all accesses to a website. Data that is usually logged includes the date and time of access, the filename(s) accessed, the user's IP address, the referring web page, the user's browser software and version, and cookie data.

Long-tail search queries

These are longer queries (i.e., queries containing more words, such as four or more) that are often more targeted than shorter broad queries. For example, a search for *widgets* would be very broad, while *red widgets with reverse threads* would be a more specific long-tail search. A large percentage of all searches are long-tail searches.

LSI

An acronym for latent semantic indexing.

Made for AdSense (MFA)

MFA sites are websites that have been designed from the ground up as a venue for Google AdSense advertisements.

Meta description

A meta tag hidden in the HTML that describes the page's content. Should be relatively short; around 12 to 20 words is suggested. The meta description provides an opportunity to influence how your web page is described in the search results, but it will not improve your search rankings.

Meta keywords

A meta tag hidden in the HTML that lists keywords relevant to the page's content. Because search engine spammers have abused this tag so much, it provides no benefit to your search rankings.

Mobile-friendly

A phrase used to describe web pages that render in a user-friendly way on smartphones or tablets, which have smaller screens and don't have access to a full-sized keyboard or a mouse.

Mobilegeddon

The phrase the industry used to label Google's planned Mobile-Friendly algorithm update prior to its launch on April 21, 2015.

mod_rewrite

A module or plug-in for Apache web servers that can be used to rewrite requested URLs on the fly. It supports an unlimited number of rules and an unlimited number of attached rule conditions for each rule to provide a flexible and powerful URL-manipulation mechanism that can be used to offer search engine–friendly URLs, thus increasing indexing chances for a dynamic database-driven website.

Monetization

The process by which a publisher derives income from a site. For example, placing AdSense ads on your site is one strategy for monetizing it.

`nofollow`

A link attribute (`rel="nofollow"`) that instructs the search engines not to pass any link juice on to the linked-to page. NoFollow can also be specified in the head section of a web page, with the result that none of the links on the page will pass link juice to the linked-to pages.

`noindex`

A value found in the head section of a web page that instructs robots to not index the page.

Nonreciprocal link

If site A links to site B, but site B does not link back to site A, then the link is considered nonreciprocal. Search engines tend to give more value to nonreciprocal links than to reciprocal ones because they are less likely to be the result of collusion between sites (a.k.a. *link exchange* or *reciprocal linking*).

Organic search results

Search engine results that are not sponsored, or paid for in any way.

Outbound link

A link that points “off-site” to another website. Also known as an *outlink* or *outgoing link*.

Page load time

The amount of time it takes for a page to load. With the intent to improve search quality, Google considers page load time as a ranking factor in its algorithm.

Pagejacking

Stealing high-ranking web page content from another site and placing it on your site in the hopes of increasing your site’s search engine rankings.

Pagejacking is yet another shady way of gaming the search engines, and as such it is strongly discouraged.

PageRank (PR)

A Google algorithm for measuring the importance of a web page based on its inbound link profile. PageRank is a purely numerical measurement and does not take into account the relevance or the linking page, or how trusted or authoritative that page might be.

PDF

Adobe's Portable Document Format, a file format that renders the page exactly as intended regardless of the computer used. Typically used for creating documents that will be printed. PDF is used instead of HTML when the content creator wants absolute control over the display of the document. In contrast, the display of an HTML document depends on the computer and web browser software used.

PHP

An "open source" programming language for building dynamic websites, PHP can be used to write server-side programs that access databases. PHP is the most popular web programming language—more popular than JSP (Java Server Pages), Microsoft's ASP (Active Server Pages), and Macromedia's Cold Fusion. PHP is especially well suited to web development and can be embedded into HTML. It is secure, easy to learn, efficient, fast to code, and fast to deploy. As of 2011, PHP is used by over 9 million websites (over 24% of the sites on the Internet), due largely to benefits it offers such as quicker response time, improved security, and transparency to the end user.

Probabilistic model

A classic model of document retrieval based on a probabilistic interpretation of document relevance to a given user query.

Pull-down list

A web form where the user chooses from a list of items. For example, you might be asked to identify which country you are from using a pull-down list. A pull-down list usually displays the default choice within a box and a down arrow immediately to the right. Clicking on the down arrow will display the full list to choose from.

Query

A keyword or phrase inquiry entered into a search engine or database. In search, a person types in the word or words she is looking for information on, and the search engine database returns results that match the user's query.

Reciprocal linking

The practice of trading links between websites (a.k.a. link exchange). Search engines usually don't see these as high-value links, unless the relevance of the two swapping sites to each other is very high.

Redirect

Any method used to change the address of a landing page, such that the user is automatically taken to another URL without clicking on anything.

Referrer

When a web page receives traffic as a result of a user clicking on a link on another web page, this is considered a referred visitor. The page containing the link the user clicked on is called the *referrer*. Seeing which pages on the Web send traffic to your website can provide a lot of useful data for your SEO (and other marketing) efforts.

Relative link

A link that does not specify the full web address. Because the *http://* part of the link is not specified, the web server knows that the link is relative to the current location of the user and looks for the specified page in the same directory as the current page.

This is in contrast to an absolute link, which specifies all aspects of where

the content is to be found.

Relevance

The likelihood that a given web page will be of interest or useful to a search engine user for a keyword search.

Repeat visitor

A single individual or browser who accesses a website or web page more than once over a specified period of time.

Return on investment (ROI)

The benefit gained in return for the total cost of investing budget into advertising or some other project. ROI can be measured by one of the following two calculations:

- ROI = (Revenue – Cost) / Cost or, more simply:
- ROI = Revenue / Cost

Robot

Another name for a search engine's program for traversing the Web by following hyperlinks. This program can also be referred to as a *crawler*, *bot*, or *spider*.

robots.txt

A text file placed in a website's root directory and linked in the HTML code that controls the behavior of search engine spiders on the site, or even denies them access.

Sandbox

An outdated concept that referred to Google putting new sites into a "holding pen," preventing them from ranking well for anything competitive until a period of time has passed.

Search Console

A web application offered by Google that provides publishers with information about how Google sees their website. This program was formerly known as Google Webmaster Tools.

Search engine marketing (SEM)

A term often used to describe acts associated with efforts to achieve maximum exposure of a website within the SERPs. SEM may include search engine optimization, PPC campaigns (paid listings), social media campaigns, and other search engine–related services and functions that will increase the exposure of and traffic to your website and increase the amount and quality of sales leads. Note that many people mistakenly use this term to refer only to pay-per-click advertising on search engines.

Search engine optimization (SEO)

The process of increasing the number of visitors to a website by achieving high rank in the search results returned by a search engine. The higher a website ranks in the results pages, the greater the chance is that users will visit the site.

Search engine results page (SERP)

A page of search results delivered by a search engine.

Search engine spam

The result of manipulation techniques that violate search engine guidelines—for example, pages created to cause search engines to deliver inappropriate or less relevant results. SEO professionals are sometimes unfairly perceived as search engine spammers.

Search intent

The intention of the end user who is searching for information on the Internet. Search engines are focused on improving the user experience by better understanding search intent so they can deliver consistent and relevant

search results. On the other hand, advertisers strive to better understand users' search intent in order to sell them relevant products and services.

Search term

A keyword or phrase used to conduct a search engine query.

SEM

An acronym for search engine marketing.

SEO

An acronym for search engine optimization.

SERP

An acronym for search engine results page.

Sitemap

There are two types of sitemaps, HTML sitemaps and XML Sitemaps. An HTML sitemap is a page or structured group of pages that link to every user-accessible page on a website, and hopefully improve site usability by clarifying the data structure of the site for the users. An XML Sitemap is a file that is not visible to users, but is designed to help search engine spiders to find all of the site's pages.

Social media marketing (SMM)

The practice of website or brand promotion through social media. SMM is a subset of SEM.

Spam

See *search engine spam*.

Spamglish

Keyword-rich gibberish used as search engine fodder instead of thoughtfully

written, interesting content. Spamglish often includes meaningless sentences and keyword repetition.

Spammer

A person who uses techniques that violate search engine guidelines to obtain high search rankings for his website(s).

Spider

Another name for a search engine's program for traversing the Web by following hyperlinks. This program can also be referred to as a *crawler*, *bot*, or *robot*.

Spider trap

An infinite loop that a spider may get caught in if it explores a dynamic site where the URLs of pages keep changing—for example, a crawlable calendar program that goes on infinitely.

Static page

A web page that always shows the same content. Pages whose content is dynamically generated when they are requested may still be considered static pages if the actual content eventually rendered does not change from access to access.

Stemming

A process used by search engines to reduce words to their grammatical roots and deliver results based on the spelling of the root. For example, a search for *clothes* will also return results for clothing.

Stickiness

A measure of a site's ability to retain visitors' interest, by getting them to view more pages or spend more time on the site.

Stopword

Certain words—such as *the*, *a*, *an*, *of*, and *with*—are so common and meaningless that a search engine won't bother including them in its index, or database, of web page content. So in effect, the stopwords on your web pages are normally ignored as if those words weren't on your pages in the first place. Including a lot of stopwords in your `<title>` tag waters down the `<title>` tag's keyword density. Note that there are important exceptions that search engines track, though, such as references to the TV show *The Office*, where the word *the* is very important to the context. In addition, searchers can implement search queries in double quotes to require that the stop words not be ignored.

Supplemental index

Google's supplemental index is a secondary database containing supplemental results pages that are deemed to be of lesser importance by Google's algorithm or are less trusted. These are pages that are less likely to show up in search results.

Tagging, tags

Simple word descriptions used to categorize content.

Target audience

The market to whom advertisers wish to sell their products or services.

Taxonomy

Classification system of controlled vocabulary used to organize topical subjects, usually hierarchical in nature.

Text link

A plain HTML link that does not involve graphic or special code such as Flash or JavaScript.

Theme

The main keyword focus of a web page.

Thin affiliate

An affiliate site that provides little value-added content.

Thin content

Content that offers little material value. This type of content is considered undesirable by search engines as they prefer not to offer it to their users in the search results.

Time on page

The amount of time that a user spends on one page before moving to a different page. This is an indication of quality and relevance.

`<title>` *tag*

The text displayed at the very top of the browser window, above the address bar and the Back, Forward, Print, and other buttons. Although inconspicuous to the user, the `<title>` tag contains the most important bit of text on a web page as far as the search engines are concerned.

Toolbar

A browser add-on usually including a search box. A popular example is the Google Toolbar.

Toolbar PageRank (PR)

A value between 0 and 10 assigned by the Google algorithm, which very loosely quantifies page importance and is not the same as PageRank.

Traffic

The number of users that surf to a site.

TrustRank

A measurement of how much a website or web page is trusted. It is likely that the search engines measure this in some form, but there are many

different potential algorithms for doing so.

Unique visitors

Individual users who have accessed your website over a period of time. If one person visits your site multiple times during that period, that user still only counts as one unique visitor. For example, the number of daily unique visitors is equal to the number of different people who visit the site during that day.

URL (Uniform Resource Locator)

Used interchangeably with *web address*, a URL can specify the location of a web page, an email address, or a file on an FTP server, among other things.

URL rewrite

A technique used to help make website URLs more user-and search engine–friendly.

Usability

A reference to how easy it is for a user to accomplish what she wants on a website.

User agent

A text string exchanged between a browser or a robot and a web server that identifies the source of the visit to the website. It allows the web server to get some basic information on the visiting person or bot. For example, `Googlebot/2.1` is a user-agent string that Googlebot used for a period of time in the past.

User intent

In the context of the search engines, user intent refers to the objective of a user conducting a search query. This may not always be immediately clear just by reading the query string itself.

User-generated content (UGC)

Content created and published by the end users online. UGC is composed of videos, podcasts, and posts on discussion groups, blogs, wikis, and social media sites. UGC allows for a wider content provider base and the chance for all users to share their opinions online. Criticism of UGC includes credibility and quality issues.

Webmaster Tools

Google offers a web application that can be used by publishers to get information about how Google sees their website. This used to be called Google Webmaster Tools, but as of 2015 was renamed to Search Console. Bing also offers a similar program, called Bing Webmaster Tools.

User session

An instance of an Internet user accessing your website for a length of time, then leaving the site. During a user session any number of pages may be accessed. A user session is also considered finished once an arbitrarily chosen period of inactivity occurs—typically 30 minutes—even if the user never leaves the site.

Web address

Used interchangeably with *URL*; can specify the location of a web page, an email address, or a file on an FTP server, among other things.

Web browser

Software installed on the Internet user's computer that allows him to view web pages. Popular web browsers include Microsoft Internet Explorer, Firefox, Safari, Chrome, and Opera.

White hat SEO

A term that refers to ethical SEO practices that conform to the guidelines of the major search engines and do not attempt to "game" or manipulate the SERPs.

XML (eXtensible Markup Language)

A scripting language that allows the programmer to define the properties of the document.

XML Sitemap

An XML file that lists the URLs available on a site for the search engines to crawl. The aim is to help site owners notify search engines about the URLs on a website that are available for indexing. XML Sitemaps are not viewable by end users; they exist solely to help search engines better understand what content is available on a website.

Index

Symbols

 # (hash mark), Single-Page Applications

#! (hashbang), Single-Page Applications

*** (asterisk) character, Common usage and phrase combinations**

+Post ads, Social content amplification steps

.com vs. other TLDs, When to Use a TLD Other Than .com, Optimizing Domains

.edu domain, Source TLDs

1-by-1 GIFs, Glossary

10 blue links, Overview of Changes in Search Complexity

301 status code, Search engine–friendly navigation guidelines, Using the rel="canonical” link element, Good and Bad Redirects, Content Management System Issues, The Basics of Moving Content, Glossary

302 status code, Technology Decisions, Search engine–friendly navigation guidelines, Good and Bad Redirects, Glossary

400 status code, Glossary

401 status code, Glossary

403 status code, Glossary

404 status code, Redirects, Glossary

410 status code, Glossary

500 status code, Glossary

501 status code, Glossary

A

A/B tests, Landing Page Optimization, Showing Different Content to Engines and Visitors

absolute links, Glossary

accessibility (for disabled persons), Understanding What Content Search Engines Can “See” on a Web Page, Glossary

accessibility (to search engines)

- following sitemaps protocol, XML Sitemaps
- importance of, Accessibility/spiderability, Developing an SEO-Friendly Website
- search-friendly site navigation, Search-Friendly Site Navigation-Search engine–friendly navigation guidelines
- supplying indexable content, Indexable Content
- supplying spiderable link structures, Spiderable Link Structures, Internal Link Structure and Site Architecture

action tracking, Taking a Deeper Look at Action Tracking, Action tracking by referral source

ad density, Ad Density and Offensive Ads, Summary of Major Google Algorithms

adaptive search, Adaptive Search, Other Ranking Factors

adCenter, Glossary

AdSense, Glossary

advertising

- **ad density, Ad Density and Offensive Ads, Summary of Major Google Algorithms**
- **consumer dissatisfaction with, Searcher Intent**
- **contextual, How Google and Bing Collect Engagement Metrics**
- **data brokers, Psychographic targeting**
- **discovering audiences through, Get to Know Other People's Audiences**
- **offensive, Ad Density and Offensive Ads**
- **PPC results ordering, Understanding the Layout of Search Results Pages**
- **shift toward mobile ad spend, The Mobile Shift, The Mobile Landscape**
- **social content amplification steps, Social content amplification steps**

advertorials, Links Google Does Not Like

AdWords, Ad Campaign Runs and Third-Party Search Data, Promoting Products in AdWords, Traffic by keyword, Glossary

affiliate sites, A word of caution to affiliates, Glossary

Agent Rank, A Brief History of Google Authorship, Is Author Authority Dead for Google?

aggregators, Directories and aggregators

Ahrefs Site Explorer, Ahrefs

AJAX (Asynchronous JavaScript and XML), What search engines cannot see, AJAX and JavaScript-AJAX and JavaScript, Glossary

Alexa, Alexa

algorithmic ranking criteria, Retrieval and Ranking, How Social Media and User Data Play a Role in Search Results and Rankings

algorithmic search results

- failures in, Why These Algorithms Sometimes Fail
- ranking systems for, Algorithm-Based Ranking Systems: Crawling, Indexing, and Ranking-What search engines cannot see
- SERP section for, Understanding the Layout of Search Results Pages
  - (see also organic search results)

algorithms

- analyzing differentiation across engines/search types, Analysis of Algorithmic Differentiation Across Engines and Search Types
- changes/updates to, Diagnosing the Cause of a Traffic Loss, Establishing a Proper Baseline
- DMCA/pirate penalty, Summary of Major Google Algorithms
- exact-match domain (EMD), Summary of Major Google Algorithms

Farmer, Panda

Hummingbird, Document Analysis and Semantic Connectivity, Understanding the Impact of Google Hummingbird, Google's Hummingbird, Summary of Major Google Algorithms

mobilefriendly determination, SEO for Mobile

Panda, Content Quality and User Engagement, Panda-Successful Panda recovery example, Establishing a Proper Baseline

payday loan/spammy sites, Summary of Major Google Algorithms

Penguin, Google's Penguin Algorithm, Penguin, Establishing a Proper Baseline

Top Heavy update, Summary of Major Google Algorithms

Allow directive, Syntax of the robots.txt file

alt (alternative text description), Glossary

audit review of, Image alt attributes

content management systems and, Content Management System Issues

indexable content and, Indexable Content, Java, images, audio, and video

keyword targeting and, Image Filenames and alt Attributes

vs. search by image feature, Image search

uses for, Understanding What Content Search Engines Can "See" on a Web Page

analytics tools, Glossary

(see also web analytics)

anchor text, Crawling and Indexing, Link Analysis, Use anchor text, intuitively, Content Management System Issues, Content Management System Issues, Anchor text, Target Areas of Penguin, Links Google Does Not Like, Optimizing Your Anchor Text, LinkBased Tracking of Content Marketing, External Link Attraction Analysis, Glossary

AngularJS, What search engines cannot see, Angular.js: Making it SEO-friendly

annotations, Optimize your annotations

antispam mechanisms, Content Management System Issues

App Indexing API, App Indexing

apps (see mobile app content)

article directories, Content Syndication, Social news sites, Target Areas of Penguin, Links Google Does Not Like, Glossary

article syndication, Implementing Content Marketing Strategies, Content Syndication, Glossary

asterisk (*) character, Common usage and phrase combinations

attribution gap , How Search Engines Drive Ecommerce, Managing Attribution

audience segmentation, Segmenting Your Site's Audience, Segmenting Your Audience, Identifying Personas, and Targeting Content

audiences, building

other people's audiences, Get to Know Other People's Audiences-The

indirect benefits of building reputation and visibility

through social media, The Indirect Influence of Social Media Marketing (see also target audiences)

audio files, What search engines cannot see, Java, images, audio, and video

audits

elements of, Elements of an Audit-Code quality

internal links correction, Example: Fixing an Internal Linking Problem

keyword cannibalization, Keyword Cannibalization

keyword reviews, The Importance of Keyword Reviews

link profiles, Negative Link Building

role in SEO, Auditing an Existing Site to Identify SEO Problems

server/hosting issues, Server and Hosting Issues, Search engine–friendly navigation guidelines

strategic reviews, Conduct regular strategic reviews

for website assets, Review your website assets

authenticated spam reports, Other Algorithmic Approaches to Fighting Link Spam, Penalties

Author Authority, The Future of Author Authority at Google

author profile pages, Author Profile Pages

authority site, Authority, Guest Posting, Role of Authority in Rankings, Not

enough link authority to remain in main indexes, Glossary

AuthorityLabs, AuthorityLabs

automated link dropping, Using automated link dropping

automated submitting, Glossary

automatically generated content, Keyword Targeting in Content Management Systems and Automatically Generated Content, Target Areas of Panda

awards, Awards and Badges

AWStats, Log file Tracking

B

B2B sites

action tracking for, Taking a Deeper Look at Action Tracking

leveraging LinkedIn for, LinkedIn, LinkedIn

search engines for, More specialized vertical search engines

backlinks, Execute aggressively, Perform research and analysis to pick potential targets, Swapping links, Search engine–supplied tools, Glossary

badges, Awards and Badges

Baidu, Country-Specific Search Engines

ban, Glossary

bandwidth limitations, Server and Hosting Issues

beacons, Glossary

benchmarking

- **current indexing status, Benchmarking Current Indexing Status, Site Indexing Data**
- **current traffic, Benchmarking Current Traffic Sources and Volume, Google Search Console**
- **establishing a baseline, Tracking Results and Measuring Success, Establishing a Proper Baseline**
- **organic rankings, Benchmarking Organic Rankings, Benchmarking Organic Rankings**

biased data, Common analytics mistakes

Bing, Glossary

- **advanced search operators, Bing Advanced Search Operators**
- **app indexing by, App Indexing**
- **experiments with social signals, Bing's Experiments with Social Signals**
- **market share of, An Evolving Art Form: The Future of SEO**
- **move to HTTPS, Bing Webmaster Tools**
- **user engagement metrics, How Google and Bing Collect Engagement Metrics**
- **vertical search products, Vertical Search from the Major Search Engines**

Bing Keyword Research, Bing Keyword Research-Cost, Bing Webmaster Tools

Bing Places for Business, Bing Places for Business

Bing search results

- **compared to Google, Analysis of Algorithmic Differentiation Across Engines and Search Types**
- **page layout, Understanding the Layout of Search Results Pages**
- **ranking score example, Retrieval and Ranking**
- **vertical result examples, Understanding How Vertical Results Fit into the SERPs**

Bing Suggest, Brainstorming Content Ideas and Being Creative

Bing toolbar, How Google and Bing Collect Engagement Metrics

Bing Webmaster Tools

- **Crawl Error data, Bing Webmaster Tools**
- **data on log in screen, Bing Webmaster Tools**
- **defined, Glossary**
- **Diagnostics & Tools section, Bing Webmaster Tools**
- **importance of using, Using Search Engine–Supplied SEO Tools**
- **Index Explorer data, Bing Webmaster Tools**
- **indexing data on, Site Indexing Data**
- **Keyword Research tool, Bing Webmaster Tools**
- **SEO Reports data, Bing Webmaster Tools**
- **setup, Search Engine Tools for Webmasters**

verifying site ownership, Search Engine Tools for Webmasters

Bingbot, Mobile user agents, Glossary

black hat/gray hat SEO, Glossary

automated link dropping, Using automated link dropping

awards and badges, Awards and Badges

customer discounts/incentives, Customer Discount/Incentives

hosting giveaways, Hosting giveaways

link farms, Using link farms/link networks

negative link building, Negative Link Building

noscript tag and, <noscript>

popular tactics, Gray Hat/Black Hat

purchasing links, Purchasing links for SEO

rating yourself, Business Website Optimization for Local Search

swapping links, Swapping links

three-way link swaps, Using link farms/link networks

blacklists, Glossary

blended search, The History of Eye Tracking: How Users Scan Results Pages, Universal Search/Blended Search, Overview of Changes in Search Complexity, Optimizing for Vertical Search

(see also Universal Search)

blocking

by IP address range, Blocking and cloaking by IP address range

by user agent, Blocking and cloaking by user agent

blog search, Blog search

blogs, Blogging for Links, Blogs and forums, Optimizing for Blog Search-Links, Taking a Deeper Look at Action Tracking, Tracking the Blogosphere, Glossary

body copy, Glossary

boldface text, Boldface Text

book search, Book search

bookmark sites, Links Google Does Not Like

Boolean searches, Document Analysis and Semantic Connectivity

bots, Server and Hosting Issues, Search Engine Robot Traffic Analysis, Glossary

bounce rate, Content Quality and User Engagement, Potential User Engagement Signals, Ad Density and Offensive Ads, Glossary

brand pages, Google+ Brand Pages in the Search Results

branded search traffic, Branded versus nonbranded traffic

branding

benefits of SEO for, SEO for Mindshare and Branding

claiming key profiles, Claiming Key Profiles, Local Listing Submissions

critical search terms and, Business Factors That Impact Your SEO Strategy

impact of social media on, The Indirect Influence of Social Media Marketing

keyword choice and, Integrating Emotional Triggers, Titles, and Images

sales campaigns and, Microsites

top placement in search results, Visibility (Branding)

unique monikers and, Optimizing Domains

voting mechanisms and, Voting Mechanisms

(see also reputation management)

breadcrumbs, Use breadcrumb navigation, Content Management System Issues, Glossary

bridge pages (see doorway pages)

business information, submitting, Local Listing Submissions

bylined articles (see guest posting)

C

cache, Content Management System Issues, Glossary

Cached link, Uncovering subscriber-only or deleted content

call-to-action (CTA), Social content amplification steps

cannibalization (see keyword cannibalization)

canonical redirect, Duplicate content checks, URL redirect checks, Using

the rel="canonical" link element, Content Syndication

CAPTCHA boxes, Using login/password protection, Multimedia Is Becoming Indexable

Carousel (see Google Knowledge Graph Carousel)

cascading style sheets (CSS), Spiderable Link Structures, Frames, Page segmentation, CSS and Semantic Markup, Content Management System Issues, Glossary

case-sensitivity, Picking the Right URLs

CC (see Creative Commons)

CDNs (see content delivery networks)

celebrity endorsements, Business Website Optimization for Local Search

centroid, Business Website Optimization for Local Search

Change of Address tool , Expectations for Content Moves

charities, Integrating Emotional Triggers, Titles, and Images, Business Website Optimization for Local Search

Chrome browser

- Incognito mode, Other Ranking Factors, Google+ Personalization
- viewing source code in, Understanding What Content Search Engines Can "See" on a Web Page

Chrome Personal Blocklist Extension, Content Quality and User Engagement, Voting Mechanisms, Panda

citation analysis, Retrieval and Ranking, The Indirect Influence of Social Media Marketing, Links Google Does Not Like, Business Website Optimization for Local Search

click tracking

analytic software setup for, Setting Up Analytics Software to Track Conversions

organic vs. paid results, Click Tracking: How Users Click on Results (Organic Versus Paid)

search result/traffic distribution, Distribution of Search Results and Traffic

click-through action, Glossary

click-through rate (CTR), Glossary

as user engagement signal, Potential User Engagement Signals

by query length, Distribution of Search Results and Traffic

by SERP position, Ad Campaign Runs and Third-Party Search Data

in market opportunity assessment, Understand Your Market Opportunity

meta tags and, Understanding What Content Search Engines Can “See” on a Web Page

problems associated with, An alternate theory and calculation

ROI determination via, Determining Project ROI

cliques, Using link farms/link networks

cloaking, Negative Ranking Factors, Controlling Content with Cookies and

Session IDs, Cloaking and Segmenting Content Delivery, Displaying Different Content to Search Engines Versus Visitors, Blocking and cloaking by IP address range, Other types of manual penalties, Glossary

cloud computing, increased reliance on, Growing Reliance on the Cloud

CMS (see content management systems)

co-occurrence analysis, Understanding the Impact of Google Hummingbird, Understanding Keyword "Not Provided" and Co-Occurrence Analysis, Co-occurrence, phrase-based indexing, and entity salience

code quality, Code quality

cold calls, Use email to contact the influencers

comment spam, Target Areas of Penguin, Other types of manual penalties, Links Google Does Not Like, Glossary

common phrases, discovering, Common usage and phrase combinations

Compete, Compete

competitive research

- analyzing algorithmic differentiation, Analysis of Algorithmic Differentiation Across Engines and Search Types
- analyzing top-ranking sites/pages, Brainstorming Content Ideas and Being Creative, Analysis of Top-Ranking Sites and Pages
- backlink analysis, Find out where your competitors get links
- competitive analysis, Including Competitive Analysis, Competitive Analysis-Competitive Analysis for Those with a Big Budget

link analysis tools, Execute aggressively

natural search and, Understanding Context: Market Competitiveness

overall SEO strategy and, Business Factors That Impact Your SEO Strategy

smaller players in, Determining Top Competitors

comScore, comScore Search Planner

Conductor Searchlight , Tying SEO to Conversion and ROI, Conductor Searchlight, Competitive Analysis for Those with a Big Budget

conferences, Build Offline Relationships, Participation in Conferences and Organizations

contact information, obtaining, Get contact information

contact methods, direct vs. indirect, Identify methods for contacting potential partners, Use email to contact the influencers

"Contact us" forms, Search and web forms, Business Website Optimization for Local Search, Setting Up Analytics Software to Track Conversions

content

attracting viewers to, Integrating Emotional Triggers, Titles, and Images

audit review of, Content review

automatically generated, Keyword Targeting in Content Management Systems and Automatically Generated Content, Target Areas of Panda

categorically structured, Category structuring, Content Management System Issues

checking for duplicate, Duplicate content checks

competitive analysis of, Content Analysis

controlling with cookies and session IDs, Controlling Content with Cookies and Session IDs

cross-linking, Cross-link relevant content

curated, Target Areas of Panda

customizing, Customizing Your Content Types to Your Audience

database-generated, Target Areas of Panda, Directories and aggregators

dealing with copyright infringement, Content Theft

deep, App SEO: Deep Linking and Indexing for Mobile Search

deleted, Uncovering subscriber-only or deleted content

delivered by geographical region, Showing Different Content to Engines and Visitors, Google Search Console

determining quality of, Content Quality and User Engagement, Content Marketing

duplicate, Duplicate Content Issues-How to Avoid Duplicate Content on Your Own Site, Showing Different Content to Engines and Visitors, Duplicate Content, Duplicate Content

expanded sources available, Engines Are Getting New Content Sources

fair use of, Fair Use?

hidden behind forms, Technology Decisions, Spiderable Link Structures,

Search and web forms, Forcing form submission, Form submission requirement

high-quality, Understanding What Makes Content High Quality, Target Areas of Panda

holiday-related, Trending, Seasonality, and Seasonal Fluctuations in Keyword Demand

invisible to search engines, What search engines cannot see

judging value of, Evaluating Content on a Web Page

keyword targeting and, Document Text, Effective Keyword Targeting by Content Creators

language/localization issues, Multiple-Language Issues

link-attracting, How Are Links Earned?, Using Content to Attract Links, Link-Worthy or Viral Content

mobile app content, App SEO: Deep Linking and Indexing for Mobile Search

moving/migrating, The Basics of Moving Content-Expectations for Content Moves

preventing copyright infringement, Copyright Infringement, Intellectual property law, Image Optimization Tips

preventing hidden, Hidden Content

pulled in via frames/iframes, What search engines cannot see, Spiderable Link Structures, Using iframes

reading level of, Content Quality and User Engagement, Reading Level

recommended number of clicks to access, Flat Versus Deep Architecture

repurposing, Repurposing Content

retiring, Mapping Content Moves

rich, Implementing Content Marketing Strategies

role in SEO, SEO Planning: Customizing Your Strategy, Business Factors That Impact Your SEO Strategy, Understanding That Content Is King

shareable, Link-Worthy or Viral Content

subscriber-only, Uncovering subscriber-only or deleted content

supplying indexable, Indexable Content

syndication of, Implementing Content Marketing Strategies, Content Syndication

theft of, Content Theft

thin, Content Uniqueness and Depth, Duplicate Content Issues, Target Areas of Panda, Thin-content penalties, Glossary

unique, Evaluating Content on a Web Page, Content Uniqueness and Depth, Duplicate Content Issues

unoriginal, Target Areas of Panda

viral, Implementing Content Marketing Strategies, Link-Worthy or Viral Content, The Indirect Influence of Social Media Marketing

visible to search engines, Understanding What Content Search Engines

Can “See” on a Web Page-What search engines cannot see

content delivery

cloaking disallowance, Cloaking and Segmenting Content Delivery

cloaking examples, Cloaking and Segmenting Content Delivery

cloaking penalties, Displaying Different Content to Search Engines Versus Visitors, Other types of manual penalties

cloaking process, Cloaking and Segmenting Content Delivery

methods of segmentation, Displaying Different Content to Search Engines Versus Visitors-Removing URLs from a search engine’s index

multivariate and A/B split testing, Showing Different Content to Engines and Visitors

of unspiderable content, Showing Different Content to Engines and Visitors

reasons for segmented, Showing Different Content to Engines and Visitors

registration-only and First Click Free, Showing Different Content to Engines and Visitors, Paywalls and Subscription Sites

content delivery networks (CDNs), Search engine–friendly navigation guidelines

content formatting, Content Structure

content management systems (CMS), Glossary

benefits of, Technology Decisions

determining need for, Content Management System Issues

drawbacks of, Technology Decisions

Flash coding best practices, Flash Coding Best Practices-Scalable Inman Flash Replacement

keyword targeting in, Keyword Targeting in Content Management Systems and Automatically Generated Content

selecting, CMS Selection

SEO considerations, Content Management System Issues-Content Management System Issues

third-party plugins/modules, Third-Party CMS Add-Ons

content marketing

additional link evaluation criteria, Additional Link Evaluation Criteria-Source TLDs

automated outreach management, Establish yourself on Google+

basics of, Understanding Content Marketing Basics

brainstorming ideas, Brainstorming Content Ideas and Being Creative

brief history of, Content Marketing

building audiences, Building an Audience-Build Offline Relationships, The Indirect Influence of Social Media Marketing

building relationships, Relationships and Outreach-Using Direct Email Pitches Effectively

campaign types, Types of Content Marketing Campaigns

case study on, Establish yourself on Google+

charity sponsorships, Integrating Emotional Triggers, Titles, and Images, Business Website Optimization for Local Search

content syndication, Content Syndication

conversion rates and, The content marketing bump

creative help, Getting Creative Help

culture of, Create a content marketing culture

customization by audience, Customizing Your Content Types to Your Audience

defined, Introduction to Content Marketing, Glossary

developing content that works, Developing Content That Works

earning links, Other Ways to Earn Links-Customer Discount/Incentives

emotional triggers and, Integrating Emotional Triggers, Titles, and Images

guest posting, Guest Posting-Perform research and analysis to pick potential targets, Implementing Guest Posting Successfully, Links Google Does Not Like

guidelines for, Introduction to Content Marketing, Integrating Emotional Triggers, Titles, and Images

high-quality content, Understanding What Makes Content High Quality

historical influence of links on rankings, How Links Historically Influenced Search Engine Rankings-The Original PageRank Algorithm

link value determination, How Search Engines Determine a Link's Value

linkbased tracking of, LinkBased Tracking of Content Marketing-Measuring the value of a link

link-worthy content, How Are Links Earned?, Using Content to Attract Links, Link-Worthy or Viral Content

measuring viewer engagement, Measuring Engagement in Content Marketing

using memes, Leveraging the Power of Memes-Intellectual property law

overview of, Conclusion

primary goal of, Content Marketing, Introduction to Content Marketing, Putting It All Together

purpose of, Content Marketing

repurposing content, Repurposing Content

social networking for links, Social Networking for Links-Putting It All Together

speedstorming, Speedstorming

strategies for, Implementing Content Marketing Strategies

strategy selection, Choosing the Right Content Marketing Strategy-Identify methods for contacting potential partners

user-generated content, User-Generated Content, Other types of manual

penalties

content optimization

content length and word count, Content length and word count

content structure, Content Structure

content themes, Content Themes

content uniqueness and depth, Content Uniqueness and Depth

CSS and semantic markup, CSS and Semantic Markup

impact of rich snippets, Impact of rich snippets

importance of, Content Optimization

visual layout, Visual layout

contextual advertising, How Google and Bing Collect Engagement Metrics, Glossary

continuing education (see industry resources)

conversion optimization (see landing page optimization)

conversion rate, Glossary

determining, Determining conversion rates

segmenting campaigns and SEO efforts by, Segmenting Campaigns and SEO Efforts by Conversion Rate

conversions

analytic software setup, Setting Up Analytics Software to Track Conversions

attribution gap management, Managing Attribution

defined, Glossary

increasing, Increasing Conversion

placing value on, Tying SEO to Conversion and ROI, Conversion tracking strategy

tracking strategy, Conversion tracking strategy

types of, Tying SEO to Conversion and ROI

cookies, Search engine–friendly navigation guidelines, Controlling Content with Cookies and Session IDs, Glossary

copyright infringement, Copyright Infringement, Fair Use?, Intellectual property law, Image Optimization Tips, Content Theft

Copyscape, Content Theft

corporate reputation management, SEO for Reputation Management, Optimizing for Image Search, Tracking the Blogosphere

cost per action (CPA), Glossary

cost per click (CPC), Understand Your Market Opportunity, Glossary

cost per thousand (CPM), Glossary

cost vs. benefit analysis, Review your website assets

cost-per-click model, The Mission of Search Engines

country-specific search engines, Country-Specific Search Engines

CPA (see cost per action)

CPC (see cost per click)

CPM (see cost per thousand)

crawl budget, Consequences of Duplicate Content

crawl delay, Syntax of the robots.txt file

crawl efficiency, What is crawl efficiency and why is it important?

crawl rate/frequency, How Search Engines Use Links, Google Search Console

crawlers, Glossary

- **alt (alternative text description) and, Understanding What Content Search Engines Can "See" on a Web Page**
- **code analyzed by, Understanding What Content Search Engines Can "See" on a Web Page**
- **code ignored by, Understanding What Content Search Engines Can "See" on a Web Page**
- **content invisible to, What search engines cannot see-What search engines cannot see, Single-Page Applications, Hidden Content**
- **content via frames/iframes and, What search engines cannot see, Spiderable Link Structures, Using iframes**
- **controlling access with robots.txt, Leveraging the robots.txt file**
- **discovering/diagnosing crawl errors, Crawl Errors-Third-party tools to check for crawl errors, Identifying the Cause of Non-Spidering, Google**

Search Console, Bing Webmaster Tools

example page analysis, Understanding What Content Search Engines Can "See" on a Web Page

HTML analysis by, Understanding What Content Search Engines Can "See" on a Web Page

image tags and, Understanding What Content Search Engines Can "See" on a Web Page

KWIC (keywords in context) and, Understanding What Content Search Engines Can "See" on a Web Page

meta tags and, Understanding What Content Search Engines Can "See" on a Web Page

noscript tags, Understanding What Content Search Engines Can "See" on a Web Page, SWFObject library and <noscript> tag

order of crawl, Crawling and Indexing

planned improvements to, Engines Will Make Crawling Improvements

Trip Advisor analysis, Understanding What Content Search Engines Can "See" on a Web Page

Creative Commons (CC), Finding Creative Commons (CC) licensed content

critical mass, achieving, Establish yourself on Google+

CRM (see customer relationship management)

cross-links, Cross-link relevant content

CSS (see cascading style sheets)

CTA (see call-to-action)

CTR (see click-through rate)

curated content, Target Areas of Panda

Curveball, An Evolving Art Form: The Future of SEO

customer discounts, Customer Discount/Incentives, Submit a special offer with a promotion feed

customer relationship management (CRM) , Establish yourself on Google+

customer retention, Optimizing User Experience to Improve SEO

customer service, The Indirect Influence of Social Media Marketing

Customers Who Bought This Item Also Bought section, Cross-link relevant content

customers, identifying, Segmenting Your Audience, Identifying Personas, and Targeting Content

(see also target audience)

D

dashboards, Using Custom Analytics Dashboards

data brokers, Psychographic targeting, Directories and aggregators

Data Highlighter tool, Local Listing Submissions, Google Knowledge Graph Carousel, Google Search Console

database-driven websites, Target Areas of Panda, Directories and aggregators, Glossary

deep linking, App Deep Linking

demographics, Segmenting Your Audience, Identifying Personas, and Targeting Content, Psychographic targeting

density (see ad density; keyword density)

digital media platforms (DMPs) , Psychographic targeting

Digital Millennium Copyright Act (see DMCA (Digital Millennium Copyright Act))

direct contact, Identify methods for contacting potential partners, Use email to contact the influencers, Using Direct Email Pitches Effectively

direct marketing, SEO for Lead Generation and Direct Marketing

directories

- **local search optimization and, Directories and aggregators**
- **to avoid, Target Areas of Penguin, Links Google Does Not Like**
- **vs. link farms, Using link farms/link networks**
- **vs. search engines, Glossary**
- **web directories, Web Directories**

disambiguation, Problem Words, Disambiguation, and Diversity

discounts, Customer Discount/Incentives, Submit a special offer with a promotion feed

discussion forums (see forums)

disruptive events, Search: Reflecting Consciousness and Connecting

Commerce

diversification, Source diversity

diversity, Problem Words, Disambiguation, and Diversity, Importance of Diversity in Rankings

DMCA (Digital Millennium Copyright Act), Summary of Major Google Algorithms, Content Theft

DMPs (see digital media platforms)

DNS (see domain name system)

document analysis

ad density, Ad Density and Offensive Ads

areas analyzed, Document Analysis and Semantic Connectivity

best applications of, Document Analysis and Semantic Connectivity

common searches used, Document Analysis and Semantic Connectivity

content quality, Content Quality and User Engagement, Document Analysis

examples of, Document Analysis and Semantic Connectivity

foundations of, Document Analysis and Semantic Connectivity

keyword stuffing, Keyword Stuffing/Lack of Synonyms

lack of synonyms, Keyword Stuffing/Lack of Synonyms

offensive ads, Ad Density and Offensive Ads

page speed, Page Speed, Mobile site speed

reading level, Content Quality and User Engagement, Reading Level

sameness, Sameness, Target Areas of Panda

semantic connectivity, Document Analysis and Semantic Connectivity

spelling/grammar, Poor Editorial Quality

document retrieval, probabilistic model of, Glossary

Domain Authority, How Search Engines Determine a Link's Value, Find out where your competitors get links, Refine the list, Establish yourself on Google+, Search engine–supplied tools, Measuring the value of a link, MozBar

domain extensions, Root Domains, Subdomains, and Microsites

Domain Insights, Setting Up Analytics Software to Track Conversions

domain name system (DNS) , Changing Servers

domain squatting, Optimizing Domains

domain-level brand metrics, Analyzing Ranking Factors

domain-level link authority, Analyzing Ranking Factors, Not enough link authority to remain in main indexes

domains

advantages of microsites, Microsites

assessing authoritative value of, Source TLDs

changing domain names, Large-Scale Content Moves, Expectations for

Content Moves

domain name optimization, Optimizing Domains

domain selection tool, Optimizing Domains

evaluation of exact-match, Summary of Major Google Algorithms

indicating preferred in Google Search Console, Google Search Console

links across, Links across domains

low-quality, Low-Quality Domains and Spam Sites

maintaining visibility during changes, Maintaining Search Engine Visibility During and After Domain Name Changes

multilanguage/country targeting, Problems with Using Your Existing Domain

other than .com, When to Use a TLD Other Than .com, Optimizing Domains

root vs. sub, Root Domains, Subdomains, and Microsites

sale of, Microsites

separate root domains, When to Use a Separate Root Domain

subdomain use, When to Use a Subdomain

subfolder use, When to Use a Subfolder, Bing Webmaster Tools

"young" domains, Unique Challenges of Domain Name Changes

doorway pages, Microsites, Thin-content penalties, Glossary

duplicate content, Duplicate content checks, Duplicate Content Issues-How to Avoid Duplicate Content on Your Own Site, Showing Different Content to Engines and Visitors, Duplicate Content, Duplicate Content, Glossary

dwell time (see time on page)

dynamic serving, Same URL approaches: responsive web design and dynamic serving

dynamic URLs, Technology Decisions

E

ecommerce

action tracking for, Taking a Deeper Look at Action Tracking

search engine-driven, How Search Engines Drive Ecommerce

SEO strategy for, SEO for Ecommerce Sales

ecommerce sites, Glossary

Editor's Picks, Editor’s Picks

editorial links, Glossary

editorial quality, Poor Editorial Quality

education (see industry resources)

embed tag, What search engines cannot see

EMD (see exact-match domain)

emotional triggers, Integrating Emotional Triggers, Titles, and Images

engagement, measuring, Measuring Engagement in Content Marketing

(see also user engagement)

entity salience, Co-occurrence, phrase-based indexing, and entity salience, Entities and Semantic Search, How to Use Schema.org

exact-match domain (EMD), Summary of Major Google Algorithms, Unique Challenges of Domain Name Changes

Exalead, Exalead

Experian Hitwise, Experian Hitwise-Cost

eXtensible Markup Language (see XML (eXtensible Markup Language))

external links, External linking, Sources of Data, External Link Attraction Analysis

eye tracking

blended search results, The History of Eye Tracking: How Users Scan Results Pages

Golden Triangle, The History of Eye Tracking: How Users Scan Results Pages

Google Instant results, The History of Eye Tracking: How Users Scan Results Pages

Google Maps results, The History of Eye Tracking: How Users Scan Results Pages

history of, The History of Eye Tracking: How Users Scan Results Pages

F

Facebook

"Login via Facebook" option, Setting Up Analytics Software to Track Conversions

additional resources for, Additional Social Media Resources

continued growth of, An Evolving Art Form: The Future of SEO

effective use of, Facebook

Facebook Insights, Facebook

generating content ideas with, Brainstorming Content Ideas and Being Creative

improving SEO through, Step 2: Send It to Your Customers/Potential Customers

increasing exposure through, Get Active in Social Media

indexing by Google, Does Google Use Facebook as a Ranking Signal?

indirect form of obtaining links, Leveraging Major Social Media Platforms

manual link creation, Manual Social Media Link Creation

number of users, The Indirect Influence of Social Media Marketing

relationship with Bing, Bing’s Experiments with Social Signals

social content amplification steps, Social content amplification steps

tracking user interaction on, Facebook

facial recognition, Image Optimization Tips

fair use, Fair Use?

Farmer update, Panda

Fast-Break, An Evolving Art Form: The Future of SEO

FCF (see First Click Free program)

Fetch and Render, Spiderable Link Structures, Other types of manual penalties

FFA (free for all) sites, Glossary

Firefox browser, viewing source code in, Understanding What Content Search Engines Can “See” on a Web Page

First Click Free program, Showing Different Content to Engines and Visitors, Paywalls and Subscription Sites

Flash, Glossary

- challenges to search engines, What search engines cannot see, Indexable Content, Third-Party CMS Add-Ons
- coding best practices, Flash Coding Best Practices-Scalable Inman Flash Replacement
- limitations of, Technology Decisions
- mobilefriendly web pages and, General SEO guidelines for mobile
- spiderable content and, Showing Different Content to Engines and Visitors

Flesch-Kincaid Grade Level Readability Formula, Content Quality and User Engagement, Reading Level

Flickr, Image Optimization Tips

fonts, selecting, Content Structure

footer links, Links Google Does Not Like

forums, Blogs and forums, Taking a Deeper Look at Action Tracking, Glossary

FQDs (see fully qualified domains)

frames, What search engines cannot see, Spiderable Link Structures, Frames, Glossary

free for all sites (see FFA (free for all) sites)

Frequently Bought Together section, Cross-link relevant content

freshness, Where freshness matters

FTC regulations, Pay for reviews

fully qualified domains (FQDs), Root Domains, Subdomains, and Microsites

fuzzy logic, Document Analysis and Semantic Connectivity

Fuzzy Set Theory, Document Analysis and Semantic Connectivity

G

gateway pages (see doorway pages)

generic top-level domains (gTLDs), New gTLDs

geolocation, Other Ranking Factors, Geolocation

geotargeting, Glossary

giveaways, hosting, Hosting giveaways

GMB (see Google My Business)

Golden Triangle

history of, The History of Eye Tracking: How Users Scan Results Pages

progression away from, The History of Eye Tracking: How Users Scan Results Pages

Goo.gl URL shortener, Content Quality and User Engagement, How Google and Bing Collect Engagement Metrics

Google

advanced search operators, Advanced Google Search Operators-Firefox plugins for quicker access to Google advanced search queries, Frequency of recent usage

analytics packages, Setting Up Analytics Software to Track Conversions

annoying add detection by, Ad Density and Offensive Ads

App Indexing API, App Indexing

dominance of, Google's Dominance-Google's Dominance

First Click Free program, Showing Different Content to Engines and Visitors, Paywalls and Subscription Sites

Hummingbird, Document Analysis and Semantic Connectivity, Understanding the Impact of Google Hummingbird, Google's Hummingbird, Summary of Major Google Algorithms

incognito search, Google+ Personalization

knowledge graph, The Knowledge Graph

links frowned upon, Links Google Does Not Like-Links Google Does Not Like

manual actions, Diagnosing the Cause of a Traffic Loss, Penalties, Manual spam actions

market share of, Country-Specific Search Engines

mission of, The Knowledge Graph, Structured Data, Overview of Changes in Search Complexity

paid link policy, Google's policy on paid links

paid links reporting form, Other Algorithmic Approaches to Fighting Link Spam

Panda algorithm, Panda-Successful Panda recovery example

Penguin algorithm, Google's Penguin Algorithm

Universal Search, Universal Search/Blended Search

use of social signals in rankings, Correlation Between Social Signals and Google Rankings, Does Google Use Facebook as a Ranking Signal?

user engagement metrics, How Google and Bing Collect Engagement Metrics

vertical search products, Vertical Search from the Major Search Engines

voice search, Document Analysis and Semantic Connectivity, Google's Hummingbird

Google AdWords Display Planner, Google AdWords Display Planner

Google AdWords Keyword Planner, Understanding Context: Market Competitiveness, What the Keyword Planner provides-Cost

Google Alerts , Google Alerts

Google Analytics, Content Quality and User Engagement, Web Analytics, How Google and Bing Collect Engagement Metrics, Diagnosing the Cause of a Traffic Loss, Traffic by search engine, Setting Up Analytics Software to Track Conversions

Google Authorship

- **Author Authority and, The Future of Author Authority at Google**
- **current state of, Is Author Authority Dead for Google?**
- **discontinuation of, Google Authorship and Author Authority**
- **Google+ authors in personalized search, Google+ Authors in Personalized Search**
- **history of, A Brief History of Google Authorship**
- **program overview, Google Authorship and Author Authority**
- **publisher tag and, Google's Publisher Tag, Optimizing for News Search: Google News**
- **reasons for ending support for, Why Did Google End Support for rel="author"?**

Google Autocomplete, Google Autocomplete

Google Blog Search, Blog search

Google bombing, Anchor text

Google Books, Book search

Google bowling, Glossary

Google Fiber, How Google and Bing Collect Engagement Metrics

Google Instant, Other Ranking Factors, Google Instant

Google Knowledge Graph Carousel, Google Knowledge Graph Carousel

Google Maps, Local search/maps

Google Merchant Center, Submitting a Product Feed

Google My Business (GMB), Google My Business

Google News

- **acceptance criteria, Acceptance Criteria**
- **application process, Application Process**
- **benefits of, Optimizing for News Search: Google News**
- **Editor's Picks feature, Editor’s Picks**
- **paywalls and subscriptions, Paywalls and Subscription Sites**
- **Publisher Center, Google News Publisher Center**
- **recrawling, Recrawling**
- **Sitemaps protocol, Google News Sitemaps**
- **technical content requirements, Technical Requirements**
- **thumbnail images, Thumbnail Images in Google News**

videos in, Videos in Google News

Google Product Search (see Google Shopping)

Google Product Taxonomy, Use the correct Google product category

Google Search Console

Crawl section data, Google Search Console

Data Highlighter tool, Google Search Console

data on log in screen, Google Search Console

defined, Glossary

detailed keyword data on, Traffic by keyword

discovering traffic loss via, Diagnosing the Cause of a Traffic Loss

examining tags with, Google Search Console

HTML Improvements report, Google Search Console

importance of using, Using Search Engine–Supplied SEO Tools

indexing data on, Site Indexing Data

manual spam action notifications, Manual spam actions

Search Appearance data, Google Search Console

Search Traffic data, Google Search Console

Setting section, Google Search Console

setup, Search Engine Tools for Webmasters

sitemap statistics, Google Search Console

verifying site ownership, Search Engine Tools for Webmasters

Google search results

compared to Bing, Analysis of Algorithmic Differentiation Across Engines and Search Types

diversity in, Problem Words, Disambiguation, and Diversity, Importance of Diversity in Rankings

document reading level in, Reading Level

freshness in, Where freshness matters

Google+ brand pages in, Google+ Brand Pages in the Search Results

Google+ posts in, Google+ Posts in the Search Results

for images, Indexable Content

nonpersonalized rankings, Google+ Impact on Nonpersonalized Rankings of Content

page layout, Understanding the Layout of Search Results Pages

providing direct answers in, The Knowledge Graph

query refinement via Orion Technology, Understanding the Layout of Search Results Pages

refinement via knowledge graph, Google's Knowledge Graph, Google's Knowledge Graph and the Knowledge Vault-Fair Use?, Google Knowledge Graph Carousel

vertical result examples, Understanding How Vertical Results Fit into the SERPs

Google Shopping

importance of optimizing for, Optimizing for Google Shopping Search

main steps, Optimizing for Google Shopping Search

product feed optimization, Optimizing a Product Feed

product feed submission, Submitting a Product Feed

product promotion in AdWords, Promoting Products in AdWords

Google Suggest (see Google Autocomplete)

Google supplemental index, The Importance of Experience, Glossary

Google Toolbar, Content Quality and User Engagement, How Google and Bing Collect Engagement Metrics

Google Trends, Google Trends-Cost, Google Trends for Video

Google Trusted Store, Become a Google Trusted Store

Google Video, Video SEO for Google

Google Webmaster Tools (see Google Search Console)

Google+

building relationships with influencers on, Build an initial list

correlation of +1s with rankings, Evaluating Social Media Signals, Content Marketing, How Might Google Use Google+ as a Ranking Factor?

effective use of, Google+

establishing yourself on, Establish yourself on Google+

generating content ideas with, Brainstorming Content Ideas and Being Creative

Google+ Hangouts, Creating a Value Proposition for a Relationship

increasing exposure through, Get Active in Social Media

indirect form of obtaining links, Leveraging Major Social Media Platforms

manual link creation, Manual Social Media Link Creation

number of users, The Indirect Influence of Social Media Marketing

repurposing content for, Repurposing Content, Create templates

social content amplification steps, Social content amplification steps

tracking user interaction on, Google+

use as a ranking signal, Does Google Use Google+ as a Ranking Signal?-How Might Google Use Google+ as a Ranking Factor?

using as a warmup, Use Twitter as a warmup

Googlebot, Glossary

feature phones vs. smartphones, Mobile user agents

robots.txt instructions, Syntax of the robots.txt file

grammatical errors, Content Quality and User Engagement, Poor Editorial Quality

graph theory, Using link farms/link networks, What Is the Value of Social Signals?

gray hat SEO, Strategies that are not considered buying links, Glossary

gTLDs (see generic top-level domains)

guest posting, Links across domains, Implementing Content Marketing Strategies, Guest Posting-Perform research and analysis to pick potential targets, Get to Know Other People's Audiences, Implementing Guest Posting Successfully, Links Google Does Not Like

H

hash mark (#), Single-Page Applications

hashbang (#!), Single-Page Applications

hashtag memes, Other types of memes

heading tags, Heading Tags, Glossary

hidden text, Other types of manual penalties, Glossary

high-value pages, Step 2: Site architecture

hijacking, Glossary

holiday-related content, Trending, Seasonality, and Seasonal Fluctuations in Keyword Demand

horizontal navigation, Understanding the Layout of Search Results Pages

hreflang, hreflang for multiple languages/no specific location

HTML (HyperText Markup Language), Glossary

HTML 5, What search engines cannot see, Content Management System Issues

HTML sitemap, Using the meta robots tag, Glossary

HTML status codes

301 (permanent redirect), Search engine–friendly navigation guidelines, Using the rel="canonical" link element, Good and Bad Redirects, Content Management System Issues, The Basics of Moving Content

302 (redirect), Technology Decisions, Search engine–friendly navigation guidelines, Good and Bad Redirects

404 (page not found), Redirects

diagnostic reports, Tools from the search engines

overview of, Glossary-Glossary

HTML tags, Content Management System Issues

HTTP (HyperText Transfer Protocol), Glossary

HTTPS (HyperText Transfer Protocol Secure), Glossary

hubs, Authority, Glossary

Hummingbird, Document Analysis and Semantic Connectivity, Understanding the Impact of Google Hummingbird, Google's Hummingbird, Summary of Major Google Algorithms

HyperText Transfer Protocol (see HTTP)

HyperText Transfer Protocol Secure (see HTTPS)

hyphens, Optimizing Domains, Picking the Right URLs

hypotheses, forming, Start with a hypothesis

I

IA (see information architecture)

Identifier Exists attribute, Use correct product identifiers

ideological influence, SEO for Ideological Influence

iframes, What search engines cannot see, Spiderable Link Structures, Frames, Using iframes

image recognition, What search engines cannot see, Java, images, audio, and video, Optimizing for Image Search

image search

- drag and drop feature, Indexable Content
- keyword-rich filenames, Image Filenames and alt Attributes
- optimizing for, Optimizing for Image Search
- search engine support for, Image search

image tags

- defined, Understanding What Content Search Engines Can “See” on a Web Page
- indexability of, Indexable Content
- sitemaps and, Image sitemaps

images

attracting viewers with, Integrating Emotional Triggers, Titles, and Images

of celebrities, Business Website Optimization for Local Search

purchasing, Image Optimization Tips

sharing, Image Optimization Tips, Thumbnail Images in Google News

tips for optimization of, Image Optimization Tips

impressions, Glossary

inbound links, Glossary

incognito searches, Other Ranking Factors, Google+ Personalization

indexed pages, Glossary

indexing, Glossary

benchmarking current status, Benchmarking Current Indexing Status, Site Indexing Data

of Facebook pages, Does Google Use Facebook as a Ranking Signal?

first step, Crawling and Indexing

of images, Image Optimization Tips

link value and, How Search Engines Use Links

measuring effective, Number of pages getting search traffic

of mobile apps, App Indexing

phrase-based, Co-occurrence, phrase-based indexing, and entity salience

real-time, Where freshness matters

removing URLs from indexes, Removing URLs from a search engine's index

of structured data, Structured Data

supplying indexable content, Indexable Content

of Twitter posts, Does Google Use Twitter as a Ranking Signal?

indirect contact, Identify methods for contacting potential partners

industry resources

analyzing algorithmic differentiation, Analysis of Algorithmic Differentiation Across Engines and Search Types

analyzing top-ranking sites/pages, Analysis of Top-Ranking Sites and Pages

competitive analysis, Competitive Analysis-Competitive Analysis for Those with a Big Budget

conferences and organizations, Participation in Conferences and Organizations

employee communications, Commentary from search engine employees

forums and blogs, Commentary from search engine employees, The SEO Industry on the Web

metrics commonly analyzed, Other useful SEO metrics

overview of, SEO Education and Research

proxying process, Other useful SEO metrics

search-engine supplied tools, Using Search Engine–Supplied SEO Tools-Bing Webmaster Tools

SEO testing and experiments, SEO Testing

social network communities, Communities in Social Networks

websites and periodicals, Websites

influencers (online)

building relationships with, Building the relationship, Building Relationships with Influencers-Establish yourself on Google+

contacting Pinterest influencers, Leverage Pinterest group boards

contacting via email, Use email to contact the influencers, Contact the influencers, Using Direct Email Pitches Effectively

defined, Leverage Influencers and Influencer Marketing

determining impact of, Refine the list

encouraging content sharing by, Building the relationship

engaging, Engaging the influencer

identifying, Tracking the Blogosphere

influencer marketing, Leverage Influencers and Influencer Marketing

paid product reviews, Pay for reviews

power of, Leverage Influencers and Influencer Marketing

weight in Google indexings, Does Google Use Facebook as a Ranking Signal?

infographics, Hosting giveaways, Links Google Does Not Like

information architecture (IA)

categorically structured information, Category structuring, Content Management System Issues

competitive analysis of, Internal Link Structure and Site Architecture

creating optimal, Creating an Optimal Information Architecture

design principles of, Site Architecture Design Principles

filing cabinet analogy, An analogy

flat vs. deep architecture, Flat Versus Deep Architecture

logical flow, The Importance of a Logical, Category-Based Flow

obtaining current document, Structural Decisions, Mapping Content Moves

redirects and, Redirects

search-friendly site navigation, Search-Friendly Site Navigation-Search engine–friendly navigation guidelines

site architecture protocol, Site architecture protocol

static URLs, URLs, Picking the Right URLs

subdomains, Subdomains

taxonomy/ontology and, Taxonomy and ontology

- topically relevant information, Topical relevance
- usability/search friendliness, Usability and search friendliness

information retrieval (IR), Determining Searcher Intent and Delivering Relevant, Fresh Content, Glossary

informational queries, Informational Queries, Understanding Search Engine Traffic and Visitor Intent

inheritance, How to Use Schema.org

Instagram

- effective use of, Instagram, Vine, and Snapchat
- increasing exposure through, Get Active in Social Media
- number of users, The Indirect Influence of Social Media Marketing
- tracking metrics for, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

instant answers (see vertical results)

intellectual property, Intellectual property law, Content Theft

intent (see search intent; user intent)

interlinking, Glossary

internal links, Glossary

- anchor text, Use anchor text, intuitively, Content Management System Issues, Content Management System Issues, Anchor text, Target Areas of Penguin, Links Google Does Not Like, Optimizing Your Anchor Text, LinkBased Tracking of Content Marketing, External Link Attraction

Analysis

audit review of, Internal linking checks

best practices for, Structural Decisions, Spiderable Link Structures

breadcrumb navigation, Use breadcrumb navigation, Content Management System Issues

competitive analysis of, Internal Link Structure and Site Architecture

correcting problems with, Example: Fixing an Internal Linking Problem

link depth, Minimize link depth

relevant cross-links, Cross-link relevant content

target keywords, Target keywords

Internet protocol (IP), Glossary

Internet service providers (ISP), Glossary

inventory, Glossary

invisible GIFs, Glossary

invisible Web, Glossary

IP (Internet protocol), Glossary

IP addresses, Server and Hosting Issues, Search engine–friendly navigation guidelines, Blocking and cloaking by IP address range, Changing Servers, Glossary

IR (see information retrieval)

ISAPI_Rewrite module, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting-mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting, Large-Scale Content Moves, Glossary

ISP (see Internet service providers)

items, How to Use Schema.org

J

Java applets, Spiderable Link Structures, Java, images, audio, and video, Showing Different Content to Engines and Visitors, Hiding text in Java applets, Glossary

JavaScript, Glossary

browser support for, Spiderable Link Structures

users disallowing, Understanding What Content Search Engines Can "See" on a Web Page

visibility to search engines, AJAX and JavaScript-AJAX and JavaScript, Showing Different Content to Engines and Visitors

web analytics packages, Selecting the Right Analytics Package

JSON-LD, Step 3: Choose your implementation technique

K

key performance indicators (KPIs), Social distribution, Why Measuring Success Is Essential to the SEO Process, Separating the Analytics Wheat from the Chaff, Glossary

keyword cannibalization, Keyword Cannibalization, Keyword Cannibalization

keyword density, KeywordDiscovery

keyword research

ad campaign runs/third-party data, Ad Campaign Runs and Third-Party Search Data

Bing Webmaster Tool for, Bing Webmaster Tools

(see also keyword targeting)

brainstorming original ideas, Traditional Approaches: Domain Expertise and Site Content Analysis

competitive analysis, Including Competitive Analysis

in content marketing, Brainstorming Content Ideas and Being Creative

data analysis, Keyword Research Data Analysis

data from search engines, Keyword Research Data from Search Engines-Frequency of recent usage

data from tools, Keyword Research Data from Tools-Visuwords

defined, Glossary

impact of Google Hummingbird, Understanding the Impact of Google Hummingbird

impact of keyword "not provided", Understanding Keyword “Not Provided” and Co-Occurrence Analysis, Traffic by keyword

impact of longtail search queries, Understanding the Long Tail of the Keyword Demand Curve

landing page optimization, Landing Page Optimization

leveraging longtail search queries, Leveraging the Long Tail of Keyword Demand-Applying User-Generated Content Strategies for LongTail Targeting

in market opportunity analysis, Understand Your Market Opportunity

options for, Keyword Research Options

searcher behavior and, Thinking Strategically

strategy for, Integrating Keyword Research, Co-Occurrence Analysis, and Knowledge of User Intent

structural decisions and, Target keywords

theory behind, The Theory Behind Keyword Research

trending and seasonality, Trending, Seasonality, and Seasonal Fluctuations in Keyword Demand

keyword review

importance of, The Importance of Keyword Reviews

keyword mapping, Step 3: Keyword mapping

site architecture and, Step 2: Site architecture

site reviews, Step 4: Site review

keyword targeting

boldface text, Boldface Text

by content creators, Effective Keyword Targeting by Content Creators

in content management systems, Keyword Targeting in Content

Management Systems and Automatically Generated Content

document text, Document Text

heading tags, Heading Tags

HTML title tags, HTML <title> Tags

image filenames and alt attributes, Image Filenames and alt Attributes

importance of, Keyword Targeting

in automatically generated content, Keyword Targeting in Content Management Systems and Automatically Generated Content

initial steps, Keyword Targeting

keyword cannibalization, Keyword Cannibalization

longtail search queries, LongTail Keyword Targeting

measuring impact of, Establishing a Proper Baseline

meta description tags, Meta Description Tags

(see also keyword research)

KeywordDiscovery, KeywordDiscovery-Cost

keywords, Glossary

audits and, Keyword health checks

branded, Visibility (Branding)

charts, Searcher Intent

determining ranking difficulty, Determining keyword difficulty,

Understanding Context: Market Competitiveness, Determining conversion rates

determining value of, Determining keyword value

domain names and, Optimizing Domains

during search process, How People Search

in breadcrumb navigation, Use breadcrumb navigation

in URLs, Picking the Right URLs

including/excluding in search, Using Advanced Search Techniques

keyword cannibalization, Keyword Cannibalization, Keyword Cannibalization, Glossary

keyword density, Glossary

keyword popularity, Glossary

keyword prominence, Glossary

keyword stuffing, Keyword Targeting, Keyword Stuffing/Lack of Synonyms, Other types of manual penalties, Glossary

keywords in context, Understanding What Content Search Engines Can “See” on a Web Page

measuring traffic by, Traffic by keyword

meta keywords, Understanding What Content Search Engines Can “See” on a Web Page, Glossary

as ranking factors, Analyzing Ranking Factors

target, Target keywords

theme, Evaluating Content on a Web Page, Content Themes, Content Management System Issues, Glossary

triggering Google Knowledge Graph Carousel, Google Knowledge Graph Carousel

KeywordSpy, KeywordSpy

Klout, Refine the list, Bing's Experiments with Social Signals

Klout Score, Refine the list

knowledge graph

changes in search complexity, Google's Knowledge Graph and the Knowledge Vault

concept of fair use and, Fair Use?

future of, The Future of Semantic Search and the Knowledge Graph

knowledge extraction and, Overview of Changes in Search Complexity

local search optimization and, Google Knowledge Graph Carousel

original databases used for, The Knowledge Graph

providing direct answers via, The Knowledge Graph

relationship mapping via, Google's Knowledge Graph

search result example, The Knowledge Graph, Overview of Changes in Search Complexity

vs. vertical results, The Knowledge Graph

Knowledge Vault

community information resources, How the Knowledge Vault Works

concept of, How the Knowledge Vault Works

future of, The Future of the Knowledge Vault, The Future of Semantic Search and the Knowledge Graph

machine reading process, How the Knowledge Vault Works

web-based questions and answers, How the Knowledge Vault Works

KPIs (see key performance indicators)

Kred, Refine the list

KWIC (keywords in context), Understanding What Content Search Engines Can "See" on a Web Page

L

landing page optimization, Landing Page Optimization, Optimizing User Experience to Improve SEO

language preferences, Other Ranking Factors

latent semantic indexing (LSI), Glossary

layout, Visual layout

lead generation, SEO for Lead Generation and Direct Marketing

link analysis, Link Analysis, Analyzing Ranking Factors, Authority, Execute aggressively, Using Competitive Link Analysis Tools

link attraction, Link-Worthy or Viral Content

link bait, Link-Worthy or Viral Content, Glossary

link brokers, Other Algorithmic Approaches to Fighting Link Spam

link building, The Indirect Influence of Social Media Marketing, Glossary

link depth, Minimize link depth

link exchange, Glossary

link farms, Using link farms/link networks, Glossary

link graph, What Is the Value of Social Signals?

link juice, Glossary

link neighborhoods, Link Analysis, Identifying neighborhoods, Using link farms/link networks

link networks, Using link farms/link networks

link popularity, Glossary

link spam, How Search Engines Fight Link Spam-Other Search Engine Courses of Action, Glossary

link text, Glossary

link tracking

- **Ahrefs Site Explorer, Ahrefs**
- **Conductor Searchlight, Conductor Searchlight**
- **Exalead, Exalead**
- **LinkBuilder/LinkGalaxy, LinkBuilder and LinkGalaxy**

LinkResearchTools.com, LinkResearchTools.com

Majestic, Majestic

measuring link value, Measuring the value of a link

Open Site Explorer, Open Site Explorer

purpose of, LinkBased Tracking of Content Marketing

Raven Link Manager, Raven Tools

search engine-supplied tools, Search engine–supplied tools

LinkBuilder/LinkGalaxy, LinkBuilder and LinkGalaxy

LinkedIn

generating content ideas with, Brainstorming Content Ideas and Being Creative

improving SEO through, Step 2: Send It to Your Customers/Potential Customers

indirect form of obtaining links, Leveraging Major Social Media Platforms

leveraging for B2B sites, LinkedIn, LinkedIn

LinkedInPublishing, Repurposing Content

manual link creation, Manual Social Media Link Creation

number of users, The Indirect Influence of Social Media Marketing

repurposing content from, Repurposing Content, Create templates

social content amplification steps, Social content amplification steps

tracking user interaction on, LinkedIn

linkerati, Glossary

LinkResearchTools.com, External linking, Source diversity, Find out where your competitors get links, Execute aggressively, Refine the list, Blogs and forums, Sources of Data, LinkResearchTools.com, Analysis of Top-Ranking Sites and Pages, External Link Attraction Analysis

links

across domains, Links across domains

additional link evaluation criteria, Additional Link Evaluation Criteria-Source TLDs

app deep linking, App Deep Linking

audit review of external, External linking

audit review of internal, Internal linking checks

automated link dropping, Using automated link dropping

on blog platforms, Links

categorizing potential value of, Placing a Value on the Sites, Measuring the value of a link

cleanup and removal process, Link Cleanup Process-Expected reconsideration request timeline

competitive analysis of, Internal Link Structure and Site Architecture, Using Competitive Link Analysis Tools

context/relevance of, Context/relevance, Target Areas of Penguin, Links Google Does Not Like

costs to pursue, Review your website assets

cross-links, Cross-link relevant content

determining most powerful, Find out where your competitors get links

devious link schemes, Identifying Spam, Spam Filtering and Penalties

disavowing, Negative Link Building, The Disavow Links tool

earning, Other Ways to Earn Links-Customer Discount/Incentives

evaluation by search engines, How Search Engines Determine a Link's Value, Penguin

external, External linking, Sources of Data, External Link Attraction Analysis

factors affecting ranking value, Additional Factors That Influence Link Value

footer, Links Google Does Not Like

gained through social advertising, Social advertising can drive links to your site

historical influence on rankings, How Links Historically Influenced Search Engine Rankings-The Original PageRank Algorithm

inbound, Glossary

incentive-based, Hosting giveaways, External Link Attraction Analysis

interlinking, Glossary

internal, Internal linking checks, Example: Fixing an Internal Linking Problem, Internal Link Structure and Site Architecture, Glossary

link-building vs. quality content, Content Marketing, Anchor text, Other Ways to Earn Links, External Link Attraction Analysis

maximum number of, Spiderable Link Structures, Flat Versus Deep Architecture

negative, Negative Link Building

nofollow attribute, Spiderable Link Structures, Using the rel="nofollow" attribute-Using the rel="nofollow" attribute, Content Management System Issues, No direct links, Glossary

nonreciprocal, Glossary

outbound, Glossary

paid, Negative Ranking Factors, The Original PageRank Algorithm, Get to Know Other People's Audiences, Paid social media, Pay for reviews, Purchasing links for SEO, Other Algorithmic Approaches to Fighting Link Spam

partial, Partial link penalties

rate of acquisition links, Other Ranking Factors

reciprocal, How Are Links Earned?, Swapping links, Glossary

relative, Glossary

sitewide, Links across domains, Other Algorithmic Approaches to

Fighting Link Spam, Sitewide link penalties

social networking for, Social Networking for Links-Putting It All Together

spiderable, Spiderable Link Structures

swapping, Swapping links

temporal factors of, Temporal factors

temporal link growth measurements, Temporal Link Growth Measurements

text links, Glossary

text-based navigational structure, Search engine–friendly navigation guidelines, Content Management System Issues

that Google frowns upon, Links Google Does Not Like-Links Google Does Not Like

three-way swaps, Using link farms/link networks

tracking blog links, Blog links

unnatural, Unnatural Links Messages, Other types of manual penalties, Spam Filtering and Penalties

use of by search engines, How Search Engines Use Links

ways to earn, How Are Links Earned?

listing age, Determining listing age

local search

Bing Places for Business, Bing Places for Business

business website optimization for, Business Website Optimization for Local Search

directories and aggregators, Directories and aggregators

Google Knowledge Graph Carousel, Google Knowledge Graph Carousel

Google My Business, Google My Business

impact of presence in, Local search/maps

increasing importance of, Local Search

local listing submissions, Local Listing Submissions

opportunities afforded by, Optimizing for Local Search

optimizing for, Optimizing for Local Search

Yahoo! Local, Yahoo! Local

localization issues, Multiple-Language Issues

locations, multiple, Business Website Optimization for Local Search

logfiles, Glossary

tracking, Log file Tracking, Selecting the Right Analytics Package

logins, Using login/password protection

longtail search queries, Understanding the Long Tail of the Keyword Demand Curve, Leveraging the Long Tail of Keyword Demand-Applying User-Generated Content Strategies for LongTail Targeting, LongTail Keyword Targeting, Key Performance Indicators for LongTail SEO,

Glossary

looping, avoiding in redirects, How to Redirect a Home Page Index File Without Looping

LSI (see latent semantic indexing)

M

machine reading, How the Knowledge Vault Works

Made for AdSense (MFA), Glossary

Majestic, Majestic

malware, Negative Ranking Factors, Links Google Does Not Like

manual actions, Diagnosing the Cause of a Traffic Loss, Penalties, Manual spam actions

map search, Local search/maps

market opportunity

assessing, Understand Your Market Opportunity

understanding, Understand Your Market Opportunity

marketing, direct, SEO for Lead Generation and Direct Marketing

markup

hreflang, hreflang for multiple languages/no specific location

Schema.org, CSS and Semantic Markup-Summary

in search results, Markup in search results

semantic, CSS and Semantic Markup, Schema.org

for structured data, Structured Data

supported types of, Supported types of markup

measuring success (see tracking and measurement systems)

media contact databases, Establish yourself on Google+

Meetup.com, Meetup.com

memes

benefits of, Leveraging the Power of Memes

defined, Leveraging the Power of Memes

examples of, Leveraging the Power of Memes, Meme marketing in action

generating tools for, Integrating Emotional Triggers, Titles, and Images

intellectual property laws and, Intellectual property law

in mainstream advertising, Memes in advertising

in marketing campaigns, Memes in your marketing campaigns

non-image memes, Other types of memes

purpose of, Meme marketing in action

targeted to audiences, Memes targeted to your audience

meta description, Glossary

audit review of, Meta tag review

customization of, Content Management System Issues

examining tags in Google Search Console, Google Search Console

keyword targeting and, Meta Description Tags

optimizing for Google Shopping, Optimize the description

ranking systems and, Understanding What Content Search Engines Can "See" on a Web Page

semantic search and, Structured Data

meta keywords, Understanding What Content Search Engines Can "See" on a Web Page, Glossary

meta robots tag, Spiderable Link Structures, Using the meta robots tag-Using the meta robots tag, Blocked by the robots meta tag

MFA sites (see Made for AdSense)

microsites

advantages of, Microsites

drawbacks of, Microsites

vs. root and sub domains, Root Domains, Subdomains, and Microsites

when to use, Microsites

mindshare, SEO for Mindshare and Branding, Microsites

misspelled search terms, Document Analysis and Semantic Connectivity, KeywordDiscovery

mobile advertising spend, shift toward, The Mobile Shift

mobile app content

- **app deep linking, App Deep Linking**
- **app indexing, App Indexing**
- **app-specific search, App SEO: Deep Linking and Indexing for Mobile Search**
- **opportunities afforded by, App SEO: Deep Linking and Indexing for Mobile Search**

mobile subdomain, Same URL approaches: responsive web design and dynamic serving

mobilefriendly web pages, Glossary

- **additional resources for, Additional mobile SEO resources**
- **guidelines for, General SEO guidelines for mobile**
- **increasing importance of, Mobile Search**
- **mobile landscape, The Mobile Landscape**
- **mobile site speed, Mobile site speed**
- **mobile user agents, Mobile user agents**
- **phone operating systems, How Google and Bing Collect Engagement Metrics**
- **same URL approach to, Same URL approaches: responsive web design and dynamic serving**
- **SEO for mobile, SEO for Mobile**

separate URL approach to, Separate URL approach

sitemaps for, Mobile sitemaps, Mobile Friendliness

smartphone effect on ad spending, The Mobile Landscape

MobileDetect.net, Mobile user agents

Mobilegeddon, Glossary

mod_rewrite module, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting-mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting, Large-Scale Content Moves, Glossary

monetization, Glossary

“moving from strings to things”, Overview of Changes in Search Complexity

Moz

action tracking by, Taking a Deeper Look at Action Tracking

Crawl Test tool, Third-party tools to check for crawl errors

glossary provided by, Optimizing for Local Search

history of algorithm updates, Establishing a Proper Baseline

Keyword Not Provided and, Number of pages getting search traffic

Mega-SERP infographic, Universal Search = Blended Search

metrics used by, Measuring the value of a link

MozRank, Search engine–supplied tools

Open Site Explorer tool, Action tracking by referral source, Open Site Explorer, Conductor Searchlight, Measuring the value of a link, Analysis of Top-Ranking Sites and Pages, Using Competitive Link Analysis Tools

tools and resources provided by, Websites, SEO News Outlets, Communities, and Forums, Making the Most of Limited Resources or Budgets, Making the Most of Limited Resources or Budgets, Selecting an SEO Firm/Consultant

use of Google Analytics by, Segmenting Campaigns and SEO Efforts by Conversion Rate

MozBar, MozBar

multilanguage/country targeting

challenges of, Best Practices for Multilanguage/Country Targeting

major approaches to, The Two Major Approaches

multiple-language issues, Multiple-Language Issues

targeting specific countries, How to Target a Specific Country

using existing domain, Problems with Using Your Existing Domain

multimedia, indexability of, Multimedia Is Becoming Indexable

(see also image search; video search)

multivariate testing, Landing Page Optimization, Showing Different Content to Engines and Visitors

N

NAP (name, address, and phone number), Local Listing Submissions

natural language, Google's Hummingbird

natural search results (see organic search results)

Naver, Country-Specific Search Engines

navigational elements

AJAX and JavaScript, AJAX and JavaScript-AJAX and JavaScript

best practices for, Search engine–friendly navigation guidelines

breadcrumb, Use breadcrumb navigation

frames/iframes, Frames, Using iframes

non-text files, Java, images, audio, and video

role of, Evaluating Content on a Web Page

search and web forms, Search and web forms, Forcing form submission

search-friendly navigation, Search-Friendly Site Navigation

navigational queries, Navigational Queries, Understanding Search Engine Traffic and Visitor Intent

negative link building, Negative Link Building

neighborhoods (see link neighborhoods)

news search, News search

nofollow, Spiderable Link Structures, Using the rel="nofollow" attribute-Using the rel="nofollow" attribute, Content Management System Issues, No direct links, Glossary

noindex, **Syntax of the robots.txt file**, **Using the meta robots tag**, **Content Management System Issues**, **Content Syndication**, **Identifying Content That Search Engines Don't See**, **Glossary**

nonbranded search traffic, **Branded versus nonbranded traffic**, **Understand Your Market Opportunity**

nonreciprocal links, **Glossary**

noscript tags, **Understanding What Content Search Engines Can "See" on a Web Page**, **SWFObject library and <noscript> tag**

numbers, **Optimizing Domains**

O

OCR (see optical character recognition)

ontology, **Taxonomy and ontology**, **How to Use Schema.org**

OPA (see other people's audiences)

Open Site Explorer, **Open Site Explorer**, **Using Competitive Link Analysis Tools**

opinion influencing, **SEO for Ideological Influence**

optical character recognition (OCR), **What search engines cannot see**, **Image Optimization Tips**

organic search results, **Glossary**

- **benchmarking**, **Benchmarking Organic Rankings**
- **impact of drop in**, **Panda, Penguin, and Penalties**
- **importance of top spots**, **Click Tracking: How Users Click on Results**

(Organic Versus Paid)

SERP section for, Understanding the Layout of Search Results Pages

understanding your audience and, Determining Searcher Intent: A Challenge for Search Marketers and Search Engines

Orion Technology, Understanding the Layout of Search Results Pages

other people's audiences, Get to Know Other People's Audiences-The indirect benefits of building reputation and visibility

outbound links, Glossary

outgoing links (see outbound links)

outlinks (see outbound links)

outreach (see relationships)

P

page load time, Page load time, Page Speed, Mobile site speed, Glossary

page post ads, Social content amplification steps

page segmentation, Page segmentation

page titles, Understanding What Content Search Engines Can "See" on a Web Page

page views per visitor analytic, Content Quality and User Engagement, Potential User Engagement Signals

page-level ranking factors, Analyzing Ranking Factors

pagejacking, Glossary

PageRank (PR), Glossary

external links and, Internal linking checks

fundamental flaw in, The Original PageRank Algorithm

implementation of, How Links Historically Influenced Search Engine Rankings

original algorithm, The Original PageRank Algorithm-The Original PageRank Algorithm

refinements to, The Original PageRank Algorithm

pages printed analytic, Potential User Engagement Signals

pagination, Flat Versus Deep Architecture, Content Management System Issues

paid links, Negative Ranking Factors, The Original PageRank Algorithm, Get to Know Other People's Audiences, Paid social media, Pay for reviews, Purchasing links for SEO, Other Algorithmic Approaches to Fighting Link Spam

Panda algorithm

Chrome Blocklist Extension and, Panda

curated content target, Target Areas of Panda

database-generated content target, Target Areas of Panda

diversity target, Importance of Diversity in Rankings

impact of weak content, Impact of Any Weak Content on Rankings

measuring impact of, Establishing a Proper Baseline

measuring user engagement with, Content Quality and User Engagement

nondifferentiated content target, Target Areas of Panda

penalty recovery, Path to Recovery

poor-quality content target, Target Areas of Panda

recovery example, Successful Panda recovery example

recovery timeline, Expected timeline to recovery

release history, Panda

release of, Panda

role of authority in, Role of Authority in Rankings

SEO aspects of, Panda

target areas of, Target Areas of Panda

thin content target, Target Areas of Panda

thin slicing target, Target Areas of Panda

unoriginal content target, Target Areas of Panda

Panguin tool, Diagnosing the Cause of a Traffic Loss, Establishing a Proper Baseline

passwords, Using login/password protection

pattern matching, Syntax of the robots.txt file, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting

pay-level domains (PLDs), Root Domains, Subdomains, and Microsites

pay-per-click model, The Mission of Search Engines, Understanding the Layout of Search Results Pages

payday loan algorithm, Summary of Major Google Algorithms

paywalls, Paywalls and Subscription Sites

PDF (Portable Document Format), Spiderable Link Structures, Glossary

penalties

- **algorithmic vs. manual, Penalties, Manual spam actions**
- **for cloaking/redirects, Other types of manual penalties**
- **for hacked sites, Other types of manual penalties**
- **for hidden text/keyword stuffing, Other types of manual penalties**
- **for partial links, Partial link penalties**
- **for sitewide links, Sitewide link penalties**
- **for spam, Other types of manual penalties**
- **for thin-content, Thin-content penalties**
- **for unnatural links, Other types of manual penalties, Spam Filtering and Penalties**
- **for user-generated spam, Other types of manual penalties**
- **penalty messages, Diagnosing the Cause of a Traffic Loss, Penalties**
- **reconsideration requests, Path to Recovery, Search Engine Penalties and Reconsideration Requests**

recovering from Panda penalty, Path to Recovery

recovering from Penguin penalty, Path to Recovery

recovery example, Successful Panda recovery example

recovery timeline, Expected timeline to recovery, Penguin

types of manual, Types of Manual Penalties

Penguin algorithm

focus of, Google's Penguin Algorithm

measuring impact of, Establishing a Proper Baseline

penalty recovery, Path to Recovery

release history, Penguin

target areas of, Target Areas of Penguin

Personal Blocklist Extension, Content Quality and User Engagement, Voting Mechanisms, Panda

personalization, Other Ranking Factors, More Personalized, Localized, and User-Influenced Search

personas, identifying, Segmenting Your Audience, Identifying Personas, and Targeting Content

phone operating systems, How Google and Bing Collect Engagement Metrics

PHP programming language, Glossary

phrase-based indexing, Co-occurrence, phrase-based indexing, and entity

salience

physical memes, Other types of memes

pingbacks, Content Management System Issues

Pinterest

effective use of, Pinterest

increasing exposure through, Get Active in Social Media

indirect form of obtaining links, Leveraging Major Social Media Platforms

number of users, The Indirect Influence of Social Media Marketing

Pitchbox, Establish yourself on Google+

Piwik, Web Analytics

Plagium, Content Theft

playlists, Create and optimize playlists

PLDs (see pay-level domains)

Portable Document Format (see PDF (Portable Document Format))

post slugs, Content Management System Issues

potential partners, contacting, Identify methods for contacting potential partners

PowerPoint files, Spiderable Link Structures

PPC (see pay-per-click model)

PR (see PageRank (PR))

press releases, Get to Know Other People's Audiences, Links Google Does Not Like

probabilistic model, Glossary

product feeds

optimizing, Optimizing a Product Feed

submitting, Submitting a Product Feed

product reviews, paid, Pay for reviews

product search, Product search

product taxonomy, Use the correct Google product category

products

images of, Optimizing for Image Search, Use appealing images

mapping of, Mapping Your Products and Services

tracking sales of, Setting Up Analytics Software to Track Conversions

profiles, claiming, Claiming Key Profiles, Local Listing Submissions

progress, assessing, Assessing Historical Progress

(see also tracking and measurement systems)

promoted tweets/posts, Social content amplification steps

properties, How to Use Schema.org

proximity searches, Document Analysis and Semantic Connectivity,

Business Website Optimization for Local Search

psychographic targeting, Psychographic targeting

public domain, Image Optimization Tips

public opinion, SEO for Ideological Influence

publicity, SEO for Reputation Management

publisher tag, Google's Publisher Tag, Google+ Brand Pages in the Search Results, Optimizing for News Search: Google News

pull-down lists, Technology Decisions, Glossary

Q

QDD (see query deserves diversity)

QDF (see query deserves freshness)

Quantcast, Quantcast

queries, Glossary

ambiguous, Problem Words, Disambiguation, and Diversity

depersonalizing, Other Ranking Factors

informational, Informational Queries, Understanding Search Engine Traffic and Visitor Intent

length of, The Human Goals of Searching

longtail search queries, Understanding the Long Tail of the Keyword Demand Curve, Leveraging the Long Tail of Keyword Demand-Applying User-Generated Content Strategies for LongTail Targeting, Key

Performance Indicators for LongTail SEO, Glossary

multistep process of, How People Search

navigational, Navigational Queries, Understanding Search Engine Traffic and Visitor Intent

number performed, Who Searches and What Do They Search For?

refinement suggestions on SERPs, Understanding the Layout of Search Results Pages

term weighting in, Document Analysis and Semantic Connectivity

transactional, Transactional Queries, Understanding Search Engine Traffic and Visitor Intent

triggering a news box, Optimizing for News Search: Google News

query deserves diversity (QDD), Problem Words, Disambiguation, and Diversity, Importance of Diversity in Rankings

query deserves freshness (QDF), Where freshness matters

Quora, Claiming Key Profiles

R

ranking systems

applying rank data to SEO, Ranking

benchmarking organic rankings, Benchmarking Organic Rankings

crawling and indexing, Crawling and Indexing

data tracking, AuthorityLabs

determination of rank order, Retrieval and Ranking

determining keyword ranking difficulty, Determining keyword difficulty, Determining conversion rates

development/tweaking of, How Social Media and User Data Play a Role in Search Results and Rankings

for trustworthiness, Become a Google Trusted Store

historical influence of links on, How Links Historically Influenced Search Engine Rankings-The Original PageRank Algorithm

impact of social media on, Content Marketing

importance of understanding, Algorithm-Based Ranking Systems: Crawling, Indexing, and Ranking

increased weight on social signals, How Social Media and User Data Play a Role in Search Results and Rankings

international considerations, Optimizing for Specific Countries, Best Practices for Multilanguage/Country Targeting, Links Google Does Not Like, Increased Market Saturation and Competition

key to success, How Social Media and User Data Play a Role in Search Results and Rankings

limitation of, Understanding What Content Search Engines Can "See" on a Web Page-What search engines cannot see

meta tags and, Understanding What Content Search Engines Can "See" on a Web Page

for mobilefriendly web pages, Mobile Friendliness

negative ranking factors, Negative Ranking Factors

PageRank (PR), The Original PageRank Algorithm-The Original PageRank Algorithm

ranking factor overview, Analyzing Ranking Factors

ranking factors in, Retrieval and Ranking, Study on Google+ as a Ranking Factor

rate of acquisition links, Other Ranking Factors

retrieving and ranking, Retrieval and Ranking

for trustworthiness, Trust

user data, Other Ranking Factors

user engagement, Content Quality and User Engagement, How Social Media and User Data Play a Role in Search Results and Rankings, User Engagement as a Measure of Search Quality

web page content evaluation, Evaluating Content on a Web Page

Raven, Tying SEO to Conversion and ROI, Raven Tools

reading level, Content Quality and User Engagement, Reading Level

reciprocal linking, How Are Links Earned?, Swapping links, Glossary

reconsideration requests, Path to Recovery, The Disavow Links tool, Filing reconsideration requests, Search Engine Penalties and Reconsideration Requests

recursive directs, avoiding, How to Redirect a Home Page Index File Without Looping

reddit

effective use of, reddit

tracking user interaction on, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

redirect methods, Glossary

avoiding looping , How to Redirect a Home Page Index File Without Looping

best use of, Redirects, Good and Bad Redirects

canonical, Duplicate content checks

checking with server header checker, URL redirect checks

default document redirect, The default document redirect solution

ISAPI_Rewrite, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting-mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting, Large-Scale Content Moves

for large-scale content moves, Large-Scale Content Moves

mod_rewrite, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting-mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting, Large-Scale Content Moves

overview of, Methods for URL Redirecting and Rewriting

purpose of, Redirects

why and when to use, Why and When to Redirect

referrer, Referring Sites, Action tracking by referral source, Glossary

regular expressions, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting

related terms, Related terms, KeywordDiscovery

relationship mapping, Google's Knowledge Graph, Document Analysis and Semantic Connectivity, Category structuring

relationships

- **building offline, Build Offline Relationships**
- **building with influencers, Building the relationship, Building Relationships with Influencers-Establish yourself on Google+**
- **creating value propositions for, Creating a Value Proposition for a Relationship**
- **using direct email effectively, Using Direct Email Pitches Effectively**

relative links, Glossary

relevance, Glossary

- **algorithm failures, Why These Algorithms Sometimes Fail**
- **content quality, Content Quality and User Engagement**
- **cross-linking content, Cross-link relevant content**
- **defined, Retrieval and Ranking**
- **determining, The Mission of Search Engines, Determining Searcher Intent and Delivering Relevant, Fresh Content**
- **document analysis for, Document Analysis and Semantic Connectivity**

in keywords, Identifying relevant keywords, Keyword Targeting

link analysis, Link Analysis, Authority, Target Areas of Penguin

link ranking value and, Relevance, Context/relevance

PPC results order by, Understanding the Layout of Search Results Pages

social media signals, Evaluating Social Media Signals, What Is the Value of Social Signals?

topical, Topical relevance

repeat visitors, Potential User Engagement Signals, Glossary

reputation management

indirect benefits of, The indirect benefits of building reputation and visibility

via product images, Optimizing for Image Search

SEO for, SEO for Reputation Management

tracking software for, Tracking the Blogosphere

request for proposal (RFP), Preparing a Request for Proposal

response times, Server and Hosting Issues

responsive web design (RWD), Same URL approaches: responsive web design and dynamic serving

return on investment (ROI), Glossary

determining by project, Determining Project ROI

initial experimentation cycle and, Deciding on a New Social Network

link building, Review your website assets

measurement tools, Tying SEO to Conversion and ROI

SEO vs. other methods, High Return on Investment

rewrite rules, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting

RFP (see request for proposal)

rich content, Implementing Content Marketing Strategies

rich snippets, Schema.org, Google Knowledge Graph Carousel, Google Search Console

Rio SEO Search Analytics, Rio SEO Search Analytics, Competitive Analysis for Those with a Big Budget

Rio SEO Website Optimizer, Rio SEO Website Optimizer

robots, Search Engine Robot Traffic Analysis, Glossary

robots.txt, Glossary

Allow directive, Syntax of the robots.txt file

cautions for use, Syntax of the robots.txt file

controlling spider access with, Leveraging the robots.txt file

crawl delay method, Syntax of the robots.txt file

example of, Syntax of the robots.txt file

example of use, Leveraging the robots.txt file

filename and location, Leveraging the robots.txt file

full instruction for, Syntax of the robots.txt file

HTTP/HTTPS and, Syntax of the robots.txt file

impact on search engines, Leveraging the robots.txt file

Noindex directive, Syntax of the robots.txt file

pattern matching and, Syntax of the robots.txt file

restrictions on, Syntax of the robots.txt file

spiderable link structures and, Spiderable Link Structures

syntax of, Syntax of the robots.txt file

use during site redesigns, Identifying Content That Search Engines Don’t See

XML sitemap autodiscovery, Syntax of the robots.txt file

ROI (see return on investment)

RSS pinging, Content Management System Issues

RWD (see responsive web design)

S

sales, tracking, Setting Up Analytics Software to Track Conversions

sample size, Common analytics mistakes

sandbox, Glossary

Scalable Inman Flash Replacement (sIFR), Scalable Inman Flash Replacement

Schema.org

- **benefits and challenges of, Overview**
- **best approach to using, How to Use Schema.org**
- **core concepts of, Schema.org**
- **establishment of, Structured Data**
- **general plan for use, How to Use Schema.org**
- **hierarchical inheritance in, How to Use Schema.org**
- **impact on SERP construction, Schema.org**
- **local search optimization and, Business Website Optimization for Local Search**
- **meaning communicated via markup, Overview**
- **modeling and markup in, How to Use Schema.org**
- **overview of, Summary**
- **properties in, How to Use Schema.org**
- **semantic web movement and, Schema.org**
- **specific steps for use, Step 1: Determine Schema.org types-Step 5: Test**
- **specific vs. general type descriptions, How to Use Schema.org**

types and items in, How to Use Schema.org

Search Console (see Google Search Console)

manual spam action notifications, Unnatural Links Messages

search engine marketing (SEM), Glossary

search engine optimization (see SEO (search engine optimization))

search engine results page (SERP), Glossary

10 blue links, Overview of Changes in Search Complexity

additional advertising links, Understanding the Layout of Search Results Pages

advertisements on, Searcher Intent

continually shifting formats of, Understanding the Layout of Search Results Pages

engine-specific formats, Understanding Search Engine Results

Google+ brand pages on, Google+ Brand Pages in the Search Results

Google+ posts on, Google+ Posts in the Search Results

horizontal navigation section, Understanding the Layout of Search Results Pages

impact of rich snippets on, Schema.org

information contained on, Understanding How Vertical Results Fit into the SERPs

layout of Bing results, Understanding the Layout of Search Results Pages

layout of Google results, Understanding the Layout of Search Results Pages

layout of Yahoo results, Understanding the Layout of Search Results Pages

natural/organic/algorithmic results section, Understanding the Layout of Search Results Pages

offering diverse results on, Problem Words, Disambiguation, and Diversity

paid search results, Understanding the Layout of Search Results Pages

query refinement suggestions, Understanding the Layout of Search Results Pages

results information section, Understanding the Layout of Search Results Pages

search query box, Understanding the Layout of Search Results Pages

user scanning of, The History of Eye Tracking: How Users Scan Results Pages

vertical navigation section, Understanding the Layout of Search Results Pages

vertical results, Understanding How Vertical Results Fit into the SERPs-Understanding How Vertical Results Fit into the SERPs

search engine spam, Glossary

search engines

ability to read Flash files, What search engines cannot see, Technology

Decisions, Indexable Content, Showing Different Content to Engines and Visitors

algorithm development/tweaking, How Social Media and User Data Play a Role in Search Results and Rankings

vs. alternate tools, Determining Searcher Intent: A Challenge for Search Marketers and Search Engines

audits and, Search engine health checks

common steps used, The Human Goals of Searching

comparing, The Importance of Experience

comparing algorithms, Analysis of Algorithmic Differentiation Across Engines and Search Types

content invisible to, What search engines cannot see-What search engines cannot see, Single-Page Applications, Hidden Content

country-specific, Country-Specific Search Engines-Optimizing for Specific Countries

determining unique page content, Evaluating Content on a Web Page

ecommerce sales and, How Search Engines Drive Ecommerce

elements read by, Understanding What Content Search Engines Can "See" on a Web Page

employee communications concerning, Commentary from search engine employees

frames/iframes and, What search engines cannot see

identifying duplicate content, How Search Engines Identify Duplicate Content

importance of understanding, Conclusion

institutional biases among, The Importance of Experience

keyword research data from, Keyword Research Data from Search Engines-Frequency of recent usage

limitations of, Search Engine Basics, Understanding What Content Search Engines Can "See" on a Web Page-What search engines cannot see

link value determination by, How Search Engines Determine a Link's Value

market share of, The Market Share of Search Engines, Country-Specific Search Engines, An Evolving Art Form: The Future of SEO

measuring traffic by, Traffic by search engine

mission of, The Mission of Search Engines

planned improvements to, Engines Will Make Crawling Improvements

removing URLs from indexes, Removing URLs from a search engine's index

revenue generation by, The Mission of Search Engines

search-friendly site navigation, Search-Friendly Site Navigation-Search engine–friendly navigation guidelines

spam penalties and, The Mission of Search Engines, Other Search Engine

Courses of Action

spam prevention by, How Search Engines Fight Link Spam-Other Search Engine Courses of Action

tasks executed, Algorithm-Based Ranking Systems: Crawling, Indexing, and Ranking

use of links by, How Search Engines Use Links

use of semantic connectivity, Document Analysis and Semantic Connectivity

use of social signals, How Social Media and User Data Play a Role in Search Results and Rankings

user profiles, Who Searches and What Do They Search For?

user satisfaction, Searcher Intent

vertical, Vertical Search Engines-More specialized vertical search engines, Optimizing for Vertical Search

vs. directories, Glossary

web page content evaluation, Evaluating Content on a Web Page

web structures recognized, Root Domains, Subdomains, and Microsites

(see also crawlers)

search intent, Glossary

adaptive search, Adaptive Search, Other Ranking Factors

challenge of determining, Determining Searcher Intent: A Challenge for Search Marketers and Search Engines

disambiguation, Problem Words, Disambiguation, and Diversity

informational queries, Informational Queries

keyword "not provided" and, Understanding Keyword “Not Provided” and Co-Occurrence Analysis, Traffic by keyword

keyword use and, Searcher Intent

navigational queries, Navigational Queries

transactional queries, Transactional Queries

search metrics (see web analytics)

search query box, Understanding the Layout of Search Results Pages

Search Query Terms Report, Bing Webmaster Tools

search results

impact of drop in, Panda, Penguin, and Penalties

organic, Determining Searcher Intent: A Challenge for Search Marketers and Search Engines, Click Tracking: How Users Click on Results (Organic Versus Paid), Glossary

website alterations affecting, Domain Changes, Redesigns, and Troubleshooting

search terms

common usage/phrase combinations, Common usage and phrase combinations

defined, Glossary

effect on traffic, Traffic by keyword

extracting from relevant web pages, Extracting Terms from Relevant Web Pages

frequency of recent usage, Frequency of recent usage

misspelled, Document Analysis and Semantic Connectivity, KeywordDiscovery

related terms, Related terms, KeywordDiscovery

seasonal fluctuations, Frequency of recent usage, Trending, Seasonality, and Seasonal Fluctuations in Keyword Demand

searches/searching

advanced techniques with Bing, Bing Advanced Search Operators

advanced techniques with Google, Advanced Google Search Operators-Firefox plugins for quicker access to Google advanced search queries

Boolean searches, Document Analysis and Semantic Connectivity

branded vs. nonbranded, Branded versus nonbranded traffic, Understand Your Market Opportunity

business building through, Who Searches and What Do They Search For?

by query length, The Human Goals of Searching

common steps followed, The Human Goals of Searching

determining searcher intent, Determining Searcher Intent: A Challenge for Search Marketers and Search Engines-Searcher Intent

for images, Image search

for maps, Local search/maps

for news, News search, Optimizing for News Search: Google News-Editor's Picks

for videos, Video search, Optimizing for Video/Multimedia Search-Use recognized file types

fuzzy logic, Document Analysis and Semantic Connectivity

high demand for, Search: Reflecting Consciousness and Connecting Commerce

human goals of, The Human Goals of Searching

incognito, Google+ Personalization

increased complexity of, Overview of Changes in Search Complexity, The Ongoing Evolution of Search-The Growth of Search Complexity

of blogs/RSS feeds, Blog search, Optimizing for Blog Search-Links

of books, Book search

organic vs. paid clicks, Click Tracking: How Users Click on Results (Organic Versus Paid)

product features, Product search

proximity, Document Analysis and Semantic Connectivity, Business Website Optimization for Local Search

result page scanning, The History of Eye Tracking: How Users Scan Results Pages

search operators, Using Advanced Search Techniques, More Advanced Search Operator Techniques, Frequency of recent usage

search process, How People Search-How People Search

semantic, The Knowledge Graph, The Future of Semantic Search and the Knowledge Graph

shift toward mobile ad spend, The Mobile Shift, The Mobile Landscape

term weighting, Document Analysis and Semantic Connectivity

Universal/blended, Universal Search/Blended Search, Overview of Changes in Search Complexity, Optimizing for Vertical Search, Optimizing for Image Search

user profiles, Who Searches and What Do They Search For?

Searchmetrics, Searchmetrics, Traffic by keyword, Searchmetrics Essentials, Competitive Analysis for Those with a Big Budget

seasonal data, Frequency of recent usage, Trending, Seasonality, and Seasonal Fluctuations in Keyword Demand

second-level domains, Root Domains, Subdomains, and Microsites

segmentation, Segmenting Your Site's Audience, Page segmentation, Displaying Different Content to Search Engines Versus Visitors, Segmenting Your Audience, Identifying Personas, and Targeting Content, Segmenting Search Traffic, Segmenting Campaigns and SEO Efforts by Conversion Rate

SEM (see search engine marketing)

semantic connectivity, Document Analysis and Semantic Connectivity,

Category structuring

semantic maps, Evaluating Content on a Web Page

semantic markup, CSS and Semantic Markup, Schema.org

semantic search

basis of, Google's Hummingbird

entities and, Entities and Semantic Search

future of, The Future of Semantic Search and the Knowledge Graph

Hummingbird, Google's Hummingbird

vs. knowledge graph, The Knowledge Graph

SEO factors of, Semantic Search and SEO

structured data and, Structured Data

vectors concerning, Semantic Search

seminars, Build Offline Relationships

SEMrush, SEMrush-Cost, Traffic by keyword, SEMrush

SEO (search engine optimization)

advanced search operators for, Advanced Google Search Operators-Finding Creative Commons (CC) licensed content

black hat, <noscript>, Gray Hat/Black Hat-Hosting giveaways, Negative Link Building, Glossary

for blogs/RSS feeds, Optimizing for Blog Search-Links

business of (see SEO support)

country-specific, Optimizing for Specific Countries

definition of term, Glossary

for vertical search, Optimizing for Vertical Search-The Opportunity Unleashed

future of, An Evolving Art Form: The Future of SEO-The Future of Semantic Search and the Knowledge Graph

for Google News, Optimizing for News Search: Google News-Editor's Picks

for Google Shopping, Optimizing for Google Shopping Search-Reporting Results of Shopping Ads

gray hat, Strategies that are not considered buying links, Glossary

importance of, Search: Reflecting Consciousness and Connecting Commerce, SEO Planning: Customizing Your Strategy

for local search, Optimizing for Local Search-Business Website Optimization for Local Search

major objectives of, SEO as an Enduring Art Form

measuring success of (see tracking and measurement systems)

for mobile apps, App SEO: Deep Linking and Indexing for Mobile Search

for mobile, SEO for Mobile, General SEO guidelines for mobile

as process vs. project, SEO Planning: Customizing Your Strategy

vs. social media marketing, Monitoring, Measuring, and Improving Social Media Marketing: Best Practices

testing and experiments, SEO Testing

(see also SEO support)

ultimate objective of, An Evolving Art Form: The Future of SEO

for videos, Optimizing for Video/Multimedia Search-Use recognized file types

white hat, Glossary

SEO for Firefox, SEO for Firefox

SEO implementation

accessing server statistics, Identifying Current Server Statistics Software and Gaining Access

assessing historical progress, Assessing Historical Progress

auditing for potential problems, Auditing an Existing Site to Identify SEO Problems-Server and Hosting Issues

benchmarking current indexing status, Benchmarking Current Indexing Status

benchmarking current traffic, Benchmarking Current Traffic Sources and Volume

benchmarking organic rankings, Benchmarking Organic Rankings

creating detailed documentation, SEO Documentation for Actions and Progress

determining top competitors, Determining Top Competitors

development platform/information architecture, Development Platform and Information Architecture-Server rendering versus client-Only rendering, Creating an Optimal Information Architecture-Search engine–friendly navigation guidelines

for single-page applications, Single-Page Applications-Server rendering versus client-Only rendering

identifying development process/players, Identifying the Site Development Process and Players

importance of planning, The Importance of Planning

(see also SEO strategy)

leveraging business assets for, Leveraging Business Assets for SEO

overview of, SEO Implementation: First Stages

testing and experiments, SEO Testing

SEO problems, discovering (see audits)

SEO Quake, SEO Quake

SEO research (see industry resources)

SEO strategy

audience segmentation, Segmenting Your Site’s Audience

business factors affecting, Business Factors That Impact Your SEO Strategy

competitive analysis of, What Is Their SEO Strategy?

customizing, Every SEO Strategy Should Be Customized

finding your niche, Understanding Your Audience and Finding Your Niche

for ecommerce sales, SEO for Ecommerce Sales

for ideological influence, SEO for Ideological Influence

for lead generation/direct marketing, SEO for Lead Generation and Direct Marketing

for mindshare/branding, SEO for Mindshare and Branding

for raw traffic, SEO for Raw Traffic

for reputation management, SEO for Reputation Management

integrated approach to, Developing an SEO Plan Prior to Site Development

mapping products/services, Mapping Your Products and Services

market competitiveness, Understanding Context: Market Competitiveness

psychological basis for, The Human Goals of Searching

role of content in, Business Factors That Impact Your SEO Strategy, Understanding That Content Is King

SMART objectives, SMART Objectives

strategic goals, Strategic Goals SEO Practitioners Can Fulfill

SWOT analysis for, SWOT Analysis

understanding search engines/queries, Understanding Search Engine Traffic and Visitor Intent, Business Factors That Impact Your SEO Strategy

SEO support

benefits of outside experts, The Case for Working with an Outside Expert

collaborating with internal teams, How to Integrate SEO Knowledge in the Organization

creating SEO culture, Building a Culture of SEO into Your Organization

hiring SEO firms/consultants, Selecting an SEO Firm/Consultant-Making the Decision

hiring SEO talent, Hiring SEO Talent

impact of site complexity on workload, The Impact of Site Complexity on SEO Workload

implementing recommendations, How to Implement Your Expert's Recommendations, Applying SEO Recommendations Intelligently

in-house vs. outsourced teams, Using In-House SEO Talent Versus Outsourced Support: Dynamics and Challenges

integrating internal and external teams, Mixing In-House SEO with Outside SEO Help

integrating SEO knowledge, How to Integrate SEO Knowledge in the Organization

laying groundwork for, Lay the Groundwork

leveraging outside help, How to Best Leverage Outside Help

market opportunity assessment, Understand Your Market Opportunity

motivating resources, Motivate Resources That Don’t Share Your Goals to Help You

obtaining organizational buy-in, Get Buy-In Across the Organization

preparing request for proposal (RFP), Preparing a Request for Proposal

SEO as business focus, The Business of SEO

solutions for large organizations, Solutions for Large Organizations

solutions for small organizations, Solutions for Small Organizations

stages of maturity, Progress Through the Stages of SEO Maturity

team building, Building an SEO team, Hiring SEO Talent

SEO vendors

changing, Changing SEO Vendors or Staff Members

selecting, Selecting an SEO Firm/Consultant-Making the Decision

SEO-Browser, Third-party tools to check for crawl errors

seoClarity, Competitive Analysis for Those with a Big Budget

SERP (see search engine results page)

server issues

changing servers, Changing Servers

dynamic serving, Same URL approaches: responsive web design and dynamic serving

dysfunctional servers, Search engine–friendly navigation guidelines

server geography, Server and Hosting Issues

server timeouts, Server and Hosting Issues

tracking and measurement systems, Identifying Current Server Statistics Software and Gaining Access

services, mapping of, Mapping Your Products and Services

session IDs, Technology Decisions, Search engine–friendly navigation guidelines, Controlling Content with Cookies and Session IDs, Session IDs

Seznam, Country-Specific Search Engines

shareable content, Link-Worthy or Viral Content

shelf space, Shelf space

shingles, Duplicate Content Issues

sIFR (see Scalable Inman Flash Replacement)

single-page applications (SPAs) , Single-Page Applications

site search elements, Search and web forms

sitemaps, Glossary

adding video information to, Adding video information to a sitemap

audit review and, Step 2: Site architecture

examining with Google Search Console, Google Search Console

for Google News, Google News Sitemaps

for Google video, Use a Google video sitemap

XML sitemaps, XML Sitemaps, Glossary

sitewide links, Links across domains, Other Algorithmic Approaches to Fighting Link Spam, Sitewide link penalties

SMART objectives, SMART Objectives

smartphones (see mobilefriendly web pages)

SME (see subject matter experts)

SMM (see social media marketing)

Snapchat, Instagram, Vine, and Snapchat

snippets, Duplicate Content Issues, Schema.org

social media

additional resources on, Additional Social Media Resources

applications for, Social Networking for Links

audience building through, Get Active in Social Media, The Indirect Influence of Social Media Marketing

benefits and drawbacks of, Putting It All Together

best practices, Using Social Media Networking Effectively: Dos and Don’ts

Bing Places for Business, Bing Places for Business

Bing's vs. Google's relationship with, Bing’s Experiments with Social Signals

blogging for links, Blogging for Links

contacting Pinterest influencers, Leverage Pinterest group boards

in content marketing strategy, Implementing Content Marketing Strategies, Brainstorming Content Ideas and Being Creative

cost of, What Is the Value of Social Signals?

determining content relevancy via, Evaluating Social Media Signals, What Is the Value of Social Signals?

discovering audiences through, Get to Know Other People's Audiences

Facebook, Facebook, Bing's Experiments with Social Signals, Does Google Use Facebook as a Ranking Signal?, The Indirect Influence of Social Media Marketing, Facebook, Setting Up Analytics Software to Track Conversions, An Evolving Art Form: The Future of SEO

Flickr, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube, Image Optimization Tips

getting mentioned on, Use email to contact the influencers

Google+, Establish yourself on Google+, Google+, Does Google Use Google+ as a Ranking Signal?-How Might Google Use Google+ as a Ranking Factor?, The Indirect Influence of Social Media Marketing, Google+

guest posting, Implementing Guest Posting Successfully

image sharing sites, Image Optimization Tips

impact on ranking, Content Marketing

indirect influence of, The Indirect Influence of Social Media Marketing

Instagram, Instagram, Vine, and Snapchat, The Indirect Influence of Social Media Marketing, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

leveraging major platforms, Leveraging Major Social Media Platforms

LinkedIn, The Indirect Influence of Social Media Marketing, LinkedIn

links gained via advertising, Social advertising can drive links to your site

manual link creation, Manual Social Media Link Creation

Meetup.com, Meetup.com

NAP information consistency and, Local Listing Submissions

paid vs. organic activity, Paid social media

Pinterest, Pinterest, The Indirect Influence of Social Media Marketing

psychographic targeting, Psychographic targeting

Quora, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

as ranking factor, Analyzing Ranking Factors

reddit, reddit, Social news sites

return on investment in, Deciding on a New Social Network

Snapchat, Instagram, Vine, and Snapchat

social content amplification steps, Social content amplification steps

social distribution, Social distribution

StumbleUpon, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

targeting influencers on Twitter, Use Twitter as a warmup

Techmeme, Social news sites

Twitter, Twitter, Does Google Use Twitter as a Ranking Signal?, The Indirect Influence of Social Media Marketing, Twitter

Vine, Instagram, Vine, and Snapchat

WhatsApp, The Indirect Influence of Social Media Marketing

Wikipedia, Wikipedia

wikis, Wikis

Yelp, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

YouTube, Using YouTube Successfully for Content Marketing, The Indirect Influence of Social Media Marketing, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube, Video SEO for YouTube-Leverage paid advertising

(see also social signals))

social media marketing (SMM), Glossary

blogs and forums, Blogs and forums

claiming key brand profiles, Claiming Key Profiles

developing influencer status, Monitoring, Measuring, and Improving Social Media Marketing: Best Practices

effective use of, Monitoring, Measuring, and Improving Social Media Marketing: Best Practices

Facebook, Facebook

Google+, Google+

indirect influence of, The Indirect Influence of Social Media Marketing

key to success, Monitoring, Measuring, and Improving Social Media Marketing: Best Practices

LinkedIn, LinkedIn

news sites, Social news sites

performance metrics, Tracking Social Media

selecting social networks, Deciding on a New Social Network

sites worth considering, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

social network structure, Monitoring, Measuring, and Improving Social Media Marketing: Best Practices

Twitter, Twitter

social proof signals, Content Marketing

social signals

Bing's experiment with, Bing’s Experiments with Social Signals

correlation with Google rankings, Correlation Between Social Signals and Google Rankings

Google's use of Facebook, Does Google Use Facebook as a Ranking Signal?

Google's use of Google+, Does Google Use Google+ as a Ranking Signal?-How Might Google Use Google+ as a Ranking Factor?

Google's use of Twitter, Does Google Use Twitter as a Ranking Signal?

use in ranking systems, How Social Media and User Data Play a Role in Search Results and Rankings

value of, What Is the Value of Social Signals?

Soovle, Soovle

source code, viewing, Understanding What Content Search Engines Can "See" on a Web Page

source diversity, Source diversity

source independence, Source independence

SPA (see single-page applications)

spam

basic rules for spam-free SEO, Basic Rules for Spam-Free SEO

comment spam, Target Areas of Penguin, Other types of manual penalties, Links Google Does Not Like, Glossary

content management systems and, Content Management System Issues

identifying, Identifying Hidden Content That May Be Viewed as Spam, Spam Filtering and Penalties-Duplicate Content

identifying in competitors, Identifying Spam

penalties for, The Mission of Search Engines, Other Search Engine Courses of Action, Other types of manual penalties

preventing, Content Management System Issues, How Search Engines Fight Link Spam-Other Search Engine Courses of Action

search engine spam, Spam Filtering and Penalties, Glossary

signals of low-quality domains, Low-Quality Domains and Spam Sites

spamdexing, Glossary

spamglish, Glossary

spammers, Understanding What Content Search Engines Can “See” on a Web Page, Glossary

special offers, Submit a special offer with a promotion feed

speedstorming, Speedstorming

spelling errors, Content Quality and User Engagement, KeywordDiscovery, Poor Editorial Quality

spider traps, Search engine–friendly navigation guidelines, Glossary

spiders (see crawlers)

SpyFu, SpyFu

static pages, Glossary

static URLs, Technology Decisions

stemming, Glossary

stickiness, Glossary

sticky posts, Sticky Posts

stopwords, Glossary

strategic reviews, Conduct regular strategic reviews

structured data

- **semantic search and, Structured Data**
- **SEO best practices for, Structured Data**

StumbleUpon, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

subdomains

- **logical flow and, Subdomains**
- **trust level of, Avoidance of unnecessary subdomains**
- **vs. root domains and microsites, Root Domains, Subdomains, and Microsites**
- **vs. subfolders, When to Use a Subfolder**
- **when to use, When to Use a Subdomain, Picking the Right URLs**

subfolders

- **examining with Bing Webmaster Tools, Bing Webmaster Tools**
- **metrics assigned to, Root Domains, Subdomains, and Microsites**
- **URLs and, Picking the Right URLs**

when to use, When to Use a Subfolder

subject matter experts (SME), Developing Content That Works

success, measuring (see tracking and measurement systems)

supplemental index, Glossary

SWFObject, SWFObject

SWOT analysis, SWOT Analysis-SWOT Guidelines

syndication, Implementing Content Marketing Strategies, Content Syndication

synonyms, Synonyms

T

tags, Glossary

in blog posts, Structural Blog Optimizations

Data Highlighter tool, Google Knowledge Graph Carousel, Google Search Console

examining with Google Search Console, Google Search Console

heading tags, Heading Tags, Glossary

publisher tag, Google's Publisher Tag, Google+ Brand Pages in the Search Results, Optimizing for News Search: Google News

title tags, HTML <title> tag review, HTML <title> Tags, Content Management System Issues, Integrating Emotional Triggers, Titles, and Images, Business Website Optimization for Local Search, Optimize the title, Google Search Console, Glossary

target audience, Business Factors That Impact Your SEO Strategy, Segmenting Your Site's Audience, Segmenting Your Audience, Identifying Personas, and Targeting Content, The Indirect Influence of Social Media Marketing, Glossary

target keywords, Target keywords

targeted traffic, High Return on Investment

targeting data, obtaining, Psychographic targeting

taxonomy, Taxonomy and ontology, Glossary

Techmeme, The Indirect Influence of Social Media Marketing, Social news sites

temporal link growth measurements, Temporal Link Growth Measurements

temporary (302) redirects, Technology Decisions, Good and Bad Redirects

Teoma, Authority

term frequency-inverse document frequency (see TF-IDF)

term weighting, Document Analysis and Semantic Connectivity

text

- boldface, Boldface Text
- hidden, Other types of manual penalties, Glossary
- hiding in images, Hiding text in images
- hiding in Java applets, Hiding text in Java applets

text links, Search engine–friendly navigation guidelines, Glossary

text-mining, Extracting Terms from Relevant Web Pages

TF-IDF (term frequency-inverse document frequency), Term frequency—Inverse document frequency

theme, Evaluating Content on a Web Page, Content Themes, Content Management System Issues, Glossary

thin affiliates, A word of caution to affiliates, An actual penalty situation, Glossary

thin content, Content Uniqueness and Depth, Duplicate Content Issues, Semantic Search and SEO, Target Areas of Panda, Thin-content penalties, Glossary

thin slicing, Duplicate Content Issues, Target Areas of Panda, Thin-content penalties

three-way link swaps , Using link farms/link networks

thumbnails, Image Optimization Tips, Thumbnail Images in Google News

time on page, Potential User Engagement Signals, Glossary

time on site analytic, Content Quality and User Engagement, Potential User Engagement Signals

Time to Live (TTL), Changing Servers

title bars, Understanding What Content Search Engines Can "See" on a Web Page

title tags, HTML <title> tag review, HTML <title> Tags, Content Management System Issues, Integrating Emotional Triggers, Titles, and Images, Business Website Optimization for Local Search, Optimize the title,

Google Search Console, Glossary

TLD (see top-level domain)

Toolbar PageRank, Glossary

toolbars, How Google and Bing Collect Engagement Metrics, Glossary

Top Heavy algorithm, Summary of Major Google Algorithms

top-level domain (TLD)

- fully qualified, Root Domains, Subdomains, and Microsites
- generic top-level domains, New gTLDs
- other than .com, When to Use a TLD Other Than .com, Optimizing Domains
- ranking value and, Source TLDs

trackbacks, Content Management System Issues

tracking and measurement systems

- action tracking, Taking a Deeper Look at Action Tracking
- baseline data points, Why Measuring Success Is Essential to the SEO Process
- for blogs, Tracking the Blogosphere
- common mistakes, Common analytics mistakes
- discovering/diagnosing crawl errors, Crawl Errors-Third-party tools to check for crawl errors

establishing goals, Tracking Results and Measuring Success

for longtail search queries, Key Performance Indicators for LongTail SEO

free web analytics tools, Identifying Current Server Statistics Software and Gaining Access

linkbased, LinkBased Tracking of Content Marketing-Measuring the value of a link

log file tracking, Log file Tracking

measuring search traffic, Measuring Search Traffic-Common analytics mistakes

metrics commonly analyzed, Tracking Results and Measuring Success

ranking, Ranking

search engine robot traffic analysis, Search Engine Robot Traffic Analysis

search metrics, Competitive and Diagnostic Search Metrics-Temporal Link Growth Measurements

SEO platforms, SEO Platforms

shelf space, Shelf space

temporal link growth measurements, Temporal Link Growth Measurements

time required to see improvement, The Tracking Cycle: Produce, Launch, Measure, Refine

tools for, Other Third-Party Tools

tracking cycle, The Tracking Cycle: Produce, Launch, Measure, Refine

tying SEO to conversion and ROI, Tying SEO to Conversion and ROI-An alternate theory and calculation

using as a business case for SEO, Using Analytics as a Business Case for SEO, Determining Project ROI

web traffic comparison, Web Traffic Comparison

trademark infringement, Optimizing Domains

traffic, Glossary

benchmarking, Benchmarking Current Traffic Sources and Volume, Google Search Console

branded vs. nonbranded, Branded versus nonbranded traffic, Understand Your Market Opportunity

examining with Google Search Console, Google Search Console

measuring, Measuring Search Traffic-Common analytics mistakes

segmenting, Segmenting Search Traffic

SEO strategy for raw, SEO for Raw Traffic

strategic goals for, Website Traffic

targeted, High Return on Investment

web traffic comparison, Web Traffic Comparison

traffic loss

algorithms to examine, Summary of Major Google Algorithms

confirming, Diagnosing the Cause of a Traffic Loss, Penguin

confirming date of, Diagnosing the Cause of a Traffic Loss, Penguin

diagnosing cause of, Diagnosing the Cause of a Traffic Loss

due to domain name changes, Pre-Move Preparations

due to Panda algorithm, Panda-Successful Panda recovery example

due to Penguin algorithm, Penguin-Path to Recovery

due to poorly-valued links, Links Google Does Not Like-Links Google Does Not Like

effect on baseline measurements, Establishing a Proper Baseline

impact on business, Panda, Penguin, and Penalties

penalties resulting in, Penalties-Other types of manual penalties

recovering from, Path to Recovery

recovery examples, Successful Panda recovery example

recovery timeline, Expected timeline to recovery

transactional queries, Transactional Queries, Understanding Search Engine Traffic and Visitor Intent

transfer limitations, Server and Hosting Issues

Trellian's KeywordDiscovery (see KeywordDiscovery)

Trojans, Negative Ranking Factors

TrueView advertisements, Social content amplification steps

trust, Trust, The Indirect Influence of Social Media Marketing, Become a Google Trusted Store

TrustRank, Glossary

TTL (see Time to Live)

Twitter

- **building relationships with influencers on, Build an initial list**
- **effective use of, Twitter**
- **generating content ideas with, Brainstorming Content Ideas and Being Creative**
- **Google's use of, Does Google Use Twitter as a Ranking Signal?**
- **improving SEO through, Step 2: Send It to Your Customers/Potential Customers**
- **increasing exposure through, Get Active in Social Media**
- **indirect form of obtaining links, Leveraging Major Social Media Platforms**
- **manual link creation, Manual Social Media Link Creation**
- **number of users, The Indirect Influence of Social Media Marketing**
- **social content amplification steps, Social content amplification steps**
- **tracking user interaction on, Twitter**
- **using as a warmup, Use Twitter as a warmup**

Twitter memes, Other types of memes

types, How to Use Schema.org

U

Ubersuggest, Ubersuggest

UGC (see user-generated content)

Uniform Resource Locator (see URL (Uniform Resource Locator))

Unique Page Content, Evaluating Content on a Web Page, Understanding What Content Search Engines Can “See” on a Web Page

unique product identifiers (UPIs), Use correct product identifiers

unique visitors, Glossary

Universal Analytics, Setting Up Analytics Software to Track Conversions

Universal Search, The History of Eye Tracking: How Users Scan Results Pages, Universal Search/Blended Search, Optimizing for Vertical Search, Optimizing for Image Search

unnatural links, Unnatural Links Messages, Other types of manual penalties, Spam Filtering and Penalties

UPIs (see unique product identifiers)

URL (Uniform Resource Locator), Glossary

adding to sitemap files, XML Sitemaps

best practices for, URL checks

for blog posts, Structural Blog Optimizations

canonical, Consequences of Duplicate Content, Using the rel="canonical" link element, Content Syndication

content management systems and, Content Management System Issues

custom creation of, Content Management System Issues

discovering blocked, Google Search Console

dynamic vs. static, Technology Decisions, URLs, Picking the Right URLs

mapping content moves using, Mapping Content Moves

migrating, Expectations for Content Moves

rankings assigned to, Root Domains, Subdomains, and Microsites

redirect audit review, URL redirect checks

removing from indexes, Removing URLs from a search engine's index

rewriting and redirecting, mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting-mod_rewrite and ISAPI_Rewrite for URL rewriting and redirecting

same URL approach to mobile SEO, Same URL approaches: responsive web design and dynamic serving

selecting optimal, Picking the Right URLs

separate URL approach to mobile SEO, Separate URL approach

session/user IDs in, Technology Decisions, Session IDs

superfluous parameters in, Technology Decisions

user-friendly, Single-Page Applications

URL rewrite, Glossary

usability, Usability, Glossary

user agent, Glossary

user engagement

application to SEO, Content Quality and User Engagement

Chrome Personal Blocklist Extension, Content Quality and User Engagement, Voting Mechanisms, Panda

determining content quality through, Content Quality and User Engagement

determining search quality through, User Engagement as a Measure of Search Quality

vs. document analysis, Document Analysis

Goo.gl URL shortener, Content Quality and User Engagement, How Google and Bing Collect Engagement Metrics

Google +1 button, Content Quality and User Engagement, Google+ Personalization

Google Analytics, Content Quality and User Engagement, How Google and Bing Collect Engagement Metrics

Google Toolbar, Content Quality and User Engagement, How Google and Bing Collect Engagement Metrics

measuring, Measuring Engagement in Content Marketing, How Google

and Bing Collect Engagement Metrics

search result interaction, Content Quality and User Engagement, Managing Attribution

signals of, Potential User Engagement Signals

voting mechanisms, Voting Mechanisms, Panda

user IDs, Technology Decisions, Blocking and cloaking by user agent, Session IDs

user intent, Optimizing User Experience to Improve SEO, More Personalized, Localized, and User-Influenced Search, Glossary

(see also search intent)

user session, Glossary

examples of, How People Search

initiation of, Retrieval and Ranking

optimizing user experience, Optimizing User Experience to Improve SEO

user-agent detection, Mobile user agents

user-generated content (UGC), Applying User-Generated Content Strategies for LongTail Targeting, User-Generated Content, Other types of manual penalties, Taking a Deeper Look at Action Tracking, Glossary

V

vertical navigation, Understanding the Layout of Search Results Pages

vertical results

defined, Understanding Search Engine Results

examples of Bing search, Understanding How Vertical Results Fit into the SERPs

examples of Google search, Understanding How Vertical Results Fit into the SERPs

examples of Yahoo search, Understanding How Vertical Results Fit into the SERPs

impact on traffic, Understanding How Vertical Results Fit into the SERPs

vs. knowledge graph, The Knowledge Graph

opportunities afforded by, Understanding How Vertical Results Fit into the SERPs, Vertical Search Engines

vs. web results, Understanding How Vertical Results Fit into the SERPs

vertical search

defined, Optimizing for Vertical Search

for images, Image search, Optimizing for Image Search-Image Optimization Tips

for news, News search, Optimizing for News Search: Google News-Editor's Picks

for videos, Video search, Optimizing for Video/Multimedia Search-Use recognized file types

Google Shopping, Optimizing for Google Shopping Search-Reporting Results of Shopping Ads

local/maps, Local search/maps, Optimizing for Local Search-Business Website Optimization for Local Search

of blogs/RSS feeds, Blog search, Optimizing for Blog Search-Links

of books, Book search

optimizing for, Optimizing for Vertical Search-The Opportunity Unleashed

overview of products, Vertical Search from the Major Search Engines

product features, Product search

specialty websites offering, Vertical Search Engines, More specialized vertical search engines

universal/blended, Universal Search/Blended Search, Optimizing for Vertical Search

video files

crawlability of, What search engines cannot see, Java, images, audio, and video

in Google News, Videos in Google News

sitemaps for, Video sitemaps

video search

optimization for, Optimizing for Video/Multimedia Search

recognized file types, Use recognized file types

search engines focusing on, Video search

SEO for Google.com, Video SEO for Google

SEO for YouTube, Video SEO for YouTube-Leverage paid advertising

video distribution services, Use a video distribution service

viewer engagement, measuring, Measuring Engagement in Content Marketing

(see also user engagement)

Vine, Instagram, Vine, and Snapchat

viral content, Implementing Content Marketing Strategies, Link-Worthy or Viral Content, The Indirect Influence of Social Media Marketing

viruses, Negative Ranking Factors

visibility (branding) goals, Visibility (Branding), The indirect benefits of building reputation and visibility

vision-impaired users, Understanding What Content Search Engines Can "See" on a Web Page

visitors

repeat, Potential User Engagement Signals, Glossary

unique, Glossary

visual layout, Visual layout

Visual Thesaurus, Visual Thesaurus

Visuwords, Visuwords

voice search

examples of, Document Analysis and Semantic Connectivity

increased demand for, An Evolving Art Form: The Future of SEO, Voice Recognition Search

technology behind, Google's Hummingbird

volume, Semantic Search

voting mechanisms, Voting Mechanisms, Panda, Business Website Optimization for Local Search

W

W3C validation, Code quality, Basics of search engine friendliness

Web (World Wide Web)

complex nature of, Search Engine Basics

increasing scale of, Search Engine Basics

invisible Web, Glossary

structures recognized by search engines, Root Domains, Subdomains, and Microsites

Web Accessibility Initiative, Understanding What Content Search Engines Can "See" on a Web Page

web address, Glossary

web analytics, Taking a Deeper Look at Action Tracking

applications for, Measuring Search Traffic

for blogs, Tracking the Blogosphere

common mistakes, Common analytics mistakes

discovering/diagnosing crawl errors, Crawl Errors-Third-party tools to check for crawl errors

extracting SEO data, Extracting Valuable SEO Data in Web Analytics

for longtail search traffic, Key Performance Indicators for LongTail SEO

free tools, Identifying Current Server Statistics Software and Gaining Access, Basic Overview

(see also tracking and measurement systems)

(see also web analytics)

hidden content and, Identifying Content That Search Engines Don't See

linkbased tracking, LinkBased Tracking of Content Marketing-Measuring the value of a link

older vs. newer methodology, Basic Overview

package selection, Selecting the Right Analytics Package, Using Custom Analytics Dashboards

ranking, Ranking

search engine robot traffic analysis, Search Engine Robot Traffic Analysis

SEO platforms, SEO Platforms

shelf space, Shelf space

site indexing data, Site Indexing Data-Site Indexing Data

staying on track with, Separating the Analytics Wheat from the Chaff

temporal link growth measurements, Temporal Link Growth Measurements

tools for, Other Third-Party Tools

tracking conversion rates with, Setting Up Analytics Software to Track Conversions

using custom analytics dashboards, Using Custom Analytics Dashboards

web traffic comparison, Web Traffic Comparison

web beacons, Glossary

web browsers, Glossary

web bugs, Glossary

web directories, Web Directories

web forms, Search and web forms, Forcing form submission, Setting Up Analytics Software to Track Conversions, Form submission requirement

web pages

content evaluation, Evaluating Content on a Web Page

doorway pages, Microsites, Thin-content penalties, Glossary

editorial quality of, Poor Editorial Quality

example crawler analysis, Understanding What Content Search Engines Can “See” on a Web Page

extracting search terms from, Extracting Terms from Relevant Web Pages

frames on, What search engines cannot see, Spiderable Link Structures

gateway pages, Glossary

grammatical errors on, Content Quality and User Engagement

high-value, Step 2: Site architecture

measuring number being indexed, Site Indexing Data

measuring number receiving traffic, Number of pages getting search traffic

mobilefriendly, Mobile Friendliness, The Mobile Landscape-App Indexing, Glossary

noindexing, Syntax of the robots.txt file, Using the meta robots tag, Content Management System Issues, Content Syndication, Identifying Content That Search Engines Don't See, Glossary

page load time, Page load time, Page Speed, Mobile site speed, Glossary

page segmentation, Page segmentation

page-level ranking factors, Analyzing Ranking Factors

pagejacking, Glossary

PageRank (PR), Internal linking checks, Root Domains, Subdomains, and Microsites, Glossary

recommended number of clicks to access, Flat Versus Deep Architecture

referrer pages, Referring Sites, Action tracking by referral source, Glossary

repairing weak content on, Ways to address weak pages

slow, Negative Ranking Factors

static, Glossary

titles of, Understanding What Content Search Engines Can “See” on a Web Page

Unique Page Content, Evaluating Content on a Web Page, Content Uniqueness and Depth

web results, vs. vertical results, Understanding How Vertical Results Fit into the SERPs

web spam teams, The Mission of Search Engines

Webalizer, Log file Tracking

weblogs (see blogs)

webmaster tools (see Bing Webmaster Tools; Google Search Console)

website redesigns

changing SEO vendors/staff, Changing SEO Vendors or Staff Members

changing servers, Changing Servers

Google guidelines for site moves, Domain Changes, Redesigns, and Troubleshooting

maintaining search engine visibility, Maintaining Search Engine Visibility During and After a Site Redesign

moving/migrating content, The Basics of Moving Content-Expectations for Content Moves

preventing copyright infringement, Content Theft

preventing hidden content, Hidden Content

spam filtering and penalties, Spam Filtering and Penalties-Search Engine Penalties and Reconsideration Requests

websites

alterations affecting search performance, Domain Changes, Redesigns, and Troubleshooting

approach to getting links, How Can Sites Approach Getting Links?

building high-quality, Target Areas of Panda

competitive analysis of, Competitive Analysis-Competitive Analysis for Those with a Big Budget

database-driven, Target Areas of Panda, Directories and aggregators, Glossary

evaluating trust of, Trust, Become a Google Trusted Store

examining with Bing Webmaster Tools, Bing Webmaster Tools-Bing Webmaster Tools

examining with Google Search Console, Google Search Console-Manual spam actions

identifying low-quality, Low-Quality Domains and Spam Sites

improving usability of, How People Search

MFA sites, Glossary

new domains vs. subfolder structure, Root Domains, Subdomains, and Microsites

optimizing for local search, Business Website Optimization for Local

Search

promoting through social media, The Indirect Influence of Social Media Marketing

reviewing assets of, Review your website assets

search engine-friendly vs. spammy, Spam Filtering and Penalties

SEO-friendly development, Developing an SEO-Friendly Website

SEO-related structural decisions, Structural Decisions, Step 2: Site architecture

SEO-related technology decisions, Technology Decisions

site architecture protocol, Site architecture protocol

site search elements, Search and web forms

targeting for content marketing, Identifying Types of Sites That Might Link to a Site Like Yours

traffic goals, Website Traffic

Webtrends, Selecting the Right Analytics Package

WhatsApp, The Indirect Influence of Social Media Marketing

white hat SEO, Glossary

widgets, Hosting giveaways, Links Google Does Not Like

Wikimedia Commons , Image Optimization Tips

Wikipedia, Wikipedia

word count, Content length and word count

WordStream, WordStream-Cost

World Wide Web (see Web)

X

XHTML, Content Management System Issues

XML (eXtensible Markup Language), Glossary

XML pinging, Content Management System Issues

XML sitemaps, Glossary

- **autodiscovery with robots.txt, Syntax of the robots.txt file**
- **benefits of, XML Sitemaps**
- **browsers supporting, XML Sitemaps**
- **determining which URLs to include, Deciding what to include in a sitemap file**
- **generator for, Laying out an XML sitemap, Content Management System Issues**
- **highlighting images with, Image Optimization Tips**
- **for image files, Image sitemaps**
- **importance of, XML Sitemaps**
- **laying out, Laying out an XML sitemap**
- **managing/updating, Managing and updating XML sitemaps**

for mobile devices, Mobile sitemaps

purpose of, XML Sitemaps

uploading sitemap files, Uploading your sitemap file

for video files, Video sitemaps

Y

Yahoo

adoption of Bing technology, An Evolving Art Form: The Future of SEO

vertical search products, Vertical Search from the Major Search Engines

Yahoo search results

page layout, Understanding the Layout of Search Results Pages

vertical result examples, Understanding How Vertical Results Fit into the SERPs

Yahoo! Answers , Brainstorming Content Ideas and Being Creative

Yahoo! Local, Yahoo! Local

Yandex, Country-Specific Search Engines

Yelp, The Indirect Influence of Social Media Marketing, Claiming Key Profiles, Instagram, StumbleUpon, Quora, Yelp, Flickr, and YouTube

YouTube, Video search, Social content amplification steps, Using YouTube Successfully for Content Marketing-Using YouTube Successfully for Content Marketing, The Indirect Influence of Social Media Marketing, Video SEO for YouTube-Leverage paid advertising

YouTube Autocomplete, YouTube Autocomplete

About the Authors

Eric Enge is the founder and CEO of Stone Temple Consulting, a leading SEO firm that serves a wide variety of companies, including a number of the world's largest brands. He writes regular columns in Search Engine Land and Moz. He also writes for Forbes, Copyblogger, and Social Media Today.

On the "Digital Marketing Excellence" blog, Eric publishes industry-leading research studies that have been covered in the *Wall Street Journal*, the *Huffington Post*, NBC News, and many other high-profile sites, as well as interviews of senior people from Google and Microsoft. Eric is also a highly regarded speaker, and he regularly speaks on digital marketing topics at conferences such as Pubcon, Search Engine Strategies, Search Marketing Expo, the American Marketing Association, SEMPO, Internet Retailers, O'Reilly Web 2.0, and many others. You can also see content published by Eric on a regular basis by following him on Twitter (@stonetemple) or putting him in one of your circles on Google+ (+Eric Enge).

Stephan Spencer is an internationally recognized SEO expert, Internet entrepreneur, sought-after professional speaker, and bestselling author. He is the author of *Google Power Search* and coauthor of *Social eCommerce*, both published by O'Reilly. He is the host of the podcasts Marketing Speak and The Optimized Geek. Stephan produces and presents the three-day training intensives Traffic Control and Passions into Profits. He is producing a video-based SEO training and coaching program at The Science of SEO.

Stephan founded the interactive agency Netconcepts in 1995 and grew it into a multinational SEO firm with dozens of staff in each of its U.S., New Zealand, and China offices. Stephan invented a pay-per-click SEO technology called GravityStream which contributed millions annually to the bottom line; indeed, the technology was responsible for the majority of Netconcepts's revenue. Netconcepts was acquired in 2010 by Covario. Covario was in turn acquired by the multibillion dollar ad agency conglomerate Dentsu Aegis in 2014. Stephan's clients since the acquisition have included Zappos, Sony, Chanel, Quiksilver, Bed Bath & Beyond, Best Buy Canada, Fluor, CNBC, Bloomberg BusinessWeek, Despegar, and IEEE, among others.

Stephan has spoken at hundreds of Internet marketing events, including all the major conferences in the search and ecommerce industries (SES, SMX, PubCon,

Internet Retailer, Shop.org, eTail, etc.). He's been a contributor to Search Engine Land, The Huffington Post, Multichannel Merchant, Practical Ecommerce, DM News, and MarketingProfs, to name a few. Stephan is an avid blogger. He blogs at StephanSpencer.com, ScienceOfSEO.com, Natural Search Blog, Business Blog Consulting, and Changes For Good.

Stephan is on the board of Impact Network, a nonprofit that builds and operates schools in Zambia. Stephan holds an MS in biochemistry from the University of Wisconsin-Madison.

Connect with and follow Stephan on Facebook (StephanSpencerSEO), Twitter (@sspencer), Google+ (+StephanSpencer), and Pinterest (stephanspencer).

Jessie C. Stricchiola is the founder and CEO of Alchemist Media, Inc., a San Francisco-based digital consulting firm with core expertise in search marketing. She began her digital marketing career in 1997 by reverse engineering search results to understand how to rank websites more prominently within them—a practice that later became known as SEO. Since 2002 she has been a guest speaker at leading search and digital marketing conferences, including Search Engine Strategies, WebmasterWorld's PubCon, the O'Reilly Web 2.0 Summit, Third Door Media's Search Marketing Expo, ad:Tech, MarketLive Ecommerce Summit, and Shop.org Digital Retail events.

Jessie has served as a qualified subject matter expert on Internet-related litigation matters since 2005 and is one of the original nine cofounders of the Search Engine Marketing Professional Organization (SEMPO). She has been interviewed by numerous trade publications and media outlets for her work on click fraud within the pay-per-click (PPC) industry, including NPR, The BBC, the *New York Times*, the *Wall Street Journal*, *Wired Magazine*, the *Washington Post*, SearchEngineWatch.com, CNET.com, and CNBC.

Colophon

The cover fonts are URW Typewriter and Guardian Sans. The text font is Meridien; the heading font is Akzidenz-Grotesk; and the code font is Dalton Maag’s Ubuntu Mono.