

The background of the cover is a street map with a yellow and orange color scheme. A hand is shown in silhouette, pointing its index finger towards a red circular location marker on the map. The map shows a grid of streets with various labels in Spanish, such as 'BOULEVARD DE MADRID', 'BOULEVARD DE VALLEJO', and 'BOULEVARD DE LOS ANGELES'.

Audience, Relevance, and Search

Targeting Web Audiences
with Relevant Content

James Mathewson
Frank Donatone
Cynthia Fishel

Foreword by Mike Moran
Author of *Do It Wrong Quickly*
Coauthor of *Search Engine Marketing, Inc.*

The authors and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

© Copyright 2010 by International Business Machines Corporation. All rights reserved.

Note to U.S. Government Users: Documentation related to restricted right. Use, duplication, or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corporation.

IBM Press Program Managers: Steve Stansel, Ellice Uffer
Cover design: IBM Corporation

Associate Publisher: Greg Wiegand
Marketing Manager: Kourtnaye Sturgeon
Publicist: Heather Fox
Acquisitions Editor: Bernard Goodwin
Managing Editor: John Fuller
Designer: Alan Clements
Project Editor: Elizabeth Ryan
Copy Editor: Ed Winograd
Indexer: Ted Laux
Compositor: LaserWords
Proofreader: Mike Shelton
Manufacturing Buyer: Dan Uhrig

Publishing as IBM Press

IBM Press offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales, which may include electronic versions and/or custom covers and content particular to your business, training goals, marketing focus, and branding interests. For more information, please contact:

U. S. Corporate and Government Sales
1-800-382-3419
corpsales@pearsontechgroup.com.

For sales outside the U. S., please contact:

International Sales
international@pearson.com

The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both: IBM, the IBM logo, IBM Press, BladeCenter, and DB2. Microsoft, Bing, and Windows Live are trademarks of the Microsoft Corporation in the United States, other countries, or both. Flash is a registered trademark of Adobe Systems Incorporated in the United States, other countries, or both. JavaScript is a trademark of Sun Microsystems, Inc. in the United States, other countries or both. Other company, product, or service names may be trademarks or service marks of their respective owners.

Library of Congress Cataloging-in-Publication Data

Mathewson, James, 1963–

Audience, relevance, and search : targeting Web audiences with relevant content / James Mathewson, Frank Donatone, Cynthia Fishel.

p. cm.

Includes index.

ISBN-13: 978-0-13-700420-1 (pbk. : alk. paper)

ISBN-10: 0-13-700420-6 (pbk. : alk. paper)

1. Online authorship. 2. Web usage mining. 3. Web sites—Design. 4. Web search engines. 5. Internet users. 6. Internet marketing. I. Donatone, Frank. II. Fishel, Cynthia. III. Title.

PN171.O55M38 2010

006.3—dc22

2010006751

All rights reserved. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permissions, write to:

Pearson Education, Inc.
Rights and Contracts Department
501 Boylston Street, Suite 900
Boston, MA 02116
Fax: (617) 671-3447

ISBN-13: 978-0-13-700420-1

ISBN-10: 0-13-700420-6

Text printed in the United States on recycled paper at Courier in Stoughton, Massachusetts.
First printing, April 2010

Foreword

Although in recent years the Internet has been overrun with images, audio, and video, the Web remains, at its heart, a writer's medium.

Why do I say this? Because the written word is at the heart of every Web experience. No matter how much video you watch, images you look at, and audio you listen to, you read a lot more words. While Web sites demand skilled people to create all of these kinds of content, the demand for Web writers has never been greater.

And that demand keeps growing, both because more companies each day are using the Web as a marketing tool, and because messages are virtually unlimited in size. Unlike every other medium, there is no natural limit to the amount of space that can be used for a Web message. Advertising increases in cost as size goes up. Direct mail costs increase for paper and stamps. On the Web, the words can just go on and on with almost no incremental cost—you can always create one more page that explains one more idea.

So, yes, it's a writer's medium. But not just any writer need apply. If you are an accomplished writer, schooled in creating advertising copy, magazine articles, direct mail pieces, newspaper stories, books, or any other kind of printed media, you have the basic skills required to succeed on the Web. But you don't have *all* the skills you need.

The book you hold in your hands can help any experienced writer adapt those skills for Web writing. Web writing is a challenge for even the most gifted print writer, for several reasons.

- **Web writers must be brief.** The act of reading on an electronic device causes readers to skim content rather than read in depth. Even though this user behavior might change over time, as people get more used to reading from a computer screen, the relatively cramped experience of reading from mobile devices will always demand brevity for online writers. While you can always create another page, you must be brief and to the point within each page.
- **Web writers must appeal to search engines.** Because Google and other search engines are so important in driving traffic to Web sites, no writer can afford to

overlook search engines as an audience equally important as Web users. While many elements go into successful search marketing, the words on the page are the most important.

- **Web writers must create pass-alongs.** Web sites have always needed links from other sites, and well-written content makes that happen. Nowadays, social media allows customers to pass along good content to other Web users, using blogs, Twitter, social networks, and many other means. This new stream of minute-by-minute content creates a nearly endless demand for good writers.

If you are feeling a bit overwhelmed at the thought of adapting your writing style for all of these new demands, don't be. This book leads you through these concepts, and many more, in step-by-step fashion. The authors are experts in the practical approaches needed to succeed as a Web writer for any company.

You'll learn everything, from determining the relevance of your message to your audience, through measuring the results of your efforts. You'll especially learn how to use search engines and social media to ensure that your message is seen by the maximum possible audience.

Don't assume that everything you know about writing is wrong when it comes to the Web. It's not. You need not be intimidated from branching out to this new online medium. But neither should you assume that there's nothing to learn. This book will help you leverage the writing skills you've always had to enter the brave new world of Web writing success. So stop reading this foreword and dive into what you really need to know!

—Mike Moran
Author of *Do It Wrong Quickly*
Coauthor of *Search Engine Marketing, Inc.*

Preface

Many books focus on how to publish content on the Web, or how to measure its success, or how to take existing text from print venues and make it suitable for the Web. But no book adequately focuses on creating text exclusively for the Web. Perhaps everyone thinks that the topic of writing has been adequately covered in such works as Strunk and White's excellent little book, *The Elements of Style, Third Edition* (1979). Our view is that this assumption misses an important fact: Writing for the Web is fundamentally different from writing for print. We will unpack this fact to delve into how to write successfully for the Web.

In so doing, we fill an important gap in the literature. Other books about Web publishing, such as Mike Moran and Bill Hunt's book *Search Engine Marketing Inc., Second Edition* (2009), focus on using the Web medium as a marketing tool, but spend only a few pages on the key success factor on the Web: writing. The old saying "content is king" rings ever clearer as the Web evolves. Search engines such as Google, and social media venues such as Facebook, accentuate the need for Web-centric writing. This book helps its readers write effectively for the Web by taking into account how search and social media usage affect readership and audience. This is not just a challenge; it's an opportunity. Search and social media can help Web writers learn common audience attitudes to better engage their audience with relevant content.

The target audience for this book includes writers, editors, and content strategists. Though many of the examples and case studies apply to Web marketing (the field in which the authors have 45 years of accumulated Web experience), the book is intended to have a broader scope than just marketing writing. Writers for blogs, wikis, and various online media outlets can also benefit from the insights contained in this book. Web marketing is just an excellent source of rich Web publishing examples, because it clearly shows how effective writing has changed on the Web.

In print media, readers have already chosen to pick up the publication. This simple act implies a certain level of consent as to its relevance. Print writers can assume that the

reader finds the publication at least nominally relevant and can get on with the business of presenting a compelling flow of information. But on the Web, visitors often don't choose the specific page they land on through search or social media referrals. They must first determine the relevance of the page to their needs. For this reason, Web writers must demonstrate relevance before they can start engaging Web visitors in the flow of information. This step is often missed by Web writers, who wonder why so many Web visitors leave their pages without engaging in the content at all. This book shows writers, editors, and content strategists how to attract a target audience to content that is relevant to them, how to demonstrate this relevance when their target audience arrives on the site, and how to measure the depth of the audience's engagement.

After an introductory chapter, the book spends the next two chapters discussing foundational concepts about how print and Web media differ. It focuses on how relevance determination and audience analysis differ on the Web, as opposed to print. This is necessarily deep stuff, because it depends on a rich body of literature in three different fields. We do not want to present the literature in a breezy fashion, out of respect for the great minds who have studied media determinism, relevance theory, and audience analysis over the centuries. Still, we do our best to make the topics accessible to those without backgrounds in these fields.

After the fundamentals are covered, Chapters 4 and 5 focus on how to write for the Web using a search-first perspective. Search-first writing is based on the premise that search engines provide a lot of insights into what is relevant to your target audience. Unpacking this assumption is central to this book.

This discussion has two major themes. The first theme relates to word choice: If you know what words resonate with your target audience members, you can write more effectively for them. You learn these words through keyword and social media research. The second theme is the subject of Chapters 6 and 7: how links determine the structure of Web information. Search engines use links to determine the relevance and credibility of content on the Web. They then use that information to sort search engine results for their users—the writers' audience. Designing your site (Chapter 6) and collaborating with other sites to enable search engines to determine the relevance and credibility of your content (Chapter 7) are writing tasks unique to Web publishing. Some of the insights developed in these chapters are unique to this book.

The Web is a social medium—ever more so with new applications that connect like-minded people to communities centered on conversations. This is not a new phenomenon: The Web has always had a social element, as the value of Web content is directly proportional to the quantity and quality of links to it. The best way to get links to your content is by collaborating with the community of experts in your topics of interest. This is as true for traditional static Web pages as it is on Facebook. Still, there are aspects of social media writing that differ from traditional Web publishing. Chapter 8 delves into these distinctions.

Unlike print publications, Web sites are never done. The more you change them to accommodate your audience, the more effective they become. Measuring site effectiveness and making intelligent changes to better adapt to audience needs is the subject of Chapter 9.

Though we must call this book done and published, the Web will continue to evolve, requiring our continued updates and new insights. We will add to the gift of knowledge this book represents by maintaining a related Web site (www.writingfordigital.com) containing blog posts around our particular areas of expertise and links to the references you will find in our bibliography.

The Story of This Book

This book had its genesis in my M.S. thesis of the same title in scientific and technical communication at the University of Minnesota. When the thesis was published in August of 2008, I was collaborating with Frank Donatone and Cynthia Fishel on five search engine optimization courses for writers, editors, and content strategists at IBM. Education was part of my role as editor-in-chief of *ibm.com*. Frank and Cynthia brought fresh examples of the best practices to our course development as they consulted with their clients as Web effectiveness leads. The concept of marrying the content in the thesis with the coursework to create a comprehensive book on search-first writing for the Web was mutually agreeable to us.

What we didn't realize when we entered into the coauthoring relationship was how much we would learn along the journey of writing this book. Like a jigsaw puzzle emerging before our eyes, the missing pieces presented themselves as fresh insights amidst the lessons we taught in our search courses. We could not have predicted the almost daily eureka moments as we wrote page after page and chapter after chapter. We hope you enjoy reading this book as much as we enjoyed writing it.

—James Mathewson
Faribault, Minnesota

Writing for Web Users Implies Writing for Search Engines

Writing for the Web is fundamentally different than writing for other media, such as print. Sure, some good writing habits for print also apply to the Web. You should use an engaging tone and fresh word choices. You should organize your information clearly. And, most importantly, you should understand your audiences and write in ways that make it easy for them to understand your content. However, analysis shows that readers approach Web content far differently than print content. This book seeks to use this insight to provide a practical guide for Web writers and content strategists.

This book is about understanding the content needs of Web users to do a better job of presenting relevant content to them. It assumes a good working knowledge of how to write for print and therefore will not delve into the mechanics of quality writing. But it *will* focus on the distinction between how print and Web media differ, which requires some explanation of how the print medium works. From this foundation, we can understand how the Web as a medium differs from print. We can then develop practical guidance on how to do a better job of engaging Web readers.

You might be skeptical about this. Whether for Web or print, text is text, right? In this book, it is our job to counter this skepticism. In subsequent chapters we will cite numerous case studies and deep research into user behavior that clearly demonstrate how Web readers behave and why they do. For the time being, however, we ask that you suspend your skepticism so that we can introduce the content of this book. What follows is a brief sketch of the chapters in this book, which we hope will convince you to read on. We promise that by understanding what is covered in these chapters, you can truly master a field that is crying out for competent practitioners: Web content writing.

How the Web Medium Has Evolved from Its Print Origins

The basic difference between print and Web media is in the reader/writer relationship. In print contexts, you typically invite an audience to journey with you through your prescribed content path. The best print writers encourage their readers to surrender control and let the

writer lead them by the hand through the material. Often, print readers will readily concede this control, trusting that the writer knows how best to organize and present information.

On the Web, readers (if we may call them readers for the present) will not cede control over the information path. They navigate through paths of their choosing, cutting corners and trying to get to the most relevant content as quickly as possible. On the Web, it is the writer's job to provide a multitude of clearly marked paths, letting readers find the relevant nuggets of information that they seek. How to write to let readers sift through your content and find those nuggets is a considerable challenge that deserves a book of its own.

A particularly salient example of how Web writing differs from print is the way Web readers use search engines. Web users are impatient with content providers, because they can be. If they can't find the information they're looking for by navigating to it, they will use search engines. This impatience with information retrieval shows up in their reading habits. As a study of Web users by Weinreich et al. (February 2008) has demonstrated, Web "readers" do much more skimming and scanning than print readers. The study shows that on average, people spend 82% less time actually *reading* Web pages than they do when they read print pages, assuming average print reading speeds of 250 words per minute.

As Jakob Nielsen (June 2008) shows, Web users usually don't read pages in the conventional way, line by line, serially. They scan for keywords in the heading and short descriptions and only read after deciding that some content is relevant. With this in mind, Google has designed its search crawler to mimic how Web users behave. The crawler scans pages for keywords and captures the pages with the strongest placement of those keywords to include in its index. When a user enters a keyword phrase into Google's search field, Google returns the results that its algorithm deems relevant to those search terms. The design of its crawler is one reason that Google has become the search engine market leader in the United States and elsewhere. It tends to return highly relevant results for users, and it displays those results in ways that users can easily digest, given their extreme impatience. Because Google and other search engines strongly cater to Web user behavior, learning to write for the Google algorithm is an essential aspect of writing for Web readers.

Though our book relies on much of the information provided in Mike Moran's and Bill Hunt's excellent book *Search Engine Marketing, Inc., Second Edition* (2009), this is one point where our approach diverges from theirs. Hunt and Moran claim (2005, 309) that "The best philosophy for writing for search is: Write for people first, not for search engines." Our claim is that writing for search engines approximates writing for people. Also, Web writers often lack audience knowledge—readers can come from anywhere using search. Because Web writers often lack audience knowledge, writing for search engines is often the best way to understand how to write for people. So we take a "write-for-search-engines-first" approach.

Writing Relevant Content for the Web Audience

How do you analyze your audience for print publications? Suppose you write for academic periodicals. If so, you have a good sense of the history of debate within each one. And you know that readers of a given periodical are professors or graduate students in the field. Perhaps you have a demographic survey of its subscribers. From this, you form a mental representation of a typical reader (maybe even someone you know, like your advisor), and as you write or review your own work, you imagine that person reading it. In the print world, this is the closest you will ever get to addressing audience members based on known facts about them.

In most print contexts, you know significantly less about your audience than you do in academic periodical contexts. Magazine writers might know rough demographics about subscribers, but they never know who might pick up a given publication at newsstands. Book writers know even less about their audiences. You might write a book for a particular audience, in the sense that you define its topic and purpose so that audience members who are interested in those things will buy it or check it out. But you don't always know exactly how to address them, either as individuals or as a group. Print audiences are typically much more diverse than subscribers to an academic journal. It is simply not possible to address all possible readers with one print product. You just don't know them well enough to do this.

For this reason, many print writers invoke their audiences by using storytelling and other compelling techniques to draw them into the book's world. When readers start down the path of a particular story, they leave all expectations of being addressed for who they are at the trail head and follow the writer into unexplored territory. The more richly the writer creates that territory, the more readers will feel compelled to take the journey.

On the Web, you know a little more about your audiences, but your knowledge is fairly generic: You might know their service providers and perhaps which search engines they came from, and which paths they take through your site. But on an individual level, you don't know much at all. (Unless they sign in to your site. But let's leave those cases behind, since very few users take the time to do so. And even if they do sign in, what you know about them can't help you tailor messages for them.) Because users take unpredictable paths through your information, you can't connect with your audience as you do in print, such as by addressing a tightly defined audience or by appealing to a diverse audience. You have to find some way to connect with them in order to deliver content that they will find relevant. There are no perfect solutions to this problem, but we have developed some strategies and tactics to help you better connect with anonymous audiences on the Web. All of them center on search.

When you write content explicitly for search engines rather than for your users, in a sense, you invoke the search engine users with an effective mix of keywords and links that draws them to your pages. The challenge is to craft your pages in ways that attract specific

users from search engines, especially Google. In so doing, you can present relevant information to your audience. This book is about attracting your audience with keywords and links and thereby providing relevant information to them. As mentioned earlier, because Google and other search engines cater to the way users scan and retrieve content on the Web, writing for Google is also an effective way to write relevant content for your audience.

Discovering and Using Popular Keywords

The first thing to do if you want to optimize your pages for search is to find out what keywords related to your theme or topic are most often searched for. These keywords become your site's nomenclature. If you use these words in prominent ways on pages in your site, you will have a better chance to get traffic.

But traffic volume is not the end game. The end game is *targeted* traffic. You want to engage with your visitors. You want your target audience to come to your site and find that your content is relevant to them. Visitors who find your content irrelevant typically click the back button or “bounce” off your site without clicking any of its links. If you try to get high traffic without taking care to also target your audience in the process, you will get a lot of traffic; but most of it will bounce. What you want, instead, is low bounce rates with relatively high traffic. How can you achieve this? It's not as easy as it might seem. Popular keywords that are used by many people in a variety of contexts will yield mixed results, if all you do is optimize your pages for those keywords. You will get high traffic volume, but also high bounce rates. The first step is to develop a set of related keywords—or a **keyword cloud**—which your target audience uses frequently. Then you need to develop pages that use the words in this cloud.

So how do you develop these keyword clouds? One way is to use keyword research tools to find related keywords. These tools can help you identify not only the most often searched-on words, but also related words and how often they are searched on. Once you get a sense for the number of users who search on a keyword—or its search demand—you can use the most relevant, high-demand words as the building blocks for your content.

When we use the term *keyword*, we do not merely refer just to single words. Most keywords that users enter into search engines consist of *phrases*. A keyword cloud typically contains not only related single words, but also related phrases. Many users search on so-called **long-tail keywords** to zero in on the exact content they are looking for. These are not just longer strings of words and phrases, though they typically are longer than high-demand single words. The phrase **long-tail** refers to the demographics of users who search for very *particular* content, rather than searching on more *generic* topics. Users who enter long-tail queries tend to be more search savvy. If your content ranks well for these, you will attract a highly targeted audience. But no single long tail will garner much traffic. The number of long-tail keywords in your content will need to be enough to drive targeted traffic to your

site. For these keywords, you have to understand the language of your target audience at the *sentence* level, rather than at the *phrase* level. One way to develop this understanding is to research your target audience's social media hangouts, such as blogs, communities, and forums. Because users tend to use the same sentences in their long-tail keyword searches that they use in social media contexts, knowing the writing habits of your target audience will help you know the best long-tail keywords to use.

There are a variety of tools you can use to better understand the writing habits of your target audience. Very effective ways to do this include Google Alerts, Yahoo Alerts, and a method that uses Yahoo Pipes to track mentions of your company via **RSS feeds**—subscription feeds that automatically update when the source content is changed. You can use these methods to guarantee that whenever someone mentions a particular phrase (such as your company name) in a blog post, you get an RSS notification and can look and see what that person has said. This not only helps you get a sense of how users feel about your offerings, but also about what kind of language that blog's readers use for them. Later, we will show how to use free tools like Yahoo Pipes to monitor social media for common keyword-related activity.

Engaging with Web Visitors through More Targeted Search Referrals

Until now, we have focused on using keywords to attract a target audience to your content. But keywords are not the only parts of Web content that determine whether your content is relevant. A user can find the content on a page (with the same keywords) relevant one week and irrelevant the next. You might ask: How can that happen? If keyword usage determines relevance, how can attracting users to your pages through keywords drive users to information that they find irrelevant? Well, language is a lot more complex than creating a simple matching algorithm between keywords and users. If it were that simple, Web writing and editing would be a matter for technology and would not require human decisions. Fortunately, making good content decisions based on a variety of variables, including keyword usage, requires humans.

Suppose that a visitor to one of your pages has viewed all the information on it, but there have been no new updates since then. The content might still be relevant to the visitor's interests, but no longer relevant enough to reread. In a marketing context, users might come to your site one week to see what you are offering, and the next week to see how those offerings fit their needs. Once they are aware of your offerings, if you simply drive them back to the same page through search, you're creating an irrelevant experience—one that could end in a bounce and a bad user experience.

There are many more variables that affect relevance than we have space to list. Those are for linguists to determine, rather than writing instructors. But we can point you to one

important variable beyond keywords that affects whether visitors will find the content on your pages relevant: *purpose*. If you tune your pages to the activities your users hope to accomplish when they arrive on your pages, you have a much better chance of getting them to engage with your content. *Engagement* is one way we measure relevance: If people click a link on your page, we can at least say that link is relevant to them. Sometimes engagement is merely a question of reading the content. Sometimes it's clicking a link. Sometimes it's getting users to comment or fill out a form. There are countless calls to action that a Web site can have—another key difference between print and Web. In print, you simply want to get your reader to read and comprehend your information. Perhaps you want the reader to be entertained or merely informed. However, you never want the reader to interact with a book—to write in the book in hopes that you will write back.

But on the Web, engagement or interaction is typically exactly what we want users to do. If all we want them to do is read and comprehend, we can provide PDFs for printing and reading offline. But good Web content is interactive. It compels the visitor to take action. So, in addition to tuning your page's content for keywords, you also must tune it for the interactive purpose of that page. The real trick is mapping the keyword phrase to the page's purpose. This isn't as hard as it might seem. It is merely a question of adding a verb to your keyword phrase that describes what you want users to do when they get to a particular page. For example, one page can have keyword phrases with *learn* in them, another page might have keyword phrases with *shop* in them.

The question is: How do we help users land on one of our pages that is relevant to a keyword and related to their purpose in seeking the information? We will attempt to give some answers to this question in this book. But a word of caution: There are very few general answers. Different environments demand different answers to complex search questions. You will never achieve perfect engagement levels; because of the complexity of Web interactions, some visitors will bounce. But we can give you a framework for answering the question, and improving your engagement rates with writing that is more focused on the purpose or user goal of the page.

Developing an Optimized Site Architecture

We think of print publications as mostly self-contained units. Sure, we refer to related works in the bibliography or source list of a publication. But we expect print publications to be consumed whole. This is another key difference between print and the Web. Web users do a lot more skimming and scanning than print users. Only after they determine that the information matches what they're looking for do they bother to read. This is a central insight of this book, and it affects every aspect of Web publishing, including design and architecture.

For our purposes, **Web architecture** is the practice of designing information experiences that help users find the information they're looking for. In a sense, writing for search engines is part of this practice. If you write in a way that helps users find information more easily through search engines than if they navigated to it from a home page, you are approaching an optimal user experience. But search is not enough, either; you also need to *design for navigation*. The goal is to create engagement, and in some cases, conversions. It's not enough to get a visitor to click a link; that click should land the user on a relevant page. Search can draw users to lower-level pages and encourage them to navigate up; or it can draw users to higher-level pages and enable them to navigate down. And horizontal navigation is also part of a good architectural plan. In any event, the content experience doesn't end with getting the user to come and click on something; you must get the user to engage with it.

Many architectural discussions at IBM and other companies focus on designing a hierarchy of pages that enables users to easily move from the top level to the specific information they're looking for. This is a good approach if you are designing a user experience for navigation. But it leaves out key considerations that can help users find information from search. Some search engines rely on **metadata**—extra-linguistic information hidden from view in the code of pages—to help determine the relevance of those pages. Though Google does not use metadata as part of its ranking, it does analyze how pages are interlinked with the rest of the Web to help determine relevance. Thus, architecting a set of pages for search engines necessarily includes paying attention to metadata and linking.

In this book, we focus on Google as the most popular search engine in the United States and elsewhere. From an architectural perspective, this means a thorough discussion of linking. This is the subject of an entire chapter later in the book.

Before we do that, we will first discuss how the standard practice of designing a hierarchy of related pages around a central topic or theme relates to keyword usage for search engines. To reach a more general audience, most architects design information around topics starting at a high level. This architecture should help users drill down into the information in the hierarchy, according to their individual needs. To enable this experience, you should choose different related keywords for each page in the topic, following the hierarchy. For example, on the IBM Web site, we might start with a high-level page on IBM Servers and then develop pages related to a specific product line, like the BladeCenter and other IBM offerings within this topic, such as BladeCenter hardware (Figure 1.1).

You might think this is a fairly straightforward process. The architect designs the hierarchy of pages, and the writer picks keywords to fit into it. We suggest that this process rarely works the way it is drawn up on paper. At IBM, we struggle with pages in a hierarchy that do not produce search referrals because users simply do not search on the keywords we chose for those pages. Imagine a hierarchy of pages in which the third level down the tree is



Figure 1.1 A hierarchy of pages in ibm.com.

the highest ranking page because it uses a popular keyword, yet the top-level page in this hierarchy gets little or no search engine traffic because few users ever search on its keyword.

We suggest (and will demonstrate later in the book) that the best practice is to optimize the top pages of the hierarchy with the most competitive keywords—the ones that draw the broadest audience. Pages lower in the hierarchy need to be optimized for narrower, more targeted audiences, who typically use long-tail keywords. You do this by choosing related keywords that are more likely to appeal to specific segments of the broader audience. The point of our chapter on the relationship between architecture and writing is that *architecture reflects writing practices*: When we create search-first architectures, we do a better job of creating an information experience for users for the whole hierarchy, not just for isolated pages within it.

The goal of information architecture is to serve users with relevant information. But how do you know what information is relevant to them? We suggest that the search-first architecture does a good job of creating relevant information for a large set of users, with

less guesswork and less trial-and-error than common architecture processes. Keyword data is the best information available from which to design your information. It not only helps you isolate specific keywords that will draw higher volumes of users to specific pages, but it helps you understand how keywords and phrases are related to one another. If you optimize a hierarchy of pages for a set of related keywords, you not only direct users to specific pages they might be interested in; you can also get them to navigate to other pages in the hierarchy from their initial search referral. If they find the top-level pages relevant, chances are they will find pages targeted toward more narrow audiences even more relevant.

In large organizations with a complex matrix of Web sites, content creators can unwittingly compete for the same keywords, thus harming visitors' overall experience. For this reason, large organizations need to manage keywords across their whole site, not just within specific areas.

Suppose you own a part of a company's Web content—perhaps the marketing pages related to a portion of your company's product portfolio. You do all the keyword research related to that portfolio, develop a site architecture that maps pages to popular keywords and desired visitor interaction, and write optimized content for each page in the architecture. In short, you do everything right to attract the most targeted audience possible to your pages. However, suppose that a colleague owns the Web content for a related set of offerings in your company's portfolio. She does everything you do and optimizes her pages for the same keywords, all the way down to the long-tail ones. You are now unwittingly competing with your colleague for the same targeted audience.

As unlikely as this scenario is, competition is actually quite common in a company such as IBM that has a large and diverse portfolio of offerings, and many Web pages related to them. Even for themes such as Green IT, several efforts might spring up at the same time and could unwittingly compete with one another for the same users, unless these efforts are coordinated. For this reason, we recommend corporate-wide keyword management systems, which enable content owners to reserve specific keywords for specific pages. These systems can spread keyword usage across an enterprise in way that is similar to, but more pervasive than, what you do when you develop a keyword-based architecture around your theme or topic. With such a tool in place, you can optimize your enterprise for popular keywords and attract targeted audiences to the most important pages for your business.

Gaining Credibility through PageRank

As important as keyword usage is for your search efforts, it is less than half of the Google algorithm. The Google algorithm takes two primary things into account: **relevance**, which is a function of how keywords are used on a page; and **PageRank**, which is a function of how your page is interwoven into the Web's vast map of links. (There are other factors as well, such as the prominence of your site in the scheme of the Web, but those are beyond

your influence as a Web writer. In this book, we will focus on keyword relevance and PageRank.) You can do everything right to improve the relevance of your pages by using keywords, yet still fail to get your page listed on the first page of Google results. If a particular keyword is highly competitive, meaning that lots of other sites are doing everything in their power to rank high for it, you will also need to improve your PageRank to get listed on that first page of Google, Bing, and other search engines.

For example, the keyword *Service Oriented Architecture*, or *SOA*, is highly competitive in Google—a lot of very smart companies spend a lot of money developing optimized pages around a cloud of keywords related to *SOA*. The only way to get on the first page of results in Google for *SOA* is to get credible pages around the Web to link to your *SOA* page. Google counts each link to your page as a vote of confidence for the content on your page. This is how Google overcomes the problem of relevance on the Web. Because the text alone cannot determine whether users will find the content relevant to them, and there are no better contextual cues to content than links, Google uses linking as its main contextual cue. All evidence indicates that Bing also uses links, and, if anything, gives them even greater weight than does Google.

As complicated as keyword usage can get, linking is much more complex. Search engines don't count every link to your pages as equally valuable. Some links get more **link juice**, or value, than others, depending on how valuable and relevant Google deems the site that links to your page.

We will explain some of this complexity later in the book. For now, suffice it to say that you need to develop a plan to promote your pages to high-value sites. Many sites have value simply because they serve as **link aggregators** related to a given topic. One example is Wikipedia: It contains relatively little original content, but it is an excellent place to begin researching a topic, because every source is a link to further research. (If you attempt to publish original content on Wikipedia in the sense that it lacks original sources, the content gets flagged as in need of support.) However, for technical reasons to be explained later, Wikipedia is not a great site from which to get link juice. It merely illustrates the point that sites like it, which link to deeper dives into research topics, are highly valued by users. And because users value them highly, search engines do as well. For this reason, these hubs of authority typically find themselves listed near the top of search results.

Our approach is to determine the best path to becoming a **hub of authority** on your topic. This requires a lot of deep thinking about the state of the art of your topic and a lot of research into the existing hubs of authority on it. How does your content fit into the matrix of authority on your topic? What contribution does your content make to the interlinked conversation related to your topic? How do you engage with other subject matter experts in your field to promote this contribution to the discussion? We will explore these and other questions to help you improve your PageRank in Google. And who knows, if you develop a good relationship with the recognized authorities in your field and your content makes a

unique contribution to the field, you too can become a hub of authority and thus rank highly in search engines.

There are no shortcuts to becoming a hub of authority. To do so, you must gain the trust of other authorities for your topic. That takes time. Still, there are some tactics that can help you promote your content to authority sites. We will cover them in Chapter 8.

Capturing Social- and Rich-Media Opportunities

Print habits die hard. Chief among them is the habit of wanting to control the conversation. When you write for print, it's your tale and you control the telling—you control how the reader consumes the information. But when you write for the Web, the reader controls the pace and flow of information. Trying to control it is a surefire way to get users to bounce off your pages. The best you can do is to give users options they will want to choose. Part of gaining their trust to choose your content options—to click your links—is making it clear that it's their choice. Part of this is demonstrating that your content is not isolated, that it doesn't claim to have all the answers, and that it is but a small part of a bigger conversation.

The Web is an evolving medium, and users' expectations evolve with it. At this time, the fastest growing practice on the Web is sometimes called **social media**. Loosely defined, it is a set of practices that engage Web users to participate in the conversation rather than merely consume static information. These practices include blogs, wikis, forums, persona sites such as Facebook, microblogging sites such as Twitter, and the like. Social media is rapidly evolving to include graphics and video sharing and sophisticated syndication.

Nowhere is the Web more distinct from print publishing than in social media contexts. Users control the flow of information as they navigate through static Web pages; but they actually *contribute* to information in social media contexts. Here, their control of the information is complete. In extreme cases, users who make frequent comments on a blog are as important to the blog's success as its author. And a blog is not much more than an online opinion column if no one comments on its entries.

As the Web becomes more of a collaborative medium—more of a space for symposiums rather than lectures—users' expectations for the whole Web change. Even publishers of static Web content need to adjust their practices to engage users in conversations. Whether your site explicitly enables users to comment on its content or not, users expect it to. Social media has accelerated the need to give users more control over their information paths, even if your content is not intended to be shared and commented on.

For example, at one time IBM had a number of Web producers who preferred to get permission to republish content from other sources on their sites, rather than to link directly to the sources of that information. Fortunately, across the industry, this practice has long ago become a distant memory of how things used to be done in Web publishing. That memory is all that remains of writing habits that stem from print: "Let's keep users on our site.

Let's not let them leave." This attitude is self-defeating on the Web. You lose users' trust when you try to control their experiences; and you lose potential PageRank, as well.

Because page rankings are a large part of Google's PageRank (and of Bing's similar algorithm) and are determined by how a community "votes" on the credibility of content, search is a social medium. Not surprisingly, improving your standing in the community by engaging users in social media contexts is a very effective way to gain credibility and PageRank. If you encourage users to pass your links, podcasts, and videos around to their friends, the PageRank for the pages on which those assets are accessed will grow virally. Since users expect this pass-along content to be ever more visual, search effectiveness is not just about text anymore. It is crucially about rich media as well.

The irony is that the more control you give users, the more they will want to visit your site. The more open you are in your reciprocal linking relationships to and from rich media on other sites, the more users will treat your site as a hub of authority and will keep coming back to hang out and explore new paths through your content. Later on, we will demonstrate in more detail how to take advantage of the social attitudes and habits of Web users.

This letting go of control of your information is a necessary, though sometimes painful, cultural shift in Web content practices. Because this shift is so necessary and fundamental to successful Web content efforts, we believe it deserves a chapter all its own.

Measuring Web Effectiveness

How do you know how well you are doing on Google? That is a rather easy question: You simply search on your keywords and see where your page ends up in the results. How do you know the volume of visitors you get from Google? This question is a little harder: You use a Web analytics tool such as Unica Netinsight to find out where your visitors are coming from and then filter the results to show only those who come from Google. You can also run reports in modern Web analytics tools that show what keywords brought users to your pages, and in what volume.

The more detailed you want to get in understanding the effectiveness of your search efforts, the harder the questions become. You can get a lot of information related to data such as Google ranking, Google referrals, raw traffic, no-results keywords, click-throughs, and customer feedback specific to search, but correlating that data is quite difficult. The goal of this kind of research is to measure engagement: How many users are doing the things you want them to do on a page, such as downloading a white paper, filling a survey, taking a poll, or making a comment? We will discuss a best practice in gathering and correlating the various search effectiveness later in the book.

For the purpose of this discussion, suffice it here to say that how you define and measure search effectiveness is a fundamental facet of using search to target audiences with

content. Search efforts rarely work exactly as they were drawn up on paper. More often than not, you will find aspects of your search efforts that underperform expectations. In these cases, you will have to go back to your pages and fine-tune your content to better target your desired audience. If you find a high bounce rate, for example, you might need to change the language on the page to develop better engagement with your target audience. The practices outlined above and detailed later in this book can help you do a better job of search optimization when you first develop and publish content. But language on the Web is a complex medium, and you can expect to have to do a certain amount of reworking to achieve better engagement. For this reason, measurement is a critical phase in search optimization work.

Also, to compete for users, search engines are evolving rapidly. If an engine changes after you publish information, it might make sense to go back and tweak your pages in ways that can help you rank better for it. To do this, you need analytics tools and best practices. We will describe how to use available Web analytics tools to measure search effectiveness and develop action items to improve existing content in the process.

I N D E X

- Abstract thought, 16
Acrolinx IQ tool, 157
Acronyms, 68
Adaptive targeting, 159
AdSense tool, 157
Advanced Keyword Research tool, 54, 56–57
Advanced Search screen, 47
AdWords
 competition keywords, 65
 international
 keywords, 74
 keyword discovery, 53–55, 61
 social keyword, 91
Aggregators, link, 10
AI functions, 110
Albertus Magnus, 19
Alerts tool, 5
Alexandrou, Marios, 72
Alta Vista Babel Fish, 72
Alterian SM2 tool, 91–92, 159
Alumni organizations, 122
Analytics tool, 156–157
Anderson, Chris, 82
Aquinas, Thomas, 19
Architecture. See Search-optimized site architecture
“As We May Think” (Bush), 22
At signs (@) for Twitter, 141
Audience. See Target audience
Authoritative information, links from, 113
Automatic translation tools, 72
Autonomy Optimost Adaptive Targeting, 159
Autoping function, 138
Awareness as purpose, 76–78
Babel Fish tool, 72
Back button usage, 25, 35–36, 150
Bacon, Roger, 19
Bad neighborhoods, 118
Baidu search engine, 34
Basic survey module, 80
Berners-Lee, Tim, 23, 125, 135–136
Bing search engine, 34
 anchor text, 128
 link emphasis, 110–111
 PageRank, 10, 12
bit.ly tool, 141
BladeCenter hardware, 7–8
Blogger site
 keyword research, 91
 listening tools, 159
Blogs
 for audience analysis, 81
 for linking relationships, 123
 optimizing, 138–139
 overview, 136–140
 post headlines in, 131
Boardreader tool, 91
Bolding for emphasis, 70
Bookmarks in social media, 89
Bounce rate as engagement indicator, 150
Bush, Vannevar, 22–23
Cailliau, Robert, 23
Case studies
 engagement/relevance relationship, 153

- Case studies (*continued*)
 new product releases, 120–122
 search effort
 improvements, 51–53
 site redesign, 99–102
- CERN, 23, 135
- Characteristics, audience, 44–46
- Chatting factor in
 engagement, 160
- Class CIP addresses, 118
- Cleverness limitations, 26
- Clusters, 92–93
- Clusty tool, 91–93
- Cognitive terms, 36
- Collecting factor in
 engagement, 160
- Comments on blogs, 137–138
- Commitment for social
 media, 140, 144
- Competition
 in audience analysis, 82
 keyword research, 65
- ComputerUser magazine, 115
- Concise writing for Twitter, 141
- Consideration as purpose, 76–78
- Content
 behind firewalls, 116
 free and open, 116
 freshness, 79, 137–138
 management systems, 106–107
 vs. purpose, 77
 relevant. *See* Relevant content
 search-optimized site
 architecture, 103
 transparency, 110, 131–132
- Content experiences in
 design, 100
- Content plans, 97–99, 103–104, 106
- Content quality metrics tools, 157
- Context
 keywords, 109
 in relevance, 33, 148–149
 in translations, 72
- Contextual effects, 36–39
- Controversy on blogs, 137
- Converseon service, 159
- Copyright issues, 115–116
- Core terms, 61
- Coremetrics tools, 156
- Covario tool, 53, 157
- Crawlers
 and Flash content, 100
 navigation behavior, 96
 and page optimization, 66
 page scanning by, 2
 site maps, 105–106
- CRC (Customer Relationship Management) analysis, 81
- Creating factor in
 engagement, 160
- Creative Commons licenses, 116
- Credibility
 earning, 26–27
 on forums, 134–135
 PageRank for, 9–11, 113
- Critiquing factor in
 engagement, 160
- Customer Relationship Management (CRM) analysis, 81
- deGeyter, Stoney, 84–85
- Delicious.com
 linking relationships, 123
 social tagging, 128–129
 tagging and bookmarking sites, 91
- Density, keywords, 68–69
- Description meta tags, 151–152
- Descriptive headings in
 social media, 131
- Descriptive link text, 69
- Design
 navigation, 7, 96–97
 redesign, 99–102
 visual, 97
- Digg.com
 tagging and bookmarking sites, 91
 user ratings, 129–130
- Diggs, 129–130
- Discovery, keywords, 4–5, 50–61
- Diversity in audience, 45
- Donatone, Frank, 142
- doshdosh.com site, 82
- Duplicate content, 106
- Dylan, Bob, 16
- Edutainment format, 132
- Effectiveness measures, 12–13, 147–148
 audience semantics and sentiment, 157–161
 engagement, 148–157

- Effort in contextual effects, 38
- 80/20 rule, 150
- Elements of Style (Strunk and White), 141
- Elliance.com, 63
- Emphasis, bolding for, 70
- Encyclopedia Britannica, 116
- Engagement
 - in audience analysis, 81
 - as audience characteristic, 45
 - audience semantics and sentiment, 157–161
 - measuring, 148–157
 - page purpose for, 76–79
 - PostRank for, 160
 - relevance relationship, 153
 - targeted search referrals for, 5–6
- Engagement rate metric, 152–153
- Engelbart, Douglas, 23
- epinion site
 - keyword research, 91
 - listening tools for, 159
- eResponder tool, 159
- Etsy.com site, 82
- Evergreen content, 79
- Exclusions in searches, 47
- Extent conditions, 36, 38
- External referrals as engagement indicator, 150
- Facebook
 - for keyword research, 91
 - linking relationships, 123
 - listening tools for, 159
 - overview, 142–144
- Failures, search, 35–36
- Fiction, 20
- Field studies, 81
- File sharing sites, 125
- Filtrbox tool, 91
- Firefox SearchStatus plug-in, 124
- Firewalls, content behind, 115–116
- Fishkin, Rand, 111–112
- Flash content and crawlers, 100
- flickr.com, 91
- Flow diagrams in task analysis, 97–98
- Flow of information, 11–12
- Focus groups, 81
- Forums, 134–135
- Free and open content, 116
- Free Keyword List Generator, 53
- Fresh content, 79, 137–138
- Gaps in subject matter, blogs for, 137
- Gender differences in translations, 72–73
- Gerund keyword forms, 69
- Google
 - anchor text, 128
 - keyword indexes, 46–47
 - PageRank. *See* PageRank ranking algorithm, semantics in, 83
 - relevance on, 34
 - SERP page listing criteria, 111
 - site maps, 105–106
 - Google AdSense tool, 157
 - Google AdWords tool
 - competition keywords, 65
 - international keywords, 74
 - keyword discovery, 53–55, 61
 - social keywords, 91
 - Google Alerts tool, 5
 - Google Analytics tool, 156–157
 - Google Blog Search site, 138
 - Google crawler
 - and Flash content, 100
 - navigation behavior, 96
 - and page optimization, 66
 - page scanning by, 2
 - site maps, 105–106
 - Google Insights, 54–56
 - Google Knol, 91, 159
 - Google Sitemap tool, 105
 - Google Traffic Estimator tool, 74
 - Governance system for search-optimized information, 106–108
 - Grice, Paul, 37
 - Grocery store analogy, 103–105
 - Hashtags for Twitter, 141
 - Headings in social media, 131
 - Headlines for blog posts, 131
 - Heat maps, 154–155
 - Hierarchy of pages, 7–9
 - High-demand keywords, 104
 - High keyword priority, 65
 - Homeric epic poems, 15

- Homographs, 48–49
- Homonyms, 48–49
- Homophones, 48–49
- Hubs of authority, 10–12
 - as goal, 27
 - PageRank for, 113–117
 - Wikipedia as, 123
- Hunt, Bill, 2
- Hypertext
 - development of, 23
 - hypertext-matching analysis, 111–112
- IBM Smarter Planet site
 - engagement/relevance relationship, 153
 - flow chart, 97–98
 - redesign, 99–102
 - search efforts, 51–53
- Ice Rocket tool, 91
- IDOL (Intelligent Data Operating Layer), 159
- Impatience, user, 2
- Indexes
 - Google, 46–47
 - latent semantic indexing, 85–86, 158
 - Scholastic libraries, 19
- Information Architecture (IA)
 - clear, 96–99
 - developing, 6–9, 103–106
 - goal, 8
 - governance system, 106–108
 - Smarter Planet Site redesign, 99–102
- Information retrieval basics, 24–25
- Intellectual property, 115
- Intelligence as audience characteristic, 45
- Intelligent Data Operating Layer (IDOL), 159
- Interest as purpose, 77
- Internal links
 - PageRank effect, 96
 - search-optimized site architecture, 105
 - site maps, 105
- Internal referrals as engagement indicator, 150
- International keyword research, 71–74
- JavaScript navigation, 100–102
- Keyword clouds, 4, 61–62, 69, 90
- Keyword Density and Prominence Analyzer tool, 54, 58
- Keyword Discovery Tool, 53–54, 56
- Keyword tags, 91, 141
- Keywords
 - context, 109
 - density, 68–69
 - discovery, 4–5, 50–61
 - as engagement indicator, 151
 - governance systems, 106
 - international, 71–74
 - linguistic considerations, 46–50
 - for long tails optimization, 82–89
 - proximity, 69
 - purpose in, 76
 - research on, 61–65, 82–89, 91–92
 - search-optimized site architecture, 103–105
 - Smarter Planet Site redesign, 100
 - social media, 91–92, 131
 - social tags as, 128–129
 - spamming, 117
 - for target audience, 43–46
 - user scanning of, 2
 - for Web page optimization, 66–71
- KingPing site, 138
- Kleinberg, Jon, 110
- Knol, 91, 159
- Latent semantic indexing (LSI), 85–86, 158
- “Latest Headlines—Your Vote Counts” study, 132
- Lexical FreeNet tool, 88
- Libraries, Scholastic, 19
- Linguistic considerations
 - keywords, 46–50
 - in relevance, 37–40
- Link aggregators, 10
- Link bait, 114–115, 129, 131
- Link farms, 118
- Link text for keywords, 69
- Linkbacks on blogs, 138
- LinkedIn, 142–144
- Links and link juice, 10
 - new product releases, 120–122
 - obtaining, 113–117
 - PageRank effect, 96

- relationships for, 118–123
- search engine attention to, 110–113
- search-optimized site architecture, 104–105
- site maps, 105
- Smarter Planet Site redesign, 100
- social media, 123–125, 129
- text, 69, 120
- in Twitter, 142
- warnings, 117–118
- Web for, 125–126
- Wikipedia, 136
- Listening in social media, 133, 158–159
- Literacy, 16
- Literal translations, 72
- Long-tail keywords, 4–5, 46
 - optimization, 82–89
 - search-optimized site architecture, 104
- Low keyword priority, 65
- LSI (latent semantic indexing), 85–86, 158
- LSI Insights tool, 157–158
- Lurking on forums, 134
- Magnolia tool, 91
- Maps, site, 105–106
- Marketing new product releases, 120–122
- Mathewson, James, 115, 134
- McGee, Matt, 86–87
- McLuhan, Marshall, 15, 127
- Meaning as use, 25–26
- Measuring effectiveness, 12–13, 147–148
 - audience semantics and sentiment, 157–161
 - engagement, 148–157
- Media determinism, 15–17
- Medium keyword priority, 65
- Memex system, 22–23
- Message board threads for audience analysis, 81
- Meta tags, description, 151
- Metadata
 - search engine reliance on, 7
 - spamming, 110
- Microblogging sites, 81
- Mining social media, 159
- Moderated forums, 134
- Moran, Mike, 2
- Mouse-over navigation, 100–101
- Navigation
 - design for, 7, 96–97
 - mouse-over, 100–101
- Nelson, Ted, 23
- NetInsight tool, 12, 147–148, 157
- New product releases, 120–122
- New visitors as engagement indicator, 149
- Niches, blog, 137
- Nielsen, Jakob
 - relevance, 149
 - user behavior, 2, 25, 95
 - Web usability, 96
- nofollow links, 123–124
- Noise on Twitter, 140
- Notes in social media, 89
- Notifications, RSS, 5
- Noun keyword forms, 69
- Offsite engagement measurements, 160
- Off-topic forum posts, 134
- Omigly tool, 91
- Omniture Discover tool, 156
- Omniture path capture package, 155
- On-page factors
 - hypertext-matching analysis, 111
 - limitations, 113
- 140-character Twitter limit, 140
- Ong, Walter, 15
- Onsite engagement measurements, 160
- Optimization
 - blogs for search, 138
 - long tails, 82–89
 - site. See Search-optimized site architecture
 - Web pages, keywords for, 66–71
- Oral contexts in relevance, 33
- Oral tradition, 15–16
- Pace of information, 11–12
- Page-level keyword research, 51
- Page views metric, 149
- PageRank
 - credibility through, 9–11
 - for hubs of authority, 113–117
 - internal links effect, 96

- PageRank (continued)
 links for. See links and link juice
 overview, 109–113
 sitemaps, 105
 and social media, 12
- Pages
 audits, 151–152
 hierarchy, 7–9
 optimizing, 66–71
 perpetually unfinished, 27–28
 purpose, 76–79
- Paid placements, 65
- Paper, invention of, 20
- Partners, links, 122
- Pathing data, 155–156
- Pay-per-click (PPC) tool, 61
- Permutations of keyword phrases, 49
- Perpetually unfinished Web pages, 27–28
- Personality for blogs, 137
- Phrases, keyword, 4–5
- Pike, Walter, 130–131
- Ping-o-Matic tool, 138–139
- Pingbacks on blogs, 138
- Pinging blogs, 138–139
- Polarity of threads, 158
- Polls for audience analysis, 81
- Ponzi schemes, link-equity, 117
- Popular discovering, keywords, 4–5
- Post-publishing audience analysis, 81–82
- PostRank metric, 160
- PPC (pay-per-click) tool, 61
- Pragmatics, 36
- Pre-publishing audience analysis, 81
- Present perfect keyword forms, 69
- Print media
 audiences for, 3
 contexts in, 33, 109
 evolution from, 1–2
 limitations, 22–23
 media determinism, 15–17
 spaces between words, 17–22
 vs. Web content, 23–24
- Prioritizing keywords, 65
- Professional contacts on LinkedIn, 143
- Professional groups, 122
- Proximity, keyword, 69
- Purchase as purpose, 77
- Purpose
 audience engagement, 76–79
 considerations, 77–78
 in relevance, 6, 75
- Qualified visitors, 70
- Quintura tool, 87–88, 107
- Quotation marks (“”) in searches, 47
- Ranks.nl, 58
- “Rate This Page” surveys, 76
- Raw traffic data metric, 150
- Readership studies, 81
- Really Simple Syndication (RSS) feeds, 5
 for audience analysis, 81
 for blogs, 137–139
 purpose, 125–126
- Reciprocal links, 117–118
- ReddIt tool, 91
- Redesign of Smarter Planet Site, 99–102
- Referral data
 audience analysis, 81
 as engagement indicator, 150–151
 tools for, 159
- Related words for keywords, 69
- Relationships
 engagement/relevance, 153
 link, 118–125
 text, 112
- Releases, new product, 120–122
- Relevance
 for audiences, 79–82
 content. See Relevant content
 crawlers for, 66
 engagement relationship, 153
 inferring, 148–157
 linguistics of, 37–40
 PageRank, 9–11
 purpose in, 6, 75
 search-optimized site architecture, 103
- Relevance (Sperber and Wilson), 36–39
- Relevance survey module, 80
- Relevant content, 31–32
 approach for, 39–40

- audience considerations, 32–33
- determining, 33–34
- for search, 34–39
- writing, 3–4
- Remix feature, 92
- Renaissance, 19
- Repeat visitors metric, 149
- Reputation on forums, 135
- Restrictions on content, 116
- Retrievability of social media, 128–129
- Retweets, 142
- Reverse IP look-up function, 81
- Rhetlist forum, 134
- Rich media opportunities, 11–12
- Root words, 61
- RSS (Really Simple Syndication) feeds, 5
 - for audience analysis, 81
 - for blogs, 137–139
 - purpose, 125–126
- Saenger, Paul, 18–19
- Saleem, Muhammad, 131–132
- Scams, link, 117–118
- Scanning on Web, 24–25
- Scholastic period, 17–19
- Scotus, Duns, 19
- Search Engine Marketing, Inc. (Moran and Hunt), 2
- Search engine results page (SERP), 67–68
 - description meta tags for, 152
 - page listing criteria, 111
- Search metrics tools, 157
- Search-optimized site architecture, 95–96
 - developing, 6–9, 103–106
 - governance system, 106–108
 - information architecture, 96–99
 - redesign, 99–102
- Search referrals
 - as engagement indicator, 151
 - targeted. See Targeted search referrals
- Search-relevant content, 34–39
- SearchStatus plug-in, 124
- Seed words, 61–63, 65, 69–70, 72, 90–91
- Self reference in Twitter, 142
- Semantics
 - audience, 157–161
 - in Google ranking algorithm, 83
- Sentence level, keywords at, 5
- Sentiment, audience, 157–161
- SERP (search engine results page), 67–68
 - description meta tags for, 152
 - page listing criteria, 111
- Service as purpose, 77
- Sessions metric, 149
- Signal-to-noise ratio in Twitter, 140
- SIGs (Special Interest Groups), 134
- Site architecture. See Search-optimized site architecture
- Site Catalyst tool, 156
- Site maps, 105–106
- Site overlays, 154–155
- Site traffic
 - for audience analysis, 81
 - as engagement indicator, 149–150
- Sitemap protocol, 105
- Skeptical Web readers, writing for, 26–27
- Skepticism, 45
- SM2 tool, 91–92, 159
- Smarter Planet site
 - engagement/relevance relationship, 153
 - flow chart, 97–98
 - redesign, 99–102
 - search efforts, 51–53
- Social measurements, Webtrends, 156
- Social media
 - for audience analysis, 81
 - for audience understanding, 89–93
 - blogs, 136–140
 - Facebook and LinkedIn, 142–144
 - file sharing sites, 125
 - forums, 134–135
 - linking relationships, 123–125
 - listening tools, 158–159
 - opportunities, 11–12, 127–133
 - retrievability, 128–129
 - sustainable practices, 144

- Social media (continued)
- Twitter, 140–142
 - wikis, 135–136
- Spaces between words, 17–22
- Spam bots, 138
- Spamming
- keywords, 117
 - metadata, 110
- Special Interest Groups (SIGs), 134
- Sperber, Dan, 36–39
- Spiders and page optimization, 66
- Stemming, 69, 72
- Stenotype, 22
- Strunk, William, Jr., 141
- StumbleUpon tool, 91
 - linking relationships, 123
 - listening tools, 159
- Subscriptions for RSS feeds, 5, 125–126
- Success, search, 35–36
- Support as purpose, 77
- SurfAid Analytics tools, 156
- Surveys
- for audience understanding, 80–82, 160
 - Rate This Page, 76
- Swapping links, 117–118
- Synonyms
- international, 74
 - keywords, 69
- Tag clouds, 87–88, 90–91
- Tagging
- for blogs, 139
 - social, 89–91, 128–129
- Target audience
- common characteristics, 44–46
 - defining, 43–46
 - keyword research, 61–65
 - keywords used by, 50–61
 - linguistic considerations, 46–50
 - page optimization, 66–71
 - page purpose for, 76–79
 - relevance for, 32–33, 79–82
 - semantics and sentiment, 157–161
 - social media for, 89–93
 - vocabulary, 133
- Targeted search referrals, 5–6, 75
- long tails optimization, 82–89
 - page purpose for audience engagement, 76–79
 - relevance for audiences, 79–82
 - social media, 89–93
- Targeted traffic
- as engagement indicator, 150
 - vs. traffic volume, 4
- Targeting, adaptive, 159
- Task analysis
- navigation design, 97
 - page purpose, 76
- Technorati tool, 91
- Test & Target tool, 157
- Text
- anchor, 128
 - links, 69, 120
 - relationships, 112
- TextSTAT tool, 91–92
- Themes, 61
- Time challenged audiences, 45–46
- Time on page metric, 154
- Timeliness in Twitter, 142
- Topics, 61
- Trackbacks on blogs, 138
- Traffic
- for audience analysis, 81
 - as engagement indicator, 149–150
 - volume vs. targeted, 4
- Traffic Estimator tool, 74
- Transparency of content, 110, 131–132
- Trellian Keyword Discovery tool, 57
- Trust
- earning, 26–27
- [EM]in Facebook and LinkedIn, 143
- Twing tool, 91
- Twitter
- for audience analysis, 81
 - character limits, 128
 - Facebook and LinkedIn connections, 143
 - for keyword research, 91
 - for linking relationships, 123, 125
 - listening tools for, 159
 - overview, 140–142
 - writing for, 133
- UIMA framework, 110
- Unica NetInsight tool, 12, 147–148, 157

- Unica path capture package, 155
- Unicode, 72
- Unique visits metric, 149
- Uniqueness as audience characteristic, 46
- Updates for social media, 144
- URL shorteners for Twitter, 141
- Usability, Web, 96–97
- Use, meaning as, 25–26
- Use cases, 76
- User groups, 134
- User ratings in social media, 129–130
- UTF-8 characters, 72

- Virtual worlds, 81
- Visits metric, 149
- Visitstat tool, 157
- Visual design, 97
- Vocabulary, audience, 133
- Votes of confidence
 - links as, 113–114
 - social media, 129

- Wall, Aaron
 - Advanced Keyword Research tool, 53, 56–57
 - Free Keyword List Generator, 53
 - Google ranking algorithm, 83
- Web
 - architecture, 7
 - information retrieval basics, 24–25
 - for links, 125–126
 - meaning as use on, 25–26
 - origins, 22–23
 - pages. See Pages
 - vs. print media, 23–24
 - usability, 96–97
- Web Analytics Association, 156
- Web analytics tools, 156
- Web effectiveness, measuring, 12–13
- Webmetrics tool, 157
- Webtrends social measurements, 156
- Weinreich, Harald, 2
- Wetpaint tool, 136
- White, E. B., 141
- Wikipedia
 - article traffic statistics tool, 54, 59–60
 - free and open content on, 116
 - integrity, 136
 - linking to, 116–117, 136
 - links from, 10
- Wikis, 116, 135–136
- William of Ockham, 19
- Wilson, Deidre, 36–39
- Wireframes in task analysis, 97, 99
- Word Frequency Calculator, 62

- WordPress
 - for blogs, 138
 - LSI Insights plug-in, 158
- Words, spaces between, 17–22
- Wordtracker tool, 54, 58–59
 - credibility through PageRank, 9–11
 - evolution from print, 1–2
 - keyword discovery, 4–5
 - optimized site architecture, 6–9
 - relevant content, 3–4
 - for skeptical Web readers, 26–27
 - social media and rich media, 11–12, 127–133
 - targeted search referrals, 5–6
 - for Twitter, 141
 - web effectiveness, 12–13
 - for wikis, 136
 - write-for-search-engines-first approach, 2
- Writing overview, 1
- Writing Quality survey module, 80

- XML-based sitemaps, 105

- Yahoo, link emphasis by, 110–111
- Yahoo Alerts, 5
- Yahoo Pipes, 5
- YouTube for linking relationships, 123