



**6TH**  
EDITION



# FIND IT FAST

Extracting Expert Information  
From Social Networks, Big Data,  
Tweets, and More

**ROBERT I. BERKMAN**



**6TH**  
EDITION

# FIND IT FAST

*A wonderful book . . . highly recommended.*

—Library Journal

*A valuable reference work for anyone digging for information.*

—Journalism Quarterly

*For serious researchers, this book is a gold mine.*

—Choice

**6TH**  
EDITION

# FIND IT FAST

Extracting Expert Information From  
Social Networks, Big Data, Tweets, and More

**ROBERT I. BERKMAN**



**CyberAge Books**

Medford, New Jersey



**Information Today, Inc.**

First printing

***Find It Fast, Sixth Edition: Extracting Expert Information From Social Networks, Big Data, Tweets, and More***

Copyright © 2015 by Robert I. Berkman

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means including information storage and retrieval systems without permission in writing from the publisher, except by a reviewer, who may quote brief passages in a review. Published by Information Today, Inc., 143 Old Marlton Pike, Medford, New Jersey 08055.

Publisher's Note: The author 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 contained herein.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book and Information Today, Inc. was aware of a trademark claim, the designations have been printed with initial capital letters.

### **Library of Congress Cataloging-in-Publication Data**

Berkman, Robert I.

Find it fast : extracting expert information from social networks, big data, tweets, and more / Robert I. Berkman. — Sixth edition.

pages cm

ISBN 978-1-937290-04-7 (paperback)

1. Information retrieval. 2. Research. 3. Internet searching. 4. Search engines. 5. Information resources—Evaluation. I. Title.

ZA3075.B47 2015

025.0425—dc23

2015019680

Printed and bound in the United States of America

President and CEO: Thomas H. Hogan, Sr.  
Editor-in-Chief and Publisher: John B. Bryans  
Associate Editor: Beverly M. Michaels  
Production Manager: Tiffany Chamenko  
Indexer: Nan Badgett

Interior Design by Amnet Systems  
Cover Design by Alison Gallagher-Rehn

**infotoday.com**

To Sol, Pat, Budd, and Don



# Contents

List of Figures.....	viii
Acknowledgments .....	xi
Preface to the Sixth Edition .....	xiii
Introduction.....	xix
Getting Started.....	xxiii

## Part I: Sources

1. Super Sources: The Cream of the Crop ..... 3
2. Libraries—Still Valuable in the Digital Age ..... 51

## Part II: Searching

3. Search Engines, Precision Search Strategies, and Taming Information Overload ..... 89
4. The Social Search: Tapping into Your Networks..... 131
5. Truth, Lies, and Influence: Determining Credibility in a Trending Social Media World ..... 157

## Part III: Experts Are Everywhere

6. Identifying Experts: Who They Are, Where to Find Them ..... 225
  7. Making the Connection: Getting Access to an Expert ..... 235
  8. Talking with Experts: Strategies for Getting Inside Information ..... 253
  9. Wrapping It Up: Organizing and Writing up Your Results and the Expert Review..... 271
  10. Troubleshooting: Typical Questions Information Seekers Ask ..... 281
- 
- Appendix A: Digital Forensic Tools..... 291
  - Appendix B: Sources of Further Information ..... 293
  - About the Author ..... 303
  - Index ..... 305



# List of Figures

- 1.1 Results of a search on “hydraulic fracturing” and Oklahoma and Earthquakes on Google Scholar retrieves results from academic journals, documents, and proceedings .....26
- 1.2 SSRN provides free searching of thousands of published and unpublished academic papers, conference presentations, articles, and other scholarly materials.....28
- 1.3 This page retrieved via the Wayback Machine shows the front page of the September 15, 2008 *New York Times* at the height of the financial crisis when Lehman Brothers announced its Chapter 11 bankruptcy.....40
- 1.4 Vanderbilt University offers free archives of television broadcast shows back to 1968 .....42
- 2.1 A photo of Leonard Bernstein with Stephen Sondheim at the piano, rehearsing music from the Broadway production of *West Side Story*, from the Library of Congress’s Music Division.....76
- 2.2 Albert Einstein’s application to become an American Citizen is one of many thousands of historical images available from the Library of Congress’s Digital Collections.....77
- 2.3 The James B. Hunt Jr. Library at North Carolina State University is state-of-the-art, and includes a wall sized curved video display.....81
- 3.1 When Google first launched in 1998, it offered an appealing search page that was much more uncluttered than those of other search engines at the time .....98
- 3.2 Boolean AND, OR, and NOT search operations .....102
- 3.3 One of the most useful but lesser known Google features is the ability to limit results to a certain date. This search retrieved pages about Lehman Brothers just for the month of September 2008, when the company filed for bankruptcy .....105
- 3.4 The DuckDuckGo search engine does not track users or retain their search history.....109
- 3.5 The Internet Public Library (IPL2) is a high-quality directory of pages on the web, grouped by subject .....113

3.6	Newsle alerts you to news, blogs, and web pages on people you want to keep up with and track.....	120
3.7	An alternative method for displaying results is to cluster similar results together, as provided by the experimental search engine SearchPoint.....	121
3.8	Subscribe to an RSS feed when you want news and information on specific high-interest topics delivered to you automatically. Environmental issues are the topic of the pictured feed.....	122
3.9	An example of a Scoop.it curator’s aggregation of recent online news items on the topic of solar energy.....	124
3.10	This Google search turns up “Scoop.it” curations on the topic of farmers markets .....	125
4.1	The profile of the author on the scholarly social network, ResearchGate .....	135
4.2	Academia.edu is another leading scholarly social network where users can share their work and connect with others in their field .....	136
4.3	A search on “Big Data” on Academia.edu identifies scholars who have knowledge, credentials or shared work for that topic .....	136
4.4	The author created three focused lists for following people on Twitter with a certain expertise .....	139
4.5	Using the free TweetDeck dashboard helps organize different tweets from different people or themes I follow on Twitter .....	142
4.6	TweetDeck provides an “at-a-glance” display of incoming messages, new tweets on a designated keyword, and more .....	142
4.7	TweetDeck allows users to not only search for tweets on a certain word or phrase, but to limit results just to those tweets that initiated a high level of engagement .....	143
4.8	Twitter Advanced Search Operators. Twitter provides a wide range of advanced search operators to make a search more precise .....	144
4.9	If you search for a word linked to a hashtag, you are more likely to retrieve a tweet that is “about” that particular topic .....	146

**x Find It Fast**

4.10 The result of a search for tweets that include both the words “earthquake” and “hydrofracking” .....146

4.11 Topsy allows for searching on tweets and assigns a “sentiment” score that attempts to show whether the overall sentiment of a tweet is more positive or negative ..148

4.12 Searching Topsy for “hydrofracking” references occurring between June 15 and July 15, 2014 shows the highest number of tweets on June 30th ..... 149

4.13 A keyword search on Facebook retrieves matches for a person’s name, employer, group names, company pages, cities, and other elements in Facebook .....151

4.14 Facebook has retrieved a listing and description of groups concerned with the topic of hydrofracking .....151

4.15 It is possible to do a search of all posts and comments on a Facebook group by clicking on the magnifying glass in the upper right hand corner of the group’s page .....152

4.16 It is possible to search conversations by keyword on LinkedIn .....155

5.1 Facebook identifies news items being shared most frequently and describes these items as “Trending” .....158

5.2 NPR Journalist Andy Carvin has relied on his followers on Twitter to verify uncertain information .....204

5.3 You can search past tweets on Google by limiting your search results just to comments posted directly on Twitter .....217

5.4 TweetCred uses a variety of signals to assign a “credibility score” to a particular tweet.....219

7.1 Rapportive is a free add-on extension to Gmail that automatically displays a snapshot and data about the person you are emailing based on their social network profiles .....247

7.2 Once someone follows you on Twitter, you can send that person a private message .....249

**Tables**

U.S. Government Sources by Topic .....7–10

Wilson Indexes by Interest Area .....63

Key Directories and Databases That Include Them.....64

Six Steps for Project Planning .....285

# Acknowledgments

I'd like to express my gratitude to the many people who in one way or another played a big part in making this book come to pass, from the first edition published in 1987 through this one, the sixth.

My wife, Mary Walsh, gave me encouragement at the most critical moments—without it the book would not have been written. George Finnegan has been a true mentor. I'll always be thankful for the chance to have worked for him which was a once-in-a-lifetime learning opportunity. Janet Goldstein, my first editor at HarperCollins, believed enough in the book to give it a chance and always provided support and great ideas.

Nancy Brandwein, Ken Coughlin, and Lilli Warren all offered critical insights, suggestions, and critiques of the original manuscript. Those who assisted with past editions include Sandy Gollop, Debbie Cohen, Ginny Fisher, Leigh Woods, Hella Rader, Cindy Pereiri, Jeff Goodman, Pam Goodman, Marie Kotas, and Trisha Karsay. Thanks are also due to Sasha Mishkin for her valuable help with the new edition.

I'm very appreciative of my colleagues at Information Today, Inc., particularly my editor John B. Bryans for helping get this 6th edition published after a 15-year hiatus from the 5th. I've been fortunate to have a helpful and encouraging team to work with for many years, publishing my books and the monthly research journal I co-edit, and supporting my other professional endeavors. I also want to note my appreciation for the excellent work of Beverly Michaels for her careful reading of the latest manuscript and her vital observations and questions that helped shape it into a better book.

A special thank you goes to everyone in the Norris family for providing me with such a beautiful and relaxing spot in which to write the first edition, in North Truro on Cape Cod.

Finally, thanks to my parents and brothers for their love, support, and invaluable ideas.



# Preface to the Sixth Edition

It's simple to go on the web and find information. So why would you need a book about doing research?

Perhaps if you Google that question you'll find the answer.

Actually . . . no. Googling that question won't get you a good answer. There are, in fact, still countless queries and information problems for which the answer won't be found via Google. Or Wikipedia. Or Twitter. And that's one of the points of this newly revised book on how to find credible information and perform good research.

*Find It Fast*, first published in 1987, was last revised for its fifth edition in the year 2000. That's eons ago in the information world, and an awful lot has happened in that world since then. For example Web 2.0, Wikipedia, Twitter, citizen reporting, sentiment detection, real-time search, social media, Google+, Pinterest, Instagram, Tumblr, YouTube, Vine, tagging, cloud collaboration, the wisdom of crowds, collaborative filtering, LinkedIn, Klout, Siri, Snapchat, social media monitoring, Facebook, Foursquare, and Flickr—and more!

Back in the year 2000, all those concepts and sites were either not around or barely emerging. That's an enormous amount of change in the information world.

But a lot has *not* changed—specifically, the elements of doing good research. For example, understanding how sources differ, knowing which resources to use, finding and evaluating experts, verifying information, understanding surveys, controlling one's biases, asking the right questions, assessing source credibility, disregarding the noise, understanding what's truly significant, and knowing when to conclude a project.

## The More Things Change . . .?

Can we say, then, that doing good research has not fundamentally changed, but that the sources, tools, and strategies have changed?

Well, not quite. Not only have the sources, tools, and strategies clearly changed, they have, in fact, changed *so* dramatically that they have produced a revolutionary rethinking in many aspects of what constitutes good research.

Why are these changes truly *revolutionary*? We have to start by looking back to the mid-1990s, with the arrival of the internet, of course, which demolished the high barriers to entry for becoming a publisher. The last time we had a change *that* revolutionary for creating and distributing information was over 600 years ago when a German by the name of Johannes Gutenberg came up with the concept of movable type.

By the early 2000s we saw the emergence of the social web—or Web 2.0 as it was then known. Initially, interaction on the web was primarily in the form of blogs, but this innovation was quickly followed by an explosion of other two-way information-sharing platforms: Flickr, YouTube, Tumblr, MySpace, Facebook, and many others. The social web permitted anyone to share his or her opinions, knowledge, and content, and to become part of a larger global *conversation*. As a result, we have all had to rethink some fundamental assumptions about what constitutes a credible source and, therefore, what it means to do “good” research.

## **Who’s an Expert?**

Who counts as a trusted source today? Before the social web, an individual established his or her authority on a subject through credentials (e.g., a PhD in history, an MA in design, etc.) or with their published works, professional presentations, or other means that demonstrated recognition by institutional gatekeepers such as universities, publishers, and editors.

If a person could manage to clear those hurdles—write the dissertation, get a presentation approved, have an article published, be quoted in a news story—then he or she was given the mantle of legitimate authority. The rest of us could then feel certain that he or she was in fact a credible source of expertise.

The internet and, more recently, social media have forced us to reconsider, or at least expand, our criteria for who should be considered an authority and whose views and knowledge are worth taking seriously.

Today we live in an information-rich environment, where the views of countless people without any formal—or even informal—credentials, who have never been peer reviewed or quoted in a

newspaper, never presented at a conference or even created a piece of work, are given a great deal of credence. Consider, for example, the popularity and sometimes even prominence of certain bloggers, “citizen-journalists,” YouTube producers, prolific Wikipedia entry writers, or just someone with a high “Klout” score. These people are sometimes trusted as credible sources. Should they be given that recognition just because they are popular? Should popularity really confer authority and expertise?

These critical questions will be discussed later in this book.

And what about the whole “wisdom of the crowd” meme: that, in certain circumstances, a larger group of ordinary people is smarter and more likely to come to a more accurate decision than any one individual—even if that individual is an expert?

These are the kinds of questions that are creating major—in fact revolutionary—changes in how we all create, find, and assess information sources, and they will be treated in detail.

## **Big Data and Big Challenges**

The question of expertise—who currently deserves to be considered a legitimate authority and credible source—is not the only dramatic change faced by information searchers today. Of course, there is little difficulty these days in finding a simple fact or answer, or even turning up a good description of a place, person, or thing. Link to a search engine, key (or speak) a few words into Google, and get matching results instantly—that’s as easy as it gets.

But significant challenges are still faced by those with more involved research tasks. One is simply how to sort through and make sense of the countless streams of incoming real-time data continually washing over us. The fact that we are *all* becoming publishers and broadcasters means that more and more data is generated—from tweets to video clips—making it that much harder to discern what truly deserves our precious and limited attention.

One increasingly popular approach for making sense of the growing ocean of information is to rely on “big data,” perhaps the hottest buzzword in the business and technology world. The term describes the use of super-powerful computers to collect, crunch, and analyze the enormous and ever-increasing quantities of



information being produced these days. There are countless sources of data, ranging from people's social media profiles to sensors on household objects, supermarket scanners, weather station reports, and smartphone geolocation data.

The promise of big data lies in its ability to deeply analyze so much disparate data, find correlations, and surface previously hidden patterns to provide new insights and drive more informed decisions. The big data phenomenon also presents challenges for the researcher, described later in this book.

## **The Value of *Find It Fast* in the Digital Age**

Finally, you may be wondering about the value of some of the items included in this book or even the concept of the book itself.

Why, for instance, have I bothered to include certain seemingly old-fashioned print directories and indexes in the library chapter? While there are a few reasons print is still valuable, as discussed in Chapter 2, one subtle but quite important reason is that using these types of sources serves as an important check on your internet searching, which generally gives the highest rankings to the most popular sites. This is fine as far as it goes (much more on that process in Chapters 3 and 5), but what it means is that while you are likely to get a lot of results published by the best known and most used news and information sources, such as *The Huffington Post*, *The New York Times*, *CNN*, the *Guardian*, *TMZ*, etc., you are much less likely to turn up articles published in very small, niche, and lesser-known journals and news sites. And sometimes these turn out to be your best sources.

By adding an old-fashioned print index and directory to your set of research tools, you can at least complement this search-engine ranking method by using a source with a different mechanism for discovery, and thereby override the built-in biases of the search engines. You then have a more rounded set of results than you would receive from a Google search alone—and that's a good thing.

Given the speed of change in this arena, you may also wonder how long a book like this can remain relevant and useful. That's a good question.

It is certainly true that the websites and search tools referenced in this book are bound to change, so I've addressed this issue in two ways: first, by including only sources that have been around—or are most likely to be around—for a long time. But even more importantly, the heart of this book is not a list of sites or the technologies of the moment. It is all about cultivating a deeper understanding of the critical principles of research and what it means to be a good researcher in the digital age.

Finally, a few more words on the journey this book has taken since its first edition in 1987.

The introduction to that initial edition began with the remark that “it may come as a surprise to you, but for virtually any subject, facts and information are out there—by the truckload.” The book promised to provide the necessary know-how so the reader would “be able to get the kind of information that’s normally available only to a select few.”

But by fifth edition, published in 2000, the power of the internet—as a source of an unfathomable amount of information to anyone, anywhere, anytime—was becoming more fully realized. And that edition of the book concluded with some thoughts on the potential of this amazing new technology:

The internet provides explanations for problems or concerns that may have stumped us for years . . . The sum total of all of our knowledge and thoughts, in fact, is increasingly being made public, linked and accessible. All sorts of artifacts of knowledge are being digitized, made searchable and available . . .

Of course, the internet will never answer the deepest human questions . . . or will it? And finding answers to one set of questions allows for, and even compels, the emergence of a new set . . . and so our questioning will surely continue.

And so it continues today.

Here's to getting your questions answered—and then generating new and even better ones!



# Introduction

## Who's in Charge of Your Research, Anyway?

If someone were to ask you who's in charge when you do some casual searching on the web, your response would likely be something like, "Well, I decide what I'm looking for online. Most of the time I go to Google, enter the words for what I want to find, and I get my results. So I'd say that *I'm* in control."

Well, not to get too Matrix-like on you, but there are a whole host of hidden entities that are jockeying for your attention and working to control what you are able to do, how you are able to do it, and what results you will view. Among the most powerful behind-the-scenes forces are the following.

### Google Itself

As will be examined in detail in Chapters 3 and 5, Google prioritizes the pages that are most popular—these are the ones with lots of incoming links. It also prioritizes web pages that, according to its calculations, are supposed to be of highest interest to you personally, based on your past search history, geographic location, and other individual factors.

### Search Engine Optimizers (SEOs)

SEOs use a variety of techniques—some approved by Google, others more iffy, to try to drive their clients' pages to the top of the returned list. This means that what you see from your research not only reflects your search statement and Google's own ranking algorithm, but also the skill and intent of other entities trying to get their pages ranked highest.

### Clickbait Creators

Clickbait is the term used to describe those ridiculous yet oddly alluring headlines crafted just for the purpose of getting you to click and read the story. Unfortunately, research shows that these absurd techniques often do work, due to creating what's called a "curiosity gap" in us, causing our minds to seek the missing information.<sup>1</sup>

## **xx Find It Fast**

So what's a good researcher to do?

First we might ask, Why is any of this a big deal? Haven't there always been external forces directing how we do research?

No, not really; at least not in the same way.

In "the old days"—let's put that before the mid-1990s or so—when a researcher went to consult his or her sources, whether it was an encyclopedia, reference book, journal, magazine, primary document such as a letter or photo, or a traditional searchable online database, the primary forces that determined what the researcher would find included only the following:

- The scope, content, and timeliness of that resource (archive, timeliness, quality controls, etc.)
- The methods and limitations of the tool employed for searching a set of data or sources (ability to find a citation vs. abstracts vs. full text; ability to do advanced keyword searches, etc.)
- The researcher's own skills and abilities to construct an effective research strategy and search statement
- The method selected by the publisher/database to define what constitutes a relevant result (keyword match, recency of item, etc.)

But now we all must contend with external forces whose larger goals don't necessarily coincide with our own as researchers.

The fact is, unlike a library database or other traditional information provider, these entities do not have the same overriding goals as we do. They include a larger mission that typically means attracting more users, luring in paid subscribers, selling advertising, collecting data, and ultimately turning a profit. And so what we ultimately end up seeing as searchers in some manner serves those overriding goals.

A primary purpose of this book is to help put you back in control of your own research. This means having a level of digital information literacy so you are aware of the various forces that impact what you find and why, and enough knowledge and understanding of today's information environment so you are in the driver's seat—finding what you need and seeing and working with sources that are going to best fulfill your research needs.

What can you do to ensure that you are in charge, that you're steering the ship and finding the best—most relevant, insightful, timely, and credible—information?

That's where this book comes in.

## **What's Inside This Book**

A preliminary section, “Getting Started,” will help you define what information you're really after and organize your plan of attack. Part I consists of two chapters that identify specific information sources and provides tips on how best to find and use each of them. Chapter 1 is a compilation of my favorite all-purpose and powerful “super sources,” ranging from associations to museums to government data, which are among the best resources for virtually any research endeavor. This chapter also identifies some of my favorite open access sources. These are scholarly books, journals, and other resources that were created to be free and available to anyone, reflecting the goals of the open access community to make information more freely available and outside of expensive and restrictive paywalls. Chapter 2 explores how to find the right library, why libraries are still valuable in the age of the internet, some of the best ways to use your library, and even when it is most effective to use print sources.

Part II looks at what it really means to be an effective information searcher on the internet. Chapter 3 examines the birth and development of search engines: how they work, getting the best results out of your search, alternatives to Google, and a discussion of lesser-known internet discovery tools that are not even search engines at all. Chapter 4 explores the recent rise of “social searching”—using your social network as a source of information, when it makes sense to do so, and practical search strategies you can employ for the most popular social networks including Facebook, Twitter, and LinkedIn. Finally, Chapter 5 takes on the particularly tough and perplexing issue of source credibility, examining both traditional scholarly and journalistic methods for assessing a source, and also the latest thinking and methods for verifying citizen journalism and social media posts. This chapter also analyzes the notion of the wisdom-of-crowds, explaining how that principle works, under what circumstances it can be relied on, and when it does not apply.

The last major segment of the book, Part III, is about finding and interviewing experts. In Chapters 6 through 8 you'll discover how to locate the best type of experts for your research, and the pros and cons of each type. You'll also learn how to make contact with the experts and how to get them to open up and share their knowledge.

Finally, the last chapters offer advice on how to tell when it's time to wind up your project and some tips on writing up your results. Chapter 10 includes a "Researchers Road Map," a guide to choosing and understanding the varying types of information sources.

At the end of this book are two appendices. One lists useful research and verification tools, and in the other I share my favorite books, journals, and other resources on doing research and understanding how to better leverage knowledge in the digital age.

## **Note**

1. For a detailed analysis on how the site Upworthy refines this process of creating the most attention-grabbing headlines possible, see this piece in *New York* magazine: [nymag.com/daily/intelligencer/2014/03/upworthy-team-explains-its-success.html](http://nymag.com/daily/intelligencer/2014/03/upworthy-team-explains-its-success.html).

# Getting Started

Before plunging into your project, take a few minutes to consider your endeavor. Try to determine for yourself exactly what kind of information you need and why. What are your reasons for wanting to find this information? What are your overall goals? Try to be as specific as possible, even though it may be difficult to do so at the earliest stage. The more you can narrow your scope and break your task into subprojects, the easier your search will be and the more likely it is that your project will be a success.

## **TIP: Deciding What to Research**

How do you decide what to research when the research topic is up to you? The best guide that I've come across is called "Identifying Your Interests and Establishing a Research Plan" (<http://tinyurl.com/p8w8sjp>) written by Professor Shannon Mattern at the School of Media Studies at the New School for Public Engagement. While geared for media studies students, Mattern's inspiration and advice is invaluable for any researcher.

For example, say your goal is to find information on how newspapers are rethinking their business models in the digital age. With a little reflection, you can break that topic into its major components: What were the traditional business models before the internet? What elements provided revenue, and how much? Which of those elements are no longer viable? What new methods are being tried out to generate revenue? Who are the leaders in creating a workable business model—what did they do, and how did they do it? Are there differences in other countries or lessons that can be learned—what are they? What do readers say they want and like in a newspaper today? Will they pay for those features, and if so, under what circumstances? And so on.



Now you have some specific and concrete subtopics to zero in on. If during your research you discover that your subject is too broad for you to adequately handle it within the strictures of your plan, you can decide whether to choose one or more of your subtopics instead. And if you are not familiar enough with your subject at the outset to identify the subdivisions, you'll find that you will discover them once you begin your research.

The first step in the information-gathering process is to find the best “nonhuman” published sources—whether in print, in databases, or on the web. Although some of your best results will eventually come from talking to experts, you don't want to begin your project by contacting them. It's much better to first read and learn about your subject, and then speak with the experts. This way, when you do eventually talk to the authorities in the field, you'll be knowledgeable enough to ask the right questions and get the most out of your conversation.

Before you actually start your research, you should also come up with a method for capturing the information you'll be gleaned from your published and expert sources. Your approach to note taking and organization is important because it will affect the course of your entire project. If you need advice on this, see Chapter 9. If, however, you feel confident enough to jump right in to the information search, continue on to the first chapter.

Part I

# Sources



# Super Sources: The Cream of the Crop

The resources described in this chapter are the cream of the information-source crop. They range from museums to the federal government, from the state business filing registry offices to other storehouses of information, but they all have a few things in common. Each contains information on an enormous scope of subjects. Each can easily be tapped for answers and advice. And each provides answers either free of charge or nearly free.

I've organized these sources into the following broad categories:

- The Best of the Library Sources
- The Best of the U.S. Government Sources
- Business Super Sources
- Statistical Sources
- Scholarly Databases, Theses, and Journals
- Open Access and Public Data Sets
- Other Super Sources

## The Best of the Library Sources

In Chapter 2, I examine the continuing value of libraries in the digital age and how to make the best use of libraries when doing research. Here I'll simply list a handful of the very best individual libraries and library-oriented resources that you can turn to in order to find information on virtually any topic you are researching. This section is broken up into two subcategories: "The Best of the Best" and "The Rest of the Best."

## 4 Find It Fast

### The Best of the Best

**Source:** *Library of Congress*

The U.S. Library of Congress in Washington, D.C., is the largest library in the world. Its collection includes millions of volumes and pamphlets, technical reports, maps, manuscripts, photographs, negatives, prints, and slides. The library is also known for its collection of rare books and foreign publications.

Using the Library of Congress's vast resources can be tricky—not only because there is so much information available but also because the library's policy discourages extensive reference usage when the same materials are available on a more local level. However, it does assist users in researching topics unique to the library (such as copyright, legislative research, and international law).

There are also certain services and sections of the Library of Congress that are set up specifically to help the public find and use its resources:

- **Ask a Librarian.** Click on a specific subject area and see a page with the range of the Library of Congress's resources, and have the opportunity to pose your research question to a librarian. You should receive a response within five cases. In some cases, you can even have a live chat.
- **Virtual Reference Shelf.** A list of sites and resources recommended by the Library of Congress on topics ranging from architecture to statistics.
- **Searchable databases.** While many of these powerful databases are available only for in-library use at the Library of Congress itself, others are free and available remotely.

There's a lot more—spend time browsing and searching the site and you'll see for yourself!

**Source:** *WorldCat (OCLC)*

Do you have need to find the library closest to you that has a copy of a particular book (or music CD or DVD)? Just enter the name of the

item and your zip code into WorldCat and you'll get an immediate listing of which libraries nearest you have that item on their shelves. (You can do a lot more on WorldCat too, including finding journal articles and downloading ebooks, but its fundamental purpose is to seamlessly search about 10,000 of the world's library catalogs to help people find the books or other content they seek.)

**TIP:**

Remember that WorldCat only surfaces titles and other bibliographic information about a book or content, so you'll need to track down the original item yourself. However, Hathi Trust, a consortium of major leading research libraries around the world, is another excellent resource that not only allows searching across millions of public domain and private books, journals and other resources, but also permits a certain amount of free fulltext searching as well.

Note that when you find a book on Google Books, the dropdown menu under "Get this book in print" includes a "Find in a library" link to the WorldCat entry for that title.

**Source:** *The New York Public Library*

The New York Public Library (NYPL) is a tremendous source of all kinds of information. The library's mid-Manhattan branch is especially rich in its holdings and regularly answers inquiries from around the country via its ASK NYPL reference service, which accepts phone calls at 917-ASK-NYPL from 9:00 AM to 6:00 PM, Monday through Saturday. Email requests are also accepted. Its collections cover the fields of art, business, education, history, literature and language, and science. In addition, the library contains an extensive image collection.

Specialized research collections of the NYPL include the George Arents Collection on Tobacco; Berg Collection of English and American Literature; Dorot Jewish Division; Jean Blackwell Hutson Research and Reference Division, which focuses on peoples of

## 6 Find It Fast

African descent throughout the world; Jerome Robbins Archive of the Recorded Moving Image; Lionel Pincus and Princess Firyal Map Division; Rodgers and Hammerstein Archives of Recorded Sound; Schomburg Center for Research in Black Culture; Science/Industry and Business Library; Spencer Collection of illustrated word and book bindings of all periods and all countries and cultures; and the Theatre on Film and Tape Archive. There is also the Performing Arts Research Center, which answers inquiries regarding music, dance, and theater at no charge.

Here are just a few of the ways you can start learning about the resources of the NYPL:

**Research Guides.** The NYPL has created detailed step-by-step instructions on how to begin research in the library on many different subject areas, ranging from historical photographs to maps to patents.

**Digital Collections:** The NYPL has digitized over 800,000 items from its holdings and made them available online. There are digitized photographs, audio, and other formats, on topics ranging from the arts to maps, birds, immigration, and social conditions.

**Articles and Databases:** All of these are free to use when visiting the NYPL in person; online, there are a mixture of free and fee-based databases.

## The Rest of the Best

**Source:** *The Center for Research Libraries (CRL)*

The Center for Research Libraries (CRL) is an international consortium of university, college, and independent research libraries headquartered in Chicago and it makes its catalog of holdings available for free searching on the web. Documents themselves are available on loan free to member libraries or to nonmember libraries for a fee. CRL's holdings include international newspapers, doctoral dissertations, and much more.

**Source:** *CiteSeer*

CiteSeer is a digital scientific literature library that focuses primarily on computer and information science.

## The Best of the U.S. Government Sources

The U.S. government is a gold mine of information. Although some government resources and information services have been eliminated over the years, an awesome amount of advice, data, and information is still available—and since the mid-2000s or so the focus of the government has been to make much of that information freely and easily available over the internet. Information is available from the government on a vast number of topics. The following table of departments and agencies, along with selected topics each covers, should give you a good feel for what’s available.

Department/Agency	Topics
Department of Agriculture	animal and plant health; consumer affairs; family nutrition; food safety and inspection; human nutrition; veterinary medicine
Department of Commerce	business outlooks; economic and demographic statistics; engineering standards; imports and exports; minority-owned businesses; patents and trademarks; technology; travel; weather and atmosphere
Department of Defense	atomic energy; foreign country security; mapping; military history; nuclear operations and technology; tactical warfare
Department of Education	adult education; bilingual education; civil rights; educational statistics; elementary and secondary education; handicapped services; higher education; libraries; special education
Department of Energy	coal, gas, shale, and oil; conservation; energy emergencies; fusion energy; inventions; nuclear energy; nuclear physics; radioactive waste; renewable energy
Department of Health and Human Services	AIDS; alcohol abuse; disease; drug abuse; drug research; family planning; food safety; minority health; occupational safety; smoking; statistical research data; toxic substances; veterinary medicine; NIH institutes for cancer; eyes; heart, lung, and blood; arthritis and musculoskeletal and skin diseases; dental and craniofacial research;



## 8 Find It Fast

Department/Agency	Topics
Department of Health and Human Services (cont.)	diabetes and kidney diseases; environmental health; neurological and communicative disorders and stroke; human development services; aging; children, youth, and families; developmental disabilities; Native Americans
Department of Housing and Urban Development	block grants; elderly housing; energy conservation; fair housing; urban studies
Department of Homeland Security	border security; citizenship and immigration services; civil rights and civil liberties; cybersecurity; disasters; economic security; human trafficking; immigration enforcement; terrorism prevention; transportation security
Department of the Interior	archaeology; fish and wildlife; geology; mapping; minerals; Native Americans; natural resources; water
Department of Justice	antitrust; civil rights; drug enforcement; immigration; justice statistics; juvenile justice; prisons
Department of Labor	employment training; labor-management relations; labor statistics; occupational safety and health; pension and welfare benefits; productivity and technology; veterans' employment; women
Department of State	African affairs; arms control; Canadian and European affairs; East Asian and Pacific affairs; human rights; inter-American affairs; international environmental affairs; international narcotics; Near Eastern and South Asian affairs; nuclear and space arms negotiations; passport inquiries; prisoners of war/missing in action; refugees; visa inquiries
Department of Transportation	automobile safety; aviation safety; aviation standards; boating; hazardous materials transportation; highway safety; mass transit; National Highway Traffic Safety Administration; railroad safety; shipbuilding; vehicle accident statistics; vehicle crashworthiness
Department of the Treasury	coin and metal production; currency production; currency research and development; customs; saving bonds; Secret Service protection; taxpayer assistance; tax return investigation

<b>Department/Agency</b>	<b>Topics</b>
Consumer Product Safety Commission	burn hazards; product safety assessment; mechanical hazards; injury information; electrical shock hazards; safety packaging; chemical hazards
Environmental Protection Agency	air and radiation; pesticides and toxic substances; acid deposition; climate change; environmental monitoring and quality assurance; solid waste and emergency response; water; noise control
Federal Communications Commission (FCC)	cable television; broadcast stations; radio regulation
Federal Trade Commission (FTC)	advertising practices; competition and antitrust matters; consumer protection; financial statistics
General Services Administration (GSA)	consumer information; government audits and investigations; fraud hotlines; federal property; purchasing of equipment and supplies; information management; public buildings management
National Aeronautics and Space Administration (NASA)	aeronautics and space technology; life sciences; astrophysics; earth sciences; solar system exploration; space shuttle payload; Mars observer program; microgravity science; upper atmosphere research; solar flares
National Archives and Records Administration	naturalization records; census data; military records; land records; passenger lists; passport applications; selected vital statistics; presidential documents; audio and other media
National Endowment for the Arts and Humanities	literature; museums; folk arts; visual arts; music arts; theater arts and musical theater; opera; media arts; history; language
National Science Foundation	atmospheric/astronomical and earth–ocean sciences; mathematical and physical sciences; Arctic and Antarctic research; anthropology; engineering; biology; genetic biology; chemistry; computer science; earthquakes; economics; ethics in science; meteorology; galactic and extragalactic astronomy; geography; geology; history and philosophy of science; nutrition; linguistics; marine chemistry; metallurgy;