

Ready Reference on the Web: Resources for Patrons

Quick! What is the best Web resource for finding a good vacuum cleaner? Getting information for a grade-school “state” report? Finding a life-saving clinical trial?

People just walk up to our desks and ask us these questions every day. And we are expected to answer them in seconds, not minutes or hours.

For heaven’s sake! How are *we* supposed to know?

Well, as Robert W. Winter, Arthur G. Coons Professor of the History of Ideas, Emeritus at Occidental College in Los Angeles once told our class, “A liberal arts education will teach you what you don’t know.” In other words, as educated people, as *librarians*, we know that we don’t know everything, yet we are armed with tools to find out about anything.

My hope is that the first section of this book can serve as one of those tools. I have tried to cover some of the most asked subject areas that I encounter every day on the job. I also point to directories of high-quality sites that can quickly connect librarians and their patrons to the answers that they need.

These listings may not answer every question. But I guarantee you, if you can learn to turn to them first, your patrons will hail your alacrity and brilliance.

And isn’t this glory what we librarians *live* for?

Searching and Meta-Searching the Internet

The topic of searching the Web puts me in an elegiac mood. I think back to 1994 when I was first playing with the Web during my library school internship at the Getty Research Institute. For 15 hours every week, I scrolled and clicked, calling in new home pages from around the world. My desk faced a window seven stories above the Santa Monica Bay, with its whitecaps and sailboats and Catalina Island in the distance. Yet, I found the baby Web so arresting that I rarely looked up from my task to admire the view.

In that year, Yahoo! (www.yahoo.com/) was just a little hyper-linked directory maintained by a couple of guys from Stanford University. The search engine of choice was Lycos (www.lycos.com/), which seemed at the time to be a marvel of software engineering, able to find anything that existed on that sparsely populated World Wide Web. Later, Lycos foundered, started selling its listings, and lost its edge to such newer engines as Infoseek and AltaVista (www.altavista.com/). Now, Lycos is back, InfoSeek is gone, and AltaVista is a fading presence.

Yes, even as the seasons change, so do the search engines. Disney bought Infoseek, renamed it GO.com (www.go.com/), and turned it into a portal for Disney press releases. Another search engine, Norway's AllTheWeb.com (www.alltheweb.com/), appeared and purports to index more Web pages than anyone else.

Amid all this chaos of bursting and fading search tools, two Stanford Ph.D. candidates, Larry Page and Sergey Brin, came up with a technologically advanced method to analyze links on Web

pages. In 1998, they applied their paradigm through a new search engine called Google (www.google.com/). It became an instant success, famed for both the simplicity of its interface and the almost spooky relevance of its results. Google bases its relevancy rankings on the number of links pointing to a site. It also considers the importance, based on linkage, of the pages that point to the site. The result, especially for common word searches, is spectacular and highly resistant to “spamdexing.” (Spamdexing, that is, the application of false or misleading meta-tags to Web content, corresponding to sending unwanted messages via e-mail, aka spam, continues as a serious problem for search engines, so much so that the struggle between spamdexers and search engines is referred to as an “arms race.”)

Of course, Google's relevance comes at a price. Its default Boolean “and” cuts out most results. Admittedly, many of those results would have been irrelevant. Still, Google might miss obscure sites that contain a rare treasure.

Because of its honing features, Google is a great library assistant if you need precise results from the Internet. The other day, a patron came up to me and asked if we had that book by, “... oh, you know, that lawyer in the O.J. trial, the one that always wears buckskin.”

I knew exactly who he was talking about. I could see the man's face doing lawyerly analysis on TV. But, danged if I could remember the guy's name. So, I pulled up Google and typed “lawyer O.J. buckskin” into the search box. “Gerry Spence” came up in the first hit. Holy cow!

Google is also great for finding poems based on half-remembered lines quoted by patrons. “I want that poem that starts with ‘I have slipped the surly bonds of earth,’” they will ask. Imagine using a book to find the first line of a poem for which you have no frame of reference. But type “I have slipped the surly bonds of earth” into the Google search box, and up pops the text of the

poem “High Flight” by John Gillespie Magee, Jr., who was killed in World War II during a dogfight at age 19. Just what our patrons asked for!

Google will even check your spelling. For instance, if you search for a file-sharing service by typing in “morpheus,” Google will try, but will also helpfully suggest, “Did you mean: *morpheus*?” Yes, thank you, Google.



Figure 1.1 Google automatically checks your spelling.

Sometimes, though, Google will fail you, Lycos's results will seem off target, and AllTheWeb will dish up too many repetitive, irrelevant hits. What do you do? When we are not having much luck right out of the gate, it's time to turn to the meta-search engines.

Meta-Search Sites

Web sites that access several search engines from one place or at one search command have been around for a while. They're a great idea, really, but they haven't been that practical. Many force

users to search one engine at a time, the only advantage being that users don't have to retype their search terms. Others will parallel search several engines, but will not "de-dupe," i.e., remove duplicate results. Then, of course, the user ends up with a huge result set full of repeats. Bummer.

But now, like the Web itself, some meta-search sites are maturing. Here are some of the best. Which one you use depends on the features you like the most.

Ixquick

www.ixquick.com/

David Bodnick has created this most technologically sophisticated of meta-search engines. It is one of the few meta-search tools that can understand both natural language and advanced Boolean searches, and knows which engines can handle each syntax. It ranks its results by relevance and gives details about the ranking in the engines it queries. Search the Web, or just search news, images, or MP3 music. Try your search in English, German, Spanish, French, Italian, Dutch, or Portuguese.

Query Server

<http://queryserver.dataware.com/>

This handsome site actually contains a suite of meta-search engines. For general searches, Query Server allows you to query up to 10 of the largest search engines in one swoop. As an alternative, you can use one of three companion engines for subject-oriented meta-searches in finance, medicine and health, or U.S. government. Query Server allows you to set such options as search time, number of results, result clustering, and Boolean set.

Vivísimo

<http://vivisimo.com/>

Scientists at Carnegie Mellon University have conspired to cook up this meta-search engine that de-dupes, parses, and clusters its results. The end result looks similar to Northern Light's; that is, one page frame holds general results and another lists subject clusters.

Choose the cluster that is most relevant to your query to be rewarded with a nest of appropriate results. Vivísimo differs from Northern Light in that it queries many engines at the same time, then delivers search results that are the cream of the crop.

ProFusion.com

www.profusion.com/

Intelliseek, the company that gives us InvisibleWeb.com (www.invisibleweb.com/), has developed this meta-search site that can search over 1,000 sources divided into more than 200 search groups. ProFusion also searches at least 500 sources from the "Invisible Web," that is, databases and other information neglected by traditional search engines.

Metor Search

www.metor.com/

German company Metor Development Group brings us this fast and clean meta-search engine. Metor handily deletes redundant results from up to 10 general search engines. It also features 24 subject meta-search channels including sports, news, games, and health. Because of its country of origin, Metor Search has a European slant, and can be searched entirely in German.

qbSearch

<http://qbsearch.com/>

qbSearch will scour up to 18 regular ol' search engines at the same time. It lets users display up to 20 result pages at once, all joined together into one long page. "That way," says company founder Marc Fest, "you can scan results much faster." Search results open up in new windows to make it easy to pop back to the original search page.

SurfWax

www.surfwax.com/

California-based SurfWax has a laid-back, mystical feel to it that some love and others hate. It operates in two frames, the right

showing the databases searched, the left showing the results, ranked by relevance. You can ask that SurfWax “focus” on a word in your search, thus giving it more weight. You can also set up personal preferences to make SurfWax work well for you. Ohmmmm.

InfoGrid

www.infogrid.com/

Roger Lee Heath, an inventor from Glenview, Illinois, thought up this meta-search engine. Type in a search for known items, or choose links to subject search sites to do a thorough Web search.

Karnak Library

www.karnak.com/

Oh, you’ve got a lot of work to do on a long-term project. Register at the Karnak Library and have it leisurely search the Web. It will e-mail its results to you within 8 to 20 minutes. For monthly fees ranging from \$10 to \$40, Karnak will store your search results on its server, allowing access from wherever you log on. If you need in-depth results and have time to wait, this is the meta-research tool for you.

Deep Web Search Engines

Not all Web-enabled data is lying out there in HTML format. Much information lies concealed in databases or in collections of .pdf files (written in Adobe Acrobat). Traditional engines cannot mine this data; it is “invisible” to them. Thank goodness there are folks who have fashioned directories for us so we can get access to these obscure gems.

Invisible Web

www.invisibleweb.com/

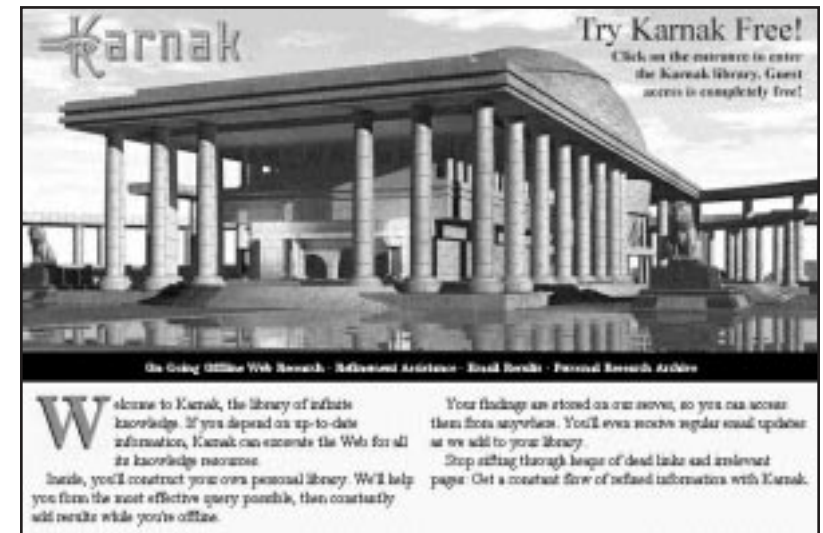


Figure 1.2 Let Karnak take its time collating research for you.

Here is the original directory of “invisible Web” resources in many different categories: finance, the sciences, and shopping. Browse it by subject category, or search it by keyword or phrase.

Invisible Web Directory

www.invisible-web.net/

Chris Sherman and Gary Price wrote the book about the invisible Web. No, I really mean that they *really* wrote the book called *The Invisible Web*. Here is their subject directory of data-rich resources untapped by normal search engines.

CompletePlanet

www.completeplanet.com/

Want to know what’s on the “Deep” or “Invisible” Web? *CompletePlanet* offers this list of subject-specific search engines and searchable databases. When we last scanned it, we found links to the seachable database on a site called “Albania Today” (www.centraleurope.com/albaniatoday/) from “Central Europe Online” and the archives of the “rec.food.cuisine.jewish”

newsgroup (www.cyber-kitchen.com/rfcj/), featuring recipes for cauliflower kugel and "Borscht, Cold Beet: Microwave"). Now that's *deep* Web!

Browser Searchbots for the Desktop

Maybe now you are so excited by meta-searching that you want to do it from your desktop! Here are a couple of outstanding free-ware programs that allow users to search many engines at once.

Copernic 2001

www.copernic.com/

Copernic 2001, offered by Copernic Technologies Inc., offers concurrent searches of up to 80 search engines in seven categories. A few seconds later, it will start to display results on the screen as it finds them, without duplicates. Then, it sorts them in relevance order. Go ahead and use fuzzy Boolean. It lets you know which of the engines you picked doesn't support it. It works great. And, it's free.

A more extensible version of Copernic, Copernic 2001 Plus, is available for \$39.95. This version allows users to add up to 93 categories of specialty search engines to their meta-searches.

Another advantage? No advertising banners!

WebFerret

www.zdnet.com/ferret/index.html

WebFerret searches nine major engines and de-dupes the results quickly and simply. Version 4.0 is a cute little thing, only 590 KB. It's tiny, free, and it packs a lot of power.

Keeping Up with the Engines

As the Internet is always in flux, so are the search engines. Where can you go to keep pace with the engines?

Search Engine Watch

<http://searchenginewatch.com/>

Fortunately, there is a fellow who makes it his job to stay on top of the changes. His name is Danny Sullivan, and you are probably familiar with his site, Search Engine Watch. Sullivan covers everything remotely search-engine related, including techniques for successfully submitting your own site to the engines. He even covers specialized search engines, those tuned to certain subjects such as news, shopping, children, medicine, legal topics, and the invisible Web.

SearchDay Newsletter

www.searchenginewatch.com/searchday/searchday.html

Chris Sherman, associate editor of Search Engine Watch and co-author of the book, *The Invisible Web: Uncovering Information Sources Search Engines Can't See*, keeps you informed about Web searching trends in a friendly, companionable way. Read his daily column on the Web, or subscribe to have it pushed to your inbox (www.searchenginewatch.com/searchday/index.html).

Search IQ

www.zdnet.com/searchiq/

ZDNet helps out with a page of in-depth reviews of meta-search engines. These are listed in descending order based on the IQ power of the site assigned by the reviewer.

Search Engine Showdown

www.notess.com/search/

Research librarian Greg R. Notess offers analyses of search engines, including charts showing their relative strengths and weaknesses. He points the way to several subject-specific engines, and offers searching tips to get the most out of all of them. This is a great site for the analytically inclined.

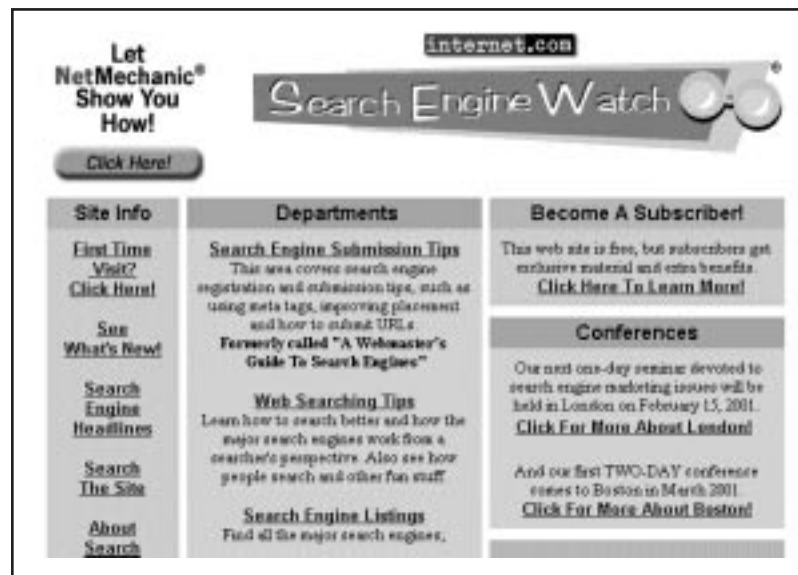


Figure 1.3 Keep up with changing search tools at Search Engine Watch.

Searching Rule of Thumb

Sometimes, what we want to find on the Web just isn't there, or isn't listed in a search engine. For example, our city manager recently asked me to find her daughter's soccer schedule on the Internet. She had seen the URL on a sign at the playing field; she knew it existed. But I looked for an hour and still couldn't find it.

Whether we use Google or the most powerful meta-search engine, a good rule of thumb for both our patrons and ourselves is to devote no more than 20 minutes to any search. If we can't locate the page we want on the Web within that time, it probably isn't there. Or maybe it just hasn't been written properly, with powerful meta-tags that attract search engines. In any case, it's out of our control. This can be hard to explain to our patrons, who think that everything is easily accessible on the Web—or should be.

The city manager returned to her daughter's soccer field, and this time she wrote down the URL on the sign. Sure enough, it pulled up the Web page that gave her the information she wanted. I felt especially embarrassed because I, supposedly the searching expert, was unable to fulfill my city manager's simple request. Oh well, we can't win them all!

