Mary Ellen Bates
Bates Information Services
Washington, DC USA

Advanced Search Techniques:
How the Super Searchers Mine for Gold on the Web

Two of the books I have written have focused on the research techniques of "super searchers" – *Super Searchers Do Business: the online secrets of top business researchers* (Information Today, 1999) and *Super Searchers Cover the World: the online secrets of international business researchers* (Information Today, 2001). While I cannot cover all aspects of how to become a super searcher during this presentation, I will provide you with at least some of the tools that the super searchers I have interviewed use to find information more efficiently.

First Item: Think!

Before you even start searching, be sure you know what you are looking for.
- White papers, statistics or position papers?

- Articles?

- Conference papers? Book chapters?

- News?

- Images? Sound files?

- Web site material or only more traditional sources?

- Plain ASCII, Excel spreadsheet, HTML, XML or other format?

- Material too current to appear in a CD-ROM or web search engine?

- Material too old to appear online at all?

- Full text or abstracts?

- Introductory or advanced material

- Lots of information or just a few good articles

Then think about turning your "searching" into "hunting", as Chris Sherman as said. Assume that you will be able to find something. If you keep coming up with nothing, try broadening your search by eliminating some search terms, eliminating date or field restrictions, and so on. We researchers often have higher relevance requirements than

our clients. Often, the client wants the best you can find, even if it isn't a perfect match. Assume that not everything is online. We (and our clients) sometimes forget that sometimes the best sources are NONline.

Watch for references to "gray literature". Government agencies, international organizations and professional associations often publish reports and other literature. Try a search with the words *report, white paper, study,* and *survey* to identify this hidden material.

Be sure to value your time. Does your client really want *everything* on the subject? We can see how far out the information horizon extends, but often our client does not want us to go all the way to the horizon.

Next Item: Move Beyond Search Engines

Search techniques need to vary, depending on what resource you are using; what works best for the professional online services does not necessarily work best with web searching.
- Complex searches do not work on the web
- The order of the search terms matters on the web
- Forget precision on the web, and go for what will most likely float to the top

Use tools, not search engines. Much of the open web is not captured by search engines, for many reasons – the content has never been linked-to, it is too new, or it has changed recently. Think about other ways to find the information, including subject-specific directories and other finding tools.

Change your approach to search for sources, not just information. You can use the web to identify experts who you can call for more information, government agencies that would collect information, and so on. The search techniques you use to find sources can be much broader than those used to find "the answer."

Follow the clues. In Google, for example, note the *Similar pages*: links and note pointers to the Google Directory (also known as the Open Directory Project). Also use

---

the *related:www.whatever.com* syntax to find similar pages in Google. Likewise, Teoma.com lists "Refine" terms to help you focus your search, and "Resources" sites that are link-rich sources. Alta Vista uses its Prisma technology to provide suggested terms to further define your search.

Use reverse web link look-up for "more like this". This tool lets you find pages that link to a particular page, the assumption being that if one web page is relevant, pages that point to it are probably also going to be relevant. The syntax for Google and AltaVista is `link:`*www.BatesInfo.com* (where www.BatesInfo.com is the URL you want to find links to). The syntax for AllTheWeb is *www.BatesInfo.com*, from the Advanced Search page, pull down [in the link to URL].

Remember the invisible and opaque web. This includes material on a site that requires registration; any information that is in a database; much of the content that is not in HTML, ASCII, PDF or Word; information that is time sensitive; material that is buried deep in a web site and so not spidered by search engines.

Use type-of-document search terms to focus your search. If you are looking for market research on the web, use words like *units shipped, forecast, projected, market share,* and *shipments* as well as words that indicate the industry that you are researching. You might also try the words *graph or table or chart.* To find white papers and opinion pieces, search for *.pdf, .rtf* and *.doc*; often, opinion papers are published on the web in a format that preserves the formatting.

ALWAYS use several tools for the same task. No two search engines have exactly the same index. In fact, there is surprisingly little overlap among search engines, so you are likely to find different results when searching with the same words in two different search engines. HotBot makes this easy; it lets you switch among AllTheWeb, Google, Inktomi and Teoma.

Every six months, purge your "assumptions cache". The resources that were ideal six months ago may be outdated or superseded by better, more complete, more updated sources today. Always watch for new resources in your subject area.

Know what you are looking for. Searching the web can be distracting, so be sure to stay focused and keep in mind what you are researching.

Keep a history of where you have been. Save a list of the web sites you visited – it is useful for the next time you do research on the topic, it is an excellent way to promote your expertise to your client, and it helps you from going around and around in circles. If it was a particularly successful search, consider building a web-bibliography for others to use.

Knowing When to Stop

I have always believed that what makes a good researcher or librarian is the presence of the "finder gene" – that urge to dig deeper, to try one more approach, to look through one more source to find the very best answer. It is a valuable trait but, like many skills, can be taken to an extreme.

Since I charge for my time, I have to be careful to stay focused and to concentrate on the sources and research techniques that are most likely to be fruitful. But I have that finder gene too, and once I am on the chase, it is hard to stop. I have developed a checklist to remind myself of when I need to end my search. The following items are from my own checklist and from the interviews with the 30 super searchers I interviewed.

It's time to say "enough!" if:
- You run out of time. You have to have the answer by 10:00 and it is now 9:59.
- You start seeing references to the same sources over and over again; you start feeling like you're walking in circles.
- You've found enough information, and doing an exhaustive search is not appropriate for the project.
- You ask a colleague or subject expert for suggestions, and he or she does not suggest anything you have not already seen.
- You are doing research in a field or geographic region that simply does not have that many information sources, and you have exhausted those sources.

- You sense that the 80:20 rule is in effect. That is, you think that you have found 80% of the information, but that the time, effort and expense to find that last 20% would match or exceed that required to find the first 80%.
- You have done everything in your checklist of usual sources (a library catalog, a web search in three search engines, a search of the major web directories and portals, a search in the professional online services).

And as one of my super searchers, Desmond Crone at Schroder Salomon Smith Barney, said, "I don't ever really stop. I file away the projects that didn't come to a satisfactory conclusion, and just keep an eye out for anything else that might be useful. Oddly enough, there does seem to be kind of a serendipity about these things. If you have looked for something, and bashed your head against all the brick walls you could find, then two weeks later you'll find a little snippet in *The Wall Street Journal* or in the *Financial Times* or in an industry magazine. That's how life is."

# # #