Knowing Where to Look: an Overview of Search Engines

by David A. Sweet, MLS

"The next best thing to knowing something is knowing where to find it," said eighteenth-century English lexicographer Samuel Johnson. This axiom is as true today as it was in Dr. Johnson's time; perhaps more so, due to the ever-growing store of knowledge on the Internet.

Search engines are the primary tools we use to find the information we need or want on the Web. Although complete books can and have been written on how to search the Web, we will briefly discuss types of search engines, give an overview of some of the major engines, how to search them, and suggestions for more in-depth search engine usage.

A study of Internet users conducted last year by Berrier Associates showed that more than 75 percent of Web users use search engines to navigate the Web.1 The study also reported that 44 percent of users say they are frustrated with Web navigation and search engine use.

Single Search Engines

Search engines generally fall into two categories: single search engines and meta search engines (which search across several single engines). Among single search engines, the Information Advisor reported that Google (www.google.com) is the top-rated business search engine.2 According to test results, Google consistently ranks as the most useful and relevant sites of all engines tested.

Greg Notess, an authority on search engine evaluation, reports in his search engine evaluation that Google's strengths include its size as well as the relevance of its results, which is based on sites' linkages and authority.3 Google also now indexes portable document files (.pdf). Notess adds, however, that Google has limited search features: no nesting (the ability to build a complex search by nesting or joining simple search concepts into a single, more complex search), no truncation (the ability to search all forms of a word, e.g., comput* would search for computers, computing, computerized, etc.), and does not support full Boolean searching. Boolean searching is the ability to combine concepts or words in a search using the Boolean connectors AND, OR, and NOT. For example, searching A AND B would result in items that contain both topics. Searching A OR B would give you all of topic A plus all of topic B. Searching A NOT B would give you all of A that does [not] contain B.

Northern Light (www.northernlight.com) is another very popular and powerful search engine. Introduced in 1997, its strengths include rich search features such as Boolean, proximity, and field searching, truncation, and a searchable current news section.

One of its best features is the Custom Search Folders. Each time you perform a search, the search results are not only displayed in relevancy order, they are also automatically categorized into subtopic folders so you can drill down to more specific areas within a broad search. Other engines have attempted to do this, but none as successfully as Northern Light. A recent addition is an alert service, which runs a specified search on the Northern Light database and then sends you e-mail when the search finds new material. The alert service requires a Northern Light account, but it can be set up for free.

Meta Search Engines

Meta search engines perform a single search across several other engines. Some of the more popular are Cyber 411, DogPile, and SavvySearch. Cyber 411 (www.c4.com) searches AltaVista, DejaNews, Excite, Galaxy, GoTo, HotBot, LookSmart, Lycos, Magellan, PlanetSearch, search.com, Snap, Thunderstone, WebCrawler, What-U-Seek, and Yahoo. The default Boolean or search term connector is AND, and it does support full Boolean searching using AND, OR, and NOT. It does not search Northern Light or Google. In addition to using keywords or phrases, it allows you to perform a search using natural language. Search results are listed with "preferred" items first with a relevancy percentage listed for the remaining hits.

DogPile (www.dogpile.com) searches InfoSeek, Lycos, Direct Hit, LookSmart, Thunderstone, Yahoo, Open Directory, About.com, and AltaVista. It offers other searches for news wires, business news, weather, yellow pages, and white pages. Searches can be customized to include or exclude specific engines.

SavvySearch (www.savvysearch.com) searches All the Web, Google, AltaVista, Lycos, WebCrawler, Thunderstone, InfoSeek, Direct Hit, HotBot, Excite, Galaxy, and National Directory. It has a detailed site map, a key feature for a good Web site, but it retrieves only a limited number of hits from each search engine.

Meta engines can be convenient and time saving, but there are some important limitations. The total number of hits retrieved will probably be considerably less (usually only the top 10-15 hits) than if you did a search on each engine separately. Meta engines can also have a tendency to time out. Further, advanced search features are usually not available from a given engine if it is searched with a meta search engine.

Who Has What You Need?

How to use and search each engine can be as important as which engine to use. Whether the engine allows for Boolean searching or the use of truncation, what fields are searched and how the results are sorted are all important features. Notess' excellent comparison chart of these and other features can be found at www.notess.com/search/features.

The Search Tools Chart at www.infopeople.org/search/chart.html may also be helpful. Further, Search Engine Watch, a monthly Web-based newsletter that covers the latest developments and changes in search engines, is available free at http://searchenginewatch.com.

Many experts agree the best place to begin searching the Internet is by using one or more of these general engines. However, there are others that may be useful when searching healthcare or HIM topics. One of the better health-care engines is Health A-Z (www.healthatoz.com). Others include On Health (www.onhealth.com), Medical Web Search (www.mwsearch.com), and Medical Explorer (www.medexplorer.com).

Though the Internet puts Dr. Johnson's adage on knowing where to find knowledge in a bright new light, the changing and growing ways of searching the Web through search engines makes the process much more manageable.

Note: For more HIM resources, try the following search services: PubMed (www.ncbi.nlm.nih.gov/PubMed), a service of the National Library of Medicine, indexes the Journal of AHIMA, as well as MEDLINE and numerous other health journals. The National Information Center for Health Services Administration (www.nichsa.org) provides AHIMA members library and research services. The NICHSA Web site features searches on topics vital to HIM professionals.

Notes

1. "Internet Usage High, Satisfaction Low; Web Navigator Frustrates Many Consumers." Study conducted by Berrier Associates, April 2000.

2. "Popularity Engines: They're Hot But Do They Work?" Information Advisor 12, no. 7 (2000), pp. 1-4.

3. "Search Engine Showdown: The Users' Guide to Web Searching." Available at searchengineshowdown.com.

David Sweet (david.sweet@ahima.org) is the manager of the FORE Library of AHIMA.

Article citation:
Sweet, David A. "Knowing Where to Look: an Overview of Search Engines." *Journal of AHIMA* 72, no.8 (2001): 28,30.
