Broadband unlimited: Search like an expert

Expert searching

There's more to searching than just typing a word into Google. Belinda Weaver explains how to get the information you need first time, every time.

These days people search the Internet for everything from cheap flights to prospective partners. Yet how many people make the most of everything Google has to offer? Enter the query "diet" and Google will use that word stem to find variations — dietary, dieters. Use the tilde symbol (~) to get synonyms. Searching for ~food returns nutrition and cooking results as well. For many queries, you can browse in the Google Directory. This by-subject guide allows you to zero in on what's of interest — Commodore cars, say, not Commodore computers.

If you only ever want Web results in English, save that preference. Set the desired number of results per page to 50 or100 while you're at it. You'll save time otherwise wasted waiting for results to load. If a promising page seems to have vanished, click on Google's "cached" link in Results.

Power search options are available in basic search, though the step-by-step Advanced Search is easier to use. Power shortcuts include `allintitle:`, `allinurl:` and others designed to narrow searches down. `Allintitle:digital television` will return sites with 'digital television' in the title. Use `intitle:` if you have only a single word in mind. `Site:` followed by a domain name allows you to search within one specific site. For example, entering `site:www.uq.edu.au courses` should get you to study options at the University of Queensland. Using `site:` with a domain type alone, `site:.gov`, for example, will retrieve only sites of that domain type. Google also offers `allinurl:` and `inurl:` if you are seeking phrases or words within a Web address.

Found one terrific site? Use Google's `related:` option to find others — `related:mydoctor.com`, for example. Google now offers calculations — simply type them in and hit return for the answer. It also provides dictionary look-up via its `define:` option — for example, `define:bioscience`; `define:genetics`, and so on.

Enclose any exact phrase in quotes, such as "herpes zoster", the medical term for shingles. Add other words (such as "treatment" or "symptoms") as necessary. Words are linked by Google's implicit `AND`. Enter the most important words or phrases first, as word order influences what results you get. Should you want to include a stop word or numeral to a search, just add a plus sign (+) before it, as in `World War +1`.

In Advanced Search, exclude older material by limiting date parameters. Consider limiting searches to specific file formats, such as PDF. The query 'herpes zoster' treatment `filetype:pdf` turned up a goldmine of patient information sheets.

Google Answers (answers.google.com/answers) offers a fast turnaround Q&A service. You have to pay to use it, but it's one option for those in a hurry. Alternatively, try AskNow (www.asknow.gov.au). It's free, it's online 10 hours a day, works in real time via chat, and you can ask just about anything. A trained librarian will help you with your query.

Although loads of useful information is stored in thousands of different databases online, for a range of reasons many search engines can't index it. Most database information is served up 'on the fly' — generally in response to typed in queries — so is too transient to be indexed. Information protected by firewalls, fees or passwords is also 'invisible'. Search engines are also poor with non-text material, such as Flash animations. To find something that currently falls outside the Google universe, investigate the world of the invisible Web.

The invisible web

The invisible Web is not only around four times the size of the visible Web, it also has some of the most useful stuff online. Bibliographic databases are searchable files of journal articles. They are a fantastic source of information on new research, and essential for academic study. Scirus (www.scirus.com) is a giant gateway to article searches in the sciences. It provides links to high quality science sites as well. Profusion (www.profusion.com) is another iWeb tool.

The inability of even the biggest search engines to mine this invisible Web is their biggest limitation. Gary Price's Direct Search (www.freepint.com/gary/direct.htm) shows the scale of what search engines such as Google, AltaVista and AllTheWeb miss.

Virtual reference desks

Whenever you search, think first about what you hope to achieve. Are you looking for a fact? A brief introduction to a subject? Deep background? You will need different strategies for each.

In the first two cases, an online encyclopaedia or factbook may be more useful than Web searching. Such tools structure information into manageable chunks, and cover an enormous range of subjects. Topic-specific ones such as the PC Webopedia (www.pcwebopedia.com) can also be handy.

To find look-up tools, use Virtual Reference Collections. VRCs contain everything you'd expect a library to have — dictionaries, directories, thesauri, encyclopaedias, maps, atlases, factbooks, biographical dictionaries — as well as currency converters and translation tools. Go here for basic, factual information. It's quicker than searching and the information is more reliable. Try the Internet Public Library's Ready Reference collection (www.ipl.org).

Defining terms

The right search words can make all the difference. With any kind of medical question, always look for both common and clinical names, "glandular fever" and "mononucleosis", for example. Master the use of the Boolean operators AND, OR and NOT, as these provide a structured way to link terms effectively. AND allows you to narrow down searches. If you search for "economy AND Australia", results will include only sites where both terms appear. The NOT operator excludes terms, and must be used with caution.

Use OR when you aren't sure which terms will be successful — "asylum seekers OR refugees OR illegal immigrants OR boat people". You can narrow things down from there. Clustering engines can be good for that, as they suggest new concepts. On new tool Mooter (www.mooter.com), a search for "music downloads" gave clusters for MP3,

software, bands, sheet music, lyrics and more. Vivisimo, another clustering tool, can also give you what's top in several engine databases (www.vivisimo.com).

AllTheWeb (www.alltheweb.com) and HotBot (www.hotbot.com) simplify Boolean by offering drop-down options. "All the words" equals the Boolean AND search; "any of the words", the Boolean OR. "Exact phrase" (these words in this precise order) is useful for proper names or concepts always expressed as a phrase, such as "open source programming".

Information strategy

Though the right terms make a difference, think about exactly what it is you need to know. Then try to identify what kind of organisation could tell you. Finding organisations which have names, goals and mission statements is easier than searching for abstract information.

Want the facts on immigration? If you think: Who would be gathering this kind of data?, it's not a huge leap to the answer — a statistical agency. Australia's agency is the Australian Bureau of Statistics. It has exhaustive tables of information not just on immigration, but about every aspect of Australian life. Much of the data can be compared over time. Many state and university libraries provide free access to ABS publications.

When the SARS scare was at its height, information was everywhere. But bodies such as the World Health Organisation (www.who.int) had the latest, most authoritative information. For the latest on digital television, try the regulator, the Australian Broadcasting Authority (www.aba.gov.au).

Finding the organisation which publishes the information you seek really pays off. Use the directories at www.journoz.com/orgs.html to identify them. It's always quicker to search a site that has information you want than to search more generally and waste time sorting good links from bad.

The article was featured in ***APC***, February 2004, and is online at

http://www.apcmag.com/apc/v3.nsf/0/E82D60CE91E22FBBCA256E3F00218D6E