

Savvy searching

Web-based metasearch engines

I wrote more than a year ago in this column that “savvy searchers know very well that they have to search several databases for quasi-comprehensive results on a subject” when they are using traditional databases. MEDLINE is wonderful, but you still need to search at least EMBASE for more comprehensive results about, say, the treatment of some disease. The same is true when searching the open access databases built by the crawler programs of Web-wide search engines. Google is awesome, but it does not cover more than 20 percent of even the free (open access) parts of the Web universe. You can increase the recall of relevant results when you also use the FAST (AlltheWeb), AltaVista and Teoma search engines. Searching for a topic such as treatment of acquired ocular toxoplasmosis using several search engines is not a walk in the park. You need to invoke each engine and type in your query, observing the various syntax rules of the different search engines. At that time I recommended desktop-based metasearch engines which take your query and forward it to search engines which you could select. Web-based metasearch engines were quite irritating, acting more like meta-ad engines that flooded you with entries from engines touting sites based on the paid inclusion system. The situation was particularly bad when InfoSpace went on a shopping spree buying up WebCrawler, Metacrawler, Dogpile and Excite and made meta-searchers feel like night guards who keep switching among TV stations but keep getting late night commercials on every click of the remote. The situation has changed considerably in the past 15-16 months, and there are now several good Web-based meta-search engines.

Vivísimo

This brainy metasearch engine deserves credit for introducing novel features and combining them with useful but rarely used existing features. It is perhaps best known for

clustering results retrieved from search engines. Result clustering was introduced by Northern Light several years before and remained a unique and valuable feature (Figure 1).

It organised results into folders by subject, language and source and offered a way to look at various aspects of the search topic by selecting one of the folders that included, for example, documents dealing with toxoplasmosis and pregnancy without the need for the user to refine the original query with additional words to restrict the search such as (toxoplasma OR toxoplasmosis) AND (pregnancy OR pregnant women OR congenital). Vivísimo uses similar clustering but it also shows immediately how many items there are in each cluster, and with another click displays the name of the sub-clusters and the number of items therein (Figure 2).

Vivísimo added to this feature with several innovations. It shows in the last line of each entry how many of the search engines found that page, and which rank position each search engine gave the page. Apparently, the pages which got the best ranking by the most engines were ranked as the top entries in the Vivísimo result list. It is not only a great idea, but is also easy to understand and verify visually. (I rush to add that it is not yet perfectly implemented as Vivísimo does not always recognise identical pages collected by two search engines, and does not aggregate the results, as is the case with the page from the *British Journal of Ophthalmology* where item No. 1 (retrieved from AltaVista) is the same as Item No. 4 from MSN in spite of the difference in using uppercase versus lower case in the last segment of the URL, and should have been aggregated. This can be critical when an item's ranking is much lower without aggregation than with aggregation) (Figure 3).

Another novel feature is the option to sneak a peek (Figure 4) of the item without leaving the results list, let alone opening a new window, in a way similar to changing the angle of the slats of a blind instead of rolling it up.

Vivísimo offers meta-searching of the databases of five Web-wide search engines, and nearly 30 other databases from a variety

Figure 1 Northern Light

The screenshot shows the Northern Light search interface. At the top left is the Northern Light logo, a stylized ship, with the text "powered by COMPACT OpenVMS AlphaServer systems". Below the logo is a section titled "Organized by topic into Custom Search Folders™" with a list of folders: Toxoplasmosis, Pharmaceuticals industry, Biotechnology industry, Medical equipment industry, Pregnancy, Veterinary services, Health & medical services, Management, Vaccination, Dogs, wolves, foxes, coyotes & dingos, and Feline diseases & disorders. The search bar contains "toxoplasmosis" and a "Search" button. Below the search bar, it says "829 items in 583 sources for: toxoplasmosis". The results are listed as "Special Collection" items:

- Special Collection** Washington University in St. Louis Medical Science Tips: Swaddling Babies, Brain... 90% - Articles & General info: ST. LOUIS, Oct 02, 2002 (ASCRIBE NEWS via COMTEX) -- The following are news and feature story ideas from Washington University in St. Louis. For more ... 10/03/2002
AScribe Newswire: Available at NorthernLight.com
[More results](#) from this publication
- Special Collection** Toxoplasmosis in Pregnancy: Cook the Meat and Avoid Soil 89% - Articles & General info: Synopsis: Eating undercooked meat, contact with soil, and travel outside Europe, the United States, and Canada--but not contact with cats--are the major risk ... 08/15/2000
Infectious Disease Alert (magazine): Available at NorthernLight.com
[More results](#) from this publication
- Special Collection** CNS Toxoplasmosis.... 89% - Articles & General info: A week of progressively worsening weakness of his right arm sent a 22-year-old man with a history of HIV infection to his physician.... 02/01/2000
Consultant (magazine): Available at NorthernLight.com
[More results](#) from this publication

Figure 2 Vivísimo

The screenshot shows the Vivísimo search interface. At the top left is the Vivísimo logo. Below it is a navigation menu: company | products | demos | partners | press. The search bar contains "toxoplasmosis" and a "Search the Web" button with a "Search" button. Below the search bar, it says "Top 153 documents retrieved for the query toxoplasmosis". The results are listed as:

- Toxoplasmosis in Cats - Cornell Feline Health Center** [New Window] [Full Window] [Preview]
Cornell University's Feline Health Center reviews the life cycle of the parasite, the signs of illness and the public-health concerns.
URL: web.vet.cornell.edu/Public/FHC/toxo.html
Source: MSN 1st, Looksmart 3rd, Netscape 3rd, AltaVista 3th, Lycos 3th
- Division of Parasitic Diseases - Toxoplasmosis Fact Sheet** [New Window] [Full Window] [Preview]
What is **toxoplasmosis**? **Toxoplasmosis** is an infection caused by a single-celled parasite named *Toxoplasma gondii*. It is found throughout the world.
URL: www.cdc.gov/ncidod/dpd/parasites/toxoplasmosis/factsht_toxop
Source: MSN 2nd, Netscape 2nd, AltaVista 4th, Lycos 3th
- CONGENITAL TOXOPLASMOSIS** [New Window] [Full Window] [Preview]
Infection of the pregnant mother could be harmful for unborn babies. Prophylaxis, early detection and treatment can avoid Congenital **Toxoplasmosis** and many long term effects
URL: www.trojovsky.net/toxo/
Source: Lycos 3rd, Netscape 1th, MSN 7th, Looksmart 8th, AltaVista 9th
- Division of Parasitic Diseases - Toxoplasmosis** [New Window] [Full Window] [Preview]
Toxoplasmosis (TOX-o-plaz-MO-sis). Fact Sheet Toxoplasmosis. DPDx Lab

On the left side of the screenshot, there is a sidebar with a tree view showing categories like "Pregnancy (20)", "Disease (4)", "Toxoplasmosis - understand", "Abbott Diagnostics Division", "Pregnancy Conditions Toxo", "Toxoplasmosis", "Cats, Infants (4)", "Parasite, Toxoplasma gondii", "Pregnant women and their", "Diagnosis (2)", "Preventing toxoplasmosis", "Other Topics (5)", "Risk (9)", "Children (7)", and "Care (7)".


of good sites, like BBC News, CNN, Britannica, MEDLINE, etc. You may choose the databases and thus eliminate databases which you are not interested in. One deficiency of Vivísimo is that it does not include the newer search engines like AllTheWeb, Teoma or WiseNut, and – similar to most other metasearch engines – Google. But enter the youngest Turk: Fazzle (Figure 5).

Fazzle

This is a not a real word, but it rhymes with dazzle (which was already used as a URL), and in terms of letter pattern also with Google, and is the name of a good metasearch engine which I do not think will fizzle as so many other metasearch engines did.

Undoubtedly, it took some pages from Vivísimo's book, but it also added its own ideas to metasearch engines.


Figure 3 Duplicates

"acquired ocular toxoplasmosis" Search the Web  Search


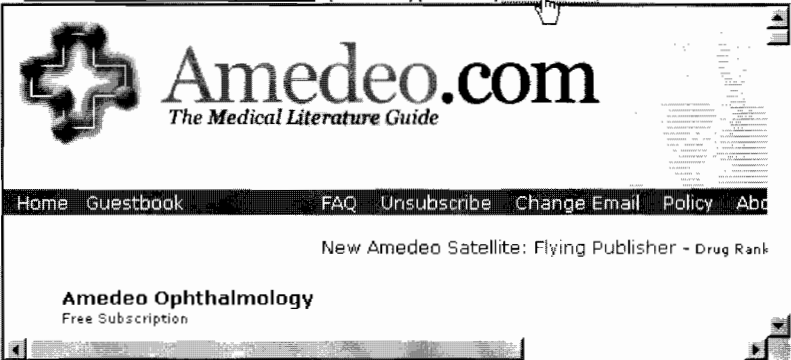
▶ [Advanced Search](#) ▶ [Help!](#) ▶ [Tell Us What You Think!](#)

1. **OPHTHALMOLOGY: Br J Ophthalmol** [New Window] [Full Window] [Preview]
 Medical Literature Review
 URL: www.amedeo.com/medicine/eye/BRJOPHTH.HTM
 Source: [AtaVista 13th](#)
2. **OPHTHALMOLOGY: Eye** [New Window] [Full Window] [Preview]
 Medical Literature Review
 URL: www.amedeo.com/medicine/eye/EYE.HTM
 Source: [AtaVista 14th](#)
3. **portal.htm** [New Window] [Full Window] [Preview]
 ... **Ophthalmology**, 1993; 100:1177-1182. 8. Hausmann, N.; Richard, G.: **Acquired ocular toxoplasmosis**. A fluorescein angiography study. ...
 URL: www.smo.org.mx/smo/75-5-7.htm
 Source: [Netscape 20th](#)
4. **OPHTHALMOLOGY: Br J Ophthalmol** [New Window] [Full Window] [Preview]
 New Amedeo Satellite: Flying Publisher - Drug Rankings in Clinical Medicine Br J Ophthalmol
 Retrieve all available abstracts of the following 370 articles: PANDA A, Bajaj MS, Sethi HS, Pushker N, Kumar H, Dada VK.
 URL: amedeo.com/medicine/eye/brjophth.htm
 Source: [MSN 14th](#)

Figure 4 Sneak a peek

"acquired ocular toxoplasmosis" Search the Web  Search

▶ [Advanced Search](#) ▶ [Help!](#) ▶ [Tell Us What You Think!](#)

1. **OPHTHALMOLOGY: Br J Ophthalmol** [New Window] [Full Window] 


Medical Literature Review
 URL: www.amedeo.com/medicine/eye/BRJOPHTH.HTM
 Source: [AtaVista 13th](#)
2. **OPHTHALMOLOGY: Eye** [New Window] [Full Window] [Preview]
 Medical Literature Review
 URL: www.amedeo.com/medicine/eye/EYE.HTM
 Source: [AtaVista 14th](#)

Its most important novel asset is that it searches AllTheWeb (although it was not available in the second round of my tests), Teoma and WiseNut, as well as the Yahoo! version of Google, and the databases of more than 100 other search engines (including some which are based on the paid inclusion system).

The GoogleYahoo! search does not produce the same result (22 items) as the direct Google search (38 items), partly because of hit counting differences, partly because PDF files are not included,

and it does not offer trademark Google features like looking up a cache version of the item, or suggesting the correctly spelled phrase if you type "acquired ocular toxoplasmosis" instead of "acquired ocular toxoplasmosis", but it does not promise the full Google, unlike many other, sloppily compiled metasearch engines which promise but never deliver anything from Google.

Its other big asset is that you can desegregate the results with the click of a

Figure 5 Fizzle

Home → Advanced Search → The Web → Web's Best

Search for: And Or Phrase Title

Your search for **acquired ocular toxoplasmosis** returned **851** matches.
GoogleYahoo: We picked the top 22 from 22 domains.

Download MSN Explorer now.
It's fast. It's free!

Show results from: [All](#) [Lycos](#) [Teoma](#) [WiseNut](#) [MSN](#) [ePilot](#) [GoogleYahoo](#)

1 2 3 [next](#) **Sort Relevance: Original Order Title Source**

My Results: [Email All](#) [Email Selected](#) [Report](#) [Options](#) [Help](#)

1 [search within this site](#) [Preview](#)
<http://search.yahoo.com/search?p=acquired+ocular+toxopla...> **100%**

2 [Experimental ocular toxoplasmosis induced in naive and ...](#) [Preview](#)
 - ... Microbiology, Lebanon, USA. We have developed a murine model to investigate the pathogenesis of acquired ocular toxoplasmosis. Tachyzoites of ...
<http://www.szp.swets.nl/szp/journals/oi071017.htm> **56.99%**

3 [ANTIBODY AVIDITY IN OCULAR TOXOPLASMOSTIS](#) [Preview](#)
 - ... With the antibody- avidity, in two cases a recently acquired ocular toxoplasmosis was proven, in three cases we found evidence for different plasmacell clones ...
<http://www.dog.org/1998/e-abstract98/444.html> **56.99%**

4 [Frequency of Specific anti-Toxoplasma gondii IgM, IgA and IgE in ...](#) [Preview](#)
 - ... An alternative strategy to distinguish a recently acquired from a congenitally acquired ocular toxoplasmosis is by measuring T. gondii-specific immunoglobulins ...
<http://www.dbbm.flocruz.br/www-mem951/3794dis.html> **56.88%**

button and show only results from a specific search engine. This is a swift and deft process, and several of the target search databases are starter team members for many savvy searchers. In a short time there will be super metasearch engines which will metasearch the metasearch engines, and create aggregate results from the

metasearch engines' results, offering clusters and sneak a peek options, on the fly aggregation and segregation of results. With the growth rate of new Web sites even the savviest searchers can benefit from the good metasearch engines.

Péter Jacsó

University of Hawaii at Manoa