

Search engines

From HLWIKI Canada

Are you interested in contributing to **HLWIKI International**? contact: dean.giustini@ubc.ca

To browse other articles on a range of HSL topics, see the **A-Z index** (http://hlwiki.slais.ubc.ca/index.php?title=Special:AllPages&from=2007_Sabbatical_-_Dean_Giustini) .

Contents

- 1 Last Update
- 2 Introduction
- 3 Searching for medicine information
- 4 Browsing & precision trends
- 5 Challenges faced by search engines
- 6 Search engines
- 7 Data search tools
- 8 Science search engines
- 9 Health specific search tools
- 10 References



Search engine wheel

Last Update

- **UPDATED** 15 September 2015

Introduction

See also **EBooks** | **FindZebra** | **Google scholar** | **Microsoft Academic Search** | **Open access** | **Search engine optimization (SEO)**

Google (<http://www.google.com>) , **Bing** (<http://bing.com>) and **Yahoo** (<http://www.yahoo.com>) are three of the top Internet search engines in 2014. As such, they have a direct impact on everyone's ability to find things. Reference librarians and information specialists are affected by search engines which are linked to other trends such as open access and web 2.0. Search engines are popular because they offer a quick search of the web. Search engines may introduce problems of various kinds in information retrieval, especially in subject-based searching. Some health librarians say that poor recall, consistency and authority control in search engine content make them unacceptable for some queries because but acknowledge that for known items they facilitate quick access and point to popular content.

In 2014, in attempt to improve the relevance of its search results, Google is exploring a combination of personalized results, localization, social networking and semantic search. In other words, two people can search for the same topic using the identical keywords and they may come up with different results.

Searching for medicine information

Medical information is of interest to a variety of user groups including patients and families, researchers, general practitioners and clinicians, as well as specialists. Despite the popularity of search engines, the development of search and access technologies in health and medicine remains particularly challenging. One of the central issues is the diversity of user groups of these services. In particular, they have varying information needs, levels of medical knowledge and language skills. Format, reliability and quality of information varies considerably in the area. A single health record may contain clinical notes, technical pathology data, images and patient-contributed histories, and may be linked to research papers. The importance of health and medical topics to the everyday lives of patients makes the need for efficient and effective retrieval of best evidence especially important. Determining the reliability of information is challenging. As with information retrieval in general, evaluating medical search technologies is a perennial challenging.

Browsing & precision trends

It was inevitable that searching would take on some of the features of web 2.0. Many of the earliest algorithms were based on link popularity, a kind of wisdom of the crowds. Now, information needs in workplaces arise in the context of collaboration and participation. At one end, health teams write collaboratively and participate in trials where comprehensive (high recall) searches are needed (though they may tolerate low precision to begin). At the other end, medical students and nurses work together to retrieve a few good articles. These health professionals do not require high recall but high precision and algorithms coupled with other forms of recommended websites produce acceptable results. As the web scales in size, new requirements emerge. Customized *social search* - offered by Google health (<http://www.google.com/coop>) - is likely to become more important as a means of offering targeted searching among a group of sites. These websites will be recommended or compiled by other workers and collaborators or other experts. However, it must be said that with the rise of search tools and social search, there is a concomitant decline in the use of traditional databases. This may be due to the simple truth that high recall and precision in searching is not required for most queries. Health library users, for example, have different requirements for literature reviews and basic information. Precision tolerance of health professionals is directly related to recall. In the Internet age, the notion of complete recall as an indicator of success seems outdated and unrealistic. Exhaustive searching is not always needed. The idea of leading users to an acceptable number of papers has led some librarians to suggest proportional recall (or relative recall) where success is expressed as the number of relevant documents retrieved, over relevant documents required. A pharmacist may need five relevant documents, but her search retrieves only three. Proportional recall is therefore three-fifths, or 60%. This measure, while appealing, is artificial in that few health library users can specify what they really need before searching, let alone how many documents they will need.

Challenges faced by search engines

- The web is growing faster than present-technology can index it. Major search-engines are slower to index new Web pages, according to anecdotal evidence. The overlap in the top three search engines of sites indexed is less than 10%.
- Many web pages are updated frequently, which makes it necessary to revisit them daily.
- Queries are limited to searching for key words, which may result in false drops, especially using the default page-wide search. Better results might be achieved by using a proximity-search option with a search-bracket to limit matches within a paragraph or phrase, rather than matching random words scattered across large pages. Another alternative is using human operators to do the researching for the user with organic search engines.
- Dynamically-generated sites may be slow or difficult to index, or may result in excessive results, perhaps generating 500 times more Web pages than average. Example: for a dynamic Web page which changes content based on entries inserted from a database, a search-engine might be requested to index 50,000

static Web pages for 50,000 different parameter values passed to that dynamic web page.

- Many dynamically generated sites are not indexable by search engines; this phenomenon is known as the deep, "dark" or invisible web.
- Some search-engines do not rank results by relevance, but by the amount of money the matching Web sites pay.
- In the past year, search engine optimization (SEO) has become big business and some techniques conspire to undermine organic results. This leads to linkspam or bait-and-switch pages which contain little information about matched phrases. Some observers suggest that relevant Web pages are being pushed down in results due to SEO.
- Google is exploring a combination of personalized results, localization, social networking and semantic search. In other words, two people can search for the same topic using the identical keywords and they may come up with different results.

Search engines

- Alta Vista <http://altavista.digital.com/>
- Ask.com <http://ask.com>
- Bing <http://bing.com>
- Excite <http://www.excite.com/>
- Gapminder <http://www.gapminder.org/>
- Google <http://www.google.com>
- Google Scholar <http://scholar.google.com>
- Infochimps <http://www.infochimps.com/>
- OAIster Worldcat <http://oaister.worldcat.org>
- Quandl <http://www.quandl.com/>
- WebCrawler <http://webcrawler.com/>
- Yahoo <http://www.yahoo.com>

Data search tools

- Gapminder <http://www.gapminder.org/>
- Infochimps <http://www.infochimps.com/>
- Quandl <http://www.quandl.com/>
- Wolfram Alpha - <http://www.wolframalpha.com>

Science search engines

- Citeseerx (computer science, engineering)
- Scirus <http://scirus.com/>
- Web of Science vs. Scopus

Health specific search tools

- Bing health - http://en.wikipedia.org/wiki/Bing_Health
- Google Coop - Health <http://www.google.com/coop/topics/Health>
- Hakia - <http://hakia.com>
- Healia <http://www.healia.com/healia/>
- Healthline – <http://healthline.com/>
- iMedix – <http://www.imedix.com/>
- intute <http://www.intute.ac.uk/>

- MEDgle – <http://medgle.com/>
- Mednar - Federated Search - <http://mednar.com/mednar/>
- MeraMD – <http://meramd.com/>
- OrganizedWisdom - <http://organizedwisdom.com/>
- PogoFrog <http://www.pogofrog.com/>
- PubMed <http://pubmed.gov>
- RightHealth – <http://www.righthealth.com/>
- SearchMedica <http://searchmedica.com>
- SUMSearch <http://sumsearch.uthscsa.edu/>
- Scirus <http://scirus.com/>
- Turning Research Into Practice (TRIP) <http://www.tripdatabase.com/index.html>
 - TRIPAnswers - <http://www.tripanswers.org/>
- Wolfram Alpha - <http://www.wolframalpha.com/examples/HealthAndMedicine.html>

See also **NNLM Super Searcher: Enhancing Your Online Search Super Powers** (<http://nnlm.gov/training/supersearcher/>)

References

- **Google scholar bibliography**
- Eureka (http://www.eureka.com/) build mini-search engines that aggregate information on particular topics from sites they choose. In the Search Engine Journal, Greg Sterling said: "Eureka... serves up progressively more relevant results (than general Web search) on the basis of communal search behavior..." about (http://eureka.com/about)
- Duran-Nelson Aa, Gladding S, Beattie J. Should we Google it? resource use by internal medicine residents for point-of-care clinical decision making. *Acad Med*. 2013. (http://journals.lww.com/academicmedicine/Abstract/publishahead/Should_We_Google_It_Resource_Use_by_Internal.99411.aspx)
- Higgins O, Sixsmith J, Barry MM, Domegan C. A literature review on health information-seeking behaviour on the web: a health consumer and health professional perspective. Stockholm: ECDC; 2011. (<http://ecdc.europa.eu/en/publications/Publications/Literature%20review%20on%20health%20information-seeking%20behaviour%20on%20the%20web.pdf>)
- Johnson BE. Bing or bust: can Microsoft cure 'search overload syndrome'? *Computers in Libraries*. 2009;29(10):36–40.
 - Microsoft's Bing is promoted as a new kind of search engine. A departure from Google's minimalist user interface, it instead integrates several navigational features that bring human- and computer-generated metadata to the front end of the search process. This article examines the extent to which the new approach is an advance over existing search engines, and the impact on information professionals.
- Callery A, Proulx DT. Yahoo! cataloging the web. *J Internet Cataloging*. 1997;1:1. (http://www.tandfonline.com/action/showCitFormats?doi=10.1300%2FJ141v01n01_06)
- Macneily A. Paging dr. Google. *Can Urol Assoc J*. 2013 Mar;7(3-4):106-7. (<http://journals.sfu.ca/cuaj/index.php/journal/article/view/571/615>)
- O'Hanlon LH. Teaching students better online research skills: improving web research tactics is a priority. *Education Week*. 9 June 2013. (<http://www.edweek.org/ew/articles/2013/05/22/32el-studentresearch.h32.html?tkn=QWCCgXpStXBSdGy+RabLBT9BSWvJPfQ47w2&cmp=clp-sb-ascd>)
- Shariff SZ, Bejaimal SAD, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, Garg AX. Retrieving clinical evidence: a comparison of PubMed and Google Scholar for quick clinical searches. *J Med Internet Res*. 2013;15(8):e164. (<http://www.jmir.org/2013/8/e164/>)
- Shenker BS. The accuracy of Internet search engines to predict diagnoses from symptoms can be assessed with a validated scoring system. *Int J Med Inform*. 2014 Feb;83(2):131-9. (<http://www.ncbi.nlm.nih.gov/pubmed/24315497>)