Effective Utilization of Web Search Engines: An Approach for Medical Practitioners

RP Pareek

INTRODUCTION

We are in the age of technological excellence. High-speed internet devices and connectivity are no more a fantasy. From clicking and uploading a picture in an instant to transferring gigabits of data across continents, we are involved in using the web for various activities of our day-to-day life. For a medical practitioner, this also includes using the World Wide Web (WWW) for application of the best available evidence to clinical practice. Medical practitioners often face unmet information needs during their work. The web provides a wealth of relevant medical information and using the web for question answering is becoming increasingly accepted by medical professionals. It is as important to the general practitioner as it is to the specialist. Previous studies of information seeking and use of information by clinicians in general practice focus on the types of questions that arise in practice and the sources used to answer them (resources used are mainly textbooks, colleagues, and journal articles held in the office). General practitioners make little use of medical libraries because of problems of access, lack of skill in using catalogs and databases, and difficulties in applying research literature to clinical situations. The web seems to provide a new opportunity to overcome problems of access and provide clinically appropriate information to practitioners. However, while use of the web for clinical information has grown substantially in recent years, problems of access, lack of skills, and applicability of information remained barriers to effective use of the internet as a source of information in general practice. An important part of utilizing the power of internet is to use web search engines to arrive at the desirable results and procuring the required information, and not just any information but critically appraised information which would yield results in clinical practice. This article tries to identify an approach for medical practitioners to effectively utilize the web search engines in the quest for practicing evidence-based medicine and improving their overall skills.

INTERNET, THE WORLDWIDE WEB AND SEARCH ENGINES

The two terms, internet and the WWW (or simply the web) is generally used interchangeably but while the internet links computers, the web links pages. Medical research and practice rely heavily on the academic publications and libraries but since the advancement of web in the early 1990s, it has become a valuable tool alongside these traditional sources.

Search engine is a program that searches documents for specified keywords and returns a list of the documents where the keywords were found. Although search engine is really a general class of programs, the term is often used to specifically describe systems like Google and AltaVista that enable users to search for documents on the web.

The web has become an important resource for health-related decisions, making it likely that the quality of information available on the web can have a noticeable effect on public health leaving us with the question. “How to increase the usefulness of the web for medical practitioners?” The solution lies in strategic search and understanding search engines.

There are, basically, three types of search engines: (1) those that are powered by robots (called crawlers, ants or spiders); (2) those that are powered by human submissions; and (3) those that are a hybrid of the two.

Crawler-based search engines are those that use automated software agents (crawlers) that visit a web site, read the information on the actual site, read the site’s meta tags (a special HTML tag that provides information about a webpage) and also follow the links that the site connects to performing indexing on all linked web sites as well. The crawler returns all that information back to a central depository where the data is indexed. The crawler will periodically return to the sites to check for any information that has changed. The frequency with which this happens is determined by the administrators of the search engine.

Human-powered search engines rely on humans to submit information that is subsequently indexed and cataloged. Only information that is submitted is put into the index.

In both cases, when we query a search engine to locate information, we are actually searching through the index that the search engine has created—we are not actually searching the Web. These indices are giant databases of information that is collected and stored and subsequently searched. This explains why sometimes a search on a commercial search engine, such as Yahoo! or Google, will return results that are, in fact, dead links. Since the search results are based on the index, if the index hasn’t been updated since a webpage became invalid the search engine treats the page as still an active link.

Companies that are dedicated to indexing the web for rating of content and prioritized listings based on rigorous review from expert panels of professionals are known as general web directories, e.g. Yahoo.com and Magellan.cc. There are also specific web directories which have indices maintained by interested individuals, academic departments, or commercial sponsors, and provide as easy and quick way of finding a large number of relevant sites. These directories are current and accurate due to the level of competition they face.
One such directory is Medical Matrix list (Source: www.medmatrix.org). If the desired results are not obtained with one search engine or directory, other search engines and directories should be tried because the quality and number of webpages varies.

**APPROACHING THE WEB: WHERE TO LOOK AND HOW?**

Strategically, searching the web to find the desirable results provides the searcher with a plethora of options to go ahead with their search and end up with better results as opposed to general search methods. The searcher can make use of the information retrieval process discussed below.

The information retrieval process is the science and practice of identification and effective use of recorded media. It can be decomposed into four processes:

1. **Indexing**
2. **Query formulation**
3. **Retrieval** (Flow chart 1)
4. **Evaluation and refinement.**

*Indexing* is the process by which the content is represented and stored in a computer database. The index allows the user to find the relevant content immediately.

*Query formulation* is the process of stating the information needed in a query, e.g. the information needed on “Can patients above the age of 60 years be prescribed fluoxetine?” can be queried as individual items joined by combination operators (AND here): prescribing Fluoxetine AND elderly.

Once the information needed is represented in a query and content is represented in an index, the query can be compared against the index, in a process called retrieval. This process ranks or sorts the output as per some criteria and gives the result to the user.

In the final process of evaluation, the user inspects the results and evaluates for relevance to the original information need. If the search results are inadequate, the user attempts refinement by rewording the query and again matching against the index. This process is repeated until the user arrives at the desired search result.

There are plenty of other operators available to use while formulating a search query to refine the results like OR, NOT (Boolean operators) or wildcard operators like *, $, - , "", ~, etc. which are applicable on some or all search engines over the internet and assist in adding or omitting results as desired.

Boolean operators form the basis of mathematical sets and database logic.

**Flow chart 1**: A graphic representation of the information retrieval process. The user retrieves content by formulating a query that contains index terms. The user may then modify the query to improve the results of the search.

- They connect your search words together to either narrow or broaden your set of results.
- The three basic Boolean operators are AND, OR and NOT.

**The Boolean Operators are Used to**

- Focus a search, particularly when your topic contains multiple search terms.
- Connect various pieces of information to find exactly what you are looking for.

Use **AND** in a search to:

- Narrow your results
- Tell the database that ALL search terms must be present in the resulting records
- Example: cloning AND humans AND ethics.

The purple triangle in the middle of the Venn diagram (Figure 1) represents the result set for this search. It is a small set using AND, the combination of all three search words.

Be aware: In many, but not all, databases, the AND is implied.
- For example, Google automatically puts an AND in between your search terms
- Though your search terms are included in the results, they may not be connected together in the way you want
- For example, this search: college students test anxiety is translated to college AND students AND test AND anxiety. The words may appear individually throughout the resulting records
- You can search using phrases to make your results more specific
- For example, “college students” AND “test anxiety”. This way, the phrases show up in the results as you expect them to be.

Use **OR** in a search to:

- Connect two or more similar concepts (synonyms)
- Broaden your results, telling the database that ANY of your search terms can be present in the resulting records
- Example, cloning OR genetics OR reproduction.

All three circles (Figure 2) represent the result set for this search. It is a big set because any of those words are valid using the OR operator.

Use **NOT** in a search to:

- Exclude words from your search
- Narrow your search, telling the database to ignore concepts that may be implied by your search terms
- Example, cloning NOT human.
Databases follow commands you type in and return results based on those commands. Be aware of the logical order in which words are connected when using Boolean operators:

- Databases usually recognize AND as the primary operator, and will connect concepts with AND together first
- If you use a combination of AND and OR operators in a search, enclose the words to be "ORed" together in parentheses.

**Examples**

- Ethics AND (cloning OR reproductive techniques)
- (ethic* OR moral*) AND (bioengineering OR cloning).

The usage of search operators can be illustrated with the help of Table 1.

Screening/filtering of search results is the processes by which search results occurring on a specific search engine can be refined by applying constraints like date of publish/upload popularity, relevance, country of origin, type of subject matter (article, book, research paper etc.) and other such filters. Filtering of search results can also be carried out for desirable file types (pdf, doc, ppt, etc.) or across specific webpages. Filtering helps the searcher to remove clutter from search and approach towards the desired information.

**Knowing the Right Place to Search**

While search engines like Google would provide the user with generally results on everything they would like to search, there are numerous webpages, databases and dedicated search engines available for medical practitioners and those interested in medicine.

Medical or Health Care search tools are not traditional search engines, but share some of the same search features that are offered by search engines.

Medical/Health Care search tools also offer access to free and/or fee-based services and collections of healthcare resources. Many medical search tools have resources for both healthcare professionals and consumers.

It is very important to pick up the right database for the information we seek to gather better results and to better choose from only relevant areas of interest. For medical search there are the well-established web databases like Pubmed, Medline, National

### Table 1: Usage of search operators

<table>
<thead>
<tr>
<th><strong>Search type</strong></th>
<th><strong>Example</strong></th>
<th><strong>Description</strong></th>
</tr>
</thead>
</table>
| **Search for an exact word or phrase “search query”** | Use quotes to search for an exact word or set of words in a specific order, without normal improvements such as spelling corrections and synonyms. This option is handy when searching for song lyrics or a line from literature (bicuspid aortic valve are more common in males).  
**Tip:** Only use this if you are looking for a very precise word or phrase, because otherwise you could be excluding helpful results by mistake. |  |
| **Exclude a word–query** | Add a dash (–) before a word to exclude all results that include that word. This is especially useful for synonyms like Cancer the disease and Cancer the astrological sun sign (cancer-astrology).  
**Tip:** You can also exclude results based on other operators, like excluding all results from a specific site [Asthma-site:wikimedia.org]. |  |
| **Include similar words~query** | Normally, synonyms might replace some words in your original query. Add a tilde sign (~) immediately in front of a word to search for that word as well as even more synonyms (~stomach ulcer) includes results for “gastric ulcer”. |  |
| **Search within a site or domain site: query** | Include “site:” to search for information within a single web site like all mentions of “Thyroid” on the New York Times web site [Thyroid site:medicalnewstoday.com/].  
**Tip:** Also search within a specific top-level domain like .org or .edu or country top-level domain like .de or.jp. [Pituitary site:edu] |  |
| **Include a “fill in the blank” query”** | Use an asterisk (*) within a query as a placeholder for any unknown or “wildcard” terms. Use with quotation marks to find variations of that exact phrase or to remember words in the middle of a phrase [“a “in blood” levels was observed”]. |  |
| **Search for either word query OR query** | If you want to search for pages that may have just one of several words, include OR (capitalized) between the words. Without the OR, your results would typically show only pages that match both terms. You can also use the symbol between words for the same effect. [Hepatitis A OR C]  
**Tip:** Enclose phrases in quotes to search for either one of several phrases. [“APICON 2011” OR “CSI 2011”] |  |
| **Search for a number range number..number** | Separate numbers by two periods (with no spaces) to see results that contain numbers in a given range of things like dates, prices, and measurements. [camera $50..$100]  
**Tip:** Use only one number with the two periods to indicate an upper maximum or a lower minimum. [world cup winners 2000] |  |

Source: www.google.com
Chapter 162  Effective Utilization of Web Search Engines: An Approach...

Library of Medicine (NLM), health on net (HON), and WebMD, etc. as well as various local and worldwide web journals, publications and search repositories.

The advantages of using Medical/Health Care search tools are:

- Their focus is specific to medicine, so users do not have to sort through non-medical resources when searching for healthcare information.
- They usually provide indices or organized subject guides that you can use instead of doing a general search. These tools understand medical terminology and are designed around such databases to handle queries of that sort hence reducing ambiguity and discrepancies from search results and thus providing the searcher with accurate medical information as called upon.
- Another advantage of using dedicated search engines or databases lies with the fact that while searching for legal, technical, social or other correlated implications of any medical condition or medicine, viz. clinical trials, public healthcare, telemedicine etc. using general search engines would provide results based on all the specified terms while the dedicated search tools will generate results from their specific indices thus removing huge amounts of unwanted information in medico-implication cases.

Disadvantages of Using Medical/Health Care search tools

- Their collections of resources and links are usually small, so a search may not generate many results.
- Some medical/healthcare search tools are fee-based only.

A few key web-based medical database and search engines/tools are listed below:

**Worldwide**

- www.omnimedicalsearch.com/
- www.webmd.com/
- www.ncbi.nlm.nih.gov/pubmed/
- www.healthline.com/
- www.healthfinder.gov/
- www.hon.ch/
- ipubmed.ics.uci.edu/(Interactive search for MedLine/PubMed)
- www.ohsu.edu/clinweb/search.html
- www.medworld.stanford.edu/medworld/medbot
- www.medspace.com
- www.gretnar.com/webdoctor/home.html
- www.pslgroup.com/DOCGUIDE.htm
- www.mwresearch.poly.edu
- www.imedsearch.com
- www.medmatrix.org

**Indian**

- www.icmr.nic.in/Publications/IJMR.html
- indmed.nic.in/indmed.html
- medind.nic.in/
- www.medconnect.in
- ctrl.nic.in (Clinical Trials Registry)
- www.indiamedicalsearch.com/

This is simply a sample from a vast pool of medical web sites on the internet which indicates clearly towards the potential that the internet holds for medical practitioners worldwide searching for information on clinical practice, research, evidence-based medicine, continuing medical education (CME), etc. Browsing through these web sites, search engines and databases further opens new paths leading towards more resources and as the user acquaints himself with these tools, it becomes more and more easier to be able to refine their information needs and querying techniques to arrive at the most accurate information effectively.

Further, for search enthusiasts who do not want to leave any page unturned there are more ways arriving as the internet grows further still. There are social networking sites like Facebook and LinkedIn to search topics and people of interest.

The Micro-blogging site Twitter has its special feature for discovering topics, the # symbol, called a hashtag and is used to mark keywords or topics in a Tweet. People use the hashtag symbol # before a relevant keyword or phrase (no spaces) in their Tweet to categorize those Tweets and help them show more easily in Twitter search, e.g. searching for #Diabetes would instantly bring up Tweets from people talking about diabetes, news, articles, current trends on diabetes, and people associated with diabetes, etc. This opens up another dimension into the world of web search and enables the user to find anything with a real-time emergence.

Google also keeps up with the pace and brings in new developments and makes searches even better every passing day with its development features. Some of these are enlisted below:

**Google Scholar**

http://scholar.google.com is the search engine giant’s dedicated engine for scholarly literature. You can search across many disciplines and sources: articles, theses, books, abstracts and court opinions, from academic publishers, professional societies, online repositories, universities and other web sites. Google Scholar helps you to find relevant work across the world of scholarly research.

**Google Books**

Google Books is the search engine for books, on any discipline, with the feature of free text search bringing up results when certain portion from a book is searched (http://books.google.com). Google Books bring up almost every book relevant to a searched topic and can help medical practitioners to browse books on desirable topics.

**Google Similar Pages**

This is an add-on for Google Chrome, the web browser from Google. This add-on helps the user to find web sites similar in content delivery based on the web site or topic searched. There are also other web sites like similarpages.com which produce somewhat similar results.

**Peer-to-Peer Sharing**

Peer-to-peer (P2P) computer network is one in which each computer in the network can act as a client or server for the other computers in the network, allowing shared access to various resources such as files, peripherals, and sensors without the need for a central server. P2P networks can be set up within the home, a business, or over the internet. Each network type requires all computers in the network to use the same or a compatible program to connect to each other and access files and other resources found on the other computer. P2P networks can be used for sharing content such as audio, video, data, or anything in digital format. P2P networks over the internet can be utilized to access databases which are “free” from influence of major players and “open” as they are referred to as. There are P2P search engines like Faroo.com and Yacy.net but these are not as developed as mainstream search engines. P2P sharing is more commonly represented with BitTorrent. BitTorrent is a protocol that underpins the practice of peer-to-peer file sharing and is used for distributing large amounts of data over the Internet. BitTorrent is one of the most common protocols for transferring large files and it has been estimated that, collectively, peer-to-peer networks have accounted for approximately 43–70% of all internet traffic. The beauty of BitTorrent is that it does not rely on the main server location for obtaining files as it can procure them from other users connected over the net. It can be utilized to search for public domain,
General

free, or open source literature, articles, books, audio-video, etc. and a medical practitioner can gain from this non-traditional source of information.

CONCLUSION

As the world progresses and mankind advents into the future with innovations and inventions the need for every individual to keep up with the times is essential. Learning from the past and building up on it to create new knowledge has been the norm of the human society and would remain to be so. With the power of internet and WWW today we can access any type of information we seek and in any form, all it takes is to search for that needle of information in the haystack of data available over the net. Medicine and the medical science have seen many developments in the course of their age and continue to do so. They are improved with inputs from researchers, practitioners, and scientists, etc. who collate information from all over the world to enhance the efficiency and effectiveness of medical practice and to do so require a lot of researching, experimenting and learning. We learn from our own experiences and from others’ and this is what makes it very important for a medical practitioner to be in touch with the ongoing work world over. The reliance on journals, books and other traditional sources needs to be supplemented with contemporary sources which are e-journals, web sites, internet databases, repositories, webpages, articles, books, publications, blogs, tutorials etc. in electronic format and also accessing the traditional sources of information separated by geographical boundaries. They also require collaborating with peers, experts, students, patients, corporations, manufacturers, and intellectuals over the matters pertaining to developing medical science and its practice for the ultimate benefit of the human race.

In order to do so, the medical practitioners of this age need to acquaint themselves with the internet and the web and to constantly strive to gain the maximum out of their usage of these tools. Efficiently searching the web using search-engines, search tools, social networks, and their ever improving adaptations is of paramount concern for them. Modeling their approach to reduce time spent and increase the quality of results obtained from these sources not only helps the practitioner, scholar or student of medicine to increase their knowledge but also adds up to the overall well-being of each and every person on this planet. Utilizing the methods discussed in this paper would assist in obtaining that efficiency and the excellence to which we all look forward to.

BIBLIOGRAPHY

5. en.wikipedia.org/wiki/
7. https://www.google.com/