searching the evidence:

Search Engines

Training Guide 9
Edition 2: June 2011
Contents:

1. How do search engines work? 3
2. Alternative Search Engines 4
3. Using Advanced Search Facility 7
4. Google Scholar 10
5. Results Comparison between an advanced search (AS) and a healthcare databases advanced search (HDAS) 11

To help you use this guide,

Indicates a step in the process of searching and retrieving articles.

Indicates a tip, or an extra piece of information.

Edition 1: January 2011
Next Revision due: June 2011

This guide was produced by Laura Wilkes and Liza Alderman on behalf of EEHIST: East of England Healthcare Information Skills Trainers
How do search engines work?

Since the creation of the internet, people have been storing and sending (posting) huge amounts of information online.

There is no mechanism for indexing and displaying the entire contents of the internet, which of course change on a continual basis. The challenge for any internet user, therefore, is to find and retrieve relevant information as quickly as possible.

The four main functions of a search engine are to locate (crawl), sort (index), store and rank (results) that information on the web.

The most popular search engines, such as Google, Yahoo and Bing, use these functions to find relevant information and make it available to you.

Functions of a search engine

Crawling (locating information)

This is achieved through the use of software known as crawlers or web spiders. The crawler sorts through the internet to find website addresses and the contents of a website for storage in the search engine database.

Crawlers have the ability to search a wide range of websites at the same time, for both new and old information, and collect large amounts of information simultaneously.

The main drawback of using search engines for research purposes is that much of the information held on specialist medical, nursing or healthcare databases will not be available to crawlers, either because there are insufficient links to those sites, or there is a robots.txt or nofollow tag attached to the databases.

For example, Google will pick up some results from PubMed and Medline, but not the other medical and nursing databases.

Indexing (sorting information)

Once the crawler has obtained the content from the internet, it sorts that content based on the occurrence of keyword phrases in each individual website. Keyword phrases are the particular group of words used by an individual to search a particular topic, for example “falls prevention in the elderly”.

Storage
Web content is stored within the database of the search engine. The amount of content available to the user is dependant on the amount of storage space available. Larger search engines like Google are able to store huge amounts of data and are able to offer a larger source of information for the user.

**Ranking (results)**

Results are the hyperlinks to websites that display in the search engine page when a certain keyword or phrase is queried. Algorithms created by the search engine designers are used to provide the most relevant data first. Each search engine has its own set of algorithms (for example, Google uses a patented and secret system called PageRank) and therefore returns different results.

**Alternative search engines**

Whilst Google has the largest index of all the search engines, it is worth remembering that other search engines exist that may produce additional/alternative results. The following is not an exhaustive list but for those keen to try, these search engines can be broken down into five categories:

**Keyword Search engines**

If you know what you're looking for, and can describe it with some key words or phrases, the following make a useful alternative to Google:

- **Yahoo Search** This was once the web's most popular search service and is still a good alternative to Google. Yahoo is the oldest major web site directory, having launched in late 1994.
- **Bing** may provide results if Google and Yahoo Search don't work.
- **Exalead** an alternative to the big three.
Index or Directory based search engines

These search engines arrange data in hierarchies from broad to narrow and are useful if you need an **overview of a subject** or you're not entirely sure of what you want.

- **Yahoo Directory** provides 14 main categories.
- **Google Directory** provides access to 16 main categories.
- Virtual libraries from [Pinakes](#). Drill down for the content/sites you need.

Multi or Meta search engines

These search engines are useful if you need to run a **comprehensive search** quickly across a number of different engines, to compare results or to suggest search engines that you may not have tried before. The majority do a Google, Yahoo, MSN, Ask search (GYMA, or GYM search depending), but there are differences.

- **Browsys** 18 search engine options. Formerly intelways.
- **Ixquick** has a number of UK based engines in its collection.
- **Joongel** 10 engines in multiple categories.
- **Searchlo** wide variety of different engines.
- **Searchboth** Compare 2 search engines at once, eight options.
- **Symbaloo** visual and multi engine, add your own engines as well.
- **Trovando** - Wide range of different resources.
- **Whonu?** Wide variety of resources, lots of options.

Visual results search engines

Rather than a simple textual list of results some search engines will provide content in a **visual format**. This is great if you want a change, or to view results differently.

- **Cluuz** provides network results.
- **Quintura** is an engine offering a tag cloud based on results.
- **Search-cube** uses a cube of thumbnails.
- **TouchGraph** for Google.
Category search engines

Some search engines will create categories for you to narrow or expand your search criteria. This is good if you don't want to think, or need some help in areas that you don't know that well.

- **Ask** (formerly Ask Jeeves) suggests different categories.
- **Carrot Clustering** is a newish engine which provides options.
- **Exalead** has related term suggestions and other methods of narrowing your results.
- **Gigablast** provides some suggested terms/phrases.
- **Google sets** creates terms from a few examples.
- **iZito** provides topics to narrow a search.
- **Zapmeta** gives users basic options to narrow searches.

Blended results

There are some search engines that will try and blend a variety of results onto one page for you - websites, news, video, images and so on. Good for an overall view of a subject area. Unfortunately there are not very many of them!

- **Allplus** for clusters, web results, news, images, video.
- **MSE360** Blogs, web pages, wikipedia, images all on one page.

* Reproduced in part with the kind permission of Phil Bradley from his “Which search engine when?” webpage: [http://www.philb.com/whichengine.htm](http://www.philb.com/whichengine.htm)
Using the Advanced Search facility

Most search engines will offer an advanced search option, which is always preferable over a basic search because you can use multiple search terms that will deliver more relevant results than a basic search.

This Guide has used the advanced search facility from Google to illustrate the above point, but most search engines offer a similar facility. We have performed a search on the topic – **falls prevention in the elderly** using Google Scholar and an Advanced Healthcare databases Search on Medline and we will compare the results later in this guide.

[Diagram showing advanced search features on Google]

- **Domain box**: Allows you to search for a particular kind of website, e.g. for NHS sites enter “nhs.uk” or for government sites enter “gov.uk”.
- **This box is the same as the Google basic search**: This box is the same as entering a phrase in inverted commas in the basic Google search.
- **File Type**: Allows you to retrieve results in a given format, e.g. if you were looking for presentations, you could restrict your results to Microsoft Powerpoint (.ppt).
- **Result box**: This box will retrieve results that contain one or more of the words that you enter. Useful for searching for synonyms – words with similar meanings.
- **This box allows you to restrict your results to pages written in the specified language**: This box allows you to restrict your results to pages written in the specified language.
More offers you access to more Google search functions including translate, and searching specific types of site, e.g. blogs, photos and Google Scholar.
### Basic search shortcuts and some other Google features

Many of the options available in Google advanced search can be duplicated using "shortcuts" in the basic search, just as you might use inverted commas to duplicate the effect of the "exact phrase" box in the advanced search.

To search for "at least one of the words", you can combine words with "**OR**", e.g. `cancer OR neoplasms`. The **OR** must be in upper case, otherwise Google will treat it as a stop word.

To search "without the words", you can exclude words by preceding them with a "-", e.g. `-children`.

To search by Domain, you can enter the appropriate URL, preceding it with "site:“. For example, `site:dh.gov.uk` would search for pages on the Department of Health website.

Search for dictionary definitions by entering a word or phrase, preceded by "**define:**", for example `define:oncology`.

### Other general tips

If you want to try and find words quickly in a large web page, hold down the CTRL key and press **F** on your keyboard. A Find box will appear that you can use to search through the page and save time scanning manually. This is exactly the same as the Find option in other Microsoft products such as Word and Excel.

If you want to visit a ".*com" site, enter the centre part of the URL into Explorer's address box, hold down the CTRL key, and press enter/carriage return on your keyboard. For example, to visit [http://www.bupa.com](http://www.bupa.com), just enter "bupa" in the address box, and press CTRL+Enter.
Google Scholar

Google Scholar will search only for academic articles. There are basic and advanced search screens very similar to the normal Google screens.

- Search in either title, or text and title of article
- Change number of displayed results per page
- Limit search by subject area, choose more than one if necessary
- Limit search to a particular journal, e.g. British Journal of Nursing
- Limit search by date range

Find articles
- with all of the words
- with the exact phrase
- with at least one of the words
- without the words
- where my words occur

Author
- Return articles written by

Publication
- Return articles published in

Date
- Return articles published between
- e.g., 1996

Collections
- Articles and patents
- Search articles in all subject areas (include patents).
- Search only articles in the following subject areas:
  - Biology, Life Sciences, and Environmental Science
  - Medicine, Pharmacology, and Veterinary Science
  - Business, Administration, Finance, and Economics
  - Physics, Astronomy, and Planetary Science
  - Chemistry and Materials Science
  - Social Sciences, Arts, and Humanities
  - Engineering, Computer Science, and Mathematics

Advanced Search Tips
Results Comparison between an advanced search (AS) and a healthcare databases advanced search (HDAS)

Full details of how to carry out a healthcare advanced databases search can be found in Training Guide 5

How an advanced search in Google Scholar displays

How a HDAS displays
Free access to full text articles

AS results will display an article citation, but it is not immediately clear whether or not you will have free access to the full text of the article.

HDAS – the link under the result confirms you have free access to the full text of the article simply by clicking on the link.

Data available

AS, particularly a search on GS, will only search for academic articles and only those which allow access to crawlers or web spiders.

HDAS – will search 8 dedicated healthcare databases and search for many different types of resources, such as systematic reviews, rcts, literature reviews, articles from peer-reviewed journals, current opinion and best practice, guidelines and government documentation and legislation.

However, many of these resources will prevent access to crawlers, so if you only perform an AS search, you will not retrieve results from many of these relevant sources.

Combining search terms

AS will search for separate concepts, e.g. elderly AND falls prevention and retrieve one set of results for both concepts.

HDAS will search each concept separately and retrieve a set of results for each concept, e.g. one set for elderly and one set for falls prevention.

Those results can be combined into one search to produce one final set of relevant, specific and focused results.
For more information please contact Basildon Healthcare Library
T: 01268 524900 ext 3594
E: library@btuh.nhs.uk
W: www.btuheks.nhs.uk