Exploring Advanced Search Features on the Web

Introduction
This document assumes that you are familiar with the use of a computer keyboard and mouse, have a working knowledge of Microsoft Windows and are familiar with using the World Wide Web.

Contents

1	Including & Excluding Terms	3
2	Phrase Searching	5
3	Field Searching	7
4	Google Advanced Search	8
5	Boolean Searching	10
6	Search Engines: What's Under the Bonnet?	12
7	MSN Search	14
8	Further References	16

About this Document

Words in **bold**	Will need to be typed or chosen from a menu or window
Small capitals – e.g. ALT	Indicate keys that you press
Press KEY1 + KEY2	Press both keys together
Press KEY1, KEY2	Press each key consecutively
Bulleted lists	Are guidelines on how to perform a task
Choose **Insert** - **Picture**	Show menu commands – in this case, choose the option Picture from the Insert menu at the top of the screen

Aims & Objectives

The aim of this document is to provide experience of using different search techniques and use advanced search features in some of the major web search engines

When you have completed these exercises you should be able to:

- Understand how to use various search strategies on search engines, such as phrase searching, truncation and using Boolean operators
- Make effective use of web search engine features and functions

Pre-requisites

This document assumes that you are familiar with the use of a computer keyboard and mouse, and Microsoft Windows based products. It also assumes some familiarity with using a web browser such as Netscape or Internet Explorer.

Copyright

© Netskills, University of Newcastle
Copyright in the whole and every part of this Courseware whether in the form of a written manual, document, software program, service or otherwise belongs to the University of Newcastle upon Tyne ("the Owner") and may not be used, sold, licensed, transferred, copied or reproduced in whole or in part in any manner or form or in or on any media to any person other than in accordance with the terms of the Owner's Licence Agreement or otherwise without the prior written consent of the Owner.
All use of this material is governed by the Owner's Standard Licence Agreement together with the appropriate Schedule. The following are available:

- A Standard Licence Schedule to cover all use including all for-profit use by any type of organisation and all use by non-educational establishments
- An Educational Licence Schedule for not-for-profit internal use only by a recognised educational establishment

The Netskills logo and this copyright notice must be included in any copy or adaptation.
Netskills is a trademark of Netskills, University of Newcastle

Contact Information

Central Computing Services, Birkbeck, Malet Street WC1E 7HX
Email: **ccs-helpdesk@bbk.ac.uk**
Phone: **UK 020 7631 6543**

1 Including & Excluding Terms

Objectives To learn how to include or exclude terms from your search engine results.
Method You will access a search engine and use the plus and minus signs to control your search.
Comments Often, when you enter keywords into a search, it is difficult to know what exactly the search is doing with those keywords – is it searching for results containing all the terms, or just one? Simple search methods can ensure you can control your search.

Activity 1.1 Go to the Google home page at the address:

`http://www.google.co.uk/`

Figure 1-1 Google UK Home Page

Beneath the search box leave the default search option set to `the web`.

Activity 1.2 In the search box shown in Figure 1-1, enter the following terms:

`russian revolution`

Click on `Google Search` and make a note of the number of results returned:

Notes

- Notice that this search has retrieved both upper and lower case versions of these search terms – Google is ***not*** case-sensitive.
- You don't need to use a plus sign in front of your terms, as Google automatically searches for both terms.

Activity 1.3 To exclude a word from a search the minus sign can be used. Try the following search:

`russian revolution -industrial`

This should reduce the number of results and make them more relevant to the Russian Revolution rather than the Industrial Revolution.

Note Be careful when excluding words – because search engines index so much text, you may unwittingly exclude results that are relevant, but which just happen to mention your exclusion term in passing.

Activity 1.4 Now try the phrase search:

`"russian revolution"`

Note the number of results:

© Netskills, University of Newcastle

The phrase search retrieves the words placed between the quotes in the order you have specified, exactly next to each other.

Using a phrase search should have once again reduced the number of results retrieved quite dramatically and improved the relevance. In this example, Russian Revolution is a commonly used phrase, and a phrase search would be the best technique to try initially.

You will learn more about phrase searching and its advantages in the next task.

Note Phrases can also be included or excluded by prefixing them with the plus or minus signs.

2 Phrase Searching

Objectives To target a search more effectively using phrase searching and wildcards.
Method You will use the Google search engine to carry out a phrase search.
Comments Phrase searching is a type of proximity search. It requires that the words you enter as a phrase are contained within your search results in the exact order you have specified.

Activity 2.1 Using Google, enter the following search:

```
meat free recipes
```

Press the Google Search button and review the results – how relevant do you feel they are? Are there any unexpected results?

See if you can work out why you might be retrieving irrelevant results.

Note Notice that this search does not use the plus signs to require words. This is because in the previous task you established that Google's default is to require all words be contained in results, so the plus signs are not necessary.

Activity 2.2 Now enter the following search:

```
"meat free" recipes
```

Press the search button and then examine your results (make a mental note of the number of results retrieved).

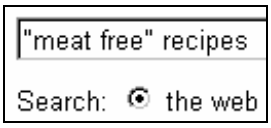


Figure 2-1 Phrase Searching in Google

You should find that the results are more relevant – this is because Google is now searching for "meat free" as a phrase, i.e. only returning results where the words appear next to each other in that order.

Previously, you may have found results that contained meat recipes, but not necessarily meat-free recipes!

Note You should also find that the number of results has decreased. This is because fewer pages are likely to contain "meat free" as a phrase.

Activity 2.3 Try the following search:

```
to be or not to be
```

Examine your search results – do they look at all relevant to Shakespeare? Several of the words in the above quote are in fact **stop words** – words that are ignored in the search.

Look at the words that have been highlighted in your results – you should find that only the word ***not*** is highlighted, and this is the only word that has been searched for.

© Netskills, University of Newcastle

Scroll further down the page – you should see Google has actually run a second search for you automatically: a search on **`"to be or not to be"`**

These results should be a lot more relevant to Shakespeare

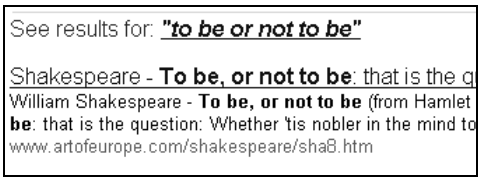


Figure 2-2 Google's automatic phrase searching

Note You can force Google to search for stop words, either by phrase searching, or by using the plus sign in front of a stop word.

3 Field Searching

Objectives To understand how to use field searching options to search more effectively.

Method You will use the title search at Google.

Comments Field searching can be useful in searching the web as it enables you to narrow a search down to specific parts of a web document, such as a title.

Activity 3.1 In Google, try the following search:

```
search engine tutorials
```

Press Search and make a mental note of the number of results returned. Review the listing and you will find that most sites are already relevant to search engine tutorials, but you can increase relevance by restricting the search to the title of the web page.

Activity 3.2 Now try the search:

```
allintitle:search engine tutorials
```

This time you are searching for all of the words in the title of a page (the title of a web page is the text that appears in the bar at the top of the browser window). The title is also listed as the main heading to each site in the search results list. You can see the search terms highlighted in the title in the figure below.

Figure 3-1 Results for Title Searching

You should find that you retrieve fewer results and the results should be more targeted to your search term.

Notes It is likely that a document which contains your term in the title will be more relevant for you than one which just mentions your keyword somewhere in the body of the document.

Activity 3.3 Try out some of the following search options:

`link:`	Restricts the search to hypertext links in documents. It can be useful if you want to find out how many sites link to a particular page. For example, `link:www.bbk.ac.uk`
`intitle:`	This will search for *any* of the words in the title of the web page. This differs to `allintitle:` used above where *all* words must be in the title.
`inurl` or `allinurl:`	Searches for a particular URL (URL = Uniform Resource Locator = web address) or part of a URL on a web page. For example, `inurl:pinakes` will retrieve web addresses containing the term `pinakes`.

© Netskills, University of Newcastle

4 Google Advanced Search

Objectives To explore the advanced search features at the Google search engine.
Method You will use the advanced search page at Google.
Comments Google generally works well using simple search queries, but it does offer several advanced search options to help you construct and restrict your search.

Activity 4.1 From the main Google home page, select the link to Advanced Search.

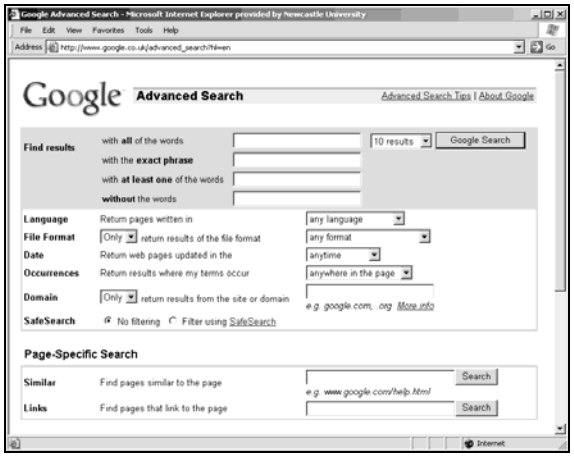


Figure 4-1 Google Advanced Search

Google's advanced search features are relatively new to the site and are continually being developed and expanded.

Activity 4.2 Under the Advanced Web Search, examine the different search options available. This search interface allows you to do a limited type of Boolean search operation using a form instead of search syntax.

Notice also that there are options to restrict by Language, File Format, Date and Domain. The Occurrences option allows you to restrict your search to certain parts of the web page, for example title or URL.

Use the form to enter the following search as shown in Figure 4-2:

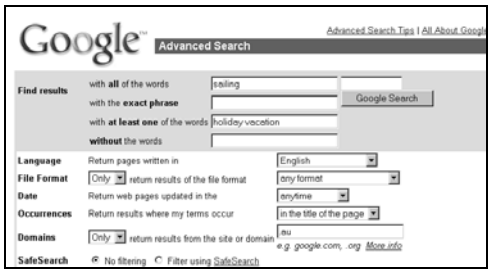


Figure 4-2 Using Google Advanced Search

Find results with all of the words: **sailing**

Find results with at least one of the words: **holiday vacation**

Language Return pages written in: **English**

Occurrences Return results where my terms occur: **in the title of the page**

Domains Only return results from the site or domain: **.au**

When you have entered the search, press the Google Search button.

Activity 4.3

Examine your search results.

You should find that your search retrieves Australian sites which have titles containing the word sailing and either the word holiday or vacation.

Note

On your search results page, examine the search box. Google has automatically entered the search query it has constructed from the information you entered in the forms, as in Figure 4-3.

Figure 4-3 Google's Search Query

Notice that this query doesn't need the AND operator – this is because by default, Google automatically combines search terms using AND.

Activity 4.4 (Optional)

Go back to the Advanced Search page and experiment with some of the other search options, such as restricting the date or file format and trying out Page-specific searches.

© Netskills, University of Newcastle

5 Boolean Searching

Objectives To learn how to structure a Boolean search query.
Method You will use Boolean search features at Google.
Comments Boolean operators and brackets can be used to create complex search queries. Not all search engines support full Boolean searching.

Activity 5.1 Go back to Google's home page at:

http://www.google.co.uk/

Google does not support full Boolean searching using the AND, OR and NOT operators and brackets. It does offer the OR operator which can be used with the plus and minus signs.

In the Google search box, enter the following search:

"meat free" recipes chinese OR indian

The search should retrieve meat free recipes for either Chinese or Indian food.

Note It is very important that the OR operator is entered in capital letters – otherwise it will be treated as the normal word or which is a stop word and as such will be ignored.

Reference The Search Engine Showdown site contains a good overview of how Boolean can be used at Google:

http://www.searchengineshowdown.com/features/google/googleboolean.html

Activity 5.2 Complex Boolean searches cannot be carried out on Google as it does not support the full range of features. AltaVista supports full Boolean.

Go to AltaVista UK at:

http://uk.altavista.com/

From the options to the right of the search box select Advanced Search (see Figure 5-1).

Figure 5-1 Selecting AltaVista Advanced Search

Activity 5.3 On the Advanced Web Search page shown in Figure 5-2, select to Search with… this boolean expression.

In the large search box enter the following search:

(roman OR greek) AND mythology

The brackets group together certain categories of words and the Boolean operators are used to control how these words are combined.
Press the Find button and review the results.

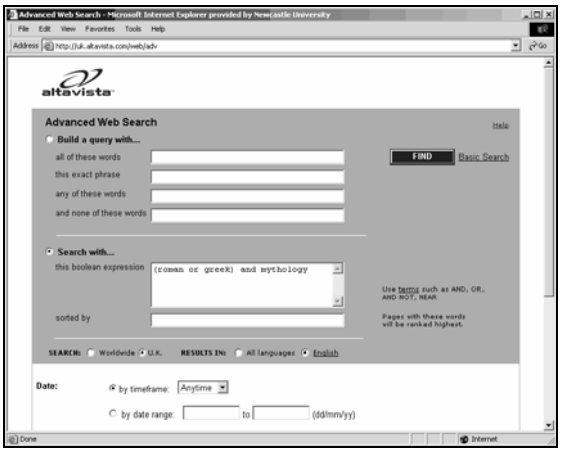


Figure 5-2 Selecting a Boolean Phrase

Note The Boolean operators are written in capital letters here for clarity – Boolean operators do not have to be written in capitals at AltaVista, but they will only work in the Advanced Search page.

Activity 5.4 The previous search was quite general – just looking for any pages on Roman or Greek mythology. The search can be refined further by using nested brackets. Alter the search to read:

((roman OR greek) AND mythology) AND dictionary

Run the search again. This time you should retrieve sites referring to dictionaries of Greek or Roman mythology.

References Opinions on the effectiveness of Boolean searching on the internet vary – the following references give both sides!

Boolean Searching on the Internet – a conventional view of Boolean searching

http://library.albany.edu/internet/boolean.html

Answering Quick Reference Queries using Google – suggests Boolean might not be as effective on Google:

http://www.freepint.com/issues/080802.htm#feature

© Netskills, University of Newcastle

6 Search Engines: What's Under the Bonnet?

Objectives To understand which search engine databases are unique, and which are owned by the same company.

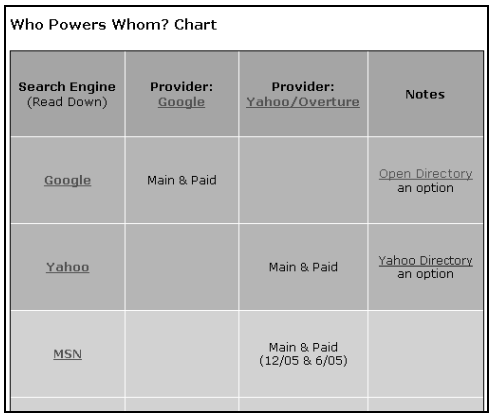
Method You will go to the Searchenginewatch.com site and learn more about search engines.

Comments Understanding which database is underlying each search engine can help you quickly identify which search engines are the best alternatives to the one you regularly use.

Activity 6.1 In your web browser go to:

`http://searchenginewatch.com/reports/article.php/2156401`

This will take you to the SearchEngineWatch "Who Powers Whom? Search Providers Chart". The chart explains the (rather complicated!) nature of commercial search engines – read through the introductory text that explains the chart.



Who Powers Whom? Chart

Search Engine (Read Down)	Provider: Google	Provider: Yahoo/Overture	Notes
Google	Main & Paid		Open Directory an option
Yahoo		Main & Paid	Yahoo Directory an option
MSN		Main & Paid (12/05 & 6/05)	

Figure 6-1 Searchenginewatch.com "Who Powers Whom? Chart"

Note This chart has not yet been updated to include the new MSN Search (launched at the end of January 2004). After previously using results provided by Yahoo!, Microsoft has now developed its own database.

Activity 6.2 Examine the chart - key points are that:

- At the time this chart was last updated, there were only 2 main search engine database *providers* – Google and Yahoo! Before February 2004, Yahoo! used to use Google's database, but after buying Inktomi, AltaVista and AlltheWeb, it launched its own database.
- Although Google and Yahoo! are the major providers, there are some others, including Teoma (owned by Ask). See the `Notes` next to each search engine.
- As well as providing normal search engine results, Google also provides paid listings to other search engines (its `Sponsored Links`). search engines, for example Teoma, may use Google's paid listings while using their own database for their main web listings.
- Yahoo's paid listings are provided by Overture, a company which it

also owns.

Note Although the same search engine database may underly 2 different search engines, there may offer different search features. For example, Yahoo! owned AltaVista offers full Boolean searching, which is not offered at Yahoo!.

© Netskills, University of Newcastle

7 MSN Search

Objectives To learn about the MSN Search engine and the features it offers.
Method You will access the UK MSN search.
Comments Surprisingly, Microsoft did not have its own search engine up until the end of January 2005. Previously, it relied on results being provided from other search companies. In January 2005 it dropped results from Yahoo! in favour of its own database.

Activity 7.1 In your browser go to the UK MSN Search page at:

`http://search.msn.co.uk/`

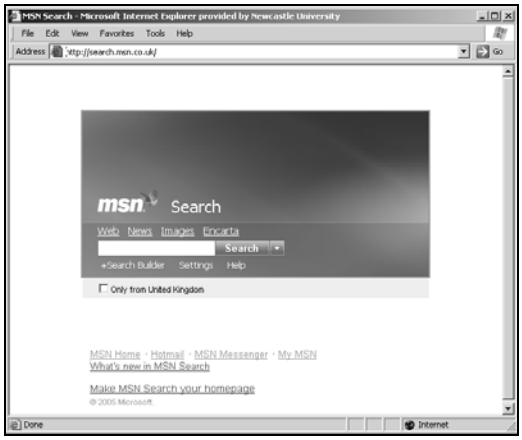


Figure 7-1 MSN Search (UK)

From the links at the bottom of the page select `What's new in MSN Search?` Watch and read the animated tour of the search features offered at MSN. (This requires the Macromedia Flash Player to view.) Once you have watched this, click on `Back` to return the MSN Search home.

Activity 7.2 One of the strengths of MSN Search is the inclusion of information from the Encarta encyclopaedia. Accessing Encarta via MSN Search means that a 'free pass' is automatically given to the user to access the content of Encarta for 2 hours – this is renewed each time a link is followed from MSN Search to Encarta.

From the options on the MSN Search home page, click on the link `Encarta`.

Now run enter and run a search for: **`wine history`**

Your results will be retrieved, similar to Figure 7-2.

Explore the results.

MSN Search thus manages to allow easy general web searching and quick access to reliable information via Encarta, through the same interface.

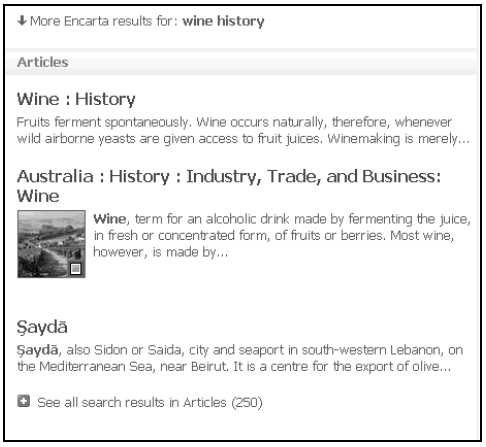


Figure 7-2 Encarta results

Activity 7.3

Go back to the MSN Search home page.

From under the search box, click on Search Builder.

This is how you access MSN Search's advanced features. The search builder allows you to select various search options and build your search step-by-step.

Click on each of the menu options and explore what search features are available.

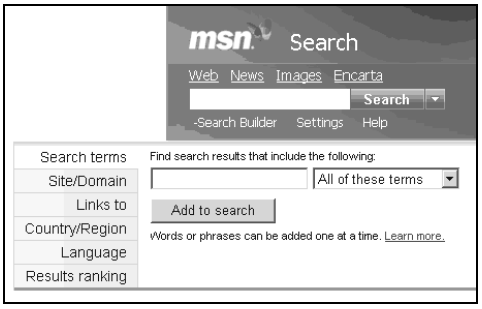


Figure 7-3 MSN Search Builder

Activity 7.4

Most of the options offered in the Search Builder are also available at other search engines such as Yahoo! and Google. However one feature that is more unusual, is the Results ranking. This allows you to influence the relevancy ranking that is done by the search engine using controls.

Click on the Results ranking option to see the controls:

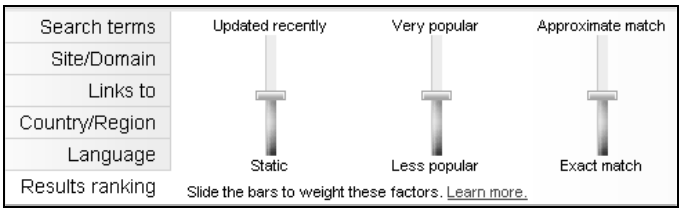


Figure 7-4 Results ranking controls

Click on the link Learn more... to find out how this works.

Activity 7.5

Create a search of your own using the Search Builder.

8 Further References

Objectives To be aware of useful sites on the internet to learn about web searching.
Method You will explore a list of references.

Searching and Evaluation, from University of Cal at Berkeley

http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/FindInfo.html

Search Engine Showdown (inc review of search features), run by Greg Notess

http://www.searchengineshowdown.com/

Search Engine Watch

http://www.searchenginewatch.com/

Including " Search Features Chart" from the following Web Search Tips page:

http://www.searchenginewatch.com/facts/

Sherman, Chris. Why Search Engines Fail

http://searchenginewatch.com/searchday/article.php/2160661

InfoPeople Search Guides

http://www.infopeople.org/search/guide.html
http://www.infopeople.org/search/chart.html

MetaCrawler MetaSpy – search engine voyeur

http://www.metaspy.com/

7 Stupid search mistakes

searchenginewatch.com/searchday/article.php/2159561