

Using the Internet for...

Research



Using the Internet for Research

by Catherine Comiskey, An Chomhairle Leabharlanna (The Library Council).

Lifesteps offers you a range of online and printed guides. Written by staff from Irish public libraries, each guide contains easy to follow, step by step explanations of how to use the Internet to achieve a specific goal or 'life step'.

Lifesteps is an initiative of Irish public libraries, managed by An Chomhairle Leabharlanna and funded by the Department of the Environment, Heritage and Local Government through the Information Society Fund.

For updates, news on the project and for a full listing of all the guides, log on to www.lifesteps.ie

You can contact the Lifesteps team at lifesteps@librarycouncil.ie

ISBN: 1873512236

The publisher would like to thank all relevant institutions for permission to reproduce their material. Every care has been taken to trace copyright holders. However, if anyone has been omitted, we apologise and will, if informed, make corrections in any future edition.

An Chomhairle Leabharlanna (The Library Council) is not responsible for the content of the websites included in this guide.

Contents

Introduction	2
Need to Know	2
Need to Have	2
Related Guides	2
Quick Guide	2
Step-by-Step: Getting Started	3
Worked Example: Search Engine	9
Worked Example 2: Directory	11
Glossary	12
Links	14
Final Notes	14

Note: words in red appear in the glossary at the end of this module.

Introduction

This guide will advise you on the best options available for your intended research. It will explore the resources available on the **Internet**, with a worked example. When you have finished with this guide you will be able to research any topic online.

Researching on the Internet has endless possibilities. Instead of spending hours manually searching through a library catalogue or indexes, travelling to and fro or making numerous phone calls, the **World Wide Web** has millions of pages filled with information available to us at the touch of a button.

Information online is constantly changing and updating. Researching on the Internet allows you to keep up to date with these changes and affords you the convenience of doing it from your home PC.

Need to Know

You will need to know the basics on how to use the Internet and a computer.

Need to Have

You need to have thought out what it is you are trying to research and **keywords** that will relate to the topic.

Related Guides

'*Internet Basics*' will cover the 'need to know' requirements for this guide (copy in your local library).

Quick Guide

1. Why research online?
2. Directories
3. Search Engines
4. Specialised Databases
5. The Invisible Web or the Deep Web
6. Referencing

Step-by-Step: Getting Started

Why research on the Internet?

The Internet is often referred to as the 'information society'. In saying this, it is meant that the Internet is full of vast amounts of information on a current scale. By accessing the Internet, you are not only able to access web pages across Ireland but also on a global scale. This allows the researcher to have access to a volume of information that would, ten years ago, have been unheard of.

Researching online is a fast and effective tool. But as with all varieties of information, there are several things that must be considered.

Who created this web page?

With the Internet you must remember that in order for you to be able to read the information someone had to place that information on line. Take time to note the author of the **web page**, if possible. Remember that the source may be biased.

Also, take a look at the **URL**. Does it end in edu; gov; org or; ie? Internet addresses ending in gov are government authored, org is an organisation that will most likely have a point of view, edu are usually education authorities, and ie/com tend to be commercially based.

One way to validate a source is to check if it has any links from the web page, and where these links lead. Is it a reputable site? Another possible answer is to try and cross reference the information that you have found on the web site with other web sites. Or maybe do a search on the author to see if they have any other web pages.

Is it up to date?

Check the date that the web page was created or last updated. Sometimes the creator will state the date at the very beginning or at the end of the page. This is especially important if it is quoting statistics or current news stories. Generally, anything that is more than five years old would only be useful as a general overview, rather than a given fact.

Remember all information is put on the World Wide Web for a purpose.

Searching the Web

There are two main ways of viewing information on the Web. The first is to **browse** the Web. This would be in the case where you are unsure of what you are looking for and therefore are quite happy to browse from page to page. The second is a direct search, when you know what it is you are looking for. The web is broken down into two search formats: directories and **search engines**.

Directories

Directories differ to search engines in that they are compiled by humans rather than computers. The web pages in a **directory** are organized into categories. This way you can search through the categories to locate your information. Most directories also have a searching facility that will search through its list of directory pages for your chosen topic.

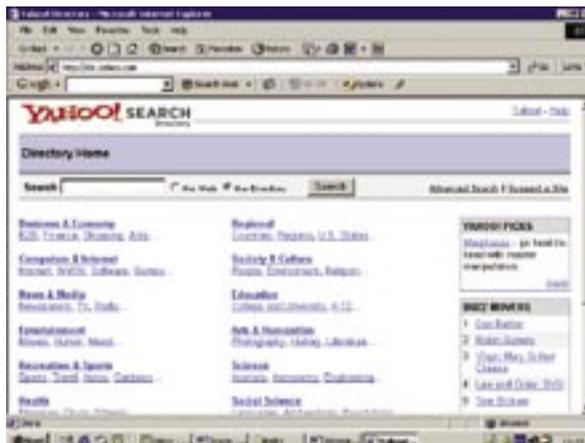
However, it is important to remember that this search facility will only search the title and introductory text, rather than the whole text.

Searching through directories does have its limits, and unless your search topic is very broad it might be best not to choose this option. It is important to remember, though, that because these pages have been filtered by human eyes, the content of the web page results will be more reliable.

Examples of directories

<http://dir.yahoo.com>

(One of the most popular general directories that contains a plethora of popular everyday categories e.g. Entertainment, Government, Health, News & Media, Science, Society & Culture).



www.academicinfo.net

(An academic directory that categorises online educational information into areas such as an educational subject directory, a list of online degree programs, and resources for examination preparation).

www.google.com/dirhp

(Similar to Yahoo! in that it is also a very popular general directory, that is categorised under headings such as Arts, News, Recreation, Shopping, and Sports)

www.joeant.com

(Categorised in the same manner as the Yahoo! and Google directories)

www.lycos.co.uk/dir/

(A British directory site that is similarly organised to the other general directories mentioned above)

www.dmoz.org

(This directory is one of the largest and most comprehensive directories available online. It is constructed and maintained by a global community of volunteer editors. The categories are similar to the other general directories)

www.about.com

(This is another general directory that is again categorised under everyday topics)

Search engines

Search engines are compiled by a computer programs called **spiders**. The spider travels across endless web pages, indexing the pages into its database as it goes. Unlike directories, search engines will index the entire text of a document and not just the title. This is called full-text indexing.

This information and its corresponding links are retrieved when you enter keywords or phrases into the **form** of a search engine home page. Each search engine has a different way of storing the information in its database. That is why when you use a search engine, what might turn up as the first result in one search engine could be the fifth result in another. Try and become familiar with a couple of search engines and their workings.



Examples of Search Engines...

www.google.ie

(This is one of the most popular extensive search engines. It permits you to search web pages from Ireland or from throughout the world. It also has the advantage of having little advertising).

www.scholar.google.com

(Google Scholar is an academic search engine permits you to search for scholarly work, including academic peer-reviewed papers, theses, books; or publications from professional societies or academic publishers).

www.altavista.co.uk

(Similar to Google, this is an extensive search engine that usually provides you with plenty of possible answers to your search queries).

www.webcrawler.com

(Webcrawler is a search engine that functions by transforming accumulated results from other major search engines onto one search results page).

www.lycos.co.uk

(This is another search engine that also provides search categories [similar to a directory], and linked commercial features on its homepage).

www.wisenut.com

(Wisenut is another example of a search engine that offers you the opportunity to run a detailed search on Web resources for a particular topic).

Boolean search

When it comes to entering keywords in a search, some search engines use a particular formula to read your command. One such format is the Boolean search.

The name Boolean comes from mathematician George Boole (1815-1964) who came up with a method to express specific queries. It is devised as a list of words or **operators** that link keywords together to help a search.

Not all search engines use Boolean operators, so again understanding how a couple of search engines work would be useful. Others automatically place these words invisibly into your search so that you do not have to. The main Boolean operators are AND, OR, NOT. They are always written in CAPITAL LETTERS.

AND: When you place AND between two words it tells the search engine that you want all of these words in the results page. For example: Galway AND Ireland, would bring up only the web pages containing Galway and Ireland.

OR: If you put OR between your keywords, you are telling the search engine to come back with results that will have either of your keywords or all. For example: cats or felines would retrieve pages that contain the word cats or felines or both.

You can increase your search further by using a combination of Operators. For example: (cats or felines) and dogs, would be telling the search engine to look for pages containing the words either cats or felines but had to include dogs.



NOT: By putting NOT in your search, you are asking the search engine to look up pages that contain one keyword but not the other. For example: recipes NOT desserts, is asking for web pages that contain recipes but not those for desserts.

You can also use the plus (+) and minus (-) signs between words, but without any spaces. For example, recipes-desserts. It will have the same results as before.

If you are looking for information on a particular subject but are unsure of how to phrase a keyword search, you can use a process known as 'stemming'. For example, if you are looking for web sites on houses or housing, in the search box type hous*. This will bring up all the web pages containing house, houses, housing, etc. However, this does not work on Google.

Specialised/proprietary databases

Some web sites exist that are generated specifically towards one area, for example medical journal, financial information and statistical data. These web pages may sometimes be overlooked by search engines or you may need to pay a subscription fee in order to access the information.

Other web sites may only be accessible through your work or an organisation. For example, most college libraries assign each student a user name and password. This allows them to have privileged access to selected web pages that ordinary users would not have.

The Invisible Web

There are millions of web pages and links across the Internet. What we access through search engines only covers a fraction of what is actually available to us.

Search engines are unable to index all these pages. Non-textual files such as multimedia files, graphic files, and documents in non-standard formats such as PDF are sometimes overlooked by the spiders of search engines (although Google now includes an option to search for PDF files that it has re-written to HTML). This area that is not data based is known as the Invisible or Deep Web. It is thought that the size of the Invisible Web could be up to 500 times larger than the surface Web according to Bright Planet Corporation.

How to access the Invisible Web?

They are many specialised search engines that claim to search the web pages of those that are 'lost' to the Deep Web. These include www.invisibleweb.com, www.profusion.com, www.incywincy.com

Referencing the Web

There are a number of formats to referencing web pages that you have used for research, or may have quoted. However, the principle (apart from some differences) remains the same as if you were sourcing a book or journal. You should include the author of the site, the date it was posted on the webpage or last revised, the date you accessed the information, the URL and the title of the work (if there is one).

You may not have all these references but try to include as many as you can. This way, people who are interested in what you have written know where to follow up on your research, if they want to get more information.

Worked Example: Search Engine

For this example, we are going to try and look for an evening course in Cork using www.google.ie. We will practice Boolean searching. Our search topic is evening classes, and we have decided to narrow down our search to Cork only.

On the Google homepage type into the search form (evening+classes) AND cork. A list of possible matches will be shown.



Look through these results to see which would suit you. Let us say for this time that the first option www.nightcourses.com/search/search_cork.html is the best option. Click on the link on the Google results to access the page.

On this website you can search for a course by keyword, type of class or county, or all of these options if you choose. There is also a section to search for distance learning. This site is very informative offering links to other colleges and courses. There is an added option whereby you can decide to receive a free weekly newsletter from Nightcourse by email.

The course we are going to search for is under languages – Italian. Click on **Languages** on the left hand side column.



Alternatively you can use the keyword section of the search form. In this case, as we are looking for Italian classes, we enter **Italian** here. Choose the county **Cork**. Click on **Go**.

We are brought to a list of our chosen results on Italian courses in Cork. It found three courses that might interest us in Douglas, Carrigaline and Ashton.

This is an excellent web site, easy to navigate and understand. It also has a features section with links to articles and advice on what language to choose for you.



Worked Example 2: Directory

For our directory search, we are going to use www.zeal.com/directory/full_directory.shtml, and our search will be for **winter gardening**.

On the homepage we are shown a list of categories to choose from. It also has a search form. Under the heading of Lifestyles, click on **gardening**. This brings you into another page that contains subheadings in relation to gardening. We can then pick **Advise by Subject**, since we know the subject we will be looking for.

Next, click on **Seasonal**, and finally **Winter**. A list of sites dedicated to winter gardening are brought up.

Alternatively, instead of browsing through the catalogues in the directory, on the homepage you can choose to search the directory. Simply type **Winter Gardening** in the search form and click on **search**. You will get similar results as if you had browsed the catalogue. This is just a quicker way of getting to the end result.

With browsing it helps if you are unsure of what you are looking for. Otherwise I would recommend that you use the search facility of the directory.

Glossary

- Browse:** This is another way to find what you are looking for online. Browsing a website is like browsing in a shop – you look at what is displayed, and see if it fits your requirements. Browsing does not require that you can describe what you are looking for; instead you take advantage of the layout of the website to find what you are looking for. Websites may use categories, headings and site maps to help you to browse successfully.
- Directory:** A database compiled by humans available for web searching.
- Form:** A form that you fill in as if it were on paper. It can be entering details for a web based search or registering with a website.
- Home page:** The main or introductory page of a website.
- HTML:** A language recognised by computers used to write web pages.
- Internet:** A large collection of computers, connected together to allow them to share information with one another.
- Keyword:** Words that are related to or can describe your intended search.
- PC:** Personal computer.
- PDF:** **(Portable Document Format)**
A PDF is a type of computer file (in the same way that an MP3 or a Microsoft Word file is a type of file), which is commonly used on the Internet for presenting documents. PDFs are excellent for printing out – a printed PDF generally looks much better than a printed (HTML) Web page. PDFs are thus mainly used for material which the user is expected to print out (such as Government forms, brochures, reports), rather than material which is meant to be viewed online. PDFs require a specific piece of software, a PDF viewer, to be viewed and printed out. By far the most common PDF viewer is Adobe Acrobat, which is available for free online.

- Search engine:** A program running on a web server computer which creates and maintains a directory of web pages on the Internet. The engine has a web page with a form; you type in words describing what you are looking for; the engine looks up its directory and sends back a list of web pages that it thinks will fit your search criteria.
- Spider:** A computer program that searches the Internet and locates web pages for indexing in a search engine database.
- URL:** The Uniform Resource Locator, or address, of a web page. Like the phone number of a particular person, this is unique and goes direct to the web page it is associated with.
- Web page:** A web page is made up of text and images mostly, and is what we see when we access a website. Think of the web pages of a website, as the pages of a book.
- World Wide Web: (WWW)**
Also called the WWW or the Web. A service that sends information over the Internet from web server programs to web browser programs. The browser then displays this information to the user on his computer screen.

Links

For further reading:

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

(This is a tutorial from the University of California that provides a good overview on undertaking research on the Internet).

www.sofweb.vic.edu.au/internet/research.htm#looking

(This is a very detailed guide to undertaking research on the Internet from the State of Victoria in Australia. It provides excellent information on the different resources for research, along with advice on evaluating the results of your research from the Web).

www.journalist.com/engines.html

This resource provides information on search engines. Although originally intended for journalists, it also provides useful information for the general public on the various search engines and on the techniques to use for using search engines.

www.webliminal.com/search

This is another useful site that provides information on searching and researching on the Internet. It is organised under various simple relevant headings.

www.bbc.co.uk/webwise/learn/

This site from the BBC is a basic online course on using the Internet. It contains an easy to understand section on finding information on the Internet.

<http://library.albany.edu/internet>

This site from the library of the University of Albany, New York, provides a research guide along with a guide to search engines, subject directories and The Deep Web.

Final Notes

Websites and the URLs are constantly being updated and modified. For every website that becomes obsolete, another one is created. The links used in this guide are only a small percentage of what is available. The more familiar you become with search facilities, the more efficient and quicker it will be to find what you are looking for.

