Why use the internet for research?

- The internet is a worldwide system of computer networks – a network of networks in which users at any one computer can, if they have permission, get information from any other computer.
- The world wide web (WWW) consists of all the resources and users on the Internet that are using the Hypertext Transfer Protocol (HTTP).
- The biggest disadvantage of the internet is that not all of the data you find will be accurate or come from a recognized authority.
Learn the language

- **Domain**
  - The last part of an internet address which indicates either the type of website, the country where the website is based, or both.

- **Domain name, URL, website address**
  - These terms all refer to the address of a website.

- **Host**
  - The computer where a web page is located.

- **Meta search**
  - A search that queries several different resources at the same time.
Learn the language (2)

- Wildcard
  - A symbol used to denote a number of missing letters in a search word or phrase, usually a ‘*’.

- Spider
  - A software program which indexes a website by following all the hyperlinks it finds on each page. This is the mechanism used by a search engine to populate and update its index.

- Bot
  - A software program that performs some job automatically.

- Hyperlink
  - Text that is highlighted on a web page and when clicked on will take you to another website, another page on the same website, a document, image, sound file or video clip.
Learn the language (3)

- **Dead link**
  - A hyperlink that no longer works because the resource it points to is no longer available.

- **Search engine**
  - A software program that will locate a website or other resource by searching an index.

- **Directory**
  - An index of websites and other resources compiled and maintained by human beings instead of a spider.

- **HTML**
  - Stands for Hypertext Markup Language, the commands used to create a web page.

- **http**
  - Stands for Hypertext Transfer Protocol, the mechanism used to get and display a web page in a browser.
Learn the language (4)

- **Web browser**
  - A software program installed on your computer that enables you to display web pages.
- **Usenet newsgroup, forum**
  - A discussion list where people come together to discuss a specific subject.
- **Mailing list**
  - Similar to a usenet newsgroup or forum but each message is sent to the discussion group participants via email.
- **Keyword**
  - A single word used in a search engine to find resources.
- **Search phrase**
  - Two or more words used in a search engine to find resources.
- **Portal**
  - A website that links to several resources covering a specific subject.
- **Results page**
  - A list of resources found by a search engine in response to a keyword or search phrase.
Before you start:

- Determine exactly what you are searching for (statistics, background information, sources, information about a particular person).
- Decide whether you are trying to confirm a fact, find new information for the first time or add to your knowledge of a subject.
- Work out what type of resources would be most useful to you (reports, articles, government records).
- Write down the name of any leading experts in your subject area. Can you consult them via email? Do they have a website?
- Determine whether the information you are seeking is likely to be available online. If it is confidential, you may not be able to find it.
- Decide which search tool is most appropriate to your needs.
- Make sure you understand how to use the advanced search capabilities of the tool.
A search strategy

- Choose your search phrases carefully
  - Choose words that can quickly narrow down your search.
  - Select words that are unique or peculiar to your subject.

- Use advanced search techniques
  - NOT, AND, OR

- Execute your search step by step
Step by step search

- Can any particular company, association, or specialist organization hold the information you need?
- If you have found a site that sort of covers what you want but does not satisfy you totally, check to see who is linking to the site you have found.
  - [http://www.linkpopularity.com](http://www.linkpopularity.com)
- If you were not able to guess the name of a suitable site, go to a directory such as Yahoo! And browse by your subject category.
- If you are looking for product information, use a specialized subject directory instead of a general directory.
- Go to Ask Jeeves at [http://www.ask.com](http://www.ask.com) and enter your search as a question, such as ‘Where can I find sites on genealogy?’.
- If you still have not found what you are looking for, then enter your search keywords (written down from your directory browsing) in the major search engines.
  - Do not enter single words in the search box.
  - Enter a search phrase.
A search engine is simply a database containing thousands of links to websites.

There are over 2500 search engines available and they all vary according to the types of websites they index, the amount of coverage of the web that they say they have, the features they have for searching, and their speed of operation.

- [http://www.altavista.com](http://www.altavista.com)
- [http://www.excite.com](http://www.excite.com)
- [http://www.northernlight.com](http://www.northernlight.com)
- [http://www.hotbot.com](http://www.hotbot.com)
- [http://www.go.com](http://www.go.com)
- [http://www.google.com](http://www.google.com)
Directories

- A directory differs from a search engine due to the fact that it does not simply consist of a long list of links. A directory takes this one stage further by grouping the links into meaningful categories.
  - [http://www.yahoo.com](http://www.yahoo.com)

- Meta directories
  - A meta directory is not a directory as such in that it does not consist of a database and it does not directly index websites. Instead a meta directory allows you to search multiple search engines and directories from one place.
  - [http://www.metacrawler.com](http://www.metacrawler.com)
Interactive Web Search Wizard

- It guides you to an appropriate search engine or directory based on one of several predefined categories.

Smartborg

- The idea behind this tool is that it analyses your search keyword or phrase and then suggests the search engine that would return you the most relevant results.

AltaVista

- AltaVista indexes around 10 million web pages a day and large parts of the database are refreshed every night.
- Perhaps AltaVista’s strongest feature is its ability to retrieve web pages based on specific search phrases.
- You can search using specific fields such as:
  - ‘title:’
  - ‘host:’
  - ‘link:’
AltaVista (2)

- AltaVista advanced search
  - AND, OR, AND NOT, NEAR
- Sorting your results
- Removing duplicates
- Using case in your search phrase
- Using wildcards
  - You can only use the wildcard after a minimum of three characters.
  - AltaVista will match the wildcard with between zero and five characters only.
- Forcing AltaVista to read your search phrase in a particular order
Excite

- Excite is best used when searching for information on a specific company and background material about that company. This is because Excite will always return the company website first if you enter a specific company name.
  - Excite precision search
  - Excite advanced search
Northern Light

- Northern Light does things slightly differently. Many librarians work for Northern Light and are responsible for organizing its site.
- It is very strong at supplying newspaper and magazine stories, which are sold from the site. This, combined with its news-filtering technology in its search engine, makes it a useful resource if you are looking for news items. You can perform industry-specific, publication-specific, or news-specific searches.
Northern Light (2)

- Using wildcards
- Field searching
- Power search
- Other useful features of Northern Light
HotBot

- HotBot provides focused search resources that allow you to quickly find resources in a specific subject area.
  - Using wildcards
  - HotBot advanced search
Go, which used to be known as InfoSeek, grades the websites and resources available in its directory as follows:

- **Best sites**
  - These are the ‘best of the best’ according to Go.

- **Very good**
  - These are sites that are rich in content and contain information that is easy to access.

- **Good**
  - These sites are functional, informative, and relevant to the topic.
Go (2)

- Search within
- Boolean searches
- The Go power search
Google

- Google claims to have features that make the results returned from your searches more relevant.
  - Boolean operators
  - Google advanced search
Most useful directories

- Yahoo! – the most popular directory
  - [http://www.yahoo.com](http://www.yahoo.com)
- INFOMINE – a scholarly collection of internet resources
- eBLAST – Britannica’s internet guide
  - [http://www.eblast.com](http://www.eblast.com)
- The internet Sleuth – a good tool to find more subject directories
  - [http://www.isleuth.com](http://www.isleuth.com)
- Argus Clearinghouse – strong in pop culture, politics, academic resources and other general subjects
  - [http://www.clearinghouse.com](http://www.clearinghouse.com)
- LookSmart
  - [http://www.looksmart.com](http://www.looksmart.com)
- Magellan Internet Guide
  - [http://mckinley.com](http://mckinley.com)
Most useful directories (2)

- About.com – covers hundreds of subject areas, very high quality
  - [http://www.about.com](http://www.about.com)
- Webcrawler
  - [http://www.webcrawler.com](http://www.webcrawler.com)
- Librarian’s Index to the internet – geared towards academic resources and maintained by professional librarians
  - [http://sunsite.berkeley.edu/internetindex](http://sunsite.berkeley.edu/internetindex)
- Open Directory Project
  - [http://www.dmoz.org](http://www.dmoz.org)
Meta tools

- Meta tools access several search engines and subject directories at the same time and can provide huge productivity gains for anyone wanting to find information.

- Meta tools do not use their own databases; instead they act as an interface to several search engines and directories, resulting in more broad-based searches which can often reveal websites and resources you may not have come across before.
More powerful meta tools

- All-in-one – uses dozens of search tools. A particularly useful feature is the ‘Other Interesting Searches’ function which provides access to an amazing collection of specialized search engines dedicated to specific topics
  - [http://www.allonesearch.com](http://www.allonesearch.com)
- Beaucoup! – access to more than 1200 search engines
  - [http://www.beaucoup.com](http://www.beaucoup.com)
- Dogpile – one of the most thorough meta tools
  - [http://www.dogpile.com](http://www.dogpile.com)
- The Internet Sleuth
  - [http://www.isleuth.com](http://www.isleuth.com)
More powerful meta tools (2)

- SavvySearch – good for searching in different languages
  - [http://www.savvysearch.com/search](http://www.savvysearch.com/search)

- Ixquick
  - [http://www.ixquick.com](http://www.ixquick.com)
Using natural language

- A good way of starting your research is to frame your search query as a question and then use a search tool that recognizes questions (or natural language queries as they are known).

- To get the best out of the natural language engine you should:
  - Keep your sentence structure simple and clear.
  - Make sure you spell all the words correctly.
  - Search for one thing at a time.
Ask Jeeves

- Ask Jeeves was probably the first major search engine to interpret natural language queries and to match the meaning of those queries to highly relevant resources.
- It provides answers from a mixture of human editorial judgment and popularity technology that learns from each question asked.
- The more questions Ask Jeeves answers, the more it learns about the popularity of those answers and their relevance to the question originally asked.
- [http://www.ask.com](http://www.ask.com)
Simpli

- Simpli is a natural search engine that goes one stage further. The search engine utilizes principles of linguistics and cognitive science in its interactions with users to place search terms in context. The company claims this provides improved search results for both novice and advanced users.
- [http://www.simpli.com](http://www.simpli.com)
Archie

- Archie is a system that will allow you to search for software, data, or text files that are sitting on public FTP servers.
- Archie covers around 2000 servers and more than 3 million files, many of which cannot be found through a search engine.
Tunneling with Gopher

- Gopher is a lookup tool that allows you to browse through thousands of documents and other resources using nothing but menus.
- You do not have to worry about how to get to a piece of information that interests you – you simply select an item from a menu and Gopher will ‘go fer’ it!
Finding people

- Write down the name of the person you are looking for, including his/her middle name, their maiden name, parents’ names, children’s names.
- Write down the last known addresses, telephone number and job.
- Select the best search tool. Generally, search engines are not a good tool for finding people, unless they provide a specialist people-finder tool.
- When entering the name of the person you are searching for always capitalize each name.
- Make sure you use the correct spelling of a surname
Better people-finding tools

- International White and Yellow pages
  - This site provides links to global telephone directories organized by continent.
  - [http://www.wayp.com](http://www.wayp.com)

- Meta-Email-Search-Agent
  - This site allows you to search multiple people finder resources in parallel. The site is based in Germany but provides searches across countries.
  - [http://mesa.rrzn.uni-hannover.de](http://mesa.rrzn.uni-hannover.de)
Better people-finding tools (2)

- World Email Directory
  - From this site you can search more than 180 million email addresses and 140 million business and telephone directories worldwide
  - [http://www.worldemail.com](http://www.worldemail.com)

- AnyWho
  - This comes from AT&T and is one of the best, containing more than 90 million US listings. It also has a reverse lookup where you can locate a person or business from their telephone number.
  - [http://www.anywho.com](http://www.anywho.com)

- Yahoo! People Search
  - A good people and email address finder.
  - [http://people.yahoo.com](http://people.yahoo.com)
Better people-finding tools (3)

- **InfoSpace**
  - Very good resource for locating government numbers. Includes telephone directories for Canada and UK.
  - [http://www.infospace.com](http://www.infospace.com)

- **Switchboard**
  - US names, telephone numbers and email addresses
  - [http://www.switchboard.com](http://www.switchboard.com)

- **WhoWhere?**
  - A good email address locator that works in English, French and Spanish.
  - [http://www.whowhere.com](http://www.whowhere.com)

- **Ultimate White Pages**
  - This is a meta directory that uses many of the people finders mentioned above.
  - [http://www.theultimates.com/white](http://www.theultimates.com/white)

- **WorldPages Global Find**
  - An international directory covering more than 60 countries.
Finding email addresses

- Internet Address Finder
  - [http://www.iaf.net](http://www.iaf.net)
Finding subject experts

- **FACSNET**
  - This site puts you in touch with think tanks, special interest organizations, government, academic and private sector experts
  - [http://facsnet.org/sources_online/main.htm](http://facsnet.org/sources_online/main.htm)

- **Sources & Experts**
  - [http://metalab.unc.edu/slanews/internet/experts.html](http://metalab.unc.edu/slanews/internet/experts.html)

- **Ask an Expert**
  - Helps you locate experts in various subject categories
  - [http://www.askanexpert.com](http://www.askanexpert.com)

- **ProfNet**
  - [http://www.profnet.com](http://www.profnet.com)
Finding reference documents

- Internet Public Library
  - A great place to start looking for reference material on all sorts of subjects.
  - [http://www.ipl.org](http://www.ipl.org)

- Digital Librarian
  - A great starting place for all sorts of subjects.

- eBLAST: Britannica’s Internet Guide
  - An online encyclopedia
  - [http://www.eblast.com](http://www.eblast.com)

- Encyclopedia.Com
  - Provides reference material in all major fields of research.
  - [http://www.encyclopedia.com](http://www.encyclopedia.com)

- Research-It!
  - Contains dictionaries, thesauruses, biographies, maps and other useful information.
  - [http://www.itools.com/research-it](http://www.itools.com/research-it)
Finding quotations

- Barlett’s Quotations
  - This site is great for finding quotes prior to the 20th century, including the Bible and Shakespeare
  - [http://www.cc.columbia.edu/acis/bartleby/bartlett](http://www.cc.columbia.edu/acis/bartleby/bartlett)

- The Commonplace Book
  - [http://metalab.unc.edu/ibic/Commonplace-Book.html](http://metalab.unc.edu/ibic/Commonplace-Book.html)

- Yahoo! Reference: Quotations
Finding FAQs

- **Usenet FAQs**
  - Usenet newsgroups are the discussion lists that thousands of people get involved in every day to discuss any subject under the sun. A quick way of learning exactly what these groups are discussing and the terms of how you can participate is to consult the FAQ for the newsgroups you are interested in.
  - [http://www.faqs.org/usenet](http://www.faqs.org/usenet)
Researching current events

- **Moreover.com**
  - The headlines on this site are refreshed up to four times an hour. They are placed into 150 categories, enabling you to view just those headlines pertaining to your interest
  - [http://www.moreover.com](http://www.moreover.com)

- **NewsHub**
  - This site is similar to Moreover. The headlines are updated every 15 minutes and you can specify your area of interest to reduce the number of headlines you see.
  - [http://www.newshub.com](http://www.newshub.com)

- **NewsNow**
  - If you are interested only in UK news, this site is for you.
  - [http://www.newsnow.co.uk](http://www.newsnow.co.uk)

- **CNN News**
  - You can configure the web page to show only news items covering your subject area.
  - [http://www.cnn.com](http://www.cnn.com)

- **Yahoo! News Coverage**
The Informant

- The Informant is an agent that will use the search engines to find websites that cover your chosen topic. It will then email you with details of the number of new websites it has found together with other websites that have been updated since you last visited them.

- [http://informant.dartmouth.edu](http://informant.dartmouth.edu)
TracerLock

- TracerLock is a free service that allows you to monitor AltaVista for the occurrence of up to five sets of keywords. Every day TracerLock will search AltaVista for pages that match your search term and were indexed three days earlier. The first ten results will then be emailed to you.

- [http://www.peacefire.org/tracerlock/](http://www.peacefire.org/tracerlock/)
Spyonit

- Spyonit is a very powerful bot. In fact you can think of it as a collection of bots some of which perform predefined searches and others which are more generic.
- The best feature is that other Spyonit users can share the bots they have defined.
- [http://www.spyonit.com](http://www.spyonit.com)
Keeping track of the news

- PR Newswire
  - http://www.prnewswire.com
- NewsBytes
  - http://www.newsbytes.com
- PR Web
  - http://www.prweb.com
- Web Wire
  - http://www.webwire.com
- Internet News Bureau
  - http://www.newsbureau.com
Copernic

- Copernic is a search tool that will perform searches on multiple search engines simultaneously. It falls into the category of a meta search engine but instead of running from a website you run it from your PC.

- [http://www.copernic.com](http://www.copernic.com)
Alexa

- Alexa is a product that is integrated with Microsoft Internet Explorer and can be a useful aid to online researchers.
Zapper

- Zapper is a computer-based search tool similar to Copernic.
- First you have to install the product on your computer by visiting the Zapper website and installing from there.
- Once installed Zapper can be popped up in a separate window and used to perform searches across the World Wide Web.
- [http://www.zapper.com](http://www.zapper.com)
Quality and usefulness of the information

- Look for clues as to the original source of the data.
- Ask lots of questions.
- Look at the website address in your browser for clues of origin and authority.
- Establish what type of resource you are looking at.
- Ask whether it is a primary or secondary source.
- Decide how relevant it is to your purpose.
Quality criteria to use to evaluate an internet resource

- Validity
- Accuracy
- Reasonableness
- Authority
- Uniqueness
- Completeness
- Comprehensiveness
- Information and site integrity
Validity

- Has the information been filtered by a third party?
- Is the information well researched?
- Is the resource available in a book format or some other medium?
- Is there any bias?
- Are there any references to documents used when researching the information?
- Is there a bibliography?
Accuracy

- Has the information been checked by someone else such as an editor?
- Can the content be cross-checked with another resource?
- Is the author motivated to want to provide accurate information?
- Are there spelling mistakes or other typographical errors?
- Are there any references or a bibliography?
- Has the information been cross-checked by the author’s peers?
- Is there a date on the document?
- Does the content include sweeping generalizations?
- Does the information change rapidly?
- Is the information one-sided or does it acknowledge opposing views?
Reasonableness

- Is the content offering a balanced, reasoned argument or is it slanted towards some view or organization?
- Is the resource objective or are the views presented linked to the sponsorship of some organization or association?
- Are there conflicts of interest between the views put forward and the author’s background (will the author benefit in some way)?
- Are wild claims being made or is emotive language being used?
- Do any arguments discussed within the document contradict themselves?
Authority

- Who is the author?
- Who has published the resource? Is it the author, a well-known publisher, or a knowledgeable webmaster?
- Can a cross-check be made of the author’s authority?
- Is the author’s email address readily available?
- Why should I believe this author over another?
- What is the author’s education, training, and/or experience in a field relevant to the information?
- Can the author be contacted (has he/she provided email or postal address, phone number)?
- What is the author’s reputation or standing among peers?
- What is the author’s position (job function, title)?
Uniqueness

- Does the resource contain anything other than a list of links to third-party sites?
- Does the resource contain any primary information? The ‘About Us’ link will give you a clue as to the amount of primary information the author has provided.
Completeness

- Are there any dead links or empty pages?
- Does the information made available have anything missing?
- Is the full text of the information made available, or are you looking at an abstract?
Comprehensiveness

- Does the resource go into sufficient depth or does it leave many points unexplained?
- Are there any obvious gaps?
- Look at the contents list. Does it cover everything you would expect it to?
- Are sources cited? If so, are they reputable?
- Is there a hyperlinked bibliography of cited sources?
- Are external links for additional information included?
- If there are any quoted statistics, have their sources been cited?
Information and site integrity

- Is the resource being updated regularly? Look for the date it was last updated.
- Are there any statements on the site about how frequently the information is supposed to be updated?
- Is there any archive information on the site?
- How frequently are documents moved to the archive area of the site?
- Is there a website version number or last updated date?
- Are all the links on the site still active?
- Slides by: Shayan PourAhmadi
  - pourahmadi@ce.sharif.edu
- Reference: