Course 6: Selecting and Verifying Information on the Internet

Outline

1. Quality of the information
2. Formalize a search in keywords
3. Boolean operators
4. Proximity operators
5. Truncations
6. Evaluation of information

Introduction

Searching for scientific information on the Internet is a complex process.

The problem of determining the reliability of information on the Internet has become increasingly important as more and more people get their information from the Internet.

Determining the reliability of some information can be a complicated matter, and it can sometimes involve energetic investigation, sophisticated instrumentation, and extensive background knowledge. Nevertheless, there are various considerations and criteria of reliability that are a function of the content of the information itself.

In this context, this course helps the student to:

- understand what determines the quality of information on the internet,
- learn to select their search information by playing on keywords and Boolean operators,
- make (better) aware that it is necessary to work with vigilance and to look above all for the channels leading to quality information.
- exercise critical thinking by working on the evaluation of information.

1. Information quality

Data quality can be one of the major challenges for research laboratories or higher education establishments today.

Four factors that qualify information, namely:

- **Relevant:** if the information is useful and if it allows you to better understand a situation, form an opinion, or make a decision.
- **Accurate, valid:** Accurate or valid information describes reality well and comes from a trustworthy source.

- **Complete**: the information is complete if it includes the main facets of the research defining the object of the research. Provides several answers.
- **Recent:** Information is recent if it is up to date and still valid.

2. Formalize a keyword search

2.1. Definition

A keyword is most often **a proper** or **common noun** or a **noun phrase without an article** which characterizes the content of a document or a subject being searched for.

Example " **cell membrane** " is a keyword for a search on "cell membrane functions".

2.2. Objective of formalization of research

The goal of formalizing keyword research is to enable research **on a subject** in the research tools that we saw in the previous course, namely: in the library **catalog , in a database** , and on the **web.**

Formalizing a search in keywords **also allows you to avoid** finding nothing (**documentary "emptiness or silence"**) or finding too many document references (**documentary "noise"**).

2.3. Principle of formalization of research

In the formalization of research:

- Do not use articles, pronouns, verbs...
- Avoid terms that describe relationships between ideas, such as "causes," "consequences," "exchanges," "operation," "functions," etc.
- Search for synonymous or equivalent terms.
  Example: "role" for "function"
- Adapt your vocabulary to the databases used.
  Example : translate words into English to query a database in English
- Use scientific terms instead of common names.
  Example: Buccal epithelium instead of mucosa
- Beware of polysemic words (which have several meanings)
  Example ***:*** **Cell:** fundamental unit of life and **cell** : room used to lock someone up.
- Specify the research topic
  Example: "plant cell" for research on the cellulose wall, "tectonic plate" for research on the lithosphere,

- Start by searching for specific terms and, if the search doesn't yield anything, for more general terms.

  Example: Start with " Template Strand" and expand the search to "Cell Cycle".

3. Boolean operators

Now that you've created a list of search terms and keywords, you need to organize them into meaningful search statements.

When you search electronic databases (such as a library catalog or an electronic index), you cannot just type in your research question ("What risk factors are associated with polluted drinking water?"). Unlike some big, commercial search engines like Google, the electronic databases won't understand what you are looking for. You need to speak in a way they will understand.

One concept to formulate effective search statements is to use ***Boolean Operators*** **(AND, OR, NOT)** to connect your keywords and concepts, like this:

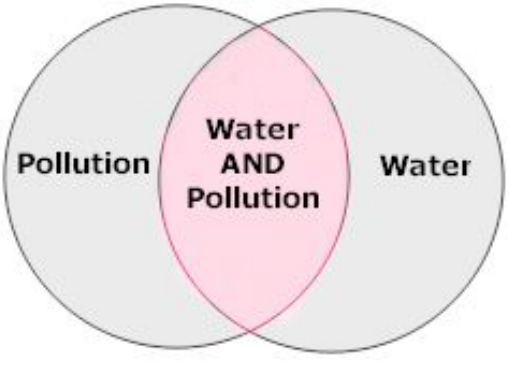
AND - *narrows* a search; Use to combine key concepts, for example: **Water AND Pollution**.

AND - *narrows* a search; Use to combine key concepts, for example: **Water AND Pollution**.

AND - *narrows* a search; Use to combine key concepts, for example: **Water AND Pollution**.

"AND"

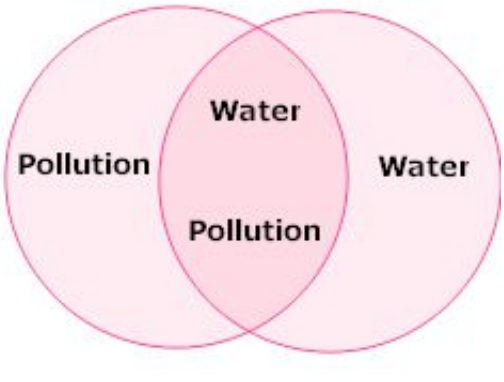
- Connects two or more words (or phrases)
- Narrows the search
- Will produce information on water and pollution



OR - *broadens* a search; Use to add concepts, for example: **Pollution OR Water**.

"OR"

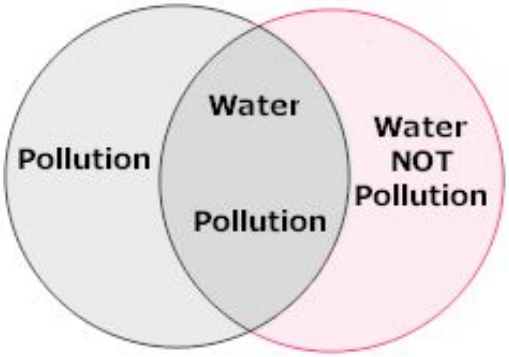
- Allows either word (or phrase) to be included somewhere in the search results
- Broadens the search
- Will produce information on either water or on pollution or maybe water pollution



NOT - *excludes* search term(s). Use to eliminate a concept, for example: **Water NOT Pollution**. **Use sparingly!**

"NOT"

- Excludes select terms (or phrases) from a search
- Warning! Use Cautiously
- Will produce information on everything to do with water except water pollution



4. Truncation and wildcard searching approach

This involves the use of truncated and wildcard searches to find variations to widen or reduce the scope of searches. ***Truncation allows for finding singular and plural terms or keywords with variant endings***. Applying truncations and wildcards is easy when using ***Boolean logic*** to combine search terms.

Asterisking keyword endings: Inserting an asterisk *(*)* at the word-ending of a keyword will automatically produce a search result for all the possible endings for that word. Many databases use an asterisk *(*)* as their truncation symbol. It is necessary that researchers apply specific truncations in their search. For *Example:* the search " Cyto *" will give: cytology, cytoskeleton, cytosol, cytosine, cytoplasm…

Using variant spellings: Using *OR* to capture variant spellings (e.g., neighbour *OR* neighbor) will lead to searching for the variant keywords inclusively.

Exacting phrases*:* Enclosing terms in quotation marks *("")* will lead to a search for that specific term or quote.

5. Information evaluation

This involves applying some evaluation criteria that will help determine whether the information is trustworthy. The four main groups of criteria are:

- the credibility of the source,
- the quality of the content,
- the quality of the intellectual organization of information,
- and the presentation of information.

5.1. Credibility of the source

The credibility of a web page is measured by the possibility of identifying ***the person responsible*** or the source. It is necessary to know what the expertise of the site's ***host is, whether the latter is linked to an institution or an organization***, and whether the objectivity of the information is evident.

Some questions to ask yourself to assess the credibility of the source:

- Who is responsible for the site?
- Is the page host recognized?
- Is there proof of validation of the information (editor, reviewer, editorial board, etc.)?
- Is the site trying to inform or sell?

5.2. Quality of content (mentioned in point 2 of this course)

5.3. Quality of the intellectual organization of information

Some questions to ask yourself to evaluate the intellectual organization of information:

- Is the site easily accessible?
- Is it easy to get to the desired page?
- Is the site well structured?
- Does it have a map, a search engine?
- Is navigation easy?

5.4. Presentation of information

Some questions to ask yourself to evaluate the presentation of information:

- Are the pages pleasant to read?
- Do you move around easily?
- Can the information be printed or downloaded in a convenient format?

If viewing the site requires special software, does it provide the means to obtain it?