Search techniques and strategies
In this session we’ll speak about

- Scientific information
- Information sources – Categories
  - Primary sources
  - Secondary sources
  - Tertiary sources
- Information sources - types
  - Databases – different types
  - Repositories
  - Aggregating systems
- Search techniques and strategies
  - What and how to search – defining your topic
- Scientific publishing and Bibliometrics
What is Scientific Information?

It is current, relevant, accurate information whose author is qualified to speak on the subject and whose objective is to be impartial, objective and to promote the evolution of science.

It’s information certified through the process of peer review, and published in scientific journals.

What is Peer Review?

It is the process by which scientific publishers ensure the quality of their publications. It consists in the submission of the works proposed for publication (articles) to the evaluation of one or more specialists in the field, designated by the publishers and invited among the world's leading specialists in the various scientific / disciplinary areas.
What are scientific articles?

- It’s the main medium used for the formal communication of science
- Allow researchers to communicate to peers the results of an investigation
- They are written by researchers
- Have peer review
- They are published in journals with mechanisms of knowledge certification
Types of Articles

**Scientific article:** (also known as original research articles)
It describes first hand the results of a study / research.

**Review article:**
It is a type of article that organizes and critically evaluates previously published studies in that field of expertise.

**Rapid Communications ou Letters**
Smaller than regular articles, they are a mean of briefly disclosing the work in progress on a ongoing project.
The publishing process is faster.
More appropriate for areas where information has a short lifespan.
Proceedings are publications through which the various papers / communications that have been submitted, approved and presented at a given congress or conference are known.

Nowadays it is common for International Conferences, workshops or Seminars to submit the communications to a review process.

They may or may not have been previously published.
Types of Information Sources

**Primary sources** - contain original information about the subject, that is to say, when the information is expressed by the 1st time:

- Theses
- Research articles reporting new results published in scientific journals
- Scientific and technical reports
- Conference proceedings
- Statistics, interviews, surveys
- Books and articles presenting original ideas

In some cases, the primary sources are documents that constitute themselves the research object, like: letters, diaries, comics, etc. They are also referred to as background material so as not to be confused with academic publications introducing new research in the curriculum area.
Secondary sources:
These analyze, interpret and comment the primary sources;
They summarize and structure information from primary sources

- Books and articles (review articles) that report or summarize the findings of others, i.e. a summary of existing knowledge
- Library catalogs are included in this category
Types of Information Sources

Tertiary sources - these are specialized works that cover «a set of knowledge or concise explanations related to themes, authors, works, associations, resources, etc.». These sources catalog, select and organize information from primary and secondary sources. "(Faria, Pericão, 1999)

- Reference books from the different scientific areas - allow familiarization with subject-specific terminology and help you to get a general idea of a subject (handbook, textbook)
- Encyclopedias
- Dictionaries
To ensure that the source material is reliable, it is best to use primary sources in your academic work. As the information is analyzed, it becomes less accurate...
Where to start?

Retrieved from: Seleção de Fontes de Informação Científica – 2016 http://www.slideshare.net/bibliotecasUA/selecao-de-fontes-de-informacao-cientifica-2016
Different types of databases

- Multidisciplinary
- Thematic
- Specific or from 1 publisher only
Aggregators

B-On

Nova Discovery

Google?
What are aggregators?

- They’re tools that join in a single search point several databases and resources.
- Advisable for a first approach to the most recent publication in the various disciplinary areas.
- They have the great advantage of saving time as they are very comprehensive.
- They don’t always present a structured indexing language, since they collect information from different information systems.
- They are not appropriate when you need specific/specialized information.
# Databases versus search engines

<table>
<thead>
<tr>
<th>Data Bases</th>
<th>Search Engines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both are tools that serve to answer an information need</td>
<td>Searchable websites databases</td>
</tr>
<tr>
<td>Subscribed, in open access, repositories, OPAC's = online catalogs</td>
<td>Searchable websites databases</td>
</tr>
<tr>
<td>= (books, scientific journals, etc.)</td>
<td></td>
</tr>
<tr>
<td>They are organized by information specialists to meet the needs of</td>
<td>The information is automatically managed by computer</td>
</tr>
<tr>
<td>researchers</td>
<td>programs</td>
</tr>
<tr>
<td>They contain subscribed information (inaccessible to the general public)</td>
<td>They contain information accessible to the general</td>
</tr>
<tr>
<td>and open access (accessible to the general public)</td>
<td>public</td>
</tr>
<tr>
<td>Where we can find quality information specifically directed to</td>
<td>There is no quality control, information is not</td>
</tr>
<tr>
<td>researchers</td>
<td>always complete and is not always reliable</td>
</tr>
</tbody>
</table>
Subscribed databases

- Academic institutions pay to provide the best resources to their users
- These resources are not accessible to the general public
- In-campus access is direct
- From the outside the “Campus” it is done by authentication (VPN)

As a member of the FCT, you have access to B-on where you can find, among other resources, databases in full-text and reference databases that being similar can operate differently
Science and Technology Databases

Scopus®

IOP Publishing

American Institute of Physics

IEEE Xplore Digital Library

ScienceDirect™ makes sense.

Engineering Village

SAGE journals online

ProQuest Information and Learning

Taylor & Francis

EBSCO

Taylor & Francis Group

ASCE

American Society of Civil Engineers

ASME

Nature

Blackwell Publishing

WILEY ONLINE LIBRARY
Open Access databases and resources

- RCAAP - Repositórios das várias universidades portuguesas
- OpenDOAR (The Directory of Open Access Repositories)
- DOAJ (Directory of Open Access Journals)
- Doab (Directory of Open Access Books)
- PubMed
- WorldCat
- Scielo (Scientific Electronic Library Online)
- PLoS (Public Library of Science)
- BioMed Central
Statistics databases

- Eurostat
- INE (Instituto Nacional de Estatística)
- Pordata
It is a subscribed resource, made available by NOVA University of Lisbon for all its members.

NOVA Discovery is a content aggregator that integrates the catalogs of the various UNL libraries, B-On and Google Scholar.

Searching in this platform within the Campus does not require authentication, outside the Campus it is done by authentication (VPN)
NOVA Discovery
Repositories
What are Repositories

- They are information systems with scientific and academic content available in Open Access.

- They include scientific articles, as well as the so-called gray literature such as master's and doctoral theses, preprints, reports, conference communications

- They are associated with a teaching and / or research institution

- In the repositories you can find the digital file of the referenced document, in full-text format and in open access.

- They allow greater (international) visibility to the scientific production of teaching and / or research institutions.
Where to start

Start with B-on and from there go to specific information sources
Why start with B-On

➢ It is a federated search engine that will retrieve information within the various databases it has access to

➢ It allows you to identify from which platform the best results are retrieved

➢ Allows us to “go” to 1 specific resource

➢ It is a multidisciplinary resource

➢ It is a trusted resource

➢ The information is current
But, … Not everything is online!

Don’t forget Books!

➢ Many are not available in full text on the Internet.
➢ Many contain historical content not found elsewhere

Search Library Catalogs!

Don't forget **scientific journals in paper format**, as not all of them are available in digital format.

**Use the Interlibrary loan** service to find articles that you can’t access directly through the subscribed databases
Interlibrary loan

➢ It is a service that allows the user to obtain documents or copies of them from another institution.

➢ This service can be used by all members of the NOVA community (internal users) and similar domestic and foreign institutions (external users).

➢ For details contact your Library
Search for information
The first steps of a scientific project
A concept map is a simple way to visually display the concepts and relationships among ideas/subjects.

This will help you to organize your ideas and define your topic.

They work as a brainstorming exercise and can be used to capture ideas/thoughts, or to take notes.

They allow you to organize your work, structure a presentation, review notes...

They should be done manually, with colors to make visible and explore in more detail the relationships between ideas and/or concepts.
How it’s done ...

➢ By hand (although there are softwares to do them) and with colors to make visible and explore the relationships between concepts and stimulate our imagination and our ability to associate.

➢ First, write down the main idea in the center.

➢ Draw a circle around the idea. This will be the starting point for the concept map.

➢ A partir do centro e em forma radial adicionam-se ramos de acordo com as possíveis subdivisões temáticas.

➢ A partir dessas linhas adicionam-se ramos secundários para especificação dos subtemas.

➢ Para além das palavras pode conter imagens, post its ...

• Like this:
Mind map or Concept map
Mind map on vegetarianism

- Varieties of Vegetarianism
- Ethics and Diet
- Health Benefits and Concerns
- Environment and Diet
- Religion and Diet
- Demographics
Mind map on vegetarianism

- Veganism
- Varieties of Vegetarianism
- Demographics
- By Country
- Animal Treatment
- Ethics and Diet
- Food Safety
- Health Benefits and Concerns
- Nutrition
- Longevity
- Local Food
- Environment and Diet
- Sustainability
- Buddhism
- Christianity
Mind map softwares

Chrome Web store

MindMapr
A Google Chrome tool

MIND MAP MAKER

For Macs

The basic version is free

MindNode
Web based.
Maps can be saved as images

The basic version is free; available for Windows, MAC OS and Linux

bubbl.us
Search techniques and strategies
When searching a database, you need to question the system as clearly as possible.

Most databases have simple and advanced search features.

Simple search allows the use of natural language and / or keywords but can retrieve many results because the system searches all fields of the bibliographic records.

Advanced search gives us the possibility to use controlled language, associate terms, and search by specific field…
Natural language vs Controlled vocabulary

<table>
<thead>
<tr>
<th>Natural language</th>
<th>Controlled vocabulary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Words taken from natural language (a good way to start)</td>
<td>Pre-defined controlled vocabulary to describe the content of the document in a library catalogue</td>
</tr>
<tr>
<td>More flexible to search (allows multiple combinations)</td>
<td>They allow more exact searches but they’re less flexible, because you need to know the exact term (you can always ask the librarian)</td>
</tr>
<tr>
<td>The database looks for the keywords anywhere in the record (not necessarily connected together)</td>
<td>You are looking for material only in the subject heading or descriptor field</td>
</tr>
<tr>
<td>It may yield too many or very few results depending on whether it is a more or less common theme.</td>
<td>When you have too many results, you need to add other subject heading to filter (date, type of document, author)</td>
</tr>
<tr>
<td>May yield many irrelevant results (you need to add filters)</td>
<td>Results are usually very relevant to the topic</td>
</tr>
</tbody>
</table>

Results are usually very relevant to the topic.
Searching by keyword?

Not always the best way

WHY?
Databases use controlled vocabulary!

... and give better results if properly searched
Start by defining the terms that best describe your topic, set the limits to your search, and build a “search expression”.

Analyze the results

In the articles abstract check the field Descriptors and / or Subjects, see the terms used there to describe the document and compare it with your search terms.

If necessary, redo the search using these controlled terms.

See if the database has an online thesaurus to browse or subject headings that match your topic (check the Help screens)
Search techniques

- Boolean operators
- Truncations or “wild cards”
- Quotation marks
- Parentheses
- Search Limits
- “Stop words”
- E no Google
  - / + / OR (the space works as an AND)
Traditional search model: pyramid

1. Start with more general terms.

2. Add other search terms by linking them with boolean operators.

3. Are the retrieved results interesting for your search?

4. If not, further refine the search by adding + terms and / or limits.
The inverted pyramid search model

1. Start with specific terms

2. Use related or broader terms

3. If you still don't get results, move up the hierarchy of concepts.

This model became possible with the new technologies.
Search strategies

5 basic strategies:

- Help button (in databases)
- Boolean operators
- Truncations
- Nesting or term associations (use parentheses)
- Controlled vocabulary (thesaurus/subject headings)
The importance of good research skills

Simple search – allows you to search by

  - Keyword
  - Author
  - Title
  - Subject

Returns many results, but the records retrieved may not match your specific needs,

Advanced search

  - Using the Boolean operators (AND, OR, NOT)
  - Truncation or Wild cards
  - OR other techniques

It’s a more precise search
Build up your search expression using:

**Boolean operators**

- “Reading disorder” OR Dyslexia AND University OR Academic students NOT PhD Students

**Truncation or wild cards**

- like * $ ? Usefull for replacing characters
- Ex:
  - know* = know, knowing, knowledge, knows
  - genetic* = genetic, genetics, genetically
Simbols that replace characters

- The asterisk and question mark should be placed at the end of the word, immediately following the root of the word, making it possible to retrieve singulars, plurals, and derivations of the term.

  The asterisk is considered the most flexible (replaces a maximum of 5 to 7 characters)

  Ex: know* = know, knowing, knowledge, knows

  ➢ genetic* = genetic, genetics, genetically

- The dollar sign and question mark when placed within the word (at any point) replace individual characters. They are indicated to help you search for the same term with different spellings.

  (ex: colour = Ing; color = EUA)

Há bases de dados que utilizam os símbolos ! # como truncatura.
Veja a função HELP da base de dados
Quotation marks

➢ When using an expression with multiple words, we must enclose it in quotation marks so that the database understands it as an “exact expression”.

➢ Ex: “Total quality management”
Parentheses (nesting)

- They are used to compose search expressions and combine related terms or search synonyms.
- Clarify the order of search terms
- Assess the results

Ex: (“Reading Disorder” OR Dyslexia) AND Students
OR
Ex: Students AND (“Reading Disorder” OR Dyslexia)
The use of acronyms

Don’t use them!

Unless you associate them with the full expression, so that the database “understands” what you want to find (in which scientific or disciplinary area)

Ex: Total quality management OR TQM

It is common for different subject areas to have similar acronyms!
Refine your searches

All databases offer the possibility to refine searches using limits such as:

- Specific subjects
- Publication date
- Author
- Type of publication: case studies, reports, peer-reviewed, conference proceedings, etc.
- Journal Title
- And others …
Stop Words

Words that are considered irrelevant to the search!

<table>
<thead>
<tr>
<th>Stop Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A</td>
</tr>
<tr>
<td>• An</td>
</tr>
<tr>
<td>• Are</td>
</tr>
<tr>
<td>• Be</td>
</tr>
<tr>
<td>• In</td>
</tr>
<tr>
<td>• If</td>
</tr>
<tr>
<td>• Into</td>
</tr>
<tr>
<td>• Of</td>
</tr>
<tr>
<td>• On</td>
</tr>
<tr>
<td>• The</td>
</tr>
<tr>
<td>• Witch</td>
</tr>
</tbody>
</table>

In Portuguese, they are equivalent to definite and indefinite articles, demonstrative and possessive pronouns, ....
Pearling or pearl growing

**Pearling** is the act of analyzing the bibliography or reference list (sources cited) of articles that we retrieved from our searches and found relevant to our study.

It aims to identify more articles relevant to our study that were not found in our searches. It also refers to the analysis of the articles that cited those we identified as important for our study. (cited by).
Snowballing: (amostragem em bola de neve)
It refers to the use of the citations of an article that you found interesting to retrieve others that might be relevant to your research.
Google works differently

Google analyzes:

➢ the full content of a page
➢ the exact location of each word
➢ the content of neighboring pages
➢ links between pages

Results may be similar to subscribed databases, but unlike these it doesn’t mean that it’s possible to access the full text of the articles
Google tools and shortcuts

➢ To associate more than one search term, use + or space
➢ To exclude terms to search on, use -
➢ You can also use boolean OR
➢ To search for a concept consisting of multiple terms, type the phrase in quotation marks.
➢ Some words for Google have a special meaning:
  ➢ Website - used to locate a concrete web page
  ➢ Link - show all pages that link to a specific url

➢ With the translation tool, Google allows you to find pages with similar content in English if the search was done in Portuguese and vice versa
“Think full-text”

➢ Be precise about what you are looking for
➢ Don’t use generic terms
➢ Use controlled language
➢ Choose combinations of the terms you are looking for with the + symbol
➢ Use exact quotes in quotation marks
➢ When defining your search delete or add options:

Ex: proliferation - nuclear or bush legacy + environment

You can limit your search to web page titles by placing the phrase “in title” or title + the term
Ex: title hybrid cars
“Think full-text”

When researching any topic that reflects a current social concern, prefer official sites (eg edu “global warming”)
You can also define the file type you want:
**Ex.:** filetype: ppt site: edu “global warming” (in this case only powerpoint files of this subject will appear)

You can also choose advanced search right from the start by creating the limits you consider relevant
EXERCISE
How to create a search expression

“How to write a research paper or report in health sciences”

- Determine Concepts / Terms to Search
- Build your search expression using boolean operators and truncations
- Define limits to your search
- Test your search equation in a database or search engine...
Good Work!

Maria do Rosário Duarte

mrd@fct.unl.pt