Chapter II - Bibliographic Research

Methodology of searching scientific documents

II.1. Bibliography

Bibliographic research is the first step in any research or scientific writing project. It is the process of carrying out a relevant research of useful information, about a scientific subject, that are available in the various sources of literature. In fact, this activity begins even before the conception of the project or the writing, and continues throughout all the stages of their realization.

Bibliographic research mainly aims to establish a "literature review", to have clear insight on the state of the art on what has been done or written on a given subject. Moreover, it also allows the analysis of the results obtained by others and their comparison with own results, to have an idea on the techniques and methods used to deal with a given question, and to have the conclusions of other researchers about the treated topic. Consequently, it must be undertaken in a methodical manner to be effective, and with a good knowledge of bibliographic research tools.

II.1.1. Bibliography¹ and Literature

Bibliography refers to the set of information sources such as books, journals, articles, etc., published on a subject, presented in different conventional forms. From this definition, the bibliography is synonymous with what is commonly called "Scientific Literature", often abbreviated in "Literature".

Scientific literature comprises scholarly and academic publications that report original empirical and theoretical work in the natural and social sciences, and within an academic field. Academic publishing is the process of contributing the results of one's research into the literature, which often requires a peer-review process to be accepted and published.

II.1.1.1 Types of Bibliography (Literature)

Bibliographic sources can have several forms, the most known are:

<u>- Articles (scientific papers, scholarly articles):</u> they are scientific writings that are published in scientific journals. Different types of articles can be defined: Original articles (research articles) which are considered as the highest level of scientific writing, Review articles (synthesis articles), Short communication, Notes, Scientific letters, etc.

<u>- Books:</u> (sometimes called monographs) they are scientific publications that differ from others in volume, content and structure. They present a synthesis of knowledge in a given topic, field or discipline, and they can be of different forms (Books, textbooks, Manuals, etc.)

There are popular scientific works geared towards a wide audience (books for the general public), books for students (teaching materials) and others for researchers and specialists.

- <u>- Conference proceedings, abstracts and reports:</u> they present interventions as well as discussions in a conference between scientists whether it is a congress, seminar or others. A conference article is similar to a scholarly article insofar as it is academic.
- <u>- Theses and dissertations:</u> they are scientific reports that are prepared in the end of academic cycles (graduation and post-graduation). Theses and Dissertations contain extensive bibliographies and can present original research; they provide inspiration for the formatting and presentation of ideas, graphs, charts, and other components of a document.
- <u>- Grey literature:</u> The term grey literature refers to research that is either unpublished or has been published in non-commercial form. Examples of grey literature include: Researchers notes or communications, research reports, government reports, etc.



Figure 1. Example of some bibliographic sources. From left to right: Book, Manual, Original (research) article and Review article.

Scientific literature is also classified on the basis of the originality of the approached topic or question, we note:

- <u>Primary literature:</u> Results from research conducted by an individual scientist or collaboration by a group of other scientists. Most primary literature is published in scientific journals. Original scientific research published for the first time, patents and technical reports, for minor research results and engineering and design work; are all considered primary literature.

- <u>Secondary literature</u>: is a type of literature that relies on primary sources of information (primary literature). Its objective is to summarize and synthesize information in a specific area. These publications include reviews, books (whether monographs, textbooks or handbooks) and manuals. They are particularly useful for having a general overview over a topic.

-<u>Tertiary literature:</u> emerges from primary and secondary scientific literature and aims for a more lay-audience or researchers in completely different fields of knowledge. These publications are written in a more simplified and popular-language format versus a scientific style. Tertiary sources might include encyclopedias and similar works intended for broad public consumption.

II.1.2. Bibliography², References and Citation

The bibliography (or List of Bibliography) is also defined as the list of sources (books, articles, scientific notes, scientific reports, governmental reports, etc.) consulted for the development of a work (thesis, internship report, article, synthesis, etc.) that they may or may not be cited in the body of the work. By this definition, the bibliography is in some way synonymous with "References". However, Bibliography is larger in meaning than References (or List of References) which is defined as the list of sources (books, articles, scientific notes, scientific reports, governmental reports, etc.) used for the development of a scientific work, and which should be cited in the body of the work.

The sources used to elaborate a scientific work (writing) are first "cited" every time that they are used in the body of the work, then they are all listed generally in the end of the work (writing) as Liste of Bibliography or List of References. Citation and References are made up of bibliographical referencing which is a set of elements that describe a given source (author, title, publisher, year, pages, *URL*, etc.) and which allow it to be identified.

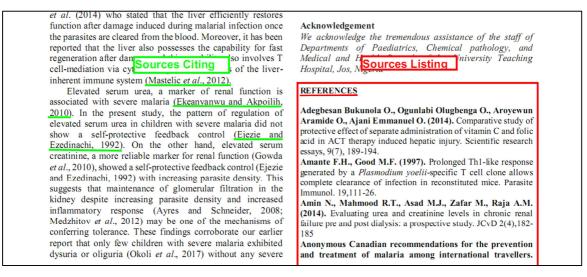


Figure 2. Citing and listing sources in scientific writing (image taken from Okoli et al., 2019).

II.1.3. Importance of Bibliography

- The bibliography is a key tool for work and reflection in scientific research. It presents an essential support for the construction of scientific work or writing projects.
- It allows collecting information around a theme to be developed, and to have a background on what to do and what remains to be done in this theme.
- It allows synthesizing knowledge on a given scientific subject, which facilitates the understanding of scientific themes and allows their progression.
- -It allows having an insight on the techniques and the strategies used to treat the different scientific questions, as it allows the comparison between the different results obtained by different researchers about a given question.
- The bibliography allows also giving more credibility and value to the work, and to show respect for the researchers in the different field of science.

II.2. Bibliography research

Bibliography research is a systematic approach, which consists of identifying, retrieving and processing various elements (figures, bibliography, texts, etc.) on a given subject. This identification of information is an essential step in any synthesis of knowledge and review of the literature in the biomedical field, as in other fields. This approach must be as relevant as possible and aim to be exhaustive. It therefore requires:

- Perfect knowledge of multiple sources of information;
- Perfect use of research tools and strategies.

II.2.1. Resources of Bibliographic documents

Bibliographic documents can be searched in different resources, which can be grouped in two categories:

I.2.1.1. Academic Libraries

Classic resources of documents found in universities, institutions, colleges and research centers; they contain books, Manuals, journals and reviews in their original format (paper format). However, these documents are not accessible all the time; their accessibility depends on the availability of the document and the opening hours of the library. Bibliographic documents of libraries are generally listed in Catalogs that can be consulted in library and sometimes on the website of the institution, in order to

facilitate the search. Catalogs allow identifying the document, its biographic parameters, writing its reference correctly and locating it.

II.2.1.2. Internet resources

Internet resources of bibliography are multiple and of various types. These resources may be, data banks or research engines, specialized or general, open access or with fees. In contrast to library resources, internet resources are digitized (in electronic format), and are available on the internet all the time which makes them easier to consult at any time.

We can distinguish the following types:

a. <u>Bibliographic databases</u>

A database (or data banks) is a Data set relating to a defined area of knowledge and organized to be offered to users.

A bibliographic database is a database of bibliographic records, an organized digital collection of references to published literature, including journal and newspaper articles, conference proceedings, reports, government and legal publications, patents, books, etc. In contrast to library catalogue entries, a large proportion of the bibliographic records in bibliographic databases describe articles, conference papers, etc., rather than complete monographs, and they generally contain very rich subject descriptions in the form of keywords, subject classification terms, or abstracts.

A bibliographic database may be general in scope or cover a specific academic discipline like Biomedical science. The database is:

- produced by an institution, a commercial company or a documentation service;
- stored on a server;
- and displayed to users in response to a request.

Different types of information are taken into account in the databases: bibliographical references, full text, knowledge bases (factual data or digital), chemical structures, images.

Bibliographic Databases offer:

- *Abstracts*: almost bibliographic databases offers abstracts of journal articles plus the citation information (e.g. author names and affiliations, the journal title, volume, and page numbers).
 - Full-text articles: bibliographic databases can also offer free access to some complete articles.
- *Books and text-rich documents:* some bibliographic databases can occasionally offer online publication of textbooks, as well as new, information-rich documents.

Resource	Produced by	Examples of access	Free access*	URLs
PubMed/MEDLINE	The National Library of	PubMed	Yes	http://www.pubmed.gov
	Medicine (NLM)	BioMedNet	Yes	http://research.bmn.com/medline
		Ovid	No	http://www.ovid.com/
		BIDS	Yes	http://www.bids.ac.uk/
ISI Citation Database	Institute for Scientific	Web of Science	No	http://www.isinet.com/isi/journals/
(Web of Science)	Information (ISI)			
Current Contents®	Institute for Scientific	Current Contents Connect	No	http://www.isinet.com/isi/journals/
	Information (ISI)	Ovid	No	http://www.ovid.com/
BIOSIS Previews®	BIOSIS	BIOSIS	No	http://www.biosis.org/
(comprising biological abstracts		Ovid	No	http://www.ovid.com/
and biological abstracts/RMM®				
Pascal	Institut de l'Information Scientifique et Technique	BIDS	Yes	http://www.bids.ac.uk/
EMBASE	Elsevier Science	EMBASE.com	No	http://www.embase.com/
		Ovid	No	http://www.ovid.com/
The Cochrane Reviews	The Cochrane Library	The Cochrane Library	Yes	http://www.update-software.com/
(abstracts)				abstracts/crgindex.htm

Table I. Some scientific Databases (Barton and Trawick, 2004)

Example of Databases:

1- PubMed/MedLine (pubmed.ncbi.nlm.nih.gov)

PubMed is one of biggest databases offered for biomedical sciences. It was developed at the National Center for Biotechnology Information (NCBI), within the National Library of Medicine (NLM), USA.

MEDLINE includes over than 32 million abstracts in, and currently covers about 8,500 journals from at least 90 countries, dating back to 1966. MEDLINE abstracts have a controlled vocabulary associated with them known as **Medical Subject Heading (MeSH) terms**. Several terms are assigned to each MEDLINE abstract, and are used for indexing articles to provide a consistent way to retrieve information. As well as enabling bibliography searches, PubMed offers the following additional functions:

- 1- Links to biological sequences information, including data such as GenBank protein and nucleotide sequences, and macromolecular structures.
- 2- Links to the full-text of journal articles (about 8,500 journals are currently linked in this way). Whether the full text can be viewed without purchasing the journal depends on the journal policy.
- 3- Links to 'Related articles'. For each abstract, similar articles in the database have been identified, based on a statistical analysis of words and phrases found in the abstract text. This is an easy way to expand on a PubMed search when a useful abstract has been found.
- 4- Links to resources outside of the NLM. The 'LinkOut' feature allows other providers of information, such as organism-specific databases like FlyBase, to link to related abstracts.
- 5- Links to textbooks, linking the content of textbooks to PubMed abstracts to serve as background information.

^{*}In cases where access to the database is not free, consult your library for subscription information.

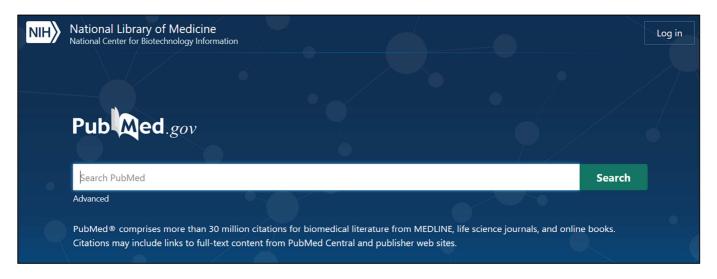


Figure 3. Homepage screenshots of PubMed.

PubMed is a free search engine accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics. Actually, PubMed is the main bibliographic data search engine for all areas of specialization in biology and medicine.

User's Guide is available on the link: https://pubmed.ncbi.nlm.nih.gov/help/, to help researcher and students for a good and effective use of PubMed/Medline.

2- Cochrane Library (www.cochrane.org)

Cochrane Library is a collection of 6 databases containing evidence-based data on the effects of health care.

The aim is to help healthcare professionals make good decisions when their clinical interventions:

- By collecting all the information on the effectiveness of therapies in the different medical fields.
- By criticizing, sorting and synthesizing this information.
- By establishing a database of the results of these syntheses.
- By disseminating its summaries and their messages to prescribers.

The Cochrane Library is a collection of databases that contain different types of high-quality, independent evidence to inform healthcare decision-making. The Cochrane Library includes the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Central Register of Controlled Trials (CENTRAL), and Cochrane Clinical Answers, and a federated search feature that incorporates results from external databases.

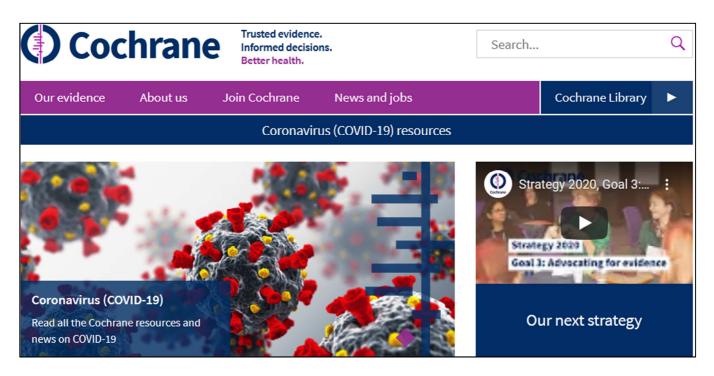


Figure 4. Homepage screenshots of Cochrane Library.

b. Journals websites and Journals packages

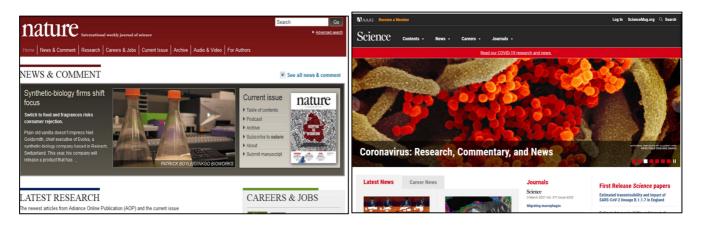
Currently, most *scientific journals* (reviews) have their own websites, on which researcher can find all the published volumes and issues of the journal with the possibility of carrying out searches on the site for a given article or for a given issue of the journal.

Examples: Nature review (www.nature.com), Science journal (www.sciencemag.org)

Journals packages are a collection of electronic journals posted online, whether they already exist in printed form or not. They can be of the same editor or not, private or public, with free or paid access.

Example: ScienceDirect wich is an online service from the publisher of scientific, technical and medical journals and books, *Elsevier-Masson*. It offers more than 3,200 French and English journal titles.

Scince Direct is accessible on the link: www.sciencedirect.com, it offers multiple research options (Keywords, Authors, Titles, Books, etc.) and advanced research option. However, most of its services are with charges (must pay).



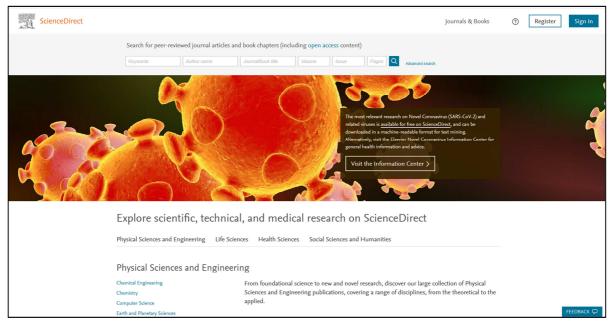


Figure 5. Homepage screenshots of two famous journals, Nature and Science (top); and (bottom) Homepage screenshot of ScienceDirect website.

c. Specialized Search Engines

Search Engines are research tools which make it possible to carry out research on scientific articles approved or not by peer-reviewed committees, university-type theses, citations or even scientific books.

Specialized Search Engines indexes contain most of the online journals submitted to review committees, journals from major publishers of scientific literature. However, the quality of search results does not match that of conventional databases. Although, attention has been drawn to the fact that Research Engines are not completely a suitable tool for bibliometric applications, or for evaluating the impact of researchers and journals.

Examples: Google Scholar: <u>scholar.google.com</u>

Scirus: www.scirus.com

II.2.2. Keywords and research equations

The effectiveness of a bibliographic search is based on three essential elements:

- The good understanding of the treated subject or question.
- The good use of bibliography resources, in particular bibliographic databases and search engines.
- The relevant choice of words and phrases used for the search, called the keywords.

II.2.2.1. Keywords

Keywords are very important relevant words or expressions that describe and represent a subject, a field and in particular the content of an article. They have an important role and place in the bibliography, and are essential for documentary research.

Most online science databases (PubMed, Web of Science, Scopus, OpenEdition, etc.) use keywords to index articles by topic. Some journals even ask for a list of keywords in English or other languages to promote international Search Engines Optimization (SEO), to facilitate access to the bibliographic resources.

Currently, keywords are organized in Thesauruses (A thesaurus is a list of synonyms and can help to identify alternative keywords) in order to facilitate bibliographic research; for example when a searcher use one keyword like "apoptosis" he will get results for all articles containing apoptosis or one of its synonyms such as apoptosis, apoptoses, programmed cell death..., or related words. One of the important thesauruses in biomedical research is MeSH (Medical Subject Headings) which is the National Library of Medicine's controlled vocabulary thesaurus, used for indexing articles for the PubMed/Medline database.

For documentary research, researcher can use a single key word (e.g. *apoptosis*); however, to improve bibliographic search, it is preferable to use several keywords (e.g. *apoptosis*, *caspases*, and *cell death*), or even part of whole sentence sentences (e.g. *neutrophil apoptosis*, or *caspases in neutrophil apoptosis* ...). Moreover, in addition to being scientific or technical words, the keywords can also be names of authors (e.g. *J Watson*), journals (e.g. *J hepatology*), or a combination of all these keywords (e.g. *mTOR* [MeSH] *J Brown* [Author] *Blood* [Journal]).

II.2.2.1. Research equations

Bibliography search equations are combinations of keywords or search phrases, formulated using the use of Boolean operators. Boolean operators are used to link ideas, concepts, keywords used during the search in order to find the most relevant results as possible.

There are three Boolean operators; AND, OR, NOT:

- AND: narrows a search by finding records that contain all the terms entered. Example: *cloning* AND *humans* AND *ethics*: only references containing all these words together will be found.
- OR: broadens a search by finding records that contain either or all of the terms that have been entered. Example: *Neutrophil degranulation* OR *granules exocytosis*: all documents containing one of the two keywords or both will be found.
- NOT: narrows a search by finding records that contain one term but not another. Example: *hepatic diseases* NOT *carcinoma*: in this case, all documents containing *hepatic diseases* will be found except those treating of *carcinoma*.

I.2.3. Bibliographic research methodology

I.2.3.1. Bibliographic research process

Bibliographic research consists of several steps which allow an efficient literature references finding, this process can be summarized in following steps:

1. Define the research subject

This step consists in defining the subject, based on an initial theme, it is a question of precisely defining the problem, and this will allow the exploration of all the dimensions to be treated in order to complete the answer the question. Thus, when information is needed, asking certain questions can help to clarify precisely what kind of information to look for.

2. Define the appropriate keywords:

This step is about defining the terms to be used for the query of different resources, based on the defined topic of in first step. It's necessary to identify also synonyms, more generic or conversely more specific terms.

3. Define the type of information to find: this step aims to determine the nature of the information required for the development of the project undertaken (research, writing). Determining the type of information required allows identification of the nature of queried bibliographies; research articles, review articles, methodological articles, technical manuals, books, etc.

The type of bibliography searched can be specified on databases and search engines using the "Advanced Search" bar. The advanced search parameters allow researcher to specify the search criteria making it more relevant and more efficient, and prevent unsuccessful searches. Researcher can therefore carry out a search according to the nature of the publication (research articles, review articles, book, etc.),

specify the domain for common keywords, carry out a search according to the authors or the years, etc. This greatly facilitates bibliographic research and saves a lot of time.

4. Conducting the search and

This step consists of interrogation bibliographies resources, databases and search engines. Databases and Search engines are multiple and diverse, consulting the right database is crucial for effective research. This is very important since the databases differ in their specifications and do not contain the same types of references. In the case of health science, the choice of which bibliographic database to use is not always easy, as some of the material may be in other bibliographic databases that focus on different disciplines. This means that restricting the search to a certain bibliographic database may result in omission of important references. Therefore consulting different databases would be of great interest to increase the chances of finding most of the published references about a question or an area.

5. Evaluate the results

This step consists of an evaluation of the quality of the research carried out and determining the relevance and reliability of the found references. Researcher must evaluate the relevance of the results to ensure that found references correspond to the topic and the needs previously defined. The researcher must also evaluate the quality and reliability of the results to ensure the scientific value and quality of the content. This questioning is particularly necessary for documents found on the Internet, especially in the era of the explosion of info-obesity (abundance of every kind of information). Researcher must pay special attention to what he can find as bibliography. Finally, researcher could not collect more documents than he can use, especially for subject widely covered by the literature, he must be realistic in assessing the workload.

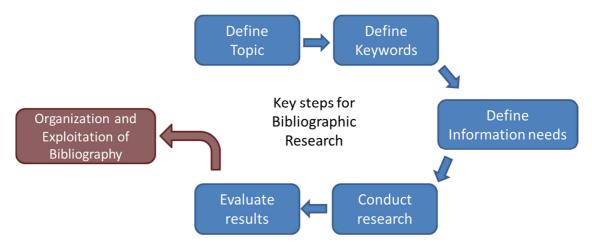


Figure 6. Key steps for Bibliographic Research

Once the literature search is complete, the next step is to organize, analyze and use the collected literature.

a. Organization of bibliography

At this stage, it is important to first organize the bibliography, which sometimes gives to collect a significant number of documents. These documents could be stored in files and organized in folders with meaningful names. Thus, it would be useful to group the references according to their type, for example: practical references, review articles, books, technical manuals, etc. in a more practical way, the references can be grouped according to their relevance use for a given part of the work, example: Introduction, bibliographical review, Material and methods, chapter I ..., discussion, experimental protocols, etc. It is also of interest to compile these references in a file (Excel) in which can be notes the parameters of the document (title, source, date, name of the downloaded file, interest, etc.).

a. Annotation of bibliography

In addition to their organization and modification, bibliographic references can be the subject of a sort of reading sheet (sheet per document) this process is called annotation of the bibliography. An annotation is a summary and/or evaluation. Therefore, an annotated bibliography includes a summary and/or evaluation of each of the sources. Depending on the project or the assignment, annotations may do one or more of the following: referencing the bibliography, summarizing the essential of the source, a short evaluation of the source and a reflection on use of the reference. Annotation notes for each document the complete reference, a summary, the ideas main, quotes, and comments. This process helps to save time and facilitate any subsequent using if sources, but also in formatting the bibliography.

