

How to search for information

In this module we will focus on the ways to find required documents. Adopting basic search habits will help you search more effectively - you will find more good-quality information within a shorter space of time.

Module objective

- to acquaint you with the process of searching for information and its stages
- to connect topic definition while working on a paper with search request formulation and analysis
- to introduce the use of search tools (operators, wildcards etc.)
- to acquaint you with search strategies and tactics
- to teach you to work with your query while searching for information
- to present the most frequent searching mistakes

Keywords

search for information, search request, search query, Boolean operators, proximity operators, wildcards, search strategies (the strategy of building blocks, the strategy of growing pearls, the strategy of chipping)

1 MODULE INTRODUCTION

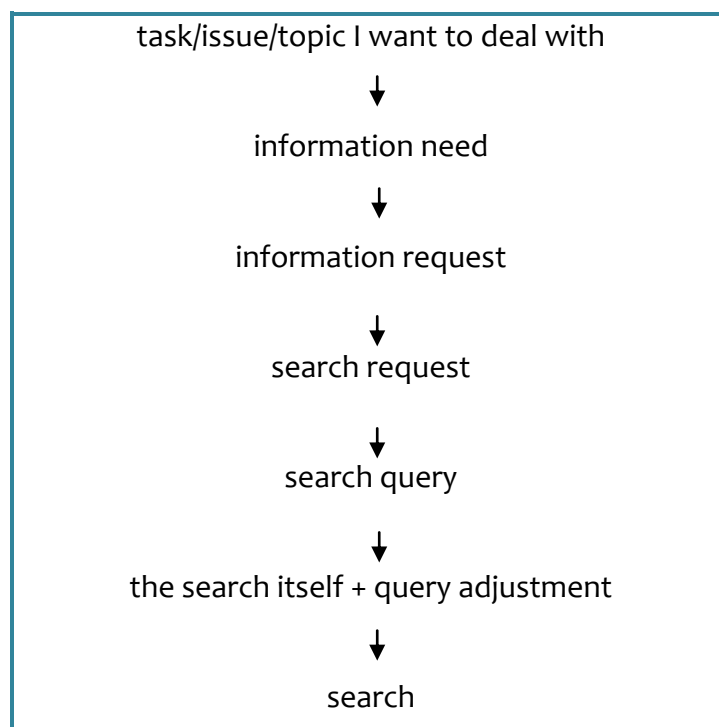
Now you will get acquainted with the way to find required documents. We will acquaint you with individual stages of searching for information, their content, strategies and tactics you may use in practice. In order for your search to be successful, you have to know **WHAT**, **WHERE** and **HOW** to search.

2 SEARCH FOR INFORMATION

If you have a **task/issue** ahead of you, you **need** adequate **information** to solve it. You are therefore in the so-called **information need**. If you express this subjective need, you may speak of a particular “**information request**”. As soon as this request becomes the subject of your search, it is called a **search request**.

If you subsequently express it in a relevant query language of a particular system (database, index, search machine), it becomes a **search query**. **In most cases you proceed with the search for information or documents yourselves**. The result of the search activity is a **search**, which is a list of bibliographic entries, factual entries or full texts of documents corresponding with the given search request.

We will summarize the search process in the following table:



Search for information is therefore an activity **the objective of which is to identify relevant documents or information** in information sources.

In order to be satisfied with the search result, it is advisable to **think and proceed in the following steps/stages**:

- to clarify *what we need* to find, realize a possible *connection* with other topics etc., that is *formulate and analyse our search request*,
- to select (a) *suitable source(s)* of information and *search tool(s)*,
- to search effectively - *formulate our search query* well,
- to assess the *relevance* of found results,
- to *adjust, specify our query* if need be (the so-called *search query adjustment*),
- to *select the most suitable documents* from the results.

3 SEARCH REQUEST FORMULATION AND ANALYSIS

This stage mainly comprises an understanding of the **search subject** (WHAT I want to find) and definition of **key concepts** and their mutual **relations**.

3.1 IDENTIFICATION OF KEY CONCEPTS

- **nouns** are the basis of a successful search;
- omit the so-called **stop words** (e.g. prepositions, conjunctions and other function words); omit **adjectives, pronouns, adverbs** (the so-called modifiers), **verbs**; modifiers serving to identify a word to avoid ambiguity are an exception (e.g. semantic web, machine translation, Meissen china, Pythagorean theorem...);
- write down **synonyms** and concepts related to your key concepts.

Practical example

Based on the above recommendations, words which have been crossed out below will not be included in the keyword set. Does information training of university students ~~affect their ability to work with~~ information sources?

In addition to topic definition by means of keywords and synonym identification, think about the following aspects:

- what **kinds of documents** you will want to search for (monographs, scientific articles, reviews, diploma theses...),
- in what **language**,
- from what **time period**,
- whether you will only search for documents related to a particular **geographic area** etc.

4 SELECTION OF SUITABLE SOURCES AND TOOLS FOR INFORMATION SEARCH

As soon as you think of a topic and determine other delimiting aspects, you will choose a suitable information source or search tool in which you will carry on your search.

You should take the following into consideration:

- **incorporating** your topic in a particular field - is there a **suitable field source** (e.g. a specialized database)?;
- **types of documents** which you want to use;
- the selection will also be affected by your **knowledge** of available scientific sources you could use and the fact whether you can get them or not;
- if you need **scientific** information or if **general/popular** information is enough etc.

*Each time you have to choose your sources according to the **objective of your search**.*

5 SEARCH QUERY FORMULATION

While formulating your own search query, you will further work with **key concepts/search concepts**. Via **search tools** you will **connect** and combine them, and work with their various **forms**.

Your search query may have more forms: it may be a single query or the so-called multiple query.

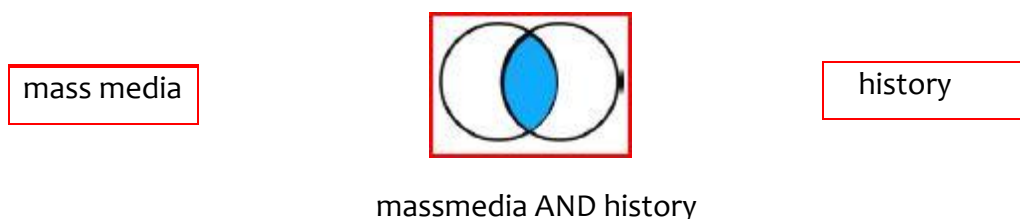
5.1 SEARCH TOOLS

5.1.1 BOOLEAN (BOOLE'S) OPERATORS

Logical relations between search concepts may be expressed by means of the so-called Boolean operators. Among the most frequent operators are **AND**, **OR** and **NOT**.

Operator AND - a logical product, an intersection of two sets of concepts

AND between two terms means that only those documents/entries will be found which contain **both**, or **all terms** connected with this operator. This operator *narrows* the search result. It serves to *connect concepts with different meanings*. Example:

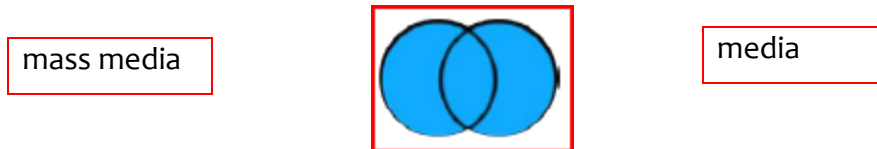


*The found set of entries/documents must contain both terms - mass media **as well as** history (it is an intersection of two sets highlighted in blue).*

Operator OR - a logical sum total of two sets of concepts

Those documents/entries are found which contain **at least one** of the terms connected with OR. This operator *expands* the query. It serves to *connect synonyms and related concepts*.

Example:

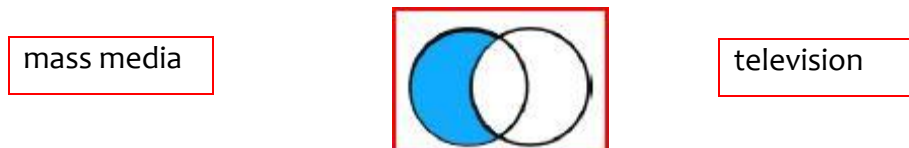


mass media OR media

The found set of entries/documents contains at least one of the terms or both terms (it is a sum of both sets highlighted in blue).

Operator NOT - logical negation

Those entries/documents containing the concept **preceded by NOT** will be **excluded**. This operator *narrows* the search result.



mass media NOT television

The found set of entries/documents does not contain the term television; however, it must contain the term mass media (it is a difference between both sets highlighted in blue).

Practical example

The following sets of identified keywords and related concepts with the following operators may be determined in relation to the topic Information training of university students:

information training **OR** information education **OR** information literacy

AND

university student **OR** undergraduate student **OR** student **OR** college student

5. 1. 3 TOOLS FOR VARIOUS WORD FORMS EXPRESSION

Truncation

Truncation enables you to search according to **word parts** or only **word roots**. You may replace 1 or more characters at the beginning or at the end (prefixes, suffixes) with an **asterisk *** or a **question mark ?**. If you enter e.g. *student?* or *student** in the Aleph system, the results will comprise entries containing *students* etc. (it is an alternative for the operator OR).

5. 2 BASIC SEARCH STRATEGIES

In case one keyword or phrase is not enough and we need a more complicated query to get a better result, it is advisable to select one of the basic search strategies (methods) which will help you create your search query.

Among the most frequent are:

5. 2. 1 THE STRATEGY OF BUILDING BLOCKS

A query will consist of more individual **subqueries** (building blocks). We will divide the original query into several subqueries which will express the **central concepts**. Then we will connect the subqueries via the operator **AND** get the final query.

Practical example

Topic: **information training** of **university students**
first central concept (subquery 1), second central concept (subquery 2)

In order to formulate subqueries, we will use the set of related concepts defined in the previous steps. We may get the following subqueries.

Subquery 1:

“information training” OR “information education” OR “information literacy”

Subquery 2:

“university student” OR “undergraduate student” OR “student”

We will connect the subqueries via the operator AND into the final query.

(“information training” OR “information education” OR “information literacy”)

AND (“university student” OR “undergraduate student” OR “student”)

The brackets separate the individual subqueries.

Hint! In the case of multiple queries in which we use both AND (product) and OR (sum total), the operator AND takes precedence over the operator OR. If we need to connect concepts in logical groups and ensure/change the order of evaluation, we have to use brackets. Try to enter e.g. *shakespeare AND (julie OR henry OR dream)* and *shakespeare AND julie OR henry OR dream* in the MU union catalogue. Compare the sets of found entries.

5. 2. 2 THE STRATEGY OF GROWING PEARLS

With this method we start from

- **the narrowest concept** and proceed to individual concepts according
- to the most specific terms.

The objective is to find **at least one relevant entry**.

Then we use **descriptive data** from the database/index, e.g. keywords, subject word entries, authors (the so-called **selective data**), for the subsequent adjustment of our query according to found terms.

5. 2. 3 THE STRATEGY OF CHIPPING

It is the opposite of the strategy of growing pearls. The **query** is **formulated widely** on purpose, and **gradually narrowed**. The objective is to find more entries/documents.

The set of found concepts is larger than we need, therefore we “chip it off” and narrow it by means of **narrowing tactics**:

- limitation to a particular document type,
- NOT in order to exclude some entries,
- selection of entries in which keywords occur in the “title” field,
- language limitation,
- a particular time period,
- combination of sets of data expressing another query concept,
- a particular point of view (operator AND), etc.

Practical example You are interested in library publications dealing with **nation and ethnic minorities**. You also know that the location number of the books on this topic starts with **323.1** in the Central Library of the Faculty of Arts of MU. You will start your search with the number 323.1. You will use the function **indexes** and check the field **Universal decimal classification**. In the next step you will narrow your search, e.g. formal or thematic limitation.

5. 3. 1 SEARCH VIA KEYWORDS

Search via words or phrases is the most frequent way of searching in library catalogues and databases. It enables you to search according to words entered **in all fields** or **in several different fields at once** (e.g. in the title + subject word entries + annotation), or **in one selected field** (e.g. words from the title, words from the author data, words from subject word entries etc.).

Practical example You want to find documents containing the terms mass media or media, either in the title or in other entry fields. Check “All fields” in the MU catalogue basic search, and enter *mass media* and *media* in the search line, connected with OR.

6 REDUCING AND ADJUSTMENT OF THE SEARCH QUERY

In case you have found too **many entries**, it is advisable to use one of the **narrowing tactics** and reduce the set of found entries:

- combine keywords with controlled terms (subject word entries...),
- limit your search to a particular entry field (title, subject matter, abstract...),
- limit your search to a particular document type (e.g. articles from reviewed magazines...),
- use another limitation - time, language, geographic,
- use proximity operators (the proximity of words),
- use the operator NOT to exclude some entries,
- use even subordinate terms for your search (narrower concepts),
- use other options which databases offer (some databases have e.g. a function called “limiters”).

In case you have found too **few entries**, try to apply one or more of the following **expanding tactics**:

- add synonyms, various word forms (operator OR, wildcards, truncation according to word roots),
- use controlled terms such as keywords (e.g. for the search in all fields),
- add broader controlled terms (i.e. those which are superior to the used terms (descriptors, subject word entries); you will find them in the relevant controlled database dictionaries, catalogues...),
- use general terms, i.e. those with a frequent occurrence,
- cancel all preliminary limitations.

7 MODULE SUMMARY

In this module you have got acquainted with the basic principles of information search. In the following modules we will deal with particular search sources (electronic indexes, databases, search machines, subject directories etc.), and the use of search strategies and tactics will thus cross to a practical dimension.

Points to remember:

- first you should clarify what you want to find - the **defined topic**/issue of your paper becomes the **subject of the search** for relevant information
- determine **keywords**, select **synonyms** and **word forms**, define their mutual relations
- clarify the **document types** which you will use (monographs, dictionaries, encyclopaedias, magazines, websites...) and other limitations (language, time...), choose **suitable sources** accordingly (indexes, bibliographic or full-text databases...)
- **specify keywords** via the **thesaurus/controlled dictionary** of the given system (index, database...) - notice broader and narrower terms
- put together a **search query** via the **Boolean operators** (AND, OR, NOT), proximity operators (NEAR, WITHIN, ADJ...) or **wildcards** (?, * etc.) according to the options of the particular search system. (see Help).
- based on the result, **adjust your search query**: **too few found entries** ☞ expand the query (synonyms, OR, superior terms...); **too many found entries** ☞ narrow the query (other terms, AND, subordinate terms, formal limitations - language, time, etc.)