

## 4. Fundamentals of use of information search systems

The specific purpose of each of the information retrieval systems is different, but all of the two have similar basic features:

- Possibility of navigation/search
- Using free/controlled language
- Simple and advanced search
- Using operators

### 4.1. Navigation and search

A system based on "navigation and search" allows you to search, through the system, for information without knowing the terminology used either through alphabetically organized lists of authors, titles and topics, or through systematically available information.

And it also allows you to search through a search expression, either in a single field (author, title, etc.), or in multiple fields at the same time.

### 4.2. Information retrieval techniques

There are several techniques for information retrieval, however, the most common are ,

- Diffuse Logic Recovery Systems: Selecting natural language terms and performing the query using freely chosen terms (free text research)
- Language techniques: selecting terms using a controlled vocabulary (thesaurus or other type of list of controlled terms) to avoid problems that cause variants to exist for the same term.

In library catalogs and databases, it is common to use standardized terminology (documentary languages) that serves users to understand the terms they can use to conduct their research: classifications, topic titles, indexes, and thesaurus.

- Term weighting techniques: Weighting provides ofrece an indicative value appropriate to search criteria (Boolean concepts and operators), depending on the user's interests, therefore, information retrieval depends on the assigned value. The most relevant search document would be one that has represented all search terms and plus the one with the most repeated value more times, regardless of where you are in the document.

### 4.3. Simple search and advanced search

Automated recovery systems that use weighting techniques typically allow for two types of searches:

- Simple Search: Search either in a single field (title, author, subject, etc.), or in all fields simultaneously. Multiple terms can be combined within the search field using Boolean operators.
- Advanced Search: Allows you to combine multiple simple searches by using Boolean operators

### 4.4. What are operators and types?

Operators are symbols that are used in the elaboration of a search equation and allow you to combine different terms with each other and establish logical relationships between terms.

There are several types of operators:

- Logical or Boolean operators
- Synthetic or proximity operators
- Truncation and masks or wildcards
- Other operators

## 4.5. Logical or Boolean operators

### 4.5.1. Intersection: AND

Select only documents that contain both terms at once.

Indicates that the words that precede and follow the operator must be in the search result. Example: "teens" AND "depression" will show results containing the words "teens" and "depression".

We want to locate those works in which "advertising" and "women" appear..

The search equation could be: "advertising" AND "women"

### 4.5.2. Denial: NO, NOT

This operator selects documents that contain the first term, but not the second one.

It indicates that the previous keyword in the operator should appear, but not the subsequent keyword. Example: NO "teen" "depression" will show the results in which the word "depression" appears and that does not appear "teen" in them.

We want to locate documents that contain the term advertising and do not contain the term women. The search equation we would perform would be: NO woman advertising

### 4.5.3. Logical OR, OR

It selects documents that contain either or both terms separately or both at the same time. When we do not indicate any operators between words, the systems interpret it as if we had placed OR.

Indicates that it arrives with that only one of the words is present. In most search tools you can replace it with white space. Example: Young depression gold will give results that contain any of the words.

### 4.5.4. XOR

Specifies that, of both keywords, only one

We want to locate articles that contain the term advertising or the term women or the term marketing. The search equation we would perform would be: advertised OR women OR marketing

## 4.6. Operators without proximity

Proximity or position operators (NEAR, ADJ, SAME, WITH) measure the distances between the terms you are trying to locate. They can be used to connect words or phrases within a search field, but not between search fields.

- NEAR: Used to specify searches because they allow you to indicate how close they should be to other terms; the closer they are, the more feared they are. Example: NEAR global climate change.
- ADJ: Bring the two words, one followed by the other without other characters between them. You can type ADJ or use a hyphen between words without leaving room. Example: If

you search for "Chicago ADJ History", only records containing "Chicago" and "history" will be retrieved together in the same field and with "Chicago" retrieved first.

- SAME: Used to locate records where the bibliographic record field contains all the specified terms. All search terms are in the same field, but not necessarily in the same sentence. Example: If you search for "Chicago SAME history", only records that contain "Chicago" and "history" within the same field will be retrieved.
- Operator 'WITH': Displays the results containing the key terms entered in the same field (title, summary, content...) and in the same sentence.

These operators are interleaved between descriptors to establish a proximity relationship between them, that is, they allow you to specify the position of two or more words in the search equation and therefore in the record. Each database has its own proximity operators, there are small differences between them, so it is always advisable to consult the help they offer.

#### 4.6.1. Adjacency: ADJ

This operator allows you to retrieve records in which the terms listed in the question appear side by side within the document, without the words detaching them.

We want to locate a document that shows that "total quality is starting to be implemented in service companies"

The search equation we would perform would be: total quality of ADJ

We would not recover documents that read: "Quality has replaced the total in recent years"

#### 4.6.2. Presence of terms at a maximum distance of words "n": W

This operator allows you to retrieve records in which the question terms appear in the document separated by a number of words equal to or less than the indicated one.

In the search equation: total quality of 6w

We would recover documents in which it was said: Total quality has come a long way in recent years; Quality has advanced more than total quality

We would not recover: Quality has advanced less in the last decade than total quality

#### 4.6.3. Presence of two terms located at a certain distance "in words": P

This operator allows you to retrieve documents in which the two terms are separated by a fixed number of words between them.

2P company vision

We would get documents declaring "vision of the company".

We wouldn't get documents stating, "Coca Cola's vision."

#### 4.6.4. Presence within a sentence: NEAR, S

This operator is used to locate documents that contain both terms in the same sentence.

We want to locate those articles that contain the same quality and excellence of the phrase.

The research we would carry out would be: QUALITY CLOSE to excellence

This operator allows you to specify the number of words that can exist between the two terms.

Use NEAR/x to search for records in which operator-joined terms appear with a distance between them of a specific number of words.

Replace the x with a number to specify the maximum number of words that separate the terms.

If you use NEAR without / x, the system will find records in which the terms joined by NEAR appear with a distance of 15 words between them. For example, these searches are equivalent:

near the salmon virus

near salmon viruses / 15

We want to locate articles that in the same sentence record quality and excellence, maximum separated by two words.

The research we would do would be: QUALITY of NEAR2 excellence

4.6.5. Two terms in the same field regardless of the order in which they are located.

The WITH operator, W allows you to retrieve records in which the two terms are in the same field.

We want to locate the documents in which the total quality appears in the title or in the summary. The research we would carry out would be: total quality W

It will not allow you to recover a document in which it appears in the title: "quality in Spain" and in keywords "total".

Displays the results containing the key terms entered in the same field (title, summary, content ...) and in the same sentence.

Reduce the number of results and nuance your searches.

Example: "Alzheimer's WITH treatment". Displays the results that are included in the same sentence the two specified terms

4.6.6. SAME operator

In certain search services (WOS database for example) this advanced operator is allowed to be used in address searches. Used to restrict search to terms that appear at the same address in a full record. Use parentheses to group the terms of your address.

For example:

(McGill Univ SAME Quebec SAME Canada) searches for records in which McGill University appears in the Directions field of a complete record along with "Quebec" and "Canada".

(Portland SAME Oregon) searches for records showing "Portland", "Oregon", or "OR" (short for the state) in the Directions field of a record.

Note that the SAME operator works exactly like AND when used in other fields (for example, in the Subject and Title fields) and when terms appear in the same record. For example:

(Cat SAME mouse) retrieves the same results as (cat and mouse).

## 4.7. Truncation and masks or *wildcard* characters

Presence operators can override a character or set of characters, located to the right of a term. They are used when it is necessary to use not a simple term, but also its derivatives (with prefixes, suffixes, lexical variants ...).

Truncation or precision operators (\*, ?, \$) They are used to take into account variations in word endings (plurals, common word roots, language variations, etc.). The asterisk (\*) is a sign that replaces a character or group of characters, which are very useful when the full way of typing the term is unknown. The asterisk is used when one or more letters are omitted and the (?) When a single one is omitted, either in the middle or end of the word. Examples: bibliot \* or bibliot? will be retrieved: library, librarian, librarians, librarians, librarians, librarians, librarian.

The most common are as follows.

- Operator \$
- Operator?
- Symbol \*

### 4.7.1. Operator \$

Allows truncation of a multiple number of characters in the middle or end of a search term.

### 4.7.2. Operator ?

Override a character in the middle or end of a search term. Cannot be used at the beginning of a word.

### 4.7.3 Symbol \*

It is used only as a truncation character on the right to find all the shapes of a word.

Example: *econom \** you will find *economy, economy, econometrics, econometrics*

### 4.7.3 Important things to consider with the use of truncation

- It is advisable in long and unequivocal terms because important forms of the same term (gerunds, participle, plural, etc.) are obtained as well as substantive or adjectives of verbs.

- It is not advisable to truncate with less than four letters at the root.

Truncations that result in the selection of a large number of terms run the risk of generating noise and blocking the computer, as it will take a long time to perform all the logical sums.

- It should be used with caution to prevent noise from being introduced with unthinkable terms.

- Allows to increase the completeness of the recovery.

## 4.8. Other operators

### 4.8.1. Parenthesis

They are used when the search statement includes three or more terms to control how it will be searched, because without parentheses the search is performed from left to right while with parentheses the included words are searched first, and the result of this search does the following.

We want to locate those articles that contain, education or teaching and technology or computers or (education or teaching) and (technology or computers) In the first case research recovers everything that has technology in education and education, as well as all about computers, refer or not the articles to education. In the second example we use parentheses to locate only articles about technology or computers in relation to teaching or education.

#### 4.8.2. Quotes

"Quotes" are used to search for an exact phrase in the exact order of the words to be searched.

#### 4.8.3. Relational operators

They are not generally used in general bibliographic research; perhaps useful for very specific areas and situations. However, it is important to keep them in mind in case we ever lose them. Its use is limited to establishing relationships between numeric expressions to find statistical and mathematical data.

- <Smaller than
  - Greater than
- • Equal to
- Different from
- <- Less than or equal to
  - ➔ • Greater than or equal to

#### 4.8.3. Priority of the search operator

If you use different operators in your search, the search will be processed in this order of priority:

- CLOSE / x
- SAME, NEW
- No
- And
- Gold

Join the parentheses to override the priority of the operators. For example:

FLU AND AVIAN OR influence research records that include the word influenza. It will also look for records that include flu and avian.

(Flu Gold Influenza) And avian searches for records that include both flu, avian or flu and avian.

If you write copper or lead and algae, you will get all the records in which the terms "lead and algae" are present, as well as all records in which the word "copper" appears.

If you write (copper or lead) and algae, you will get all the records in which the word "algae" is present along with the words "copper" or "lead".