Department of Library and Information Science

Bharathidasan University

Tiruchirappalli-620024

Name of the Programme: M.Lib.I.Sc

Course - 2.2 Information Processing and Retrieval Systems

Course Code:P21MLS7

- **Unit-IV:** Query formulation - search process; Search Techniques and strategies in Web of Science, Scopus, BLAISE, INSPEC, MEDLINE.

Course Teacher: Dr.C.RANGANATHAN

Professor

e-mail: cranganathan72@gmail.com

Search Strategy Techniques

Search Strategy

- Search strategy is the action plan which is drawn to conduct a search.
- It encompasses several steps and levels of work in information retrieval.
- There are many issues that need to be considered while formulating an appropriate search statement.
- These are:
  - the concepts or facets to be searched and their order;
  - the term(s) that appropriately represent(s) the search concept;
  - the feature(s) of the retrieval system concerned; and
  - the measures to be taken in revising a search statement.

Basic Search Strategy: Steps in-

- Identify the important concepts of your search.
- Choose the keywords that describe these concepts.
- Determine whether there are synonyms, related terms, or other variations of the keywords that should be included.
- Determine which search features may apply, including truncation, proximity operators, Boolean operators, and so forth..

Cont…

- Choose a search engine/Database.
- Read the search instructions on the search engine's home page. Look for sections entitled "Help," "Advanced Search," "Frequently Asked Questions," and so forth.
- Evaluate the results. How many hits were returned? Were the results relevant to your query?

Categories of Search strategy

- **High recall search:** when the user needs to find out all the relevant items on the stated topic. Recall is a parameter used to measure the performance of information retrieval systems; it is measured as the proportion of relevant items retrieved from a collection in a given search session;
- **High precision search:** when the user needs only relevant items, i.e., as small a number of non-relevant items as possible. Precision is a parameter used to measure the performance of information retrieval systems; it is measured as the proportion of the retrieved items that are relevant in a given search session; and
- **Brief search**: when the user wants only a few relevant items as opposed to all the relevant items.

Search Strategies

- Keyword Search
- Boolean Operators
- Truncations
- Wild Cards
- Domain Search
- Phrase Search
- NEAR
- Parentheses
- Limiting
- Synonym Search
- Range Search
- Dictionary Definitions etc.,

Boolean Operator – AND

- AND is the small overlap where both terms occur
- Indicates that only those Web pages that have both words in them will be retrieved or will result in Web pages that have either term

Example:

“Global Warming” AND “Sea level rising”

Boolean Operator – OR

- OR allows more than one term
- OR requires at least one of the terms joined by it to appear somewhere in the document in any order.

Example

"Digital Library" or "Electronic Library"

Boolean Operator – NOT

- Excludes documents containing whatever follows it.

Example:

Global Warming NOT Srilanka

Animals NOT Dogs

Phrase Search/Quotation Search

- Requires words to searched as phrase, in the exact order you type them.

Example:

"Digital Library"

"Global Warming"

NEAR

- Near operator is used when we want to require that certain terms appear in the same sentence or paragraph of the documents.

- Requires the two phrases to occur within 16 words of each other in either direction.

Example:

"Global Warming" NEAR "sea level rise"

Parentheses ( )

- Parentheses also must be used with NEAR

Example:

("Global Warming" NEAR "Sea level rise") AND (California OR "Pacific coast")

Domain Search

- To search only within one specific website by entering the search terms you're looking for, followed by the word "site" and a colon followed by the domain name.

- For example, here's how you'd find admission information on bdu site:

  Admission Site: www.bdu.ac.in

Truncation Searching

- Truncation is a search facility whereby a search can be conducted for all the different forms of a word having the same common root.
- Truncation allows the computer to search for multiple forms of a word.
- EXAMPLE:

  music* will retrieve as:

  musical, musician, musicians, musicality, music's
- bank* will retrieve as :

  bank or banks or banking or banker or bankruptcy, etc.

  Librar* will retrieve as:

  Library, Librarian, Libraries, Librarianship,

Wild Cards

- Some databases allow for wild cards to be embedded within a word to replace a single character. For instance, in InfoTrac, you can also use the question mark (**?**) within a word to replace a character.

For example:

wom?n **will retrieve** woman or women

Limiting

- Limiting by Date, Language, fileformate etc.,

Digital Library + PDF

Semantic Web + PPT

Web Technology + html

Synonym Search

- If you want to search not only for your search term but also for its synonyms, place the tilde sign (~) immediately in front of your search term.

Example:

~ Fast Food

Range Search

- Range search is very useful with numerical information.
- It is important in selecting records within certain data ranges.
- The following options are usually available for range searching, though the exact number of operators, their meaning etc., differ from one search system to another:
  - Greater than (>)
  - Less than (<)
  - Equal to (=)
  - Not equal to (|= or <>)
  - Greater than or equal to (>=)
  - Less than or equal to (<=)

Dictionary Definitions

- To see a definition for a word or phrase, simply type the word "define" then a space, then the word(s) you want defined. To see a list of different definitions from various online sources.
- Example:

Define: Virtual Library

INFORMATION SEARCH PROCESS

Basic Features

- a) matching of information needs with the information available in IR system;
- b) conducting the search with some criteria such as surrogates, exact form, etc.; and
- c) subject surrogates are the main approach of information seekers.

Search Tactics

The success of any search depends upon the tactics effectively adopted in conducting a search.

Search tactics is concerned with adopting techniques in maneuvering the actual search.

The tactics described by Bates can be grouped in four categories [Bates, 1979]:

- Monitoring tactics (to keep the search on a track)
- File structure tactics (for traversing information within the system)
- Search formulation tactics (to aid in the process of designing and redesigning the search formulation)
- Term tactics (to select and reuse search terms)

ONLINE SEARCHING

- Online information retrieval involves searching remotely located databases through interactive communication with the help of computers and communication channels [Chowdhury, 2004].
- The database can be accessed by the user directly or through a vendor (supplier of online services); in each case through the computer and communication network.
- The term online retrieval can thus be used to indicate the information retrieval services available from producers of databases, or vendors of these databases.
- Although online information retrieval systems have existed for more than four decades, emergence of the Internet and world wide web have brought significant changes and improvements in the online information retrieval environment.

Online Search-Tasks

- Decide which particular database(s) is/are to be searched;
- Guess the words that might have been used by the authors and the indexers in a database of potentially relevant documents;
- Use the thesaurus of the chosen database in order to translate the query terms in the appropriate way (is in the language of the system);
- Coordinate the terms (often using Boolean operators) to formulate the search statement;
- Use the search features and search operators appropriate for the chosen database;
- Input the search statement;
- Repeat steps (e) and (f) until a desirable output is obtained or the search fails altogether; and
- Identify the actual relevant items from among those retrieved.

Online Search Services

- There are various components of an online search service such as:
  - **information providers** or database producers who provide databases to be accessed in an online mode;
  - **a search service provider** or vendor, which provides access to the databases and software for conducting the search;
  - **communication links** that connect the user with the host and the database(s); nowadays users can communicate with the service providers through the web and therefore an Internet connection is necessary; and
  - **a local workstation** through which the user is linked to the service.

Examples of online search services:

Dialog (http://www.dialog.com/about/): A pioneer in online search services, Dialog provides online access to over 800 million records in 900 databases in different disciplines.

OCLC FirstSearch (http://www.oclc.org/firstsearch/): This provides library users with instant online access to more than 72 databases, including these valuable OCLC databases: OCLC WorldCat, OCLC FirstSearch Electronic Collections Online, OCLC ArticleFirst, OCLC PAIS International, OCLC PapersFirst, OCLC ProceedingsFirst and OCLC Union Lists of Periodicals.

Ovid (http://www.ovid.com/site/index.jsp): Ovid provides access to hundreds of full text journals, renowned textbooks and premier bibliographic databases in various disciplines.

STN (http://www.cas.org/stnonline.html): STN offers current and archival information from over 200 scientific, technical, business and patent databases covering a broad range of scientific fields, including chemistry, engineering, life sciences, pharmaceutical sciences, biotechnology, regulatory compliance, patents and business.

Example-Full Texts of Journals and Books

- EBSCO Information service (http://www.epnet.com/default.asp): This provides access to a large collection of full text and bibliographic databases suitable for all kinds of libraries.

- Ingenta (http://www.ingenta.com/): Ingenta provides access to the full text of over 5,400 publications from over 230 academic and professional publishers.

- ProQuest (http://proquest.umi.com/pqdweb): ProQuest is a resource of electronic collections containing millions of articles originally published in magazines, newspapers and journals.

Free Web Resources on Internet

- E-Journals
- E-Books
- Electronic Theses and Dissertations
- Bibliographic Databases
- Scholarly Publications
- Video sharing Web Sites etc.,

E-Books

- Guttenberg - http://www.gutenberg.org/wiki/Main_Page
- *O'Reilly's CD bookshelf* - http://docstore.mik.ua/orelly/
- *Google* - http://docstore.mik.ua/orelly/
- *Childrens Books* - http://en.childrenslibrary.org/
- *DLI* - http://en.childrenslibrary.org/

Electronic theses and Dissertations

- Scirus ETD – http://www.ndltd.org/serviceproviders/scirus-etd-search
- Open Thesis - http://www.openthesis.org/
- Australian digital Thesis - http://adt.caul.edu.au/
- Vidyanidhi - http://www.vidyanidhi.org.in/
- INFLIBNET – *indcat.inflibnet.ac.in/indcat/thesis.jsp*
- Indian Institute of Science – http://etd.iisc.ernet.in

E-Journals

- *Indian Academy of Sciences* - http://www.ias.ac.in/
- *NISCAIR Journals* - http://www.niscair.res.in/sciencecommunication/ResearchJournals/rejour/rejour1.htm
- *DOAJ* - *www.doaj.org/*
- *Stanford University Journals* - http://highwire.stanford.edu/lists/freeart.dtl
- Free Medical Journals - http://www.freemedicaljournals.com/
- Infolibrarian
- http://www.infolibrarian.com/eart.htm#eletonic

Scholarly Publications

- ***Indian Academy of Sciences -*** **http://www.ias.ac.in/**
- ***NISCAIR Journals -*** **http://www.niscair.res.in/sciencecommunication/ResearchJournals/rejour/rejour1.htm**
- ***DOAJ -*** ***www.doaj.org/***
- ***Stanford University Journals -*** **http://highwire.stanford.edu/lists/freeart.dtl**
- **Free Medical Journals -** **http://www.freemedicaljournals.com/**
- **http://www.infolibrarian.com/eart.htm#eletonic**

Bibliographic Databases

- Pubmed

  www.ncbi.nlm.nih.gov/pubmed

- LISTA

  http://www.libraryresearch.com/

Subscribed Web Resources

- E-books
- E-Journals
- Bibliographic Databases
- Numeric Databases
- Aggregators Databases (EBSCO)
- Protocols
- Patents and Standards

E-Books

- E-library
- Springer
- Elsevier Science
- EngNetBase
- Oxford University Press
- Sage etc.,

E-Journals

- Springer
- Emerald
- Elsevier Science
- ACS
- Wiley Publications
- JSTOR
- Blackwell etc.,

Bibliographic Databases

- Web of Science
- Scopus
- INSPEC and COMPEDEX
- Mathscinet
- Biosis and
- Scifinder Scholar etc.,

cranganathan72@gmail.com