Subject: Library and Information Science

Production of Courseware
e-Content for Post Graduate Courses

Paper No : 04 Information Sources, Systems and Services

Module   : 19 Literature search

Development Team

Principal Investigator & Subject Coordinator	Dr. Jagdish Arora, Director INFLIBNET Centre, Gandhinagar
Paper Coordinator	Mrs Renu Arora Former Head, Education and Training, NISCAIR
Content Writer	Dr M Natarajan Sr Principal Scientist, NISCAIR
Content Reviewer	Dr R K Verma Former Chief Scientist, NISCAIR

Module 19: Literature Search

I. Objectives

After reading this module, you will be able to:

- Learn to search professional online databases and the Web efficiently and effectively, with emphasis on their use as part of reference service in libraries and information centres;
- Gain familiarity with the characteristics of bibliographic and non-bibliographic databases from a professional searcher's point of view;
- Learn the basics of searching the most widely used professional online information systems in libraries;
- Define Boolean Logic and know about the Boolean operators;
- Understand the search process and design a search query; and
- Raise awareness of the deficiencies in the expensive professional online information systems.

II. Learning Outcome

After reading this module, you will be able to learn the fundamentals of online information systems, the need, and sources of databases. You will be acquainted with the basics of searching the most widely used professional online information systems in libraries. This module will also give you a functional understanding of information searching and retrieval systems. You will learn how to formulate search strategy, planning the search and developing a search strategy. Lastly, this module will discuss the aspects of search techniques for information retrieval.

III. Module Structure

1. Introduction

Literature search is the activity of looking thoroughly literature in order to find suitable information for a specific purpose. In library and information science, searching refers to looking through all kinds of records thoroughly in order to find desired information. In this module you will learn the need and ways of searching organized information.

This module will give you a functional understanding of information searching and retrieval systems, the way they are implemented in a diverse array of Web and professional online databases, and how to search and use those effectively in research and reference work. Research scholars and other students who are the focus have to select an original research topic and conduct rigorous database searches to support their studies. However, they are often unfamiliar with the various types of sources, databases, and search methodology required for such in-depth research.

Literature has shown that users' information searching skills are initially inadequate, even at the research level. It would, therefore, be interesting to find a good way to help users acquire the needed skills set in order to retrieve necessary information. Over the past four decades, numerous studies have examined the differences between experts and novices in different domains, a research area known as expertise or novice-expert research. A good understanding of how people become experts may help novices such as beginners as research scholars shorten their learning curve toward becoming experts in information searching and retrieval. With regard to methods employed by searchers, past research has investigated search tactics; subject searches; keyword searching; Boolean searching; proximity searching; author and title searching; and searching by browsing. With the above in view, this module will also be discussing the aspects of search techniques for information retrieval.

2. Fundamentals of Online Information Systems: Literature Review

Information retrieval is now seen as an interactive or social activity with the various situations and aspects of the user influencing overall system performance. Lancaster has clearly explained the concepts that are often unfamiliar or confusing to readers, including Boolean logic, file structures, evaluation criteria, and vocabulary control. The need for information professionals to know these fundamentals has not changed, nor has continued reliance on these basics in information retrieval systems. While the underlying technology of inverted file structure has improved dramatically to provide efficient retrieval of massive full text databases, the importance was established in early online systems. Although often criticized and now faced with many alternatives, Boolean logic remains the standard for information retrieval systems. The most common criticism of Boolean logic systems throughout the 1980s and 1990s was that end users had trouble understanding Boolean logic and thus query formulation is too difficult. Lancaster and others, notably (Salton & McGill, 1986) anticipated these concerns as early as the 1960s by recognizing that Boolean systems were difficult for users to understand. Frants and et al. believeed that criticism of the difficulty of Boolean query formulation is mere criticism of existing operational systems and interfaces, rather than of Boolean logic as the underlying foundation of a system However, Lancaster was an early advocate of partial match systems coupled with relevance ranking of partial match results. This allows the user to make the decision of when he or she has found enough relevant documents, rather than presenting results as a complete unranked set that must be examined in total. Frants, et al. pointed out that Boolean logic-based information retrieval system do not preclude relevance ranking, and, indeed, in 1968 Lancaster described the use of weighted index terms to rank documents from a Boolean query. Many experimental systems that use statistical, linguistic, or other approaches to partial match, however, are more typically associated with relevance ranking (Belkin & Croft, 1987; Kinnucan et al., 1987; Sparck Jones & Willet, 1997). One thing that had to be changed to make online systems friendlier or easier to use was to improve the interface (Ahmed, McKnight, & Oppenheim, 2006).

Marchionini and Komlodi (1998) traced the development of interfaces for information retrieval systems from the 1970s through the 1990s; from interfaces "designed mostly for users who were highly specialized professionals" to those that "support casual, literate end users (i.e., average educated citizens) to the current emphasis on highly technical areas such as medical and scientific research to now include all areas of human interest". Ten years after Marchionini and Komlodi's descriptions, different interfaces continue to help a wide variety of users navigate and find a wide variety of textual, numeric, and graphical information. Interface development has paralleled user-centered research and development in information retrieval and the Web. Looking ahead, Marchionini and Komlodi predicted today's ever-present access that is "embedded in the larger information activities of life and customizable to individual preferences and abilities". Best practices for future user interfaces as described by Resnick and Vaughan (2006) include considerations about the structure and metadata of the corpus, automatic vocabulary matching, user control in browsing and searching, search assistance in the interface, and special considerations for mobile devices. Many of these were considerations even in Lancaster's early work, but even he did not anticipate the ever-presence in his lifetime of mobile information retrieval devices smaller than a deck of cards! Information system's basics have gotten more complex, mingling the components of the past with new structures, features, and design considerations made possible by development in hardware, software, and communications technologies. In turn the information industry itself has gotten more complex.

3. The Information Industry

In the 1970s and into the 1980s, the information industry was a world of secondary publishers of indexes and abstracts who leased their bibliographic databases to third party vendors or large library systems. The bibliographic databases and early search systems served as pointers to primary publications that remained in print containers such as printed journals. Today secondary publishers and third party vendors both still exist, but primary publishers are also electronic publishers and the lines between the three are less sharply drawn. Bibliographic databases pointed to printed content; today's content is most often completely in digital. Linking through technologies such as OpenURL and cooperative initiatives such as CrossRef draws all parties together for a unified search experience (Grogg, 2006). A library user may search on a bibliographic database such as PsychInfo that is made searchable by a third party vendor such as H. W. Wilson or ProQuest, and click on a "full text" button to be seamlessly taken to a selected article held on a primary publisher's full text e-journals platform.

Major scientific primary publishers, such as Elsevier, Wiley, Springer, etc. all have their own search and retrieval platforms in addition to participating in the search and retrieval systems of others by linking and other agreements. Their articles are likely searchable from their own platform, from various secondary indexes, and by major search engines such as Google with links back to their own repository of articles. The July 2007 issue of Full text Sources Online lists nearly 35,000 periodical titles available on average from nearly six different e-sources, including aggregators, primary publishers, and other online sources (Glose, Currado, &Orbanus, 2007). The biggest drivers of traffic to e-articles today are Web search engines, but the behind the scenes links to full texts are often a result of library and CrossRef linking (Grogg, 2006). As of 2004, the Gale Directory reported on over 18,000 databases (up from 301 in 1975), made available by nearly 2,000 database vendors. It was conceivable in 1973 for an online searcher to know the characteristics of every available online database; today they may know well just those few in a specific subject area or on selected search services. While government agencies still produce major databases and search systems (for example, the National Library of Medicine), the database industry now includes a majority of commercial organizations and professional societies.

Currently Full text Sources Online (FSO) is a directory of periodicals accessible online in full text through 30 aggregators and content providers. Published biannually in January and July, FSO lists over 45,000 newspapers, journals, magazines, newsletters, newswires, and transcripts. Each title entry comprises the aggregators and databases that provide the publication online in full text. Coverage dates, frequencies, and lag times of titles appearing online, as well as ISSNs and document types, are included. Also provided are more than 39,000 publisher's URLs indicating free archives, selected coverage, and Open Access Journals. Subject, geographic, and language indexes are supplied as well. FSO is also available on the Internet as FSO Online and for license as FSOe.FSO Online provides the same complete information from FSO, and is updated weekly online. FSOe is the licensed text version of the FSO database for network or intranet use, with quarterly updates.

Not only is the number of databases growing, the amount of information within each is also growing. By Williams' (2004) calculations, the number of records in databases increased by "a factor of 403" from 1975 to 2003; from a total of 52 million records to nearly 21 billion. There is, of course, much variation in both the number of records in databases and the average size of a record.

According to Williams: "The entities counted as database records vary widely but generally range from 200 to 2,000 words (or, in the case of non-word-oriented records, they require a comparable number of bytes for storage). Records may be citations, abstracts, news stories, magazine articles, biographical records, unique names of chemicals, unique chemical structures, property data, recipes, time series, software programs, images, or descriptions or listings of virtually anything."

The impressive growth of the information industry does not include the whole of the massive Web and does not begin to touch the annual production of information. Major recent trends include the continued consolidation of the information industry within a handful of major commercial players that are responsible for primary journal and book publications (Tenopir et al., 2007) and an acceleration of innovative search features, automatic indexing and abstracting tools, search platforms, and other software tools. Personal files, as envisioned by Lancaster, are now a reality, with a number of software tools that help researchers download and maintain personal files (Tenopir et al., 2006). Databases of today often have millions of records and extensive full texts. Visualization and clustering of search results help searchers cope when they retrieve thousands or tens of thousands of potentially relevant items. Many commercial online systems have added clustering or visualization techniques to their system displays recently after years of testing and development (Zhu &Hsinchun, 2005).

4. Literature Search

A literature search is a systematic and thorough search of all types of published literature in order to identify a breadth of good quality references relevant to a specific topic. The success of a research project is dependent on a thorough review of the academic literature at the outset. A literature search is, therefore, a methodical analysis of all printed and electronic sources for information on the desired, usually a scientific or technological topic. It can also mean a wide-range of sourcing of published information from in-house and exterior records.

A formal definition of literature search is, 'a well thought out and organized search for all of the literature published on a topic. A well-structured literature search is the most effective and efficient way to locate sound evidence on the subject one is researching. Evidence may be found in books, journals, government documents and the internet.'

4.1 Need for Literature Search

As literature search involved search and evaluation of the available literature in a given subject area. It is needed due to the following:

- Literature search is a core part of the academic communication process,
- It connects a researcher's work to wider scholarly knowledge as it demonstrates understanding, and puts any research that has been done in a wider context,
- Identifies potential issues with the work a researcher plans to do,
- Helps to avoid unnecessary duplication in research work, and
- To understand that the work to be undertaken is relevant, worth doing, and might add to body of knowledge.

4.2 Purpose of a Literature Search

The purpose of a literature search is to identify the existing information sources (including books, journal articles, and Web documents) most relevant to the research question being studied. In other words, the purpose of a literature search for any research is not to identify every existing resource related to the topic of the research, but rather to identify the most relevant resources. Literature search also helps to:

- Broaden searcher's and researcher's knowledge on a topic
- Helps in decision making as vast amount of information is published and available on a topic
- Enables information specialists to show skill at finding relevant information
- Allows for critical appraisal of research
- Locate what suggestions have been made for future research

4.3 Why Literature Search?

Research, especially scientific research is a process that needs to be developed gradually with present research building upon a knowledge base of information that resides in the scientific literature. The following are the three reasons why one needs to find, evaluate and use this literature:

- Literature Review
- Practical or Everyday Needs
- Current Awareness

4.4 Role of Literature Review

The literature review is an evaluative report of studies/information found in the literature related to a selected subject area. The review should describe, summarize, evaluate and clarify this literature. It should give a theoretical basis for the research and help to determine the nature of research to be carried out. It is required that a limited number of works that are central to selected subject area be selected rather than trying to collect a large number of works that are not as closely connected to the topic area.

The goals of a literature review are:

- To demonstrate a familiarity with the body of knowledge and establish credibility.
- To show the path of prior research and how a current project is linked to it.
- To integrate and summarize what is known in an area..
- To learn from others and stimulate new ideas.

5. Databases

Databases (DBs) are available in different forms, for example, Table of contents like Current Contents DBs, Full text DBs in the form of e-books, e-journals, e-theses, e-dissertations. These are available from publishers like Cambridge University Press, Elsevier, Emerald, Maney Publishing, Royal Society of Chemistry, Sage Publications, Taylor & Francis, ACS, AIP, ASME, ASCE, ASTM, ACM,IEEE, IOP, NPG, OSA, Oxford University Press,

Springer, Thomson Innovation, Wiley-Blackwell and others depending on searchers' subject requirements. There is a DIALOG DB vendor who has more than 2500 DBs at one place on payment basis to have access. STN is available from American Chemical Society for chemistry related information.

6. Formulating the Search Strategy

6.1 General Guidelines and Requirements

Searching is a very complex and time-consuming process. Therefore, use the databases intensively and critically. It is advisable to consult database help files, readings, etc. often. Searcher should work on his/her own, then reach consensus with the group on the best solutions. A digital diary should be maintained of search steps, rationale and results. A back up of search files including screenshots are helpful. The time required to develop an optimal search strategy is often underestimated.

For an appropriate search, certain guidelines and requirements are listed below:

a. There should be a clear description of the topic and the search strategy used.

- An explanation of the scope of the literature search with a clear understanding of the implications for searching
- Search topic broken down into main 'facets' or 'concepts'
- Rationale for the approach to searching and techniques used
- Explanation of decisions taken during search process.
- Results examined for relevance and revised as required
- Keep a record of each search

b. A wide range of relevant databases and sources of information explored

- Attempt to use a wide range of potentially relevant sources

c. A wide range of relevant search terms employed

- Appropriate use of synonyms
- Wide range of terms
- Imaginative use of synonyms
- Effective use of thesauri/controlled vocabulary if available
- Effective use of keyword index if available

d. Use of full range of appropriate search techniques

- Wide range of search operations
- Correct use of truncation and wildcards
- Limiting searches by field, if appropriate
- Taking into account alternative spellings
- Using Boolean operators effectively

e. Relevant references found covering all aspects of the topic or identification of gaps in evidence

- If a 'gap' is suspected, has a systematic approach been taken to confirm this?
- Discussion with other colleagues
- Contacting key organisations and experts
- Searching for unpublished and 'grey' literature
- Research being carried out currently that hasn't been published yet

f. References recorded accurately and consistently

- Consistent use of appropriate citation methods and referencing styles

6.2 Planning the Search

Regardless of the search tool being used, the development of an effective search strategy is essential if one hopes to obtain satisfactory results. A simplified, generic search strategy might consist of the following steps:

- Formulation of the research question and its scope
- Identification of important concepts within the question
- Identification of search terms to describe those concepts
- Consideration of synonyms and variations of those terms
- Preparation of the search logic

This strategy should be applied to a search of any electronic information tool, including library catalogues, CD-ROM and online databases. However, a well-planned search strategy is of especially great importance when the database under consideration is one as large and amorphous as the World Wide Web. Another factor that underscores the need for effective Web search strategy is the fact that most search engines index every word of a document. This method of indexing tends to greatly increase the number of results retrieved, while decreasing the relevance of those results, because of the increased likelihood of words being found in an inappropriate context. When selecting a search engine, one factor to consider is whether it allows the searcher to specify which part(s) of the document to search (e.g. URL, title, first heading) or whether it simply searches the entire document by default.

The most productive searches are those where the information seeker has spent time working out a search strategy before going online. The strategy is a pre-requisite for anyone attempting exhaustive searching, such as those embarking on research, and recommended practice for any user wishing to conduct an efficient search and avoid frustration caused by low retrieval. In situations where connect time is charged for a search strategy it is essential to prevent escalating costs.

Searcher and user should work out specific information needs and identify the different major concepts and alternatives. For example, the topic Inorganic fertilizers have two main concepts:

- inorganic fertilizers
- soil fertilization

Put ideas on paper in natural language.

Examine each concept to find as many synonyms and terms as one can think of, and group the related items together to provide the basis of a structure for searching:

Inorganic fertilizers	fertilization	soil
Soil fertilizers	fertilized plants	
Fertilizers	producing	factories

Consider the levels - the amount of information required, any limitations by date, language, etc. and add these qualifications to the structure.

6.3 Controlled Vocabularies and Thesauri

Controlled vocabularies and thesauri include lists of keywords which are "authorized terms" or descriptors used to organize subjects in a defined and standardized methodology to describe the contents of a work. There are multiple terms or synonyms applicable to a subject and controlled vocabularies and thesauri serve as a means of standardizing subjects into keywords that represent the concepts of that subject. This reduces ambiguity among subjects with multiple terms or synonyms and ensures that, most if not all works on the same topic will be indexed using the same keyword. Use of controlled vocabularies and thesauri enhances standardization of how works are described and indexed, promotes consistency of search results and allows for replication of search results using the same query.

A controlled vocabulary or thesaurus often includes a definition and some include scope notes to provide context for the keyword or as well as qualifiers or subheadings to allow for more precise searching. Some controlled vocabularies and thesauri offer additional keywords for searching in order to refine a search strategy. Most major databases utilize controlled vocabularies and thesauri for indexing of their works, with some using multiple controlled vocabularies and thesauri.

7. Developing the Search Strategy

The development of the search strategy includes conceptual formulation of query, translation of conceptual formulation into the language of keywords, descriptors or facets, identification of synonym and associated terms, etc. The concept of facet analysis (PMEST), given by Ranganathan as well as the concept of specific subject can be used as an effective tool for designing a query. After this, it is important to select the information domain to be searched like, the OPAC of a library, database or likewise, depending upon requirements. The search string or query, is the combination of terms, keywords or descriptors, which represent the information. As search strings contain vocabulary, the linguisic features and their implications on the search and retrieval of information have to be analyzed.

Here, three aspects, namely, syntactic, semantic and Boolean operators are to be understood. Syntactics of a search string deals with the kind of formula or connecting symbols through which keywords or terms are connected to represent the concept to be searched by the search engines. The semantics in a search string deals with the meaning of the string in the context of the required informtion and the interpretation by the search engine. The Boolean operators are explained in the subsequent section.

7.1 Boolean Logic

Boolean logic is the term used to describe certain logical operations that are used to combine search terms in many databases. The basic Boolean operators are represented by the words **AND**, **OR** and **NOT**. Boolean Operators are simple words (AND, OR and NOT) used as conjunctions to combine or exclude keywords in a search. These are used to connect and define the relationship between the search terms. Thus, resulting in more focused and productive results. These three terms are widely accepted by the designers of the search engines. They have well defined meaning while used as operators in information search. The three operators of Boolean logic are the logical sum (+) **OR**, logical product (X) **AND**, and logical difference (-) **NOT**. All the information retrieval systems allow the users to express their queries by using these operators. Let us now understand the implications of these three operators.

7.1.1 AND

If you need to pose a more specific query, use the Boolean operator **AND**, which limits results to those items that contain both (or all) of the search terms in your query. Again using the two words from the example above, the search query would retrieve only those items containing both words in the same item: "**Inorganic fertilizers AND Soil fertilization**".

This search query would return a much smaller set of hits, and the items would be more applicable to the field of inorganic fertilizers. To demonstrate the difference between the OR and the AND operator, run the two searches above using Internet. For example, the search query **Inorganic fertilizers OR Soil fertilization** returns over 31,000 items, while the query **Inorganic fertilizers AND Soil fertilization** returns only 176 items.

7.1.2 OR

The OR operator is useful for the first phases of a search, when one is not exactly sure what information is available on the topic or what words are used to categorize it. When used between two words, the OR operator instructs the search tool to retrieve any record containing either of the words. For instance, the search query would retrieve items containing either the word "fertilizers" or the term "fertilization": **Inorganic fertilizers** OR **Soil fertilization**

Once searcher views the types of items containing either word, one might want to narrow the search by dropping one term and confining the search to the other. For instance, one might find that the records indexed under the term "fertilizers" are more relevant to the research question than those indexed under "fertilization". Or, as in the example below, one might find that the items related to the specific field of "soil fertilization" must contain both words, not simply either one. Because OR is the Boolean operator that returns the most "hits" (items meeting the search criteria), search queries containing OR are very broad and sometimes return items that are not relevant.

7.1.3 NOT

The last of the three most common Boolean operators is the word **NOT**. The **NOT** operator is used to eliminate records containing a particular word or combination of words from your search results. For instance, if one is performing a general search on soil fertilization, one

might wish to exclude items dealing with the very specific discipline of "fertilizers production". To make this exclusion, one could construct the search query as: **Fertilizers** NOT **organic**

This search would return all items containing the word "fertilizers" except for those that also contain the word "organic.

Another example for Boolean operators searching is provided in the following figure:
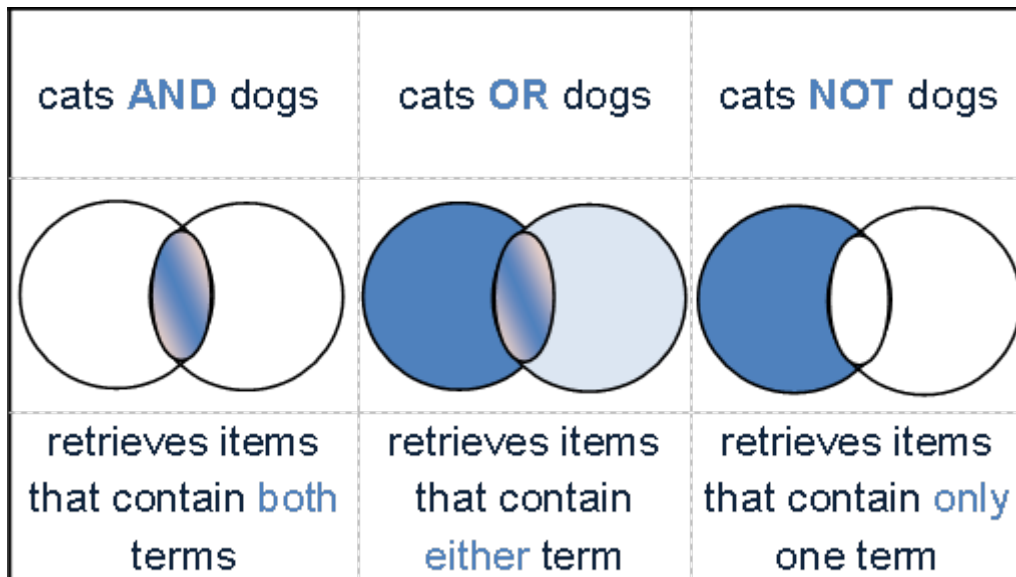
Fig.1: AND, OR, NOT

When we visit a search site, we should always read the instructions or help file before beginning the search. Each search engine has different parameters for using upper- and lower-case letters and combining Boolean operators. Another good method for refining the search is to run a few searches experimentally to see what results are returned. By browsing through the results list, we can determine whether or not the strategy is returning relevant items. Then, we can construct a search strategy using the Boolean operators **OR**, **AND**, and **NOT** to improve our results.

8. Choosing the Database and Host

If we are unsure which database to choose, help is at hand online. Some major hosts provide the facility for comparison of the number of occurrences of input search terms within each database they hold. However, it is advisable to ascertain names of the major databases in the area of search before committing to accessing a particular host which may not provide those particular databases.

The Search Process has the following steps:

- Choosing and developing a topic
- Designing the search
- Carrying out search and evaluating the results
- Handling the products of Search

8.1 Choosing and Developing a Topic

The first stage of any information search is to know what one is looking for as for behind any search, there should be a good, well-defined topic. But to know whether or not the topic is a good one, there are some general rules to follow:

a. If possible, choose a topic that interests you. There are fewer things more difficult than trying to write about a topic in which you have little or no interest.

b. Be sure your topic is neither too broad nor too narrow for the assignment you have been given. Check available time you have and the requirements to see how much you are expected to write about the topic.

c. Choose a topic about which there is likely to be information available in the library and/or on the Internet. You should do some preliminary checking for potential sources before you decide on your topic.

d. If you are selecting your own topic (rather than the one required to be searched by a user) make sure your topic is feasible before you start your research.

8.2 Designing the Search

There are a number of methods for finding a research topic. Depending on the available time and scope of the topic, the search can be designed. The searcher should have knowledge of data structure adopted by the database or the information system that stores data before executing a search. The system based search engines are designed to search information in a database according to its architecture. Depending upon the need and purpose of the search and expertise of the searcher, the search may be conducted using the features of the search engines. Hence a searcher should know the types of search and implications to get effective output.

8.3 Carrying out Search and Evaluating the Results

Carry out the search using the various search tools available. The quantity of published scientific material continues to grow exponentially. Fortunately there are tools (secondary sources: Encyclopaedias and dictionaries, Reviews, Databases, Abstracts and Indexes) which help you to search for information on a given topic. Evaluate the references you find, for relevance to your task. If necessary modify the search strategy.

8.4 Handling the Products of your Search

Having found interesting references, your next task is to make good use of them. This involves obtaining the corresponding full-text documents, critical examination of the material, organization of the information, possibly in some form of personal database, and incorporation into your personal frame of knowledge. This provides the starting-point for further work.

9. Subject Searching

In order to carry out a successful subject search, it is necessary to use the exact subject headings adopted by the database. Most searchers, at times, do not know these precise subject headings, and they use incorrect terminology and find no results. A successful search strategy starts with a keyword search, which would search most (if not all) of the fields in the records. Such an inclusive search is certain to retrieve some useful material. The searcher can then click on any relevant subject headings in the found records to conduct a proper subject search. Subject searches can be performed in two ways, firstly, keying in search terms, and secondly, selecting (clicking) one or more subject terms available in the database. The second type is easier, whereas the first requires some understanding of the database design. Initially, many searchers had claimed that they were familiar with subject searching, but results showed, after analyzing their search statements, that this was not the case.

10. Keyword Searching

New researchers, being novice researchers, would usually find keyword searches useful as, in most databases, keywords search most of the fields in the records. Analysis of search statements of searchers of keywords indicate that these can be categorized into two types: namely, simple keyword search, using one single search term; and complex keyword search, using a search statement of two or more search terms connected by one or more search operators.

10.1 Steps in Composing Complex Keyword Search Statements

Constructing suitable complex keyword search statements is crucial to developing searching expertise. There are at least two major steps in composing an effective complex keyword search statement: (i) constructing search terms, and (ii) using appropriate search operators to combine the search terms to form a search statement. In most searches, subject knowledge as well as searching skills are necessary to attain expertise in searching. Expert searchers are either knowledgeable in their domain, or they work together with experts in that particular subject area. This can most easily be observed in the construction of search terms. Without sufficient domain knowledge, searchers would have to carry out much background exploration, browsing through sites and clicking on links to arrive at the specific terms needed to conduct the actual search. Domain experts, on the other hand, could come up with the specific terms readily in composing search statements.

10.2 Constructing Search Terms

Constructing search terms comprises three processes: choosing key search terms, using related search terms, and considering various forms of the search terms.

10.3 Review Search Results

The best reviewer of the search results is the user. But the searcher or the information professionals should also review the search results on the basis of criteria given for evaluating information retrieval systems.

10.4 Edit Search Results

The editing of search results involves transformation of the search results into a user friendly format. This may involve arranging the results into a well-organised package, hightlighting the important entities, adding more information to the entities and reformating of information to suit the user's requirements.

10.5 Evaluation and Feedback

The evalution of search results involves participation of both, the users and the searchers. The quality and quantity of the results are assessed and if needed, the process may be redefined and restarted if the final result does not satisfy the users' needs.

11. Future Trends

Lancaster and Fayen (1973) made fourteen predictions of what the future of online systems might be. They recognized the danger of predicting the future and that "we may be just beginning to scratch the surface on the possibilities of applying technological advances to problems of information transfer". Danger aside, they were remarkably persistent in their predictions, which included:

- A great increase in the number of information services that can be accessed from around the world, including large general purpose systems and systems for specialized subjects;

- Specialized systems will be more "user oriented," easily accessible, and will require "comparatively little effort" to use;

- Systems will exploit the interactive, heuristic, and browsing powers of the online computer more fully for practitioners in a field, rather than information professionals;

- They should be oriented to natural language rather than controlled vocabularies;

- Vocabulary search aids at the time of searching will be incorporated, bringing together synonyms and semantically related terms;

- Computer aided instruction should be incorporated into systems;

- Systems should be capable of being searched by techniques other than formal Boolean expressions (including English language input, relevance ranking, fractional retrieval (partial match),

- On-line retrieval systems must certainly permit the ranking of output;

- Future on-line systems must require less effort to use. They should adapt to the user rather than expecting the user to adapt to them;

- Online systems and the equipment to use them must be more widely accessible;

- Systems will provide online support to personal files;

- Ultimately, on-line systems must interface with systems capable of retrieving and displaying complete text;

- Informal channels of communication will remain important and new communications technologies will "facilitate the transfer of information among scientists; and

- Online systems will interface with other systems, such as statistical packages, text editing programs, etc.

None of these predictions is controversial anymore, indeed, for those developments that are still only partially achieved; most researchers would wonder why progress has not been swifter. The Internet, developments in computing and telecommunications technology, and great leaps forward in software, standards, and digitization, have made the online information world of today remarkably similar to Lancaster's predictions. Stephen Arnold, an information industry thought-leader, remarked that "Many of the present developments in online systems built on ideas of the past, with hardware, software, and telecommunications advances making all of Lancaster's predictions at last possible."

Of course, not every development in today's online systems was predicted. The domination of large commercial Web search engines is changing user expectations and leading the way for system developments on an unexpected scale. Joining people and the power of online communication can merge the formal and informal information networks in ways that are just beginning. Physicist Paul Ginsparg (2000), founder of the physics e-print server now at arXiv.org, articulates the future vision of a "global knowledge network." He prefers this term to "electronic publishing," which connotes cloning a paper-based world rather than inventing a new way to communicate. In 2000 Ginsparg predicted: "In the next 10 to 20 years, it is likely that many research communities will move to some form of global unified archive system, without the current partitioning and access restrictions familiar from the paper medium, for the simple reason that it is the best way to communicate knowledge and hence to create new knowledge." This vision incorporates many elements that Lancaster foresaw nearly thirty years previously.

11.1 Some of the obstacles to consider

- Many databases available
- Each covers different types of information and subject areas
- Each has its own unique organization - Subject headings, indexing, limits are all different
- Conference proceedings
- Many access points available to locate conference proceedings – Databases & Websites
- Databases backup / unavailability time
- Many databases are only on charge basis (too costly)

12. Summary

Lancaster, with several different co-authors, was an early visionary and teacher in the practical aspects of online search and retrieval systems. From the earliest days of commercial online systems in the late 1960s and early 1970s he advocated better systems that would make online searching easier and more effective for those who have the information need. It took over three decades for online systems to begin to fully live up to the expectations described by Lancaster and Fayen and another decade for systems to begin to move into realms and ideas that expand on their expectations. The underlying structure and content of online searching laid in the 1960s and 1970s (and before) still serve online systems today. But this underlying structure, coupled with great advances in hardware, software, and telecommunications, is allowing growth of online systems into much more than the systems described by Lancaster (1973). End users not only have their hands on today's systems, their needs and experiences are driving developments and the future of information creation and retrieval as never before. Now most of the users' try to get the full text on their own from search engine sites and from the author's themselves. Only when there is a demand, online searching is done with the help of publishers DBs and DIALOG / STN, as per the users' request.

13. References

1. Ahmed, S. M. Z., McKnight, C., & Oppenheim, C. (2006). A user-centered design and evaluation of IR interfaces. Journal of Librarianship and Information Science, 38(3), 157-172.

2. Belkin, N. J., & Croft, W. B. (1987). Retrieval techniques. Annual Review of Information Science and Technology, 22, 109-145.

3. Frants, V. I., Shapiro, J., Taksa, I., & Voiskunskii, V. G. (1999). Boolean search: Current state and perspectives. Journal of the American Society for Information Science, 50(1), 86-95.

4. Ginsparg, P. (2000). Creating a global knowledge network. BMC News and Views I (9). Retrieved Jan 12, 2014, from Biomed Central http://www.biomedcentral.com/1471-8219/1/9.

5. Glose, M. B., Currado, T. D., & Orbanus, C. (Eds.). (2007). Fulltext sources online. Medford, NJ: Information Today.

6. Grogg, J. (2006). Linking and the open URL. Library Technology Reports, 42(1).

7. Kinnucan, M. T., Nelson, M. J., & Allen, B. L. (1987). Statistical methods in information science research. Annual Review of Information Science and Technology, 22, 147-178.

8. Lancaster, F. W., &Fayen, E. G. (1973). Information retrieval on-line. Los Angeles: Melville Publishing.

9. Marchionini, G., &Komlodi, A. (1998). Design of interfaces for information seeking. Annual Review of Information Science and Technology, 33, 89-130.

10. Resnick, M. L., & Vaughn W. V. (2006). Best practices and future visions for search user interfaces. Journal of the American Society for Information Science, 57(6), 781-787.

11. Salton, G., & McGill, M. (1986). Introduction to modern information retrieval. New York: McGraw Hill.

12. Sparck Jones, K., & Willett, P. (Eds.). (1997). Readings in Information Retrieval. San Francisco: Morgan Kaufmann.

13. Tenopir, C. (2001, May 1). Why I still teach dialog. Library Journal, 126, 36, 38.

14. Tenopir, C., Baker, G., & Grogg, J. (2007, May 15). The database marketplace 2007: Not your family farm. Library Journal, 132, 34-40, 42+.

15. Tenopir, C., Baker, G., Robinson, W., & Grogg, J. (2006, May 15). The database marketplace 2006: Renovating this old house. Library Journal, 131, 32-36.

16. Wang, Y D., & Forgionne, G. (2006). A decision-theoretic approach to the evaluation of information retrieval systems. Information Processing and Management, 42(4), 863-874.

17. Williams, M. E. (2004). The state of databases today: 2004. Gale Directory of Databases 2004, 1 (1). (Alan Hedblad, Ed.). Detroit: Thomson Gale.

18. Zhu, B., & Hsinchun, C. (2005). Information visualization. Annual Review of Information Science and Technology, 39 (1), 139-177.

19. http://books.infotoday.com/directories/fso.shtml#ixzz2qjigeX9Q