

African Research Review

An International *Multidisciplinary Journal, Ethiopia*

Vol. 5 (6), Serial No. 23, November, 2011

ISSN 1994-9057 (Print)

ISSN 2070--0083 (Online)

DOI: <http://dx.doi.org/10.4314/afrr.v5i6.10>

Information Retrieval Methods in Libraries and Information Centers (Pp. 108-120)

Onwuchekwa, Edeama O. - Information Services and Training

Personnel, National Open University Library, Lagos

Email: edemau@yahoo.co.uk

Tel No: +2348033765845

Jegade, Olumakinde Richard - College Library, Michael Otedola

College of Primary Education, Noforija, Epe, Lagos State

Email: olumakindejegede@yahoo.co.uk.

Abstract

The volumes of information created, generated and stored are immense that without adequate knowledge of information retrieval methods, the retrieval process for an information user would be cumbersome and frustrating. Studies have further revealed that information retrieval methods are essential in information centers for storage and retrieval of information. The paper discusses the concept of Information retrieval, the various information retrieval methods. It examines the users of these information methods and their information behavior. The conclusion emphasizes the need for a continuous evaluation of the information retrieval methods to make for an effective and efficient information retrieval system

Key Words: Information Retrieval, Libraries, Information Centers.

Introduction

For Centuries libraries have been organizing reading materials on shelves for easy access. However, systematic methods that have been widely adopted for

the organization of library materials and their recordings for use by readers came into being a little more than a century ago.

Today's information professionals should know and be conversant with the traditional information retrieval tools and methods like classification, cataloguing, and vocabulary control as well as the traditional manual indexing systems. This is because these traditional methods show the process of evolution of information retrieval and most importantly, many recent developments in information retrieval in web and digital library environments have their roots in these traditional tools and methods.

Different measures are currently taken for informing users about various materials accessible through a given digital or hybrid library.

The concept of information retrieval

According to Spack and Willet (1997), the term Information retrieval was coined in 1952 and gain popularity in the research communities from 1961. At that time, Information retrieval's organizing function was seen as a major advance in libraries that were no longer just storehouses of books, but as places where information was catalogued and indexed.

The concept of information retrieval presupposes that there are some documents or records containing information that have been organized in an order suitable for easy retrieval.

An information retrieval system is designed to retrieve the documents or information required by the user community. It should make the right information available to the right user. Thus, an information retrieval system aims at collecting and organizing information in one or more subject areas in order to provide it to the user as soon as it is asked for.

Lancaster (1968) comments in Chowdhury (1999) that an information retrieval system does not inform i.e change the knowledge of the user on the subject of his enquiry; it merely informs him of the existence or nonexistence and whereabouts of documents relating to his request.

The information retrieval system serves as a bridge between the world of creators or generation of information and the users of that information. Two broad categories of information retrieval have been identified:

- In-house Information retrieval
- Online Information retrieval

In-house Information retrieval systems are set up by a particular library or information centre to serve mainly the users within the organization. An example of an in-house database is the library catalogue. Online public access catalogue (OPAC) provides facilities for library users to carry out online catalogue searches, and then check the availability of the item required.

By online information retrieval systems, we mean those that have been designed to provide access to remote databases to a variety of users. Such services are available mostly on commercial basis, and there are a number of vendors that handle this sort of service.

Writers in times past have suggested that an effective and reliable information retrieval system must have provision for:

- Prompt dissemination of information
- Filtering of information
- The right amount of information at the right time
- Browsing,
- Getting information in an economical way
- Current literature
- Interpersonal communication and
- Personal help.

In a typical library environment there are two categories of users, the library and information personnel and the end users. Library and information personnel often act as intermediaries and they may also act as end users seeking for information for their own use or for decision making.

All information retrieval systems should be user oriented. As such the interest of every user should be given due importance at every point of information storage and retrieval.

Information retrieval methods

Libraries have been in existence since the beginning of writing and have served as a repository of the intellectual wealth of society. As such, libraries have always been concerned with storing and retrieving information in the

media it is created on. As the quantities of information grew exponentially, libraries were forced to make maximum use of information retrieval methods to facilitate the storage and retrieval process.

Cataloguing and metadata

In 1876 Melvin Dewey developed a systematic scheme of library classification, which became a unique tool for organizing library materials on the shelves. In the same year Charles A. Cutter brought out rules for a dictionary catalog, which enabled librarians to record systematically the library holdings in the form of catalogue entries that could be consulted easily by the user community.

Since then a number of schemes of library classification and catalogue codes have been developed to aid the process of organizing library materials systematically and to make for easy retrieval of information materials.

Svenonius (2000) lists the following major objectives of a catalogue, identified by cutter, as being:

To enable a person find a book by:

- Author
- Title
- Subject

To show what the library has:

- By a given author
- On a given subject
- In a given literature

To assist in the choice of a book

- By edition
- By character

But the above objectives as stated by Cutter were slightly modified by Lubetzky and then formally adopted at an International Conference on Cataloguing Principles held in Paris in 1961. In 1997, the objectives of a catalogue were reformulated as follows by an IFLA (International Federation

of Library Associations and Institutions) study group to suit the automated cataloguing environments.

- To find entities that corresponds to the user's search criteria
- To identify an entity
- To select an entity that is appropriate to the user's needs
- To enquire or obtain access to the entity described.

In the works of Su (1994), Large and Behesti(1997), Online public catalogues(OPAC) were first used in mid 1970s, it was only at the beginning of the next decade that a significant number of libraries switched from card catalogues to automated catalogues. However, those first catalogues were usually modules linked to the automated circulation system and had brief catalogue records and very limited functionality.

Online public catalogues (OPACs) are the interfaces that help users communicate with the collection(s) of a library. Typically OPACs allow users to search the library's catalogue, and also provide other facilities, such as checking borrower records, reserving reading materials, library news bulletins, and so on. Several changes have taken place and OPACs have improved significantly since then.

Internet resources vary significantly in terms of their content (text, numeric, audio, image video, etc), file format, availability, URL (uniform resource locator) or the address of a web page and so on. Some new rules and guidelines are required to help cataloguers make information retrieval easy and effective.

Metadata

Schwartz (2001) mentions that the term metadata, thus far used primarily in the field of database management, began to appear in LIS literature in the mid- 1990s. However, within a short period the topic became very popular area of research concentration, giving rise to several hundred publications.

Metadata has been classified into five categories on basis of their use:

- administrative metadata used in managing and administering information resources
- descriptive metadata used to describe or identify information resources

- preservation metadata related to the preservation management of information resources
- use metadata related to the level and types of use of information resources

Metadata supports a variety of operations and the users of the metadata may be human beings or computer programs. The primary functions of metadata are to facilitate the identification, location, retrieval, manipulation and use of digital objects in a networked area.

Metadata has become an important issue for information organization since the advent of the internet and the web. Digital Libraries have led to an increased awareness of the need for metadata for diverse categories of items available in digital format.

Classification

The first library classification Scheme was developed by Melvil Dewey in 1876. Universal Decimal Classification (UDC) was the second major classification scheme to appear. The Library of Congress Classification Scheme (LC) is another classification scheme; it is an example of a semi enumerative scheme of classification which provides a long list of all the classes in the universe of subjects.

Although classification schemes were mainly designed for organizing bibliographic items on the library shelves, many librarians and information professionals have also used library classification schemes for organizing information resources electronically. Typical examples are:

BUBL LINK: which provides access to catalogue of over 11,000 selected internet resources is catalogued according to DDC on all academic sources. Users can search the catalogue by selecting a Dewey class for example, 300 Social Sciences or by selecting a term /phrase from the alphabetical index. Thus a user can gain access to the digital resources by a classified list or through an alphabetical list of subjects

CYBERDEWEY: Is another example of the use of DDC in organizing digital information resources and dates from 1995 when David Mundie used DDC to organize internet information. The Cyber Dewey has similar features like the Bubl link where the user can select a specific Dewey class or term or phase from the alphabetical index to search the catalogue.

EEVL: Enhanced and Evaluated Virtual Library uses the EI Classification Scheme.

It is created and coordinated by a team of information specialists from a number of universities and institutions to provide access to digital information in engineering, mathematics and computing. From the main webpage of EEVL, Users can select a subject and once a subject is chosen a list of subclasses appears; the user can go down the hierarchy and finally get a list of items on a specific class or topic.

Indexing

Xin Lu (1990) writes that in the ideal document retrieval environment, a document or query statement is represented by a group of distinct index terms as well as the semantic relationships between these terms, so that retrieval could be based on a structure of semantic relationship.

Macleod (1990) also adds that documents are retrieved on the basis of the correspondence between search terms expressed in the query and the index terms in the document.

Indexing systems designed to assist in the retrieval of documents operate by assigning index terms to the analyzed subject of each document either manually or automatically.

The most difficult part of indexing is that phase where two different indexers analyze the content of a given document in two different ways resulting in two different index entries.

On the other hand, it has been observed by information retrieval experts that indexing tends to be more consistent when the vocabulary used is controlled, because indexers are more likely to agree on the terms needed to describe a particular topic if they are selected from a pre established list than if a free hand is given. And it is also easier on the users/ searchers part to identify the terms appropriate to the information need if the terms must be selected from a definitive list.

Both subject heading list and thesauri contain alphabetically arranged terms with necessary cross references and notes that can be used for indexing or searching in an information retrieval environment.

The different kinds of vocabulary control tools include subject heading list, thesauri and the thesaurifacet .

Rowley (1992) defines a thesaurus as a compilation of words and phrases showing synonyms and hierarchical and other relationships and dependencies, the function of which is to provide a standardized vocabulary for information systems. The function of the thesaurus is to exert terminology control in indexing and to aid in searching by alerting the searcher to the index terms that have been applied.

The thesaurus is a specialized kind of retrieval language with both a thesaurus type and classification type each containing some unique item to itself and not found in the other. The obvious advantage of a thesaurus is that it is used for arranging books on the shelves of a special Library as well as for indexing the terms in a database.

Some of these tools especially the subject heading list in catalogues have also been used to organize internet resources like the Infomine, Biome, Sosig etc.

With the development of technology, an automatic indexing is when the assignment of content identifiers is done with the aid of modern computer facilities. In automatic indexing environment; the lack of human expertise can be overcome by intelligent use of frequent vocabularies in stored records and information request.

Other advantages of automatic indexing are the maintenance of consistency in indexing, indexing time is saved, index entries are produced at a lower cost and better retrieval effectiveness is achieved.

Abstracting

A lot of authors have defined the abstract from different points of view. Lancaster (2003) defines an abstract as a brief but accurate representation of the contents of a document and he opines that an abstract is different from an extract, an annotation or summary

Rowley (1996) defines an abstract as a concise and accurate representation of the content of a document in a style similar to that of the original document. She adds that an abstract covers all the main points made in the original document and usually follows the style and the arrangement of the parent document.

Abstracts as documentary products always take the form of short texts either accompanying the original document or included in its surrogate.

Different criteria have been used by other information scientist to categorize the different kinds of abstracts and they are:

- Abstract by writer: these are abstracts written by authors, subject experts, or by professional abstractors.
- Abstracts by purpose are written to serve different purposes for example informative abstract, indicative abstracts, critical and special purpose abstract.
- Abstract by form is another different kind of abstract and can be differentiated as structured abstract, mini abstract and telegraphic abstract.

With the rapid increase in the availability of full text and multimedia information in digital form, the need for automatic abstracts or summaries as filtering tool is becoming extremely important. Craven (2000) in his works proposes a hybrid abstracting system in which some task are performed by human abstractors and others by an abstractors assistance software.

Users of information retrieval methods

The user is the focal point of all information retrieval systems because the sole objective of any information storage and retrieval is to transfer information from the source to the user. This makes it very necessary that the information manager should have an understanding of the nature and number of users, their activities (information requirements, information seeking behavior) etc.

The term “user” is quite ambiguous; there are several types of users of an information system. Within the context of an organization there could be the actual users, potential users, expected user etc.

Users may be limited by the organization they work for, by the nature of their work or profession, by age, sex or other social groups. On the other hand user categories may be identified by the nature of the libraries they use:

For an academic library, the primary users are students, teachers, researchers, administrators etc.

For special or research libraries, the primary users may be categorized as researchers, planners, policymakers, scientists and so on.

In the public Library environment, anyone can be a user (members of the general public) Children, students, housewives, the literate, neo-literate etc.

Information seeking behaviour of users

Information Seeking behavior or the pattern of using information retrieval systems depends on a number of factors. Some of the general points that affect the information seeking behavior of the user are:

- the users' awareness of, and ability to access other sources of information
- the users relationship with the information unit concerned
- the information unit's ease of accessibility
- the users working conditions
- the time available to the user for consulting information systems
- the amount of competition that exist in the user's field of activities
- the users past experience or knowledge
- how easily the user gets on with other people
- how friendly, knowledgeable, and efficient are the members of the information unit.
- the various products and services of the information

Various methods have been employed in user studies over the past decades, for example Wilson (1994) and Gorman and Clayton (1997) have discussed some such as:

- Questionnaires, interviews, case studies etc.
- Observation of events, reviews if diaries etc and
- Study of internal processes.

Users Studies are very essential for prescription, improvement and efficiency of information services.

Conclusion

According to Chowdhury (1999), an evaluation is basically a judgment of worth. The two basic parameters for measuring an information retrieval system are the Effectiveness and the Efficiency.

The effectiveness may be a measure of how far it can retrieve relevant information while withholding non-relevant information. By Efficiency, it is how economically the system is achieving its objective.

An evaluation study investigates the degree to which the proposed goals and expectations have been achieved or the degree to which these can be achieved.

From time to time the information retrieval methods of an information centre should be reviewed and evaluated.

References

Anglo-American cataloging rules, (1999). 2nd ed, 1998 revision, Joint Steering Committee for Revision of AACR , Library Association, American Library Association and Canadian Library Association. Available: www.ifla.org/VII/s36/pubs/drg03.htm

BUBL Link. Available at WWW.bubl.ac.uk

Craven, T. C.(1993). 'A Computer-aided abstracting tools kit', *Canadian Journal of Information science*, vol 18, 19-31.

Craen, T.C.(1988). 'Text network display editing with Special reference to the production of customized abstracts', *Canadian journal of information science*, 13, 59-68.

Craven, T, (1996). An Experiment in the use of tools for computer-assisted abstracting. In: ASIS'96: Proceedings of the 59th ASIS Annual Meeting 1996. Baltimore, MD, October 21-24, 1996, Vol.33, Medford, NJ, Information Today , 203-8.

Craven,T.C.(2000). 'Abstracts Produced using Computer Assistance', *Journal of the American Society for Information science*, 745-56.

CyberDewey: the fast well organized internet directory. Available at www.anthus.com/cyberDewey.html

Engineering E-Library, Sweden: the Sweden: the Swedish Universities of Technology Libraries. Available at <http://eels.lub.lu.se/>.

- Frants, V.I., Shapiro ,J. and Voiskunskii ,(1997). *Automated Information Retrieval: theory and methods*, San Diego, Academic Press.
- Hildreth, C.R., (1998). 'Online Library Catalogues as IR system: what can we learn from research .In Yates-Mercer,P.A (ed.),*Future trends in information science and technology* , Taylor Graham ,9-25.
- Korfhage , R.R.,(1997). *Information Storage and Retrieval*, New York, John Wiley.
- Kowalski, Gerald, Mark T.(2000). *Information storage and retrieval systems:theory and implementation*. Hingham, MA, USA: Kluwer Academic Publishers.
- Large,A. and Behesti , J.,(1997). 'OPAC: a research review',*Library and information science research*, (19)111-33.
- Lancaster, F., Elliker, C. and Colonell, T. H., 'Subject analysis', *Annual review of information science and technology* , 24, 34-84.
- Liston.D. and Schoene, L, (1978).A systems approach to the design of information systems'. In: King D. W.(ed). *Key papers in the design and evaluation of information systems*, New York, Knowledge Industry.
- Olson, N. B. (2003). (ed). *Cataloging internet resources: A manual and practical guide*,2nd edn..
- Rasmussen, E., (1999). 'Libraries and bibliographical systems'. In Baeza-Yates,R. and Ribeiro-Neto, B., *Modern information retrieval*, ACM Press,397-413
- Rowley, J (1994). 'Aspects of a library systems methodology', *Journal of information science*, 20 (1), 41-5.
- Rowley, J.(1996).*The basics of information systems*, 2nd ed, London, Library Association Publishing.

- Rowley, J. (1998). *The electronic library*. 4th edition of Computers for libraries, London. Library Association Publishing.
- Salton, G.(1989). Automatic text Processing: the Transformation, Analysis and Retrieval of information by computer, *Reading* , MA, Addison – Wesley.
- Schwartz, C.,(2001). *Sorting out the web: Approaches to Subject Access*, Westport, Ablex Publishing.
- Sparck, K. and Willett P. (1997). ‘Overall Information’. In: Sparck Jones , K. and Willett, P.(eds), *Readings in information retrieval* , San Francisco, Morgan Kaufmann Pub.Inc.,
- Svenonius, E, (2000). *The intellectual foundation of Information Organization*, Cambridge, MA, MIT Press.
- Wilson,T. D. (1994) Information needs and Users: fifty years of Progress? .In Vickery. B.C (ed), *Fifty years of Information Progress: A Journal of documentation Review*. Aslib.
- Xin Lu, (1990). ‘Document retrieval: A structural approach’, *Information Processing and Management*, 26 (2), 209-18.