Problems in Searching Online Databases: A Case Study of Select Central University Libraries in India

Projes Roy*, Shailendra Kumar**, and M.P. Satija***

Shaheed Rajguru College of Applied Sciences for Women
Jhilmil Colony, Vivek Vihar, Delhi-110 095

**Dept of Library and Information Science, University of Delhi, Delhi-110 007*
***Dept of Library and Information Science, Guru Nanak Dev University, Amritsar-143 005*
E-mail: projesroy@gmail.com; shail3@yahoo.com; satija_mp@yahoo.com

ABSTRACT

The purpose of this paper is to investigate and determine the different challenges faced by the searchers of online database in selected central university libraries in India. Eight central universities in India were chosen on the basis of the bandwidth utilisation (higher to lowest in each group). A total of 302 library users filled in the questionnaires (out of 320 approached) in the survey all eight central university libraries across India. It finds that easy and understandable content pages are the most desirable by the users; site feasibility is directly proportional to users interaction; and the retrieval techniques vary from subject to subject. This paper affirms that the databases should be selected on the basis of their retrieval aspects and the online features. The publisher of the online databases can not overlook the retrieval features in the databases. Easy and simple retrieval features can fetch more subscribers for the vendors and publishers. This is a comprehensive study that provides statistical data on searching/retrieval problem of online databases in different central universities in India.

Keywords: Retrieval command, online databases, central university, India

1. INTRODUCTION

Ease of use and efficiency of an information retrieval system (IRS) is ordinarily measured in terms of its user friendliness, and recall and relevant ratios. Indeed it also depends on the search strategy, query formulation skills and confidence of the online searcher as well as the inherent features of the IRS and the databases.

This study finds out existing trends in IR features and different data retrieval techniques required for success in searching different databases subscribed or provided access to by the Indian central university libraries. This study is confined to a select old and established eight central university libraries (Table 1) across the length and breadth of the country:

Of these eight universities the four, namely, DU, IGNOU, JMI, and JNU are in the National Capital Region while other four represent South (PU), North–East (AU), North Central (BHU) and East (VBU). At the time of study, there was no central university in the North West India.

Table 1. List of eight universities and their years of establishment

Name of university	Abbreviation	Year of establishment	Place
Assam University	AU	1994	Silchar
Banaras Hindu University	BHU	1916	UP
Delhi University	DU	1922	NCR
Indira Gandhi National Open University	IGNOU	1985	NCR
Jamia Millia Islamia	JMI	1969	NCR
Jawaharlal Nehru University	JNU	1969	NCR
Pondicherry University	PU	1985	Puducherry
Viswa Bharti University	VBU	1951	WB

This gives fair representation to all the regions of the country. Further, these eight universities were chosen on the basis of ranking and grouping their bandwidth utilisation (as taken from the Inflibnet, Ahmadabad website (www.inflibnet.ac.in). Two universities in each group representing the highest and the lowest network

communication bandwidth were selected for this study. Table 2 shows the bandwidth utilisation ranking.

Table 2. Bandwidth utilisation ranking

Name of university and Abbreviation	Bandwidth use	%
Assam University (AU)	137. 6 kb/s (6.7 %)	0.2
Banaras Hindu University (BHU)	24.9 kb/s (1.2 %)	7.2
Delhi University (DU)	649.1 kb/s (31.0 %)	31
Indira Gandhi National Open University (IGNOU)	16.4 kb/s (0.8 %)	4.6
Jamia Mallia Islamia University (JAMIA)	647.9 kb/s (30.9 %)	30.9
Jawaharlal Nehru University (JNU)	3240.0 b/s (0.2 %)	7.7
Pondicherry University (PU)	808.1 kb/s (38.5 %)	38.5

2. OBJECTIVES OF STUDY

The aim of this study was to examine the data retrieval techniques as offered and made available by database providers, and users expectations and their feedback on these issues. The main objectives of the study were to:

• identify different accessibility features of the online databases;

• find out different aspects of the feasibility of the sites of the online databases;

• identify different browsing features of the online databases;

• recognise different menu driven or command oriented retrieval techniques of an online database;

• evaluate the data retrieval techniques on the basis of the performance of the online databases;

• propose ranking of online databases on the basis of adaptability of data retrieval techniques; and

• identify the university libraries using maximum online databases.

3. RESEARCH METHODOLOGY

A brief literature survey was done to develop a basic understanding of the problem, and to design appropriate questionnaires. Various aspects taken up by the previous researchers[1-12] helped us to approach the problem in a holistic manner and become clear in objectives of the present study.

For the present study detailed information was collected through two sets of questionnaires, one for the library staff and the other for the users. For data collection investigator directly contacted all the possible information scientists employed by these university libraries, and made a convenient sample of 320 users of databases in these libraries. Throughout the data collection process one author personally visited all the eight libraries, and the questionnaire was administered personally to the librarians/information scientists and the users. Personal visits provided a vivid and clear picture of the prevailing situation and provided a chance to explain the objectives of the study and clarify any perceived ambiguity in the questionnaires to the respondents and faculty. Out of 320 respondents approached only 298 cooperated making a high return rate of 93 per cent. In order to supplement the data informal discussions were also carried out with the all these online users. This modus operandi alone helped to have deep insights into the prevailing conditions. Respondents revealed very useful information which otherwise could not have been known. The data was analysed manually and also with the help of software packages like Excel and SPSS.

4. SCOPE

Out of the eight universities surveyed, seven are members of the UGC-INFLIBNET consortium whereas the IGNOU is subscribing to 17 databases of its own, and is not a member of the above consortium. The 116 databases are of the following form: Bibliographic-15; E-books-29; Full-text journals-72.

These databases are from the following venders/ publishers: American Chemical Society (ACS), American Institute of Physics (AIP), Annual Reviews (AR), Blackwell Publishing (BWP), Cambridge University Press (CUP), Institute of Physics (IOP), Institute of Studies in Industrial Development (ISID), J-Gate Customs Content Consortia (JCCC), JSTOR, Oxford University Press (OUP), Royal Society of Chemistry (RSC), Society for Industrial and Applied Mathematics (SIAM), SpringerLink (SL), and Taylor & Francis (T&F). Of the 116 databases, strangely enough, only one database, namely, Project Muse (muse.inu.edu) was subscribed to by all the libraries, while 69 database were subscribed by only one university each. For the academic libraries the two most important consortia are UGC-Inflibnet and AICTE-Indest. All the libraries, except IGNOU, have consortia-based databases and some have self-subscribed, too. The DU and JNU have in-house grown databases also. The DU is also a member of the AICTE-Indest consortia, though it is mostly meant for professional and engineering institutes.

All the universities use both the commercial and the open source databases. Earlier net servers and CD-ROM mirror servers were very popular among the universities, but now university libraries under study are no more offering CD-ROM services. Multimedia databases are available only in DU and JNU, while all the eight libraries own and provide access to full-text and bibliographic databases. All libraries have self-subscribed as well as consortia-based databases, whereas the DU and JNU have also developed in-house databases. Being traditional and general universities under the purview of

the UGC, all are members of the UGC-Infonet consortium. But DU in addition is also a member of the AICTE-Indest consortia.

The study found that only DU, BHU, IGNOU, and JNU are hosting open access databases on their websites. Most of the universities allow online access to databases either through their intranets or directly from the web. Experienced librarians prefer the campus-wide facility for providing access to databases.

The AU, IGNOU, Jamia, JNU, and Viswa Bharti allow access through campus-wide intranets or internet with identification and password. Desktop remote access to university resources is preferred by busy scholars. Further, it was found that AU, BHU, DU, JNU, PU and Viswa Bharti allow IP-enabled access while IGNOU databases are password-based.

5. DATA RETRIEVAL TECHNIQUES

Though a common retrieval command is easy, but it does not provide access to all the databases. Some of the universities provide common links to databases through federated search. But only the DU has the federated search engine GISTfind. In this survey, only the DU and PU revealed their online search techniques while the other libraries preferred silence on the issue.

The DU uses Boolean, interactive, and proximity searches, while the PU uses quick search, simple search, advance Boolean search, in addition to menu-driven searches, as provided by their vendors. Accordingly, both these libraries are providing facilities for further refinement of large chunks of search output to make more precise and relevant search. However, none of the universities is using data mining technique.

Indeed, the information retrieval techniques vary from university to university. It depends upon both internal and external factors of a database. In the questionnaire about 48 aspects or facilities for information retrieval were identified (Table 3) and the users of these eight libraries were asked to mark their ease of use on the Lickert scale as: Essay-3; Challenging-2; Difficult-1.

The mean value of university-wise pattern is given in Table 4.

Table 3. Performing retrieval techniques in online databases

S. N.	Retrieval features	AU	BHU	DU	IGNOU	JMIA	JNU	PU	VU
1.	Article locater	3	3	3	3	3	2	2	1
2.	Article types	3	3	3	3	3	2	2	3
3.	Automatic translation software	2	2	1	3	1	2	1	2
4.	Boolean logic	2	3	3	3	3	3	2	3
5.	Citation search	2	3	1	3	1	3	2	2
6.	Classification code	3	2	2	3	2	2	2	3
7.	Cross reference search	3	3	2	3	2	2	1	1
8.	Custom links	3	1	2	1	2	3	2	3
9.	Density of terms	3	1	3	3	3	3	2	1
10.	E-mailing an article from the result list	3	3	2	3	2	3	3	3
11.	E-mailing citations from the result list	3	3	1	3	1	3	3	3
12.	Explode/expand search	3	1	3	3	3	3	2	3
13.	Field specific searches	3	3	3	3	3	3	3	3
14.	Frequency of terms	1	1	3	3	3	3	3	1
15.	Fuzzy searching	3	1	3	2	3	3	3	1
16.	Google custom search	3	3	3	3	3	3	3	3
17.	Have rules of precedence with nested queries	2	2	3	2	2	2	2	2
18.	Help menu/online tutorial/guide	3	3	3	3	3	3	3	3
19.	Hyphen	3	1	3	2	3	2	3	1
20.	Journal browsing	3	3	2	2	2	3	3	3
21.	Lateral searching	2	1	1	3	1	3	3	2
22.	Limit field searches	3	3	3	3	3	3	3	3
23.	Mapping	3	1	2	3	1	2	3	3
24.	Match of exact words/phrases	2	3	3	3	3	3	2	3
25.	Nested queries	3	1	1	1	1	1	2	1
26.	Persistent links	3	1	1	1	3	1	2	3
27.	Phases searching	2	2	3	3	3	3	2	3
28.	Proximity search	2	1	3	2	2	2	2	1
29.	Punctuation marks	3	3	2	2	3	2	2	1
30.	Query by example	3	1	2	2	3	2	1	3
31.	Range searching	2	2	3	2	2	2	2	2
32.	Reference link	3	1	3	2	3	3	3	3
33.	Save search	3	2	3	3	3	3	3	3
34.	Search history	3	2	3	3	3	3	3	3
35.	Searching for common phrases	3	1	3	3	1	3	3	3
36.	SMART links	3	1	3	3	2	3	2	1
37.	Sort order	2	1	3	2	2	3	2	3
38.	Special characters	2	3	2	1	3	2	2	1
39.	Spell check	2	3	3	3	2	3	2	3
40.	Stemming	2	2	3	2	1	2	3	2
41.	Stop word	2	1	2	3	1	2	2	3
42.	Subject authority	2	1	3	3	1	3	2	1
43.	Subject search	3	3	3	3	3	3	3	3
44.	Suggest subject headings	2	1	3	2	1	3	2	3
45.	Table of content	3	3	3	2	3	3	3	3
46.	Times cited	2	1	3	2	1	3	2	3
47.	Truncation	3	3	3	3	3	3	3	3
48.	Use of thesaurus or permuted index for searching	2	1	3	2	1	3	2	2

Table 4. Complexity level of IR in different university libraries

Universities	AU	BHU	DU	IGNOU	JMI	JNU	PU	VBU
Mean	2.58	1.95	2.54	2.52	2.22	2.61	2.37	2.34
Std. deviation	(0.54)	(0.92)	(2.54)	(0.64)	(0.85)	(0.57)	(0.6)	(0.89)

It is clear that some aspects and component of IRS are complex and difficult. The library workers have not only to learn them thoroughly but also have to make their users familiar with them. To overcome the problem, one requires:

(a) Online training and library orientation of users

(b) Asking the vendors to supply the databases in easy formats

6. FINDINGS

Easy and understandable content page was the most desired features by the users. Most of the features are common in online databases, but the JSTOR is ranked most efficient online database in browsing features. Project Muse was rated by majority of the users as the best system in terms of retrieval features. They experienced having better control in searching using natural language terms.

Feasibility of the site of online databases is directly proportional to its interactive quality. Hence, again it was rated as the most efficient online database is JSTOR, while the last ranked is Taylor & Francis in this regard.

It was assumed and sunsequently found correct that in subject-based online databases there is a significant difference in the retrieval features among the sciences, social sciences, and multidisciplinary databases. Certainly, it is not true of multidisciplinary databases with advanced retrieval features.

It was found that retrieval techniques vary from subject to subject. There is some significant difference to apply all the retrieval features to a simple online database. Different search features vary in different online databases. For example, 'use and application of wildcard' is highest in Elsevier databases, but it is lacking in Blackwell Publishing.

Further, the feature 'use of thesaurus or permuted index for searching' is present in all the databases except in Elsevier and Institute of Physics databases. Its highest percentage was recorded in the American Chemical Society. 'lateral search' and 'fuzzy search' also widely vary in databases. The 'concept maps', 'range search', 'fuzzy search', 'special character search', and 'hyphen search' are mostly recorded best in natural science subjects. On the other hand, 'lateral search', 'free text search' are recorded very high in social sciences, arts, and humanities.

Most significant finding can be stated nearly as an axiomatic law: 'In online databases usability is mainly influenced by the ease of data retrieval features: more efficient the techniques, more the use of online databases.' There is an efficient co-relation between facilities and usability of an online database. The usability percentage was highest in Project Muse followed by Emerald, JSTOR, and Elsevier.

In addition, marketing, education, and organisation are factors to reckon with to maximise the use of library resources among the patrons.

6.1 Results

Awareness and convenience seem to be major factors in selection of resources, whether prints or electronic.

Majority of users prefer remotely-accessed online resources to the printed ones. Faculty, and to a large extent students, access the resource remotely rather than in the library. However, they use only a small portion of the available information sources. Users tend to select a limited number of databases and seem to be unaware of the availability of broad spectrum of databases other than what they use regularly.

Findings suggest that databases without links to full-text databases have lower use. Likewise librarians should also consider selecting databases that provide full-text links to their online collection in a seamless manner.

Promotion of the online catalogue as a single-window to access both online and print journals would encourage their users based on need rather than convenience.

The ogranisation of online resources is also of paramount importance for optimum use of online journals and databases.

Most respondents indicated that they were aware of their library's web home page; and they probably used it to navigate to the library resources.

6.1.1 Difficulties Faced by Users

All users mentioned that they needed some help for search query formulation, selection of search terms, and they expected the system to provide 'suggested keywords and suggested search strategy' features. Difficulties faced by the users in utilising the retrieval features were related to the applications of the retrieval features. A complete understanding of the retrieval features is not easy as these vary from system to system. For instance users, specially the freshers, faced difficulties in finding synonyms from the online thesaurus. The thesaurus only displays broader term (BT), narrower term (NT), related term (RT), and many users could not self-translate the search strategy with synonyms and

pseudo-synonyms. Users were observed to be struggling to come up with the appropriate search terms.

Majority of the online databases vendors listed some features in the help menu which were not available in the databases, e.g., automatic translation software, facility to browse subject authority file, facility to suggest subject headings, number of times an article has been cited, and the spell check. Advanced users who were enthusiastic to experiment with unique features were completely frustrated to find those features non-existing. Apart from the features 'Help' or 'Tutorial', users also expected to see some search examples before starting. They required the search Help Menu and Guide to give simple instructions and offer examples for command.

Other problem was of connectivity while conducting retrieval tasks. A database suddenly became inaccessible due to some subscription-related conditions. The users had to re-login after leaving the home-page idle for some minutes.

6.1.2 User Education

To ease their difficulties, the users especially the teachers, need guidance and training in using online databases. The information literacy unit or the library staff should conduct training sessions for academic staff and researchers on the campus. If the teachers find the tools helpful then they would promote it among their students. It would be sensible to have a guide to library home pages that are targeted to different levels of users of different depth of engagement. Ease of use also depends upon the education, training, experience, institutional settings, personal traits, and attitude of the users. Certainly, dedicated staff is needed for online training of library users. Indeed there are UGC-funded posts of information scientist in universities. But mostly they are from computer sciences and as such they are not conversant with the spirit, philosophy, and idioms of library service. Therefore, there is an urgent need of an online librarian to manage, serve, and assist in dissemination of e-resources of a library.

7. EPILOGUE

Though the design of library databases is in a permanent flux, yet hopefully, these findings may guide the databases and interface designers in the development of easily accessible products. This study may help librarians to make better procurement decisions and challenge vendors through specific questions about the functionality of their wares. Further, there is a need to periodically repeat such studies. The changing use-pattern will require libraries to re-examine their collection development policies, institutional programmes, and reference service to meet the information needs of their patrons in the online environment.

REFERENCES

1. Battleson, Booth & Weintrop. Usability testing of an academic library web site: A case study *J. Acad. Librar.,* 2001, **27**(3), 188-98.

2. Harter, P. Stephen. Online information retrieval: Concepts, principles and techniques. Academic Press Inc., 1986.

3. Horwath, Jennifer. Evaluating opportunities for expanded information access: A study of the accessibility of four online databases. *Library Hi Tech*, 2002, **20**(2), 199-206.

4. Kline, V.L. Missing links: The quest for better search tools. *Online Inf. Rev.,* 2002, **26**(4), 252-55.

5. Narendra, Vivek; Stewart, Ron & Schmetzke, Axel. Accessibility and usability of online library databases. *Library Hi Tech*, 2005, **23**(2), 265-86.

6. Othman, Roslina & Halim, Nor Sahlawaty. Retrieval features for online databases: Common, unique, and expected. *Online Inf. Rev.,* 2004, **28**(3), 200-10.

7. Robins, D. Shifts of focus on various aspects of user information problems during interactive information retrieval. *J. Amer. Soc. Inf. Sci.,* 2000, **51**(10), 913-28.

8. Shukla, K.K. Some AI techniques for information retrieval. *DESIDOC Bull. Inf. Technol.,* 1996, **10**(4), 13-18.

9. Singh, Anil & Gautam, J.N. Electronic databases: The Indian scenario. *Electronic Library*, 2004, **22**(3), 249-60.

10. Sunil, A. & Nagabhushanam, V. Information retrieval system in the academic libraries: Recent trend. *DESIDOC Bull. Inf. Technol.,* 2004, **11**(5), 17-21.

11. Wilson, T.D. Information-seeking behaviour and the digital information world. *The Indexer,* 2006, **25**(1), 28-31.

12. Xie, Hong Iris. Online IR system evaluation: Online databases versus web search engines. *Online Inf. Rev.*, 2004, **28**(3), 211-19.