IndianJournals.com
A product of Diva Enterprises Pvt. Ltd.

Empirical Study on Subject-Based Information Retrieval with Special Reference to Bengali Documents

Sambhu Nath Halder

ABSTRACT

This study has been conducted to examine users' views regarding retrieving Bengali documents using subject headings. It also examines users' views of subject access, revealing issues with retrieving Bengali documents by subject in online public access catalogue. A stratified sampling method was used in the investigation. The users met in person to collect pertinent data while visiting the library and searching the OPAC. A structured schedule, developed for the purpose, was given to library patrons, and as a result, interviews and interpretations were meticulously documented. The results indicated that subject headings do not provide a good search experience. Although keyword search performs better, due to search redundancy, it is difficult to retrieve needed documents in the shortest time. It is also found that fewer respondents execute known-item searches, and their success rates are typically high. To sustain uniformity and consistency, factors accounted for standard guidelines and principles for the formulation of Bengali subject headings. It has been suggested that relevant insertions to use controlled vocabulary be included in the OPAC interface to meet the needs of new-age users.

BACKGROUND OF THE STUDY

For centuries, the library has served as a repository of human civilization's knowledge. The advancement of technology and its use in various disciplines is accelerating. Nowadays libraries are gradually transforming themselves by using web-based tools and techniques for the organisation of information and services. Using Integrated Library Management Software, libraries can perform technically as well as administrative jobs easily. Although at present some of the libraries are running with manual and automated systems simultaneously, some of them have already terminated the dual system. The library catalogue is the surrogate of the entire collection available in a library. The subject catalogue has an important role to satisfy the unknown approach, i.e. subject approach of the user community. Alzofon and Van Pulis (1984) researched the success rate of finding information in the online public access catalogue (OPAC). Another research work of Guo and Huang (2011) provides a comparative study of subject headings and subject search.

Miller and McCarthy (2010) edited the 20th edition of Sears List of Subject Headings (SLSH).

Librarian and IQAC Coordinator, Shimurali Sachinandan College of Education, Shimurali, Nadia, West Bengal, India
Email id: sambhu.halder@gmail.com

It is a comprehensive work that provides a controlled vocabulary of preferred and non-preferred subject headings. SLSH provides patterns, perfect instructions, and examples of Subject Headings for the English language. However, the study of the literature reveals that there are more than 4000 titles published annually in the Bengali language from West Bengal. University libraries in West Bengal are found to collect these Bengali language publications. The organisation of these publications is a huge task for university libraries. Although OPAC has recently emerged from the physical form of the library catalogue, subject access has not been very popular and attractive (Sridhar, 2004). In this context, Halder and Maity (2013) have pointed out that in absence of a uniform Bengali subject headings list, subject access is the most problematic and difficult area in OPAC. According to the Bengali calendar year, Krishnamay Bhattacharyay (1370) made a pioneering attempt to compile "Bisay Shironam," a Bengali Subject Headings List, but it is no longer updated. Another work of the Bengali Subject Heading List entitled "Bijnan: Bishay Shironam" was published by Mukhopadhyay (2001). This work has prepared a Subject Headings List on the subject science only. Besides, Ratna Banyapahyay wrote a book entitled "Bishay Shironam Gathan Pahati: Darshan, Sahitya, Shilpa Kala" in 2004. Bandyopadhyay (2001) finds that the rate of publication of Bengali documents is growing and the collection of those Bengali books, journals, etc. is increasing. However, many of the libraries comprehensively lack the subject catalogue for the retrieval of documents. Halder (2015) studied different aspects associated with the retrieval of Bengali documents using subject headings in the online catalogue. Where ever Bengali subject headings are assigned, there are not complete or proper entries and these also lack standardized principles. Therefore, due to the unavailability of standard Bengali subject headings, the users'

approach is not satisfied enough when searching on OPAC.

Because of the above deliberation, it is deemed necessary to assess users' attitudes towards the effectiveness of subject headings for retrieval of Bengali documents in OPAC.

OBJECTIVE

The purpose of this study is to determine how users feel about the use and effectiveness of subject headings in the Online Public Access Catalogue for Bengali documents. To accomplish this, the study attempted to precisely understand the attitudes of the users through the following endeavours.

- to analyse the user's approach to OPAC for document retrieval;

- to determine and identify the search patterns for retrieving Bengali documents in OPAC, as well as their success rate;

- to determine the efficacy of subject headings in OPAC for retrieval of Bengali documents;

- to determine the level of satisfaction of OPAC users for retrieval of Bengali documents by subject;

- to learn the users' thoughts on how to improve the OPAC's performance for retrieval of Bengali documents by subject.

METHODOLOGY

It was decided to use the survey method to achieve the goals of this study. Nonetheless, because the survey's population was extensive and heterogeneous, it was decided to conduct a sample study of users of university library OPACs for convenience. The University of Calcutta Library, Jadavpur University Library, University of Kalyani Library, and Vidyasagar University Library were the four university libraries in India on which the survey was conducted. To make

the sample representative of the population, two institutions were chosen from metropolitan regions while the other two institutions were chosen from suburban or rural areas.

A structured schedule was prepared to record data through systematic study and observation of the use of subject headings of Bengali documents in OPACs. A total of 800 schedules (200 for each of the selected university libraries) were distributed to record the primary data, out of which 680 (85 per cent) schedules were received till the end of data collection. The categories of students, research scholars and faculty members from each of four university libraries were selected using a stratified random sampling technique.

Besides, relevant data were collected from observation of the libraries, and their annual report, etc. The quantitative data obtained by the methods described above were analysed, classified, tabulated, compared and interpreted duly keeping in view the objective of this study.

DATA ANALYSIS AND DISCUSSION

The results of this investigation are explored in further depth here. Since any information retrieval system has two components, one is the information resource and the other is the users' query. As a result, user satisfaction is a critical factor in evaluating the system. The following is the statistical analysis of the survey data, as well as the required discussion:

User's approach to OPACs for retrieving documents

Users' attitudes about the appropriateness of OPACs for document retrieval have been divided into two sub-categories.

Frequency of OPAC usage

It is revealed from Table 1 that a maximum (48.38 per cent) of the respondents uses OPAC five or

Table 1: Frequency-wise distribution of usage of library OPAC

S.No.	Frequency	User	Percentage
1	5 or more times per week	329	48.38
2	2 to 4 times per week	85	12.50
3	2 to 4 times per month	159	23.38
4	Less than 6 times per year	107	15.74
	Total	**680**	**100.00**

more times per week, followed by 23.38 per cent of the respondents uses it 2-4 times per month, 12.50 per cent of the respondents use it 2 to 4 times per week, and 15.74 per cent of the respondents use it less than 6 times per year. This finding concurs with the study of Halder (2015).

Users' preference to get access to OPAC

Table 2 given below presents the distribution of respondents by their preference to access OPAC. The result of the study clearly shows that out of 680 respondents, the maximum number, i.e., 61.03 per cent gets access to OPAC from inside the library. Besides, 33.82 per cent of them get access to OPAC from remote places through the Internet. However, the remaining 5.15 per cent of the respondents prefer to get access to OPAC through campus-wide local area networks.

Search pattern and success rate

Users' attitude towards search pattern and success rate of different searches in OPACs for retrieving Bengali documents has been presented below.

Search approaches to OPACs

Figure 1 given below shows that the maximum of the search approaches for retrieval of known items is made through the title, i.e. 17.35 per cent. Besides, 13.82 per cent of the users prefer to search by author. However, the combined search (14.56 per cent) of the author and title is a little higher than the author's search. Altogether, the known-

Table 2: Users' preference to get access to OPAC

S.No.	Preference to access to OPAC	Category of users			Total
		Student	Research scholars	Faculty members	
1	Inside the library	352(63.20%)	46(64.79%)	17(32.69%)	415(61.03%)
2	Campus-wide Local Area Network (LAN)	26(4.67%)	4(5.63%)	5(9.62%)	35(5.15%)
3	Remote access through the Internet	179(32.14%)	21(29.58%)	30(57.69%)	230(33.82%)
	Total	**557(100%)**	**71(100%)**	**52(100%)**	**680(100%)**

items search is 45.73 per cent of the total number of searches performed.

Figure 1 depicts how respondents' preferences for OPAC search strategies differ. Searching by subject headings is the most common, accounting for 29.12 per cent of all search strategies. This finding supports Halder and Maity's (2013) claim that the subject approach is the most important in OPAC searches. Furthermore, keywords represent 25.15 per cent of the search strategies of respondents, in other words, the search for an unknown document. As a result, it may be deduced that a substantial proportion of visitors approach the OPAC through subject headings or undefined keywords. Unknown document searches by subject headings or keywords account for 54.27 per cent of all searches.

Document retrieval from subject search and keyword search

The results of searches by using subject headings and keywords were compared. Concerning search by the subject headings, 42.50 per cent got zero results, followed by 30.44 per cent got 1-100 results, 20 per cent got 100-499 results, and 7.06 per cent got 500 or more results. For a keyword search, 8.97 per cent got zero results, 25.74 per cent got 1-100 results, 31.76 per cent got 100-499 results, and 33.53 per cent got 500 or more results (Figure 2).

Comparison between searches by using subject headings and by keywords reveals that 42.50 per cent of searches by subject headings has shown zero results, and 33.53 per cent of searches by keywords has retrieved 500 or more documents,

Figure 1: Search approaches to OPACs

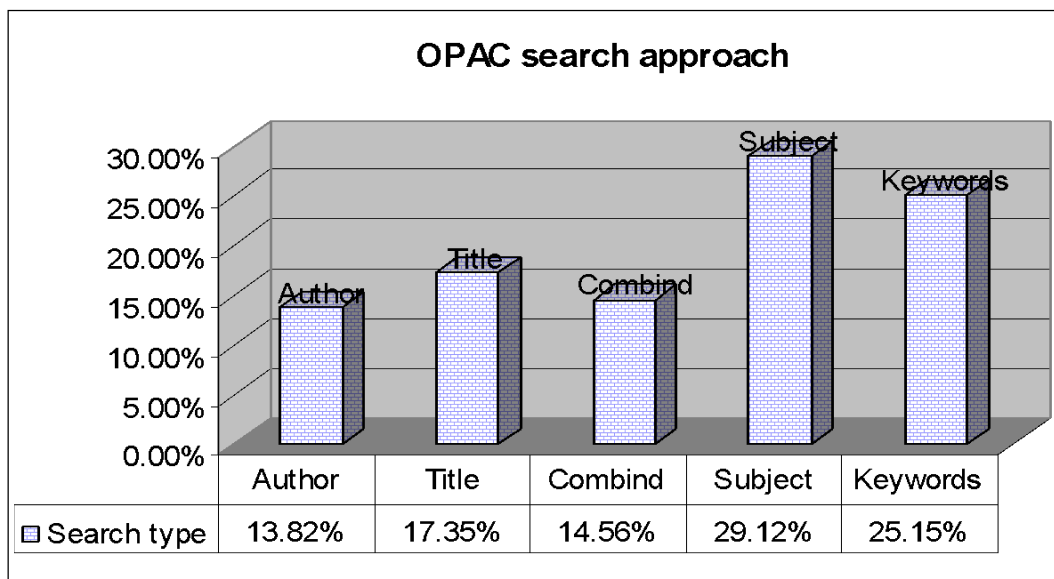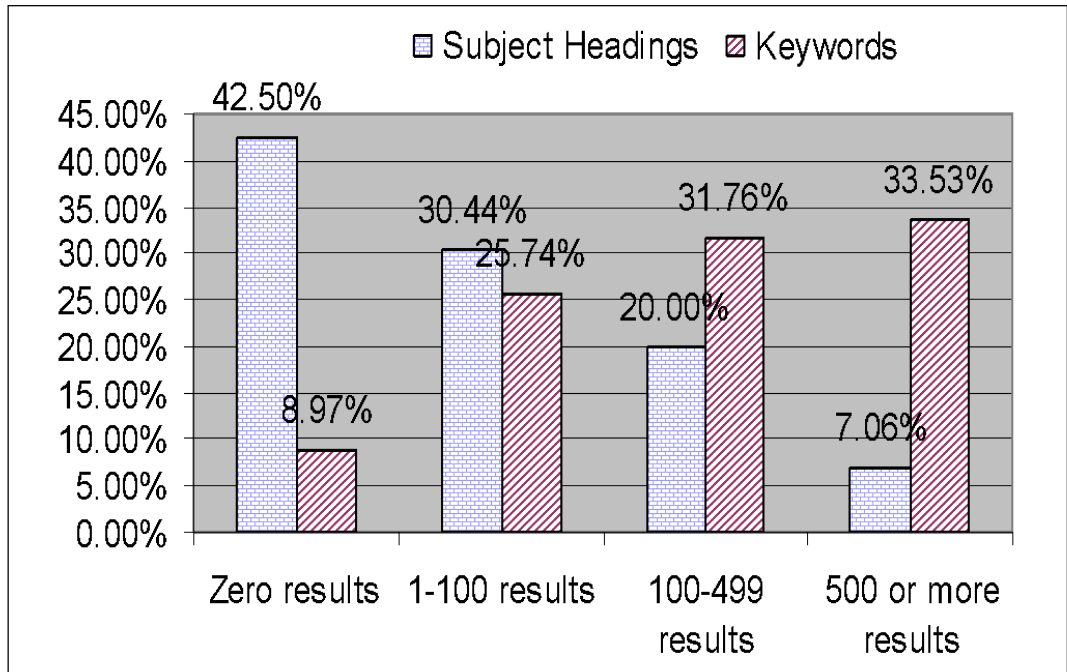|	Author	Title	Combind	Subject	Keywords
Search type	13.82%	17.35%	14.56%	29.12%	25.15%

Figure 2: Rate of document retrieval in OPAC

but only 7.06 per cent of search approaches by using subject headings have retrieved 500 or more documents. This finding is consistent with the study of Halder and Maity (2013). Therefore, it can be inferred that the retrieval rate of keyword searches is much greater than subject headings searches.

Impact of zero search results

Subject searches are identical to unknown-item searches. This implies that users do not know whether or not the database contains a certain document. To obtain documents, a subject search necessitates the use of some unique database searching strategies. This sort of search is more difficult than a known item search in terms of terminology selection abilities and database searching strategies. The examination of the data reveals that the respondents used a variety of ways to obtain documents (Figure 3).

It is found from Figure 3 that when search result becomes zero, 61.18 per cent of the respondents prefer to change search terms for searching OPAC. However, 25.88 per cent of the
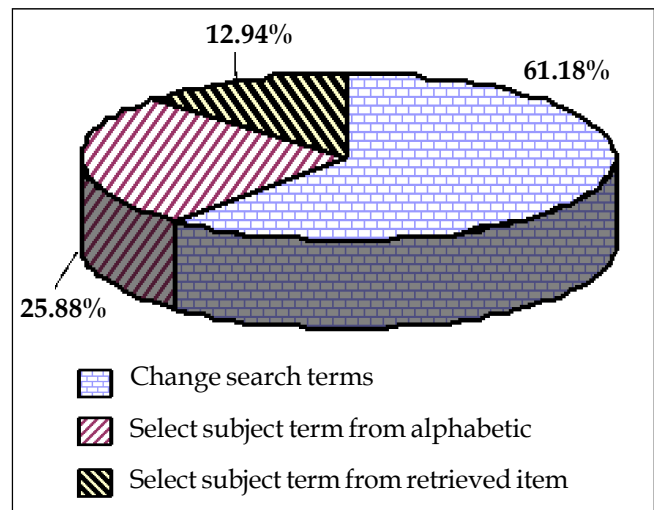

Figure 3: Adjustment of search terms when search results are zero

respondents take help from automatically suggested search terms available in the interface of the OPACs and rest 12.94 per cent of the respondents go back to use the search terms already used.

Acceptability of the Bengali subject headings

Users' attitudes towards the acceptability of the

Bengali subject headings have been presented below.

Problems in the retrieval of Bengali documents by subject

Retrieval of Bengali documents by subject in OPACs involves various difficulties. It is found from Table 3 that the respondents have marked several difficulties in searching Bengali documents by subject. The maximum of the respondents (60.44 per cent) has indicated that the OPAC is unable to retrieve the desired document(s). Besides, 59.85 per cent of the respondents have indicated that OPAC shows incomplete subject headings for Bengali documents. Due to the problems of synonyms, near-synonyms and homonyms, translation of the subject content of a document from Bengali to English word or phrase is a complex task.

A considerable number of respondents (43.24 per cent) have mentioned that equivalent English terms do not match in the course of subject searching. Another problematic area is the Romanization of subject terms from one language to another. It is revealed that 20.29 per cent of the respondents have indicated Romanization of terms does not match in course of subject retrieval in OPACs. Besides, the syntax of the subject

headings has been considered as an utmost important factor in vocabulary control. Subject headings are inconsistent due to a lack of uniform and standard controlled vocabulary device for Bengali documents. A good number of respondents (43.09 per cent) have indicated that the syntax of subject headings does not match in the course of searching. However, 20.44 per cent of the respondents have faced the problem of Bengali typewriting to enter a search term in the OPAC system.

The success of a subject search depends on the matching of subject entries existing in the OPAC database with the query term entered by the user. The users get satisfied when they can successfully retrieve their required documents. The study has revealed category-wise users' perception of the difficulties in OPAC search. It is worth noting that problems encountered by the students in the course of searching are not always the same as the problems of research scholars and faculty members. Most of the responses received from research scholars (69.01 per cent) and faculty members (86.54 per cent) have identified incomplete subject headings in OPAC as difficulty in retrieval, whereas most of the students (59.61 per cent) have indicated that they do not get desired documents in course of OPAC search.

Table 3: Difficulties faced by users in searching Bengali documents by subject

S.No.	Difficulties faced	Category of users			Total
		Students (N=557)	Research scholars (N=71)	Faculty members (N=52)	(N=680)
1	Unable to find out desired documents	332(59.61%)	41(57.75%)	38(73.08%)	411(60.44%)
2	Equivalent English terms do not match	249(44.70%)	28(39.44%)	17(32.69%)	294(43.24%)
3	Romanization of terms do not match	116(20.83%)	14(19.72%)	8(15.38%)	138(20.29%)
4	Syntax of subject headings do not match	242(43.45%)	36(50.70%)	15(28.85%)	293(43.09%)
5	Incomplete subject headings	313(56.19%)	49(69.01%)	45(86.54%)	407(59.85%)
6	Problem in Bengali typewriting	115(20.65%)	15(21.13%)	9(17.31%)	139(20.44%)

Note: Multiple responses against each of the parameters were received from each of the different categories of users. Therefore, responses did not denote any preferential order; rather they were independent as well as mutually exclusive. Responses were analysed and tabulated accordingly. Measurement of multiple responses against each of the parameters received from each of the different categories of users being shown in percentage terms is, therefore, independent as well as mutually exclusive.

It can be inferred that searching and retrieval of Bengali documents in OPACs involve several lacunas.

Results of unsuccessful subject search

Respondents were asked to respond to the question that if the search results were not satisfactory what action they would choose. Table 4 given below reveals that a maximum (43.09 per cent) of the respondents will opt to give up when search results are not satisfactory, while 22.94 per cent of the respondents will seek librarian's assistance to retrieve required documents. Besides, the study shows that 13.97 per cent of the respondents prefer to consult the card catalogue, and 13.53 per cent of the respondents will go to browse the shelves if the search results are unsatisfactory. A few (6.47 per cent) of the

Table 4: User's course of action next to unsuccessful subject search

S.No.	User's course of action	Number of respondents	Percentage
1	Give up	293	43.09
2	Browse the shelves	92	13.53
3	Ask a librarian	156	22.94
4	Ask a faculty member	44	6.47
5	Consult card catalogue	95	13.97
	Total	**680**	**100**

respondents will seek help from the faculty members.

However, in case of searching web-based OPAC from remote places, unsuccessful search results may lead the users to divert. So, the attitudes of the users were different in the course of subject search and unsatisfactory search results.
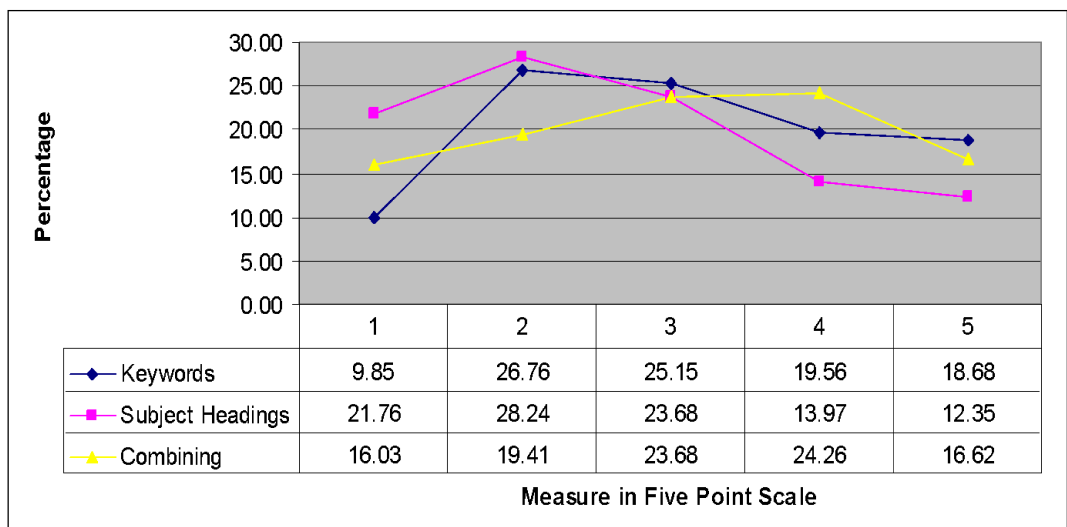
Level of satisfaction of the users

The level of satisfaction of the users in the searching and retrieval process was tried to uncover. Respondents were asked to put marks on a five-point scale, i.e. Very Unsatisfied-> 1, 2, 3, 4, 5 <-Very Satisfied. Results indicate that search performance using subject headings is not satisfactory, as only 12.35 per cent of the respondents are satisfied with this process. Rather keyword search technique is often used to obtain a required document(s), and only 9.85 per cent of the respondents are found least satisfied with this process. A combined search of subject headings and keywords is also used as a search technique; 24.26 per cent of the users have given 4 marks out of 5 and 23.68 per cent of the users have given 3 marks out of 5. Simultaneously 16.03 per cent of the users are found least satisfied and 16.62 per cent of the users are found satisfied with this search experience (Figure 4).

Figure 4: Comparison of satisfaction level of the users

Measure in Five Point Scale	1	2	3	4	5
Keywords	9.85	26.76	25.15	19.56	18.68
Subject Headings	21.76	28.24	23.68	13.97	12.35
Combining	16.03	19.41	23.68	24.26	16.62

Users' suggestions regarding the improvement of performances of OPAC

Initiatives for promoting the use of OPAC

End-users frequently experience challenges while searching in OPACs because of a lack of appropriate understanding of the concepts, tools, and procedures of searching and retrieval. However, in order to make the OPACs more effective, various user awareness programmes may be implemented. A survey of OPAC users was undertaken to obtain their thoughts on the subject. Table 5 shows how users feel about promoting the usage of OPAC.

It is revealed from Table 5 that a pretty good number of respondents (61.62 per cent) like to have a library orientation programme to know about the library. Besides, 25 per cent of the respondents like a seminar or workshop for the promotion of library services. However, 9.12 per cent of respondents prefer online instruction programmes for this purpose. Few respondents (3.97 per cent) seek personnel services from the library staff in the course of OPAC searching.

Opinions of the users about the advantages of OPAC

In this study, users were asked about OPAC's advantages for information retrieval. Table 6 presents the opinion of the users regarding the advantages of OPAC for searching for documents.

It is found from the table above 6, maximum users (56.19 per cent students, 71.83 per cent research scholars, and 86.54 per cent faculty members) feel OPAC saves time to search and retrieve required library documents. This finding concurs with the assertion of Halder (2015) that the online

Table 5: Users' opinion towards promoting the use of OPAC

S.No.	Promotional programme	Category of users			Total
		Students	Research scholars	Faculty members	
1	Library orientation programme	352(63.20%)	34(47.89%)	33(63.46%)	419(61.62%)
2	Seminar/workshop	143(25.67%)	17(23.94%)	10(19.23%)	170(25.00%)
3	Online instructions	39(7.00%)	16(22.54%)	7(13.46%)	62(9.12%)
4	Personnel services	21(3.77%)	4(5.63%)	2(3.85%)	27(3.97%)
5	Other programmes	2(0.36%)	0(0.00%)	0(0.00%)	2(0.29%)
	Total	**557(100.00%)**	**71(100.00%)**	**52(100.00%)**	**680(100.00%)**

Table 6: Advantages of OPAC

S.No.	Advantages of OPAC	Category of users			Total
		Students (N=557)	Research scholars (N=71)	Faculty members (N=52)	(N=680)
1	It saves time to search required documents	313(56.19%)	51(71.83%)	45(86.54%)	409(60.15%)
2	It can retrieve specific document at least possible time	109(19.57%)	24(33.80%)	17(32.69%)	150(22.06%)
3	It facilitates multimodal search options	123(22.08%)	35(49.30%)	15(28.85%)	173(25.44%)
4	Easy to get access and search can be done even from remote places	239(42.91%)	37(52.11%)	34(65.38%)	310(45.59%)

Note: Multiple responses against each of the parameters were received from each of the different categories of users. Therefore, responses did not denote any preferential order; rather they were independent as well as mutually exclusive. Responses were analysed and tabulated accordingly. Measurement of multiple responses against each of the parameters received from each of the different categories of users being shown in percentage terms is, therefore, independent as well as mutually exclusive.

catalogue saves valuable time for the users. Besides, 42.91 per cent of the students, 52.11 per cent of the research scholars and 65.38 per cent of the faculty members feel that it is easy to get access and search can be done even from remote places. However, 60.15 per cent of the 680 respondents believe it saves time to search and retrieve required documents, and 22.06 per cent believe it can retrieve a specific document in the least possible time. Furthermore, 25.44 per cent believe it facilitates multimodal search options, and 45.59 per cent believe it is easy to get access and search can be done even from remote places.

Opinion of the users regarding retrieval of Bengali documents by subject

A substantial majority of respondents believe that the capacity of OPACs to search and retrieve Bengali materials by subject should be improved. OPAC applications for subject search need to be more user-friendly from the users' perspective. Another set of users believes that the federated search function should be included in the OPAC system. In terms of technological advancement, library patrons desire an OPAC with new capabilities such as Web 2.0, word clouds, and relevance tagging to make searching more user-friendly.

CONCLUSION

It has been determined and established how library users feel about the use of subject headings for Bengali documents in the OPAC. In order to enhance the performance of OPACs and hence meet user requirements, it is concluded that some features need to be included. Essential instructions and assistance in using controlled vocabulary need to be made available in the OPAC. This can enable users to learn the search strategy and how to retrieve documents from the catalogue. The library must frequently provide on-site human intervention and establish plans for user education programmes. However, using the OPAC interface to carry out automated spelling and system correction will minimise the unsuccessful search results. It has been determined that there is no uniformity and consistency in assigning Bengali subject headings in the catalogue module. In order to cope with local variance, however, the decision table for cataloguing has to be correctly maintained.

Thus, the standard principles and rules for constructing Bengali subject headings have been defined to suit the requirements. Furthermore, if OPAC can give a relevancy search approach, users will be able to quickly find the information they need. These proposals are consistent with Halder's (2015) conclusion that, by taking account of the aforementioned findings and considerations, an efficient and successful OPAC system might be established.

Finally, it has been determined that the standard principles and rules for the construction of Bengali subject headings will meet the requirements. Besides, if OPAC can provide a "relevant search technique" then users can retrieve their required information easily. These suggestions concur with the arguments of Halder (2020) that by considering the above observations and deliberations, an effective and efficient OPAC system can be constructed. An effective OPAC search facility may be built to meet the needs of new-age users, with the possibility of ranking information retrieval.

REFERENCES

Alzofon SR and Van Pulis N, 1984. Patterns of searching and success rates in an online public access catalog. *College & Research Libraries*, Vol. 45, No. 2, pp. 110-115.

Bandyopadhyay R, 2001. *Suchikaran*. Kolkata: Bangiya Granthagar Parisad.

Bhattacharyay K, 1370. *Bishay Shiranam*. Kolkata: Bangiya Sahitya Parishad.

Guo J and Huang J, 2011. Subject headings and subject search: A comparative study. *Chinese Librarianship: an*

International Electronic Journal, 31. Retrieved from http://www.iclc.us/cliej/cl31GH.pdf

Halder S, 2020. Subject Approach to Information Retrieval with Special Reference to Bengali Documents: A Critical Study. *Library Philosophy and Practice (e-journal).* Retrieved from https://digitalcommons.unl.edu/libphilprac/4118

Halder SN, 2015. The use of subject headings for Bengali documents in online public access catalogue (Doctoral dissertation, Jadavpur University, Kolkata, India). Retrieved from http://hdl.handle.net/10603/199116

Halder SN and Maity G, 2013. Searching Pattern of Users of Online Public Access Catalogues for Retrieval of Bengali Documents in the University Libraries of West Bengal. *Librarian*, Vol. 20, No. 1, pp. 32-37.

Miller J and McCarthy S, 2010. Sears list of subject headings. New York: The H. W. Wilson Company.

Mukhopadhyay P, 2001. *Bijnan: bishay shironam.* Kolkata: Bona.

Mukhopadhyay PK, 1951. *Bangla granthaborgikaran: dashamik prothay Bangla granther bargikaran paddhati.* Kolkata: Orient Book Company.

Sridhar MS, 2004. Subject searching in the OPAC of a special library: problems and issues. *OCLC Systems Services*, Vol. 20, No. 4, pp. 183-191.