

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

2020

Subject Approach to Information Retrieval with Special Reference to Bengali Documents: A Critical Study

Dr. Sambhu Nath Halder

Shimurali Sachinandan College of Education, sambhu.halder@gmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Library and Information Science Commons](#)

Halder, Dr. Sambhu Nath, "Subject Approach to Information Retrieval with Special Reference to Bengali Documents: A Critical Study" (2020). *Library Philosophy and Practice (e-journal)*. 4118.
<https://digitalcommons.unl.edu/libphilprac/4118>

Subject Approach to Information Retrieval with Special Reference to Bengali Documents: A Critical Study

Dr. Sambhu Nath Halder
Librarian & IQAC Coordinator
Shimurali Sachinandan College of Education
Shimurali, Nadia, e-mail: sambhu.halder@gmail.com

Abstract: The library provides its services to satisfy the user's approach. This study is done to assess users' attitudes towards the retrieval of the Bengali document in OPAC through subject headings. It also traces users' perceptions regarding subject access highlighting the problems in the retrieval of Bengali documents by subject in OPACs. It can be inferred that some features are required to be incorporated to improve the functionality of the OPACs and thereby to satisfy users' requirements. Results revealed that search experience using subject headings is not satisfactory. Keywords search rather performs better but due to redundancy of search, it is hard to retrieve required documents in the least possible time. It is also revealed that fewer perform known-item searches and their success rates for these searches are generally good. Thus, the findings demand standard principles and rules for the construction of Bengali Subject Headings. As per the need of new-age users, an efficient OPAC search facility can be developed where there will be a chance for ranking information retrieval.

Keywords: Bengali Documents, Federated Search, Information Retrieval, OPAC, Subject Headings, User Study.

1. THEORETICAL PERSPECTIVE OF THE STUDY

For ages, the library has been established as a resource of knowledge of human civilization. The development of technology and its application in various fields is increasing steadily. Nowadays libraries are gradually transforming themselves by using web-based tools and techniques for the organization 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 catalog is the surrogate of the entire collection available in a library. The subject catalog 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 catalog (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). 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 organization of these publications is a huge task for university libraries. Although OPAC has recently emerged from the physical form of the library catalog, subject access has not been very popular and attractive (Sridhar, 2004). In this context, Halder and Maity (2013) have pointed out that the absence of a uniform Bengali subject headings list, subject access is the most problematic and difficult area in OPAC. A pioneering attempt in compiling "Bisay Shironam", a Bengali Subject Headings List, was made by Krishnamay Bhattacharyay (1370), but it is not up-to-date. Another work of Bengali Subject Heading List entitled "Bijnan: Bishay Shironam" was published by P. K. 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. Bandyapadhyay (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 catalog for retrieval of documents. Halder (2015) studied different aspects associated with the retrieval of Bengali documents using subject headings in the online catalog. 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.

2. OBJECTIVE

The present study attempts to assess the attitude of the users towards the use and effectiveness of subject headings for Bengali documents in the Online Public Access Catalogue. To achieve this, the study endeavored to comprehend the users' attitudes precisely through the following tasks.

- i). to examine the user's approach to OPACs for retrieval of documents;
- ii). to determine and identify the search patterns for retrieving Bengali documents in OPACs as well as their success rate;

- iii).to assess the effectiveness of the subject headings in OPACs for retrieval of Bengali documents;
- iv).to measure the level of satisfaction of users of OPACs for retrieval of Bengali documents by subject;
- v). to know the users' opinions towards improving the performances of the OPACs for retrieval of Bengali documents by subject.

3. METHODOLOGY

To fulfill the objectives of this study, it has been intended to adopt the survey method. Nevertheless, as the population of this study was large and heterogenous, therefore for convenience, it has been intended to carry out a sample study of the users of OPACs of university libraries. The sample on which the survey was conducted was consisting of the four university libraries in India, i.e. University of Calcutta Library, Jadavpur University Library, University of Kalyani Library, and Vidyasagar University Library. Two universities were selected from the urban areas, and the other two universities were from the suburban or rural areas to make the sample representative of the population, and for convenience.

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 percent) 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 analyzed, classified, tabulated, compared and interpreted duly keeping in view the objective of this study.

4. DATA ANALYSIS AND DISCUSSION

The findings obtained in this study are discussed in more detail below.

4.1 User's approach to OPACs for retrieving documents

Users' attitude towards the acceptability of OPACs for retrieving documents has been presented below.

4.1.1 Frequency of OPAC usage

It is revealed from table 4.1 that maximum (48.38 percent) of the respondents uses OPAC five or more times per week, followed by 23.38 percent of the respondents uses it 2-4 times per month, 12.50 percent of the respondents use it 2 to 4 times per week, and 15.74 percent of the respondents use it less than 6 times per year. This finding concurs with the study of Halder (2015).

Table 4.1: Frequency-wise distribution of usage of library OPAC

Sl. 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

4.1.2 Users' preference to get access to OPAC

Table 4.2 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 percent gets access to OPAC from inside the library. Besides, 33.82 percent of them get access to OPAC from remote places through the Internet. Remaining 5.15 percent of the respondents prefer to get access to OPAC through campus-wide local area networks.

Table 4.2: Users' preference to get access to OPAC

Sl. 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%)

4.2 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.

4.2.1 Search approaches to OPACs

Figure 4.1 given below shows that the maximum of the search approaches for retrieval of known items is made through the title, i.e. 17.35 percent. Besides, 13.82 percent of the users prefer to search by author. However, the combined search (14.56 percent) of the author and title is a little higher than the author's search. Altogether, the known-items search is 45.73 percent of the total number of searches performed.

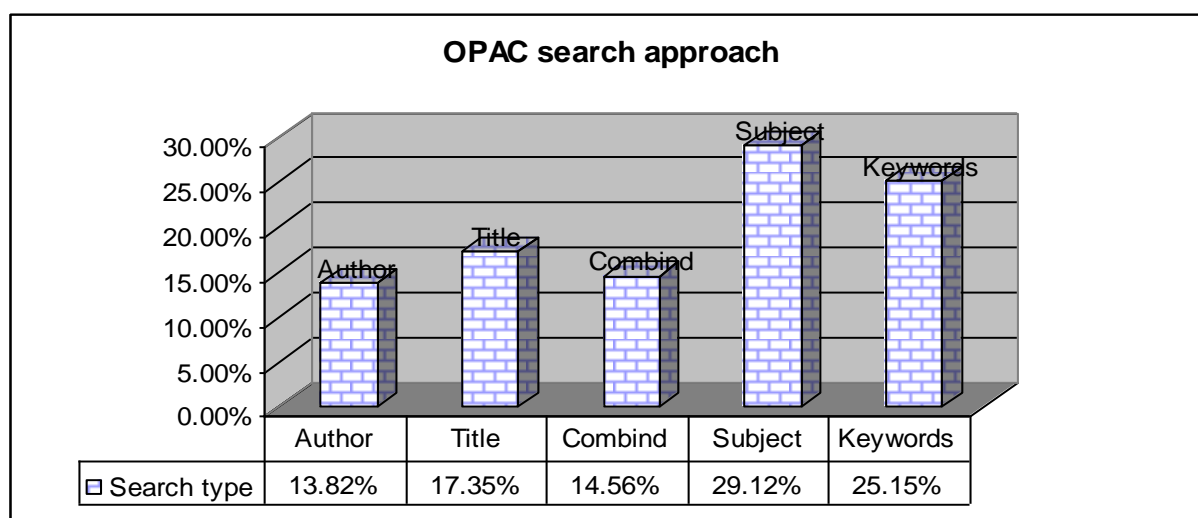


Figure 4.1: Search approaches to OPACs

As a whole, figure 4.1 demonstrates that preference to search approach to OPACs varies among the respondents. Searching by Subject Headings is 29.12 percent of the total search approaches, which is highest among all. This finding agrees with the assertion of Halder and Maity (2013) that the subject approach is highest in OPAC search. Whereas 25.15 percent of the search approaches of the respondents are made by keywords, i.e., search for the unknown document. Therefore it can be inferred that a large number of users' search approaches to OPAC are either made by subject headings or by unspecified keywords. Altogether the unknown document search by subject headings or by keywords is 54.27 percent.

4.2.2 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 percent got zero results, followed by 30.44 percent got 1-100 results, 20 percent got 100-499 results, and 7.06 percent got 500 or more results. For a keyword search, 8.97 percent got zero results, 25.74 percent got 1-100 results, 31.76 percent got 100-499 results, and 33.53 percent got 500 or more results (figure 4.2).

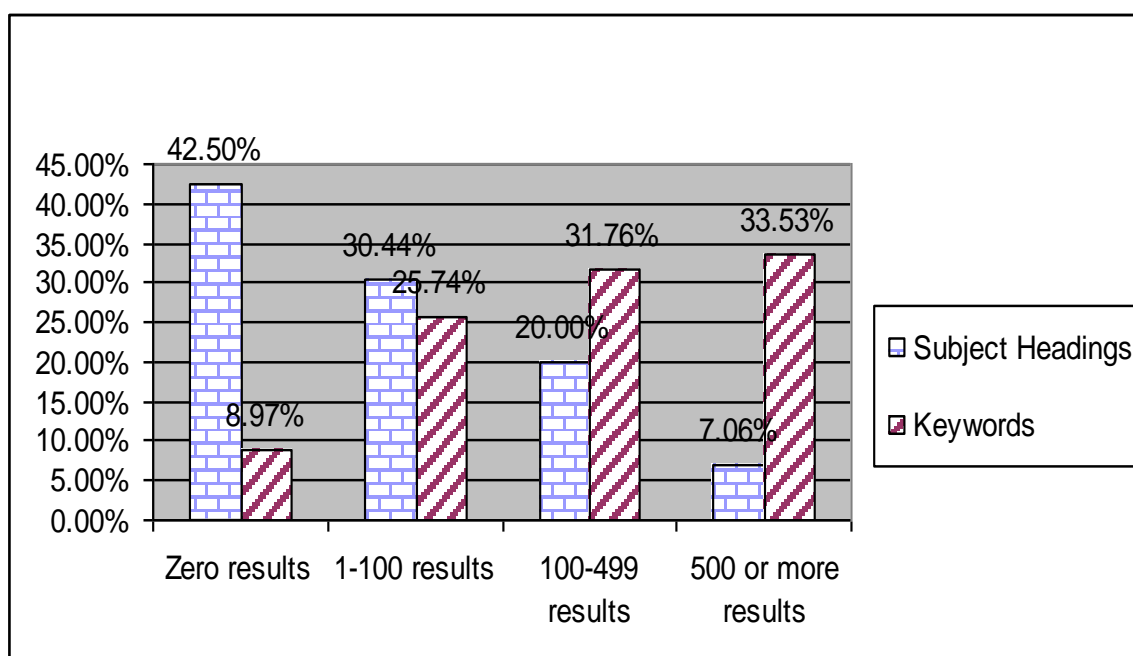


Figure 4.2: Rate of document retrieval in OPAC

Comparison between searches by using subject headings and by keywords reveals that 42.50 percent of searches by subject headings has shown zero results, and 33.53 percent of searches by keywords has retrieved 500 or more documents, but only 7.06 percent 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.

4.2.3 Impact of zero search results

Subject searches are unknown-item searches. It means that users do not know whether a specific document exists or not in the database. Subject search requires some special database searching techniques to retrieve documents. As far as skills of terminology selection and techniques of the searching database are concerned, this type of search is more

complicated than a known item search. It is revealed from the analysis of data that various techniques have been applied by the respondents to retrieve documents (figure 4.3).

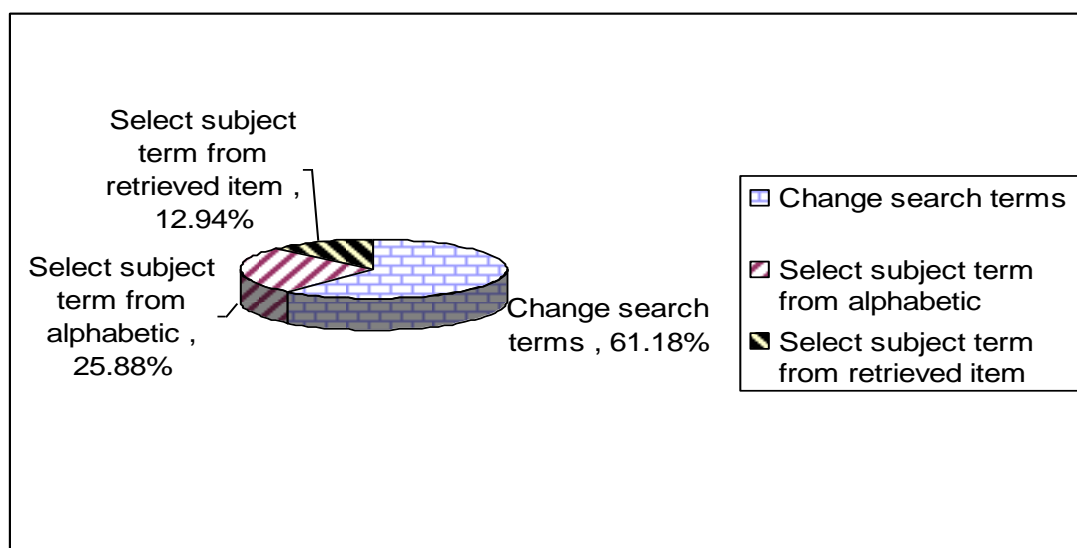


Figure 4.3: Adjustment of search terms when search results are zero

It is found from figure 4.3 that when search result becomes zero, 61.18 percent of the respondents prefer to change search terms for searching OPAC. However, 25.88 percent of the respondents take help from automatically suggested search terms available in the interface of the OPACs and rest 12.94 percent of the respondents go back to use the search terms already used.

4.3 Acceptability of the Bengali subject headings

Users' attitudes towards the acceptability of the Bengali subject headings have been presented below.

4.3.1 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 4.3 that the respondents have marked several difficulties in searching Bengali documents by subject. The maximum of the respondents (60.44 percent) has indicated that the OPAC is unable to retrieve the desired document(s). Besides, 59.85 percent 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 percent) 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 percent 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. Due to the unavailability of uniform and standard controlled vocabulary device for Bengali documents, subject headings lack consistency. A good number of respondents (43.09 percent) have indicated that the syntax of subject headings does not match in the course of searching. However, 20.44 percent of the respondents have faced the problem of Bengali typewriting to enter a search term in the OPAC system.

Table 4.3: Difficulties faced by users in searching Bengali documents by subject

Sl. No.	Difficulties faced	Category of users			Total N=680
		Students N=557	Research scholars N=71	Faculty members N=52	
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	The 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 analyzed 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.

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 percent) and faculty members (86.54 percent) have identified incomplete subject headings in OPAC as difficulty in retrieval, whereas most of the students (59.61 percent) have indicated that they do not get desired documents in course of OPAC search.

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

4.3.2 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 4.4 given below reveals that maximum (43.09 percent) of the respondents will opt to give up when search results are not satisfactory, while 22.94 percent of the respondents will seek librarian's assistance to retrieve required

documents. Besides, the study shows that 13.97 percent of the respondents prefer to consult the card catalog, and 13.53 percent of the respondents will go to browse the shelves if the search results are unsatisfactory. A few (6.47 percent) of the respondents will seek help from the faculty members.

However, in case of searching web-based OPAC form 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.

Table 4.4: User's course of action next to unsuccessful subject search

Sl. No.	User's course of action	No. of respondents	%
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 catalog	95	13.97
	Total	680	100

4.4 Level of satisfaction of the users

The level of satisfaction of the users in 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 percent 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 percent of the respondents are found least satisfied with this process. Combined search of subject headings and keywords is also used as search technique; 24.26 percent of the users have given 4 marks out of 5 and 23.68 percent of the users have given 3 marks out of 5. Simultaneously 16.03 percent of the users are found least satisfied and 16.62 percent of the users are found satisfied with this search experience (figure 4.4).

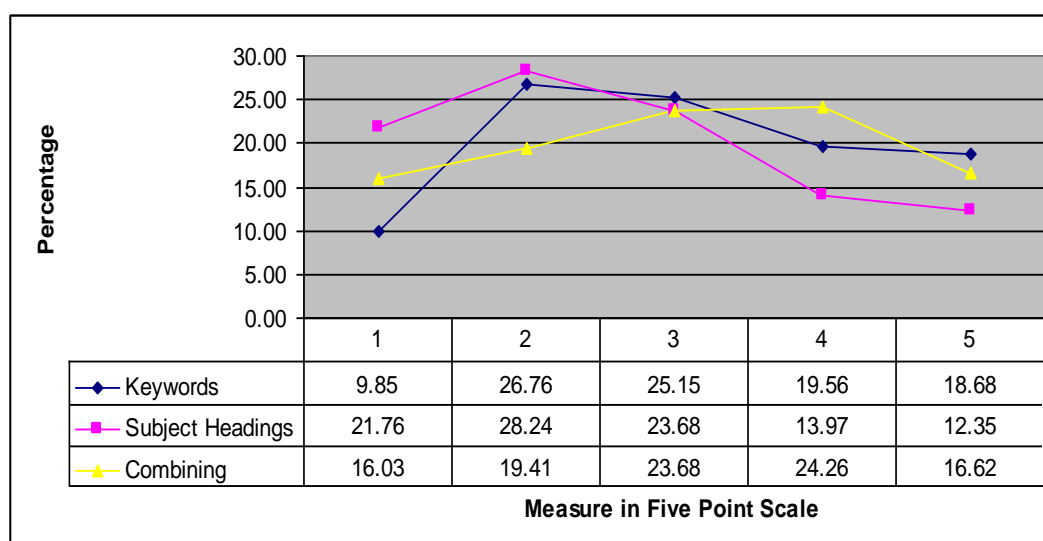


Figure 4.4: Comparison of satisfaction level of the users

4.5 Users' suggestions regarding the improvement of performances of OPAC

4.5.1 Initiatives for promoting the use of OPAC

Due to a lack of adequate knowledge about the principles, tools, and techniques of searching and retrieval, end-users very often encounter difficulties in the course of OPAC searching. However, some user awareness program may be undertaken for the effective utilization of the OPACs. A study of the users of OPACs was conducted to get the opinions of the users in this regard. Table 4.5 presents the users' opinions towards promoting the use of OPAC.

Table 4.5: Users' opinion towards promoting the use of OPAC

Sl. No.	Promotional program	Category of users			Total
		Students	Research scholars	Faculty members	
1	Library orientation program	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 programs	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%)

It is revealed from table 4.5 that a pretty good number of respondents (61.62 percent) like to have a library orientation program to know about the library. Besides, 25 percent of the respondents like a seminar or workshop for the promotion of library services. However, 9.12 percent of respondents prefer an online instruction program for this purpose. Few respondents (3.97 percent) seek the personnel services from the library staff in the course of OPAC searching.

4.5.2 Opinions of the users about the advantages of OPAC

In this study, users were asked about OPAC's advantages for information retrieval. Table 4.6 presents the opinion of the users regarding the advantages of OPAC for searching for the documents.

Table 4.6: Advantages of OPAC

Sl. No.	Advantages of OPAC	Category of users			Total
		Students N=557	Research scholars N=71	Faculty members N=52	

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 analyzed 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 is found from the table above 4.7, maximum users (56.19 percent students, 71.83 percent research scholar, and 86.54 percent 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 catalog saves user's valuable time. Besides, 42.91 percent of the students, 52.11 percent of the research scholars and 65.38 percent of the faculty members feel that it is easy to get access and search can be done even from remote places. However, out of a total of 680 respondents, 60.15 percent feel that it saves time to search and retrieve required documents; 22.06 percent feels that it can retrieve a specific document in the least possible time. Besides, 25.44 percent feel that it facilitates multimodal search options, and 45.59 percent feel that it is easy to get access and search can be done even from remote places.

4.5.3 Opinions of the users regarding the retrieval of Bengali documents by subject

A large number of respondents feel the need for improvement of the capability of the OPACs for searching and retrieval of Bengali documents by subject. In the users' point of view, OPAC applications need to be more user-friendly for subject search. Another group of users opines that the OPAC system should incorporate the federated search facility. As far as technological development is concerned, the library users want OPAC with new features like Web 2.0, word cloud, relevance tagging, to make search more user-friendly.

5. CONCLUSION

In a nutshell, the attitude of library users toward the use of subject headings for Bengali documents in OPAC has been determined and identified. It can be inferred that some features need to be incorporated to improve the functionality of the OPACs and thereby to satisfy users' requirements. Besides, relevant instructions and aid to use controlled vocabulary need to be facilitated in OPAC. This can help the users to learn the search strategy and retrieval of documents from the catalog. The library must provide on-site personnel help and make arrangements for the user education program regularly. However, implementation of automated correction of spelling and system prompts to the users in the OPAC interface will minimize unsuccessful search results. It has been determined that the assignment of Bengali

Subject Headings in the catalog module lacks uniformity and consistency. However, the decision table for cataloging needs to be maintained properly to cope up with a local variation.

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 relevance search technique then users can retrieve their required information easily. These recommendations concur with the assertion of Halder (2015) that an effective and efficient OPAC system can be developed by consideration of the above observations and deliberations.

References

- Alzofon, S. R., & Van Pulis, N. (1984). Patterns of searching and success rates in an online public access catalog. *College & Research Libraries*, 45(2), 110-115.
- Bandyopadhyay, R. (2001). *Suchikaran*. Kolkata: Bangiya Granthagar Parisad.
- Bhattacharyay, K. (1370). *Bishay Shiranam*. Kolkata: Bangiya Sahitya Parishad.
- Guo, J., & 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. N. (2015). The use of subject headings for Bengali documents in online public access catalog (Doctoral dissertation, Jadavpur University, Kolkata, India). Retrieved from <http://hdl.handle.net/10603/199116>
- Halder, S. N., & 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*. 20(1), 32-37.
- Miller, J., & McCarthy, S. (Eds). (2010). *Sears list of subject headings*. New York: The H. W. Wilson Company.
- Mukhopadhyay, P. (2001). *Bijnan: bishay shironam*. Kolkata: Bona.
- Mukhopadhyay, P. K. (1951). *Bangla granthaborgikaran: dashamik prothay Bangla granther bargikaran paddhati*. Kolkata: Orient Book Company.
- Sridhar, M. S. (2004). Subject searching in the OPAC of a special library: problems and issues. *OCLC Systems Services*, 20(4), 183-191. Retrieved from <http://eprints.rclis.org/archive/00008849/>