

May 2019

Author-Topic Modeling of DESIDOC Journal of Library and Information Technology (2008-2017), India

Manika Lamba

Department of Library and Information Science, University of Delhi, lambamanika07@gmail.com

Margam Madhusudhan

Department of Library and Information Science, University of Delhi, madhumargam@gmail.com

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

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

Lamba, Manika and Madhusudhan, Margam, "Author-Topic Modeling of DESIDOC Journal of Library and Information Technology (2008-2017), India" (2019). *Library Philosophy and Practice (e-journal)*. 2593.
<https://digitalcommons.unl.edu/libphilprac/2593>

Author-Topic Modeling of DESIDOC Journal of Library and Information Technology (2008-2017), India

Abstract

This study presents a method to analyze textual data and applying it to the field of Library and Information Science. This paper subsumes a special case of Latent Dirichlet Allocation and Author-Topic models where each article has one unique author and each author has one unique topic. Topic Modeling Toolkit is used to perform the author-topic modeling. The study further which considers topics and their changes over time by taking into account both the word co-occurrence pattern and time. 393 full-text articles were downloaded from DESIDOC Journal of Library and Information Technology and were analyzed accordingly. 16 core topics have been identified throughout the period of ten years. These core topics can be considered as the core area of research in the journal from 2008 to 2017. This paper further identifies top five authors associated with the representative articles for each studied year. These authors can be treated as the subject-experts for the modeled topics as indicated. The results of the study can serve as a platform to determine the research trend; core areas of research; and the subject-experts related to those core areas in the field the Library and Information Science in India.

Keywords Author-Topic Modeling; DESIDOC Journal of Library and Information Technology (DJLIT); Latent Dirichlet Allocation (LDA); Information Retrieval; Information Processing and Management; Text Mining

1 Introduction

Topic modeling acts as a text mining tool to process, organize, manage and extract knowledge. It is based on probabilistic modeling and is used to discover hidden structures in large archives of documents on the basis of similar patterns of word usage in each document. It is typically used to determine the underlying “*topics*” in text documents. A “*topic*” represents the broader concept shared by a document corpus and this “*topic*” evolves with time. Modeling topics without considering time will cause major problems. Topic evolution will help to identify topics within the context and how they evolve over time.

Lately, the focus has been either on a group of highly cited authors or a sample of journal articles. This leads to skew the results by incorporating highly cited works which are not necessarily representative of the works produced and including a few articles/authors which can heavily influence the results. This defying problem can be solved by performing author-topic modeling where not only the core research topics related to a discipline is identified but also the core authors which may be treated as the subject-experts for the respective modeled topics. Further, understanding the development of the discipline and the changes in topics over time will help to recognize the discipline identity. Therefore, this study provides the groundwork of identifying the identity of Library and Information Science (LIS) discipline in India by finding the main topics in DESIDOC Journal of Library and Information Technology (DJLIT) by using Latent Dirichlet Allocation (LDA) modeling technique.

2 Related literature

Few important studies which focus on author-topic modeling are by Cagliero et al. (2018) who “address the discovery of research collaborations among multiple authors on single or multiple topics. Specifically, they exploit an exploratory data mining technique, i.e., weighted association rule mining, to analyze publication data and to discover correlations between ATM topics and combinations of authors”; and Mao et al. (2017) who “propose an approach to detect Topical Scientific Communities (TSCs) with both topology and topic features by applying machine learning techniques and network theory”. This study follows the works of Sugimoto et al. (2011) where they “identify changes in dominant topics in LIS over time, by analyzing the 3,121 doctoral dissertations completed between 1930 and 2009 at North American Library and Information Science programs using author-topic modeling”; and Figuerola et al. (2017) who “offers an overview of the bibliometric study of the domain of Library and Information Science (LIS), with the aim of giving a multidisciplinary perspective of the topical boundaries and the main areas and research tendencies. Based on a retrospective and selective search, they have obtained the bibliographical references (title and abstract) of academic production on LIS in the database LISA in the period 1978–2014. They apply latent Dirichlet allocation, in order to identify the main topics and categories in the corpus of documents analyzed” with major modifications”.

Few important studies have been reviewed with regard to the application of LDA, Topic Modeling, and Text Mining. A pioneering research classic paper on LDA by Blei et al. (2003) showed the “efficient approximate inference techniques based on vibrational methods and an EM algorithm for empirical Bayes parameter estimation. They report results in document modeling, text classification, and collaborative filtering, comparing to a mixture of unigrams model and the probabilistic LSI model”. Some of the selected articles which show the implementation of the LDA in their studies are written by Momtazi (2018); Kim and Kang (2018); Zhao et al. (2016); Guo et al. (2017); Zhang et al. (2017); Chen (2017); Lu et al. (2017); Liu et al. (2016); Zhang et al. (2018); Ma et al. (2018); Woltmann and Alkærsg (2018); Chen et al. (2016); Wang et al.(2013); Koltsova and Koltcov (2013); Nichols (2014); Yan et al. (2017); Katsurai et al. (2016); and Huet al. (2014).

Few studies which applied topic modeling in the field of Library/Museum under various sub-fields are as follows: Mehler and Waltinger (2009) applied topic modeling in library classification by presenting a topic classification model using the “Dewey Decimal Classification” (DDC); whereas Bae et al. (2014) presented their “study in threefold: First, they suggested an alternative approach to real-time big data analysis, which has become an extremely important issue. Second, they applied a topic modeling technique that is used in various research areas, including Library and Information Science (LIS). Based on this, they can confirm the utility of storytelling and time series analysis. Third, they developed a web-based system, and make the system available for the real-time discovery of topics”. Lu and Wolfram (2012) “present static and dynamic word-based

approaches using vector space modeling, as well as a topic-based approach based on latent Dirichlet allocation for mapping author research relatedness. Outcomes for the two word-based approaches and a topic-based approach for 50 prolific authors in Library and Information Science are compared with more traditional author co-citation analysis using multidimensional scaling and hierarchical cluster analysis” in contrast to Efron, Organisciak, and Fenlon (2011) who applied topic modeling in Museum in which “they propose a way to improve topic modeling in large collections by identifying documents that convey only weak topical information using the corpus from the Institute of Museum and Library Services Digital Collections and Content aggregation”.

3 Research Objectives

The following are the main research objectives of the study which are both historical and methodological in nature:

- (a) To identify the core topics of DESIDOC Journal of Library and Information Technology diachronically; and
- (b) To identify the authors associated with the representative articles.

4 Methodology

A total of 393 full-text articles were downloaded from DJLIT journal for the period 2008-2017. The articles were analyzed according to the author-topic model and were divided into the ten-time slices. The author-topic modeling technique was used to determine the top five topics, top five words, top five representative articles and the top five authors associated with those articles. In this method, for each year, 5 topics were identified. Each topic contained a probability value, that is, the likelihood that the topic identified should be associated with the year. These topics were ranked by descending probability values and the top five were selected as being most representative of that year. Similarly, a probability for each word was calculated to represent the association between a word and the given topic and the top 5 words were chosen as the most representative of the topic. Lastly, the authors were assigned probability values for each topic and these too were ranked. The top five authors were chosen as highly representative authors for the given topic. In this approach, the hyperparameter α (prior parameter determining the topic distribution per document) and hyper-parameter β (prior parameter determining the word distribution per topic) are fixed at $5/T$ (where the T = number of topics taken in the study, i.e. 5 for the present study) and 0.01 respectively for each year as default. Fixing the hyper-parameters maintain the uniformity and consistency of results for each year. The author-topic modeling approach taken in this study is similar to that of Sugimoto et al. (2011) with major modifications, where they did topic modeling over 3,121 North American LIS doctoral Dissertations completed between 1930 and 2009 but this paper is restricted to full-text articles published in DJLIT journal from the year 2008 to 2017 excluding the Guest editorials.

4.1 DESIDOC Journal of Library and Information Technology

“DESIDOC Journal of Library and Information Technology (DJLIT) was started in the year 1981. It is a peer-reviewed open access bi-monthly journal that publishes original research and review papers related to library science and IT applied to library activities, services, and products. It is meant for librarians, documentation and information professionals, researchers, students and others interested in the field. DJLIT is a prestigious journal of LIS studies in India with 0.364 SJR (SCImago Journal Rank); 0.600 SNIP (Source Normalized Impact per Publication), and 0.47 Cite Score. It is abstracted and indexed in Scopus, LISA, LISTA, EBSCO Abstracts/Full-text, Library Literature, and Information Science Index/Full-text, The Informed Librarian Online, Open J-Gate, Indian Science Abstracts, Indian Citation Index, Full-text Sources Online, WorldCat, ProQuest, Google Scholar, Ulrich's International Periodical Directory, Index Copernicus, and OCLC” (DESIDOC Journal of Library and Information Technology 2018).

4.2 Latent Dirichlet Allocation (LDA)

This paper focuses on the use of Latent Dirichlet Allocation (Blei et al, 2003), which is based on Dirichlet distribution to model the topics from the corpus of LIS articles. In this study, each article gets represented as a pattern of LDA topics making every article appear. LDA automatically infers the topic discussed in a collection of articles and these topics can be used to summarize and organize the articles. LDA is based on probabilistic modeling and the observed variables are the bags of words per article whereas hidden random variables are the topic distribution per article. “The main goal of LDA is to compute the posterior of the hidden variables given the value of the observed variables” (Allahyari et al., 2017). The assumptions of LDA for the study are (i) articles with similar topics will use similar groups of words, (ii) articles are a probability distribution over latent topics, and (iii) Topics are probability distributions over words (Fig.1).

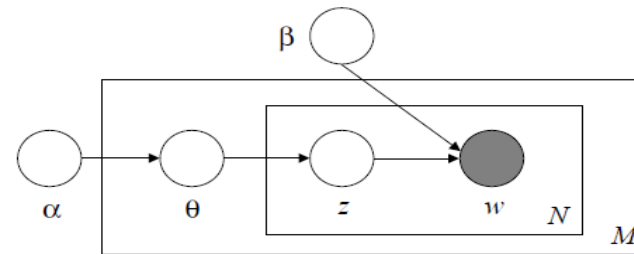


Fig.1: Graphical model representation of Latent Dirichlet Allocation
 (Source: Blei et al., 2003)

Fig.1 demonstrates the functioning of LDA where the outer box represents documents, while the inner box represents the repeated choice of topics and words within a document. The variables shown in the figure are defined as follows (Blei et al., 2003):

- α – parameter of Dirichlet prior on the per-document topic distribution
- β – parameter of Dirichlet prior on per-topic word distribution
- θ – topic distribution for the document, d
- z – atopic for the n^{th} word in the document, d
- w – is the specific word
- N – total number of words
- M – total number of documents in the corpus

4.3 Author-Topic Model

The LDA model mentioned by Blei et al. (2003) is “extended to what is called the author-topic model” (Rosen-Zvi et al. 2010). In this model, each document can be viewed as a mixture of probabilistic topics and authors. In this study, each author is associated with a single article. The author-topic model allows not only examine which topics are most salient across the various time period but also which authors are most associated with these topics.

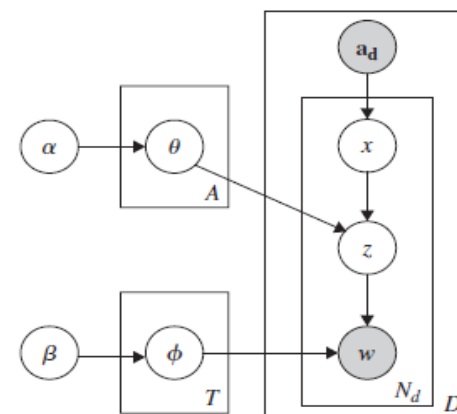


Fig. 2: Author-topic model based on Latent Dirichlet Allocation
 (Source: Rosen-Zvi et al., 2010)

Fig.2 explains the functioning of the model, “when author model as a group of authors, a_d , decide to write the document D . For each word in the document an author is chosen uniformly at random. Then, as in the topic model, a topic is chosen from a distribution over topics specific to that author, and the word is generated from the chosen topic. As in the author model, x indicates the author responsible for a given word, chosen from a_d . Each author is associated with a distribution over topics, θ , chosen from a symmetric Dirichlet (α) prior. The mixture weights corresponding to the chosen authors are used to select a topic z , and a word is generated according to the distribution φ corresponding to that topic, drawn from a symmetric Dirichlet (β) prior. The author-topic model subsumes the two models described above as special cases: topic models like LDA correspond to the case where each document has one unique author, and the author model corresponds to the case, where each author has one unique topic” (Rosen-Zvi et al., 2010).

4.4 Topic-Modeling-Toolkit (TMT)

Topic-Modeling-Toolkit (Google code archive, 2011a) is powered by Java, a graphical interface tool for LDA topic modeling. “It is a simple GUI-based application for topic modeling that uses the popular MALLET toolkit for the back-end” (Abinaya and Winster, 2014). “Topic models provide a simple way to analyze large volumes of unlabeled text. A ‘topic’ consists of a cluster of words that frequently occur together. Using contextual clues, topic models can connect words with similar meanings and distinguish between uses of words with multiple meanings. The GUI has two main windows - Basic and Advanced” (Google code archive, 2011b). All the 393 full-text articles are first converted into text format and processed using Topic-Modeling-Toolkit. A total of 10 hours were spent on the aforesaid process. In the toolkit, following parameters were being fixed for the study: (i) Number of topics: 5, (ii) Number of iterations: 200, (iii) Number of topic words printed: 5, and (iv) Topic proportion threshold: 0.05.

5 Results and Analysis

On the basis of the output files generated by the TMT, a comprehensive analysis has been performed in a chronological order in a yearly manner for the studied period. The files were available in both CSV and HTML format.

5.1 Topic Analysis

Tables 1 and 2 summarize the LDA result of the DJLIT articles. Table 1 shows the labeling of the topics, a through e for each year, and are organized in descending order according to their probability values (where a having the highest probability value). It summarizes the core topics for the corresponding years. Table 2 lists the word co-occurrence pattern over the time and summarizes the top 5 words or the high loading keywords, ranked by the probability value for each year in the descending order.

Table 1: Extended Latent Dirichlet Allocation Topic Result for Corresponding Year (2008-2017)

<i>Labels</i>	<i>2008</i>	<i>2009</i>	<i>2010</i>	<i>2011</i>	<i>2012</i>	<i>2013</i>	<i>2014</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
<i>Topic a</i>	Bibliometrics	Academic Libraries	Library Consortia	User Studies	Library Resources	Bibliometrics	Journal Evaluation	User Studies	Digital Libraries	Bibliometrics
<i>Topic b</i>	Open Access	Bibliometrics	Bibliometrics	Knowledge Management	Library Websites	Citation Analysis	Library Resources	MOOCs	Library Software	User Studies
<i>Topic c</i>	Information Literacy	Digital Libraries	User Studies	Bibliometrics	Digital Libraries	Digital Libraries	Library Services	Bibliometrics	Library Services	Academic Libraries
<i>Topic d</i>	Digital Libraries	E-Learning	Information Literacy	OPAC	Bibliometrics	Library Resources	Bibliometrics	Digital Libraries	Knowledge Management	Library Services
<i>Topic e</i>	Knowledge Management	User Studies	Library Services	Library Services	Library Services	Library Services	Knowledge Management	Library Resources	Library Resources	Mobile technology in Libraries

Table 2: Extended Latent Dirichlet Allocation Word Result for Corresponding Year (2008-2017)

<i>Labels</i>	<i>2008</i>	<i>2009</i>	<i>2010</i>	<i>2011</i>	<i>2012</i>	<i>2013</i>	<i>2014</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
<i>Topic a</i>	Research Cent Science India Journals	University Information Students Data Universities	Resources Consortium University Institutions Universities	Resources Students Cent Web Internet	Library Resources Technology Web Universities	Articles University Journal Study Total	Journals Journal Information Subject Law	Library Users User NIT Study	Search Web Citation Information Online	Research Science University Journals Papers
<i>Topic b</i>	Access OA Open Journals Research	Research Cent Papers India Services	Publications Research Indian University Cent	Library Data Libraries Knowledge Learning	Library Websites Links Http Website	Research Science Papers Citations Publications	Students Resources Information Web Internet	Http Books Subjects Open WWW	Library Software University Records Quality	Library Information Students Respondents Resources
<i>Topic c</i>	Information Library IL Learning Libraries	Digital Web Access Management System	Cent Study Table Journals Respondents	Research Papers Publications University Cent	Information Library Libraries Users Online	Library Search Users Digital Access	Library Users Services Study University	Research Papers Journal Publications University	Library Information Services Service Patent	University Web Knowledge Books Universities
<i>Topic d</i>	Virtual Online Exhibitions Digital Library	Learning Library Information Libraries Resources	Information Internet Data Learning Knowledge	Cent OPAC Search Social Documents	Research Journals Papers Journal Science	Library Journals School Resources Information	Research India Papers AL University	Information Students Social Online Metadata	Knowledge Digital Management India Ontology	Libraries Library Data Technology Research
<i>Topic e</i>	Knowledge Web Based System User	Library Information Databases Work Study	Library Libraries Users Services Technology	Library Information Libraries Services Technology	Cloud Services Libraries Services Computing	Information Services Impact Libraries Research	Information Data Libraries Management Knowledge	Library Libraries University Information Resources	Resources Library Books Journals Students	Mobile Learning Literacy Devices Information

50 articles for the year 2008 were processed. The evidence from high-loading keywords and most representative articles (Appendix-A) show that *Topic a* is about *bibliometrics* with a special emphasis to scientometrics in the geographical region particular to India. *Topic b* is about *open access* emphasizing that the research resources like journals be accessible to all. High-loading articles in *Topic c* show a focus on *information literacy* in libraries. *Topic d* displays a focus on *digital libraries* with a focus on virtual and online exhibitions. The representative articles from *Topic e* is on *knowledge management* using the web architecture for its users.

47 articles for the year 2009 were processed and reveals the dominance of *academic libraries* in *Topic a* (Appendix-A) which includes collecting data and information about the students/universities. *Topic b* indicates an interest in *bibliometrics* with an emphasis on papers published in India. *Topic c* is on *digital libraries*, with a focus on the digital access management. Research articles and keywords for *Topic d* indicate a focus on *e-learning* in libraries. *Topic e* is focused on *user studies*, particularly as a case study with respect to certain libraries.

36 articles for the year 2010 were processed and presents the keywords for *Topic a* reveals an interest in *library consortia* in particular reference to resource sharing among the libraries and institutions. A review of the representative research articles (Appendix-A) exposed *Topic b* as *bibliometrics* which is dominated by words that reflected research output in particular to Indian universities. *Topic c* has an interesting split between the keywords and representative research articles. The keywords focused on *user studies*, however, the articles share a focus on a mixture of both *bibliometrics* and *user studies* topic. Interestingly, *information literacy* dominates in *Topic d*. *Topic e* is focused on *library services*, with a particular emphasis on users, services, and technology of the libraries.

35 articles for the year 2011 were processed and shows *user studies* topic with a focus on students and the internet in *Topic a*. Representative research articles in *Topic b* (Appendix-A) suggest *knowledge management* in libraries. *Topic c* reveals an interest in *bibliometrics*. The keywords for *Topic d* indicate *OPAC* with an emphasis on searching documents. *Topic e* indicates the *library services* topic with emphasis on technology in libraries.

30 articles for the year 2012 were processed and indicates a focus on *library resources* with a special emphasis on web technology and their application in universities in *Topic a*. Research articles in *Topic b* focus primarily on *library websites* (Appendix-A). *Topic c* is an interesting split between the keywords and representative articles: the keywords focused on *digital libraries* topic and research articles share a focus on both *information literacy* and *online information*. The focus of *Topic d* is on *bibliometrics* with an emphasis on scientometrics. *Topic e* is on *library services* which focus on the application of cloud computing in libraries.

33 articles for the year 2013 were processed and reveals that *Topic a* is about *bibliometrics* with a focus on university libraries. *Topic b* is about *citation analysis* which is a method of bibliometrics. High-loading articles in *Topic c* show a focus on *digital libraries* with an emphasis on the library search, digital access and users (Appendix-A). *Topic d* displays a focus on *library resources* with an emphasis on journals, receiving information and school libraries. The representative articles from *Topic e* is on *library services* and their impact on libraries and research.

40 articles for the year 2014 were processed and shows that *journal evaluation* dominates in *Topic a* (Appendix-A). *Topic b* indicates an interest in *library resources* on Internet related to university students. *Topic c* is on *library services* with a focus on students as users and the services provided by the university libraries. Research articles and keywords for *Topic d* indicate *bibliometrics* with a focus on research and academic libraries, particularly to India. *Topic e* is focused on *knowledge management* in libraries.

36 articles for the year 2015 were processed and reveals *user studies* as *Topic a*. A review of the representative research articles (Appendix-A) shows that *Topic b* is dominated by words reflecting on *MOOCs* which is an open platform based on HTTP and is used for e-learning. *Topic c* is about *bibliometrics* with a focus on universities. *Digital libraries* dominate *Topic d* with a focus on metadata. *Topic e* is focused on *library resources* in particular to academic libraries.

36 articles for the year 2016 were processed and presents that *Topic a* is related to *digital libraries* based on the keywords and research articles (Appendix-A). Representative research articles in *Topic b* suggest *library software* as a topic with a focus on their quality and their usage in academic libraries. *Topic c* reveals an interest in *library services*, with an emphasis on patents in libraries. Keywords for *Topic d* indicates *knowledge management* in libraries as a topic with an emphasis on ontologies restricted to Indian region. *Topic e* indicates *library resources* topic with an emphasis on books and journals for the students.

50 articles for the year 2017 were processed and the emphasis that the research articles in *Topic a* indicate *bibliometrics* with an emphasis on scientometrics. Research articles in *Topic b* focus primarily on *user studies* (Appendix-A), particularly students being as respondents when querying about the library resources. *Topic c* is about *academic libraries*. The focus of *Topic d* is on *library services*. *Topic e* is on *mobile technology in libraries*.

5.2 Automated Author Analysis

Table-III summarizes the LDA result generated by TMT for each year from 2008 to 2017. It presents the top 5 authors associated with the corresponding articles in Appendix-A ranked by the probability value for each year in the descending order. It is important to consider that if same authors are coming twice under the same topic, then those authors have contributed different research articles whereas if same authors are coming under two or more topics then same research article is composed of two or more topics. This observation demonstrates that a research article can be composed of a mixture of topics.

Table 3: Extended Latent Dirichlet Allocation Author Result for Corresponding Topic a-e (2008-2017)

Labels	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
<i>Topic a</i>	1. Ganesh Surwase, B.S. Kademani and Vijai Kumar	1. B. Ramesh Babu, A.M. Narendra Kumar and S. Gopalakrishnan	1. Jagdish Arora and Kruti Trivedi	1. Sridevi Vaithianathan and Chennupati K. Ramaiah	1. Shailendra Kumar and Gareema Sanaman	1. Dines Ch. Maiti and Bidyarthi Dutta	1. Priyanka Vishwakarma and Bhaskar Mukherjee	1. Gareema Sanaman and Shailendra Kumar	1. Raj Kumar Bhardwaj and M. Madhusudhan	1. S.M. Dhawan, B.M. Gupta, Manmohan Singh, and Asha Rani
	2. B.M. Gupta and S.M. Dhawan	2. Mojtaba Sookhtanlo, Hamid Movahed Mohammadi, and Ahmad Rezvanfar	2. Jagdish Arora and Kruti Trivedi	2. Mahabaleshwara Rao Baikady and Mahesh V. Mudhol	2. Sarwesh Pareek and Dinesh K. Gupta	2. Samir Kumar Jalal	2. Jayashree Kanetkar	2. Anna Kaushik	2. Margam Madhusudhan	2. Sunaina Khanna, Neeraj Kumar Singh, Deepika Tewari, Harinder Singh Saini
	3. R. Sevukan and Jaideep Sharma	3. Mohamed A.M. Ghalib, Chandrashekara, M. and V.G. Talawar	3. B. Ramesh Babu, R. Jeyshankar and P. Nageswara Rao	3. Anwarul Islam and Keita Tsuji	3. S.F. Kattimani and Ramesh R. Naik	3. Prabir Kumar Das	3. Shri Ram and Nitin Paliwal	3. Vijayakumar Mallappa and Manoj Kumar K.S	3. Manoj Kumar Verma and Ksh. Krishna Devi	3. Vijayakumar M. and Shankar Reddy Koll
	4. G. Lalitha Kumari	4. Shyama Rajaram	4. Mukesh, G.S. Vatsa, Nikita Joshi, and Sumit Goswami	4. S. Thanuskodi	4. Sunil Tyagi	4. K.G. Pillai Sudhier	4. Priya Rai	4. Rakesh Mohindra and Anil Kumar	4. Shri Ram and Nitin Paliwal	4. Banalata Pradhan and D.B. Ramesh
	5. Subbiah Arunachalam	5. Praveen Sharma	5. R. Guruprasad and Khaiser Nikam	5. Nishat Fatima, Naved Ahmad, and Shadab Ahmad	5. Sandip Majumdar	5. Nagesh Laxman Londhe and Neela J. Deshpande	5. S.M. Pujar, G. Mahesh and Francis Jayakanth	5. Somipam R. Shimray and Chennupati K. Ramaiah	5. Margam Madhusudhan	5. Ramesh Pandita and Shivendra Singh

Topic b	<ol style="list-style-type: none"> 1. K. Nageswara Rao 2. Peter Suber 3. Peter Suber 4. Alma Swan 5. S. Shashi Nath, Sridhara B., C.M. Joshi, and Puneet Kumar 	<ol style="list-style-type: none"> 1. Priya Girap, Ganesh Surwase, Anil Sagar, B.S. Kademani and Vijai Kumar 2. B.M. Gupta, P.R. Bose, and Avinash Kshitij 3. B.M. Gupta, S.M. Dhawan, and Ugrasen Singh 4. B.S. Kademani, Ganesh Surwase, Lalit Mohan, and Vijai Kumar 5. Neerja Verma and Rajnish Tamrakar 	<ol style="list-style-type: none"> 1. B.M. Gupta, Avinash Kshitij, and Yogendra Singh 2. Lalit Mohan, E.R. Prakasan, B.S. Kademani, Ganesh Surwase, Anil Kumar, and Vijai Kumar 3. Adarsh Bala and B.M.Gupta 4. V. Kalaiappan, K. Kaliyaperumal, and V. Rajasekar 5. B. Ramesh Babu , R. Jeysankar and P. Nageswara Rao 	<ol style="list-style-type: none"> 1. Hanif Bhatt 2. Anil Kumar Dhiman 3. Shailendra K. and Namrata Rai 4. R.K. Bhatt 5. Phuldeep Kumar and Umesh Gulla 	<ol style="list-style-type: none"> 1. Vijayakumar M., B.U. Kannappanavar, and Santosh Kumar K.T 2. Kaushal Chauhan 3. Bhupendra Ratha, Leena Joshi, and G.H.S. Naidu 4. Paramjeet K. Walia and Monika Gupta 5. Sarwesh Pareek and Dinesh K. Gupta 	<ol style="list-style-type: none"> 1. B.M. Gupta 2. B.M. Gupta, B.D. Kumbar, and Ritu Gupta 3. K.R. Mulla and P.V. Konnur 4. K.C. Garg and Suresh Kumar 5. J.N. Dash, Chinmaya Rout, and Baman Parida 	<ol style="list-style-type: none"> 1. Leon James 2. Mohamed Haneefa K. and Syamili C. 3. Daulat Jotwani 4. Ngozi Blessing Ukachi, Uloma D. Onuoha, and Victor N. Nwachukwu 5. Md. Sohail and Andleeb Alvi 	<ol style="list-style-type: none"> 1. Jayashree Kanetkar 2. Subhash Khode and Sunil Singh Chandel 3. Faizul Nisha and V. Senthil 4. Fayaz Ahmad Loan and Refhat-un-Nisa 5. P. Ramesh, J. Vivekavardhan, and K. Bharathi 	<ol style="list-style-type: none"> 1. N.K. Dash and P. Padhi 2. Chandrappa and N.S. Harinarayana 3. Chidi D. Segun Adeniran, Oyeronke Adebayo, Victoria O. Itsekor and Happiness C. Michael-Onuoha 4. Mercy Ariomerebi Iroaganachi, Oluwole Durodolu, and Tosan J. Omatseye 5. Maneesh Kumar Bajpai 	<ol style="list-style-type: none"> 1. Y. Sachin, K. Divyananda 2. Prahalad G Tadasad, Deepthi K, Shobha Patil 3. Archita Nanda 4. Anasuya V. 5. Md Sohail, Shakil Ahmad
Topic c	<ol style="list-style-type: none"> 1. C.R. Karisiddappa and Iqbalahmad U. Rajgoli 2. B. Ramesh Babu 3. Intan Azura Mokhtar and Shaheen Majid 4. Michael B. Eisenberg 5. Pradeepa Wijetunge 	<ol style="list-style-type: none"> 1. Usha Mujoo Munshi 2. Lim Chee Siang Edmund, Chennupati K. Ramaiah and Surya Prakash Gulla 3. Md. Zahid Hossain Shoeb 4. M. Natarajan and Gayas Makhdumi 5. R.Balakrishna, U. Rajeswara Rao, and N. Geethanjali 	<ol style="list-style-type: none"> 1. K.G. Sudhier 2. S. Thanuskodi 3. Mohamed K. Haneefa and C.K. Abdul Shukkoor 4. Arup Kumar Mondal and Amit Kumar Bandyopadhyay 5. V. Kalaiappan, K. Kaliyaperumal, and V. Rajasekar 	<ol style="list-style-type: none"> 1. B.M. Gupta, Har Kaur, and Adarsh Bala 2. B.M. Gupta, Adarsh Bala, and Nandini Sharma 3. Anil Sagar and B.S. Kademani 4. KP Singh and Nitu Singh 5. Seema Vasishtha 	<ol style="list-style-type: none"> 1. Neena Singh and Andreas Klingenberg 2. S. Thanuskodi 3. Sunil Tyagi 4. Projes Roy, Shailendra Kumar, and M.P. Satija 5. O.M. Olaniyi, A. Omotosho, E.A. Oluwatosin, O.K. Towolawi, and G.C. Grant-Ezeronye 	<ol style="list-style-type: none"> 1. Shiv Kumar 2. Raj Kumar Bhardwaj 3. Reeta Sharma and Shantanu Ganguly 4. Anil Kumar Singh and Sumati Sharma 5. B.S. Biradar and Koteppa Banateppanavar 	<ol style="list-style-type: none"> 1. B. Sriram and M.K.G. Rajev 2. Mahabaleshwara Rao Baikady and Jessy A. 3. Imran Khan 4. Ngozi Blessing Ukachi, Uloma D. Onuoha, and Victor N. Nwachukwu 5. Tanveer Haider Naqvi 	<ol style="list-style-type: none"> 1. H. Anil Kumar, Mallikarjun Dora and Asha Desai 2. D. Sujatha and K. Padmini 3. Anil Kumar Siwach and Satish Kumar 4. Vinod Kumar Gautam and Rajani Mishra 5. S. Gopalakrishnan , S. Gopalakrishnan , A.L. Bathrinarayanan and M. Tamizhchelvan 	<ol style="list-style-type: none"> 1. Nidhi Sandal and Avinash Kumar 2. Jerome Idiegbeyan-ose, Goodluck Ifijeh, Chidi Segun-Adeniran, Michael Fagbohun and Ugwunwa Esse 3. Anand M. Dodamani and Tularam Brahma 4. N.K. Dash and P. Padhi 5. Alka Bansal 	<ol style="list-style-type: none"> 1. Manoj Kumar Verma and Krishna Brahma 2. Vanita Khanchandani and Moveen Kumar 3. K. Kumar 4. Sanjeev K Sunny and Mallikarjun Angadi 5. Naushad Ali P.M., Daud Khan
Topic d	<ol style="list-style-type: none"> 1. Leong Chee Khoon and Chennupati K. Ramaiah 2. Cèsar Carreras Monfort 3. R. V. Roberto 4. Wen-Hsi Chang 5. M.P. Satija 	<ol style="list-style-type: none"> 1. Vivek Patkar 2. H.S. Siddamallaiah and Sujin Butdisuwan 3. S.M. Pujar and R.K. Kamat 4. Pratibha A. Gokhale and Smita Chandra 5. Shantanu Ganguly 	<ol style="list-style-type: none"> 1. Indira Koneru 2. Partha Sarathi Das and Bhupendra Singh Kholia 3. Manjunath B. Hadimani and Iqbalahmad U. Rajgoli 4. G.C. Mohanta and K.P. Thooyamani 5. Sumit Goswami, Anandarup Mukherjee, Mansi Kharbanda, Abhinav Gupta and Pulkit Soni 	<ol style="list-style-type: none"> 1. K. Nageswara Rao and V.G. Talwar 2. K. Nageswara Rao and V.G. Talwar 3. Shiv Kumar and Ranjana Vohra 4. Bhaskar Mukherjee and Mohammad Nazim 5. Mohamed Haneefa K. and Sumitha E. 	<ol style="list-style-type: none"> 1. B.M. Gupta, Adarsh Bala and Avinash Kshitig 2. H. Anil Kumar and Mallikarjun Dora 3. Rajendra Kumbhar 4. Har Kaur and Preeti Mahajan 5. Anna Kaushik and Jagdish Arora 	<ol style="list-style-type: none"> 1. Chennupati K. Ramaiah and K. Nallaperumal Pillai 2. K. Kumar 3. Geetha M., Mamatha K.R. and Farhana 4. Shabahat Husain and Mohammad Nazim 5. P.K. Suresh Kumar 	<ol style="list-style-type: none"> 1. Ritu Gupta, B.M. Gupta, Avinash Kshitij and Adarsh Bala 2. Akhtar Hussain and Parvez Ahmad 3. A.Gopikuttan and Aswathy S. 4. Ritu Gupta and B.M. Gupta 5. Abhijit Chatterjee and Arabinda Maity 	<ol style="list-style-type: none"> 1. P. Ramesh, J. Vivekavardhan, and K. Bharathi 2. Shri Ram 3. Rabishankar Giri, B.K. Sen and G. Mahesh 4. Jayadev H. Kadli and Veeresh B. Hanchinal 5. Somipam R. Shimray and Chennupati K. Ramaiah 	<ol style="list-style-type: none"> 1. Abhishek Sharma 2. Ashok Kumar, Ritu Gupta, and B.M. Gupta 3. S. Dhanavandan and M. Tamizhchelvan 4. K.P.B. Moosad, Letha M.M. and S. Anantha Narayanan 5. S.Sudarshan Rao 	<ol style="list-style-type: none"> 1. Manorama Tripathi, Archana Shukla and Sharad Kumar Sonker 2. Parul Gupta and Margam Madhusudhan 3. B.B. Chand and Ramesha 4. Margam Madhusudhan and Saleeq Ahmad Dar 5. Tanveer Haider Naqvi
	<ol style="list-style-type: none"> 1. Biswanath Dutta 	<ol style="list-style-type: none"> 1. K.P. Singh, Neeru Sharma, and Nandi 	<ol style="list-style-type: none"> 1. N.K. Dash and P. Padhi 	<ol style="list-style-type: none"> 1. Namrata Rai 	<ol style="list-style-type: none"> 1. S.Y. Bansode and S.M. Pujar 	<ol style="list-style-type: none"> 1. Prakasan P.M. 	<ol style="list-style-type: none"> 1. N. Madurai Meenachi and M. 	<ol style="list-style-type: none"> 1. Soumen Kayal, Somnath 	<ol style="list-style-type: none"> 1. Kulveen Kaur and Kiran Kathuria 	<ol style="list-style-type: none"> 1. Margam Madhusudhan and

<i>Topic e</i>	2. Schubert Foo	Negi	2. K.R. Mulla , M. Chandrashekara and V.G. Talawar	2. M. Krishnamurthy and H.M. Rajashekara	2. Surinder Kumar	2. S. Ramesh and T. Yuvaraj	Sai Baba	Bandopadhyay and Swapna Banerjee	2. Rajender Kumar	Saleeq Ahmad Dar
	3. K. Nageswara Rao	2. Vijayakumar M., B.U. Kannappanavar and Mamata Mestri	3. Rajendra Kumbhar	3. Mohammad Aqil, Parvez Ahmad, and Mohammad Asad Siddique	3. Shashi Prabha Singh and R.S.R. Veralakshmi	3. Malathy S. and Kantha P.	2. Rajendra Kumbhar	2. Subhash Khode and Sunil Singh Chandel	3. Nazir Ahmad Bhat and Shabir Ahmad Ganaie	2. Karunesh Kumar Arora and Shyam S. Agrawal
	4. Pai-Hsun Chen, Jen-Shin Hong, and Sheng-Wen Shih	3. T.Y. Malliah and P.S. Yadapadithaya	4. Seema Vasishta and Maninder Kaur Dhanda	4. Manir Abdullahi Kamba	4. O.M. Olaniyi, A. Omotosho, E.A. Oluwatosin, O.K. Towolawi, and G.C. Grant-Ezeronye	4. Anil Kumar Singh and Sumati Sharma	4. Tariq Ahmad Lone and Rafi Ahmad Khan	3. Iroaganachi Mercy A., Iwu James Juliana and Esse Ugwunwa C	4. Jessy A., Mahabaleshwara Rao and Shivananda Bhat K.	3. Nidhi Sandal, Prem Chand Gupta, Rakesh Kumar Sharma, Naval Kishore Sepat, Avinash Kumar
	5. M. Natarajan	4. Dhanashree A. Date and Pratibha Walavalkar	5. Shri Ram	5. R.K. Bhatt	5. S. Yadagiri and Prashanth Vidya Sagar Thalluri	5. Sunil Bhatt	5. Archana S.N., Padmakumar P.K. and Beena C.	4. Abhijit Chatterjee and Arabinda Maity	5. Prem Chand Sharma and Raj Kumar	4. Rajul Sharma and Margam Madhusudhan
		5. Adupa Suniland K. Praveen Kumar						5. Mehar Singh and Ajay Kumar Arora		5. Sheikh Mohd Imran and Basharat Ahmad Malik

6 Discussion

16 core topics have been identified in the study and these core topics for DJLIT journal from the year 2008 to 2017 are summarized in Table-IV. Further, Table-V ranks the topics from the highest number of occurrences in the journal to the lowest over the epoch. As it can be observed from Tables IV and V, *Bibliometrics* is the most researched topic in DJLIT which occurred in almost all the years except in 2016 followed by *User Studies* and the emphasis of the bibliometric method was more towards scientometrics studies in the years 2008, 2012 and 2017. The least common research topics in DJLIT are found to be *e-learning*, *journal evaluation*, *library consortia*, *OPAC and open access* which occurred only once over the 10 year period. Moreover, topics like *Library software*; *MOOCs*; and *mobile technology in libraries* are some of the novel areas of research in the context of India which are being introduced in the DJLIT in the past three years. Lastly, the modeled topics have been compared with the bibliometrics study conducted by Garg and Sharma (2017) in LIS discipline in India during 2004-2015. The findings from the topic analysis (Tables IV and V) in the present study somehow match the sub-disciplines in LIS founded by Garg and Sharma (2017).

Table 4: Summary of Topic Analysis

S.No.	Topics	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
1.	Academic Libraries		√								√
2.	Bibliometrics	√	√	√	√	√	√	√	√		√
3.	Digital Libraries	√	√			√				√	
4.	E-Learning		√								
5.	Information Literacy	√		√							
6.	Journal Evaluation							√			
7.	Knowledge Management	√			√			√		√	
8.	Library Consortia			√							

9.	Library Resources							√		√	
10.	Library Services			√	√			√		√	√
11.	Library Software									√	
12.	Mobile Technology in Libraries										√
13.	MOOCs								√		
14.	OPAC				√						
15.	Open Access	√									
16.	User Studies		√	√	√				√		√

Table 5: Topic Ranking in DESIDOC Journal of Library and Information Technology (2008-2017)

S.No.	Modeled Topic	Number of Occurrences during 2008-2017
1	Bibliometrics	9
2	User Studies	5
3	Digital Libraries	4
4	Knowledge Management	4
5	Library Services	4
6	Academic Libraries	2
7	Information Literacy	2
8	Library Resources	2
9	E-Learning	1
10	Journal Evaluation	1
11	Library Consortia	1
12	Library Software	1
16	Mobile Technologies in Library	1
13	MOOCs	1
14	OPAC	1
15	Open Access	1

The present study determines not only the major research areas of DJLIT articles for each year but also the authors contributing to the work. As pointed by Lin et al. (2013), “expert finding is of vital importance for exploring scientific collaborations to increase productivity by sharing and transferring knowledge within and across different research areas”. Therefore, Table-6 indicates the highly-cited authors from the bibliometrics study of DJLIT journal for the period of 2011-2015 which has been compared with Tables I and III of our study to determine if the highly-cited authors of DJLIT journal are domain-experts or not. As it can be observed, out of the 17 highly-cited authors, 7 authors from the Bapte (2017) study are not present as the subject-experts in our study. This shows how the results of highly-cited authors and works can overshadow the real domain-experts in a field.

Table 6: Identification of Domain for Highly-Cited Authors in 2011-2015 by Bapte (2017)

S.No.	Highly-Cited Author	Domain Expert for Modeled Topic
1	Gupta, B.M.	Bibliometrics
2	Garg. K.C	Citation Analysis
3	Kademani, B.S.	Bibliometrics
4	Rousseau, R.	*
5	Ramaiah, C.K.	Digital Libraries and Library Resources
6	Prathap, G	*
7	Egghe, L	*
8	Sen, B.K.	Digital Libraries
9	Dhawan, S.M	Bibliometrics
10	Kalyane, V.L.	*
11	Satija, M.P.	Digital Libraries
12	Padhi, P	Library Services
13	Glanzel, W.	*
14	Schubert, A	*
15	Foo, S.	Knowledge Management
16	Lynch, C.	*
17	Sagar, Anil	Bibliometrics

The value added-feature of this study is that various librarians, stakeholders, publishers, and institutions from all around the world can recognize the Indian authors from Table-3 as subject-experts and domain-creators for their respective domain of research in LIS in India. These subject-experts can also be considered for numerous national or international collaborative projects by potential collaborators and institutions. Further, editors can acknowledge these subject-experts as potential reviewers in the field of LIS for their respective research topic. Recognizing experts along a mutual research domain can lead to better networking between the experts and researchers all around the world. Moreover, new professionals; current graduate students; and research scholars working in the LIS field can reach out to these subject-experts to receive guidance in their career. Lastly, the recognition of top-ranking authors can further help the experts to do well in their own professional career by getting more citations for their representative articles (Appendix-A) and making a strong stance in their professional society.

One of the major limitations of this study would be its small sample size but this small data size only made it possible to closely review the top core research areas, and the subject-experts related to those core areas for DJLIT on a yearly basis for the studied period. The methodological limitations of the study will include the prior identification of an appropriate number of topics for the articles before performing the Latent Dirichlet Allocation and the incompetence of the Dirichlet topic distribution to correlate among topics and manual interpretation and labeling of ‘topics’. Although some topics were fairly straightforward to label (e.g., *Topic b* of 2008, the top three loading words of which were (a) access, (b) OA, and (c) open), others proved more difficult to ascertain the content or methodological relationship that connected the words and the research articles. This work has broad application to those who want to know the research areas in LIS education which are highly researched and which are under-researched in India. From a methodological standpoint, this work may hold interest for biometricians/ scientometricians interested in new techniques for evaluating journals and articles in India.

7 Conclusion

In this study, 393 full-text articles from DESIDOC Journal of Library and Information Technology were analyzed according to topic-author modeling to identify the 16 core topics throughout the period of ten years. We also recognized the highly researched and under-researched areas published in DJLIT over the epoch. Thus, the current study provides a complementary lens to previous metric studies popular in India by exploring a form of analysis which can overcome the problem of skewing of results by highly cited works. Further, the top five subject-experts for the modeled topics have been identified with their representative articles. Such ranking can

encourage experts to compete with each other to do better quality research to maintain their rank in the top five author list and ultimately help to improve the quality of research in LIS in India. This study can be extended to perform prediction analysis by labeling the articles with the modeled topics and classify them to predict the topics of unlabeled articles in DJLIT.

REFERENCES

- Abinaya, G. and Winster, S.G. (2014), "Event identification in social media through latent dirichlet allocation and named entity recognition", in: *Proceedings of IEEE International Conference on Computer Communication and Systems ICCCS14*. pp 142–146
- Bae, J-H. Han, N-G., and Song, M. (2014), "Twitter Issue Tracking System by Topic Modeling Techniques", *Journal of Intelligence and Information Systems*, Vol. 20, pp. 109–122. doi: 10.13088/jiis.2014.20.2.109
- Bapte, V.D. (2017), "DESIDOC Journal of Library and Information Technology (DJLIT): A Bibliometric Analysis of Cited References", *DESIDOC Journal of Library & Information Technology*, Vol. 37, pp. 264–269.
- Blei, D.M. and Lafferty, J.D. (2006), "Dynamic Topic Models", in: *Proceedings of the 23rd International Conference on Machine Learning. ACM*, New York, NY, USA, pp 113–120.
- Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003), "Latent dirichlet allocation", *Journal of machine Learning research*, Vol. 3, pp. 993–1022.
- Cagliero, L., Garza, P., Kavoosifar, M.R., Baralis, E. (2018), "Discovering cross-topic collaborations among researchers by exploiting weighted association rules", *Scientometrics*, Vol. 116, pp. 1273–1301. doi: 10.1007/s11192-018-2737-3
- Chen, J., Wang, T.T., and Lu, Q. (2016), "THC-DAT: a document analysis tool based on topic hierarchy and context information", *Library Hi Tech*, Vol. 34, pp. 64–86. doi: 10.1108/LHT-07-2015-0074
- Chen, L-C. (2017), "An effective LDA-based time topic model to improve blog search performance", *Information Processing & Management*, Vol. 53, pp. 1299–1319. doi: 10.1016/j.ipm.2017.08.001
- DESIDOC Journal of Library & Information Technology. <http://publications.drdo.gov.in/ojs/index.php/djlit/index>. Accessed 30 Oct 2018
- Efron, M., Organisciak, P., and Fenlon, K. (2011), "Building topic models in a federated digital library through selective document exclusion", *Proceedings of the Association for Information Science and Technology*, Vol. 48, pp. 1–10.
- Figuerola, C.G., García, Marco F.J., and Pinto, M. (2017), "Mapping the evolution of library and information science (1978–2014) using topic modeling on LISA", *Scientometrics*, Vol. 112, pp. 1507–1535. doi: 10.1007/s11192-017-2432-9
- Garg, K.C., and Sharma, C. (2017), "Bibliometrics of Library and Information Science research in India during 2004-2015", *DESIDOC JOURNAL OF LIBRARY & INFORMATION TECHNOLOGY*, Vol. 37, pp. 221–227. doi: 10.14429/djlit.37.3.11188
- Google Code Archive - Long-term storage for Google Code Project Hosting. (2011a) <https://code.google.com/archive/p/topic-modeling-tool/>. Accessed 10 Oct 2018
- Google Code Archive - Long-term storage for Google Code Project Hosting. (2011b) <https://code.google.com/archive/p/topic-modeling-tool/wikis/TopicModelingTool.wiki>. Accessed 10 Oct 2018
- Guo, Y., Barnes, S.J., and Jia, Q. (2017), "Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation", *Tourism Management*, Vol. 59, pp. 467–483. doi: 10.1016/j.tourman.2016.09.009
- Hu, Z., Fang, S., and Liang, T. (2014), "Empirical study of constructing a knowledge organization system of patent documents using topic modeling", *Scientometrics*, Vol.100, pp. 787–799. doi: 10.1007/s11192-014-1328-1
- Katsurai, M., Ohmukai, I., and Takeda, H. (2016), "Topic Representation of Researchers' Interests in a Large-Scale Academic Database and Its Application to Author Disambiguation", *IEICE Trans Inf & Syst* E99.D:1010–1018. doi: 10.1587/transinf.2015DAP0031
- Kim, S.G., and Kang, J. (2018), "Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews", *Information Processing & Management*, Vol. 54, pp. 938–957. doi: 10.1016/j.ipm.2018.06.003
- Koltsova, O., Koltcov, S. (2013), "Mapping the public agenda with topic modeling: The case of the Russian livejournal", *Policy & Internet*, Vol. 5, pp. 207–227.
- Lee, J., Lapira, E., Bagheri, B., and Kao, H. (2013), "Recent advances and trends in predictive manufacturing systems in big data environment", *Manufacturing Letters*, Vol. 1, pp. 38–41. doi: 10.1016/j.mfglet.2013.09.005
- Lin, L., Xu, Z., Ding, Y., and Liu, X. (2013), "Finding topic-level experts in scholarly networks", *Scientometrics*, Vol. 97, pp. 797–819. doi: 10.1007/s11192-013-0988-6

- Liu, L., et al. (2016), “An overview of topic modeling and its current applications in bioinformatics”, doi: 10.1186/s40064-016-3252-8
- Lu, K., and Wolfram, D. (2012), “Measuring author research relatedness: A comparison of word-based, topic-based, and author cocitation approaches”, *Journal of the American Society for Information Science and Technology*, Vol. 63, pp. 1973–1986. doi: 10.1002/asi.22628
- Lu, K., Cai, X., Ajiferuke, I., and Wolfram, D. (2017), “Vocabulary size and its effect on topic representation”, *Information Processing & Management*, Vol. 53, pp. 653–665. doi: 10.1016/j.ipm.2017.01.003
- Ma, T., Li, R., Ou, G., and Yue, M. (2018), “Topic based research competitiveness evaluation”, *Scientometrics*. doi: 10.1007/s11192-018-2891-7
- Mao, J., Cao, Y., Lu, K., and Li, G. (2017), “Topic scientific community in science: a combined perspective of scientific collaboration and topics”, *Scientometrics*, Vol. 112, pp. 851–875. doi: 10.1007/s11192-017-2418-7
- Mehler, A., and Waltinger, U. (2009), “Enhancing document modeling by means of open topic models: Crossing the frontier of classification schemes in digital libraries by example of the DDC”, *Library Hi Tech*, Vol. 27, pp. 520–539. doi: 10.1108/07378830911007646
- Momtazi, S. (2018), “Unsupervised Latent Dirichlet Allocation for supervised question classification”, *Information Processing & Management*, Vol. 54, pp. 380–393. doi: 10.1016/j.ipm.2018.01.001
- Nichols, L.G. (2014), “A topic model approach to measuring interdisciplinarity at the National Science Foundation”, *Scientometrics*, Vol. 100, pp. 741–754. doi: 10.1007/s11192-014-1319-2
- Noh, Y. (2015), ‘Imagining Library 4.0: Creating a Model for Future Libraries’, *The Journal of Academic Librarianship*, Vol. 41, pp. 786–797. doi: 10.1016/j.acalib.2015.08.020
- Rosen-Zvi, M., et al. (2010), “Learning author-topic models from text corpora”, *ACM Transactions on Information Systems*, Vol. 28, pp. 1–38. doi: 10.1145/1658377.1658381
- Sugimoto, C.R., et al. (2011), “The shifting sands of disciplinary development: Analyzing North American Library and Information Science dissertations using latent Dirichlet allocation.”, *Journal of the American Society for Information Science and Technology*, Vol. 62, pp. 185–204. doi: 10.1002/asi.21435
- Wang, J., et al. (2013), “Unsupervised mining of long time series based on latent topic model”, *Neurocomputing*, Vol. 103, pp. 93–103. doi: 10.1016/j.neucom.2012.09.008
- Yan, J., et al. (2017), “Towards big topic modeling”, *Information Sciences*, Vol. 390, pp. 15–31. doi: 10.1016/j.ins.2016.12.014
- Zhang, Y., et al. (2017), “iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization”, *Future Generation Computer Systems*, Vol. 66, pp. 30–35. doi: 10.1016/j.future.2015.12.001
- Zhang Y., et al. (2018), “Collective topical PageRank: a model to evaluate the topic-dependent academic impact of scientific papers”, *Scientometrics*, Vol. 114, pp. 1345–1372. doi: 10.1007/s11192-017-2626-1
- Zhao, F., Zhu, Y., Jin, H., and Yang, L.T. (2016), “A personalized hashtag recommendation approach using LDA-based topic model in microblog environment”, *Future Generation Computer Systems*, Vol. 65, pp. 196–206. doi: 10.1016/j.future.2015.10.012

Appendix A: Title Corresponding to the Representative Articles (2008-2017)

Labels	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Topic a	<ol style="list-style-type: none"> 1. Scientometric Dimensions of Neutron Scattering Research in India 2. A Scientometric Analysis of S and T Publications Output by India during 1985-2002 3. Bibliometric Analysis of Research Output of Biotechnology Faculties in Some Indian Central Universities 4. Global Access to Indian Research: Indian STM Journals Online 5. Open Access to Scientific Knowledge 	<ol style="list-style-type: none"> 1. Credibility of University Websites in Tamil Nadu 2. Library Information-Seeking Behaviour among Undergraduate Students of Agricultural Extension and Education in Iran 3. Awareness and Extent of Utilisation of Web-based Open Source e-Learning Coursewares among Educators and Students 4. Application of E-Learning in Creative Writing 5. Advanced Applications of Data Warehousing Using 3-tier Architecture 	<ol style="list-style-type: none"> 1. INDEST-AICTE Consortium: Present Services and Future Endeavours 2. UGC-INFONET Digital Library Consortium: Present Services and Future Endeavours 3. Websites of Central Universities in India: A Webometric Analysis 4. Statistical Machine Translation 5. E-Journals and their Usage Patterns Amongst the Indian Aerospace Scientists and Engineers in Bengaluru 	<ol style="list-style-type: none"> 1. Animated and Hypertext User Interfaces: A Comparative Study 2. Web as a Learning Resource at the Medical College Libraries in Coastal Karnataka: Perception of Faculty and Students 3. Evaluation of Usage of University Websites in Bangladesh 4. Usage of Electronic Resources at Dr T.P.M. Library, Madurai Kamaraj University: A Case Study 5. Use of Library Portal by Engineering and Technology Students of Aligarh Muslim University: A Survey 	<ol style="list-style-type: none"> 1. Comparative Study of Multimedia Resources in Libraries of Delhi 2. Information about Services and Information Resources on Websites of Selected Libraries in Rajasthan: A Study 3. E-Learning Technology in the ICT Era: Application to Technical Education 4. Use of Web 2.0 Technology by Library Professionals: Study of Selected Engineering Colleges in Western Uttar Pradesh 5. Web 2.0 Tools in Library Web Pages: Survey of Universities and Institutes of National Importance of West Bengal 	<ol style="list-style-type: none"> 1. Comparative Study Between Words in Titles and Keywords of Some Articles on Knowledge Organisation 2. A Comparative Weblink Analysis among Top Indian, Asian and World universities 3. Journal of Informetrics: A Bibliometric Profile 4. Lotka's Law and Pattern of Author Productivity in the Area of Physics Research 5. Usage Study of UGC-INFONET E-resources at University of Pune 	<ol style="list-style-type: none"> 1. Developing Qualitative Indicators for Journal Evaluation: Case Study of Library Science Journals of SAARC Countries 2. Development of Subject Gateways: A Status Update 3. Assessment of Bradford Law's of Scattering to Psoriasis Literature through Bibliometric Snapshot 4. Electronic Repositories for Preservation of Legal Scholarships 5. An Exploratory Analysis of Messages on a Prominent LIS Electronic Discussion List from India 	<ol style="list-style-type: none"> 1. User's Perspective Towards Assistive Technologies Available in NCR Libraries of India 2. An Evaluation of National Institutes of Technology (NITs) Library Websites 3. Conflict Management in Management Library Professionals 4. User Satisfaction Regarding Quality of Library Services of A.C. Joshi Library, Panjab University, Chandigarh 5. Design and Development of an Online Exhibition on the Tangkhul Tribe Festivals 	<ol style="list-style-type: none"> 1. Online Legal Information System (OLIS) Leveraging Access to Legal Information Resources in Indian Environment 2. Use of Online Citation Tools by Students and Scholars of Research Department of Library and Information Science, University of Delhi 3. Web Content and Design Trends of Indian Institutes of Management (IIMs) Libraries Website: An Analysis 4. Management of University Research Publication: A Case Study of JUIT Publication Database (JPubDB) 5. Development of Web Directory of Aerospace Testing Facilities in India 	<ol style="list-style-type: none"> 1. Meta-materials Research: A Scientometric Assessment of Global Publications Output during 2007-16 2. Scientometric Analysis of the Research Output of Physics and Astronomy of Guru Nanak Dev University during 2006-15 3. Indian Contribution in Information Science and Library Science Research during 1991-2015: A Bibliometric Analysis 4. Scientometrics of Engineering Research at Indian Institutes of Technology Madras and Bombay during 2006-2015 5. Doctoral Theses Awarded in Library and Information Science in India during 2010-2014: A Study
Topic b	<ol style="list-style-type: none"> 1. Application Domain and Functional Classification of Recommender Systems—A Survey 2. Open Access to Electronic Theses and Dissertations 3. Open Access and Quality 4. Open Access for Indian Scholarship 5. Intellectual Property Rights: Issues for Creation of Institutional Repository 	<ol style="list-style-type: none"> 1. Publication Productivity of the Technical Physics and Prototype Engineering Division at Bhabha Atomic Research Centre 2. Science & Technology Profile of Publications Output of India and Germany during 1996-2006: A Comparative Study 3. Social Science Research in India, China and Brazil—A Comparative Study 4. Bhabha Scattering: A Scientometric View 5. Analysis of Contributions to Defence Science Journal 	<ol style="list-style-type: none"> 1. Indian Computer Science Research Output during 1999-2008: Qualitative Analysis 2. Research Trends in Nanoscience and Nanotechnology in India 3. Research Activities in Biochemistry, Genetics and Molecular Biology during 1998-2007 in India: A Scientometric Analysis 4. Scientometric Analysis of Literature Output of Prof. G.N. Ramachandran in the Subjects of Biophysics and Crystallography 5. Websites of Central Universities in India: A Webometric Analysis 	<ol style="list-style-type: none"> 1. Students' Perception of Charging Fines for Overdue Books: Case of Islamia College of Science and Technology Library 2. Knowledge Discovery in Databases and Libraries 3. Web Interface in Library Management Software Systems 4. University Libraries in India in the Digital Age 5. Corporate e-Learning: Possibilities, Promises, and Realities 	<ol style="list-style-type: none"> 1. Webometric Analysis of Web Presence and Links of SAARC Countries 2. Selected Free E-Journals in Library and Information Science in Directory of Open Access Journals 3. Webometric Study of IIT Libraries Websites 4. Web Impact Factor of Select National Libraries' Websites 5. Information about Services and Information Resources on Websites of Selected Libraries in Rajasthan: A Study 	<ol style="list-style-type: none"> 1. Bangladesh: A Scientometric Analysis of National Publications Output in S&T, 2001-10 2. Social Science Research in India: A Scientometric Analysis of Publications (2001-10) 3. Mapping of Engineering Research Trend in Karnataka: A Special Reference to Visvesvaraya Technological University 4. Citation Impact of 'Letter to the Editor' Published by Indian Scientists in Journals Indexed by Science Citation Index Expanded (SCIE) 5. Publications Productivity of Odisha in S&T: A Quantitative Study 	<ol style="list-style-type: none"> 1. Measuring the Psychological Characteristics of Titles of Journal Articles and their Subject Headings 2. Use of Information and Communication Technology by Visually-impaired Students: A Study in University of Calicut, Kerala 3. Marketing of Electronic Resources in IIT Libraries 4. Students' Attitudes as a Determining Factor to Electronic Information Resources Use in University Libraries in South-West, Nigeria 5. Use of Web Resources by Medical Science Students of Aligarh Muslim University 	<ol style="list-style-type: none"> 1. Development of Subject Gateways: A Status Update 2. Adoption of Open Source Software in India 3. MOOCs: Changing Trend Towards Open Distance Learning with Special Reference to India 4. Open Access E-books in Science and Technology: A Case Study of Directory of Open Access Books 5. Metadata Diversity, Interoperability and Resource Discovery Issues and Challenges 	<ol style="list-style-type: none"> 1. LSQA Scale: A Tool for Measuring Users' Perceptions of Service Quality in Libraries 2. Investigation into Diligence in Metadata Records of Mysore University Library 3. Process of Migrating from Alice for Windows Software to Millennium Software: Covenant University Experience 4. Migration to Viable Platform for Effective Library Operations: Millennium Experiences of Two Academic Libraries in Nigeria 5. Impact of WhatsApp on LIS Professionals 	<ol style="list-style-type: none"> 1. Usage of Library for Accessing Clinical Information by the Students of Health Science Universities in Karnataka 2. Use of Online Social Networking Services in University Libraries: A Study of University Libraries of Karnataka, India 3. Use and Awareness of E-journals by the Faculty and Research Scholars of Veer Surendra Sai University of Technology 4. Usage of Electronic Resources by the Medical, Dental and Paramedical Science Professionals in Karnataka 5. Use of Electronic Resources and Services by Faculty Members and Students of Fiji National

										University
<i>Topic c</i>	<ol style="list-style-type: none"> 1. In Search of Information Literacy Programmes and Practices: Survey of Selected Institutions at Bangalore 2. Information Literacy—Competency Standards and Performance Indicators: An Overview 3. Information Literacy Standards, Guidelines and their Implementation : An Analysis 4. Information Literacy: Essential Skills for the Information Age 5. Improving Information Skills of the Postgraduate Students of NILIS University of Colombo: A Case Study 	<ol style="list-style-type: none"> 1. Building Subject Gateway in a Shifting Digital World 2. Electronic Medical Records Management Systems: An Overview 3. Access Management for Digital Repository Safeguarding the Digital Contents: Digital Watermarking 4. Secure Authenticated Key Exchange Protocol for Credential Services 	<ol style="list-style-type: none"> 1. Application of Bradford's Law of Scattering to the Physics Literature: A Study of Doctoral Theses Citations at the Indian Institute of Science 2. Information Needs and Use Pattern of District Court Lawyers of Salem and Erode in Tamilnadu 3. Information and Communication Technology Literacy Professionals in Calicut University, Kerala 4. Application of ICT and Related Manpower Problems in the College Libraries of Burdwan 5. Scientometric Analysis of Literature Output of Prof. G.N. Ramachandran in the Subjects of Biophysics and Crystallography 	<ol style="list-style-type: none"> 1. Mapping of Indian Diabetes Research during 1999-2008:A Scientometric Analysis of Publications Output 2. Ranking of Indian Institutions Contributing to Computer Science Research, 1999-2008 3. Growth and Impact of S&T Research in Madhya Pradesh during 2000-2009 4. Depth Schedule for Common Auxiliaries of Place in Universal Decimal Classification for Rendering Indian States and Districts 5. Assessment of Academic Research Output during 1996-2009: A Case Study of PEC University of Technology, Chandigarh 	<ol style="list-style-type: none"> 1. Information Literacy in India and Germany: University Libraries as Activators of Life-long Learning 2. Assessing the Efficacy of Library Services of District Central Libraries in Tamil Nadu from Users Perception 3. Use of Electronic Information Resources at the Indian Pharmacopoeia Commission 4. Problems in Searching Online Databases: A Case Study of Select Central University Libraries in India 5. Application of Information Communication Technology to the Management of Library's Readers' Desk 	<ol style="list-style-type: none"> 1. From Clay Tablets to Web: Journey of Library Catalogue 2. Leveraging Access to E-resources through Gateway: A Case Study at St. Stephen's College, Delhi 3. Knowledge Access using DL Platform: A TERI Research Library Case Study 4. DRDO Knowledge Repository 5. Steps for Developing Digital Repository using Dspace: An Experience of Kuvempu University, India 	<ol style="list-style-type: none"> 1. Impact of Academic Library Services on User Satisfaction: Case Study of Sur University College, Sultanate of Oman 2. Awareness and Use of Evidence-based Medicine 3. A Study at Kasturba Medical College, Manipal 4. An Evaluation of Selective Foreign e-Bookshops 5. Students' Attitudes as a Determining Factor to Electronic Information Resources Use in University Libraries in South-West, Nigeria 6. Use of Collection and Services by P.G. Students and Research Scholars in GBPUAT Library, India 	<ol style="list-style-type: none"> 1. A Bibliometrics Profile of Gujarat University, Ahmedabad during 2004-2013 2. IEEE Transactions on Antennas and Propagation: A Bibliometric Study 3. Bibliometric Analysis of Research Publications of Maharshi Dayanand University (Rohtak) During 2000-2013 4. Scholarly Research Trend of Banaras Hindu University During 2004-2013: A Scientometric Study Based on Indian Citation Index 5. Uncited Publications in MEMS Literature: A Bibliometric Study 	<ol style="list-style-type: none"> 1. Searching and Analysing Patent Document to Solve R&D Problems 2. Library Professionals and Social Network Sites: Use, Relevance and Challenges from University Libraries in Ogun State, Nigeria 3. Users' Perceptions of Library Services in Tata Institute of Social Sciences, Mumbai 4. LSQA Scale: A Tool for Measuring Users' Perceptions of Service Quality in Libraries 5. Creating Awareness of Publications through Marketing: Case Study of DESIDOC Journal of Library & Information Technology (DJLIT) 	<ol style="list-style-type: none"> 1. Websites of Central Universities in North East India: A Webometric Analysis 2. Mapping of E-books in Science & Technology: An Analytical Study of Directory of Open Access Books 3. Web Impact Factor Analysis for Deemed Universities in Andhra Pradesh 4. Applications of Thesaurus in Digital Libraries 5. Investigating Knowledge Management Strategies in Central University Libraries in India

Topic d	<ol style="list-style-type: none"> 1. An Overview of Online Exhibitions 2. Voices: A Virtual Exhibition as Alternative Model to the Forum 2004 3. A Critical Look at Online Exhibitions and Online Collections: When Creating One Resource is more Effective than the Other 4. Implementation of a Virtual Archives System using Virtual Reality Technology: A Case Study of the National Archives Administration Taiwan (ROC) 5. Universal Decimal Classification: Past and Present 	<ol style="list-style-type: none"> 1. E-Learning: Liberation of Education and Training with Evolving Library and Technology Support 2. HELLS Network– Not Just a Consortium of e-Resources 3. Libraries–a Key to Harness E-Learning: Issues and Perspective 4. Web 2.0 and E-Learning: The Indian Perspective 5. Gyanoday Portal: A Virtual Reference Point for Management Knowledge Repository 	<ol style="list-style-type: none"> 1. ADDIE: Designing Web-enabled Information Literacy Instructional Modules 2. Indian Integrated Plant Taxonomic Information System: A Conceptual Framework 3. Assessing Information Literacy Competence among the Undergraduate Students of College of Agriculture, Raichur: A Case Study 4. Use of IT-related Technologies in Hyderabad-based DRDO Laboratories: An Evaluative Study 5. Visualisation of Relationships Among Library Users Based on Library Circulation Data 	<ol style="list-style-type: none"> 1. Content-Based Document Recommender System for Aerospace Grey Literature: System Design 2. Content-based Document System for Aerospace Grey Literature: Experimental Testing and User Opinion Survey 3. Online Public Access Catalogue Usage at Punjab University Library, Chandigarh 4. Open Access Institutional Archives: A Quantitative Study (2006-2010) 5. Perception and Use of Social Networking Sites by the Students of Calicut University 	<ol style="list-style-type: none"> 1. Contribution and Citation Impact of Materials Science Research in India, 2001-10 2. Research Productivity in a Management Institute: An Analysis of Research Performance of Indian Institute of Management Ahmedabad during 1999-2010 3. Trends in Classification Literature: Analysis of Literature Published during 2000 to 2009 4. Comparative Evaluation of Research Output: AIIMS vs PGIMER 5. Blogs on Marketing Library Services 	<ol style="list-style-type: none"> 1. Training Needs of School Librarians in India 2. Knowledge on ICT Skills among LIS Professionals of Engineering Institutions of Andhra Pradesh State: A Survey 3. Use of Library Portal by Research Scholars and Faculty Members at Kuvempu University: A Survey 4. Analysis of Open Access Scholarly Journals in Media and Communication 5. Marketing Strategies for the University Libraries in Kerala 	<ol style="list-style-type: none"> 1. Glaucoma Research: A Scientometric Study of Indian Publications Output, 2002-11 2. Use of Social Media by Online Newspapers in Saudi Arabia 3. Publication Productivity of University of Kerala: A Scientometric View 4. Foreign MNC R&D Centers in India: A Study of their Publications, 2003-12 5. Communication of Universities of Asia through Facebook: A Study 	<ol style="list-style-type: none"> 1. Metadata Diversity, Interoperability and Resource Discovery Issues and Challenges 2. Tag Cloud Application and Information Retrieval System: Visualisation to Create Information Literacy 3. Collection Development in Indian Academic Libraries: An Empirical Approach to Determine the Number of Copies for Acquisition 4. Information Seeking Behaviour of Law Students in the Changing Digital Environment 5. Design and Development of an Online Exhibition on the Tangkhul Tribe Festivals 	<ol style="list-style-type: none"> 1. Nanotechnology Ontology: Semantic Access to Information in the Nano World 2. A Scientometric Assessment of Global Publication Output on RFID with Reference to India During 2006-15 3. Availability of Open Access Books in DOAB: An Analytical Study 4. Relevance of In-house Publications in Knowledge Management: A Case Study from Naval Physical and Oceanographic Laboratory 5. Perception and Implementation of Knowledge Management by University Librarian in India 	<ol style="list-style-type: none"> 1. Research Data Management Practices in University Libraries: A Study 2. RFID Technology in Libraries: A Review of Literature of Indian Perspective 3. Indian Government Websites : A Study 4. Mobile Information Services and Initiatives in University Libraries: A New Way of Delivering Information 5. Use of Collection and Services: A Study of Indian Agricultural Research Institute Library
Topic e	<ol style="list-style-type: none"> 1. Semantic Web Services: A Study of Existing Technologies, Tools and Projects 2. Online Virtual Exhibitions: Concepts and Design Considerations 3. Application Domain and Functional Classification of Recommender Systems—A Survey 4. Online Exhibition Authoring System with Intelligent Affective Background Image Composition 5. Knowledge Sharing through Intranet 	<ol style="list-style-type: none"> 1. Availability, Use and Barriers to ICT in the R&D Institutions: A Case Study of the Libraries and Information Centres in Noida 2. Content Analysis of Indian Institutes of Technology Libraries Web Portals: A Study 3. Intrinsic Motivation of Librarians in University Libraries in Karnataka 4. Technology Enhanced Learning: A TCS Library Experience 5. Preservation of Library Materials: Problems and Perspective 	<ol style="list-style-type: none"> 1. Quality Assessment of Libraries 2. Usage and Performance of Various Library Software Modules in Engineering Colleges of Karnataka 3. Websites of Antiquarian Books 4. Transforming a Traditional Library to Modern Library using Barcode Technology: An Experience of Central Library, PEC University of Technology, Chandigarh 5. Information Literacy through Web 2.0 Integrated WebOPAC: An Experiment at Jaypee Group of Institutions 	<ol style="list-style-type: none"> 1. Web Interface in Library Management Software Systems 2. Current Trends in Wireless Technologies in Academic Libraries 3. Web 2.0 and Libraries: Facts or Myths 4. Implication of ICT's in Libraries of Higher Education Institutes: A Panacea 5. University Libraries in India in the Digital Age 	<ol style="list-style-type: none"> 1. Cloud Computing and Libraries 2. Establishment of Institutional Mechanism for Building National Repository in Health Sciences 3. Cloud Computing: A Promising Economic Model for Library and Information Centers 4. Application of Information Communication Technology to the Management of Library's Readers' Desk 5. Information Technology on Surge: Information Literacy on Demand 	<ol style="list-style-type: none"> 1. Information Needs and Use of Healthcare Professionals: International Perspective 2. Developing a Standardised Tool for Impact Assessment of INFLIBNET 3. Application of Mobile Technologies to Libraries 4. DRDO Knowledge Repository 5. Consortia System and Document Delivery: E-information Services of Banasthali University 	<ol style="list-style-type: none"> 1. Development of Semantic Web-based Knowledge Management for Nuclear Reactor (KMNuR) Portal 2. Academic Library's Responses to the Emerging Trends in Higher Education 3. Metadata Description Framework for Integration of Bioinformatics Information Resources: A Case of iBIRA 4. Data Mining: Competitive Tool to Digital Library Catalogue 5. Interfaces of Integrated Library Management Systems (ILMS): Experiences in a Proprietary and Open Source Software 	<ol style="list-style-type: none"> 1. Information Management Skills Required by the Minority Libraries in Kolkata and Hooghly Districts, West Bengal 2. Adoption of Open Source Software in India 3. Software Selection and Deployment for Library Cooperation and Resource Sharing Among Academic Libraries in South-West Nigeria 4. Communication of Universities of Asia through Facebook: A Study 5. Library Resources and Services in the Selected University Libraries of Haryana, India 	<ol style="list-style-type: none"> 1. Awareness and Use of E-resources: A Case Study of Mohinder Singh Randhawa Punjab Agricultural University Library, Ludhiana 2. Use of E-resources by the Medical Students of M.M. University, Ambala: A Case Study 3. Use of E-resources by Users of Dr. Y.S. Parmar University of Horticulture and Forestry 4. Citation Analysis of Academic Publications to Identify Subscribed Journals Usage: A Case Study 5. Usage Preference of E-Publications by Health Professionals of Dayanand Medical College and Hospital, Ludhiana (Punjab) 	<ol style="list-style-type: none"> 1. Mobile Information Services and Initiatives in University Libraries: A New Way of Delivering Information 2. Efficient Use of Resources for Statistical Machine Translation 3. Patent Search Tools for Competitive Intelligence for Product Design and Development of CBRN Decon Device 4. Use of Mobile Devices by Library and Information Science Students in Central Universities of Uttar Pradesh 5. Evaluation of E-Learning Web-Portals