

# APPLICATION OF TEXT MINING TECHNIQUES IN DOCTORAL THESES OF LIBRARY AND INFORMATION SCIENCE AWARDED IN INDIA

A Abstract Submitted to the  
Babasaheb Bhimrao Ambedkar University, Lucknow  
in Fulfilment of Requirement for the Award of Degree of

**Doctor of Philosophy**  
**in Library & Information Science**



BY

*Khusbu Thakur*

Enrollment No.: 168/18

SUPERVISOR

*Dr. Vinit Kumar*

Assistant Professor

DEPARTMENT OF LIBRARY & INFORMATION SCIENCE  
SCHOOL OF INFORMATION SCIENCE AND TECHNOLOGY  
BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

(A CENTRAL UNIVERSITY)

VIDYA VIHAR, RAEBARELI ROAD, LUCKNOW-226 025  
UTTAR PRADESH, INDIA

**2023**

# ABSTRACT

---

## 1.1 Introduction

Research-based literature is rapidly growing in the academic world. As time goes on, new challenges arise to replace the previous ones. Many researchers believe that being aware of emerging topics, concepts, or disciplines relevant to their specific field of study is critical if they are to identify new ideas and disciplines. In daily life, it is impossible to read an entire text on days when you are busy. In the academic world, publications for research, electronic theses and dissertations, presentations at seminars and conferences, and academic journals are all updated instantly. In higher education, there is a lot more scholarly literature in the fields of social science, humanities, and science and technology. According to the AISHE 2020-2021 report, there are 2,881 students enrolled in integrated Ph.D. in addition to 2.02 lakh students enrolled at the Ph.D. level. However, the Social Science stream has 19,965 Ph.D. students. This indicates that the number of doctoral works is increasing. Many universities, research fellows, and faculty members in many different fields have also put out research articles in electronic form. It also shows that they all work hard at research and are involved in making new discoveries and finding new ways to use the pre-existing literature. This meant that changes were happening on a regular basis in universities, whether through direct or indirect methods. In higher education, doctoral research plays a crucial role in the growth and development of universities (Hussain, et al.,2021). It is one of the most difficult programs they have ever done in their professional lives. When a person gets his Ph.D. in a certain field, they have done research that meets the high standards of the academic community and contributes to the advancement of knowledge in that field. Doctoral research is important because it gives people the chance to do original, independent research and make a contribution to their chosen field. They can find out important things about society, help to comply ways to cope with problems, and learn new things. Doctoral researchers often play a big part in expanding what we know, coming up with new ideas, and laying the groundwork for future research. They are able to do research that has real-world applications, helping to solve difficult problems and learn new things. Often, doctoral researchers help push the limits of what we know, come up with new ideas, and pave

the way for future research. It plays a critical role in educating and also educates the next generation of scholars by sharing their knowledge and experience. Indeed, universities/organizations/research institutes have to spend a significant amount of money on research equipment (Vijayakumar & Vijayakumar, 2007). So that they can concentrate on conducting excellent doctoral research.

A literature review is a part of doctoral research that gives detailed information about research from top to bottom. A significant amount of effort and time is required to select a good research topic for a Ph.D. program. For this, researchers have to follow some important approaches while choosing the research topic. Initially, the researcher must identify their area of interest by reviewing relevant literature and previous research. After that, connect them to the societal problem and talk with a supervisor about the benefits and drawbacks of the topic, as well as what resources are available in universities. Finally, researchers will select a topic that is also academically relevant. The research interest should be narrow and focused on topics that may advance scientific knowledge, solve practical problems, or benefit society (Adhikari, 2020). Once the research topic selection process is complete, clearly understand why it is interesting and significant, and then use the appropriate tools and techniques.

In the present scenario, many LIS researchers have admitted that the number of students enrolling in Ph.D. programs is constantly increasing (Kumar & Thakur, 2022; Mazumder & Barui, 2021; Nishat, et al., 2019). However, more than 100 universities in India now offer Ph.D. degree programs. LIS research advances primarily through three channels: the publication of research articles, the presentation of research papers at conferences and seminars, and the submission of theses and Dissertations (Kumar & Thakur, 2022). Finding a good research topic or understanding how research is done can be hard for new researchers. For new research scholars, choosing a research topic from the huge amount of literature and resources available can be the hardest thing to do. Because of the vast amount of published literature available in various disciplines, as well as a lack of potential and inability to read the text in depth, inexperience in searching specific topics or fields, and poor guidance and communication. With the fast growth of electronic text data in recent years, it has become very hard to analyze, manage, organize, and reach these goals. Even if it takes significant time and effort to access popular research trends, there is a risk of inaccuracy or loss of credibility in the academic world. Even identifying the field &

growth, relevance, and future research trends becomes difficult too. So, content analysis is a method used for determining research patterns and trends, conducting in-depth studies, generating themes, and determining the growth of the field. However, CA required significant time and effort which can be saved by applying tools and Techniques developed in Natural Language Processing. NLP is a method of analyzing unstructured text data that employs a variety of tools and techniques. It is a way to use machine learning to find meaningful themes and patterns in a large number of documents. Text mining tools and techniques are considered new technology in the twenty-first century that is popular in academia. So, these techniques are being used to find trends and growth patterns in research from large-text data sources like academic journals, e-newspapers, and ETDs (Oyri, 2005; Ferreira Mello, et al., 2019; Nagarkar & Kumbhar, 2015). The evolution of research trends is an important aspect of the study as can be elucidated in changing spheres of library and information science. And the finding of the study will be helpful in understanding the topic pattern, and research trends that are faced by the new researcher while choosing topics in the LIS domain.

### **1.2 Significance of the Study**

Text mining is a method for analyzing unstructured text to extract hidden patterns. The primary goals of this research are to use text mining techniques to identify topic trends in the doctoral-level study and to present the topics among them as a convenient way to assist newly admitted researchers, faculty, and scientists in determining relevant topical growth of the field in each period. Newcomers are becoming more interested in the research done in doctoral theses in LIS, which will shape the next generation of doctoral researchers. The continuity of research in this domain may increase the visibility of the LIS profession and increase the usage and impact of LIS research outputs. And finally, this study may work as a guide for the researchers and help them select their research area in the future with minimum effort.

### **1.3 Objectives of the Study**

The proposed study is based on the following objectives:

1. To conduct the Unigram, Bigram, and Trigram analysis on the Indian LIS doctoral dissertations corpus.

2. To identify the dominant themes on which doctoral-level LIS researchers conducted studies in 2002-2021.
3. To identify the trends that emerged in the dominant themes on which doctoral-level LIS researchers chose research topics from 2002 to 2021.
4. To identify the university-wise dominant topics on which researchers have undertaken study for the top ten universities on the basis of a number of theses available in shodhganga.
5. To identify the dominant topic-wise trends on which researchers from the top ten universities (based on the number of theses available in shodhganga) have undertaken research.

#### **1.4 Scope and Limitation of the Study**

The scope of the study will be the doctoral dissertations submitted to the Department of Library and Information Science at Indian universities. The present study will be limited to theses available in the Shodhganga repository and written in English.

#### **1.5 Research Methodology**

Due to the unavailability of electronic copies of all the theses awarded by LIS schools to date, data for this study were collected from the electronic version of the awarded thesis, which is accessible through the Shodhganga repository. From 2002 to 2021, 1880 doctoral theses were taken from an Indian digital repository. In this study, the LDA topic modeling technique was used with the orange data mining tools to find the hidden topics. The dataset contains words in uppercase or lowercase, stop words, punctuation, and extra characters. So, we use methods called pre-processing” to clean, tokenize, and lemmatize the text data. In addition, for topic visualization, Orange tools, and an Excel sheet were used to display topics and the frequency of their unigram, bigram, and trigram keywords. This study examined the evolution of library and information science (LIS) by analyzing doctoral dissertations in LIS. The analysis focuses on 1880 theses, using topic modeling (LDA) to identify topics.

#### **1.6 Objectives Wise findings**

In this section, the findings of the study have been discussed in light of the objectives taken at the beginning of the study. The first objective of the study was:

Objective 1: To conduct the Unigram, Bigram, and Trigram analysis on the Indian LIS doctoral dissertations corpus.

### **Unigram Terms**

It was found that the word “Library” has appeared a maximum number of times (39401) in the LIS theses followed by Information (27884), University (13641), Research (13249), Service (10520), and Knowledge (4561) appeared the least number of times. It can be said that most of the research in LIS revolves around topics covering terms such as Library and Information Services, University, Collection Development, Information resources, and the need for college resources. This information suggests that the focus of research in LIS is primarily on the development and library and information services in universities, as well as highlights the importance of college resources in this field. This information can be useful for researchers in the field of LIS to identify the most popular topics and keywords that have been extensively researched. It also highlights the importance of these topics in the field and their relevance to current research trends.

### **Bigram Terms**

It was found that the bigram word “university library” has appeared the maximum number of times (3615) in the LIS theses, followed by library information (3608), college library (2500), public library (1925), library service (1840), and library user (1145), which have appeared the least number of times. It indicated that most of the theses were in the areas of university libraries, library information, college libraries, public libraries, library services, faculty members, and information seeking. This finding suggests that the research focus in LIS theses is mainly on academic and public libraries, their services, and the information-seeking behavior of users. However, there may be a need for more research in other areas of library and information science to provide a more comprehensive understanding of the field.

### **Trigram Terms**

It was observed that the phrase Library Information Science has appeared the maximum number of times (960) in the LIS theses, followed by Information Seeking

Behavior (765), Information Communication Technology (650), Engineering College Library (590), Electronic Information Resources (375), Library Information Service (369), and University Library India (167). It indicates that most of the theses were in the areas of information-seeking Behavior, engineering college libraries, electronic information resources, and faculty member research. It indicated that most of the theses were in the areas of information-seeking Behavior, engineering college libraries, electronic information resources, and faculty member research. This suggests that there is a significant interest in improving access to information and resources within the academic community, particularly in the field of engineering. It also highlights the importance of supporting faculty research and providing access to electronic resources.

**The top 30 words/phrases frequently appear in the LIS doctoral theses from 2002-2021**

It was found that the topics of the LIS theses were mined based on the keywords of the topic for each five years period and identified the most utilized hot keywords during 2002-2021.

As indicated:

- In 2002-2006, found that the keywords revolving around the topics were library services, user studies, public library systems, ICT, and collection development.
- In 2007-2011, found the top keywords with the highest frequency rate were Library Service, collection development, public library system, LIS education, knowledge management, bibliometric/scientometric, webometric studies, and open access journal.
- In 2012-2016, it indicated that the keywords of the topic are electronic resources, digital libraries, library services, webometric studies/ evaluation of the website, LIS education, open access journal, collection development, Academic libraries, LIS professionals, and user studies.
- In 2017-2021, it was revealed that keywords of topics are library service, electronic resources, user studies, ICTs, knowledge management, collection

development, digital library, webometric study, LIS education, and open access journal preferred by doctoral researchers in their theses.

- It also found that the trigram phrase “Library Information Science” has appeared the maximum number of times (960) in the LIS theses, followed by Information Seeking Behavior (765), Information Communication Technology (650), Engineering College Library (590), Electronic Information Resources (375), Library Information Service (369), and University Library India (167).

Objective 2: To identify the dominant themes on which doctoral-level LIS researchers have conducted studies from 2002-2021.

It was found that over the past 20 years, the maximum of LIS theses (14.36%) have been awarded on topics related to "Bibliometric/Scientometric/Citation analysis followed by "Digital Library", "Awareness of library Resources/E-Resources", "Information Seeking Behaviour", "Public Library System & Services", "Evaluation of Library Service", "Knowledge Management/Total Quality", "ICT & Its Application in Libraries" and The minimum doctoral research topics selected by LIS researchers in India between 2002 and 2021 were "Job Satisfaction", "Open Access Journals and Scholarly Communication", "Case Studies on various Types of Libraries", "Content Analysis /Evaluation of Website Analysis /Webometrics", "LIS Education in India", and "Reading Habits". The provided table emphasizes that the highest number of doctoral theses being awarded are on the topic of "Bibliometric/Scientometric/Citation Analysis" and are highly preferred by researchers in the field of LIS in India. The study revealed that "Bibliometric /Scientometric/Citation analysis", "Digital Library", and "Awareness of Library Resources /E-Resources" are the top three research areas in LIS over the past two decades. However, these topics "LIS job Satisfaction", "Open-Access Journals and Scholarly Communication", and "Case studies on various types of Libraries" were the least explored topics by LIS researchers in India during the same period. The findings suggest that "Bibliometric and Scientometric analysis" has been a popular area of research in LIS over the past two decades, reflecting the growing importance of quantitative methods in the field. Additionally, the relatively low percentage of doctoral research on "Reading Habits" may indicate a need for more attention in this area in future research.

*Objective 3: To identify the trends that emerged in the dominant themes on which research topics were chosen by doctoral-level LIS researchers from 2002 to 2021.*

The study found the top ten research topics for doctoral theses in library and information science. Doctoral researchers have extensively used the dominant and most popular themes over the past twenty years as research topic trends in doctoral theses. Some topics occur in multiple times such as “Information Seeking Behaviour”, “Bibliometric/Scientometric/Citation Analysis”, “ICT & Its Application in Libraries”, “Digital Library”, “Evaluation of Library Service”, and “Knowledge Management/Total Quality Management”. These are the core topics in LIS theses from 2002–2021. However, there are changes within these topics. Awareness of Library Resources / E-Resources" occurs from 2007–2021. Between 2002 and 2011, the T2 "Case Study on Different Types of Libraries" saw its greatest growth, which shifted between 2007 and 2011. The T3 "LIS Education" was found to have little growth, specifically from 2002 to 2006; following this, researchers focused on the other T14 "Content Analysis / Evaluation of Websites / Webometrics", which occurred from 2007 to 2011. The topic "LIS Job Satisfaction" was noted to occur from the period 2012 to 2002, and the dominant topic "Open Access Journal and Scholarly Communication" occurs only in the periods 2012 to 2016. So, this study highlights the rise and fall of different topics over the last twenty years, providing insights into the changing research interests of doctoral researchers in India. It can be used to identify emerging trends and gaps in research that need to be addressed. It is also interesting to note that the popularity of certain topics has remained consistent over the years, while others have seen a significant increase or decrease in interest. This data could be useful for universities and research institutions in India to identify emerging trends and adapt their programs accordingly.

*Objective 4: To identify the university-wise dominant topics on which researchers have undertaken study for the top ten universities on the basis of a number of theses available in shodhganga.*

- It is found that the top ten leading universities in library and information science have the most Ph.D. theses awarded in fourteen major topics over the last twenty years (2002-2021).

- It is found that Alagappa University awarded 51 LIS doctoral theses in the last two decades. The highest number of theses were awarded in the dominant topics T1 and T12 each with 29.41%, followed by theses in T7 (11.76%), and T14 (9.80%). Further, it awarded a few theses in T2, T3, T8, and T9 which shows that the focus of Alagappa University is in areas such as “Awareness of Library Resources / E-Resources”, “Bibliometric/Scientometric/Citation Analysis”, “Information Seeking Behaviour”, “Content analysis/Evaluation of website Analysis/Webometrics”. The findings suggest that the doctoral students at Alagappa University are interested in exploring the use of “Electronic Resources in libraries” and analyzing “Bibliometric and Scientometric Analysis” in LIS. However, the “Public Library System” and “Open Access Journals” are areas that require more attention from researchers.
- It is found that, the Aligarh Muslim University has uploaded 62 theses out of which 22.58% were from T1, 12.90% from T4, 16.67% from T6, and 11.29 % from T11. This shows that, AMU researchers focus more on the areas of Awareness of library resources/E-Resources While T3 and T10 contributing 1.61% revealing that “LIS Education in India” and “LIS Job Satisfaction” there are some these in “Public Library System & Services”, “Digital Library” and “Knowledge Management/Total Quality Management” also. This suggests that there is a need to encourage research on “Reading Habits” among students and faculty members. Additionally, it highlights the importance of promoting awareness and utilization of “Awareness of library Resources and E-Resources” among the academic community.
- The study found that Bharathiar University has awarded 102 theses. The dominant topic T12 the most doctoral theses conducted (23.53%), followed by the T1 (18.63%), T6(16.67%), and T4 (10.78%). It demonstrates that the majority of theses produced by the Bharathiar University Department of LIS focus in the fields of “Bibliometric, Scientometric, and Citation Analysis”. Additionally, some theses contribute in the areas of “Awareness of Library Resources and E-Resources”, “Digital Library”, “Public library Systems and Services”, demonstrating the university focus. The university's emphasis on these areas of research is reflected in the high number of theses awarded and

doctoral theses conducted. This highlights the university commitment to promoting research in diverse fields and encouraging scholars to explore less-explored areas.

- The study found that Guwahati University uploaded 53 doctoral dissertations during the study period, of which 16.98% were from T8, T4, and T6 from 15.09%, T12 from 9.43%, and T7 from 7.50%. This demonstrates that GU researchers concentrate on “Bibliometric/ Scientometric/Citation Analysis” more than “Evaluation of Library Services”, “Public Library System & Services”, and “Digital Library”. Other research themes that have gained more interest among doctoral researchers at Guwahati University include topics related to environmental science, social issues, and technology. It is important for universities to encourage diversity in research themes to promote a well-rounded academic community.
- It found that Karnatak University has awarded a total of fifty-seven theses; T6 has the highest frequency of LIS theses awarded (21.05%), followed by T12 (19.30%), T1(10.53%), and T13 (10.05%). This demonstrates that Karnatak University researchers concentrate on “Digital Library” more than “Bibliometric/ Scientometric/ Citation Analysis”, “Public Library System and Services”, “Awareness of Library Resources/E-Resources” and “LIS Education in India”. This indicates that there is a significant interest in the development and implementation of digital libraries among students and researchers in the field of Library and Information Science. However, further research may be needed to explore the reasons behind the lower number of works on topics such as “Open-Access Journals Scholarly Communication”, and “Reading Habits”.
- The study found that a total of 56 theses awarded at Madurai Kamaraj University during the past 20 years, with T12 (33.93%) being the most popular, followed by T1 (23.21%), T6 (14.29%), and T4 (8.93%). This shows that MKU researchers prioritize Bibliometric/ Scientometric and Citation Analysis” more than Public Library System and Services”, Awareness of E-Resources”, and Digital Library. The trend in research topics at Madurai Kamaraj University suggests a strong focus on quantitative analysis and

evaluation of scholarly communication. However, there is a need for more research in the areas of library management, education, and user behavior to better inform library practice and policy.

- The study found that Manonmaniam Sundaranar University has awarded 101 doctoral dissertations over the past two decades. The dominant topic T1 had the highest proportion of theses with 20.72%, followed by T4 and T12 with 14.41% and 14.41%, respectively. Additionally, a few numbers of theses were awarded in T6, T3, T8, T10, T14, and T2. This indicates that Manonmaniam Sundaranar University concentrates on areas such as “Awareness of Library Resources / E-Resources”, “Public Library System & Services”, and “Bibliometric/Scientometric /Citation Analysis”. This indicates that the university places a significant emphasis on research related to “Awareness of library Resources and E-Resources” while there may be less interest or focus on studying “Reading Habits” among students and faculty. It would be interesting to explore the reasons behind this trend and how it reflects the priorities of the university.
- It is found that Savitribai Phule Pune University awarded 59 doctoral theses of LIS in this study. The highest number of doctoral works were awarded in the dominant topic T4, T7, and T11 with (15.25% each), followed by T1 (8.47%), T5 (6.78%), T6 (10.17%). Further, it awarded some theses in T2, T9, and T14. Which indicates that the focus of Savitribai Phule Pune University is in area such as Public Library System Services, Information Seeking Behaviour, and Knowledge Management/Total Quality Management. The findings suggest that there is a growing interest in research related to “Public Library Systems and Services”, “Information-Seeking Behavior”, and “Knowledge Management/Total Quality Management”. On the other hand, the lower number of doctoral theses on “Case studies of different types of Libraries” and “Reading Habits” may indicate a need for further research in these areas.
- From this study, Sri Venkateshwara University awarded 58 doctoral theses of LIS in this study. The highest number of doctoral works were awarded in the dominant topic T7 (20.69%), followed by T1 with 13.79%, and T4 with 17.24%. Further, it awarded some theses in T2, T3, T9, T10, T11, and T13

Which indicates that the focus of Sri Venkateshwara University is in areas such as “Information Seeking Behaviour”, “Awareness of Library Resources/E-Resources”, and “Public Library System & Services”. The results suggest that there is a need for more research in the area of reading habits among doctoral researchers in India. Additionally, it would be interesting to explore the reasons behind the popularity of "Information Seeking Behaviour" as a research theme at Sri Venkateshwara University.

- It is found that the University of Mysore has awarded 102 theses the reports that the highest number of LIS theses were awarded on T1 (26.32%), followed by T6 (17.11%), T12 (14.47%), and T4 (9.21%). Whereas a minimum number of doctoral students worked on T2 and T8 (1.32%) and further, it awarded some theses in T2, T3, T9, T10, T11 and T13, which indicates that the focus of University of Mysore is in area such as “Awareness of library Resources/E-Resources”, “Digital Library”, “Bibliometric/ Scientometric / Citation Analysis”. This suggests that there is a need for more research and attention on the themes of “Open Access Journal and Scholarly Communication” as well as “Content Analysis/Evaluation of website analysis/Webometrics” in the field of library and information science. In general, it is observed that from the study the theme "Awareness of Library Resources/E-Resources" maximum number of theses awarded, and the theme “Reading Habits” represents a very low area of interest among all dominant topics as well as performed very less research (PhDs) in this doctoral research topic under top ten Indian universities between 2002 and 2021. This highlights the need for further research and attention toward promoting reading habits among individuals. It also emphasizes the importance of creating awareness about library resources and e-resources to maximize their utilization.

*Objective 5: To identify the dominant topic-wise trends on which researchers from the top ten universities (based on the number of theses available in shodhganga) have undertaken research.*

- In terms of the fifty or more doctoral theses in library and information science that were available between 2002 and 2021. The top fourteen dominant topics for doctoral research in LIS in India.

- It was found that in theme one T1, that is the “Awareness of library resources/E-Resources” the top 10 universities awarded a total of 125 theses. Out of which the MSU, Baroda awarded the maximum (18.40 %), followed by UoM (16%), BU (15.20%), and AU (12%). This indicates that these universities are majorly contributing in this area in comparison to other universities.
- In this study found that in dominant T2, “Case studies on different types of Libraries” the top 10 universities awarded a total of 13 LIS doctoral theses during the last twenty years. The SPPU has awarded the most doctoral theses in this dominant topic (23.08%), followed by the SVU, AMU, GU, and MSU with 15.38% each. This suggests that SPPU has made significant contributions to the field of LIS and has a strong research focus on case studies of various types of libraries. However, over the past 20 years, other universities need to pay more attention to this area.
- This study found that last twenty years, the top ten universities dominant topic T3, “LIS Education in India” were awarded a total of 14 doctoral theses in LIS. The Most doctoral student chose this dominant topic by MSU (28.57%), followed by BU, KU, and SVU, each with 14.29%. The findings suggest that MSU has been the leading university in producing doctoral theses in LIS Education in India. However, it is worth noting that other universities such as BU, KU, and SVU have also made significant contributions to this field.
- This study revealed that the top 10 universities awarded a total of 80 LIS doctoral theses for the dominant topic T4, which is “Awareness of library Resources/E-Resources”. AMSU (20.00%), BU (13.75%), and SVU (12.50%) have the highest number of doctorate theses awarded in these areas, respectively.
- In this study found that a total of 55 doctoral dissertations on the dominant topic T5, “Open Access Journal, and Scholarly Communication” were awarded by the top ten universities. AMU and UoM were awarded the most doctoral dissertations (both 22.73%), followed by SPPU (18.18%) and GU (13.64%) during the last twenty years.

- This study found that ten universities awarded a total of 89 theses in the dominant topic T6, “Digital Library” though BU (19.10%), UoM (15%), KU (13.48%), and SVU (11.24%) awarded the maximum number of theses awarded last twenty years.
- This study found that a total of 75 doctoral theses on dominant topic T7, “Information Seeking Behaviour” awarded by the top ten universities, however, MSU (22.67%) maximum the theses awarded on this topic, followed by SVU (16.00%) and KU (10.67%) in this study.
- This study revealed that last two decades, the top ten leading universities have been awarded a total of 34 doctoral theses. Maximum theses awarded for LIS theses in the dominant topic T8, “Evaluation of Library Service” were chosen by GU (26.47%), followed by AMU (17.65%), KU, MSU, and SPPU (11.76%).
- It was found that top ten universities were awarded in the dominant topic T9, “Reading Habits” were the used maximum doctoral theses were awarded by 50.00% from both SPPU and SVU.
- It was found that in the last twenty years of LIS research in India, a total of 20 doctoral theses were awarded on the dominant topic T10, “LIS Job Satisfaction” though BU used this topic the most (40.00%), followed by MSU (20.00%), GU and MKU (10.00%) in India.
- This study found that the dominant topic T11, “Knowledge Management/Total Quality Management”. It is revealed that the topic most frequently used and awarded by AMU (22.58%), followed by GU (12.90%) and UM (9.68%) over the past two decades, when a total of 31 doctoral theses were awarded in the top ten universities.
- It was found that a total of 115 doctoral theses were awarded by the top ten universities dominant topic T12 in India. However, BU was used the most in T12, “Bibliometric /Scientometric / Citation Analysis” in doctoral LIS research (20.87%), followed by MKU (16.52%), MSU (13.91%), and AU (13.04%).

- This study found that 46 doctoral theses on T13, “ICT and its application in libraries” were awarded by the top ten universities. However, most researchers at MSU (32.61%) and BU (14.04%) chose this topic in their doctoral theses.
- This study found that the top ten universities were awarded a total of 18 theses. The dominant topic T14, “Content Analysis / Evaluation of Website Analysis / Webometrics” used maximum for doctoral-level LIS research by AU (27.78%), GU (16.67%), and UoM (16.67%). Between 2002 and 2021, LIS research (PhDs) was awarded in Indian universities.

### **1.6 Suggestions**

- The study attempts to study the availability of LIS theses on Shodhganga and applies text mining to the contents of the available theses in order to understand the research trends. The study suggests the following recommendations.
- The study suggests that central, private, and deemed universities should contribute and upload Ph.D. theses more frequently. This would help in making the research more accessible and visible to a wider audience. Additionally, it could lead to more collaborations and advancements in the field.
- The study suggests that researchers should focus on less explored topics in their research work to contribute to the advancement of the field. Additionally, exploring these less common topics can provide new insights and perspectives that can benefit the overall development of library and information science.
- This study suggests developing a dashboard and categorizing university research into broad topics and subtopics. so that researchers and faculty members can easily visualize the topic of research. This will help in maintaining the quality as well as avoiding duplicate or repetitive research. This can be achieved through the use of data visualization tools. Additionally, implementing strict guidelines for research ethics and plagiarism can help maintain research quality and originality.

- This study suggests that the search features of Shodhganga should be improved so that users can properly read and download the full documents in proper sequence. The visualization should also be user-friendly and easy to navigate, allowing users to quickly find the information they need without getting lost or confused. Additionally, incorporating interactive features such as search bars or clickable links can further enhance the user experience.
- This study suggests that a separate dashboard should be created for year-wise and language-specific theses with specific subject domains. So, that users can easily search for their preferred language as well as the year of theses without wasting too much effort. This approach would improve the accessibility and user-friendliness of the Shodhganga database, making it easier for researchers to find relevant theses. Additionally, it could potentially increase the usage and visibility of the Shodhganga database by catering to a wider range of users.
- Further, this study suggests that maintainers should check for duplicate theses, maintain proper chapters of theses in a proper sequence, and remove repetitive theses within a month. These actions can significantly improve the quality of theses and reduce the workload of reviewers. It is recommended that academic institutions implement these measures to ensure efficient and effective thesis maintenance.
- Furthermore, the study recommends that ETD should also provide an option for users to download the bibliographic records to conduct bibliometric studies.
- This study suggests that many reputable and well-known universities that conducted research and signed an MoU with the INFLIBNET Center in Shodhganga to upload their doctoral research theses to a digital Indian repository, but a number of universities have been inactive in uploading their doctoral research theses. So, all universities strictly enforce the UGC rules and regulations.

## **1.7 Chapterisation**

The present research work is divided into the following chapters:

**Chapter 1: Introduction-**This chapter introduces the background of the study and describes the problems, objectives, needs, scope of the study, and limitations of the study

**Chapter 2: Review of Literature-**This chapter critically appraises the previous studies closely related to the current research to understand the methodology, analysis tools, and results.

**Chapter 3: Methodology-** This chapter includes the details of text mining techniques and steps in details.

**Chapter 4: Data Analysis and Interpretation-**This chapter includes the analysis of the whole data represented in the form of various graphs and tables.

**Chapter 5: Findings, Conclusions, and Suggestions-**This chapter summarizes the research and includes the final results and areas for future research.