DESIDOC Journal of Library & Information Technology, Vol. 45, No. 4, July 2025, pp. 326-338, DOI : 10.14429/djlit.21029

Mapping of Emerging Technological Trends in Library and Information Science: A Computational Approach Using Sentiment Analysis, Topic Modeling and Network Analysis

Rajeevan M S[#,*], B. Mini Devi[#], V.S. Anoop[~] and C. Mallikarjuna[!]

[#]*Department of Library and Information Science, University of Kerala, Thiruvananthapuram - 695 034, India*
[~]*Amrita School of Business, Amrita Vishwa Vidyapeetham, Kochi, India - 682 041, India*
[!]*Indian Institute of Technology Hyderabad (IITH), Telangana - 502 284, India*
[*]*Email: rajeevanms_2025@keraluniversity.ac.in*

ABSTRACT

The study investigates the technological evolution of Library and Information Science (LIS) research by applying Sentiment Analysis (SA), Topic Modeling (TM) and Network Analysis (NA). The study seeks to trace sentiment changes, primary themes in the research and collaboration throughout LIS research globally. A dataset of 918 publications indexed in Scopus (2000-2024) was analysed. Sentiment analysis used VADER for sentiment scoring (positive, negative, and neutral), applied LDA and BERTopic for the topic modeling and used Neo4j based knowledge graphs to map collaboration between institutions. Overall results indicate a predominately positive sentiment regarding AI and digital libraries; while automation drew mixed sentiment. The topic modeling represent five themes depicting digital transformations within library and information science. The analysis of networks revealed which institutions contributed strongly to the body of research. Wuhan University and Florida Atlantic University emerged as collaboration hubs. Differences in cross-institutional collaboration networks were found, with different levels of centrality of other institutions across geographic contexts. Limitations include the use of only abstracts in English that does not include grey literature. The findings provide evidence of how LIS research is framed by technology trends and networks of scholars. The research study provides a conceptual framework for new studies using computational bibliometrics in LIS.

Keywords: Artificial intelligence (AI); Automation in libraries; Knowledge graph; Natural language processing (NLP); Neo4j; Network analysis

1. INTRODUCTION

The rapid advancement of digital technologies has changed the landscape for Library and Information Science (LIS) research altering how information is accessed, organised and disseminated. Despite the changing landscape of information research and professional practice, there is no systematic research that quantifies the academic community's shifting responses to technological development, through a mixed-methods perspective employing sentiment, topic focus and collaboration.

While some approaches like sentiment have been successfully employed in other domains like social media and healthcare[1,2]. Sentiment analysis has not been applied in LIS as a field[3]. Similarly, topic models have been widely used in LIS research to describe topics and thematic chnges as well as research trends[4] and network analysis has illustrated collaborative structures and shared knowledge networks[5,6].

This research work seeks to tie these approaches together in one computational framework by using sentiment analysis, topic modeling (LDA and BERTopic) and network analysis (Neo4J) to explore the published output of 918 LIS publications indexed in Scopus. The key interest is to assess what and how researchers feel about emerging technologies such as artificial intelligence, automation and digital libraries, and the parallel developments of changing modes of international collaboration.

The proposed research work adds to the LIS research by allowing us to understand the emotional tone of academic discourse, the primary and dominant themes of research we are steered towards and the global academic network we can visualise and encourage further understanding about. Finally, it offers a more comprehensive vision of how epistemic claims are evolving in the context of an increasing influence of digital transformation on LIS research. The aim of this study is to understand how emerging technologies are influencing Library and

Received : 25 February 2025, Revised : 23 April 2025
Accepted : 21 May 2025, Online published : 15 July 2025

Information Science (LIS) research through a mixed-methods computational approach, including sentiment analysis, topic modeling and network analysis.

1.1 RELATED WORKS

Advanced analytical techniques like sentiment analysis, topic modeling and network analysis heavily influence the evolution of library and information science. These have been explored for new research methodologies in detecting emerging trends, analysing research patterns and improving information retrieval systems within the field.

Ketheeswaren[7] used NLP techniques which include sentiment analysis and Latent Dirichlet Allocation for topic modeling, both applied to articles indexed in Scopus database. The user-centric services, technological integration and smart library concepts are some of the major trends that are emerging and thought to be essential in the research. Another research study analysed research articles in the LIS area for over a decade, which also applied topic modeling to understand topics such as dominant ones including bibliometric indicators, technological innovation, and resource utilisation[8].

Several thorough evaluations of sentiment and emotion analysis in computational literary studies describe the different approaches used to examine the affective aspects of texts. The potential of sentiment analysis in LIS can be noted for user engagement and information retrieval purpose[9]. Another interesting application is that of LDA to analyse dissertations in LIS and find topics changing over time that may be impacting the growth of a discipline[10].

Several works have also applied topic modeling to LIS journals, analysing LIS journal publications across the years to detect prevalent topics such as bibliometrics, ICT and user studies[11]. Topic modeling has been applied in research dynamics studies in the field of LIS, to identify trends in topic transmission and influence[6]. Sentiment analysis's growing applicability across a range of fields, including LIS, is highlighted by its correlation with the development of different research topics and the venues for their publishing[12]. A review of the literature on the use of sentiment analysis in LIS studies looked at the function and difficulties of applying sentiment analysis to comprehend how users think and feel about library services[3].

Moreover, scaling up Dynamic Topic Models (DTMs) has shown promise for discovering topics and their evolutionary trends in time-series data. The following article deals with scaling up DTMs for large datasets - a topic quite relevant in the analysis of large collections of textual data that are encountered in Library and Information Science[13].

The studies together demonstrate the transformation potential of sentiment analysis, topic modeling and network analysis in LIS. These techniques enable researchers to better understand emerging trends, user behaviour, and information science development.

2. MATERIAL AND METHODS

This study uses a mixed-methods computational approach using sentiment analysis (which will explore the attitudes of researchers), topic modelling (which will reveal the changing themes of research) and network analysis (which will reveal the relationships between institutions and authors). In this way, we triangulate these methods and be able to look at how emerging technologies are being treated in a comprehensive way.

2.1 Data Collection

To investigate the technological evolution in Library and Information Science, we systematically collected data from Scopus using the search query TITLE-ABS-KEY ("library science" AND ("technology" OR "digital" OR "automation" OR "AI" OR "machine learning" OR "data science" OR "cloud computing" OR "digital libraries")), ensuring comprehensive coverage of technology-related advancements in the field. To refine our selection, we applied filters to include only journal articles, conference papers, books and book chapters, ensuring that the dataset comprised high-quality scholarly contributions. This structured search strategy allowed us to gather diverse academic sources that reflect the evolving role of technology in Library and Information Science while maintaining relevance and credibility. A dataset of 918 publications indexed in Scopus (2000-2024) was analysed and collected on January 24, 2025.

2.2 Data Processing

The dataset was loaded into Python by using the Pandas library. Preprocessing for data quality and consistency is usually done initially. Checking for missing values was a first step in this case since incomplete records will impact the results of an analysis. Using data.info() and data.isnull().sum(), missing values were identified and handled accordingly to eliminate missing abstracts to keep sentiment analysis and topic modeling integrity.

The most pertinent columns such as 'Title', 'Abstract', 'Year', 'Authors', 'Affiliation' and 'Source' were selected. Since there is more than one metadata field involved in a typical bibliographic dataset, the presence of only indispensable columns facilitates simplified analysis while limiting computational complexity. The 'Abstract' field proved to be more important since the primary text input for sentiment and topic modeling stemmed from this very field.

Handling missing author information was another critical preprocessing step. There were some records that had no author details. This could limit the scope of author-based trends and network analysis. Instead of excluding those records, missing author names were replaced with the dummy value 'Unknown' to retain valuable abstracts in the dataset. In this way, important research contributions were retained while preventing data loss. Finally, after cleaning and preprocessing, the dataset was saved as a CSV file for further analysis.

Storing the pre-processed data ensured reproducibility and easy integration with subsequent tasks, such as sentiment classification, topic modeling and network analysis. This structured approach enhanced data usability and facilitated efficient downstream processing.

The dataset, once pre-processed, was found applicable in modeling the knowledge graph network analysis in relations between authors, their affiliations and sources for their published outputs. It becomes possible to structure data into a graph, through which the academic connections can be easily explored. The core purpose is to find out patterns like co-authorship relationships and collaborations among institutions. This graph-based approach therefore, offered a much more intuitive representation of the interaction of authors, institutions and sources to provide insight into research trends, collaboration dynamics, and flows of academic knowledge. Fig. 1 shows the complete methodological part of the research from data collection to analysis.

2.3 Sentiment Scoring

The VADER Sentiment Analysis Algorithm was used to classify abstracts based on sentiment. VADER is a lexicon-based tool for sentiment analysis that assigns a compound score to text based on the intensity of words. This is computed in the following manner:

- Positive Sentiment: If the compound score is greater than 0.05.
- Negative Sentiment: If the compound score is less than -0.05.
- Neutral Sentiment: If the compound score falls between -0.05 and 0.05.

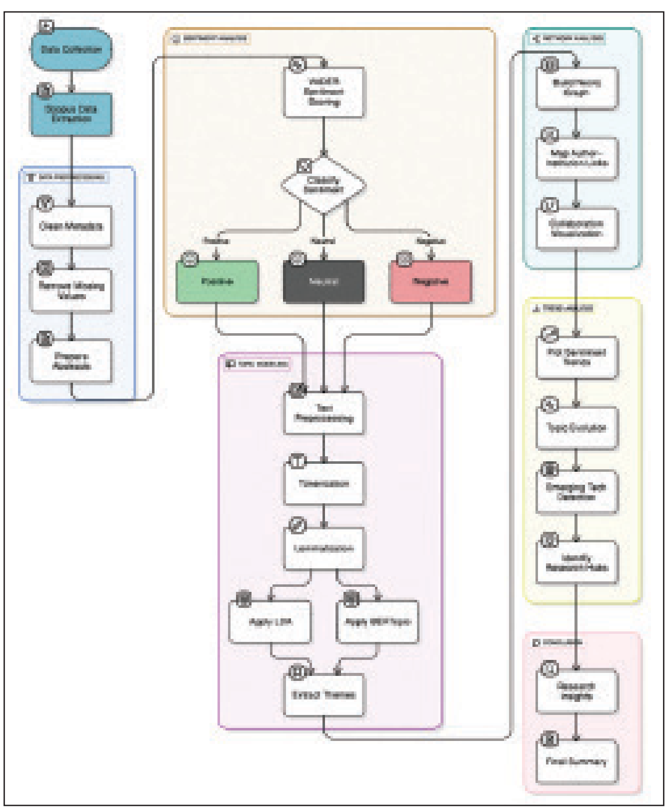

Figure 1. Data collection and analysis flowchart.

VADER is built for short texts, making it a perfect fit for the abstracts of research. It also accounts for intensity modifiers: "very good" will be more positive than "good". The implementation is achieved with the assistance of the *Sentiment Intensity Analyzer* of the NLTK library. We made use of the VADER (Valence Aware Dictionary and sentiment Reasoner) *sentiment analysis* tool, that is included with the *nltk.sentiment.vader* module (version 3.3.2) in the Natural Language Toolkit suite[14].

VADER assigns ratings to each word of the text based on a pre-defined vocabulary. Overall, the compound rating can be calculated by summing the valence ratings of each word of the text subject to the factors of negation, punctuations and intensity modifiers. The compound rating is scaled between -1 (most negative) and +1 (most positive)[1]. The compound score $S_{compound}$ is computed as:

$$S_{compound} = normalize \left(\sum_{i=1}^{n} X_i Y_i\right)$$

where:

- s_i is the sentiment score of the i_{th} word.
- w_i is the weight assigned to the i_{th} word, accounting for factors like negation and intensity.
- n is the total number of words in the text.
- The function scales the score to the range [-1, 1]. Based on the compound score, sentiments are classified as:
- Positive:
- Neutral:
- Negative

This classification allows for a nuanced understanding of the sentiment conveyed in each abstract.

2.4 Sentiment Analysis and Trend Evaluation

Once sentiment scores were assigned to each abstract, the overall sentiment distribution was analysed across different years and authors. This helped in identifying trends in research themes, such as:

- Whether research papers in certain years exhibited more positive or negative sentiments.
- Whether research sentiment changed over time, indicating shifts in perspectives or scientific discourse.

To visualise these insights, Seaborn and Matplotlib were used to generate plots, such as sentiment distribution graphs and trends over time. This provided a comprehensive understanding of how sentiment varied across research papers, helping to identify emerging patterns in scholarly work.

To analyse sentiment trends over time and among authors, the following steps were undertaken:

2.4.1 Temporal Sentiment Trends

The average sentiment score for each year was calculated to observe how the emotional tone of research has evolved over time.

$$\text{Average Sentiment per Year} = \frac{1}{N_t}\sum_{i=1}^{N_t} S_{compound,i}$$

where:

- N_t is the number of abstracts published in the year t .
- $S_{compound,i}$ is the compound sentiment score of the i_{th} abstract in year t .

This analysis helps identify shifts in the overall sentiment of research publications across different years.

The results of the sentiment trend evaluations were visualised using line plots and bar charts to illustrate the distribution and evolution of sentiments over time and among authors. These visualisations aid in comprehending the dynamics of sentiment in the research corpus.

2.5 Topic Modeling

To extract the main research themes from the dataset, topic modeling was carried out. Latent Dirichlet Allocation (LDA) was used to extract latent topics from the abstracts of research articles. LDA is a typical probabilistic generative model assuming that documents are mixtures of topics, each characterised by a bag of words and corresponding probabilities. The sequences for LDA modeling consist of several preprocessing steps, namely text tokenisation, vectorisation, LDA modelling and topic interpretation[15].

2.5.1 Text Tokenisation

Tokenisation refers to the process of breaking down a text into discrete smaller units, most commonly words, to facilitate numerical representation for computational modeling. Prior to topic modeling, the abstracts were tokenised into single words using the Natural Language Toolkit (NLTK) and spaCy libraries[16].

Further, stop word removal included the finishing touches. This involved removing from miscellaneous documents further common words of little inherent order, such as “the”, “is” and “and” from default stopword lists provided by the NLTK and spaCy libraries. The removal of stop words is expected to increase the performance of efficient topic modeling by eliminating irrelevant common words. Punctuation and special characters were removed as any further noise.

Another important preprocessing step was lemmatisation, where words were converted to their base or root form (e.g., “running” → “run”). This ensures that different variations of a word are treated as the same, reducing redundancy in the dataset.

2.5.2 Vectorisation (TF-IDF Representation)

Once the text was cleaned and tokenised, it was converted into a numerical format using Term Frequency-Inverse Document Frequency (TF-IDF) vectorisation. TF-IDF is a statistical measure that reflects how important a word is to a document relative to the entire corpus15.

The TF-IDF score for a term in a document is calculated as follows:

$$\text{TF-IDF}(t,d)=\text{TF}(t,d)\times\text{IDF}(t)$$

where:

Term Frequency (TF): The number of times term t appears in document d.

$$TF(t,d) = \frac{f_{t,d}}{\sum_k f_{k,d}}$$

where $f_{t,d}$ is the frequency of the term in document d and the denominator represents the total number of words in the document d.

- Inverse Document Frequency (IDF) measures how important a word is across all documents.

$$\text{IDF}(t) = \log \frac{N}{1+n_t}$$

where N is the total number of documents and is the number of documents containing the term t . A logarithmic function is applied to prevent extreme values.

TF-IDF helps reduce the influence of frequently occurring words that may not be topic-specific while giving more weight to rare but significant words. This ensures that the most meaningful terms contribute effectively to the topic modeling process.

2.5.3 Latent Dirichlet Allocation (LDA) Implementation

Latent Dirichlet Allocation (LDA) is a probabilistic topic modeling algorithm that assumes documents are mixtures of topics, where each topic is defined by a distribution of words. The goal of LDA is to determine the latent (hidden) topics that best explain the observed word distributions in a corpus[17].

Mathematically, LDA models the probability of the word w appearing in a document d as:

$$P(w|d)=\sum_{k=1}^{K}P(w\mid z_k)P(z_k\mid d)$$

where:

- K is the number of topics.
- Z_k represents a latent topic.
- $P(w|z_k)$ is the probability of word w belonging to topic Z_k.
- $P(z_k|d)$ is the probability of topic Z_k appearing in document d.

LDA follows a Dirichlet prior distribution, which encourages sparsity in topic assignments, meaning that each document is dominated by a few topics rather than all topics equally. This makes the model interpretable and practical for analysing research themes.

The Gibbs Sampling[18] technique was used to estimate the topic distributions efficiently. The number of topics K was chosen based on coherence scores, which measure how semantically meaningful the extracted topics are.

The Gensim library in Python was used to implement LDA, as it provides an optimised and scalable implementation for large datasets.

2.5.4 *Topic Interpretation*

Once the LDA model extracted topics, the most significant words in each topic were analysed to interpret the underlying research themes. This involved:

- Ranking the most frequent words in each topic to identify dominant themes.
- Visualising word distributions using word clouds and bar charts.
- Assigning meaningful labels to topics based on their word compositions.

For example, if a topic contained words like “machine learning,” “artificial intelligence,” “neural networks”, it could be labeled as “AI and Machine Learning Research.”

To ensure the extracted topics were meaningful, topic coherence scores were calculated. Coherence measures how often words in a topic appear together in documents[19]. A common measure is C_v coherence, defined as:

$$C_v = \frac{1}{n}\sum_{i<j} similiarity(w_i, w_j)$$

Where w_i, w_j are words in a topic, and similarity measures their co-occurrence frequency. A higher coherence score indicates that the words in a topic are more semantically related, leading to more interpretable topics.

By applying this structured approach, we successfully identified and interpreted the key research themes in the dataset.

By leveraging TF-IDF vectorisation, LDA modeling and coherence scoring, we extracted meaningful topics from the dataset. This analysis provided insights into prominent research themes, allowing for a better understanding of how research interests have evolved over time.

2.6 Knowledge Graph

A knowledge graph is a structured representation of information, organised in nodes and edges, representing entities and relationships between them. This organised form allows interconnected data to be stored, retrieved and analysed efficiently[20].

Formally, a knowledge graph can be represented as:

$$G=(V,E)$$

where V is the set of nodes (entities) and E is the set of edges (relationships).

We selected Neo4j Desktop v5.11, a native graph database platform that is optimised for connected data, to build and analyse the knowledge graph. Neo4j was selected for its unique capability to model complex relationships easily with the Cypher query language, as well as for its ability to work with Python through the py2neo and Neo4j Python Driver packages[21].

Table 1. Sentiment counts of scholarly publications.

S . No.	Sentiment	Counts
1	Positive	767
2	Negative	61
3	Neutral	84

3. ANALYSIS AND RESULTS

Rapid changes in technology have significantly impacted library science, influencing research priorities, service delivery and user engagement. This study uses sentiment analysis and topic modeling to analyse how the discussions in the field have evolved over time. A knowledge graph is also used to visualise collaborations and the dissemination of ideas among leading scholars.

3.1 Sentiment Analysis of Research Trends

3.1.1 *Overall Sentiment Distribution*

Table 1 depicts the sentiment counts of scholarly publications, in which most of the publications were found to be positive (767 articles).

Only very few of them were found to be neutral (84) or negative (61). This suggests a broad optimism about library science research.

The composite visualisations labelled in the figures allow for a comprehensive analysis of the sentiment trends within the dataset. Fig. 2 gives a general overview of the distribution of sentiment scores and shows an impressive skew toward being positive, with most scores ranging from 0.5 to 1, while negative sentiments rarely appear. Fig. 3 uses a box plot to analyse yearly sentiment tendencies, indicating consistency in positive tendencies across years; however, those small circles within the figure suggest outliers and thereby deviations or extreme cases at specific times. Fig. 4 captures the dynamics of average scores over time and shows how scores were fluctuating in early years, dropping down to negative, then becoming a steady increase towards the 1990s with stabilisation at the positive level afterwards. The aggregate graphs suggest that the dataset trends positively, evolving from early volatilities toward more stable, optimistic tone lately.

3.1.2 *Most positive and Negative Research Themes*

Table 2 shows the top 5 most positive sentiments focusing on the futuristic roles of libraries, competitive intelligence and the empowerment of library development are most positively perceived. Further, research in the areas of integrating information literacy into multidisciplinary education, high-quality mechanisms for library development and the evolving mission of libraries in the new era are also well received. The

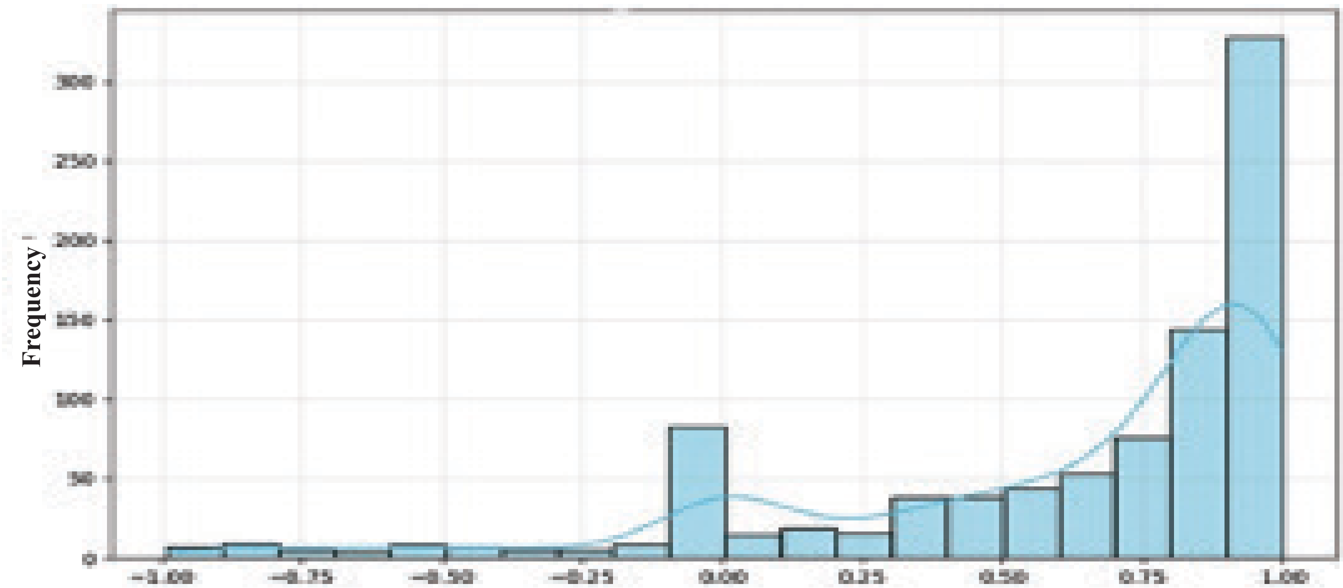

Figure 2. Distribution of sentiment scores.

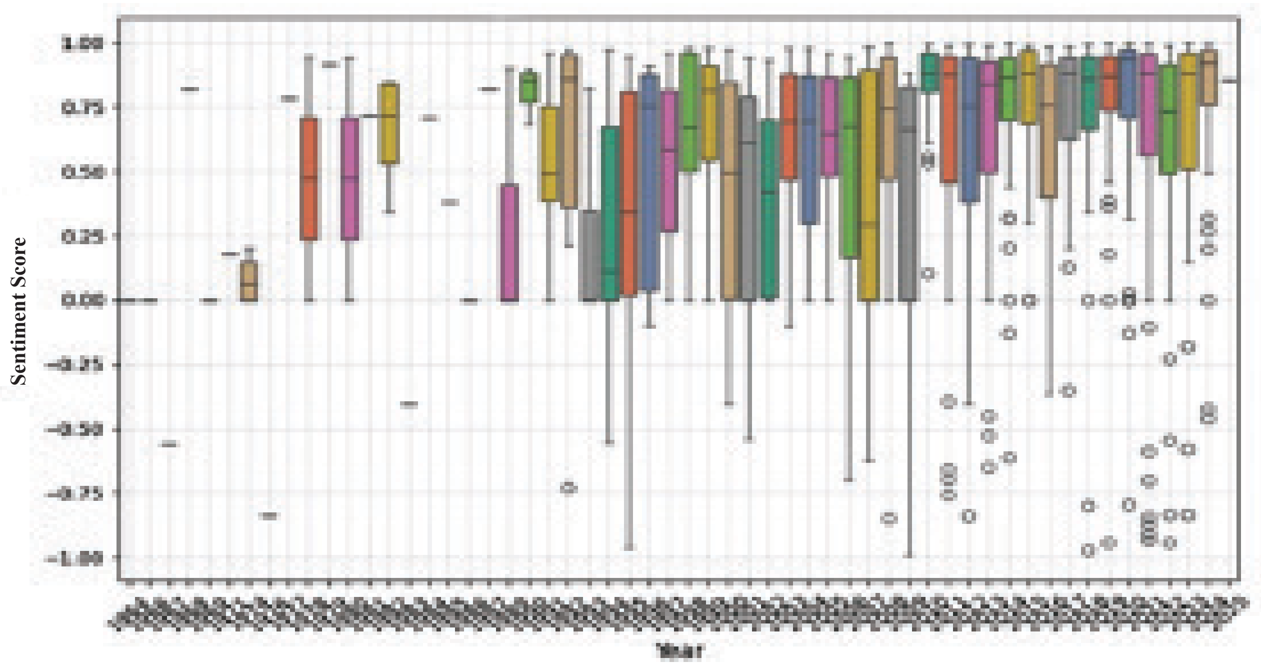

Figure 3. Sentiment trends over years.

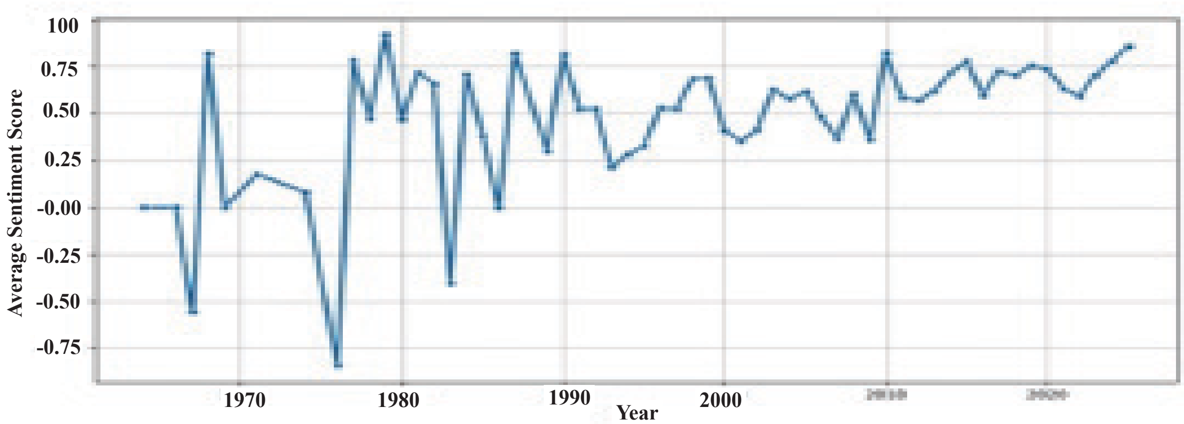

Figure 4. Sentiment trends over time.

studies on the role of libraries in the shaping of future societal structures, adoption of innovative technologies and enhancement of user experiences through AI-driven services are generally positively received.

Instead, negative sentiments exist in research focused on fraud detection, ergonomic problems in digital libraries and fake news and its implications on spreading misinformation, as displayed in Table 3. Further, studies conducted on pathology labelling errors, issues with the efficacy and safety of certain medical interventions and concerns over the morality of machine learning applications in detecting fraud are seen with scepticism as well. Most often, research in digital library ergonomics in terms of the physical and cognitive burden and how libraries can counter misinformation in the post-truth era is highly critical, because it is perhaps a very difficult area to solve.

3.1.3 Sentiment Trends Across Technologies

Table 4 illustrates sentiment on key technological themes reveals that those related to AI (+0.63) and digital libraries (+0.65) are greeted positively, indicating hope for both improving the efficiency of obtaining information, the better user experience they offer and general research capabilities. In this regard, AI-based indexing technologies, means of extracting knowledge, and recommender systems are welcomed tools. Digital libraries also evoke support in the provision of preservative and democratised access to knowledge, particularly concerning open-access programs.

Automation (0.54) again brings mixed emotions to the fore, possibly due to fears about job loss and the library's requirement for human intervention in general. Although automation may improve cataloguing of books, circulation and retrieval of information, there are issues with automated processes harming the long-standing role of librarians and system failure, a condition that demands up skilling library professionals in this regard.

The highest score of sentiment, which indicates the strong advocacy of ethical practice in library science, is that of privacy and ethics (0.85). Overall, a highly positive sentiment means that the library and information science community have great support for data privacy, user confidentiality and the implementation of ethical AI. Protecting user data, preventing bias in AI-driven decision-making and providing equitable access to digital resources all reflect core values of the profession.

Table 2. Top 5 most positive sentiments

Title	Sentiment score
Looking into the future library and society	0.9983
On the mission of library in the new era and the education of library science in the future	0.9982
How to herd cats: developing a playbook for competitive intelligence gathering at technical conferences	0.9979
Works in progress: integrating information literacy into a multi-disciplinary first-year engineering program	0.997
Empowerment mechanism of library high-quality development	0.9965

Note: Even though the third article is from a technically oriented domain rather than a traditional library and information science domain, we included it because it has been indexed in "Information Management and Systems" and relates methodologically to important themes in information analysis, competitive intelligence, and knowledge organization which are increasingly of interest to the library and information science community.

Table 3. Top 5 most negative sentiment

Title	Sentiment score
A quality initiative to decrease pathology specimen-labelling errors using radiofrequency identification in a high-volume endoscopy center	-0.9895
A state-of-the-art review of machine learning techniques for fraud detection research	-0.9729
Ergonomics in the electronic library	-0.9618
The role of libraries in the fake news era: a survey of information scientists and library science students in Greece	-0.9404
Clinical efficacy and safety of xiaoyao pill in post-stroke depression: A systematic review and meta-analysis of randomised controlled trials	-0.9403

Note: Although the first and last articles in the table are mainly concerned with medicine and pharmacy topics, they were included in the dataset because they used common keywords (RFID, data accuracy, and information systems) that were related to LIS. They were not intended for the LIS dataset, but their appearance reflects the interdisciplinary nature of applied information technologies and the elasticity of information science methods.

Overall, this analysis shows that even while people typically have positive opinions about developing technologies, automation issues and ethical considerations continue to be major topics of discussion in the library science community.

As presented, the two figures give some insight into the trend of sentiments across the different technologies, specifically in AI and Digital Libraries. Fig. 5 presents a comparative bar chart of average sentiment scores, showing that "Privacy & Ethics" has the highest sentiment followed by Digital Libraries and AI while Automation scored the lowest. This means that topics being discussed on privacy and ethics

Table 4. Sentiment on key technological themes

S . No.	Technology	Average sentiment score
1	AI	0.631149
2	Digital libraries	0.647922
3	Automation	0.538331
4	Privacy & ethics	0.846347

are more positively seen than those on automation. Fig. 6 uses sentiment tracking over time to outline fluctuating patterns in perception for AI and Digital Libraries. Digital Libraries have had quite stable, positive sentiment since at least the early 1990s, while AI's sentiment has been much more volatile, with pronounced dips around 2005 and 2020. This could reflect the burst of technological progress, ethics concerns, or societal debate over the effects of AI. Together, they reflect the very changing discourse relating to these technologies and their consequent impact upon research and diffusion.

3.2 Introduction to Topic Modeling in LIS Research

One application of natural language processing is technique used to disclose hidden themes, which are formed within large-text datasets. Within the context of bibliometric analysis, it seems to be significantly useful in pinpointing emerging themes and thematic structure in academic research. Latent Dirichlet Allocation (LDA) forms one of the most widely adapted topic modeling procedures, which applied to our LIS dataset to highlight key topics formed in LIS-based research. LDA works on the idea of co-occurrence of words found in documents and groups them according to coherent topics, based on probability distributions.

Table 5 depicts topic modeling of titles and abstracts from publications in the research, leading to the identification of five major themes. Each theme represents one distinct area of interest for LIS indicating how priorities in research have shifted over time as well as which forms of technological advancements are changing the course of the field.

3.2.1 Key Themes Identified in LIS Research

3.2.1.1 Library and Information Science Research

This topic captures the fundamental principles and theoretical frameworks of LIS. The research under this theme includes that on library science as a discipline, information-seeking behavior and the evolution of information retrieval systems. Studies under this theme also discuss the role of libraries in knowledge organisation, the impact of scientometric and bibliometric methods and trends in scholarly communication. This is a high score in the sentiment category, indicating that researchers generally

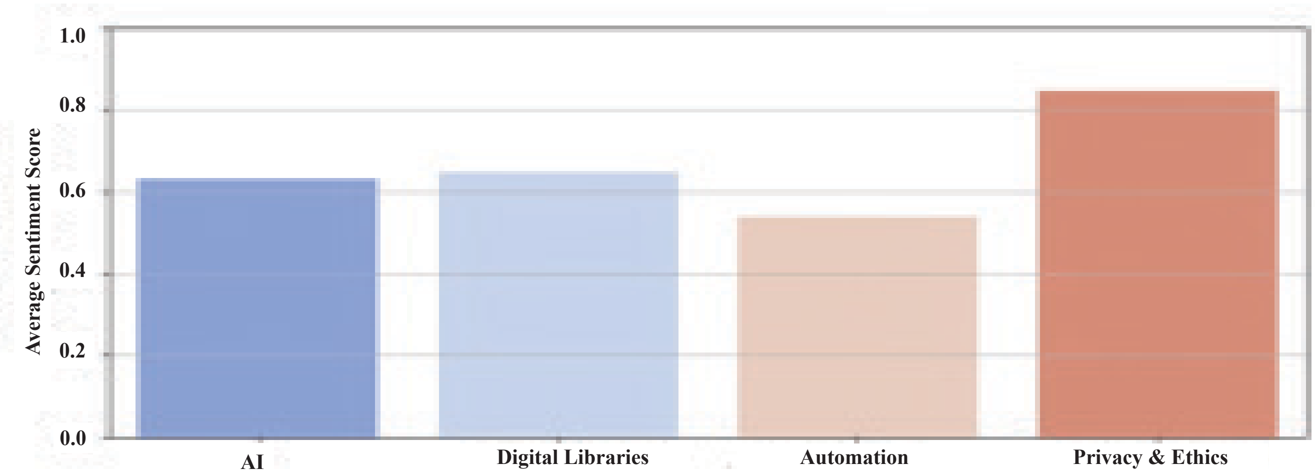

Figure 5. Sentiment scores across technologies.

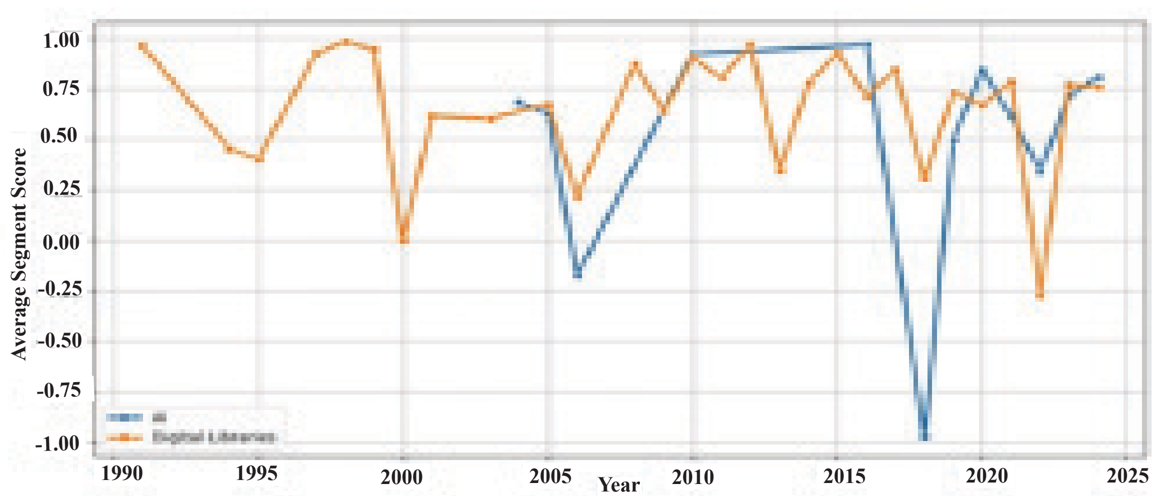

Figure 6. Sentiment trends over time for AI and digital libraries.

hold a positive view of the growth and influence of LIS as a discipline.

3.2.1.2 Study and Research in Information Science

This theme discusses methodological progress and research methods in LIS. Studies in this category include data-driven decision-making, quantitative and qualitative research methods and artificial intelligence in information retrieval. The researchers in this theme analyse how data science, machine learning, and computational tools are being utilised to enhance the services of the library and improve the user experience. The sentiment analysis is moderately positive for this topic, which indicates an optimistic view of technological progress in LIS research.

3.2.1.3 Library Services and Technology

This topic revolves around the technological evolution of services that are carried out in the libraries. Its sections include discussions of digital literacy, automation and user outreach using digital methods. Specific research deals with issues concerning the acceptance of AI-powered recommendation systems by the library, cataloguing and referencing using the RFID system and the reference desk service rendered using chat bots. Sentiment analysis shows that, while most new technologies (e.g., AI, digital libraries) yield positive attitudes, automation is a more mixed experience. Library professionals are concerned about job security, the devaluing of human knowledge and the over-reliance on machines. Library professionals express concern that automation will take the place of traditional library roles, like cataloguing and assistance to patrons, which influences their more cautious and skeptical sentiments.

3.2.1.4 Digital Libraries and Academic Research

This theme is focused on the integration of digital libraries in academic research and scholarly communication. It has institutional repositories, open-access publishing and the long-term preservation of digital resources discussed in it. The studies here discuss metadata management, interoperability between digital library systems and the role of AI in improving search and retrieval functionalities. Given the increase in reliance on digital repositories in academic research, this theme would have a higher sentiment score as it shows utmost support for the digital transformation in libraries.

3.2.1.5 Information, Library Science, and Digital Technology

This category is the broader reflection of digital technology integration in LIS. It contains topics on big data analytics, blockchain applications in libraries, and ethical issues associated with AI in information systems. Research under this theme explores the bias of search engine algorithms, concerns about data privacy, and how libraries contribute to fighting misinformation. This score shows that this theme has the strongest sentiment and that LIS will be heavily emphasised in the future in terms of ethical and technological progress.

The three visualisations, in total, provide insights about the dataset. Fig. 7 is a pie chart to show the general proportion of topics indicates the presence of Topic 3 as the dominant topic, which occurs 30 %. Topics 1 (22.4 %) and Topic 2 (20.1 %) come next in this order, but Topics 4 (18.5 %) and 5 (9.0 %) have low contributions. Fig. 8 a horizontal bar chart, emphasises the prevalence of topics in abstracts, which reiterates the dominance of Topic 3 with 29.78 %, followed by Topics 1 with 22.27 % and Topic 2 with 19.93 %, making it easier to compare the relative significance of each topic. Fig. 9 another horizontal bar chart, further confirms this trend by showing the average proportion of topics across documents, maintaining the same ranking of prevalence. Altogether, these visualisations give a holistic view of the topic modeling results, validating the prominence of key themes in the dataset.

3.2.2 Sentiment Analysis of Topics

Table 6 illustrates results of the sentiment analysis of topics further reveal how scholars perceive these research areas. The highest sentiment score is associated with discussions on privacy, ethics and digital libraries, which reflect strong positive engagement with ethical and accessibility concerns in LIS. On the other hand, topics related to automation and AI adoption show more mixed sentiments, indicating that while these technologies bring efficiency, they also raise questions about job displacement and user trust.

The Fig. 10 demonstrates the average scores for the topic's sentiment. Variations are noticed in the overall sentiment linked with each topic. Topic 1 has the lowest sentiment score, meaning a neutral or potentially negative discussion; however, topics 2, 3, and 4 display a moderate increase in the score, showing a more balanced or slightly positive tone. The highest value for sentiment score was

Table 5. Key themes identified from topic modeling of publication titles and abstracts

Topic	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8	Word 9	Word 10
Topic 1	library	science	information	research	technology	service	resource	development	librarian	digital
Topic 2	study	research	information	data	science	used	using	library	result	field
Topic 3	library	science	information	journal	article	technology	service	research	study	reference
Topic 4	library	information	science	research	digital	study	technology	paper	academic	university
Topic 5	information	library	science	librarian	data	digital	technology	study	knowledge	student

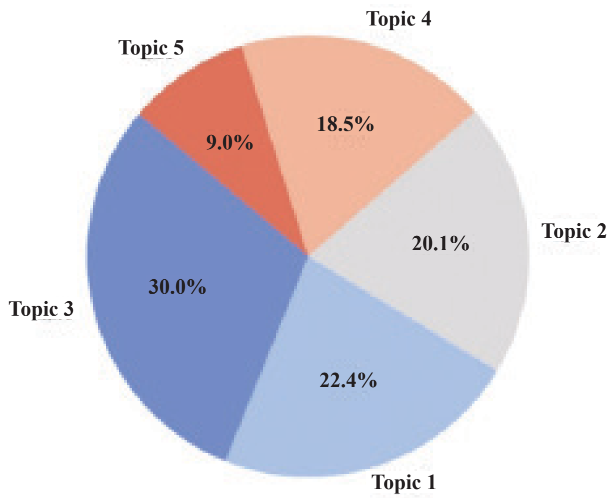

Figure 7. Proportions of topics in the dataset.

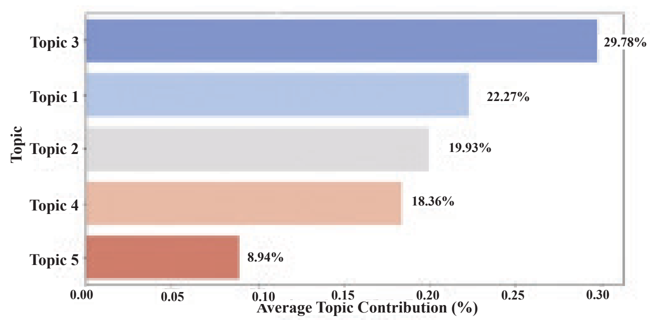

Figure 8. Topic prevalence in abstracts.

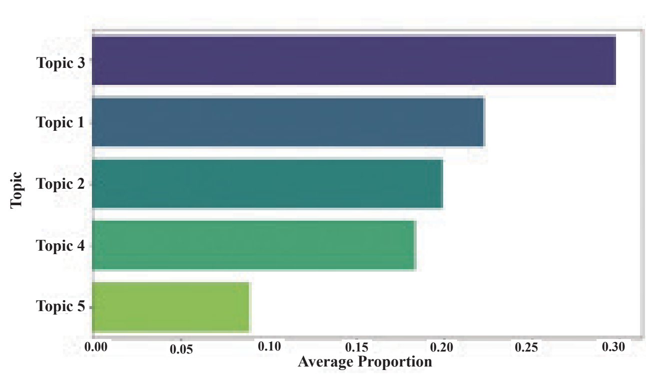

Figure 9. Topic distribution across documents.

noted in Topic 5. This means that the discussions in this category are generally more positive. The visualisation shows how the emotional tone of different research themes may vary, providing insights into how different topics are perceived within the dataset.

Topic modeling results provide an exhaustive overview of trends in LIS research, under the umbrella of technological advancements, digital transformation, and evolving methodologies. The results clearly show that the adoption of new technologies in libraries is increasing with a lot of emphasis on ethics, accessibility and user experience. Therefore, future research is bound to address these issues between innovation and ethics for the continuation of inclusive, adaptive, and forward-thinking libraries in terms of technology.

3.3 Knowledge Graph Analysis

In this section, we lay out the fruit of our exploratory analysis involving the 15 most relevant writers in the specific field, drawing attention to interplay, writings and institute affiliations. The knowledge graph represents the relationship clearly between authors and their papers as well as those institutions to which they belong, and sources at which their works are published.

3.3.1 Relationships Between Nodes

The knowledge graph establishes connections between important entities: Authors, Papers, Institutions and Sources. There are three primary relationships represented:

- WROTE: This relationship connects an author with their respective papers. For instance, YANG S has authored numerous studies focusing on AI adoption in library settings.
- AFFILIATED_WITH: This links authors to the institutions they are associated with. For example, YANG S is linked to Wuhan University, a major contributor frequently seen in prominent publications.
- PUBLISHED_IN: This ties papers to the journals or conferences where they have been published, including notable platforms like Library Hi Tech and the International Journal on Digital Libraries.

Table 6. Sentiment analysis of topics

S . No.	Topic	Average sentiment score
1	Topic-1	0.406332
2	Topic-2	0.605199
3	Topic-3	0.622903
4	Topic-4	0.644703
5	Topic-5	0.693388

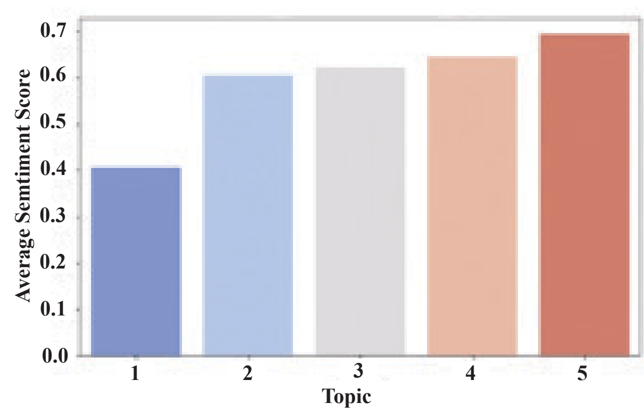

Figure 10. Average sentiment score by topic.

3.3.2 Key Insights from the Knowledge Graph

The Fig.11 knowledge graph illustrates the relationships between authors, their institutions, papers and sources. Nodes are different entities: authors, institutions, papers and sources, while edges capture relationships such as 'wrote,' 'affiliated with' and 'published in.' The graph has 153 nodes and 178 relationships, with color-coded nodes: authors in blue, papers in orange, affiliations (institutions) in yellow, sources in green, and years in violet.

The visual representation of the knowledge graph showed specific patterns related to collaboration, institutional influence and leadership and the themes of research that comprise the field of LIS:

- Institutional Influence: The graph shows that Wuhan University was the most central institution on the graph due to a high number of authors discussing a few high-sentiment papers on digital libraries and AI (e.g., YANG S). This suggests that Wuhan University is a centre of innovation in technology-driven LIS research.
- Research clusters/Thematic Direction: Authors from Florida Atlantic University and International Hellenic University have recently collaboratively published multiple papers on ethical AI and misinformation, which establishes a clear institutional thematic specialisation.
- Core Publication Venues: The journals Library Hi Tech represent most of the high-collaborative published papers, which establishes venues as the primary distribution points for LIS technology research.
- Strength of Collaboration: Several authors shared authorship on papers with institutional associations, thus establishing large clusters of collaboration. This could suggest a strong degree of intra-institutional cooperative effort and respective centres of excellence in LIS.

3.3.3 Temporal Shift in Collaborations

The network analysis further highlights a significant temporal trend in collaboration. In the early 2000s, most research in LIS was institutionally bounded with minimal cross-institutional involvement. From 2010 onwards, there has been a marked increase in multi-institutional and international collaboration and institutions like Wuhan University and Florida Atlantic University had increasingly more co-authored papers with partners in other countries, particularly in research related to AI and digital transformation, reflecting the globalisation of LIS research and its positioning with interdisciplinary research fields, such as computer science and data analytics.

In conclusion, the graph illustrates the LIS research landscape as a constellation of prominent institutions, directed themes of research and repeated collaborations. Institutions such as Wuhan University represent not only a few frequent contributors, but also direction of themes emerging in LIS technology research.

4. DISCUSSION

The findings of the work lend valuable insights into how technological advancements are shaping the field of LIS. Application of sentiment analysis found that a majority of users hold a very optimistic view about emerging trends in the form of artificial intelligence, digitisation of libraries and methodologies based on data. While automation can enhance all processes relating to the cataloguing of books, circulation and address the retrieval of information, automation has a potential problem from the perspective of the long-standing role of librarians, which may coincide with system failure, which creates a necessity to up skill library professionals. The sentiment trends over time indicate that there is a growing acceptance of technological innovations, with early fluctuations giving way to a more stable and optimistic perception of technology's role in LIS.

Topic modeling results show that research in LIS is increasingly focused on digital transformation, technological adoption and ethical considerations. The key themes identified Library Science and Information Research, Study and Research in Information Science, Library Services and Technology, Digital Libraries and Academic Research and Information, Library Science and Digital Technology show a progressive shift towards integrating digital tools and AI-driven solutions in library services. This evolution hints toward a future when libraries would operate as modern, technologically equipped institutions supporting the diffusion of knowledge in increasingly efficient ways.

The knowledge graph analysis provided relevant insights into the collaborative landscape of LIS research. Visualisation of relations among authors, institutions and sources indicates a high level of institutional influence in research directions due to academic collaborations. It has been noted that the leading institutions are Wuhan University and Florida Atlantic University as leaders in research and innovations, while the analysis indicated a great deal of continued reliance on AI, digital libraries and ethics, there are still untapped bigger challenges around automation and its disruptive action that could potentially lead to unemployment, misinformation and privacy. Hopefully future research will focus on overcoming these challenges through the use of technology in the LIS space.

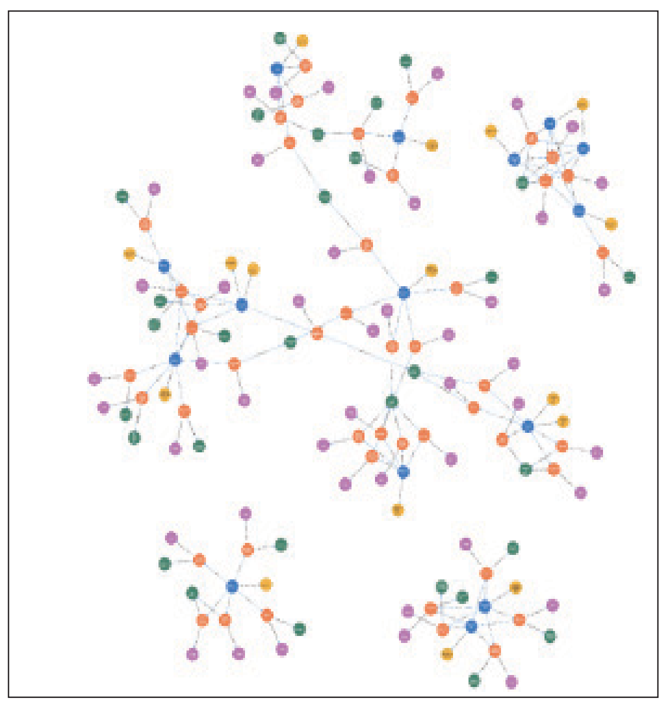

Figure 11. Knowledge graph showing connections between authors, institutions, papers, and sources.

5. CONCLUSION

The research combines sentiment analysis, topic modeling and network analysis, providing a complete picture of technological advancement in Library and Information Science. The overall results show a general acceptance of emerging technologies as the positive trend in sentiment indicates optimism regarding digital libraries, AI and research-centric inventions. However, issues with automation and ethical dilemmas continue to be a conversation we will need to maintain balance in the how we integrate technologies. The research shows a need for digital transformation in LIS studies more than time and increasingly demonstrates how sentiment trends and research themes changed over time. The inclusion of the knowledge graph analysis clearly indicates collaboration across various academics, strong institution involvement, and collaborative research which indeed portrays the blend of LIS research being interconnected and technology driven.

Overall, the research adds to the understanding of ways that LIS is addressing the digital age and highlights the need for responsible innovation, conforming to ethical practice and in service of users' needs. Future research should think about the strategic integration of emerging technologies in libraries and discuss issues related to automation, privacy, and the role of information professionals. Therefore, the research provides a way for future work on technology and LIS engagements and therefore, the future in the field could be effective and ethical.

REFERENCES

1. Hutto C, Gilbert E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proc Int AAAI Conf Web Soc Media Internet. 2014 Jun 5;8(1):21625. https://ojs.aaai.org/index.php/ICWSM/article/view/14550
2. Wang K. Community relations discovery methods for users in Fancircle based on sentiment analysis in China. 2024 Jan 29. https://www.emerald.com/insight/content/doi/10.1108/dta-09-2023-0570/full/html
3. Kumar V. Exploring the Use of Sentiment Analysis in Library User Studies: Approaches and Challenges. In 2023. p. 446–57. https://www.researchgate.net/publication/370058798_Exploring_the_Use_of_Sentiment_Analysis_in_Library_User_Studies_Approaches_and_Challenges
4. Lamba M, Madhusudhan M. Topic Modelling and its Application in Libraries: a review of specialised literature - IOS Press. https://content.iospress.com/articles/world-digital-libraries-an-international-journal/wdl15207
5. Lee YJ, Park SE, Lee SY. Machine Learning-Driven Topic Modeling and Network Analysis to Uncover Shared Knowledge Networks for Sustainable Korea–Japan Intangible Cultural Heritage Cooperation. Sustainability [Internet]. 2024 Dec 11;16(24):10855. https://www.mdpi.com/20711050/16/24/10855
6. Yan E. Research dynamics, impact, and dissemination: A topic-level analysis. J Assoc Inf Sci Technol [Internet]. 2015;66(11):2357–72. https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.23324
7. Ketheeswaren S. Evolving Landscape of Smart Libraries: A Diachronic Analysis of Themes and Trends. Tech Serv Q [Internet]. 2024 Oct 1;41(4):333–50. https://doi: 10.1080/07317131.2024.239497
8. Majhi D. Analysing Library and Information Science Articles Using Topic Modeling Approaches. https://www.academia.edu/124125107/Analysing_Library_and_Information_Science_Articles_Using_Topic_Modeling_Approaches
9. Kim E, Klinger R. A Survey on Sentiment and Emotion Analysis for Computational Literary Studies. Z Für Digit Geisteswissenschaften [Internet]. 2019. http://arxiv.org/abs/1808.03137
10. Sugimoto CR, Li D, Russell TG, Finlay SC, Ding Y. The shifting sands of disciplinary development: Analyzing North American Library and Information Science dissertations using latent Dirichlet allocation. J Am Soc Inf Sci Technol [Internet]. 2011;62(1):185–204. https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.21435
11. Lamba M, Madhusudhan M. Mapping of topics in DESIDOC Journal of Library and Information Technology, India: a study. Scientometrics [Internet]. 2019 Aug 1;120(2):477–505. doi: 10.1007/s11192-019-03137-5
12. Mäntylä MV, Graziotin D, Kuutila M. The Evolution of Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers. Comput Sci Rev [Internet]. 2018 Feb;27:16-32. http://arxiv.org/abs/1612.01556
13. Bhadury A, Chen J, Zhu J, Liu S. Scaling up Dynamic Topic Models [Internet]. arXiv; 2016. http://arxiv.org/abs/1602.06049
14. Hutto CJ. cjhutto/vaderSentiment [Internet]. 2025. https://github.com/cjhutto/vaderSentiment
15. Manning C, Raghavan P, Schuetze H. Introduction to Information Retrieval. Camb Univ Press [Internet]. 2009;496. https://nlp.stanford.edu/IR-book/information-retrieval-book.html
16. spaCy 101: Everything you need to know · spaCy Usage Documentation [Internet]. spaCy 101: Everything you need to know. https://spacy.io/usage/spacy-101
17. Blei DM. Latent Dirichlet Allocation. Journal of Machine Learning Research 3 [Internet]. 2003;993–1022. https://dl.acm.org/doi/10.5555/944919.944937
18. Bendimerad A, Lijffijt J, Plantevit M, Robardet C, De Bie T. Gibbs Sampling Subjectively Interesting Tiles. In 2020. https://link.springer.com/chapter/10.1007/978-3-030-44584-3_7
19. Mimno D, Wallach H, Talley E, Leenders M, McCallum A. Optimising Semantic Coherence in Topic Models. In: Barzilay R, Johnson M, editors. Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing [Internet]. Edinburgh, Scotland, UK.: Association for Computational Linguistics; 2011. p. 262–72. https://aclanthology.org/D11-1024/

20. Enzo. What Is a Knowledge Graph? [Internet]. Graph Database & Analytics. 2024. https://neo4j.com/blog/what-is-knowledge-graph/
21. Download Neo4j Desktop [Internet]. Graph Database & Analytics. https://neo4j.com/download/

CONTRIBUTORS

Mr. M. S. Rajeevan is Research Scholar in the Department of Library and Information Science, University of Kerala, with proficiency in digital research methodologies and computational analysis.
He conceptualised and designed the study; collected and preprocessed data from Scopus; did sentiment analysis and topic modeling; created visualizations and drafted the initial version of the manuscript

Dr. B. Mini Devi is assistant professor and Director of the Centre for Information Literacy Studies at University of Kerala, where she has also held the position of Head of the Department. She obtained her PhD from Cochin University of Science & Technology. She also serves on several academic boards, an active researcher and has published more than 180 articles, authored 11 books and 24 research scholars have been awarded PhD under her supervision.
She provided overall guidance and intellectual direction for the study, supervised the research design and analytical framework, made substantial contributions to redefining the research objectives and project outline, verified academic integrity and provided critical review and substantive editing on all versions of the manuscript.

Dr. Anoop V.S. is Assistant Professor (Senior Grade), Amrita School of Business, Amrita Vishwa Vidyapeetham, Kochi, India. He has a PhD in Artificial Intelligence from Cochin University of Science and Technology and completed his postdoctoral fellowship in Canada from Queen's University. He specializes in AI, NLP, and Computational Social Science. He has published books, book chapters and many journals and conference papers in terms of research outputs and was awarded Responsible Computing Research Grant by Mozilla Foundation and USAID.
He participated in the design and implementation of computational techniques (sentiment analysis, topic modeling, and network analysis) as well as the analytical insight of study components related to AI.

Dr. C. Mallikarjuna has over 14 years of experience as an academic librarian and is currently the Deputy Librarian at the Indian Institute of Technology (IIT), Hyderabad. He received his PhD in Library and Information Science from Sri Venkateswara University.
He contributed to the literature review, provided domain specific validation of the results, and helped to align the project with modern trends in library and information science research.