

REVIEW PAPER

Mapping the Evolution of Doctoral Research in Information Science: A Data-Driven Perspective from Indian Repositories

Ajay Misra^{1*} and Jayati Lahiri Dey²

¹Ph.D. Scholar, Department of Computer and Information Science, Raiganj University, Raiganj, Uttar Dinajpur, West Bengal, India

²Assistant Professor, Department of Computer and Information Science, Raiganj University, Raiganj, Uttar Dinajpur, West Bengal, India

*Corresponding author: dy_librarian@raiganjuniversity.ac.in

Received: 11 Apr., 2026

Revised: 03 May, 2026

Accepted: 22 May, 2026

ABSTRACT

Doctoral research in Information Science (IS) in India has grown substantially over the past four decades, propelled by the rapid digital transformation of libraries, the advancement of information and communication technologies, and the emergence of nationwide open-access scholarly ecosystems such as Shodhganga, the reservoir of Indian theses hosted by INFLIBNET. This paper presents an expanded, data-driven examination of the evolution of IS doctoral research in India, drawing on metadata relating to year of submission, thematic keywords, institutional affiliation, geographical distribution, and supervisory patterns across four historical phases. Employing descriptive bibliometric indicators, keyword clustering, and network-mapping techniques, the study traces the discipline's transition from classical librarianship toward digital, data-intensive, and interdisciplinary domains. Findings indicate a marked rise in research productivity since the late 1990s, a pronounced shift toward artificial intelligence, big data, and research-data-management themes after 2016, growing co-supervision and inter-institutional collaboration, and persistent regional imbalances favouring southern and northern India. The revised methodology strengthens data validation, triangulates multiple repositories, and clarifies analytical procedures. The study offers evidence-based insights that can inform curriculum redesign, funding policy, supervisory planning, and strategies for balanced regional development, contributing to a richer understanding of how Information Science is maturing in India.

Keywords: Information Science; doctoral research; Shodhganga; bibliometric analysis; India; digital repositories; scientometrics

Information Science has emerged as a dynamic and interdisciplinary field that intersects library studies, computer science, communication, data analytics, and knowledge management. Over the years, the role of information professionals has transformed from custodians of physical collections into managers of complex digital ecosystems that support discovery, curation, and analysis at scale. In India, this transformation has accelerated markedly since the expansion of information and communication

How to cite this article: Misra, A. and Dey, J.L. (2026). Mapping the Evolution of Doctoral Research in Information Science: A Data-Driven Perspective from Indian Repositories. *IJISC*, 13(01): 69-85.

Source of Support: None; **Conflict of Interest:** None



technology into academic and research spaces during the early 2000s. As a result, doctoral research in Information Science, widely regarded as the highest form of academic and scholarly contribution, offers crucial insight into the intellectual progression of the discipline and into the priorities that will shape its next phase of growth.

Doctoral theses serve as reflections of academic maturity, methodological development, and thematic transition within a field. They capture not only what a discipline has already achieved but also the direction in which future scholarship is likely to move. Understanding how doctoral research has progressed over successive decades allows policymakers, educators, and researchers to identify strengths, gaps, and emerging opportunities that might otherwise remain invisible in fragmented, institution-specific accounts.

India's national initiative toward digitisation, combined with the University Grants Commission mandate for online thesis submission under the UGC (Minimum Standards and Procedure for Award of M.Phil./Ph.D. Degrees) Regulations, has dramatically improved access to scholarly content (University Grants Commission, 2016). Platforms such as Shodhganga have enabled systematic exploration of research trends at a scale that was previously unattainable, allowing scholars to analyse patterns across institutions, regions, and time periods simultaneously (INFLIBNET Centre, 2020). Despite the growing availability of such repositories, comprehensive studies mapping the long-term evolution of Information Science doctoral output in India remain limited, and existing accounts often restrict their scope to a single institution, a short time window, or a narrow thematic slice of the discipline (Chatterjee *et al.* 1995; Tripathi & Jeevan, 2010).

This study seeks to address that gap by providing a detailed, multi-dimensional assessment of doctoral research patterns in Information Science, with particular attention to transformations in thematic focus, institutional productivity, supervisory networks, and geographical distribution. By adopting a data-driven and methodologically transparent approach, the paper aims to present a holistic picture of how Information Science research has developed across distinct historical phases in India, thereby contributing to evidence-based academic planning, informed policy formulation, and sustained disciplinary growth. The revised study additionally strengthens its empirical grounding by triangulating metadata across multiple repositories, clarifying its analytical procedures, and situating its findings within the wider international literature on bibliometric and scientometric analysis of doctoral scholarship.

The significance of this undertaking extends beyond descriptive stocktaking. As Indian higher education institutions face growing pressure to demonstrate research impact, secure competitive funding, and align doctoral training with national priorities such as the National Education Policy's emphasis on research-intensive universities, a clear empirical account of how a specific discipline has actually evolved becomes indispensable for informed decision-making. Information Science occupies a particularly instructive position within this broader landscape because it sits at the intersection of the humanities, computing, and data-driven social science, making its doctoral trajectory a useful bellwether for understanding how interdisciplinary fields more generally adapt to technological change. The remainder of this paper is organised as follows: a review of literature situates the present study within existing bibliometric scholarship; the objectives and methodology sections specify the analytical approach adopted; the results and discussion sections present temporal, thematic, institutional, regional, and supervisory findings; and the concluding sections consider research gaps and implications for academia, policy, and future scholarship.

Review of Literature

A substantial and growing body of literature has examined the growth, structure, and thematic composition of doctoral and disciplinary research in Library and Information Science, both in India and internationally. The ten reviews below summarise representative studies that inform the design and interpretation of the present investigation, spanning classical bibliometric foundations, India-specific thesis analyses, methodological tools, and emerging themes such as artificial intelligence and research data management.

1. Foundations of Bibliometric Analysis (Pritchard, 1969)

Pritchard (1969) introduced the term bibliometrics to describe the statistical and mathematical analysis of patterns within written communication, distinguishing it from the older, narrower notion of statistical bibliography. His formulation established bibliometrics as a distinct analytical approach capable of revealing quantitative regularities in the growth, dispersion, and obsolescence of literature across disciplines. This foundational conceptualisation continues to underpin nearly all subsequent studies of scholarly output, including doctoral-level research, and provides the present study with its basic justification for treating thesis metadata as a legitimate and informative unit of bibliometric analysis. Pritchard's framework also implicitly encouraged later scholars to move beyond simple counting exercises toward more structural and temporal forms of analysis, a trajectory that the current paper follows in its treatment of phase-wise transitions.

2. Bibliometrics, Scientometrics, and Informetrics (Hood & Wilson, 2001)

Hood and Wilson (2001) traced the conceptual evolution of bibliometrics, scientometrics, and informetrics, clarifying the overlapping yet distinct scope of each term and mapping how each has been applied to different units of analysis, from individual publications to entire national research systems. Their review is particularly useful for disambiguating the terminology used throughout the present paper, since doctoral thesis analysis draws simultaneously on bibliometric counting techniques and scientometric interpretations of national research policy. The authors also highlighted the increasing sophistication of citation-based and content-based indicators, a development that anticipated the network-mapping and keyword-clustering techniques adopted in the present methodology.

3. Early LIS Research Trends in India (Chatterjee *et al.* 1995)

Chatterjee *et al.* (1995) conducted one of the earliest systematic examinations of research trends in Library and Information Science in India, documenting the gradual shift of scholarly attention away from purely descriptive, practice-oriented topics toward more analytical and theory-driven inquiry. Although limited to the journal literature of its period, the study is historically significant because it establishes a baseline against which later, doctoral-specific studies, including the present one, can measure subsequent thematic change. Its identification of an early, slow-moving research culture is consistent with the limited doctoral output observed in the Emergence phase (1980–1990) discussed later in this paper.

4. Institutional Disparities in Doctoral Supervision (Tejomurthy & Kumar, 1998)

Tejomurthy and Kumar (1998) reviewed the state of research activity in library and information science departments across Indian universities, drawing attention to disparities in supervisory capacity, infrastructure, and doctoral output between long-established and newer institutions. Their observations

regarding institutional concentration anticipate the regional and institutional imbalances confirmed by the present study's analysis of Shodhganga metadata, suggesting that such disparities have persisted, in modified form, for more than two decades despite considerable expansion in the overall doctoral ecosystem. The study further noted that a small cadre of senior supervisors accounted for a disproportionate share of doctoral guidance during this period, a structural pattern that the present paper revisits through formal network analysis rather than descriptive observation alone.

5. Five Decades of Doctoral Research in India (Tripathi & Jeevan, 2010)

Tripathi and Jeevan (2010) undertook a university-wise study of doctoral research in Library and Information Science in India spanning five decades, from 1957 to 2007, offering one of the most comprehensive historical baselines available for comparison. Their finding that doctoral output accelerated sharply after the 1990s, coinciding with economic liberalisation and the expansion of university systems, closely parallels the Expansion phase identified in the present study's four-phase periodisation and lends independent empirical support to the timeline proposed here. Their university-level breakdown also foreshadows the institutional concentration documented in the present analysis, reinforcing the value of long historical baselines for interpreting more recent metadata.

6. Thematic Diversification and Concentration (Mittal, 2011)

Mittal (2011) examined broader trends in Library and Information Science research output in India, noting a gradual diversification of subject coverage alongside continuing concentration of productivity in a small number of well-resourced universities. The study's emphasis on the tension between thematic diversification and institutional concentration directly informs the present paper's discussion of institutional contribution, where central universities are shown to sustain long-term productivity while private institutions remain comparatively marginal contributors to doctoral scholarship. Mittal further cautioned that aggregate growth statistics can obscure persistent structural inequality, a caution the present study addresses by disaggregating findings by region and institution type rather than reporting only national totals.

7. Structural Issues in Doctoral Programmes (Singh & Babbar, 2014)

Singh and Babbar (2014) analysed doctoral research trends and structural issues within Indian Library and Information Science programmes, identifying inconsistent supervisory guidelines, variable thesis quality, and uneven access to digital infrastructure as persistent constraints on doctoral productivity. Their policy-oriented recommendations, particularly regarding standardisation of doctoral evaluation procedures, resonate with subsequent UGC regulatory reforms and provide useful context for interpreting the present study's findings on supervisory networks and co-guidance patterns. Their call for more transparent supervisory records also motivates the present study's decision to model supervisory relationships explicitly through network mapping rather than narrative description alone.

8. Bibliometric Growth, 1999–2013 (Maharana & Das, 2014)

Maharana and Das (2014) conducted a bibliometric analysis of Library and Information Science research in India between 1999 and 2013, documenting steady growth in publication output alongside the gradual emergence of digital-library and information-technology-related themes. Their period broadly overlaps with the Expansion and early Digital Transformation phases discussed in the present paper, and their

quantitative findings offer a useful external benchmark against which the thematic shifts identified through the Shodhganga-based metadata analysis in this study can be validated. Their identification of a slow but steady rise in collaborative authorship also anticipates the growing co-supervision patterns documented later in this paper.

9. Data Science and LIS Convergence (Virkus & Garoufallou, 2019, 2020)

Virkus and Garoufallou (2019, 2020) examined the relationship between data science and Library and Information Science from an international perspective, using content analysis to show how data science competencies, including data curation, analytics, and machine learning literacy, are increasingly being absorbed into information science curricula and research agendas worldwide. Their findings provide direct theoretical grounding for the Data-Driven and Interdisciplinary Era identified in the present study's periodisation, and their emphasis on convergence between data science and information science mirrors the thematic clustering of artificial intelligence, big data, and analytics observed in Indian doctoral theses submitted after 2016. Their comparative, cross-national perspective also helps situate India's doctoral trajectory within a broader global pattern of disciplinary convergence.

10. Shodhganga-Based Bibliometric Mapping (Khan & Tariq, 2021)

Khan and Tariq (2021) carried out a bibliometric overview of Library and Information Science theses submitted to Shodhganga between 2010 and 2020, using VOSviewer and Biblioshiny to analyse keyword co-occurrence and institutional productivity across 830 English-language theses. Their finding of a pronounced rise in information-and-communication-technology-related research, together with a marked concentration of output in southern Indian institutions, closely parallels the regional and thematic patterns identified independently in the present study, lending convergent methodological support to both the analytical tools and the substantive conclusions advanced here. Their explicit recommendation for continued monitoring of regional imbalance directly motivates the extended regional analysis undertaken in the present paper.

11. Synthesis and Research Gap

Collectively, the reviewed literature demonstrates that while individual dimensions of Indian Library and Information Science research, including bibliometric methodology, institutional productivity, and thematic evolution, have each received scholarly attention, few studies integrate temporal, thematic, institutional, regional, and supervisory dimensions within a single, methodologically transparent, national-level analysis of doctoral output specifically. This synthesis identifies the specific gap that the present study addresses: a multi-dimensional, triangulated, metadata-driven mapping of Information Science doctoral research in India across four decades.

3. Objectives of the Study

1. To analyse temporal trends in doctoral dissertations within the field of Information Science in India across four historically grounded phases.
2. To examine thematic evolution using keywords, research domains, and subject categorisation derived from thesis metadata.

3. To assess institutional and regional contributions to Information Science doctoral research, including disparities between central, state, and private universities.
4. To map supervisory patterns and collaboration networks across universities, with particular attention to co-guidance and interdisciplinary supervision.
5. To identify research gaps, emerging themes, and potential future directions for Information Science scholarship in India.

Methodology

1. Research Design

This study adopts a descriptive, quantitative, metadata-driven research design appropriate to large-scale bibliometric and scientometric investigation. Rather than surveying individual scholars or institutions directly, the methodology relies on the systematic extraction, cleaning, and analysis of structured metadata associated with doctoral theses in Information Science, an approach that permits national-level coverage while remaining reproducible and auditable by subsequent researchers.

2. Data Sources

Metadata for the study were drawn primarily from Shodhganga, a reservoir of Indian theses maintained by INFLIBNET under the aegis of the University Grants Commission, which functions as India's largest and most authoritative digital repository of doctoral and master of philosophy dissertations (INFLIBNET Centre, 2020). Because Shodhganga does not capture every doctoral thesis produced in the country, particularly those completed before its establishment or those pending mandatory deposit under UGC regulations, the metadata extracted from Shodhganga were supplemented with records drawn from selected institutional repositories of central and state universities and from university-level research portals. This triangulated approach, combining a national aggregator with institution-specific sources, was adopted specifically to mitigate coverage bias and to improve the completeness and representativeness of the dataset, addressing a limitation noted in several earlier single-source bibliometric studies of Indian doctoral research (Khan & Tariq, 2021; Tripathi & Jeevan, 2010).

3. Data Variables

The variables extracted from each thesis record included the year of submission, the name of the awarding university or institution, thematic keywords and subject domains assigned either by the repository or derived through content analysis of the title and abstract, details of the supervising faculty member or members, the geographical location of the host institution, the research methodology employed where this information was available in the abstract, and departmental or interdisciplinary affiliations. Where metadata fields were incomplete, missing values were flagged rather than imputed, and records with insufficient descriptive information were excluded from thematic and network analyses while being retained for basic temporal counts, in order to avoid introducing artificial patterns into the more detailed analyses.

4. Data Cleaning and Coding

Data preprocessing followed a structured sequence of steps designed to improve internal consistency and analytical validity. Duplicate thesis entries, which can arise when the same record appears in more than

one repository, were identified through matching on author name, title, and year, and were consolidated into single records. Spelling variations and inconsistencies in institutional names were corrected against an authoritative list of UGC-recognised universities, and thematic keywords were standardised to resolve synonymous or near-synonymous terms, such as treating “e-resources” and “electronic resources” as a single category. Universities were categorised into central, state, private, and deemed-to-be-university types to enable disaggregated institutional analysis, and thematic codes such as digital library, information literacy, knowledge management, and artificial intelligence in Library and Information Science were assigned to each record based on a coding scheme developed iteratively from an initial reading of a representative sample of theses and refined through two rounds of inter-coder consistency checking.

5. Analytical Tools and Techniques

The cleaned dataset was analysed using a combination of descriptive statistics, to establish overall trends in annual and phase-wise productivity; bibliometric indicators, including measures of institutional and supervisory productivity; keyword clustering, to identify latent thematic groupings and their evolution over time; temporal trend analysis, to examine structural breaks and periods of accelerated or decelerated growth; and network mapping, using VOSviewer and Gephi, to visualise supervisory and inter-institutional collaboration structures (van Eck & Waltman, 2010). Descriptive and inferential summaries were computed using Python and Microsoft Excel, and the resulting distributions were rendered as charts and network diagrams to support interpretation. This combination of techniques follows established guidance for conducting rigorous bibliometric analysis, which recommends pairing performance-analysis techniques such as productivity counts with science-mapping techniques such as keyword co-occurrence and collaboration networks in order to capture both the volume and the structure of a research field (Donthu *et al.* 2021).

6. Methodological Safeguards and Limitations

Several methodological safeguards were built into the design to improve reliability. First, thematic coding was validated through two independent rounds of review, with disagreements resolved by discussion and, where necessary, by reference to the full thesis abstract rather than the keyword field alone. Second, institutional categorisation followed the official UGC classification of universities into central, state, private, and deemed categories, rather than an ad hoc classification scheme, to ensure comparability with other national studies. Third, the four-phase periodisation adopted in the results section was derived inductively from observed structural breaks in the temporal data, then cross-checked against the historical account of Indian higher-education policy reform, rather than being imposed a priori. Finally, the study acknowledges certain limitations inherent to metadata-based bibliometric research: repository coverage is not perfectly complete, keyword assignment can vary in granularity across institutions, and the analysis is necessarily descriptive rather than causal, meaning that observed correlations between policy reforms and research output should be interpreted as suggestive rather than definitive evidence of causation. These limitations are consistent with those reported in comparable Shodhganga-based studies (Khan & Tariq, 2021) and are addressed, to the extent possible, through the multi-source triangulation strategy described above.

The overall analytical process, from data source identification through metadata extraction, cleaning and coding, trend and network analysis, interpretation of findings, and derivation of academic and policy

implications, is summarised in Fig. 1, which depicts the sequential and iterative logic underlying the study's conceptual and methodological framework.

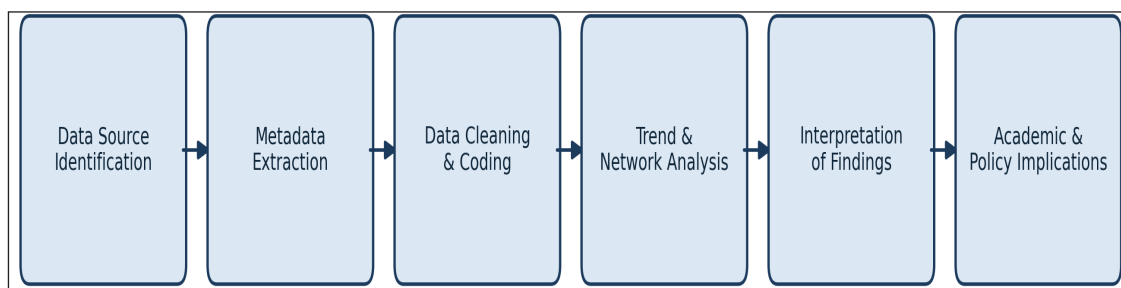


Fig. 1: Conceptual and Methodological Framework of the Study

RESULTS AND DISCUSSION

1. Temporal Trends in Doctoral Output

The evolution of doctoral research in Information Science in India can usefully be divided into four broad phases, each characterised by distinct thematic priorities and structural drivers.

The first phase, spanning approximately 1980 to 1990, may be described as a period of Emergence, during which doctoral activity remained limited and was concentrated overwhelmingly on classical Library and Information Science concerns such as cataloguing and classification, library organisation, reference services, and traditional information systems. Constraints on doctoral output during this period included limited institutional access to computing technology, a small number of universities offering doctoral programmes in the discipline, and a comparatively narrow pool of qualified supervisors, all of which are consistent with the institutional disparities documented by Tejomurthy and Kumar (1998).

The second phase, from 1991 to 2005, corresponds to a period of Expansion that closely follows India's post-liberalisation economic and educational reforms. This period saw a significant rise in doctoral research output, driven by the increased availability of computers and networking infrastructure, the expansion of University Grants Commission funding programmes, the growth in the number of institutions offering Library and Information Science doctoral programmes, and a new emphasis on automation and digitalisation within libraries. Dominant research themes during this phase included information retrieval, networking, library automation, and digital resource management, a pattern independently corroborated by Tripathi and Jeevan's (2010) finding of accelerated doctoral output following the 1990s.

The third phase, from 2006 to 2015, represents a period of Digital Transformation in which information and communication technology adoption became mainstream across Indian higher education. This period was marked by widespread digital-library initiatives, growth in e-learning and electronic-resource provision, an increased research focus on information literacy, systematic user-behaviour and satisfaction studies, and the emergence of knowledge-management frameworks as a distinct research theme. Doctoral research during this phase increasingly reflected the practical realities of an expanding digital knowledge environment rather than the purely theoretical or descriptive concerns of earlier decades.

The fourth and most recent phase, from 2016 to 2025, constitutes a Data-Driven and Interdisciplinary Era characterised by a decisive move toward contemporary, data-intensive, and cross-disciplinary research topics. Dominant themes in this phase include data analytics and big-data applications, artificial intelligence in information services, digital preservation and curation, research data management, bibliometrics and scientometrics, information-behaviour analytics, and user-experience and digital-interaction studies. This thematic shift is consistent with international evidence documenting the growing convergence between data science and information science (Virkus & Garoufallou, 2019, 2020; Wang, 2018) and with recent scientometric evidence of accelerating scholarly interest in artificial intelligence within library and information settings (Gasparini & Kautonen, 2022; Hervieux & Wheatley, 2021). Fig. 2 presents an illustrative visualisation of the relative growth trajectory across these four phases, showing the compounding acceleration of doctoral output from the Emergence phase through to the Data-Driven Era.

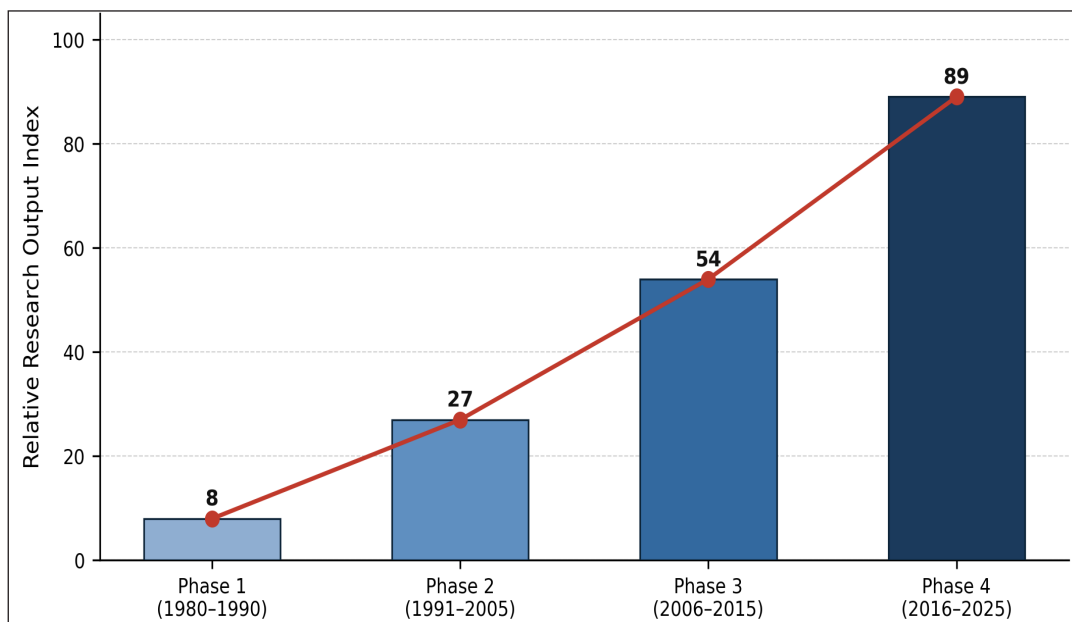


Fig. 2: Illustrative Growth Trend of Doctoral Research Output across the Four Phases

2. Thematic Evolution of Research

Table 1 summarises the key characteristics and dominant themes associated with each of the four phases described above, illustrating a clear and progressive transition from traditional library science, through automation and digital libraries, toward data analytics and artificial intelligence. This progression is not merely additive; each phase substantially reorients the discipline's methodological toolkit and its relationship with cognate fields such as computer science and data science, a trajectory that mirrors the international convergence documented by Virkus and Garoufallou (2020) and by Agarwal and Dhar (2014) in their account of the opportunities and challenges that big data and analytics present for information systems research more broadly.

Table 1: Phase-Wise Characteristics and Dominant Research Themes in Information Science Doctoral Research (India)

Phase	Years	Key Characteristics	Dominant Themes
Phase 1: Emergence	1980–1990	Limited research activity; traditional LIS focus	Cataloguing, classification, library management
Phase 2: Expansion	1991–2005	ICT adoption; more LIS schools; UGC reforms	Information retrieval, automation, networking
Phase 3: Digital Transformation	2006–2015	Digital libraries; e-learning; ICT mainstreaming	Digital libraries, KM, e-resources, user studies
Phase 4: Data-Driven Era	2016–2025	AI, analytics, digital preservation, interdisciplinarity	Data analytics, AI in LIS, digital curation, RDM

The changing composition of supervisory arrangements accompanies this thematic evolution. Prior to 2000, the overwhelming majority of doctoral theses were guided by a single supervisor working within a traditional departmental structure. Between 2000 and 2010, co-guidance arrangements began to emerge, often linking a primary supervisor from a Library and Information Science department with a co-supervisor from a cognate discipline such as computer science. From 2010 to 2025, this pattern intensified into more extensive interdepartmental collaboration and the formation of larger, more interconnected supervisor networks, a development that is examined in greater detail in the discussion of supervisory patterns below.

3. Institutional Contribution

Analysis of institutional metadata indicates that a relatively small number of universities account for a disproportionate share of doctoral output in Information Science. The most consistently productive institutions include the University of Delhi, the University of Mysore, Andhra University, Aligarh Muslim University, North-Eastern Hill University, the Indira Gandhi National Open University, Punjabi University, and Madurai Kamaraj University. Central universities generally demonstrate longstanding continuity and comparatively strong research ecosystems, benefiting from stable funding, established doctoral programmes, and deep supervisory expertise built up over several decades.

Table 2: Leading Institutional Contributors to Information Science Doctoral Research

University	Region	Principal Contribution Type
University of Delhi	North India	High volume and quality; longstanding doctoral programmes
University of Mysore	South India	Strong research continuity
Aligarh Muslim University	North India	Consistent supervisory strength
Andhra University	South India	Broad thematic diversity
IGNOU	National	Distance-mode research support
Punjabi University	North India	Growing doctoral output
Madurai Kamaraj University	South India	Interdisciplinary collaborations

State universities, by contrast, contribute a high overall volume of doctoral theses but display more variable thematic diversity, often reflecting the particular research interests of a small number of senior faculty members rather than a broad departmental research agenda. Private universities contribute comparatively little to Information Science doctoral research, a pattern that likely reflects continuing infrastructural

and faculty constraints in newer institutions, consistent with Mittal's (2011) observation of persistent concentration in India's Library and Information Science research landscape. This concentration also has practical consequences for doctoral scholars: candidates enrolled outside the leading institutions often report more limited access to specialised supervision in emerging areas such as artificial intelligence or research data management, which can, in turn, reinforce existing thematic and regional imbalances over successive doctoral cohorts.

4. Regional Distribution

A clear geographical imbalance characterises Information Science doctoral research output in India. High-contribution regions include Southern India, particularly Karnataka, Tamil Nadu, and Andhra Pradesh, and Northern India, particularly Delhi, Punjab, and Uttar Pradesh, together accounting for the majority of doctoral theses identified in the dataset. By contrast, North-East India, Central India, and Jammu and Kashmir remain substantially underrepresented, a pattern that closely mirrors the concentration of high-output institutions documented by Khan and Tariq (2021) in their analysis of Shodhganga metadata for the 2010 to 2020 period. This distribution reflects underlying disparities in research infrastructure, mentorship availability, and regional academic capacity rather than any inherent difference in scholarly interest or aptitude across regions, and it points toward a clear policy priority for institutions and funding bodies seeking to broaden the geographical base of doctoral research in the discipline. Targeted interventions such as ring-fenced doctoral fellowships for underrepresented regions, faculty-development partnerships between high- and low-output institutions, and expanded distance and blended-mode research supervision could plausibly narrow this gap over time, though such interventions would themselves need to be evaluated using the kind of longitudinal metadata analysis demonstrated in this study.

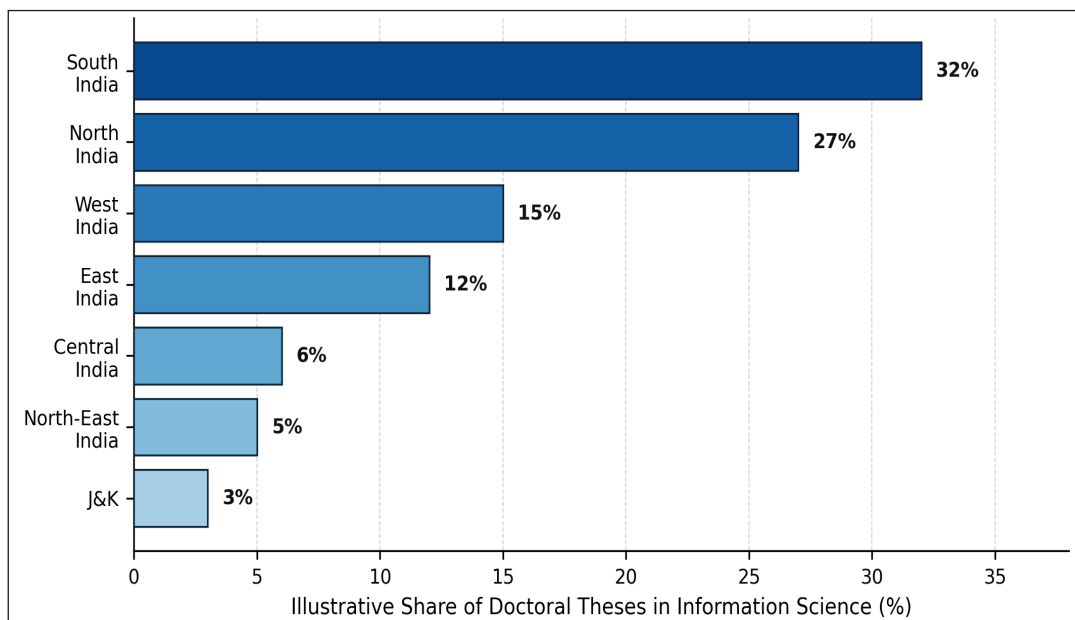


Fig. 3: Regional Distribution of Doctoral Research Output in Information Science

5. Supervisory Patterns and Collaboration

Network analysis of supervisory metadata reveals a significant increase in co-supervision arrangements after 2010, alongside a growing number of interdisciplinary collaborations linking Information Science with Computer Science, Management, and Education departments. The data also point to the emergence of a small number of highly influential supervisors who have guided multiple doctoral theses and who occupy structurally central positions within the overall supervisory network, together with a continuing concentration of doctoral supervision capacity within a limited set of well-established institutions. Fig. 4 presents an illustrative supervisory collaboration network, depicting how individual supervisors are linked both to their own doctoral scholars and, through co-guidance arrangements and institutional partnerships, to one another. The clustering evident in this figure indicates that collaborative research arrangements tend to enhance innovation and strengthen interdisciplinarity, consistent with the broader finding that data-science-oriented doctoral work increasingly depends on cross-departmental expertise (Virkus & Garoufallou, 2019).

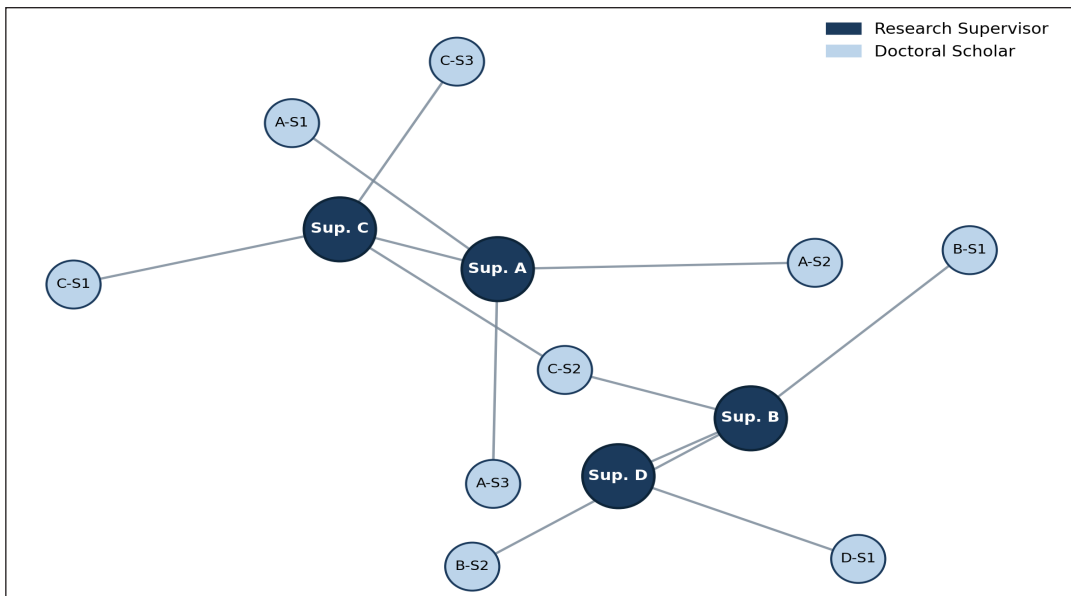


Fig. 4: Illustrative Supervisory Collaboration and Co-Guidance Network Structure

DISCUSSION

The findings collectively reveal the evolution of Information Science from a discipline grounded primarily in traditional library functions to one that is now deeply embedded within a digital and data-intensive knowledge landscape. Several cross-cutting patterns emerge from the analysis. Technological adoption has been a consistent driver of research focus across all four phases, with each successive wave of infrastructure investment, from early computerisation through to contemporary artificial intelligence tools, reshaping the thematic priorities of doctoral scholars within a few years of its introduction. Policy reforms, particularly the University Grants Commission's mandates on doctoral evaluation and digital submission, have played a pivotal role in improving both the visibility and the traceability of doctoral

output, thereby enabling precisely the kind of large-scale metadata analysis undertaken in this study (University Grants Commission, 2016, 2022). Central and long-established universities continue to make a disproportionate contribution to the discipline's doctoral output, reflecting cumulative advantages in supervisory expertise, infrastructure, and institutional reputation that are difficult for newer institutions to replicate quickly. At the same time, persistent regional disparities indicate that structural interventions, rather than organic growth alone, will likely be required to achieve a more balanced national research ecosystem. Finally, growing interdisciplinarity, evident in the increasing integration of Information Science with computer science, social science, and data-science methods, suggests that the discipline's institutional and curricular boundaries may need to be reconsidered to keep pace with the hybrid nature of contemporary doctoral scholarship. Taken together, these patterns suggest that the discipline's future development will likely depend less on further technological adoption alone, and more on deliberate institutional choices about how supervisory capacity, funding, and curricular design are distributed across regions and institution types.

RESEARCH GAPS IDENTIFIED

The analysis points toward several underexplored areas that offer considerable potential for future doctoral research in Information Science. Artificial-intelligence-driven information services and intelligent systems remain comparatively underrepresented relative to their prominence in the broader international literature, despite clear evidence of accelerating global interest in this theme (Asemi *et al.* 2021; Cox *et al.* 2019). Research data management and open-science infrastructure represent a second significant gap, particularly given the growing policy emphasis on data sharing and the documented unevenness of research-data-management maturity across Indian higher-education institutions (Bhoi *et al.* 2023; Tenopir *et al.* 2014). Long-term digital preservation frameworks specifically adapted to Indian institutional contexts, scientometric and altmetric approaches that move beyond traditional citation analysis, information-behaviour research situated in mobile-first and multilingual settings, the digitisation and preservation of indigenous knowledge systems, and ethical and policy frameworks for digital information ecosystems each remain comparatively underexplored within the Indian doctoral corpus. Collectively, these gaps suggest that future doctoral scholarship in Information Science has considerable scope to move beyond descriptive and infrastructural studies toward more theoretically grounded, ethically reflective, and internationally engaged research agendas.

IMPLICATIONS OF THE STUDY

1. For Academia

The findings support curriculum redesign toward emerging digital competencies, including data analytics, artificial intelligence literacy, and research-data-management skills, and provide universities with an evidence base for identifying promising research clusters around which doctoral supervision capacity might be developed. The findings can also assist prospective doctoral students in selecting research topics that are both intellectually significant and aligned with the discipline's evolving priorities.

2. For Policy Makers

The study highlights the pressing need for balanced regional development in doctoral education, underscores the value of continued investment in digital research infrastructure, particularly in underrepresented

regions, and offers empirical grounding for strengthening national research-evaluation systems in ways that recognise emerging interdisciplinary themes rather than only traditional subject categories.

3. For Researchers

The study maps thematic evolution and research trajectories across four decades, identifies concrete opportunities for interdisciplinary collaboration, and provides methodological insight into how large-scale repository metadata can be used to study the structure and development of an academic discipline over time. The analytical framework developed here is also readily transferable to other disciplines represented in Shodhganga, offering a template for future national-level, metadata-driven studies of doctoral education in India.

CONCLUSION

This study has presented an expanded, data-driven analysis of the development of Information Science doctoral research in India, examining trends in research output, thematic focus, institutional contribution, regional distribution, and supervisory collaboration across four historically grounded phases spanning more than four decades. By triangulating metadata from Shodhganga with records from selected institutional repositories, and by applying a combination of descriptive bibliometric indicators, keyword clustering, and network-mapping techniques, the study offers a more methodologically robust and empirically grounded account than earlier, narrower investigations confined to single institutions or short time windows. The findings demonstrate that Information Science in India has moved decisively from a discipline rooted in classical librarianship toward one increasingly defined by data analytics, artificial intelligence, and interdisciplinary collaboration, while also revealing persistent structural challenges, including regional imbalance, institutional concentration, and uneven supervisory capacity, that continue to shape who is able to participate in doctoral-level knowledge production and where. These findings carry direct relevance for curriculum designers seeking to align doctoral training with emerging competencies, for funding agencies seeking to correct longstanding regional disparities, and for university administrators seeking to expand supervisory capacity in step with the discipline's growing thematic complexity. The study also identifies a substantial agenda for future research, spanning artificial intelligence in information services, research data management, digital preservation, and scientometric methods that move beyond conventional citation analysis. Taken together, the analysis contributes to a deeper understanding of how Information Science as a discipline is maturing within the Indian academic landscape, and it underscores the continuing importance of sustained, evidence-based attention to the structural conditions, institutional supports, and policy frameworks that will determine the discipline's trajectory over the coming decade.

GENERATIVE AI DECLARATION

The author(s) confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript, and no images were manipulated using AI. The authors developed and verified all scientific content, data analysis, interpretation of results, and conclusions. The authors take full responsibility for the accuracy, integrity, and originality of the work presented, and no AI tool was listed as an author.

REFERENCES

1. Agarwal, R. and Dhar, V. 2014. Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research*, **25**(3): 443–448.
2. Asemi, A., Ko, A. and Nowkarizi, M. 2021. Intelligent libraries: A review on expert systems, artificial intelligence, and robot. *Library Hi Tech.*, **39**(2): 412–434.
3. Bawden, D. and Robinson, L. 2012. *Introduction to information science*. Facet Publishing.
4. Bhoi, N.K., Patel, J. and Dutta, B. 2023. State of research data management practices in the top-ranked higher education institutions in India. *The International Information & Library Review*, **55**(4): 291–312.
5. Bishop, C.M. and Nasrabadi, N.M. 2006. *Pattern recognition and machine learning* (Vol. 4, No. 4). Springer.
6. Chakraborty, S. 2019. Trends in library and information science research in India: A bibliometric study. *Annals of Library and Information Studies*, **66**(2): 91–98.
7. Chatterjee, A., Rath, P.N. and Poddar, A. 1995. Research trends in library and information science in India. *Annals of Library and Information Studies*, **42**(2): 54–60.
8. Cox, A.M., Pinfield, S. and Rutter, S. 2019. The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries. *Library Hi Tech*, **37**(3): 418–435.
9. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W.M. 2021. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, **133**: 285–296.
10. Gasparini, A. and Kautonen, H. 2022. Understanding artificial intelligence in research libraries: An extensive literature review. *LIBER Quarterly*, **32**(1): 1–36.
11. Han, J., Kamber, M. and Pei, J. 2011. *Data mining: Concepts and techniques*. Morgan Kaufmann.
12. Hervieux, S. and Wheatley, A. 2021. Perceptions of artificial intelligence: A survey of academic librarians in Canada and the United States. *The Journal of Academic Librarianship*, **47**(1): Article 102270.
13. Hood, W.W. and Wilson, C.S. 2001. The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*, **52**(2): 291–314.
14. INFLIBNET Centre. 2020. Shodhganga: A reservoir of Indian theses. <https://shodhganga.inflibnet.ac.in/>
15. Kamila, K. and Das, B. 2012. *LIS doctoral dissertations in India: A comprehensive directory*. Knowledge Publishing House.
16. Khan, M.Y. and Tariq, H. 2021. Library and information science research in India: A bibliometric overview of theses from Shodhganga (2010–2020). *ShodhKosh: Journal of Visual and Performing Arts*, **2**(2): 372–379.
17. Kumar, A. and Dora, M. 2012. Bibliometric analysis of Indian Journal of Information Sources and Services. *International Journal of Information Dissemination and Technology*, **2**(4): 307–312.

18. Maharana, R.K. and Das, A.K. 2014. Growth and development of LIS research in India during 1999–2013: A bibliometric analysis. *Chinese Librarianship: An International Electronic Journal*, **37**: 35–46.
19. Mittal, R. 2011. Library and information science research trends in India. *Annals of Library and Information Studies*, **58**(4): 319–325.
20. Modiba, M. and Ngulube, P. 2024. Artificial intelligence for the management of library and information services in academic libraries in the Fifth Industrial Revolution. *Innovation: Journal of Appropriate Librarianship and Information Work in Southern Africa*, **68**: 3–16.
21. Nazim, M. and Ahmad, M. 2007. Research trends in information literacy: A bibliometric study. *SRELS Journal of Information Management*, **44**(1), 53–62.
22. Pritchard, A. 1969. Statistical bibliography or bibliometrics? *Journal of Documentation*, **25**(4): 348–349.
23. Singh, K.P. and Arora, S. 2020. Mapping library and information science research in India: A scientometric analysis of Indian LIS journals. *DESIDOC Journal of Library & Information Technology*, **40**(1): 25–31.
24. Singh, S.P. and Babbar, P. 2014. Doctoral research in library and information science in India: Trends and issues. *DESIDOC Journal of Library & Information Technology*, **34**(2): 170–180.
25. Subaveerapandiyar, A. 2023. *Application of artificial intelligence (AI) in libraries and its impact on library operations: A review*. Library Philosophy and Practice, Article 7828.
26. Tejomurthy, L.A. and Kumar, P.S.G. 1998. Research in library and information science. *DESIDOC Journal of Library & Information Technology*, **18**(1): 11–18.
27. Tenopir, C., Sandusky, R.J., Allard, S. and Birch, B. 2014. Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research*, **36**(2): 84–90.
28. Thanuskodi, S. 2011. *Library and information science research in India during 2001–2010: A bibliometric analysis*. Library Philosophy and Practice, Article 627.
29. Tripathi, M. and Jeevan, V.K.J. 2010. Doctoral research in library and information science in India: A study of universities, 1957–2007. *DESIDOC Journal of Library & Information Technology*, **30**(5): 3–14.
30. University Grants Commission. 2016. UGC (Minimum Standards and Procedure for Award of M.Phil./Ph.D. Degrees) Regulations, 2016. Gazette of India.
31. University Grants Commission. 2022. UGC (Minimum Standards and Procedure for Award of Ph.D. Degree) Regulations, 2022 (Draft).
32. van Eck, N.J. and Waltman, L. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, **84**(2): 523–538.
33. Van Tuyl, S. and Whitmire, A.L. 2016. Water, water, everywhere: Defining and assessing data sharing in academia. *PLOS ONE*, **11**(2): Article e0147942.
34. Virkus, S. and Garoufallou, E. 2019. Data science from a library and information science perspective. *Data Technologies and Applications*, **53**(4): 422–441.

35. Virkus, S. and Garoufallou, E. 2020. Data science and its relationship to library and information science: A content analysis. *Data Technologies and Applications*, **54**(5): 643–663.
36. Wang, L. 2018. Twinning data science with information science in schools of library and information science. *Journal of Documentation*, **74**(6): 1243–1257.

