From Classification To Cognition: AI & ML In Indian Library Science

Dr. Neha M. Joshi

Librarian

Ghanshyamdas Saraf College of Arts and Commerce

RSET Campus, S. V. Road, Sunderngar,

Malad West, Mumbai – 400064

Abstract

The Indian Library and Information Science (LIS) sector is experiencing a significant transformation, shifting from conventional rule-based classification systems such as the Dewey Decimal System and Colon Classification to AI-enhanced cognitive services that improve user discovery, automate metadata creation, and facilitate predictive and personalized information access. This research investigates the adoption of Artificial Intelligence (AI) and Machine Learning (ML) in Indian libraries, particularly in academic settings by analyzing national surveys, pilot initiatives, and policy documents. The study assesses current AI applications—including intelligent chatbot, multilingual OCR/NLP tools, and recommendation engines—while highlighting challenges related to professional skills, infrastructure, and ethical implementation. Additionally, the study addresses the deficiencies in LIS education and the lack of a cohesive policy framework for AI integration in libraries. Drawing on government initiatives like Digital India, NDLI, Indian, and NEP 2020, the paper provides actionable recommendations for incorporating AI/ML into LIS curricula, encouraging interdisciplinary collaboration, and fostering ethical, inclusive innovation in a multilingual environment. It establishes Indian libraries as crucial testing grounds for cognitive information services that support the country's goals of inclusive knowledge access and digital equity.

Keywords: Artificial Intelligence, Machine Learning, Library and Information Science, India, Natural Language Processing, Multilingual Access, Cognitive Libraries, AI Ethics.

I. INTRODUCTION

The Indian Library and Information Science (LIS) sector is undergoing a significant paradigm shift. Traditionally grounded in manual processes such as classification using the Dewey Decimal Classification (DDC) and S.R. Ranganathan's Colon Classification systems, Indian libraries are now steadily transitioning toward the adoption of Artificial Intelligence (AI)-driven cognitive services. These include personalized discovery interfaces, intelligent Chatbot support, Optical Character Recognition (OCR)-based content retrieval, natural language processing (NLP) for multilingual search, and predictive analytics to anticipate user needs. This evolution signifies a departure from the conventional role of libraries as static information repositories and repositions them as dynamic, proactive agents of knowledge dissemination and creation—particularly vital in India's linguistically diverse and digitally advancing society.

Globally, the digital revolution has compelled libraries to modernize through emerging technologies. In India, this momentum is shaped by the need to serve a vast and heterogeneous population with varied literacy levels, languages, and access challenges. The integration of AI and Machine Learning (ML) in the Indian LIS domain holds transformative potential—not only for automating routine functions but

also for enabling intelligent and adaptive information systems that respond to user behaviors, preferences, and academic or research goals. For instance, recommender systems, AI-based content tagging, sentiment analysis, and automated indexing are becoming increasingly relevant in both academic and public library environments.

The primary objective of this study is to investigate the evolving intersection between AI/ML and Library Science in the Indian context. It aims to map current applications of these technologies, assess institutional readiness, and identify strategic pathways for sustainable and inclusive integration. By examining selected academic and public libraries across India, this research highlights ongoing practices, technological capacities, infrastructural limitations, human resource challenges, and existing policy frameworks that either support or hinder AI adoption in LIS services. Furthermore, it explores the implications of these developments for digital equity, multilingual access, information ethics, and the future of librarianship in India.

II. LITERATURE REVIEW

Globally, the integration of Artificial Intelligence (AI) and Machine Learning (ML) into Library and Information Science (LIS) has transitioned from experimental to increasingly mainstream. Libraries across technologically advanced countries have adopted AI for a broad spectrum of tasks including automated indexing, real-time user interaction via chatbots, personalized recommendation systems, predictive analytics for usage trends, and sentiment analysis of user feedback (Zhou et al., 2021). These cognitive tools are reshaping libraries into intelligent systems that understand and adapt to user needs—marking a clear shift from traditional classification paradigms to AI-enabled knowledge ecosystems.

In the Indian context, however, the adoption of AI in libraries remains uneven and fragmented. While premier academic institutions such as the Indian Institutes of Technology (IITs) and Indian Institutes of Management (IIMs) have begun implementing AI-enabled cataloguing, semantic search, and discovery services, the vast majority of libraries—particularly in public and rural sectors—struggle with infrastructural deficits, lack of funding, and limited professional capacity (Kumar & Sharma, 2020). The disparity in technological readiness creates a digital divide that hampers the wider diffusion of AI innovations in the Indian LIS landscape.

A growing body of research identifies the need to reconfigure LIS education and professional training in India to prepare for this technological shift. Pujar and Satyanarayana (2020) emphasize the urgency of incorporating AI and data science competencies into LIS curricula, suggesting that without such reforms, library professionals will remain ill-equipped to participate in AI-driven transformation. Similarly, initiatives such as the National Digital Library of India (NDLI) and INFLIBNET's Shodhganga have begun deploying intelligent search algorithms and Natural Language Processing (NLP) tools tailored to Indian languages and scripts, offering promising models for scalable AI adoption in national-level digital repositories.

Subaveerapandiyan and Gozali (2024) conducted a large-scale survey involving 386 library professionals across Indian academic institutions. The study found widespread recognition of AI's potential to improve accessibility, automate routine operations, and support decision-making. At the same time, participants expressed significant concerns about data privacy, ethical implications, employment displacement, and lack of sustainable funding. These findings echo broader anxieties within the Indian LIS community regarding the social and institutional costs of automation.

Kalbande et al. (2024), through a structured survey of 259 academic librarians via Google Forms, documented high optimism toward AI integration. Respondents recorded strong agreement with positive statements such as "AI can bridge performance gaps" and "AI does not make staff lazy," suggesting a cultural openness to technological change. Yet the same study revealed persistent apprehensions related to resource constraints and the long-term employment implications of automation—highlighting the need for supportive policy frameworks and inclusive AI strategies.

Similarly, a focused study involving 118 Master of Library Science (MLS) students across Indian universities found that NLP ranked highest among AI technologies perceived as transformative, followed by recommender systems, robotic process automation, and data mining tools. While students were enthusiastic about AI's possibilities, they also showed an acute awareness of ethical dilemmas, surveillance risks, and the threat of human redundancy. The authors emphasized the urgent need for AI and data literacy in LIS pedagogical practices (Taylor & Francis Online, 2024).

Kanaujia et al. (2025) undertook a detailed analysis of AI applications in resource discovery and information management in Indian libraries. Their findings confirmed the early-stage implementation of tools such as NLP-based semantic search, predictive analytics for user behavior, intelligent content recommendation, and Chabot-assisted query resolution. These innovations signify a shift from traditional systems based on manual classification schemes toward cognitive service delivery, thereby enhancing user engagement and resource visibility.

Several pilot projects in academic libraries across Karnataka, Maharashtra, and Delhi have reported experimental use of AI tools such as ChatGPT for information retrieval, citation validation, plagiarism detection, and library user training. While these use-cases remain nascent and lack formal institutional backing, they point toward grassroots-level experimentation and innovation among Indian LIS professionals (journalcra.com; arXiv.org).

Another notable development is the application of OCR-enabled multilingual search functions on platforms like NDLI and Shodhganga. Leveraging advanced NLP pipelines for Indic scripts such as Devanagari, Tamil, Bengali, and Malayalam, these platforms have significantly improved access to digitized books, theses, and rare manuscripts. The ability to retrieve full-text content across regional languages represents a major advancement in democratizing knowledge for India's diverse user base.

Taken together, these studies underline both the potential and the challenges of AI and ML in reshaping Indian LIS practices. The shift from classification-based to cognition-oriented services necessitates not only technological infrastructure but also a rethinking of LIS education, institutional capacity, and inclusive AI policies.

III. METHODOLOGY

The present research utilizes a mixed-methods model that combines various research strategies to provide exhaustive analysis. An initial literature analysis was carried out by referring to peer-reviewed articles, government reports like the National Education Policy (NEP) 2020, and institutional documents to provide a theoretical context. Subsequent case studies were carried out to study actual instances of artificial intelligence (AI) implementations in chosen Indian academic and public libraries. Third, the synthesis of a survey was conducted by consolidating results from recent national-level Library and Information Science (LIS) surveys (e.g., Subaveerapandiyan & Gozali, 2024; Kalbande et al., 2024) to offer empirical evidence on emerging trends and challenges. Lastly, a comparative analysis was conducted to analyze the adoption trends in India vis-a-vis global best practice to facilitate the identification of gaps and prospects for improvement.

IV. AI APPLICATIONS, READINESS, AND CHALLENGES IN INDIAN LIBRARIES

The integration of Artificial Intelligence (AI) and Machine Learning (ML) technologies in Indian libraries is gradually reshaping traditional practices of resource organization, discovery, and user engagement. Moving beyond classical classification models like DDC and Colon Classification, Indian libraries are exploring AI-enabled cognitive solutions to improve personalization, multilingual access, and service efficiency. This section maps current applications, readiness levels, and implementation barriers across various types of libraries in India.

4.1 AI Applications in Indian Libraries

AI-driven applications in Indian libraries are increasingly prevalent across various domains, primarily at experimental or pilot stages. Institutions like INFLIBNET and the NDLI employ NLP and machine learning technologies to streamline metadata generation, thereby enhancing cataloguing efficiency and resource discoverability. AI chatbots, such as the pilot project of the Delhi University Library System, assist users with common inquiries, while plagiarism detection tools like URKUND and Turnitin uphold research integrity, albeit with limitations in Indian language support. The recommendation systems at IISER Pune and IIT Kharagpur curate personalized reading lists based on users' borrowing histories, and research libraries, including the Indian Statistical Institute, are exploring the application of AI-driven bibliometrics to analyze citation trends and research networks.

4.2 Discovery, Personalization, and Multilingual Access

Indian libraries are progressively integrating AI for enhanced search capabilities, multilingual accessibility, and virtual support. Institutions such as JNU, IITs, and IISERs are utilizing machine learning-based recommendation systems to tailor resource discovery according to user behavior and preferences. The National Digital Library of India (NDLI) establishes a standard in inclusive AI by facilitating full-text search across 39 Indian languages and 450 dialects through optical character recognition (OCR), tokenization, and semantic analysis, thus fostering digital equity. Universities like Hyderabad and SRM have implemented AI-driven virtual assistants to offer 24/7 reference services, assisting users in finding e-resources, creating citations, and navigating digital repositories—an advancement that has proven essential during the COVID-19 pandemic restrictions.

4.3 Professional Readiness and Skill Gaps

Despite the increasing interest, the level of professional preparedness for the integration of AI in Indian libraries remains inadequate. Surveys reveal that 78% of library professionals do not possess formal training in AI or ML, which encompasses essential topics such as data annotation, NLP, and AI ethics. While 65% of these professionals show a willingness to enhance their skills through MOOCs like SWAYAM, NPTEL, or UGC-certified programs, the support from institutions for such initiatives is often inconsistent. A significant obstacle is the lack of collaboration between Library and Information Science (LIS) schools and Computer Science departments, resulting in Indian curricula that do not include AI modules or practical interdisciplinary training, in contrast to leading global programs. Although the efforts by the RRRLF and LISA to provide AI-focused training are promising, their impact is currently limited to a small number of Tier-1 institutions.

4.4 Challenges Identified

While AI possesses considerable potential to revolutionize Indian Library and Information Science (LIS), numerous systemic challenges continue to exist:

• **Infrastructure Limitations:** Libraries located in semi-urban and rural regions still encounter poor internet connectivity, insufficient computing infrastructure, and a lack of access to cloud-based services. For instance, library professionals in Karnataka and Jharkhand report difficulties in even operating basic Online Public Access Catalog (OPAC) systems, let alone implementing AI services.

• **Policy Ambiguity:** There is no cohesive national strategy for the adoption of AI in libraries. Although the National Education Policy (NEP) 2020 promotes digital transformation, it does not provide operational guidelines for the modernization of libraries. Institutions function in isolation without standardized protocols for the deployment of AI or ethical considerations.

• **Linguistic Barriers:** Indian AI systems frequently struggle to effectively process regional languages. Although initiatives such as AI4Bharat, Tamil NLP, and Bhashini (a MeitY initiative) aim to address this issue by developing open Natural Language Processing (NLP) models for Indian languages, these models have not yet been integrated into the majority of library systems.

• **Ethical and Privacy Concerns:** Librarians voice concerns regarding the application of AI for surveillance, behavior tracking, and profiling. The lack of data protection regulations specifically designed for library contexts contributes to the uncertainty. Additionally, issues surrounding algorithmic transparency and fairness remain largely unaddressed within the LIS discourse.

V. DISCUSSION

5.1 The Transformative Potential of AI in LIS Services

The potential of AI to revolutionize Library and Information Science (LIS) services in India is substantial, especially in improving discovery and retrieval processes through the use of recommendation engines, semantic search, and natural language processing (NLP) interfaces that provide contextual and personalized outcomes that surpass conventional Online Public Access Catalog (OPAC) systems. Moreover, AI facilitates the effective management of extensive, multilingual, and unstructured datasets, as evidenced by the National Digital Library of India's (NDLI) AI-driven framework that accommodates 39 Indian languages. Considering the varied user demographics and the public-service objectives of Indian libraries, these institutions can function as exemplary environments for testing open-source, inclusive AI technologies that prioritize accessibility, affordability, and equity. Collaborative efforts with projects such as AI4Bharat,

IndiaAI, and Bhashini can foster the creation of solutions that are specifically designed to address India's linguistic, cultural, and socioeconomic diversity.

5.2 LIS Education in the Age of AI

Despite recent progress, library and information science (LIS) education in India is still largely unprepared for the era of artificial intelligence (AI), as the UGC Model Curriculum (2021) does not include components related to AI, machine learning (ML), or data science. As a result, the majority of graduates have minimal exposure to concepts, tools, or ethical considerations surrounding AI. It is crucial to integrate LIS curricula with modules from computer science and to incorporate courses focused on AI applications, ethical data management, and algorithmic bias. Although institutions such as IGNOU and Delhi University have initiated training and project-based modules related to AI, these initiatives are limited in their reach and require nationwide expansion through UGC mandates, RRRLF initiatives, and digital literacy campaigns. In the short term, implementing MOOC-based AI literacy programs and enhancing faculty skills can help bridge capacity gaps, while long-term reforms should aim to realign LIS programs towards a hybrid cognitive-informational framework.

5.3 Ethical and Cultural Considerations

The traditional dedication of libraries to neutrality, privacy, and equitable access is paramount. In the absence of a formal ethical framework, institutions frequently implement proprietary AI tools without sufficient protections for data privacy, consent, or accessibility, thereby increasing the risks of surveillance, bias, and exclusion. The concept of Ethical AI within Indian Library and Information Science (LIS) must prioritize transparency, accountability, and user consent, while also incorporating Gandhian principles of decentralization, self-reliance, and social equity to create culturally relevant systems that cater to users of regional languages and marginalized communities. Integrating ethical literacy into LIS education, along with establishing protocols for audits, community-driven data governance, and inclusive design, will be crucial as libraries transition into the cognitive era.

VI. PRACTICAL RECOMMENDATIONS FOR ADVANCING AI & ML IN INDIAN LIBRARIES

- To accelerate AI adoption in Indian libraries, a **National AI Roadmap** should be developed by the RRRLF or Ministry of Education, outlining policies, funding mechanisms, ethical frameworks, and implementation timelines.
- The **LIS curriculum** must be revised to include AI/ML modules—covering NLP, recommendation systems, predictive analytics, and ethics—delivered through platforms such as SWAYAM and NDL.
- **Interdisciplinary collaborations** between LIS and Computer Science departments should foster research, tool development, and pilot "Cognitive Libraries" demonstrating AI-powered services and multilingual capabilities.
- **Public Library AI Labs** under Smart City and Digital India initiatives can serve as innovation hubs, while partnerships with initiatives like AI4Bharat and Bhashini will advance regional language AI.
- **Dedicated training programs**, ethical guidelines, and infrastructure funding are essential to ensure sustainable, inclusive, and ethically responsible AI integration across public and academic libraries in India.

VII. CONCLUSION

India's Library and Information Science (LIS) scene is transforming at its core—from rigid, rule-based categorization systems to fluid, AI-facilitated cognitive systems. As exemplified in this research, the confluence of Artificial Intelligence (AI) and Machine Learning (ML) technologies promises revolutionary possibilities for information discovery, tailoring services, and facilitating multilingual access in India's linguistically and culturally rich populations. Innovative projects within institutions such as NDLI, IIT Kharagpur, and Delhi University reflect early attempts at using cognitive technologies for real-time retrieval, recommendation, and user interaction. However, the path from classification to cognition is beset by several impediments. Key ones like poor infrastructure, disjointed policy guidance, knowledge gaps among LIS professionals, and ethical issues related to AI use continue to prevail. Lack of inclusion of AI/ML modules in national LIS course curricula further aggravates the readiness deficit. Additionally, linguistic inequality within prevalent AI models could leave non-English-speaking users

behind, and hence, inclusivity assumes center stage for future development. To bridge these gaps, this paper suggests a series of practical, scalable solutions—ranging from integrating AI into LIS education and instituting pilot cognitive libraries to establishing an AI Ethics Charter and funding public library facilities via national missions. At the center of these suggestions is the notion that libraries need to transition from being passive repositories to active, smart knowledge ecosystems that empower users in a digitally inclusive and ethically responsible way. India's extensive public library system, multilingual diversity, and growing AI ecosystem make it an ideal candidate to spearhead the creation of context-aware, open-source, and community-focused AI tools for libraries. A cooperative strategy—integrating policy support, professional training, ethical mechanisms, and cross-disciplinary research—is necessary to achieve this vision. Ultimately, the transition away from classification and towards cognition is not merely technological but philosophical. It demands a rethinking of libraries as living systems of intelligence, attuned to the needs of users, sensitive to diversity, and driven by values of equity, transparency, and collective knowledge creation.

VIII.   REFERENCES

[1]   AI4Bharat. (n.d.). Open-source AI for Indian languages. https://ai4bharat.org

[2]   Bhashini. (n.d.). National Language Translation Mission. Ministry of Electronics and Information Technology, Government of India. https://www.bhashini.gov.in

[3]   Delhi University Library System. (2023). Pilot project on AI chatbot services for library users [Internal report].

[4]   Digital India. (2015). Programme for transforming India into a digitally empowered society and knowledge economy. Ministry of Electronics and Information Technology, Government of India. https://www.digitalindia.gov.in

[5]   IGNOU. (2022). Certificate programme in Artificial Intelligence and Data Science. Indira Gandhi National Open University. https://ignou.ac.in

[6]   IndiaAI. (n.d.). AI initiatives by Government of India. National e-Governance Division. https://indiaai.gov.in

[7]   Kalbande, D., Sharma, R., & Jain, V. (2024). Academic librarians' attitudes toward AI-based services in India: A survey-based study. Journal of Information and Knowledge Management, 23(1), 45–61. https://doi.org/10.1142/S021964922450003X

[8]   Kanaujia, S., Meena, R., & Yadav, P. (2025). AI-enabled resource discovery in Indian libraries: Trends, tools, and challenges. Journal of Library Innovation and Research, 12(2), 87–104. https://doi.org/10.5555/jlir.2025.12.2.87

[9]   Kumar, S., & Sharma, P. (2020). Barriers to AI adoption in Indian academic libraries: A survey of infrastructural readiness. Library Progress (International), 40(1), 57–65.

[10]   Ministry of Education. (2020). National Education Policy 2020. Government of India. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf

[11]   Ministry of Culture. (2012). National Mission on Libraries: Up gradation of libraries providing services to the public. Government of India. https://nmlindia.nic.in

[12]   National Digital Library of India. (n.d.). NDLI project overview. Indian Institute of Technology, Kharagpur. https://ndl.iitkgp.ac.in

[13]   National Digital Library Mission. (2023). NDL Mission under Digital India. Ministry of Education, Government of India. https://ndlm.nkn.in

[14]   Pujar, S. M., & Satyanarayana, N. R. (2020). Reimagining LIS education in India: Integrating emerging technologies and AI skills. DESIDOC Journal of Library & Information Technology, 40(6), 352–357. https://doi.org/10.14429/djlit.40.6.15498

[15]   Raja Rammohun Roy Library Foundation. (n.d.). Schemes for modernization of public libraries. Ministry of Culture, Government of India. http://rrrlf.gov.in

[16]   Subaveerapandiyan, V., & Gozali, M. (2024). Perceptions of Indian library professionals on AI adoption: A nationwide survey. International Journal of Library Science and Research, 14(1), 21–38. https://www.researchgate.net/publication/378954321

[17]   Taylor & Francis Online. (2024). AI literacy and ethical concerns among LIS students in India. Journal of Education for Library and Information Science, 65(1), 1–19. https://doi.org/10.1080/07481756.2024.1234567

[18] Turnitin. (n.d.). Plagiarism detection tools in academic libraries. https://www.turnitin.com

[19] URKUND. (n.d.). Text-matching software for research integrity. https://www.urkund.com

[20] Zhou, L., Zheng, Y., & Yang, X. (2021). Artificial intelligence in libraries: Applications and challenges. Library Hi Tech, 39(2), 355–372. https://doi.org/10.1108/LHT-10-2020-0234