# CONTEXTUALIZING DATA LITERACY IN LIBRARIES:
## Opportunities, Challenges, and Research Prospects

**Sourav Mazumder**
Librarian, Ghoksadanga Birendra Mahavidyalaya, West Bengal
E-mail id: smazumderlis91@gmail.com

**Tapan Barui**
Assistant Professor, DLIS, University of North Bengal, West Bengal
E-mail id: tapanbarui@nbu.ac.in

***Abstract:***

***Purpose***

*This study aims to contextualize data literacy in libraries by exploring research themes, challenges, opportunities, and future research prospects.*

***Design/methodology/approach***

*A total of 160 research papers were retrieved from the Dimensions AI app. A qualitative content analysis was employed in this study.*

***Findings***

*The study identifies six major thematic areas. Findings reveal that libraries are progressively integrating data literacy to support user communities. However, there is a need for an evaluation of policies and frameworks that can reduce barriers to adopting technology, address skill gaps, and address ethical concerns. The study presents future research prospects in expanding data literacy frameworks, big data analytics, and advanced research data management.*

***Originality/value***

*This study contextualizes data literacy in libraries by analyzing research themes. It presents a conceptual framework and brings unique insights for librarians and data nerds to implement data literacy initiatives.*

***Keywords:*** *Qualitative Content Analysis; Data Literacy Skills; Library Science; Data Librarians; Big Data Analytics*

## 1 Introduction:

Literacy is not about reading or writing. It is embedded in social, cultural, and political contexts through ongoing social interactions and power dynamics (Bartlett, 2008). UNESCO (2021) defines: "...literacy is now understood as a means of identification, understanding, interpretation, creation, and communication in an increasingly digital, text-mediated, information-rich and fast-changing world." There are various forms of

literacy, such as computer literacy, scientific literacy, ecological literacy, and media literacy (Wikipedia Contributors, 2025). In the context of library and information science, information and media literacy are vital aspects (ALA, 2015; Sahu et al., 2023). In the digital era, data is being generated in large volumes in the forms of text, image, audio, and video. Logan (2018, p. 3) states, "data literacy is the new core capability of digital society." Data literacy involves skills in data collection, data management, data evaluation, and data application (Ghodoosi et al., 2023). In library environments, data literacy can be conceptualized as a vital skill in modern-day library systems for accessing, assessing, and managing data ethically and responsibly (Koltay, 2017). The present study aims to contextualize data literacy in libraries by exploring thematic areas, challenges, opportunities, and research prospects.

### *1.1 Research Objectives:*

Objectives of the study include: (1) to identify research themes in data literacy that align with the library context; (2) to explore emerging opportunities in data literacy practices; (3) to identify the challenges faced in literacy initiatives; and (4) to present research prospects for advancing data literacy.

## 2 Related Work:

Some closely related studies have been reviewed here. Khan and Du (2017) highlighted the rise of big data and new roles like data librarians by employing content analysis of job advertisements. Pinto et al. (2023) stated that data literacy remains underdeveloped compared to information literacy. Nwagwu (2024) highlighted emerging concerns, including ethics, privacy, and digital transformation, and emphasized libraries as central actors in data literacy. Oh and Mon (2024) found that public libraries rarely offer explicit data literacy programs and stated the integration of data literacy education with existing programs. Adekoya et al. (2025) revealed that regular training through workshops and conferences can develop "data-savvy" librarians. Pratama (2025) pivoted the progress of academic libraries in RDM and proposed a model for strengthening data support. These studies show various areas of data literacy in libraries. An abstract-level study might provide more information within this field of study. This paper draws on existing research publications focused on data literacy and libraries to identify major research areas, challenges, opportunities, and future research prospects.

## 3 Methods:

This study employed a qualitative content analysis. First, a dataset of 198 bibliographic records was downloaded from the Dimension AI database (for accessing the free version and significant coverage) using a search string ("data literacy" AND "library") in the Title and Abstract fields. After a thorough manual scrutiny, a total of 38 records (not aligned with the library) were excluded. Finally, 160 papers were taken for analysis. Second, the dataset includes journal articles, chapters, and conference papers; some of the chapters

have the same abstracts and blank text. In that case, each abstract was comprehended carefully. Lastly, 140 abstracts were analyzed to achieve the objectives of the study. The analysis comprises an open coding (manual) technique to categorize themes and other aspects relevant to this study. Each code was determined by keywords from abstracts. ChatGPT (OpenAI, 2023) was used to match categories for creating themes. To validate this process, intra-coder reliability was assessed and scrutinized. Finally, this study also presents a conceptual framework that links challenges, opportunities, and research prospects in data literacy. Google Sheets was used for data organization, analysis, and visualization.

## 4 Results and Discussion:

### *4.1 Trend Analysis:*

The dataset for this study revealed that the first publication in this field was traced in 2005 (Figure 1 (A)). The growth of publications was seen from 2016. It can be said that the importance of data literacy emerged during that particular period. A good number of papers were published during 2020-2023. The influencing factor could be the COVID-19 period and the post-COVID-19 period for research productivity. Other factors might align with the rise of big data. The dataset has journal articles (n = 119), book chapters (n = 29), conference proceedings (n = 6), edited books (n = 3), and preprints (n = 3). Journal articles are the primary drivers of scholarly discourse (Figure 1 (B)). Figure 1 (C) shows that the dataset has 62 Gold Open Access (OA), 20 Hybrid OA, 6 Green OA, and 72 closed-access papers. Figure 1(D) indicates that "Quality Education" (n = 59) is the most frequently represented theme. It suggests the libraries' active role in promoting data literacy.

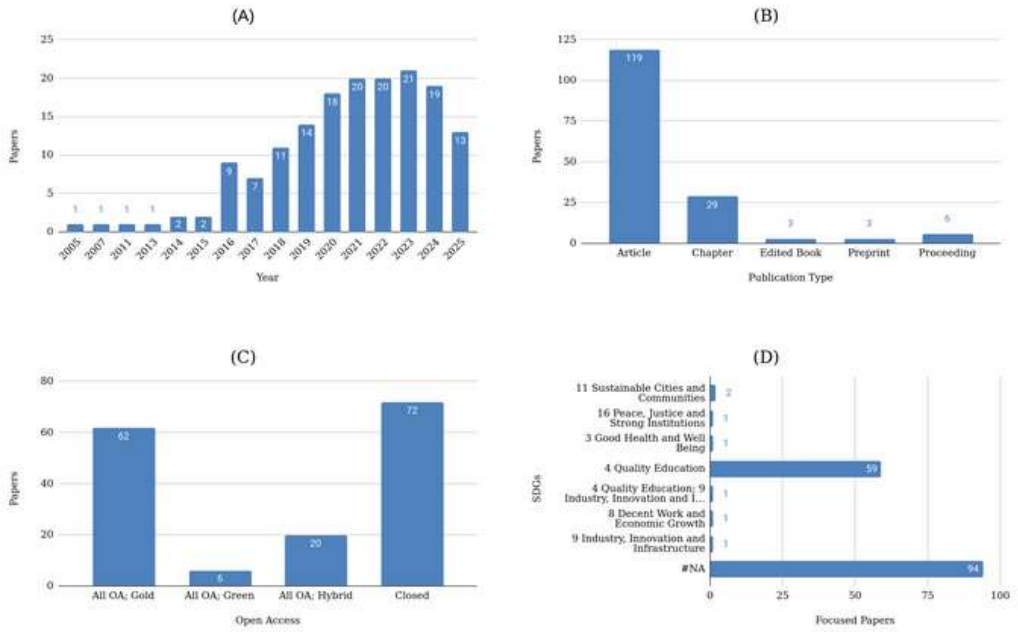



**Figure 1**: (A) Yearly publication, (B) Types of documents/papers, (C) Access type, (D) Sustainable Development Goals category

### *4.2 Major Thematic Areas:*

The analysis of the first objective of the study exhibits six themes (Figure 2) that explore how data literacy is perceived in libraries. The theme, "Integration of Data Literacy in Library Services," reflects the association between data literacy aspects and library services. It also supports the theme "Library Professionals' Skills." Das (2020) stated that library services can be enhanced through data-related collaborations. The theme "Data Management" has been found to be one of the significant research areas. It concerns RDM and the role of libraries in the strategic plan (e.g., Adekoya et al., 2025; Pratama, 2025). Next, "User Education" generally aims to train and instruct library users through programs such as workshops and seminars (Kim et al., 2024). Furthermore, "Technological Infrastructure and Policy Support," with institutional backing, digital environments, and adequate policies, plays a crucial role in sustaining data literacy programs. Lastly, "Public Libraries and Community Engagement" demonstrates how public libraries serve as facilitators for expanding the data literacy dimension and promote community participation in the information age. Overall, these themes give insights into the multidimensional context of data literacy in libraries.

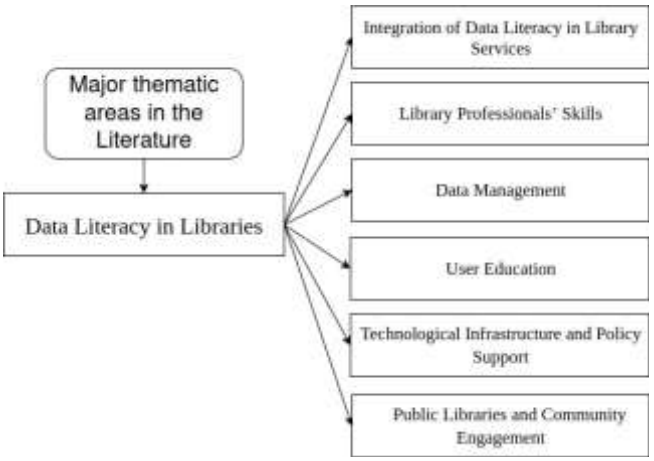



**Figure 2**: Major thematic areas

### *4.3 Conceptual Framework:*

To present a conceptual framework for data literacy, this study identifies and links relevant challenges, opportunities, and research prospects (Figure 3). First, challenges encompass several factors, including limited resources, difficulties with existing data literacy model assessments, skill gaps, policy gaps, ethical issues, and technological barriers (e.g., Goertzen, 2017). These limit the implementation of adequate data literacy and the setting of proper frameworks. The skill gaps of users, researchers, and library professionals in data management indicate an important area for improvement. Ethical issues and sustainability highlight the significance of data privacy, integrity (Nwagwu, 2024), and long-term viability.

Figure 3 also shows the opportunities that emphasize the integration of curriculum with data literacy, collaboration, open science, technology acceptance, and professional

development as key factors of progress (Khan & Du, 2017; Oh & Mon, 2024; Willaert et al., 2019). Library services can be enhanced by adopting new technologies, engaging with the community, and promoting open science (e.g., Pinto et al., 2023). Additionally, data science can help libraries and institutions align their data literacy with a broader perspective.

This study identifies research prospects. Researchers can expand and explore data literacy frameworks or models, advance RDM practices, investigate data science areas, and their applications in libraries. Assessing the impact of data literacy, cross-disciplinary collaboration, and comparative research can be essential prospects for future research. Finally, artificial intelligence (AI) and big data management have emerged as forcing factors for a better data ecosystem in libraries. These can enhance data literacy through strategic planning.

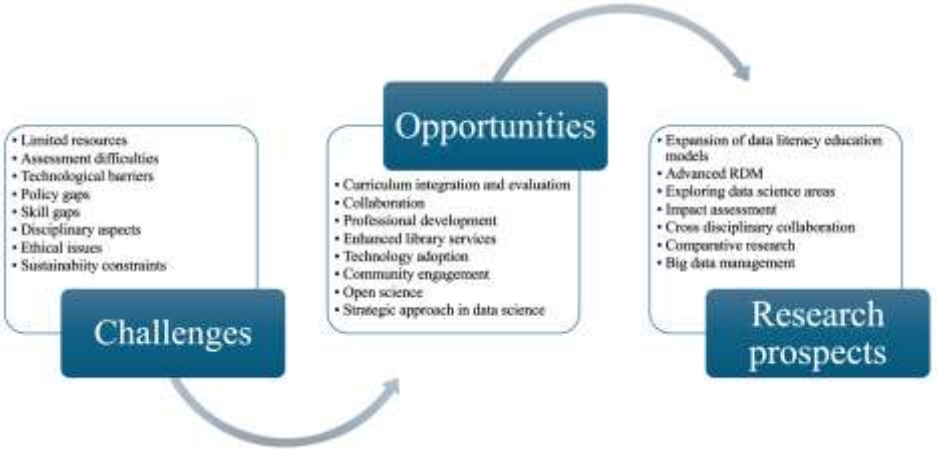



**Figure 3**: A conceptual framework: contextual factors (challenges and opportunities) and outcomes (research prospects)

## 5 Conclusions:

This study conducted a qualitative content analysis of abstracts of data literacy papers. It is limited to one database, only abstract-level analysis, and a selected search string. Further studies can execute a more comprehensive analysis. However, the findings show promising revelations regarding the current approach to data literacy in libraries. Libraries can act as data literacy platforms for data management and ethical data use. The findings signify adequate policies, technological infrastructure, and skill development for sustainable data literacy programs.

**Acknowledgment**: Grammarly was used for language enhancement.

## References:

Adekoya, C. O., Nkemdilim, I. P., Ejovwokoghene, E. R., Olajide, O., & Omolehin, J. R. (2025). Data science: Developing data-savvy librarians for effective research data management. *The Reference Librarian*, *66*(1–2), 48–64. https://doi.org/10.1080/02763877.2025.2482086

ALA. (2015). *Framework for Information Literacy for Higher Education | Association of College and Research Libraries*. https://www.ala.org/acrl/standards/ilframework

Bartlett, L. (2008). Literacy's verb: Exploring what literacy is and what literacy does. *International Journal of Educational Development*, *28*(6), 737–753. https://doi.org/10.1016/j.ijedudev.2007.09.002

Das, T. (2020). Academic Library Collaborations to Strengthen Open Government Data and Expand Librarianship. In *Open Government: Concepts, Methodologies, Tools, and Applications* (pp. 1187–1206). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-5225-9860-2.ch056

Ghodoosi, B., West, T., Li, Q., Torrisi-Steele, G., & Dey, S. (2023). A systematic literature review of data literacy education. *Journal of Business & Finance Librarianship*, *28*(2), 112–127. https://doi.org/10.1080/08963568.2023.2171552

Goertzen, M. (2017). Mixed Method Study Examines Undergraduate Student Researchers' Knowledge and Perceptions About Scholarly Communication Practices. *Evidence Based Library and Information Practice*, *12*(3), 159–161. https://doi.org/10.18438/B85W9P

Khan, H. R., & Du, Y. (2017). *What is a Data Librarian?: A Content Analysis of Job Advertisements for Data Librarians in the United States Academic Libraries*. https://repository.ifla.org/handle/20.500.14598/6403

Kim, J., Evans, S., Hong, L., Rice-Oyler, E., & Ogbadu-Oladapo, L. (2024). Data literacy in flux: Perspectives of community college librarians on evolving educational demands and library capacities. *Library & Information Science Research*, *46*(4), 101327. https://doi.org/10.1016/j.lisr.2024.101327

Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science*, *49*(1), 3–14. https://doi.org/10.1177/0961000615616450

Logan, V. (2018). *Information as a Second Language: Enabling Data Literacy for Digital Society*. Gartner, Inc. https://winfo.ro/wp-content/uploads/2020/04/Informations-as-a-Second-Language-Gartner.pdf

Nwagwu, W. E. (2024). Mapping the field of global research on data literacy: Key and emerging issues and the library connection. *IFLA Journal*, *50*(3), 491–510. https://doi.org/10.1177/03400352241257669

Oh, S., & Mon, L. (2024). Community-based learning and data literacy: The role of the public library. *Information and Learning Sciences*, *125*(7–8), 456–474. https://doi.org/10.1108/ILS-06-2023-0078

OpenAI. (2023). *ChatGPT* [Large language model]. https://chatgpt.com/

Pinto, R., Lacerda, J., Silva, L., Araújo, A. C., Fontes, R., Lima, T. S., Miranda, A. E., Sanjuán, L., Gonçalo Oliveira, H., Atun, R., & Valentim, R. (2023). Text mining analysis to understand the impact of online news on public health response: Case of syphilis epidemic in Brazil. *Frontiers in Public Health*, *11*. https://doi.org/10.3389/fpubh.2023.1248121

Pratama, A. R. (2025). A study of data-focused roles in leading North American university libraries. *The Journal of Academic Librarianship*, *51*(5), 103102. https://doi.org/10.1016/j.acalib.2025.103102

Sahu, L., Mazumder, S., & Deb, S. (2023). How students identify and evaluate fake news: A case study of a management college. *Qualitative and Quantitative Methods in Libraries*, *12*(1), Article 1. https://qqml-journal.net/index.php/qqml/article/view/794

UNESCO. (2021). *Literacy: What you need to know*. https://www.unesco.org/en/literacy/need-know

Wikipedia Contributors. (2025). Literacy. In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Literacy&oldid=1314126336

Willaert, T., Cottyn, J., Kenens, U., Vandendriessche, T., Verbeke, D., & Wyns, R. (2019). Research data management and the evolutions of scholarship: Policy, infrastructure, and data literacy at KU Leuven. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, *29*(1), 1–19. https://doi.org/10.18352/lq.10272