DCPAPERS

DCMI 2026

Proceedings of the International Conference on Dublin Core and Metadata Applications 2026

ISSN: 1939-1366

Reconstructing Metadata Literacy in the AI Era: A Conceptual Framework and Educational Reflections for LIS Education

Ba Xi[1], **Nurussobah Hussin**[2], and **Hanis Diyana Kamarudin**[2]

[1] Faculty of Information Science, Universiti Teknologi MARA Cawangan Selangor ; Hebei University of Economics and Business, [2] Faculty of Information Science, Universiti Teknologi MARA Cawangan Selangor

Abstract

This paper calls for a rethinking of metadata literacy in the age of AI. Previous discussions often defined metadata literacy as knowledge of descriptive structures or skills in creating and using metadata records. But this understanding is no longer enough — though still necessary — when AI systems create, transform, rank, summarise and recommend information at scale. In such environments, metadata is not simply about post hoc description of resources; it structures provenance, visibility, accountability, cultural representation, and the conditions under which machine outputs can be interpreted and trusted. Using metadata, metadata instruction, AI literacy, and information literacy scholarship, this conceptual paper presents a reconstructed model of metadata literacy for LIS education. The model is built on five dimensions: understanding of infrastructure, contextual description and representation, provenance and disclosure, evaluation of algorithmically mediated outputs and intervention in terms of ethics and governance. The paper also gives examples of learning tasks and assessment evidence to illustrate how the model can be applied. It argues for viewing metadata literacy not as a narrow technical specialisation but as a fundamental educational response to AI-mediated knowledge environments.

1. Introduction

Metadata literacy has generally been understood as the capacity to recognise, interpret, create, and use metadata within information environments. This understanding is still relevant in library and information science (LIS) where metadata enables discovery, organization, access, preservation, and interoperability [1], [2], [3]. It also provides students with a practical language for describing information objects, and for making records usable between systems.

But AI changes the conditions under which metadata operates. Metadata in AI-mediated environments supports more than just retrieval or back-end management. It affects which content is selected, ranked, summarised, linked, and explained. It also has implications for the traceability of provenance and representation of cultural identities, and user assessments of the trustworthiness and reviewability of machine-mediated output [1], [4]. As these functions expand, a narrow conception of metadata literacy as primarily technical or record-level competence becomes insufficient.

Research into AI literacy, meanwhile, has demonstrated that meaningful engagement with AI requires more than operational familiarity; AI literacy refers to the understanding of AI concepts, the ability to assess outputs, the ability to communicate and work with AI systems and to ethical issues [5], [6]. A related point has been made in recent

Full Paper

Available Aug 01, 2026

Correspondence to
Nurussobah Hussin
nurussobah@uitm.edu.my

discussions of information literacy: rather than taking the outputs of AI as self-evident answers [7], [8], learners need to critically interpret information generated by AI.

This paper focuses on these discussions but does not collapse metadata literacy into AI literacy or information literacy. AI literacy is defined as the general skills to use AI. Information literacy is the ability to critically use and evaluate information. Metadata literacy is one professional layer where many of these capacities are put to work: description, classification, provenance, versioning, disclosure, and the governance of representational choices. The critical question is not just whether LIS students can use AI tools, but whether they can understand and intervene in the metadata conditions through which AI-mediated information becomes findable, interpretable, and accountable.

The paper develops this argument in four stages. It begins by discussing the relationship between metadata literacy, AI literacy and information literacy. Second, it explains the necessity of reconstructing metadata literacy in the AI era. Third, it presents a five-dimension model of reconstructed metadata literacy. Finally, the paper discusses how the model can be operationalised in LIS curricula through course activities, scenarios and assessment evidence.

2. Theoretical basis: metadata literacy and AI literacy

In LIS, the metadata has long been treated not only as a neutral technical add-on. Metadata is a sociotechnical phenomenon that structures relationships between objects, contexts, systems and communities, as defined by Mayernik [1]. The work on metadata quality has also demonstrated that metadata supports not only discovery but also use, provenance, currency, authenticity, administration and long term stewardship [4]. From this perspective, metadata practice is therefore always a question of judgement: what features to record, what vocabularies to use, what relationships to foreground, what absences or ambiguities might matter later.

This broader understanding partly led to the development of metadata literacy. Mitchell's study of college students positioned metadata literacy as awareness and interpretive competence around the metadata embedded in information environments [2]. Subsequent work on metadata instruction expanded this focus from awareness to skill building, showing that metadata education requires learners to make deliberate decisions about description, subject representation, and standards-based record creation [3]. Although the term has been used inconsistently, there is a clear common emphasis: metadata literacy relates to the ability to work with structured description in meaningful, context-sensitive, and fit-for-purpose ways.

AI literacy has been born out of a different, but related, problem space. Long and Magerko define AI literacy as a set of competencies enabling people to critically evaluate AI technologies, communicate and collaborate with AI, and use AI in everyday contexts [5]. Ng et al. also show that AI literacy consists of knowledge, application, evaluation and ethical reflection rather than only technical know-how [6]. Library ori-

Table 1. *Position of reconstructed metadata literacy in relation to adjacent literacies.*

Literacy frame	Primary concern	Distinct contribution to this paper
Conventional metadata literacy	Recognizing, interpreting, creating, and using metadata records.	Gives the standards-based foundation, but can remain too record-centred for AI-mediated environments.
AI literacy	Understanding, using, evaluating, and reflecting ethically on AI systems.	Explains broad AI capacities, but not how provenance, representation, and disclosure are enacted through metadata.
Information literacy	Finding, evaluating, using, and creating information critically and responsibly.	Supports critical judgement, but does not fully address metadata as infrastructure.
Reconstructed metadata literacy	Understanding and governing the metadata conditions through which resources, AI outputs, versions, and interventions become intelligible.	Connects description with provenance, accountability, representation, and AI-mediated reviewability.

entated frameworks maintain this wider view, but translate it to professional contexts, particularly in terms of responsible use, policy awareness, evaluation of outputs and development of role specific competence [9], [10], [11], [12].

Thus their relationship is complementary, not interchangeable. In brief, AI literacy is how people understand and utilise AI systems . Information literacy is how people find, assess, and use information . Metadata literacy is how information objects and machine outputs are described, contextualised, traced, and governed . These questions come together in LIS contexts whenever a professional needs to decide how to describe, disclose, organise, evaluate, or revise an object that is AI-generated or AI-mediated.

3. Why metadata literacy requires reconstruction in the AI era

Three developments compel reconstruction. First, AI increases the number of objects with unstable status. Records can now be connected to born-digital resources, machine-generated summaries, synthetic images, enriched transcriptions, automatically extracted entities and dynamically updated outputs. In these cases the metadata is no longer simply attached to a stable document. It becomes part of how you make the status, source, and degree of human or machine intervention of an object legible.

Second, the AI magnifies the quality of metadata. Traditional concerns of completeness, consistency and standardisation still matter, but AI-mediated environments add new pressures concerning provenance, explainability, disclosure and bias. A poor or incomplete record can affect not just whether a resource is found, but how an output is ranked, summarised, linked, or trusted. Metadata choices can therefore shape interpretation before a user encounters the source object directly [4], [13].

Third, AI exposes the inextricable links between metadata work and representation and power. Metadata is not only describing the world, it is also involved in the production of what is visible, comparable, or classifiable in systems. Recent work on cultural dimensions of metadata quality suggests that metadata problems may be failures of representation, naming, geographic framing, and community recognition, rather than simply technical defects [13]. The implications may go further and faster, once those records are ingested, transformed or amplified by AI systems.

The point is made with a simple educational example. A student uses an AI tool to enrich the records for a small digital collection. The tool suggests subject terms, writes summaries and translates titles. A traditional metadata exercise might look into whether the fields are complete, and whether the values conform to a schema . A reconstructed metadata literacy exercise asks other questions: Did anyone confess to using AI? Is the

Table 2. *Proposed dimensions of reconstructed metadata literacy in the AI era.*

Dimension	Central question	Educational emphasis
1. Infrastructural understanding	How do metadata standards, schemas, platforms, and AI pipelines shape what can be seen, linked, or acted upon?	Teach metadata as infrastructure rather than as isolated fields and tags.
2. Contextual description and representation	How should resources, people, places, and machine interventions be described in contextually adequate ways?	Develop decisions about vocabularies, naming, subject representation, and scope notes.
3. Provenance and disclosure	What must be recorded about source, authorship, version, transformation, and AI involvement for later interpretation?	Strengthen abilities to document lineage, intervention, and machine contribution.
4. Evaluation of algorithmically mediated outputs	How should learners assess records or outputs that are enriched, ranked, summarized, or generated through AI?	Move evaluation beyond accuracy to include reviewability, contextual fit, and risk.
5. Ethical and governance intervention	When should metadata be revised, contested, annotated, or supplemented to reduce harm and improve accountability?	Connect metadata decisions to fairness, transparency, stewardship, and institutional policy.

proposed term verified from the source? The names, places and languages are correctly represented. Is uncertainty measured? If a subject term is problematic should it be accepted, revised, annotated or rejected? Such questions make the work of metadata a matter of judgement, not just data entry.

So the transition isn't from metadata to AI. It is from a narrower view of metadata literacy as a descriptive skill to a broader sense of metadata literacy as situated judgement in AI-mediated information infrastructures. This is an important shift for LIS education because librarians and information professionals are often the interpretive layer that links systems, records, users, and institutional values [10], [12].

4. A reconstructed model of metadata literacy

Based on the literature above, metadata literacy in the AI era can be re-constructed around five inter-related dimensions. Although the model is designed for LIS education, it is equally valuable in framing professional development in adjacent information contexts. The dimensions are not a collection of discrete skills. Rather, they outline interrelated forms of judgement that students can practise through increasingly complex learning tasks.

The first dimension, infrastructural understanding, considers metadata as one part of the building blocks of information systems. Learners need to understand not only what a field is, but also how schemas, mappings, identifiers and platform constraints impact the visibility and reusability of data. Here, the cross-section of metadata literacy and AI literacy is at the level of system awareness: an end-user cannot properly assess an AI-mediated output without some knowledge of the descriptive and infrastructural conditions that made that output possible [1], [5], [6].

The second dimension, contextual description and representation, takes on the ongoing core of the metadata work. Description is not collapsable to technical coding. It is about judgement of relevance, granularity, naming, and the relation of controlled vocabularies to lived identities. This dimension is more important, not less, in the AI era, since downstream systems can inherit and scale representational decisions encoded within metadata [4], [13].

The third dimension, provenance and disclosure, addresses one of the most pressing literacy challenges in AI-mediated environments. Users and professionals need to know more and more whether a record is based on original human description, automated

extraction, machine translation, AI-assisted enhancement or synthetic generation. Metadata literacy is about knowing what provenance information is required, how it should be recorded and why disclosure is important for interpretation and trust [1], [14], [15].

The fourth dimension, evaluation of algorithmically mediated outputs, builds on familiar evaluation practices in contexts in which metadata and AI interact directly. A user may want to assess an auto-generated summary, a relevance ranking, a suggested subject term, or an enriched record. The question is not merely whether the output is factually accurate. It is also about whether the output is contextually adequate, reviewable against the source, properly disclosed, and suitable for the service context in which it will be deployed [5], [7], [10].

The fifth dimension, ethical and governance intervention, positions metadata literacy as an active competence. In the AI era, being metadata literate is not only about identifying problems, but also about knowing when metadata needs to be reviewed, annotated, escalated or governed differently. This dimension connects metadata literacy to policy, accountability and professional stewardship. It is also where LIS education can most clearly distinguish metadata literacy from a narrow technical training model [9], [10], [13], [14].

5. Educational reflections for LIS curricula

If metadata literacy is reconstructed in this way, then LIS curricula should not consider it to be a specialised niche confined to cataloguing or metadata modules. Instead, it should be a transversal educational thread connecting knowledge organization, digital curation, information retrieval, scholarly communication, archives, and AI-related teaching.

The first implication is one of curricular integration. Introductory courses can present metadata not just as a technical language of description, but also as a social and infrastructural medium that makes information actionable. Metadata can be revisited in later courses in terms of AI-assisted discovery, machine-generated content, repository workflows, research data management and digital collections. This spiral approach aligns with the fact that information literacy and AI literacy frameworks tend to encourage iterative development rather than one-off exposure [6], [7], [8].

A second implication is the redesign of learning activities. Typical exercises require students to create records from a schema. That's still useful, but should be supplemented with evaluative and comparative tasks. For example, students may compare human-created records with AI-enriched records, document the provenance and history of interventions, identify disclosure failures, or revise problematic subject representation. These kinds of tasks can better reflect the kinds of judgement-laden conditions under which metadata is now produced and used [3], [4].

Table 3. *Sample operationalization of the five dimensions in LIS education.*

Dimension	Sample learning task	Possible assessment evidence
Infrastructural understanding	Map how a digital collection record moves from local metadata fields to a discovery interface and an AI-assisted search result.	A short workflow map explaining where metadata values are transformed, lost, or reused.
Contextual description and representation	Describe the same item twice: once using a standard schema and once with added scope notes for cultural or disciplinary context.	A record plus a justification explaining vocabulary choices, granularity, and representational risks.
Provenance and disclosure	Revise an AI-enriched record by adding source, version, tool-use, and human-review statements.	A provenance note and disclosure statement that make the record reviewable by another person.
Evaluation of algorithmically mediated outputs	Compare an AI-generated summary or subject recommendation with the original source and metadata record.	An evaluation memo identifying accuracy, omissions, hallucinations, contextual fit, and action taken.
Ethical and governance intervention	Decide whether to accept, revise, annotate, or escalate a problematic AI-suggested term or description.	A decision note linking the intervention to user needs, institutional policy, and ethical responsibility.

A third implication is a tighter connection between metadata literacy and ethics. Ethical reflection in the education of AI is sometimes taught at a high level of abstraction. Metadata literacy offers a more practical entry point. Concrete descriptive decisions can be used to teach questions of fairness, transparency and accountability: how a machine-generated abstract is labelled, how a revised subject term is justified, or how uncertainty is conveyed to users. This renders ethical training visible in everyday professional practice [9], [14].

Assessment, therefore, should not be limited to what the student can do with standards or isolated fields. A practical rubric might consist of four criteria: descriptive adequacy, provenance and disclosure, critical evaluation of AI-mediated change, and ethical justification. These criteria can be evaluated by means of records, short reflective comments, comparison memos and group reviews. The goal isn't to make every LIS student a metadata expert. The aim is to enable students to determine how metadata choices affect trust, access, representation, and accountability.

6. Discussion

The new model claims two things. First, metadata literacy is analytically distinct from AI literacy and information literacy. Metadata should not be subordinated as a subskill of either field, for it has its own conceptual history, technical practices, and professional responsibilities. Second, the AI era changes metadata literacy significantly that it needs to be rebuilt. The older focus on fields, standards, and record creation captures only part of what learners need when metadata is implicated in algorithmic mediation, synthetic content, and institutional accountability.

This distinction also points to the originality of the model. Existing AI literacy frameworks are valuable as they specify broad-based capabilities for understanding, using, evaluating and ethically responding to AI [5], [6], [9], [12]. Existing information literacy frameworks have value in that they encourage critical engagement with information and authority [7], [8]. The model proposed here is more specifically focused on the metadata layer where many of those capacities are concretised. It is asking students to look not only at the output of an AI system but also at the descriptive, contextual and provenance structures that make the output visible and reviewable.

The model also has implications for LIS education. The general call for AI preparedness may be too broad if it does not specify the concrete layers through which accountability

and interpretation are carried out. One such layer is metadata. It is where provenance is recorded, where representational categories are stabilised or contested, where disclosure can be made visible, where system-level assumptions become operational. This is how metadata literacy can translate AI ethics from the high-level to the teachable professional practice.

And there are limits to the paper, as it is conceptual. It does not empirically test the proposed model nor claim that metadata literacy is the only way to understand AI-related educational change. Other literacies like information literacy, data literacy and critical digital literacy are still very relevant. Future research could test the model using curriculum design studies, expert validation, student assessment rubrics, or case studies of AI-assisted metadata workflows in libraries and repositories.

7. Conclusion

In the era of AI, metadata literacy can no longer be viewed as merely knowing metadata fields or being able to create records that are compliant with standards. As AI is embedded in discovery, generation, enrichment and interpretation, metadata is increasingly mediating whether outputs are contextualised, reviewable and accountable. In the context of LIS education, this means re-articulating metadata literacy as a core competency for practice in AI-mediated information settings.

The framework proposed here reframes metadata literacy around infrastructural understanding, contextual description and representation, provenance and disclosure, evaluation of algorithmically mediated outputs, and ethical and governance intervention. The extra comparison with adjacent literacies and the sample learning tasks clarify the differences of the framework with existing models and the applicability of the framework in curriculum design. Taken together, these dimensions locate metadata literacy not at the technical periphery of the curriculum but near its conceptual center. In this sense, the reconstitution of metadata literacy is also an educational response to the wider transformation of libraries and information systems in the era of AI.

References

[1] M. S. Mayernik, “Metadata,” *Knowledge Organization*, vol. 47, pp. 696–713, 2020, doi: 10.5771/0943-7444-2020-8-696.

[2] E. T. Mitchell, “Metadata Literacy: An Analysis of Metadata Awareness in College Students,” University of North Carolina at Chapel Hill, 2010. doi: 10.17615/jn1v-yz04.

[3] O. L. Zavalina and M. Burke, “Assessing skill building in metadata instruction: Quality evaluation of Dublin Core metadata records created by graduate students,” *Journal of Education for Library and Information Science*, vol. 62, pp. 423–442, 2021, doi: 10.3138/jelis.62-4-2020-0083.

[4] J.-R. Park, “Metadata quality in digital repositories: A survey of the current state of the art,” *Cataloging & Classification Quarterly*, vol. 47, pp. 213–228, 2009, doi: 10.1080/01639370902737240.

[5] D. Long and B. Magerko, *What is AI literacy? Competencies and design considerations*. in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM, 2020, pp. 1–16. doi: 10.1145/3313831.3376727.

[6] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao, “Conceptualizing AI literacy: An exploratory review,” *Computers and Education: Artificial Intelligence*, vol. 2, 2021, doi: 10.1016/j.caeai.2021.100041.

[7] R. Ali and M. L. Wilson, "Information literacy in the artificial intelligence era: A proposed framework," *Communications in Information Literacy*, vol. 19, pp. 242–270, 2025, doi: 10.15760/comminfolit.2025.19.2.6.

[8] Association of College and Research Libraries, "Framework for Information Literacy for Higher Education," Association of College and Research Libraries, 2016. [Online]. Available: https://www.ala.org/acrl/standards/ilframework

[9] Association of College and Research Libraries, "AI Competencies for Academic Library Workers," Association of College and Research Libraries, 2025. [Online]. Available: https://www.ala.org/acrl/standards/ai

[10] A. M. Cox, "How artificial intelligence might change academic library work: Applying the competencies literature and the theory of the professions," *Journal of the Association for Information Science and Technology*, vol. 74, pp. 367–380, 2023, doi: 10.1002/asi.24635.

[11] S. Hervieux and A. Wheatley, "Building an AI Literacy Framework: Perspectives from Instruction Librarians and Current Information Literacy Tools," 2024. [Online]. Available: https://www.choice360.org/research/white-paper-building-an-ai-literacy-framework-perspectives-from-instruction-librarians-and-current-information-literacy-tools/

[12] M. Montesi, B. Alvarez Bornstein, N. Bautista Puig, M. Blazquez Ochando, and A. Sanchez Diez, "AILIS 1.0: A new framework to measure AI literacy in library and information science (LIS)," *The Journal of Academic Librarianship*, vol. 51, 2025, doi: 10.1016/j.acalib.2025.103118.

[13] J. Shi, M. Nason, M. Tullney, and J. P. Alperin, "Identifying metadata quality issues across cultures," *College & Research Libraries*, vol. 86, pp. 101–134, 2025, doi: 10.5860/crl.86.1.101.

[14] International Federation of Library Associations and Institutions, "Generative AI for Library and Information Professionals," 2024. [Online]. Available: https://www.ifla.org/generative-ai/

[15] Dublin Core Metadata Initiative, "DCMI Metadata Terms," 2020. [Online]. Available: https://www.dublincore.org/specifications/dublin-core/dcmi-terms/