AI-Assisted Metadata and Intelligent Catalogues: Redefining Knowledge Organization in Libraries

Shiva KanaujiaSukula
Deputy Librarian
Dr. B. R. Ambedkar Central Library
JNU, New Delhi
Email: shivasukula25@gmail.com

Abstract

The artificial intelligence and metadata are two significant components in libraries playing crucial roles for knowledge organization and retrieval. The present study is an outcome of literature review, applying scoping review method to observe the current strengths and gaps to design a futuristic approach. The findings reflect that metadata generated by AI shows variable quality, varying from domain-specific training data as well as evolution of intelligent catalogues being predictive as well as adaptive discovery systems.

Keywords: AI-assisted metadata, intelligent cataloguing, AI-assisted metadata generation, metadata enrichment and interlinking, Human–AI Collaboration, interoperability.

AI-Assisted Metadata

Artificial intelligence (AI) is rapidly transforming metadata management by automating the creation, enrichment, and evaluation of descriptive records in a wide range of domains. Platforms such as the AI-assisted archiving metadata creation framework developed by Rim et al. (2025). Rim et al. (2025) demonstrate the potential for integrating AI into archiving workflows to generate consistent and interoperable metadata at scale. Studies have shown that modern AI techniques, including natural language processing and machine learning, can significantly reduce the time and effort required to generate metadata while improving scalability ((Yang, Fu, Amin, & Kang, 2025a; Oyighan et al., 2024). Furthermore, according to professional perspectives, AI integration can improve cataloging efficiency, although human supervision is necessary to maintain contextual accuracy (Chen and Li, 2024; Deng, 2023).

In addition to production, researchers are also investigating the quality and trustworthiness of AI-generated metadata. Domain-specific approaches, such as Bagchi's (2024) generative, AI-driven metadata modeling, reflect promising results in generating rich, schema-compliant records tailored to specific collections. Liu (2025) is notable for advocating human-centered designs that strike a balance between automation and user needs. Zavalin and Zavalina (2025) emphasize the need for robust evaluation methods to determine whether AI outputs meet discoverability and reuse requirements. Bandi (2025) stressed that AI-generated metadata is crucial for improving interoperability and supporting future AI applications. Applications of AI in metadata management are expanding in specific and interdisciplinary contexts, from linked data transformation in photographic archives (Proctor and Marciano, 2021) to

verifying compliance with data protection regulations such as GDPR (Torre et al., 2020). In environmental research, AI-powered metadata can facilitate the publication of datasets in support of sustainability goals (Borja, 2025). These results underscore the importance of linking AI capabilities to quality assurance frameworks to ensure trustworthiness and ethical compliance.

Meanwhile, enterprise metadata management is leveraging AI and machine learning to improve data governance and decision-making (Bhatia, n.d.; Nandini, 2025). Collectively, this body of literature points toward a hybrid future in which AI augments human expertise, enabling faster, richer, and more interoperable metadata ecosystems that can adapt to evolving technological and ethical requirements.

Methodology:

The primary method for this study is a conceptual review of AI-assisted metadata literature set, synthesizing constructs, mechanisms, and propositions across library/archives, enterprise, and science-data contexts as literature observed (list of references).This has also performed scoping review for the literature published related to the AI-assisted metadata creation and intelligent catalogues.The literature was browsed through Google scholar during the months of May- August 2025 (with most updated and relevant data) available on the following keywords and phrases:

1. AI-Assisted Metadata
2. AI-assisted metadata creation
3. Intelligent catalogues

Inclusion and exclusion factors for finalising list of references and data involve the scholarly literature such as journal articles, books and proceedings, the broader areas coverage articles have been excluded keeping in view the above phrases in speculation for literature review.

Conceptual Review of AI-Assisted Metadata (2019–2025)

1) Scope & Definitions

AI-assisted metadata: Use of ML/LLMs + NLP/CV to generate, enrich, normalize, validate, and link descriptive/administrative/technical metadata across the data lifecycle (creation → curation → discovery → reuse).

Human-in-the-loop (HITL): Expert oversight for policy, ethics, and quality control (Liu, 2025; Chen & Li, 2024).

2) Core Conceptual Constructs

A. Automation & Coverage

AI scales description, improves throughput, and reduces backlog (Rim et al., 2025; Proctor & Marciano, 2021; Nandini, 2025; Oyighan et al., 2024).

Generative pipelines/model-driven schemas (Bagchi, 2024; Yang et al., 2025a/2025b).

B. Quality, Reliability & Bias

Mixed performance; domain drift and hallucination remain risks (Zavalin&Zavalina, 2025; Deng, 2023).

Confidence scoring, calibration, and benchmarked evaluations are central (Zavalin&Zavalina, 2025).

C. Standards, Interoperability & "AI-readiness"

Metadata as infrastructure for AI: controlled vocabularies, PIDs, FAIR, KG/linked data (Bandi, 2025; Proctor & Marciano, 2021).

Shape constraints/validation (e.g., SHACL-like ideas), crosswalks, and schema governance (Bandi, 2025; Chen & Li, 2024).

D. Human-Centered Governance

HITL design, role re-scoping, and ethics (explainability, accountability) (Liu, 2025; Chen & Li, 2024).

Organizational readiness and change management (Oyighan et al., 2024; Sandeep, 2025; Bhatia, n.d.).

E. Domain-Specific Contexts

Archives & GLAM: description at scale; provenance and authenticity (Rim et al., 2025; Mardiati et al., 2023).

Enterprise: data catalogs, lineage, policy compliance (Bhatia, n.d.; Torre et al., 2020; Hayes et al., 2019).

Science & Sustainability: publishing datasets with AI-aided metadata to accelerate reuse/impact (Borja, 2025).

F. Method Families

Extraction & Enrichment: NLP/CV for entity/subject extraction; OCR+LLM summarization (Rim et al., 2025).

Normalization & Linking: ontology alignment, authority control, KG population (Proctor & Marciano, 2021).

Validation: rule-based + ML validators, conformance checking (Torre et al., 2020; Hayes et al., 2019).

3) Mechanisms (How AI changes metadata work)

The following steps (M1 -M5) elaborate various mechanisms observed through the literature:

- M1 – Generative drafting → Expert refinement: LLMs propose fields/subjects; catalogers accept/edit with confidence cues (Rim et al., 2025; Chen & Li, 2024).

- M2 – Pattern discovery → Vocabulary alignment: Clustering/embedding maps terms to controlled vocabularies; reduces synonymy/variance (Bandi, 2025; Proctor & Marciano, 2021).
- M3 – Linked data conversion: From free-text to RDF/KG with entity disambiguation (Proctor & Marciano, 2021).
- M4 – Continuous validation: Policy/constraint checks (privacy, completeness) with automated flags (Torre et al., 2020; Hayes et al., 2019).
- M5 – Feedback loops: Human edits retrain models; governance codifies accepted patterns (Liu, 2025; Yang et al., 2025b).

4) Integrative Propositions

There has been deep insights and discussions across the literature which have been formulated as propositions (P1- P8) given as following:

- P1 (Productivity): AI-drafted metadata increases coverage/throughput; net quality depends on HITL density and domain specificity (Rim et al., 2025; Oyighan et al., 2024).
- P2 (Quality): Quality gains occur when AI outputs are constrained by standards (schemas, authorities) and validated against rules; unconstrained generation risks drift/hallucination (Zavalin & Zavalina, 2025; Bandi, 2025).
- P3 (Interoperability): Mapping to controlled vocabularies + persistent identifiers mediates portability across systems (Proctor & Marciano, 2021; Bandi, 2025).
- P4 (Trust & Adoption): Transparent confidence scores, provenance, and audit trails raise professional trust and adoption (Liu, 2025; Chen & Li, 2024).
- P5 (Ethics/Compliance): Automated policy checks (privacy/GDPR/completeness) reduce compliance risk but require up-to-date rulesets (Torre et al., 2020; Hayes et al., 2019).
- P6 (Domain Fit): Domain-tuned models and authority files outperform general LLMs on accuracy/recall in specialized collections (Zavalin&Zavalina, 2025; Rim et al., 2025).
- P7 (Organizational Readiness): Benefits scale with governance maturity (roles, workflows, metrics) and training (Oyighan et al., 2024; Sandeep, 2025).
- P8 (Societal Impact): For mission domains (e.g., ocean sustainability), AI-aided metadata accelerates discovery → policy/research impact (Borja, 2025).

5) Conceptual Model

The textual diagram for the conceptual model is given as following comprising of content, draft metadata, interoperable metadata, validated metadata and approved records.

Content (docs/images/data)
↓ [Extraction & Summarization: NLP/CV/LLMs]
Draft metadata (entities, subjects, summaries)
↓ [Normalization & Linking: vocabularies, PIDs, KG]
Interoperable metadata (RDF/JSON-LD; crosswalked)
↓ [Validation: constraints, privacy/completeness checks]
Validated metadata (+ confidence, provenance)
↓ [HITL Review & Governance: edits, acceptance]
Approved records → discovery systems &catalogs
↺ [Feedback loop retrains models; policies updated]

6) Evidence Map

The evidence map has been associated with themes and their relation with the representative sources form the literature as following:

- Automation & Coverage: Yang et al. (2025a/2025b); Bagchi (2024); Rim et al. (2025); Nandini (2025); Oyighan et al. (2024).
- Quality & Evaluation: Zavalin&Zavalina (2025); Deng (2023).
- Standards & Interoperability/AI-Readiness: Bandi (2025); Proctor & Marciano (2021).
- Human-Centered& Governance: Liu (2025); Chen & Li (2024); Sandeep (2025); Bhatia (n.d.).
- Compliance/Validation: Torre et al. (2020); Hayes et al. (2019).
- Domain Use-Cases: Archives/GLAM—Rim et al. (2025); Mardiati et al. (2023). Science/Ocean—Borja (2025). Enterprise—Bhatia (n.d.).

7) Gaps and Future Agenda

a) Gold-standard benchmarks: Shared, multilingual test sets with ground truth; task-specific metrics (field accuracy, authority match rate, link precision/recall).
b) Confidence & provenance UX: How to present uncertainty and model lineage to practitioners for safe decision-making.
c) Policy-aligned validation at scale: Maintain machine-readable policies (privacy, rights, sensitivity) and auto-update validators.
d) Cost & sustainability: TCO of LLM pipelines (compute, storage, retraining) vs. benefits in backlog reduction.
e) Bias & cultural context: Systematic studies on representation bias in subject assignment for under-described communities/languages.
f) Impact on discovery: Longitudinal A/B tests linking AI-enriched metadata to real user outcomes (findability, time-to-item, reuse).
g) Cross-domain interoperability: Playbooks for moving between library, enterprise, and scientific metadata ecosystems.

Table 1
Thematic sections table: AI-assisted metadata literature

Sl. No.	Theme	Focus Areas	Key References Year(s)
1.	AI Impact on Metadata	Management Influence of AI tools on metadata workflows, automation benefits, and challenges.	Yang et al. (2025, arXiv), Yang et al. (2025, Human-Centric Intelligent Systems), Oyighan et al. (2024), Chen & Li (2024), Sandeep (2025) 2024–2025
2.	AI-Assisted Metadata Generation Platforms	Development of platforms for automatic or semi-automatic metadata creation in archives and libraries.	Rim et al. (2025), Proctor & Marciano (2021), Mardiati et al. (2023) 2021–2025

3.	Evaluation of AI-Generated Metadata	Quality assessment, accuracy, and usability of AI-produced metadata compared to human-created records.	Zavalin&Zavalina (2025), Liu (2025) 2025
4.	Generative AI for Metadata Modelling	Use of generative AI to design or structure metadata schemas and models.	Bagchi (2024), Nandini (2025) 2024–2025
5.	Metadata for AI Readiness	Preparing datasets and resources to be machine-readable and AI-friendly.	Bandi (2025), Bhatia (n.d.), Borja (2025) n.d.–2025
6.	Ethics, Privacy, and Compliance in AI Metadata	Use of AI in metadata compliance with legal frameworks (e.g., GDPR).	Torre et al. (2020), Hayes et al. (2019) 2019–2020
7.	Sector-Specific Applications	Applying AI metadata in domains such as ocean sustainability and enterprise systems.	Borja (2025), Bhatia (n.d.) n.d.–2025

AI-Assisted Metadata Creation

A. Automation and Efficiency in Metadata Generation

A core theme across the literature is AI's transformative role in automating and accelerating metadata creation processes. Systems like the AI-assisted archival metadata generation platform developed by Rim et al. (2025) demonstrate how machine learning can reduce manual labor while improving consistency. Similarly, Magnus et al. (2025) show the effectiveness of AI in large-scale metadata enrichment for cultural heritage collections, while Tsay et al. (2020) illustrate the potential for extracting model metadata directly from AI workflows. These studies provide the potential of automation to increase both speed and coverage in metadata workflows. AI-powered frameworks facilitate linked data transformation (Proctor and Marciano, 2021) and not only highlight but also simplify the creation of web archive metadata using advanced models such as GPT-4o (Nair et al., 2024; 2025).

B. Quality, Accuracy, and Evaluation of AI Metadata

Evaluations in domain-specific contexts, such as agricultural metadata collection applications (Basir et al., 2025) and accessible audio production workflows (Steele, 2023), confirm that quality assurance remains essential for reliable metadata. Research by Oyighan et al. (2024) and Chen and Li (2024) also mention that metadata professionals express concerns about the capabilities of AI to handle nuanced cataloguing standards. Bachi et al. (2022) and Zavalin and Zavalina (2025) warn of the risk of bias inherent in legacy metadata generated by AI and argue for human oversight.

C. Human–AI Collaboration and Professional Perspectives

This theme relates strongly with demands for transparent AI processes and participatory development models in the library and heritage context. Chen and Li (2024) survey professionals who advocate a hybrid approach, while Sussmeier and Henry (2025) talk about ethical issues for bridging the gap between AI developers and the cataloguing community. Liu (2025) emphasises a human-centred design philosophy, ensuring that AI tools are in line with professional practices and user needs. Provenzano et al. (2024) and Allegrezza et al. (2024) also highlight collaboration in improving metadata discoverability in archives and repositories. Several studies emphasise that AI should complement, not replace, human expertise in metadata creation.

D. Domain-Specific Applications and Innovations

Kritika (2024) and Baburin (2025) discuss enterprise contexts, examining the impact of AI on search performance and decision-making. In cultural heritage, Jaillant and Zhao (2025) discuss the use of AI to transform raw data into accessible archival records, while Meesad and Mingkhwan (2024) explore its role in smart digital libraries. AI-powered metadata solutions are increasingly being tailored to sector-specific needs. In sustainability research, Borja (2025) demonstrates the importance and value of AI in managing ocean-related datasets. In the context of agriculture, Basir et al. (2025) introduce Meta Ag for contextual metadata capture. These applications demonstrate the adaptability of AI across sectors with different standards and metadata requirements.

E. Ethical, Policy, and Future Considerations

Integrating AI into metadata generation raises ethical and governance concerns. Together, these studies call for the creation of frameworks that ensure that AI applications respect cultural, legal, and commercial values while developing technological capabilities. Freeman (2025) emphasized upon the commodification of metadata in AI systems, while Colavizzaet al. (2021) highlight the tension between archiving principles and algorithmic processes. Bachiet al. (2022) and Sabol (2025) emphasize the importance of transparency, traceability, and bias mitigation. Nair et al. (2024) stated about the challenges associated with generative AI models, including the risk of technology and explainability issues.

Table 2
Thematic Sections Table: AI-assisted metadata creation literature

Sl. No.	Theme	Key Focus Representative Studies	Propositions (P1–P8)
1.	Automation & Efficiency - AI systems automate metadata generation, reducing manual work and increasing coverage.	Rim et al. (2025); Magnus et al. (2025); Tsay et al. (2020); Proctor & Marciano (2021); Nair et al. (2024; 2025)	P1: AI-assisted systems can significantly increase the speed of metadata creation P2: Automation improves metadata coverage but may require domain adaptation.
2.	Quality, Accuracy & Evaluation - Assessing AI-generated metadata for precision, recall, and bias.	Zavalin&Zavalina (2025); Oyighan et al. (2024); Chen & Li (2024); Bachi et al. (2022); Basir et al. (2025);	P3: Quality assurance mechanisms are essential to maintain metadata reliability.

		Steele (2023)	P4: Bias and inconsistency in AI-generated metadata require active mitigation strategies.
3.	Human–AI Collaboration -Combining AI capabilities with professional judgment for better outcomes.	Chen & Li (2024); Sussmeier& Henry (2025); Liu (2025); Provenzano et al. (2024); Allegrezza et al. (2024)	P5: Human expertise remains essential for validating and contextualizing AI outputs. P6: Effective collaboration models between AI tools and professionals enhance metadata quality.
4.	Domain-Specific Applications - Sector-specific metadata solutions in cultural heritage, sustainability, agriculture, and enterprise.	Jaillant& Zhao (2025); Meesad&Mingkhwan (2024); Borja (2025); Basir et al. (2025); Kritika (2024); Baburin (2025)	P7: AI tools must be adapted to sector-specific metadata standards. P8: Domain-specific AI models improve relevance and usability of generated metadata.
5.	Ethical & Policy Considerations - Governance, transparency, and professional values in AI metadata creation.	Freeman (2025); Colavizza et al. (2021); Bachi et al. (2022); Sabol (2025); Nair et al. (2024)	(Linked to all P1–P8 as a cross-cutting theme)

AI-Assisted Metadata Creation and Intelligent Catalogues

The following section of study has focused on scoping review of collected works on AI-assisted metadata creation and intelligent catalogues in libraries by:

a) Mapping the thematic areas covered.
b) Summarising the scope, methodologies, and insights.
c) Highlighting research gaps and future directions.

Rationale for Conducting a Scoping Review

The intersection of AI-assisted metadata creation and intelligent cataloguing systems represents a rapidly expanding and multidisciplinary research area. Current literature spans library science, archival studies, digital humanities, computer science, and domain-specific applications such as agriculture and ocean sustainability.

Given this breadth and fragmentation, a scoping review is necessary to:

a) Map the Landscape – Identify the range of AI tools, techniques, and workflows applied to metadata creation and catalogue intelligence across domains.
b) Clarify Concepts – Address inconsistent use of key terms such as "metadata enrichment," "AI-driven cataloguing," and "intelligent catalogues."
c) Capture Innovation – Track rapidly evolving AI technologies (e.g., GPT-4o, domain-specific large language models, metadata schema automation).
d) Synthesize Diverse Methods – Integrate findings from heterogeneous methodologies including technical experiments, case studies, and professional surveys.
e) Identify Gaps – Highlight areas lacking empirical evaluation, ethical frameworks, cross-domain standards, or sustainable implementation models.
f) Inform Stakeholders – Provide a structured evidence base for librarians, archivists, developers, AI researchers, and policy-makers.

A scoping review will therefore not only summarise existing knowledge but also guide future research agendas and support informed decision-making in the deployment of AI in metadata and catalogue systems.

Need for a Scoping Review in AI-Assisted Metadata Creation and Intelligent Catalogues

A. Breadth and Diversity of Research

The field spans multiple domains such as archives, libraries, digital repositories, enterprise search, and even ocean sustainability research and making it too broad for a narrow systematic review. A scoping review helps map all available evidence, methods, and use cases, without prematurely excluding emerging work.

B. Rapidly Evolving Technology

AI tools, models, and integration techniques are changing yearly (e.g., GPT-4o in 2024; domain-specific AI models in 2025). Scoping reviews capture this fluid innovation landscape and help track emerging best practices.

C. Variety in Methodologies & Evaluation Metrics

Studies use highly diverse approaches: quantitative accuracy testing, qualitative professional surveys, mixed-methods case studies, and technical proof-of-concept builds. A scoping review synthesises these without imposing rigid comparability constraints.

D. Unclear Terminology & Concept Boundaries

Terms like metadata enrichment, intelligent catalogue, and AI-assisted cataloguing often overlap or are used differently by authors. Scoping reviews clarify definitions and conceptual overlaps to avoid fragmentation in future research.

E. Identification of Research Gaps

Many studies focus on technical feasibility but lack longitudinal evaluation, ethical frameworks, or cross-domain interoperability analysis. Mapping existing work reveals such gaps for targeted future studies.

F. Stakeholder Relevance

Findings are not only academic; they inform librarians, archivists, software developers, AI ethicists, and policy makers. Scoping reviews produce structured overviews accessible to both technical and non-technical audiences.

Discerning the literature through scoping review method

A. Introduction

Artificial Intelligence (AI) is transforming metadata creation and cataloguing practices in libraries, archives, and repositories. The literature shows rapid developments in AI-assisted descriptive metadata generation, enrichment, linked data integration, and intelligent catalogues. Research spans multiple contexts—from archival collections (Rim et al., 2025; Allegrezza et al., 2024) to enterprise search (Baburin, 2025) and cultural heritage metadata (Bachi et al., 2022). The goal is to enhance discoverability, interoperability, and automation, while balancing ethical and quality considerations (Sussmeier& Henry, 2025; Freeman, 2025).

B. Thematic Mapping

Theme 1 – AI-Assisted Metadata Generation

Recent studies explore platforms for automated metadata creation in archives (Rim et al., 2025), evaluation of AI-generated metadata (Zavalin&Zavalina, 2025), and domain-specific enrichment such as intangible heritage (Bachi et al., 2022) or ocean sustainability datasets (Borja, 2025). Generative AI tools like GPT-4o are tested for web archives metadata (Nair et al., 2024, 2025), revealing both efficiency gains and quality control challenges.

Theme 2 – Metadata Enrichment and Interlinking

Linked Data approaches are integrated with AI to create provenance-rich, interoperable metadata (McKenna et al., 2022; Proctor & Marciano, 2021). AI is also used for schema mapping (Neubauer et al., 2025) and enhancing metadata to make datasets "AI-ready" (Bandi, 2025). Sector-specific applications include agricultural metadata apps (Basir et al., 2025) and digital media enrichment (Magnus et al., 2025).

Theme 3 – Human–AI Collaboration in Metadata Workflows

The literature stresses a human-centered approach (Liu, 2025; Chen & Li, 2024), arguing that metadata professionals should remain central for ethical oversight and contextual accuracy. Studies like Oyighan et al. (2024) highlight challenges of skill gaps, training, and trust in AI outputs.

Theme 4 – Intelligent Catalogues in Libraries

Work in this area spans conceptual AI catalogue frameworks (de Jager, 1995; Lancaster & Warner, 2001), intelligent web interfaces for OPACs (Poo et al., 1999), and modern implementations with AI-driven classification systems (Roy et al., 2024; Ogungbenro et al., 2025). Current trends show convergence between intelligent catalogues and AI-assisted metadata pipelines (Sabol, 2025).

C. Methodological Trends

- Experimental evaluations of AI-generated metadata quality (Zavalin&Zavalina, 2025; Baburin, 2025).
- Case studies in institutional repositories and archives (Magnus et al., 2025; Provenzano et al., 2024).
- Framework development for AI-assisted metadata-to-linked-data conversion (Proctor & Marciano, 2021).
- Survey-based studies capturing professional perspectives on AI in cataloguing (Chen & Li, 2024).

D. Key Insights

AI significantly reduces time for metadata creation and enrichment, but human review is essential for contextual accuracy.

- Metadata generated by AI shows variable quality, often dependent on domain-specific training data.
- Ethical concerns include bias propagation, loss of professional control, and black-box decision-making.
- Intelligent catalogues are evolving from search interfaces to predictive, adaptive discovery systems.

E. Gaps and Future Directions

- Lack of longitudinal studies on AI's sustained impact on metadata ecosystems.
- Limited work on cross-domain interoperability of AI-generated metadata.
- Need for ethical frameworks tailored to metadata automation.
- Insufficient exploration of user experience impacts of intelligent catalogues.
- Underrepresentation of Global South contexts in AI metadata adoption studies.

Table 3
Thematic Sections Table: With P1–P8 propositions based on AI-assisted metadata creation and intelligent catalogue literature

Following is the Thematic Sections Table with P1–P8 propositions based on AI-assisted metadata creation and intelligent catalogue literature.

Sl. No.	Theme	Focus	Representative Sources	Propositions (P#)
1.	AI-Assisted Metadata Generation	Development and evaluation of AI platforms for	Rim et al. (2025); Zavalin&Zavalina (2025); Borja	P1: AI can substantially accelerate metadata generation processes while

		automated metadata creation in archives, libraries, and domain-specific repositories.	(2025); Bachi et al. (2022)	reducing human workload. P2: Quality and accuracy of AI-generated metadata vary significantly by domain and require domain-specific training.
2.	Metadata Enrichment & Interlinking	Use of AI for metadata schema mapping, enrichment, and Linked Data integration to enhance discoverability and interoperability.	McKenna et al. (2022); Neubauer et al. (2025); Proctor & Marciano (2021); Magnus et al. (2025)	P3: AI-driven enrichment improves interoperability across repositories. P4: Schema mapping with AI reduces technical barriers but still requires human validation for semantic accuracy.
3.	Human–AI Collaboration	Integration of human oversight in AI-assisted workflows to ensure quality, ethics, and contextual accuracy.	Liu (2025); Chen & Li (2024); Oyighan et al. (2024); Sussmeier& Henry (2025)	P5: Human oversight remains essential to mitigate bias and ensure contextual relevance in AI-generated metadata. P6: Professional acceptance of AI in metadata workflows depends on trust, transparency, and training.
4.	Intelligent Catalogue Systems	AI-enhanced catalogues for adaptive discovery, predictive search, and automated classification.	de Jager (1995); Lancaster & Warner (2001); Roy et al. (2024); Ogungbenro et al. (2025)	P7: Intelligent catalogues can transform information retrieval from static search to dynamic, predictive discovery. P8: Integration of intelligent catalogues with AI metadata pipelines improves both back-end efficiency and user experience.

Conclusion

There is need of developing benchmark datasets for evaluating AI metadata quality across domains with the significance of advance explainable AI techniques for metadata decisions to improve trust. Hybrid governance models that combine human oversight and automated verification are likely to emerge, as

well as expanding research on cross-domain metadata interoperability. Practices related to assessing the legal and ethical risks of large-scale automated metadata creation may be enabled by exploring the integration of AI metadata tools into existing library, archive, and enterprise systems. Future AI depends on the long-term stability of metadata models, including the need for retraining. This encourages investigation of the role of AI in multilingual and cross-cultural metadata creation.

AI can significantly reduce the time and cost of metadata creation without compromising basic descriptive accuracy. Human-AI collaboration produces higher-quality metadata than AI-only or human-only approaches. Generative AI models can improve metadata richness by inferring implicit attributes from context. Domain-specific fine-tuning of AI models yields better metadata precision than generic models. Standardized metadata schemas and controlled vocabularies are essential for scalable AI integration. AI-assisted metadata workflows enhance discoverability and interoperability of digital resources. Quality assurance frameworks are required to systematically evaluate AI-generated metadata. Ethical guidelines and compliance checks (e.g., GDPR) must be embedded in AI metadata systems from design stage.

References:

Allegrezza, S., Caravaca, M. M., Grandi, M., Guercio, M., & La Sorda, B. (2024). The Role of AI in Identifying or Reconstituting Archival Aggregations of Records and Enriching Metadata Schemas. In *Artificial Intelligence and Documentary Heritage* (pp. 43-48). Memory of the World Sub on Education and Research (SCEaR).

Baburin, M. (2025). Effect of metadata on AI chatbot's search results in enterprise search: An experimental study.

Bachi, V., Cisneros, R., Fresa, A., Kupietzky, A., Stan, A., &Truyen, F. (2022). Metadata generation for intangible heritage: the role of crowdsourcing and AI to address bias risk. *ActesJornadesImatgeiRecerca*.

Bagchi, M. (2024). A Generative AI-driven Metadata Modelling Approach. *arXiv preprint arXiv:2501.04008*.

Bandi, P. K. R. (2025). The Role of Metadata in Making Data AI-Ready: Enhancing Data Discoverability and Usability. *Journal of Computer Science and Technology Studies*, *7*(5), 954-963.

Basir, M. S., Zhang, Y., Buckmaster, D., Raturi, A., &Krogmeier, J. V. (2025). Meta Ag: An Automatic Agricultural Contextual Metadata Collection App. *Smart Agricultural Technology*, 101073.

Bhatia, S. A. The Smart Backbone: AI and ML in Enterprise Metadata Management.

Borja, A. (2025). Publishing datasets, using artificial intelligence to help with metadata, can enhance ocean sustainability research and management. *Frontiers in Ocean Sustainability*, *3*, 1522648.

Chen, S., & Li, M. (2024). AI for cataloging and metadata creation: Perspectives and future opportunities from cataloging and metadata professionals. *Technical Services Quarterly*, *41*(4), 317-332.

Colavizza, G., Blanke, T., Jeurgens, C., &Noordegraaf, J. (2021). Archives and AI: An overview of current debates and future perspectives. *ACM Journal on Computing and Cultural Heritage (JOCCH)*, *15*(1), 1-15.

de Jager, K. (1995). The library catalogue-an artificial intelligence perspective. *South African Journal of Library and Information Science*, *53*(2), 83-88.

Deng, S. (2023). AI, cataloging& metadata.

Freeman, S. B. (2025). Ghost in the Machine: Metadata, Commodification, and the Specter of Artificial Intelligence. *Journal of Library Metadata*, 1-17.

Hayes, J. H., Payne, J., &Leppelmeier, M. (2019, September). Toward improved artificial intelligence in requirements engineering: metadata for tracing datasets. In *2019 IEEE 27th International Requirements Engineering Conference Workshops (REW)* (pp. 256-262).

Jaillant, L., & Zhao, L. (2025). Introduction: When data turns into archives: making digital records more accessible with AI. *AI & SOCIETY*, 1-5.

Kritika, M. (2024). The Smart Backbone: AI and ML in Enterprise Metadata.

Lancaster, F. W., & Warner, A. J. (2001). *Intelligent technologies in library and information service applications*. Information Today, Inc..

Liu, Y. H. (2025). Navigating the AI-Driven Metadata Landscape: A Human Centered Approach.

Magnus, B., Priem, M., Vanderperren, N., Berghe, P. V., Keer, E. V., &Vissers, R. (2025). Metadata creation and enrichment using artificial intelligence at meemoo. *Journal of Digital Media Management*, *13*(2), 110-123.

Mahmud, M. R. (2024). AI in automating library cataloging and classification. *Library Hi Tech News*.

Mardiati, W., Aisyah, S., Grataridarga, N., &Wulandari, R. (2023, October). The Potential Use of Artificial Intelligence Technology in the Process of Collecting Metadata in Photo Archive Description Activities. In *The 6th International Conference on Vocational Education Applied Science and Technology (ICVEAST 2023)* (pp. 918-931). Atlantis Press.

McKenna, L., Debruyne, C., & O'Sullivan, D. (2022). Using Linked Data to create provenance-rich metadata interlinks: the design and evaluation of the NAISC-L interlinking framework for libraries, archives and museums. *AI & SOCIETY*, *37*(3), 921-947.

Meesad, P., &Mingkhwan, A. (2024). AI-Powered Smart Digital Libraries. In *Libraries in Transformation: Navigating to AI-Powered Libraries* (pp. 391-428). Cham: Springer Nature Switzerland.

Nair, A., Goh, Z. R., Liu, T., & Huang, A. Y. (2025). Web Archives Metadata Generation with GPT-4o: Challenges and Insights. *Information Technology and Libraries*, *44*(2).

Nandini, N. K. (2025). Revolutionizing Metadata Management through Intelligent Automation.

Neubauer, F., Pleiss, J., &Uekermann, B. (2025). AI-assisted JSON Schema Creation and Mapping. *arXiv preprint arXiv:2508.05192*.

Ogungbenro, O. D., Esse, U. C., Olowoporoku, I., & Christopher, A. (2025). Revolutionizing Library Services: The Impact of Artificial Intelligence on Cataloguing and Access to Information in Nigeria Academic Libraries. *Journal of Library Metadata*, *25*(2), 99-118.

Oyighan, D., Ukubeyinje, E. S., David-West, B. T., &Oladokun, B. D. (2024). The role of AI in transforming metadata management: Insights on challenges, opportunities, and emerging trends. *Asian Journal of Information Science and Technology*, *14*(2), 20-26.

Poo, D. C., Toh, T. K., Khoo, C. S., & Hong, G. (1999, December). Development of an intelligent Web interface to online library catalog databases. In *Proceedings Sixth Asia Pacific Software Engineering Conference (ASPEC'99)(Cat. No. PR00509)* (pp. 64-71).

Proctor, J., & Marciano, R. (2021, December). An AI-assisted framework for rapid conversion of descriptive photo metadata into linked data. In *2021 IEEE International Conference on Big Data (Big Data)* (pp. 2255-2261).

Provenzano, T., Fernandez, R., Deets, C., &Kirmis, D. (2024). Using AI to Facilitate Discoverability and Curation of the ASU Library Repository Collections. In *Proceedings of the International Conference on Dublin Core and Metadata Applications*.

Rim, K., King, O. C., Lynch, K., Verhagen, M., & Pustejovsky, J. (2025, May). A Platform for AI-Assisted Archival Metadata Generation. In *International Conference on Human-Computer Interaction* (pp. 183-203). Cham: Springer Nature Switzerland.

Roy, S., Mallikraj, S. V., Moradia, S., Shantha, G., Aravind, S., & Shivaprakash, S. (2024). The Impact of Artificial Intelligence on Cataloging and Classification Systems in Modern Libraries. *Library of Progress-Library Science, Information Technology & Computer*, *44*(3).

Sabol, D. A. (2025). *Libraries' future: intelligent catalogs, digital repositories, artificial intelligence*.

Sandeep, C. R. (2025). The AI Advantage: Enhancing Metadata for Smarter Decisions.

Steele, K. (2023). *Streamlining production workflows for accessible audio using AI metadata assignment* (Doctoral dissertation, University of York).

Sussmeier, S., & Henry, J. A. (2025). Mind the Gap!: How Do we Ethically Bridge the Divide Between the Cataloging/Metadata Community and the World of AI?. *Journal of Library Metadata*, 1-19.

Titus, T. M. (2024). *Adoption of Artificial Intelligence for Improvement of Metadata Services in Libraries at Institutions of Higher Education in Namibia* (Master's thesis, University of South Africa (South Africa)).

Torre, D., Abualhaija, S., Sabetzadeh, M., Briand, L., Baetens, K., Goes, P., &Forastier, S. (2020, August). An ai-assisted approach for checking the completeness of privacy policies against gdpr. In *2020 IEEE 28th International Requirements Engineering Conference (RE)* (pp. 136-146).

Tsay, J., Braz, A., Hirzel, M., Shinnar, A., &Mummert, T. (2020, June). Aimmx: Artificial intelligence model metadata extractor. In *Proceedings of the 17th international conference on mining software repositories* (pp. 81-92).

Yang, W., Fu, R., Amin, M. B., & Kang, B. (2025). The Impact of Modern AI in Metadata Management. *Human-Centric Intelligent Systems*, 1-28.

Yang, W., Fu, R., Bilal Amin, M., & Kang, B. (2025). Impact and influence of modern ai in metadata management. *arXiv e-prints*, arXiv-2501.

Zavalin, V., &Zavalina, O. L. (2025). Are we there yet? Evaluation of AI-generated metadata for online information resources. *Information Research an international electronic journal*, *30*(iConf), 732-740.