Lala Hajibayova, Kent State University, USA
Seungmin Lee, Chung-Ang, University, Republic of Korea

Artificial Intelligence in Knowledge Representation, Organization and Discovery: Key Competencies and Design Considerations

Abstract
This paper argues that effective use of artificial intelligence technologies in knowledge representation and organization requires a distinct set of skills and competencies to navigate the ethical challenges inherent in these technologies. There are several ethical concerns associated with the application of AI technologies, including the accuracy of AI-generated information, biases embedded in AI training data, privacy and surveillance risks, and potential negative impact of AI on career paths.. Ethical application of artificial intelligence technologies requires thoughtful design of knowledge representation and organization systems to ensure transparency, explainability, contextualization, and critical reflection throughout the technology-driven processes of knowledge representation, organization and discovery. This paper examines the essential skills, competencies, and design considerations required for the ethical application of AI technologies in knowledge representation, organization and discovery.

1. Introduction

Artificial Intelligence (AI) technologies have long been an integral part of knowledge representation and organization systems: From early expert systems to contemporary machine learning models, AI has enabled more scalable, adaptive, and efficient organization of complex information resources.

The concept of Artificial Intelligence has evolved significantly since its early formulations. In their seminal 1955 proposal, John McCarthy and colleagues described AI as the science of 'making a machine behave in ways that would be called intelligent if a human were so behaving' (p. 11), emphasizing the imitation of human cognitive functions. This foundational perspective positioned AI as an attempt to replicate human-like problem-solving and decision-making in machines. Decades later, Nilsson (2009) offered a more function-oriented definition, characterizing AI as the 'activity devoted to making machines intelligent,' where intelligence is defined as 'that quality that enables an entity to function appropriately and with foresight in its environment' (p. 100). This definition broadened the scope of AI beyond human mimicry, highlighting adaptability and goal-oriented behavior within complex environments. Together, these definitions underscore AI's dual focus on human-like reasoning and autonomous decision-making, providing a conceptual foundation for understanding AI's role in contemporary systems, including those involved in knowledge representation, organization, and discovery.

From facilitating information discovery and retrieval to delivering advanced information services, AI technologies aim to provide innovative solutions for seamless access to knowledge (Cox, Pinfield, & Rutte, 2019). Provision of accurate and authoritative access to knowledge requires a deep understanding of implications of AI technologies to ensure their fair and ethical application. Library and information science researchers have investigated the role and impact of AI technologies in such areas as information representation and retrieval (e.g., Golubka 2021; Smith 1976), library and information operations and services (e.g., Cox, Pinfield, & Rutter 2019; Cox 2023),

everyday information interactions (e.g., Brewer, 2024; Hirvonen, Jylhä, Lao, & Larsson, 2024), and scholarly communications (e.g., Hosseini, Rasmissen, and Resnik 2023). For instance, Cox and Mazumdar (2022) have identified how AI can potentially improve the efficiency of library operations and user services such as knowledge discovery and references services (e.g., AI chatbots), as well as promote better understanding of users' behaviors.

The application of AI technologies in knowledge representation, organization, and discovery raises a number of ethical concerns. These include the accuracy and reliability of AI-generated information, the reinforcement of biases embedded in training data, risks to user privacy through surveillance and data misuse, and the broader societal implications, such as the displacement or transformation of professional roles and career paths. As AI becomes more deeply integrated into information systems, addressing these challenges requires more than technical proficiency—it demands a critical understanding of the ethical, social, and epistemological implications of design choices. This paper explores the essential skills and competencies that information professionals must develop to ethically engage with AI technologies, including data literacy, algorithmic accountability, and critical evaluation of automated systems. It also outlines key design considerations—such as transparency, explainability, contextualization, and inclusivity—that should guide the development and implementation of AI-powered knowledge systems. Through this lens, the paper advocates for a design consideration of application of AI in knowledge representation, organization and discovery systems that prioritizes trustworthiness, equity, and responsible innovation in the organization and discovery of knowledge.

2. Ethical Concerns Associated with the Application of AI

Library and information science scholars have long recognized the ethical and moral role of information professionals in securing fair access to the knowledge. One of the key commitments of ethical representation of organization of knowledge is to ensure diverse and multi-perspective representation of knowledge. For example, Beghtol (2002) introduced the theoretical concept of cultural warrant, incorporating a method of user-choice as a foundational approach to establishing and developing culture-neutral systems. Beghtol's intent was to create an intersection between the ethical issues of the globalization of information and its concrete implementation in knowledge organization system to enable individuals to "think globally, to act globally, and, at the same time, to think and act locally and individually" (p. 526). Along these lines, Mai (2013) argued that information professionals must embrace the challenge of making ethical choices to ensure multivocality in representation and organization of cultural objects. Further developing the idea of multi-perspective representation of knowledge, Zhitomirsky-Geffet and Hajibayova (2020) applied conceptualization of ethics of care as a moral approach to understanding values involved in care and standards that reject domination.

The ethical concerns surrounding the application and use of AI technologies are profoundly complex and multifaceted. These concerns extend beyond multifaceted representation of knowledge, encompassing issues such as privacy, accountability, transparency, and the broader societal impacts of AI systems. In a series of explorations,

Floridi and colleagues (Floridi, 2008; Floridi, 2013; Floridi & Cowls, 2019; 2022), aligned the AI ethical principles with the four fundamental principles of bioethics, i.e., beneficence, nonmaleficence, autonomy and justice, as these principles are closely applicable to digital ethics in dealing with new forms, agents, clients and environments. Floridi and Cowls (2022) have further developed these ideas and recommended a unified framework of five principles of AI in society: beneficence, nonmaleficence, autonomy, justice and explicability. In this framework, beneficence highlights the need to enhance the well-being of individuals and the planet and its inhabitants through AI; nonmaleficence, or do-not harm, underscores the importance of safeguarding individuals' privacy and security; autonomy emphasizes the importance of balancing decision-making power between humans and artificial agents; justice requires elimination of bias and discrimination, promotion of diversity, and prevention of emerging threats to fairness; and, explicability, which combines the epistemological aspect of 'intelligibility (understanding how AI works) with the ethical aspect of accountability (identifying responsibility for actions) (Floridi & Cowls, 2022).

In this regard, the International Federation of Library Associations and Institutions (IFLA) outlines several ethical and information concerns associated the with use of AI technologies:

- Inherent biases in training data, including those related to race, gender, geographical location
- Inaccuracies in generated information ("hallucinations") which lack support if valid sources, citations and consistency
- Contribution to the acceleration of content creation further exacerbating individuals' information overload
- Lack of transparency of AI systems due to lack of open information about data sources and system operability
- Privacy infringement due to lack of policies prohibiting unsolicited appropriated of individuals' data
- Copyright infringement due to use of data sources without individuals' copyright permission
- Jeopardizing of employment of content creators such as editors and journalists
- Discrimination against those who cannot afford paid AI services
- Exploitation of workers in the third world countries paid low wages to detoxify content in the development of AI systems
- Environmental concerns due to high consumption of energy sources
- Further empowerment of big tech companies

A key aspect of the fair and ethical application of AI technologies is understanding how these technologies function to ensure they do not cause or infringe upon the individual's right to access accurate and authoritative knowledge. This requires not only the relevant skills and competencies of information professionals to recognize and address ethical challenges in AI applications, but, most importantly, the ethical design of knowledge organization systems. Key considerations include accuracy

and authoritativeness of generated information, transparency and explainability of AI systems, potential biases in AI generated content, and concerns related to copyright and intellectual property rights:

- Authoritativeness and Accuracy

The concept of authoritativeness is central in evaluation and provision of access to credible information. Wilson (1983) defines cognitive authority as the extent to which a second-hand provider of information is regarded as trustworthy. From the perspective of user judgement, Rieh (2002) identifies key dimensions of authoritativeness, including trustworthiness, reliability, scholarliness, credibility, officialness, and authoritativeness. According to Farahat, Nunberg, & Chen (2007), social and textual authority are types of authority that affect information retrieval. Social authority is often measured by relying on social networks, such as citation indexes (Farahat et al., 2007), while textual authority is considered the extent to which information is considered useful, accurate and current (Farahat et al., 2007).
The aforementioned criteria of authoritativeness and accuracy are not applicable to AI-generated content, as it often lacks identifiable authorship, clear provenance, and evaluation processes that underpin traditional notions of authority. The inaccuracy in AI-generated content, coupled with the lack of reliable supporting sources and precise AI detection tools (Bhattacharyya et al., 2023), poses significant challenges to ensuring access to accurate and authoritative resources. Despite publishers' non-AI use policies, book-selling platforms like Amazon are inundated with AI-generated content (Bensinger, 2023). This necessitates extra effort from information professionals to assess the accuracy and authoritativeness of sources.

- Transparency and Explainability

The conceptualization and operationalization of the concept of transparency remain ambiguous as the term used both in a technical sense (for example, interpretable systems are considered transparent) and in aspirational sense (for example, AI transparency is associated with promoting accountability) (Ridley 2025). Technologies are often described as black boxes due to the ambiguity surrounding how they function and operate. Calls for transparency often rest on the implicit belief that making a phenomenon visible creates opportunities and obligations to hold it accountable and thus to change it (Ananny & Crawford, 2016). Ananny and Crawford (2016) suggest that "rather than privileging a type of accountability that needs to look inside systems, that we instead hold systems accountable by looking across them—seeing them as sociotechnical systems that do not contain complexity but enact complexity by connecting to and intertwining with assemblages of humans and non-humans" (p. 974).

Explainability refers to the process of making 'opaque-box' models, which are not inherently understandable and comprehensible through post-hoc methods. Explainability is often characterized by the traceability and interpretability of system behavior (Ridley 2025).Three primary approaches to explainability are identified: (i) data explainability, such as through the use of knowledge graphs or saliency maps; (ii)

model explainability, which focuses on understanding specific models; and (iii) post-hoc explainability, which employs techniques like visualizations, narrative explanations, and rule extraction to clarify system outputs (Ridley, 2025).

AI systems and platforms lack transparency and explainability regarding how they were trained, what data sources were used, and how the solutions or information they provide are generated (Long & Magerko, 2020). The lack of transparency in LLM models should be examined in conjunction with policies and regulatory measures designed to enforce transparency requirements (Liao & Vaughan, 2023). Liao and Vaughan (2023) suggest that "informational transparency" can be achieved through model documentation, publication of evaluation results, provision of explanations, and communication of uncertainty (p. 1). AI-driven relevance and recommendations are common features of mainstream knowledge representation, organization and discovery systems, yet, these systems often lack transparency and explainability. For example, WorldCat links to Goodreads, a book review platform, provide users' perspectives and contribute to a multifaceted representation of the work, however, there is no information on the criteria that applied to select the suggested reviews (Hajibayova, McCorkhill, & Bowman, 2024; Hajibayova, 2019).

- Biases

When considering the collaboration between humans and technology in the design and application of AI, it becomes evident that there is a shared recognition of two dimensions: the technical biases inherent in AI systems and the human reliance on prejudices as a foundation for understanding (Joisten & Renner, 2024). AI development and programming are inherently shaped by human input, just as human comprehension is shaped by pre-existing assumptions (Joisten & Renner, 2024). Consequently, biases in AI systems cannot be eliminated, just as human prejudices that are integral to the process of understanding and cannot be fully removed (Joisten & Renner, 2024). A closer examination of AI applications reveals that the biases, stemmed from the subjectivity of programmers and developers, play a critical role in shaping the system's behavior and outcomes (Joisten & Renner, 2024). Because AI tools are trained on data that are subject to inherent biases, the absence of careful oversight and effective mitigation strategies increases the risk of perpetuating and amplifying these biases, potentially leading to unintended consequences in information services (Nishant, Schneckenberg, & Ravishankar, 2024). This requires careful consideration of use of AI tools, in knowledge representation, organization and discovery systems including use of AI recommended data to enhance representation and findability of resources.

- Copyright and Intellectual Property Rights

AI-generated content often does not comply with existing copyright laws, and it may involve the unauthorized use of copyrighted materials in its training datasets (Hosseini, Ramussen, & Resni, 2023). This non-compliance directly impacts information services, as information professionals must navigate the legal complexities of representation and organization of such content. The lack of clear copyright

guidelines for AI-generated works could lead to legal challenges and ethical concerns regarding intellectual property rights and the ownership of AI-created content. This requires careful consideration of the use of AI tools in information services, such as cataloguing, reference services, and related activities.

3. Consideration of Ethical Design of AI Solutions in Knowledge Representation, Organization and Discovery.

A central pillar of the fair and ethical use of AI technologies lies in understanding how these systems operate, ensuring they do not obstruct or violate individuals' rights to access accurate, reliable, and authoritative information. This calls for more than just technical awareness, which requires a deep ethical sensitivity to the ways AI can shape, distort, or limit knowledge access. Information professionals must be equipped with the necessary skills and critical competencies to identify, evaluate, and respond to the ethical implications embedded in AI-driven systems. However, beyond individual expertise, the foundation for ethical AI implementation rests on the principled design of knowledge organization systems themselves. These systems must be intentionally crafted to uphold values such as transparency, accountability, inclusivity, and intellectual freedom. Through ethically informed design practices, it becomes possible to build AI-enhanced systems that not only function effectively but also reinforce the right of all users to equitable and trustworthy knowledge access.

Ethical considerations are fundamental to the design of knowledge representation, organization, and discovery systems, as these systems function within complex socio-technical contexts that intertwine technological mechanisms with social and cultural values. These systems do not merely structure information—they shape how knowledge is accessed, interpreted, and valued by diverse communities. By focusing on the users' experiences and engaging in cycles of prototyping and feedback, design thinking helps uncover hidden assumptions, address inequities, and create solutions aligned with the core ethical commitments of libraries and information institutions—particularly the promotion of intellectual freedom, equitable access, and user autonomy. This approach encourages designers and information professionals to move beyond purely technical optimization and toward the creation of systems that are socially responsive, inclusive, and ethically grounded. These skills and competencies are directly connected to ethical design considerations in knowledge organization systems, including design solutions that are based on data transparency, explainability, contextualization and reflection:

- Design Principal: Data Transparency and Explainability

Design considerations for transparency and explainability must prioritize the full disclosure of data sources and training methodologies. This involves conducting a rigorous assessment of the provenance, quality, and relevance of the data, as well as clearly articulating the processes by which data are selected, preprocessed, and used in model training. To eliminate ambiguity, detailed documentation should accompany these processes, including metadata standards, data collection protocols, and version histories. Design elements that enhance data transparency and explainability may

include visualizations and explanations of data-driven solutions to help users better understand AI applications (Long & Magerko, 2020). These tools help users understand how data inform system behavior and decisions, thereby fostering trust and enabling meaningful scrutiny of AI applications. For example, visual explanations could be provided to illustrate how relevant user-generated reviews are selected and associated with a retrieved title. Such visualizations might highlight the criteria used for selecting reviews, such as recency, reviewer credibility, thematic relevance, or alignment with user preferences, and show how these factors are weighted in the algorithmic process. By making these selection mechanisms transparent, users can better understand the rationale behind the presented user-generated reviews, assess their relevance, and develop greater trust in the system's recommendations.

- Design Principal: Data Contextualization

The KOS design should aim to empower users to examine how the data sources and methodologies were created and/or selected to understand the limitations of the data (Long & Magerko, 2020). For example, in KOS, user-generated data can be integrated in a way that allows users to filter or select the data most relevant to their needs. This may include options to prioritize content based on criteria such as popularity, recency, topical relevance, or alignment with specific user profiles or search intents. By enabling this level of customization, KOS can enhance user agency, improve the perceived relevance of retrieved information, and support more transparent and meaningful interactions with user-contributed content.

- Design Principal: Data Reflection

KOS design that facilitates access to and examination of data sources and fosters reflective and critical analysis, ensuring the use of high-quality, environmentally sustainable data. This involves implementing transparent data provenance mechanisms that allow users to trace the origin, context, and transformation of data throughout the system. Such design enables users to assess the credibility, accuracy, and relevance of the information they encounter. Moreover, by embedding evaluative tools and metadata indicators, such as data quality metrics, KOS can encourage the selection and use of high-quality, ethically sourced, and environmentally sustainable data. This approach supports responsible information practices and aligns with broader commitments to sustainability and digital ethics in knowledge infrastructure design. This approach encourages users to become critical consumers of AI technologies by questioning their intelligence and trustworthiness (Long & Magerko, 2020).

4. Conclusion

As the boundary between human and AI-generated creations becomes increasingly indistinct, it is imperative that AI applications in knowledge representation and organization are guided by robust ethical principles. These systems must be designed to ensure the integrity, accuracy, and accountability of the knowledge they produce and disseminate. This includes clear attribution of authorship, transparent documentation of AI involvement in content generation, and mechanisms to detect and mitigate bias. Furthermore, ethical design should promote inclusivity, respect for intellectual property,

and the preservation of context and meaning in knowledge structures. By adhering to these principles, AI-driven knowledge systems can support trustworthy and equitable knowledge discovery, while safeguarding against misinformation, distortion, and the erosion of epistemic standards.

This paper argues that the skills and competencies of information professionals must be cultivated through the lens of ethically grounded design in knowledge organization systems. As AI-driven and algorithmic processes increasingly mediate information access, it is essential that professionals are equipped not only with technical expertise but also with a strong ethical foundation. This includes the ability to critically engage with design decisions, assess the implications of data use, and advocate for systems that prioritize transparency, explainability, contextual awareness, and reflective practice. Training should emphasize the development of competencies that enable professionals to evaluate data provenance, interpret algorithmic outputs, recognize embedded biases, and ensure that knowledge structures support inclusivity, accountability, and informed decision-making. By embedding these ethical dimensions into professional development, information professionals will be better prepared to steward trustworthy and equitable knowledge ecosystems.

5. References

Ananny, M., & Crawford, K. 2016. "Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability." *New Media & Society*, 20, no. 3: 973-989. https://doi.org/10.1177/1461444816676645

Beghtol, C. 2002. "A proposed ethical warrant for global knowledge representation and organization systems." *Journal of Documentation*, 58, no. 5, 507–532.

Bensinger, G. (February 21, 2023). ChatGPT launches boom in AI-written e-books on Amazon. *Reuters.* Retrieved May 29, 2025, from https://www.reuters.com/technology/chatgpt-launches-boom-ai-written-e-books-amazon-2023-02-21/

Bhattacharyya, M., Miller, V. M., Bhattacharyya, D., & Miller, L. E. 2023. "High Rates of Fabricated and Inaccurate References in ChatGPT-Generated Medical Content." *Cureus*, 15 no. 5: e39238. https://doi.org/10.7759/cureus.39238

Brewer, R. 2024. 'Understanding voice-based information uncertainty: A case study of health information seeking with voice assistants. "*Journal of the Association for Information Science and Technology*, 75 no. 10, 1041-1057. https://doi.org/10.1002/asi.24854

Cox, A. 2023. "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*, 74 no. 3: 367-380.

Cox, A.M, Pinfield, S., & Rutte, S. 2019. "The intelligent library: Thought leaders' views on the likely impact of artificial intelligence on academic libraries." *Library Hi Tech*, 37 no. 3: 418-4 35. doi:10.1108/LHT-08-2018-0105

Farahat, A., Nunberg, G., & Chen, F. 2002. Augeas: authoritativeness grading, estimation, and sorting. In *Proceedings of the eleventh international conference on Information and knowledge management* (pp. 194-202). Retrieved May 29, 2025, from https://dl.acm.org/doi/abs/10.1145/584792.584827

Floridi, L, & Cowls, J. 2022. “A unified framework of five principles for AI in society.” In S. Carta, *Machine Learning and the City: Applications in Architecture and Urban Design* (pp. 535-545). John Wiley & Sons.

Floridi, L. (2008). Foundations of information ethics. In K. E. Himma & H. T. Tavan *The Handbook of Information and Computer Ethics* (pp. 3-24). John Wiley & Sons

Floridi, L. (2013). *The Ethics of Information*. Oxford: Oxford University Press.

Floridi, L., & Cowls, J. 2019. “A Unified Framework of Five Principles for AI in Society.” *Harvard Data Science Review*, 1 no. 1: 1-14. https://doi.org/10.1162/99608f92.8cd550d1

Golub, K. 2021. “Automated Subject Indexing: An Overview.” *Cataloging & Classification Quarterly* ,59 no. 8: 702 - 719. https://doi.org/10.1080/01639374.2021.2012311

Hajibayova, L. 2019. “Investigation of Goodreads’ reviews: Kakutanied, deceived or simply honest?” *Journal of Documentation*, 75 no. 3: 612-626. https://doi.org/10.1108/JD-07-2018-0104

Hajibayova, L., McCorkhill, M., & Bowman, T. D. 2024. “STEM authorship, user-generated reviews and their impact: legitimate or not?” *Journal of Documentation*, 80 no. 6: 1258-1272. https://doi.org/10.1108/JD-12-2023-0262

Hirvonen, N., Jylhä, V., Lao, Y., & Larsson, S. 2024. “Artificial intelligence in the information ecosystem: Affordances for everyday information *seeking.”Journal of the Association for Information Science and Technology*, 75 no. 10: 1152–1165. https://doi.org/10.1002/asi.24860

Hosseini, M., Rasmussen, L.M., & Resnik, D.B. 2023. “Using AI to write scholarly publications.” *Accountability in Research*, 31 no: 715-723. doi:10.1080/08989621.2023.2168535

International Federation of Library Associations and Institutions (n.d.). Generative AI for library and information professionals. Retrieved May 29, 2025, from https://www.ifla.org/g/ai/generative-ai/

Joisten, K., & Renner, T. 2024. “Technological Biases and Human Prejudices: A Technoethical Challenge," *2024 4th International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Male, Maldives, pp. 01-06, doi: 10.1109/ICECCME62383.2024.10797208

Liao, Q. V., &Vaughan, J. W. 2023. “AI transparency in the age of LLMs: A human-centered research roadmap.” arXiv. Retrieved May 29, 2025, from https://doi.org/10.48550/arXiv.2306.01941

Long, D., & Magerko, B. 2020. “What is AI literacy? Competencies and design considerations.” In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-16). https://dl.acm.org/doi/pdf/10.1145/3313831.3376727

Mai, J.-E. 2013. “Ethics, values and morality in contemporary library classification.” *Knowledge Organization*, 40 no. 4: 242-253.

McCarthy, J., Minsky, M.L., Rochester, N., & Shannon, C.E. 2006. “ A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence” August 31, 1955. *AI Magazine*. 27 no. 4, 12. Retrieved May 29, 2025, from https://doi.org/10.1609/aimag.v27i4.1904

Nilsson, N. 2009. *The quest for artificial intelligence.* Cambridge University Press.

Nishant, R., Schneckenberg, D., & Ravishankar, M. 2024. "The formal rationality of artificial intelligence-based algorithms and the problem of bias." *Journal of Information Technology*, 39 no. 1: 19-40. https://doi.org/10.1177/02683962231176842

Ridley, M. 2025. "Human-centered explainable artificial intelligence: An Annual Review of Information Science and Technology (ARIST) paper." *Journal of the Association for Information Science and Technology*, 76 no. 1: 98-120. https://doi.org/10.1002/asi.24889

Rieh, S. Y. 2002. "Judgment of information quality and cognitive authority in the web." *Journal of the American Society for Information Science and Technology*, 53 no. 2: 145–161.

Wilson, P. 1983. Second-hand knowledge: An inquiry into cognitive authority. Westport, Connecticut: Greenwood Press.

Zhitomirsky-Geffet, M., & Hajibayova, L. 2020. "A new framework for ethical creation and evaluation of multi-perspective knowledge organization systems." *Journal of Documentation*, 76 no. 6: 1459-1471. https://doi.org/10.1108/JD-04-2020-0053