



Keeping Humans in the Loop: Human Intervention, Ethical Governance, and AI Implementation in Academic Libraries

Jane Jiang

Abstract:

The integration of artificial intelligence (AI) into higher education has accelerated across teaching, research support, academic services, and administrative functions, with academic libraries emerging as important sites of implementation. While existing literature often emphasizes efficiency, automation, and scalability, this paper argues that the effectiveness of AI systems in academic environments depends on sustained human oversight and professional judgement. Drawing on examples of AI use in library instruction, chatbot-based reference services, and classroom-based learning activities, this paper demonstrates that AI systems do not independently produce reliable knowledge or services without human evaluation, correction, and interpretation. Across these cases, librarians, instructors, and academic staff play essential roles in maintaining system accuracy, supporting responsible use, and preserving educational integrity. The analysis positions human intervention not as an additional safeguard but as a fundamental component of effective AI implementation in higher education. This human-centered perspective highlights the continuing importance of professional expertise, ethical governance, and human agency in the design and management of AI-supported academic environments.

To cite this article:

Jiang, J. (2026). Keeping Humans in the Loop: Human Intervention, Ethical Governance, and AI Implementation in Academic Libraries. *International Journal of Librarianship*, 11(3), 132-139. <https://doi.org/10.23974/ijol.2026.vol11.3.660>

To submit your article to this journal:

Go to <https://ojs.calaijol.org/index.php/ijol/about/submissions>

Keeping Humans in the Loop: Human Intervention, Ethical Governance, and AI Implementation in Academic Libraries

Jane Jiang

UCNJ Union College of Union County, NJ, United States

ABSTRACT

The integration of artificial intelligence (AI) into higher education has accelerated across teaching, research support, academic services, and administrative functions, with academic libraries emerging as important sites of implementation. While existing literature often emphasizes efficiency, automation, and scalability, this paper argues that the effectiveness of AI systems in academic environments depends on sustained human oversight and professional judgement. Drawing on examples of AI use in library instruction, chatbot-based reference services, and classroom-based learning activities, this paper demonstrates that AI systems do not independently produce reliable knowledge or services without human evaluation, correction, and interpretation. Across these cases, librarians, instructors, and academic staff play essential roles in maintaining system accuracy, supporting responsible use, and preserving educational integrity. The analysis positions human intervention not as an additional safeguard but as a fundamental component of effective AI implementation in higher education. This human-centered perspective highlights the continuing importance of professional expertise, ethical governance, and human agency in the design and management of AI-supported academic environments.

INTRODUCTION

The rapid advancement of artificial intelligence (AI), particularly generative AI and large language models (LLMs), has significantly influenced the higher education sector. AI technologies are increasingly being adopted to support teaching and learning, research, academic writing, student services, enrolment management, and institutional administration, reflecting a broader digital transformation across universities (UNESCO, 2023).

As academic libraries play a central role in supporting research, learning, and scholarly communication, they have naturally become one of the earliest adopters of AI-enabled technologies (Ayinde et al., 2026). Recent studies have documented AI applications in academic library services, including technical services, information retrieval, reference support, collection development, and user education, highlighting the potential of AI to improve operational efficiency and enhance user experiences (Cox, 2024; Ayinde et al., 2026).

Although literature generally reports positive outcomes associated with AI adoption, it is predominantly framed from a technological perspective, emphasizing automation, operational efficiency, and service enhancement. Comparatively less attention has been devoted to understanding the continuing role of librarians in supervising, validating, and governing AI-enabled services. This imbalance suggests that AI implementation is frequently evaluated according to what technology can accomplish rather than how human expertise contributes to its successful and responsible use. Such a perspective is particularly problematic in academic libraries, where professional judgement, ethical decision-making, and instructional responsibilities remain fundamental to library practice.

The concept of Human-Centered Artificial Intelligence (HCAI) provides a useful theoretical lens for addressing this gap. Shneiderman (2022) argues that AI systems should be designed to augment human capabilities rather than replace human decision-making. From this perspective, the effectiveness of AI is determined not by the degree of automation achieved but by the quality of collaboration between humans and intelligent systems. Within academic libraries, this approach aligns closely with long-standing professional values, including intellectual freedom, equitable access to information, privacy protection, and accountability. Rather than diminishing the role of librarians, HCAI reinforces their responsibility for ensuring that AI technologies are implemented in ways that support learning, research, and ethical information practices.

Recent professional frameworks further strengthen this argument by suggesting that successful AI implementation depends as much on institutional readiness as on technological capability. Attain Partners (2026), for example, propose that organizational AI maturity should be evaluated through the integration of governance structures, institutional policies, and human judgement. AI systems are therefore understood as socio-technical systems that require continuous monitoring, evaluation, and ethical oversight throughout their lifecycle. This perspective shifts the focus from technology adoption to responsible implementation, highlighting the importance of human intervention in maintaining transparency, accountability, and public trust.

Empirical research further demonstrates why a human-centered approach is necessary for AI implementation in academic libraries. In their systematic review of twenty-nine studies, Ayinde et al. (2026) found that AI adoption primarily focused on reference services, cataloging, classification, information retrieval, and administrative support. However, their review also revealed persistent challenges, including inaccurate or hallucinated outputs, algorithmic bias, copyright concerns, privacy risks, and the lack of institutional AI policies. These challenges suggest that the difficulties associated with AI adoption are not simply limitations of specific technologies; rather, they reflect broader questions of governance, responsibility, and institutional readiness. Therefore, the value of AI in academic libraries depends not only on the capabilities of the technology itself but also on the human expertise required to evaluate outputs, establish appropriate policies, and ensure responsible implementation.

The importance of human intervention is particularly evident in AI-supported reference and discovery services. Conversational AI systems have demonstrated considerable potential to improve access to library services by responding to routine enquiries and providing around-the-clock assistance (Lappalainen & Narayanan, 2023). However, studies evaluating these systems consistently report limitations when users seek assistance with complex research questions requiring contextual interpretation or disciplinary expertise. Yang and Mason (2023), for example, found that ChatGPT generated inaccurate or incomplete responses to sophisticated library queries, while Dalton (2024) cautioned that the conversational style of generative AI may encourage users to overestimate the reliability of AI-generated information. These findings suggest that AI is most effective when integrated into hybrid service models in which librarians remain actively involved in monitoring AI-supported services, providing research assistance, and addressing questions that require professional judgement (Li & Coates, 2024).

Taken together, the literature demonstrates broad consensus regarding the transformative potential of AI in academic libraries while simultaneously revealing an emerging need to reconsider the role of librarians within AI-enabled environments. Although AI can enhance operational efficiency and improve access to information, its successful implementation depends on continuous human intervention throughout the AI lifecycle. The current body of research therefore supports a shift from technology-centered models of AI adoption towards a human-centered approach in which librarians remain responsible for governance, quality assurance, ethical decision-making, and user education. This shift forms the foundation for the discussion that follows, which examines how human intervention can be embedded within the implementation, governance, and educational use of AI in academic libraries.

HUMAN INTERVENTION THROUGH AI LITERACY AND INFORMATION LITERACY INSTRUCTION

The integration of generative AI into higher education has fundamentally changed the scope of information literacy instruction. Traditionally, academic librarians have focused on helping students identify information needs, locate scholarly resources, evaluate source credibility, and use information ethically. The emergence of AI-generated content, however, has expanded these responsibilities beyond evaluating traditional information sources to critically assess machine-generated knowledge. Consequently, AI literacy has become an essential extension of information literacy, requiring librarians to help students and faculty understand not only how AI tools' function but also when, why, and under what circumstances their outputs should be trusted.

Literature increasingly recognizes that librarians are uniquely positioned to guide this transition. Unlike technology vendors, whose primary focus is often on technological capabilities, librarians approach AI from an educational perspective that emphasizes critical inquiry, evidence-based evaluation, and ethical information use. Rather than encouraging dependence on AI-generated responses, librarians help users develop the skills needed to question, verify, and critically assess AI outputs within the scholarly communication process. This educational role

reflects the broader mission of academic libraries to support independent learners who can navigate increasingly complex information environments.

Several recent studies illustrate how human intervention enhances AI-supported learning rather than competes with it. Johnson et al. (2024), for example, redesigned library instruction by incorporating ChatGPT-generated essays into information literacy sessions. Instead of asking students to produce better prompts or generate more persuasive AI responses, librarians encouraged students to compare AI-generated content with peer-reviewed literature, identify fabricated citations, verify factual claims, and recognize omissions or biases in AI-generated text. AI therefore became an instructional resource rather than an authoritative source of knowledge. The educational value of the exercise was created not by the AI itself but by the librarian's guidance in facilitating critical evaluation and scholarly discussion.

This pedagogical shift reflects a broader transformation in the role of academic librarians. As generative AI becomes increasingly integrated into teaching and research, librarians are expanding their responsibilities from traditional information literacy instruction to guiding responsible AI use. Dalton (2024) argues that academic libraries should not view AI merely as another information retrieval tool but should help users understand its limitations and appropriate applications. This expanded role requires librarians to address emerging challenges related to accuracy, bias, copyright, and the reliability of AI-generated information.

Recent policy frameworks further emphasize that AI literacy should be viewed as an institutional responsibility rather than only an individual skill. UNESCO (2023) highlights the importance of human agency, transparency, and critical thinking in the adoption of AI in education. For academic libraries, this perspective reinforces their emerging role as institutional leaders in responsible AI use, building on their longstanding contributions to information literacy, research support, and digital scholarship.

The literature therefore suggests that the contribution of librarians lies not in teaching students how to use AI more efficiently, but in teaching them how to use AI more responsibly. Human intervention transforms AI from a technology that simply generates information into a pedagogical tool that supports critical thinking, scholarly integrity, and informed decision-making. Rather than replacing information literacy instruction, AI has reinforced its importance by creating new opportunities for librarians to cultivate reflective and ethically informed learners.

HUMAN INTERVENTION THROUGH AI GOVERNANCE AND ETHICAL STEWARDSHIP

Alongside teaching and learning, governance has emerged as another area where human intervention is indispensable. Although generative AI can improve efficiency and automate routine tasks, its responsible implementation requires more than technical solutions. Effective AI governance depends on professional judgement, institutional leadership, and ongoing human oversight to ensure that AI systems align with academic values and organizational responsibilities.

Recent studies further demonstrate that AI governance requires institutional planning rather than technology adoption alone. Ayinde et al. (2026) found that academic libraries face recurring challenges related to the reliability, ethical use, and management of AI systems, highlighting the need for clear policies and ongoing human oversight. Similarly, Huang, Cox, and Cox (2023) examined AI strategies in academic libraries in the United Kingdom and Mainland China and found that institutions in both contexts were moving toward formal approaches to AI governance. In the United Kingdom, discussions emphasized strategic planning and institutional approaches to AI adoption, while studies from China highlighted the influence of national innovation strategies and organizational coordination. & co Across both contexts, the authors found that successful AI implementation required leadership, policy development, and collaboration among library professionals and other institutional stakeholders. These findings reinforce that human governance is not an additional layer applied after AI adoption; it is a central condition for responsible and sustainable implementation.

Ethical frameworks for AI further demonstrate why human judgement remains central to AI governance. Floridi and Cowls (2019) propose principles such as beneficence, non-maleficence, autonomy, justice, and explicability as foundations for responsible AI development and use. However, applying these principles in academic library settings requires human interpretation because decisions about privacy, transparency, fairness, and acceptable risk cannot be determined by algorithms alone. Librarians and institutional leaders must evaluate how AI systems affect users, information access, and scholarly practices while ensuring that technological adoption remains consistent with professional values.

Recent organizational frameworks further reinforce this human-centered perspective. Attain Partners (2026) propose that institutional AI maturity should be evaluated according to the integration of technological infrastructure, governance mechanisms, and human decision-making. Their framework positions human oversight as an essential component of AI implementation rather than an external control process applied after deployment. In practice, this means that librarians should participate throughout the AI lifecycle by evaluating AI vendors, establishing institutional policies, monitoring system performance, auditing AI-generated outputs, and assessing the impact of AI services on users and scholarly communities.

Collectively, the literature demonstrates that governance should not be regarded as a constraint on technological innovation but as the foundation upon which responsible AI adoption is built. Human intervention ensures that AI systems remain transparent, accountable, and aligned with the educational mission. As AI technologies continue to evolve, the professional expertise of librarians will become increasingly important—not because they compete with AI, but because they provide the ethical judgement, contextual understanding, and institutional stewardship that AI itself cannot provide.

CASE STUDY 1: AI CHATBOTS AND HUMAN ESCALATION IN REFERENCE SERVICES

One example that illustrates the importance of human oversight in AI implementation is the use of AI chatbots in academic library reference services. Lappalainen and Narayanan (2023) describe the development of Aisha, an AI-powered chatbot designed to assist users with routine library enquiries, including opening hours, database access, borrowing services, and general navigation of library resources. By automating responses to frequently asked questions, the chatbot improved service availability and enabled users to receive immediate assistance outside normal operating hours.

However, the study also demonstrates the limitations of relying solely on AI for reference services. Questions involving complex research strategies, subject-specific resources, or contextual interpretation frequently exceeded the chatbot's capabilities. Rather than attempting to automate every aspect of reference work, the system was designed to operate within clearly defined boundaries. Queries that required professional judgement or specialized knowledge referred to librarians for further assistance, creating a hybrid service model that combined AI efficiency with human expertise.

Importantly, the researchers emphasized that the chatbot's effectiveness depended on continuous librarian involvement. Librarians maintained the knowledge base, reviewed AI-generated responses, corrected inaccuracies, updated information, and monitored user feedback. These activities helped ensure that the chatbot remained accurate, relevant, and responsive to user needs.

This example demonstrates that the value of AI in higher education does not come from automation alone but from the human practices that guide its implementation. Whether applied in academic library services or classroom learning environments, AI requires evaluation, professional judgement, and institutional oversight. The case reinforces the argument that human intervention is not an additional layer of support but a fundamental component of responsible AI implementation.

CASE STUDY 2: HUMAN-GUIDED AI USE IN CHEMISTRY LAB

The importance of human intervention in AI implementation extends beyond academic library services to teaching and learning environments. An illustrative example from an undergraduate chemistry laboratory course at an American community college illustrates how generative AI can support learning when guided by structured instructor oversight. Rather than restricting the use of ChatGPT, the instructor incorporated it into a laboratory report assignment in which students generated an initial AI-assisted draft and then revised it based on their experimental data and disciplinary knowledge.

Students were required to compare the AI-generated draft with their final reports and explain the revisions they made. The assignment required students to identify scientific inaccuracies, correct oversimplified explanations, strengthen methodological descriptions, and incorporate experimental evidence that had been overlooked or misrepresented by the AI output.

The focus was therefore not on the quality of AI-generated writing itself, but on students' ability to critically evaluate and improve machine-generated content.

This case illustrates that generative AI can function as a learning tool when human judgement remains central to the process. The instructor did not treat ChatGPT as a source of scientific authority; instead, AI served as a starting point for analysis, revision, and reflection. Although this example comes from a chemistry classroom rather than an academic library, it reflects similar principles underlying AI literacy initiatives in libraries, where students are encouraged to verify AI-generated information, evaluate sources, and critically engage with machine-produced content (Johnson et al., 2024). The case reinforces that the educational value of AI depends not on technology alone but on the human guidance and evaluation that shape its use.

CONCLUSION

The growing adoption of artificial intelligence in higher education presents significant opportunities for academic libraries to enhance services, improve operational efficiency, and support teaching and research. However, this review argues that the success of AI implementation should not be measured solely by the level of automation achieved. Instead, literature consistently demonstrates that human intervention remains essential for ensuring the accuracy, reliability, ethical use, and educational value of AI-enabled library services.

Across reference services, information literacy instruction, governance, and classroom applications, librarians play a critical role in validating AI-generated content, guiding students and faculty in the responsible use of AI, and ensuring that AI systems align with the core values of academic librarianship. Rather than replacing professional expertise, AI is most effective when it complements human judgement and supports informed decision-making. This human-centered approach enables libraries to leverage AI while preserving their commitment to intellectual freedom, academic integrity, privacy, and equitable access to information.

As AI technologies continue to evolve, academic libraries should move beyond technology-driven implementation toward Human-Centered AI practices that position librarians as active participants throughout the AI lifecycle—from planning and governance to instruction, evaluation, and continuous improvement. Future research should further explore practical models of human-AI collaboration, including Human-in-the-Loop approaches, to strengthen the relationship between librarians and AI systems. Such approaches ensure that technological innovation supports, rather than replaces, the educational mission of academic libraries while preserving human judgement, accountability, and professional expertise.

References

American Library Association. (2021). *Code of ethics of the American Library Association*.

- Attain Partners. (2026, April 14). *Human-centered AI maturity for academic libraries*. <https://attainpartners.com/blog/human-centered-ai-maturity-for-academic-libraries/>
- Ayinde, L., Ebiefung, R., & Oladokun, B. D. (2026). Adoption of artificial intelligence in academic libraries: A systematic review of current practices, challenges, and research opportunities. *The Journal of Academic Librarianship*, 52(1), Article 103185. <https://doi.org/10.1016/j.acalib.2025.103185>
- Cox, A. M. (2024). Artificial intelligence and the academic library. *Journal of Academic Librarianship*, 50(6), 102965.
- Dalton, C. (2024). Artificial intelligence, language, and humanization in the academic library. *Theological Librarianship*, 17(2), 13–16.
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1).
- Huang, Y., Cox, A. M., & Cox, J. (2023). Artificial intelligence in academic library strategy in the United Kingdom and the Mainland of China. *Journal of Academic Librarianship*, 49(6), 102772.
- Johnson, S., Owens, E., Menendez, H., & Kim, D. (2024). Using ChatGPT-generated essays in library instruction. *Journal of Academic Librarianship*, 50(2), 102863.
- Lappalainen, Y., & Narayanan, N. (2023). Aisha: A custom AI library chatbot using the ChatGPT API. *Journal of Web Librarianship*, 17(3), 37–58.
- Li, L., & Coates, K. G. (2024). Academic library online chat services under the impact of artificial intelligence. *Information Discovery and Delivery*, 53(2), 192–205. <https://doi.org/10.1108/IDD-11-2023-0143>
- Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO.
- Yang, S. Q., & Mason, S. (2023). Beyond the algorithm: Understanding how ChatGPT handles complex library queries. *Internet Reference Services Quarterly*.

About the author

Jane Jiang is the Director of Libraries at UCNJ, Union College of Union County, NJ, USA, with over 25 years of experience in academic and public libraries. She holds master's degrees in Education Technology and Library & Information Science. In addition to her library leadership, Ms. Jiang is a writer, poet, and translator, with several original works and many published translations.