



# Journal of eScience Librarianship

putting the pieces together: theory and practice

## Introduction to the Special Issue: Responsible AI in Libraries and Archives

**Sara Mannheimer**, Montana State University, Bozeman, MT, USA, [sara.mannheimer@montana.edu](mailto:sara.mannheimer@montana.edu)   
**Doralyn Rossmann**, Montana State University, Bozeman, MT, USA   
**Jason Clark**, Montana State University, Bozeman, MT, USA   
**Yasmeen Shorish**, James Madison University, Harrisonburg, VA, USA   
**Natalie Bond**, University of Montana, Missoula, MT, USA   
**Hannah Scates Kettler**, Iowa State University, Ames, IA, USA   
**Bonnie Sheehey**, Montana State University, Bozeman, MT, USA   
**Scott W. H. Young**, Montana State University, Bozeman, MT, USA 

---

### Focus

Librarians and archivists are often early adopters and experimenters with new technologies. Our field is also interested in critically engaging with technology, and we are well-positioned to be leaders in the slow and careful consideration of new technologies. Therefore, as librarians and archivists have begun using artificial intelligence (AI) to enhance library services, we also aim to interrogate the ethical issues that arise while using AI to enhance collection description and discovery and streamline reference services and teaching. The IMLS-funded **Responsible AI in Libraries and Archives** project aims to create resources that will help practitioners make ethical decisions when implementing AI in their work. The case studies in this special issue are one such resource. Seven overarching ethical issues come to light in these case studies—privacy, consent, accuracy, labor considerations, the digital divide, bias, and transparency. This introduction reviews each issue and describes strategies suggested by case study authors to reduce harms and mitigate these issues.

---

**Received:** December 14, 2023 **Accepted:** February 5, 2024 **Published:** March 6, 2024

**Keywords:** Responsible AI, artificial intelligence, privacy, consent, accuracy, labor, digital divide, bias, transparency

**Citation:** Mannheimer, Sara, Doralyn Rossmann, Jason Clark, Yasmeen Shorish, Natalie Bond, Hannah Scates Kettler, Bonnie Sheehey, and Scott W. H. Young. 2024. "Introduction to the Responsible AI Special Issue." *Journal of eScience Librarianship* 13 (1): e860. <https://doi.org/10.7191/jeslib.860>.

The *Journal of eScience Librarianship* is a peer-reviewed open access journal. © 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0>.

∞ OPEN ACCESS

## Introduction

Librarians and archivists are often early adopters and experimenters with new technologies. Our field is also interested in critically engaging with technology, and we are well-positioned to be leaders in the slow and careful consideration of new technologies. Therefore, as librarians and archivists begin using artificial intelligence (AI) to enhance library services, we also aim to interrogate the ethical issues that arise. The IMLS-funded [Responsible AI in Libraries and Archives](#) project aims to create resources that will help practitioners make ethical decisions when implementing AI in their work. The case studies in this special issue are one such resource.

The eight responsible AI case studies included here show the variety of ways in which librarians and archivists are currently using AI in their practice, with a special focus on the ethical issues and considerations that arise over the course of implementing AI tools and systems. The case studies include examples of using recommender systems—both library built systems (“[Open Science Recommendation Systems for Academic Libraries](#)”) and vendor systems (“[The Implementation of Keenious at Carnegie Mellon University](#)”), an experiment with ChatGPT (“[Ethical Considerations in Integrating AI in Research Consultations: Assessing the Possibilities and Limits of GPT-based Chatbots](#)”), a computational investigation of the Human Genome Project archives (“[Ethical considerations in utilizing artificial intelligence for analyzing the NHGRI’s History of Genomics and Human Genome Project archives](#)”), sentiment analysis of news articles about the Beatles (“[‘I’ve Got a Feeling’: Performing Sentiment Analysis on Critical Moments in Beatles History](#)”), using natural language processing to generate richer description for historical costuming artifacts (“[Automatic Expansion of Metadata Standards for Historic Costume Collections](#)”), using automated speech recognition and computer vision to create transcripts and metadata for a television news archive (“[Responsible AI at the Vanderbilt Television News Archive: A Case Study](#)”), and partnering with an AI company to extract metadata from historical images (“[Using AI/Machine Learning to Extract Data from Japanese American Confinement Records](#)”).

Seven overarching ethical issues come to light in these case studies—privacy, consent, accuracy, labor considerations, the digital divide, bias, and transparency. We review these issues further below, including strategies suggested by case study authors to reduce harms and mitigate these issues.

## Privacy

Most of the case studies in this issue consider privacy in their AI project implementation. Beltran, Griego, and Herckis, in their discussion of a library-built open science recommendation system, “[Open Science Recommendation Systems for Academic Libraries](#),” suggest that ongoing development of rules, policies, and norms can support privacy. For vendor tools, Pastva et al. describe working with a vendor to ensure that the vendor’s privacy policy and terms of service aligned with library and archives values and practices in “[The Implementation of Keenious at Carnegie Mellon University](#).” Other case studies discuss how digitization and larger-scale availability of archival records can lead to complexities related to privacy. Wolff, Mainzer,

and Drummond's case study, "[I've Got a Feeling': Performing Sentiment Analysis on Critical Moments in Beatles History](#)," analyzes historical news articles about the Beatles—high profile celebrities. By focusing on public figures, Wolff et al. reduce privacy concerns for the project at hand, but they suggest that privacy is still relevant, writing, "were the same analysis methods applied to non-public figures, privacy considerations such as the "right to be forgotten," or excluded from computational analysis of available text data, would be required." For case study authors working with more sensitive records, data security and restricted access are key considerations. Elings, Friedman, and Singh, whose case study "[Using AI/Machine Learning to Extract Data from Japanese American Confinement Records](#)" focuses on extracting metadata from images in Japanese internment records, describe building, testing, and implementing a sustainable model for integrating community input from stakeholders and people represented in the collection. Elings et al. also discuss implementing access restrictions for the data. Hosseini et al., whose case study works with genomic records, "[Ethical considerations in utilizing artificial intelligence for analyzing the NHGRI's History of Genomics and Human Genome Project archives](#)," describe reducing the number of records made available in order to "mitigate risks, ensure ethical compliance, and maintain data privacy standards while enabling valuable research outcomes." Such tradeoffs factor into responsible implementation of AI tools and projects.

## Consent

Use of data without explicit consent is of concern to the authors in this special issue. Additional challenges arise when the source data was gathered prior to the existence of AI tools. Hossieni et al. experience this ethical tension in their biometrics study of a national archive in "[Ethical considerations in utilizing artificial intelligence for analyzing the NHGRI's History of Genomics and Human Genome Project archives](#)." They observe that, even if all data used is fully de-identified and there is minimal risk of harm, they are analyzing user data without explicit consent. This approach could be viewed as undermining subjects' autonomy and a harm in and of itself. They mitigate this tension by handling the data as if the participants were being asked to participate in this new reality by conducting de-identification of the information used and rendering it to encoded data. Likewise, Wolff et al. in "[I've Got a Feeling': Performing Sentiment Analysis on Critical Moments in Beatles History](#)," acknowledge an ethical dilemma in using the work of journalists contained in a dataset with historical news articles about the Beatles. Specifically, it is unclear if this analysis falls under fair use when it is a part of a larger dataset. These questions about consent can help library and archives practitioners consider questions that may arise, despite little precedent in some of these arenas.

## Accuracy

Several case studies highlight the ethical challenges created by the accuracy of the data. Accuracy can be influenced by AI systems themselves (such as sentiment analysis tools) or can be influenced by elements of AI systems (such as OCR and named entity recognition). Wolff et al. observe that the varying accuracy of OCR can have a ripple effect in "[I've Got a Feeling': Performing Sentiment Analysis on Critical Moments in Beatles History](#)." If the source data is inaccurate, it can cause further misinterpretation by the subsequent

sentiment analysis tools used. Also, these sentiment analysis tools have been trained on a specific form of writing found in social media and are not optimized to historical writing. Anderson and Duran describe challenges brought about when named entity recognition misattributes information to individuals in “[Responsible AI at the Vanderbilt Television News Archive: A Case Study](#).” Some of these concerns can be mitigated by a human review of the OCR and named entity recognition created before sentiment analysis tools are used. Feng et al., suggest that human intervention can come into play in consideration of the quality and depth of information presented by various AI chatbots in “[Ethical Considerations in Integrating AI in Research Consultations: Assessing the Possibilities and Limits of GPT-based Chatbots](#).” Librarians and others can provide guidance of what strengths and weaknesses might be found in the different bots so researchers are more aware of potential variation in results depending on the tool used.

## Labor

Several case studies in this special issue discuss ethical issues relating to labor—both for library and archives employees and for student workers. Pastva et al. discuss how AI could negatively impact library liaison services by reducing the amount of human interaction between library employees and library users in “[The Implementation of Keenious at Carnegie Mellon University](#).” McIrvin et al. (“[Automatic Expansion of Metadata Standards for Historic Costume Collections](#)”) and Anderson & Duran (“[Responsible AI at the Vanderbilt Television News Archive: A Case Study](#)”), whose case studies focus on AI for metadata enhancement, are concerned with how AI could affect the jobs of library employees with metadata and cataloging expertise. All of these authors suggest that AI should be used to augment, rather than replace library services. Beyond displacement of workers and expertise, fair labor practices were also considered. Beltran et al. touch on the ethics of student labor. In their case study, “[Open Science Recommendation Systems for Academic Libraries](#),” they describe offering course credit to student workers in lieu of wages. To address this potential ethical challenge, Beltran and colleagues worked with unpaid students to ensure that these students’ goals were being met—co-creating learning outcomes, and drafting a collaboration agreement between the students and the library.

## Digital divide

The cost of accessing AI tools can lead to a digital divide, as Feng, Wang, and Anderson discuss in their case study “[Ethical Considerations in Integrating AI in Research Consultations: Assessing the Possibilities and Limits of GPT-based Chatbots](#).” As new AI-powered vendor tools are released, some may include free versions. But higher-quality, more accurate results are available to paid subscribers. This leads to a divide between people who can access high quality information and those who cannot. In 2012, Boyd and Crawford wrote about a divide between “the Big Data rich and the Big Data poor,” (2012, 674) and this divide continues to be a concern as new technologies are developed and turned into commercial products. Something to watch for here are the subscription models for AI that are now coming into the market. In “[The Implementation of Keenious at Carnegie Mellon University](#),” Pastva et al. briefly touch on the pay

model for Keenius, a vendor-provided resource recommender system. This seems likely to be an area where the digital divide will appear within libraries as communities with fiscal resources will be able to pay for personal recommendation bot assistants for their patrons while other communities will not have access to these technologies.

## Bias

Several of the case studies in this special issue discuss potentially biased results and strategies to help reduce that bias. In “[Automatic Expansion of Metadata Standards for Historic Costume Collections](#),” McIrvin et al. suggest that subject heading biases may be present in the controlled vocabularies used by their AI model. To reduce potential bias, the team enlisted a domain expert to review automatically-generated metadata, and engaged with diverse metadata sources to avoid misleading or culturally-insensitive terms. Pastva et al. (“[The Implementation of Keenious at Carnegie Mellon University](#)”) and Beltran et al. (“[Open Science Recommendation Systems for Academic Libraries](#)”) discuss how recommender systems may be biased. Pastva et al. point out that recommender systems and relevance ratings may show biases toward certain funding sources, legal jurisdictions, or countries of origin. They also note that students new to the research process might not be able to recognize bias when using new tools. Beltran et al. interrogate the implications of persuasive design writ large, asking, “how can we build a model that does not invite the undue or unwanted influence of library services or introduce bias but ultimately is helpful and protects the users’ autonomy?” Beltran et al. offer that one potential answer to this question is to enhance transparency—to design a recommender tool that explains why a recommendation was made. Other researchers suggest that bias can also be found in the data and the training models used. Wolff et al. worked on sentiment analysis of historical news about the Beatles. In “[‘I’ve Got a Feeling’: Performing Sentiment Analysis on Critical Moments in Beatles History](#),” they tested several models analyzing OCR text from historical newspaper reports. The group found the models adequate, but noted bias (and some false classifications) as the training data was from a current social media corpus that didn’t match historical uses of language in the newspapers. Of note here: the need to use training data that matches the data to be analyzed. Recognizing and anticipating bias in training data, and consequently, setting up custom training was proposed as a way to avoid bias. McIrvin et al. followed this solution in “[Automatic Expansion of Metadata Standards for Historic Costume Collections](#)” by incorporating metadata terms using an inclusive description model to help remove bias in the generated subject terms.

## Transparency

A number of case study authors refer to transparency and explainability as core requirements for AI systems. In “[Open Science Recommendation Systems for Academic Libraries](#),” Beltran et al. connect the tension of proprietary recommendation systems and the lack of explainability for the recommendations as a cause of bias and distrust in the system. Others (see “[The Implementation of Keenious at Carnegie Mellon University](#)”) go further in reviewing sources of the AI models and trying to draw out “algorithmic transparency” for

patrons in their work on recommendation systems. Guides that explain the decisions (algorithms) and data sources are suggested as a path towards building trust in the system. Others define their moves toward responsibility and accountability as directives that indirectly create transparency. Hosseini et al. see the responsible release of open-source code and explanations of how to use the code as part of this continuum of transparency in “[Ethical considerations in utilizing artificial intelligence for analyzing the NHGRI’s History of Genomics and Human Genome Project archives.](#)”

## Conclusion

The goal of this special issue is to provide examples of how practitioners can ethically and responsibly engage with AI tools and systems. The ethical issues raised in these case studies show that even as AI tools grow and change, our common professional values and ethical concerns as library and archives practitioners remain the same. We hope that when other practitioners read these case studies, they will be able to translate the ethical considerations and harm-reduction strategies in the case studies to their own work with AI.

### —*Special Issue Guest Editors*

Sara Mannheimer, Data Librarian, *Montana State University*

Doralyn Rossmann, Dean of the Library, *Montana State University*

Jason Clark, Head of the Research Optimization, Analytics, and Data Services, *Montana State University*

Yasmeen Shorish, Director of Scholarly Communications Strategies, *James Madison University*

Natalie Bond, Government Information Librarian and Head of Information & User Services, *University of Montana*

Hannah Scates Kettler, Associate University Librarian for Academic Services, *Iowa State University*

Bonnie Sheehy, Associate Director, Center for Science, Technology, Ethics, and Society and Assistant Professor, History & Philosophy, *Montana State University*

Scott W. H. Young, User Experience and Assessment Librarian, *Montana State University*

---

## Acknowledgements

We would like to thank the [Responsible AI](#) project advisory board: Dorothy Berry, Stephanie Russo Carroll, María Matienzo, Bohyun Kim, and Thomas Padilla. This project is made possible in part by the Institute of Museum and Library Services, through grant number [LG-252307-OLS-22](#).

## Competing Interests

The authors declare that they have no competing interests.

---

## References

boyd, danah, and Kate Crawford. 2012. “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.” *Information, Communication & Society* 15 (5): 662–679. <https://doi.org/10.1080/1369118X.2012.678878>.