

Ethical AI Tools for Critical Museum Practice

Erin Canning

University of Oxford
erin.canning@eng.ox.ac.uk

Abstract

This research examines how the practice of critical cataloguing—the power-aware interrogation of cultural heritage object records—can be supported through the development of AI tools which match the ethical framework of the intended users. I apply a feminist ethics of care to analyze the critical cataloguing domain and articulate technical requirements for the development of computational supports in this space. I test my hypotheses through technological development and research embedded in the Victoria and Albert Museum.

Introduction

The practice of *critical cataloguing* refers to work taking place in museums to understand and address the presence of offensive, biased, and outdated language in catalogue records (Canning 2026). This work is driven by concepts from bodies of critical theory such as feminist, queer, critical race, and de- and post-colonial theory (Caswell, Punzalan, and Sangwand 2017; Drabinski 2019; Watson 2020). Museums are starting to look to computational methods to support critical cataloguing work and while there is potential for this field to do so, ethical and theory-based approaches are essential for working towards justice-oriented goals in this space (Jaillant et al. 2025; Long et al. 2025). After all, critical cataloguing efforts are not about changing data for the sake of doing so, but about examining and challenging oppressive power structures and working to create just and equitable environments (Frick and Proffitt 2022). The application of computational methods in this space requires a robust understanding of the critical cataloguing domain as one with needs related to the broader goals of justice-oriented work in addition to the challenges of working with diverse and complex datasets. The question therefore is how to address the practical challenges of critical cataloguing work in ways that forefront the ethical and conceptual concerns of practitioners, in order to fit the needs of intended users.

Research Questions

This study explores the high-level question of if and how computational methods can support the goals of, and work

involved in, critical cataloguing practices in museums, taking the Victoria and Albert Museum (V&A) as the application context. In pursuit of knowledge towards this goal, this research aims to answer the following questions:

1. What is a methodology for the development of computational approaches that is informed by the theoretical foundations and concerns of critical cataloguing?
2. Can computational models be trained to model domain expert judgements regarding the presence of problematic terminology in catalogue records?
3. Can a technical system which enables critical engagement with catalogue data be developed by leveraging domain expertise regarding problematic terminology in catalogue records?

Project Description

In order to answer these questions, I developed a mixed-methods research study in which qualitative and theory-driven analysis of practitioners is used to determine engineering approach, design, and implementation decisions. Guided by an analysis of the literature, and corresponding to my research questions, my thesis covers four areas of work:

1. an interview-based study of critical cataloguing practices and practitioners;
2. the creation of an ontology and knowledge graph (KG) of critical cataloguing guidelines, workflows, and decisions;
3. the development of artificial intelligence (AI) models for major pain points of critical cataloguing work; and
4. the construction and user evaluation of a user-facing tool which uses the KG and AI models to support staff at the V&A in their critical cataloguing work.

At this stage of research, the main project work has been completed. I have conducted a series of interviews with 24 individuals who identify as “critical cataloguers” or as engaging in critical cataloguing work. The analysis of this work has resulted in an articulation of critical cataloguing as a practice in which an ethics of care is applied in a museum setting, and an articulation of domain technical requirements for working in the critical cataloguing space which reflect these ethical foundations. I have also analyzed available documentation on critical cataloguing practices and produced a formal ontology for the representation of museum-based

critical cataloguing guidelines and decision-making. I have populated a KG with the contents of three cultural heritage terminology guidance documents, three sources documenting critical cataloguing decisions and actions, and a related linked open data project vocabulary. I have worked with the V&A to create a dataset of museum records annotated for critical cataloguing tasks. I have compared different AI models using this dataset and evaluated them on a set of metrics which treat target uncertainty as signal, not noise; reward probability mass placed on plausible alternatives; and penalize overconfidence in ambiguous cases—following perspectivist data approach (Basile et al. 2023). Lastly, I have made a web-based tool for critical cataloguers at the V&A which draws on the KG and AI models to use to assist their practice. The effectiveness of this tool in supporting critical cataloguing work and improve user efficiency, confidence, and critical engagement was evaluated through user testing with V&A staff.

Contributions and Significance

Expected Contributions

This research is the first study of critical cataloguing in museums, as well as the first research on developing technical tools for practitioners working in this emerging area of practice. It proposes a novel methodology for integrating the theoretical concerns of the user community into the technical development process and presents a set of domain technical requirements. In doing so, it demonstrates how concepts from critical theory—in particular an ethics of care—can inform methodology for data engineering projects situated in the context of critically-informed work practices. Furthermore, it articulates a formal, machine-readable characterization of the critical cataloguing knowledge domain, drawing on and extending best practices for the semantic representation of cultural heritage and linguistic data. It also demonstrates how AI models can be trained using an annotated dataset of museum records and additional critical cataloguing documentation resources to model the judgement patterns of practitioners embedded the V&A context. Lastly, it evidences how a user-facing technical system can encourage critical engagement with museum data by bringing critical cataloguing guidance and expertise to users during their review of catalogue records. In sum, this research is positioned to makes several contributions to knowledge in the areas of ethical AI development and data engineering in the context of cultural heritage organizations.

Alignment with AIES

My contribution AIES will focus on the core ethical dimensions of this work: the connection between the critical ethics frameworks of cataloguers, the design of computational techniques, and the development of computational tools through which users interface with the underlying technologies. This research is grounded in real-world technology development and a distinct institutional setting. This work fits into three of the AIES conference theme categories—AI Systems, Values, and Ethical Design; AI in Social and Institutional Domains; and Critique, Consequences, and Limits

of AI—with the strongest connection being to AI Systems, Values, and Ethical Design. Although this project looks at cultural heritage institutions as the context for this work, and thus these proposed interventions, the core question is not bound to this domain: how to incorporate ethical design into the development of AI systems from the point of conception in ways that align with the ethical perspectives of the system’s intended users.

Acknowledgments

This research is taking place as part of the United Kingdom Research and Innovation (UKRI) Arts and Humanities Research Council (AHRC) funded Collaborative Doctoral Partnership scheme grant AH/X004775/1. I am grateful for the support of my supervisors, the involvement of all of the interview participants, and the participation of members of the Victoria and Albert Museum staff.

References

- Basile, V.; Cabitza, F.; Campagner, A.; and Fell, M. 2023. Toward a Perspectivist Turn in Ground Truthing for Predictive Computing. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(6): 6860–6868.
- Canning, E. 2026. Structuring Critical Cataloguing in Museums: A Taxonomy of Museum-Based Critical Cataloguing Project and Action Types. *Museum and Society*, 24(1): 17–37.
- Caswell, M.; Punzalan, R.; and Sangwand, T.-K. 2017. Critical Archival Studies: An Introduction. *Journal of Critical Library and Information Studies*, 1(2): Article 2.
- Drabinski, E. 2019. What is critical about critical librarianship? *Art Libraries Journal*, 44(2): 49–57.
- Frick, R.; and Proffitt, M. 2022. Reimagine Descriptive Workflows: A Community-informed Agenda for Reparative and Inclusive Descriptive Practice. Technical report, OCLC Research.
- Jaillant, L.; Mitchell, O.; Ewoh-Opu, E.; and Hidalgo Urbaneja, M. 2025. How can we improve the diversity of archival collections with AI? Opportunities, risks, and solutions. *AI & SOCIETY*, 40: 4447–4459.
- Long, K.; Itchuaqiyaq, C. U.; Lindgren, C. A.; Ogier, A.; Boyd, L. A.; Itchuaqiyaq, D. P.; and Yunes, E. 2025. Reimagining Archives in the Age of Automation: A Decolonial and Relational Approach. *Technical Communication Quarterly*, 34(3): 336–352.
- Watson, B. M. 2020. “There was Sex but no Sexuality*.” Critical Cataloging and the Classification of Asexuality in LCSH. *Cataloging & Classification Quarterly*, 58(6): 547–565.