

## Book Review

---

**Provo, A., Burlingame, K., & Watson, B.W., (Eds.). (2023). *Ethics in Linked Data*. Library Juice Press.**

*Ethics in Linked Data* is the first in the Series on Critical Information Organization in LIS from Litwin Books and Library Juice Press. The edited volume builds from the work of the LD4 Ethics in Linked Data Affinity Group and includes contributions from a wide variety of GLAM perspectives, personal and professional identities, and areas of practice. The volume comes at a critical moment as the GLAM fields are exploring ways to enhance the linked data stores that drive AI and wrestle with the historical harm or erasure of communities through reparative cataloging practices. Heavily influenced by Jane Sandberg’s (2019) *Ethical Questions in Name Authority Control*, also published by Litwin, the volume is centered on developing ethical frameworks for information organization and description by closely examining the values and biases embedded in the tools and standards used across the GLAM fields.

The introduction acknowledges that this text is written for an audience with some familiarity with the principles of linked data. The first few pages offer an array of resources for those who may need a primer in the technology that include approachable journal articles and texts from across the GLAM disciplines, including *Linked Data for the Perplexed Librarian* (2020). Though not overly technical, familiarity with standards such as RDF, common ontologies, and projects like BIBFRAME and DBpedia should provide sufficient context to engage fully with the text.

Much like the early days of the Internet and the development of HTML, Berners Lee’s concept of linked open data, driving the semantic web, is steeped in the utopian ideals of an open, accessible, and ungoverned technology. Linked data, like other markup languages, is written in plain text, and seeks to connect vast stores of data across the open web to freely share ideas and knowledge; however, this volume sets aside this idealistic view, “examining the darker implications or harmful consequences” of linked data (p. 5). The authors examine the values embedded in the design of the standards and ontologies used, the commercial influence on the development of linked data technologies, and call into question the ethics of representation with a specific focus on those communities historically excluded from participation in this space. However, while examining this “darker side,” the authors raise thoughtful questions about how to ethically use linked data by highlighting projects and case studies that demonstrate the promise of these open technologies for the GLAM fields. The volume also offers a critical lens for examining the full suite of tools used in GLAM practice, especially those used for cataloging and discovery. Throughout each chapter, the contributors seek to “foreground ethics” from project inception, “acknowledge and mitigate the damage caused by existing systems,” “create a space for justice,” and “enable more ethical outcomes for linked projects” (pp. 5–6).

Divided into four sections, the volume first examines the history and development of linked data tools and standards before introducing a series of case studies that demonstrate the application of ethical principles in the design and development of linked data projects and closes with the “Ethics In Linked Data Checklist” as a tool for developing, sustaining, and preserving such projects. The co-authored pieces throughout the volume highlight the collaborative nature of linked data projects and the importance of approaching knowledge organization through multiple perspectives.

Part I is a historic and philosophical examination of linked data standards and technology. This section grounds the chapters that follow in building a deeper understanding of the layers of technology and technical infrastructure necessary to support linked data stores. The chapters highlight the interdependence of linked data projects and the fragility of a web of linked data built on project-based work with limited lifespans and infrastructure. This section emphasizes the need for a reflexive examination of linked data technologies, examining the people and resources behind these tools and the values that have been embedded in the many layers of technology that these projects rely on. While focusing on linked data, these contributions aptly demonstrate methods for debunking the myth of the neutrality of technology by examining the historical context under which these tools have been created and used.

Foregrounding the next two sections of the volume, academic librarian Sam Popowich addresses one of the central challenges to linked data and AI – distilling the complexity of human knowledge into simple data expressions that then take on a fixed form in data stores. The author notes that many of the ethical challenges surrounding AI and the semantic web are because, “the messy, biased, or contested meanings in the social world of humans are made concrete and fixed or stabilized . . . within AI systems” (p. 113). All information organization within LIS fixes the way in which information resources are understood and through standardization, these constructed meanings and language perpetuate through time. The remaining chapters in the volume reflect on these issues and the way in which the dominant culture, western-settler-colonial ways of knowing, have been centered throughout descriptive practice.

Parts II and III move from theory to practice highlighting case studies from across GLAMs. The contributors in Part II highlight standards and practices, exploring ontologies, controlled vocabularies, and authority records for various identities and communities including: the “data legacies” embedded in controlled vocabularies and authority data, considerations of cultural protocols and the values of open data, creation of authority data outside of GLAM authorities, as well as identity-based vocabularies. Part III continues the examination of the ethics of knowledge representation, building on the previous case studies to highlight the wealth of different resources that can be used to contribute to linked data projects beyond the bibliographic data produced by libraries and archives. Oral history projects dominate this section, illustrating the politics of authorship and attribution for data, self-representation and data sovereignty, collaboration with stakeholders, and building authority records for and with living contributors.

The final contribution to the volume is the “Ethics In Linked Data Checklist” developed in consultation with the LD4 Affinity Groups and attendees at LD4 events and conferences. The checklist is divided into three sections that cross the various stages of development and implementation of linked data projects. Each section asks critical questions regarding the ethics and values embedded in the project intended to guide the project team as the work unfolds. Each section addresses accessibility, institutional capacity, data sources, identification of oppression or harm, inclusion and diversity, and data sovereignty. While chapters in the volume focus on issues of representation, the checklist reveals a broader range of ethical and legal questions that could be raised in future contributions to this series such as issues related to copyright and takedown requests. Interestingly, external stakeholders are not addressed until the implementation phase, suggesting a bias towards projects initiated by GLAM institutions and engaging community stakeholders post-project development, suggesting additional room for the exploration of community-driven projects where stakeholders are involved in both planning and implementation.

*Ethics in Linked Data* examines critical issues related to the development and implementation of linked data projects. The authors carefully navigate the utopian ideals of openness and broad participation, while highlighting the barriers to entry and potential pitfalls for GLAM institutions entering this space. The broad array of case studies provide insight on how to apply reflexive, ethical practices to the development and deployment of linked data projects, balancing both theoretical frameworks and practical advice. While focused on the production and generation of linked data rather than access and discovery, the volume is perhaps representative of where the field is now, as an examination of a new technology and how GLAM organizations and professionals are positioned to make an impact in this space. The contributors demonstrate the necessity of applying critical and ethical frameworks as projects are developed to prevent perpetuating the harm caused by many descriptive and discovery systems and the imperative to represent the diverse spectrum of human knowledge. In this way, *Ethics in Linked Data* recognizes the limitations of the historical tools that GLAMs have used to build and share data, acknowledging that many “best practices” become embedded and rote, threatening to become thoughtless practices without careful re-examination as the fields push technology forward, while highlighting the incredible power that GLAM professionals have to influence and impact this space.

Lindsay Kistler Mattock  
East Carolina University, USA  
E-mail: mattockl22@ecu.edu

## References

- Carlson, S., Lampert, C., Melvin, D., & Washington, A. (2020). *Linked Data for the Perplexed Librarian*. American Library Association.
- Sandberg, S. (2019). *Ethical Questions in Name Authority Control*. Library Juice Press.