

Book Review

A review of *The Politics of Mass Digitization*

Kai Li

Drexel University, Philadelphia, PA, USA

E-mail: kl696@drexel.edu

Nanna Bonde Thylstrup. *The Politics of Mass Digitization.* Cambridge, MA: The MIT Press, 2018. 216 pp. ISBN 9780262036870 Hardback, \$35.00.

Mass digitization projects, such as Google Books, Europeana, and the Digital Library of America (DPLA), have rapidly become some of the most profound and influential digital libraries in our contemporary society. Naturally, they have attracted extensive academic research by information scientists over the past few decades. Besides the fundamental debate on how these new efforts to build the “Library of Babel” (Borges, 1998) are related to the traditional concepts of libraries, books, and reading (Darnton, 2009; Kelly, 2006; Rowe, 2013), researchers have discussed extensively about how information is or should be organized, represented, and preserved in these new digital platforms (Bamman & Crane, 2008; Conway, 2010; Dempsey et al., 2014; Jacsó, 2010). However, one research topic that has not received much attention is how such mass digitization projects are intertwined with our society, culture, and politics. Constructivist studies on science and technology have informed us that no infrastructure can be produced or operated beyond its social contexts; in fact, all technical artifacts are *co-produced* along with the social order (Jasanoff, 2004). From this perspective, Nanna Bonde Thylstrup’s new book, *The Politics of Mass Digitization*, offers a timely addition to the discussions of digital libraries. Building upon a multidisciplinary approach, the author offers a solid theoretical framework on mass digitization and uses this framework to produce meaningful and rich narratives about these new digital libraries.

The central argument of this book is that the politics of mass digitization projects are rooted in the development of their technical infrastructures. On the one side, the *late-sovereign* politics framework adopted by the author makes it possible to connect these digitization projects to a new political reality where traditional political concepts (such as national representations) and new, neoliberalist logics are merged together. On the other side, the author correctly points out that mass digitization is deeply assembling and heterogeneous in nature, aligned with how infrastructure is defined in sociological terms (Star & Ruhleder, 1994). Combining these two components, the author is able to provide a more empirically-driven dissection of these

projects that considers both global and local (*glocal*), as well as political and technical, conditions. This is the essence of the author's *infrapolitical* view of mass digitization, following a classic definition of this concept developed in political science (Scott, 1990). This approach to mass digitization pays central focuses on the underlying and often invisible political forces. While the original use of *infrapolitics* is deeply connected to the dissent and resistance of subordinate groups, Thylstrup argued that the story of mass digitalization concerns at least as much "mundane and standard-seeking practices" (p. 25) as contestation, which supports her decision to extend this concept from purely resistance to also maintenance and conformity, another theme deeply inspired by sociological research on technical infrastructure (Bowker & Star, 2000).

Applying this conceptual framework, the second part of this book discusses three examples of mass digitalization projects – Google Books, Europeana, and shadow libraries – which itself deserves acclaim because of its richness of narratives and theoretical underpinning. In some ways, these three projects constitute a structure similar to Levi-Strauss's famous culinary triangle (Lévi-Strauss, 1965), which implies that contrasts always exist beyond a single dimension. The author did a great job to illustrate the subtleties of differences and similarities among these projects. For example, even though Europeana was originally proposed as a political competitor based on a sovereign paradigm (European vs. American), the author argues that it was subsequently merged into the similar late-sovereign framework in which Google Books is operated, in terms of how a mass digital library should be both built and maintained infrastructurally, economically, and politically. An important aspect of Europeana's late-sovereign nature is its goal to build "networked infrastructures of connectivity, competition, and potentially operating beyond, below, and between national societal structures" (p. 64), which suggests how difficult it is to maintain geographical boundaries in the digital realm.

Moving towards the other part of the triangle, this book focused on three shadow library projects in different countries and with distinct histories and business models: lib.ru, Monoskop, and UbuWeb. The author traced the different origins and developments of these projects. They were developed from different political, economic, cultural, and even personal factors. Despite their differences, the author claimed that all these projects have been connected to and thus are strengthening the globalized information infrastructure and infrapolitical ecosystem, rather than threatening them. This ecosystemic view of the global digital infrastructure could free up researchers from the dichotomy between legal and illegal libraries and has great potential to create new perspectives on how information and knowledge flow globally.

While not an information scientist herself, Thylstrup addresses an important topic – mass digitization – in the field of information science by drawing theories and empirical evidence from various research fields. This book opens up new possibilities for studies on not only large-scale digital library but also many other topics related to information infrastructure. For example, library systems could be put under this multidisciplinary and theoretical gaze to understand how the system and

interface used in the library setting should be rethought and redesigned in the new information and socio-political environments. Thylstrup's new book does a wonderful job of delineating the situatedness of mass digitization projects in complex technical and sociopolitical conditions, which deserves more attention by the community of information science.

References

- Bamman, D., & Crane, G. (2008). Building a dynamic lexicon from a digital library. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries* (pp. 11–20). ACM.
- Borges, J. L. (1998). The library of Babel. *Collected Fictions*.
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out: {Classification} and its consequences*. MIT press.
- Conway, P. (2010). Preservation in the age of Google: Digitization, digital preservation, and dilemmas. *The Library Quarterly*, 80(1), 61-79.
- Darnton, R. (2009). *The case for books: Past, present, and future*. PublicAffairs.
- Dempsey, L., Malpas, C., & Lavoie, B. (2014). Collection directions: The evolution of library collections and collecting. *Portal: Libraries and the Academy*, 14(3), 393-423.
- Jacsó, P. (2010). Metadata mega mess in Google Scholar. *Online Information Review*, 34(1), 175-191.
- Jasanoff, S. (2004). *States of knowledge: the co-production of science and the social order*. Routledge.
- Kelly, K. (2006). Scan this book! *The New York Times*, 14.
- Lévi-Strauss, C. (1965). Le triangle culinaire. *L'Arc*, (26), 19-29.
- Rowe, C. (2013). The new library of Babel? Borges, digitisation and the myth of a universal library. *First Monday*, 18(2).
- Scott, J. C. (1990). *Domination and the arts of resistance: Hidden transcripts*. Yale university press.
- Star, S. L., & Ruhleder, K. (1994). Steps Towards an Ecology of Infrastructure: Complex Problems in Design and Access for Large-scale Collaborative Systems. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work* (pp. 253-264). New York, NY, USA: ACM. <https://doi.org/10.1145/192844.193021>