
Online Shadow Libraries and the Future of Humanities Scholarship

Stephen Reid McLaughlin

stephen.mclaughlin@utexas.edu

University of Texas Austin, United States of America

In a publishing landscape long challenged by small audiences, tightening library budgets, and steady growth in the number of monographs produced each year, potentially illegal text sharing has become an everyday practice for many humanities scholars and students, both those with access to conventional distributions channels (libraries, online databases, inter-library loan) as well as independent scholars and those at institutions with limited library resources. Whereas many researchers in medicine and natural science disciplines have embraced open access publication practices, both in open-licensed journals and via centralized preprint repositories, access to scholarship in humanities fields poses greater challenges.

Due to humanities monographs' limited commercial value, many titles are only commercially available for a short window. As Borgman has noted, "[l]iterature in the humanities goes out of print long before it goes out of date" (2007: 241), while older works from university presses are "virtually entombed" (McGann, 2014: 133) in the stacks of academic libraries. Meanwhile, factors including rising journal prices and an increase in the number of monographs published each year have led libraries to purchase proportionally fewer monographs than in past decades (Thompson, 2005: 103–6). In turn, humanities scholars' reputation for "unreadable complexity" (Eve, 2014: 22) may be exacerbated by the unavailability of published research. By contrast, informal circulation can provide a double benefit to scholars, allowing them to secure the prestige of publishing with a name-brand university press while reaching a larger audience than would be possible in print alone (Hall 2013). In this paper, I draw on public metadata sets, download logs from Sci-Hub (Bohannon and Elbakyan, 2016), and interviews with operators and users of illicit text sharing sites to examine their place in contemporary humanities scholarship.

A decade ago, illicit text collections, also known as shadow libraries (Liang 2012; Bodó 2015, 2016), were limited in size and existed on the margins of academic culture. Today, one can find millions of books and tens of millions of journal articles spread among Library Genesis, Aaaaarg, and Sci-Hub, as well as in niche collections such as Memory of the World and Monoskop Log. And yet, these sites' coverage of literary works and literary scholarship is significantly shallower than is the case for published research in medicine, engineering, and the social sciences.

Most shadow libraries are messy, ad hoc affairs, composed of digital objects in a range of formats and quality levels, drawn from a pre-existing ecosystem of interpersonal text sharing among scholars and students. Metadata for the 1.5 million documents in Library Genesis, which is freely available, exhibits what we might call bounded messiness. A tabular dataset riddled with missing values, text encoding quirks, and duplicate entries, it clearly would not pass muster in an academic library setting. However, the simplicity and flexibility of this system make it well-suited for a text collection compiled, maintained, and mirrored by a culturally diverse community of participants. In this case, a just-good-enough database supports the goals of inclusiveness and replication by others.

In interviews with scholars who use shadow libraries, I have observed a wide range of positions with respect to copyright law and the business of scholarly publishing. While some hope for a future in which copyright is abolished and publishers are driven out of business, many are essentially satisfied with the scholarly publishing ecosystem as it stands today. Most use library resources and purchase physical books, turning to shadow libraries to fill the gaps and evaluate texts' relevance and quality. Nearly all respondents said they prefer to read printed texts when the option is available. Moreover, if the cost of these sites' continued existence is that they remain culturally marginal, my participants see this as an acceptable tradeoff.

Several interview participants described their personal document management schemes, an age-old form of scholarly labor that typically remains invisible to outsiders. What is new, however, is that some avid collectors are systematically compiling clean copies of their personal libraries, along with extensive metadata and searchable full-text indexes. They then share these collections among colleagues, either on hard drives or in private online repositories. One respondent described this text curation practice as both a pedagogical strategy and an attempt to shape the scholarly canon in his field of study.

There are no easy answers to the problems shadow libraries pose for publishers and university libraries, leading Bodó (2016) to suggest we should “bet on all horses,” throwing support behind academic publishers and shadow libraries alike. For now, exploring how where shadow libraries come from, how they are designed, curated, and maintained, and what their futures may hold — including the eventual, inevitable, closure of individual sites — may help us understand current academic culture and possible future models for humanities research and scholarly culture more broadly.

Bibliography

- Bodó, B.** (2016). “The Knowledge Singularity.” RADiCAL.PiRATiCAL symposium, Malmö, Sweden, May 2016. <https://radical.piratical.cryptonomic.net/archives/the-knowledgesingularity>.
- Bodó, B.** (2015). “Libraries in the Post-Scarcity Era.” In **Porsdam, H.** (ed.), *Copyrighting Creativity: Creative Values, Cultural Heritage Institutions and Systems of Intellectual Property*. Farnham, UK: Ashgate, pp. 75–92.
- Bohannon, J. and Elbakyan, A.** (2016). “Data from: Who’s Downloading Pirated Papers? Everyone.” <http://datadryad.org/resource/doi:10.5061/dryad.q447c>.
- Borgman, C. L.** (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press.
- Eve, M. P.** (2014) *Open Access and the Humanities: Contexts, Controversies and the Future*. Cambridge, UK: Cambridge University Press.
- Hall, G.** (2013). “The Unbound Book: Academic Publishing in the Age of the Infinite Archive.” *Journal of Visual Culture*, 12(3): 490–507.
- Liang, L.** (2012). “Shadow Libraries.” *e-flux journal* 37.
- McGann, J.** (2014). *A New Republic of Letters: Memory and Scholarship in the Age of Digital Reproduction*. Cambridge, MA: Harvard University Press.
- Thompson, J. B.** (2005). *Books in the Digital Age: The Transformation of Academic and Higher Education Publishing in Britain and the United States*. Cambridge, UK: Polity Press.