As with many commodities in the world, there often exists a knowledge gap between the consumer and the producer, whether in the context of food, clothing, or, in this case, access to digitized information. The invisible work that occurs in today's libraries is vast and complicated, perpetually balancing issues of acquisition, budgets, licensing, copyright, and privacy, among others. The complexities at play need not be known by the consumers of the information provided, and, when times are good, are arguably irrelevant to those end-users. Over time, the complexity of the information landscape has exponentially increased, as have “user expectations and professional expertise with digital technology” (p. 3). Deanna Marcum and Roger C. Schonfeld set out to capture and illuminate Google’s role in the multifaceted approaches, success, and failures that could have resulted in a universal digital library platform through which the world's information could be accessible. From the accounts, chronology, and interpretations presented, the reader gains a granular understanding of just how challenging similar large-scale library initiatives can be, and even more so when public and private entities mix despite similar goals.

The book is composed of eight chapters that I split up into three major sections, the first of which examines the push for increasing access to information, seemingly welcomed by some libraries as evident by their willingness to oblige Google in its request to digitize institutional holdings. Leading up to this moment, “no digitization of ... content had yet been undertaken, but the stage was set for thinking of libraries as a national network of information resources” (p. 32). Google saw an opportunity in working with several key libraries, including the University of Michigan, and with Google and specific university administrators being in the right place at the right time, the promise of increasing access while working with a corporate partner gained momentum. The authors intertwine the vision of a national digital library as underlying some libraries’ motivation, or justification, to work with Google to have their materials digitized. The authors touch on this...
theme throughout the book, and the complexities of such a notion become apparent and are reinforced in each chapter. In particular, the first several chapters provide a sense of optimism among the players involved in the early mass digitization, only to be thwarted later by the realities of copyright, licensing, and royalties.

The next several chapters portray the Google Book Project as the catalyst for transforming ideas of information access and library digitization into attainable realities, and they draw a very clear distinction between the worlds of academia and commercial enterprise. This contrast is touched on throughout the book in a lighthearted manner, with the authors often pointing to the cautious pace of the library world as well as the attention to detail that librarians (self-admittedly) often require in decision-making, whereas Google pursued a full-throttle approach, only to be slowed by some skeptical librarians, and eventually publishers and the courts. Ultimately the courts led to the significant reduction of Google’s efforts and limitations on the work that had been done up to that point, so much so that copyright law, publishers, authors, and librarians, in essence, put an end to what some thought “could have become a national, comprehensive digital library” (p. 151). Copyright is discussed in great detail in the context of the legal challenges brought against Google, as well as for both digital and physical library materials.

The remainder of the book looks at the aftermath of the Authors Guild v. Google, Inc. court settlement, from which “Google agreed to a settlement that limited dramatically its rights to use ... in-copyright materials” (p. 150). The void that was left was not all negative, however, as the Google Book Project itself demonstrated the possibilities of library digitization and online access to those materials. Additionally, more focus was now directed toward digital preservation and not just access, whereas Google’s focus had been on workflow efficiency and subsequent access. Efforts emanating from the University of Michigan, which continued to be a key driving force throughout the events in this book, aimed to secure institutional collaboration to both capitalize on the work that Google had completed pre-settlement and forge a path forward to build on that work, with both access and preservation in mind. HathiTrust emerged as the repository platform that could provide both access and preservation elements as well as shared governance; however, the realities of potential limitations remain; as HathiTrust executive director Mike Furlough comments, “Hathi is a preservation strategy, but it also does more with access. Preservation is a hard foundation to build on. People pay lip service to preservation, but it is hard to get people to pay for it” (p. 187, the citation indicates an interview with Furlough on May 25, 2016, conducted by Marcum). Despite all the litigation, new collaborations, and shifted priorities, “libraries and archives remain uncertain about what can and cannot be digitized and made widely accessible to the public” (p. 159).

As a side note, although chapters 3 and 5 discuss preservation standards and image-quality concerns, this book does not directly address the evolution of digitization standards or digital preservation strategies over the time period examined. This does not suggest that the book is deficient in any way, particularly as the objective is to examine Google’s impact on traditional scholarly communication. (Instead, in the context of a national digital library discussion, even though the Library of Congress did not take the lead in this regard, significant work has been undertaken by other players to at least provide pseudo-national standardization related to image quality and preservation, namely as one of many federal government agencies that comprise the Federal Agencies Digital Guidelines Initiative [FADGI].)

The authors set out to “tell the story of how Google attempted to enter, and in some senses disrupt, the traditional scholarly communications system that served the universities, their scholars
and students, and their libraries for decades” (p. 5). I would argue that the authors do much more than that here; the extent to which they bring together details from interviews, memorandums, official communications, personal accounts, and publications is impressive. This study brings to life the inner workings of major institutional agreements among themselves and with Google. The authors also provide a backdrop for much of the evolution of other major library digital projects over the last several decades, including HathiTrust, the Digital Public Library of America, Open Content Alliance, Europeana, Digital Library Federation, JSTOR, and ITHAKA, among others. Again, the complexities of these projects and their interconnectedness allow the reader to gain an improved understanding of how numerous the challenges have been to harness information access and preservation in the digital age. I believe this book to be of interest to those in the library, archives, and museum studies fields, particularly those involved with digital projects, or those interested in the history of library digitization efforts.

If there is additional discussion of this review, you may access it through the network, at https://networks.h-net.org/h-sci-med-tech


URL: https://www.h-net.org/reviews/showrev.php?id=59892

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 United States License.