Using Massive Digital Libraries

Lizzy A. Walker, Wichita State University
Overall, this book is well organized and well written with clarity of thought that is refreshing. The authors’ extensive professional experiences are clearly at play in the way they can easily and succinctly illustrate a point to the reader. Consortia are a vital part of any library relationship, and this book can serve to help the unfamiliar gain a good foundation for the role consortia play in libraries today and possibly in the future.—Ryan Litsey, Texas Tech University


In his book, Andrew Weiss takes what could be written in volumes and condenses the information into one accessible book. Massive digital libraries (MDLs) have had substantial impact on the world. Weiss provides a solid case on the potential that MDLs have for future use in libraries.

Weiss splits the book into three sections: background, philosophical issues, and practical applications. The first four chapters cover background. He presents a very short history in the first chapter on digital libraries, the merging of libraries and computer science and engineering, and other interesting topics. In the second chapter, Weiss tackles defining MDLs by outlining criteria and characteristics of an MDL, including collection size; acquisitions, collection development, and copyright concerns; collection diversity; and more. Weiss looks at several MDLs from the United States and Europe in the third chapter, such as HathiTrust, Google Books, and Europeana. He provides fact sheets on each project, giving the reader an easy-to-read overview of the particular MDL.

In the fourth chapter, Weiss takes traditional library roles, like providing access to print material and rare or old materials, interlibrary loan, and course reserves and discusses how MDLs may impact them. He includes examples, such as the Spanish Language Virtual Library of Historical Newspapers project, the Vatican library’s digitization project, and Melville’s Marginalia. This chapter includes information on MDLs’ effect on collection development, consortium opportunities, preservation and quality of digital and print assets, challenges with metadata, and other topics.

In the second part of his book, Weiss covers some important philosophical issues regarding MDLs. These projects can’t be addressed without mentioning copyright, of which the fifth chapter is devoted. Weiss argues that copyright law benefits not the authors or creators of content; instead, it benefits distributors and publishers. He discusses three important lawsuits brought forth by publishers, not authors, fighting over copyright. Weiss also covers current copyright law, the Digital Millennium Copyright Act (DMCA), the impact of MDLs on copyright law, and other topics. This chapter has the most robust references and notes section. It would be worth it for those interested in copyright to take a look at this chapter. Weiss presents an interesting comparison of the collection development approaches of HathiTrust, Google Books, and other MDLs in chapter 6. A topic that goes hand in hand, according to the author, with collection development is collection diversity, covered in chapter 7. The author argues that diversity is related to collection development policy and discusses a few communities. The author addresses future directions, European criticism and approaches to content diversity, and positive trends in MDL diversity.

Weiss looks at open access (OA) models and MDLs in chapter 8. There is a for-or-against divide between publishers on OA, which only causes more of a rift between libraries and publishers. He mentions that most of the OA discussion has focused on journal publishing primarily, so he provides extensive discussion on OA books. Again, he looks at Google Books, HathiTrust, Europeana, and the Open Library in regard to digital content and copyright and presents pros and cons of each MDL. He argues

doi:10.5860/crl.76.6.851
that, if there comes a point when copyright laws release its grip on digitized content, MDLs have the potential to be a highly effective OA tool.

In part III, practical applications, Weiss looks at real-world use of MDLs. In chapter 9, Weiss looks at how MDLs can be implemented in libraries to help user experience and to exploit their use of services such as Google and Google Scholar. He covers such services as Google’s “My Library” service, and Print-on-Demand. He makes some suggestions as to how libraries could implement MDLs into their systems. He also mentions using a hybrid approach where libraries could provide users with a physical copy of a digitized public domain, print on demand, and that libraries could charge a small fee for this service. Weiss also discusses access and print preservation, aggregation of variant texts in regard to digital humanities, how MDLs can help decision making on weeding collections, how MDLs can be a solution to demand driven acquisitions, could be useful for course reserves to link out to an MDL. Other topics include pros and cons of using MDLs as a research tool for digital humanities, the Google Books n-gram viewer, real-world applications, integrating the online catalog with the Google Books API, and more.

In chapter 10, Weiss includes four case studies of library projects that evaluate MDLs: the legibility of Google Books; the quality of Google Books’ metadata; culture and MDLs; and the Keio University Google Books digitization partnership. Each of these studies provide interesting information on MDLs in libraries and are worth a read for anyone interested in real-world applications that other libraries are using.

Weiss’ book contains valuable information for anyone interested in MDLs and how they can be useful for libraries. Each chapter includes references and notes as needed, which could provide the reader additional valuable information. The index is well organized and highlights the essentials.

Weiss is the digital services librarian at the Oviatt Library at California State, Northridge. He credits Ryan James with coauthorship in some chapters. —Lizzy Walker, Wichita State University Libraries