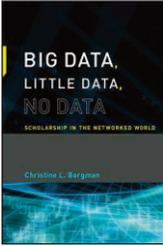


Reviews



Big Data, Little Data, No Data: Scholarship in the Networked World. By Christine L. Borgman. Cambridge, London: The MIT Press, 2015. 416 pp. ISBN 9780262529914.

In addition to ongoing media enthusiasm, “big data” has recently become the subject of numerous academic studies and publications. While the media attention is partly due to the commercial value that big data has for cultural and digital industries, it overshadows the fact that having *better* data is very often more important than having a lot of it. It also ignores the fact that “small” data can be more effective than big data, or that data can sometimes not be found or be unavailable. This is the focus of *Big Data, Little Data, No Data: Scholarship in the Networked World*, in which Christine L. Borgman examines the complexity and diversity of data for different fields of study. Starting from the premise that data are not natural objects with their own essence, Borgman rather explores the different values assigned to them, as well as their many variations according to place, time, and the context in which they are collected. It is specifically through six “provocations” that she offers a deep engagement with different aspects of the knowledge industry. These include the reproducibility, sharing, and reuse of data; the transmission and publication of knowledge; the stability of scholarly knowledge, despite its increasing proliferation of forms and modes; the very porosity of the borders between different areas of knowledge; the costs, benefits, risks, and responsibilities related to knowledge infrastructure; and finally, investment in the sustainable acquisition and exploitation of data for scientific research.

The book is comprised of three broad sections, the first of which is divided into four chapters and proposes multi-layered definitions of the concepts of data, knowledge, and infrastructure, as well as an exploration of the diversity of practices employed in current scientific research. The second section explores the respective uses of data in the sciences, social sciences, and humanities, and in each chapter here, Borgman compares an area of study that requires a large volume of data with a field requiring smaller, more localized data. These comparative analyses thus reveal the various requirements of different scientific fields in terms of their knowledge structures. They

We have some excellent books awaiting your review to be published in CJC.

Click on the Books for Review tab on cjc-online.ca's home page, choose your book, and follow the online instructions. Reviewers can be either faculty members or graduate students. Or, if you are keen to review a particular title that you have encountered and feel would benefit other readers, our Book Review Editors can request a review copy for you from the publisher. They also want to hear from authors. If you have a publication that you think the CJC should review, please contact either **Penelope Ironstone** at review_editor@cjc-online.ca or **Dominique Trudel** at redacteur_revue@cjc-online.ca.

are also an opportunity for Borgman to account for the tensions that can arise among stakeholders (researchers, assistants, teachers, employees in the private and public domains), as well as the practices and motivations behind the collection, sharing, and reuse of data. The third section explores the challenges facing the study of data with respect to politics and the different methods used for collecting data. In particular, Borgman here emphasizes the importance of acknowledgement in the practice of data collection, as well as the principles that should guide researchers in selecting the data to be retained. The author presents the book as one of the first to deal with data from social, technical, and political perspectives and to be based, moreover, on examples from various academic disciplines. In this respect, the book opens the “black box of data” to better understand how knowledge is distributed across different scientific fields. Much attention is focused on the concept of “knowledge infrastructure,” defined here as “a robust network of people, artifacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds” (p. 33), which allows Borgman to highlight all the steps in data acquisition. The comparative analyses in the second section are particularly illuminating, and reveal the actual work required of contemporary researchers, whether alone or in small groups, in acquiring data. Contemporary research practices have incorporated a whole host of tools, instruments, interpretations, and methods in order to obtain data that are as accurate and reliable as possible. Borgman also investigates the various sizes and sources (and origins) of the data resources used by researchers. These dimensions are particularly interesting in the context of the humanities (Chapter 7), which, beyond the work of defining the blurring boundaries among the different disciplines, face much greater challenges in adapting to new technologies and archival tools within institutions (universities, schools, libraries, archives, museums, governments, and private owners).

In terms of the six provocations presented early in the book, the demonstration is clear and presented in simple, accessible writing. However, the book strives for an overly comprehensive review of all forms of data and their connections with many areas of knowledge, an enterprise so ambitious that it can appear somewhat didactic. Although the renewed attention paid to small-scale data collection is justified and necessary, Borgman’s willingness to explore all systems of knowledge within scientific research means that she has to define in depth the panoply of concepts on which the study is based, including concepts of data, scientific fields, knowledge structures, and the various practices in scientific research. Not only does this work of definition—more present in the first and third parts—weigh down the reading, it also obstructs a deeper engagement with the fundamental reflection that “the enthusiasm for big data puts smaller-scale scholarship at risk” (p. 15). This idea, which appears to be the premise of the book, is only outlined through each of the thematic provocations, as she explores how academic studies are conducted today and will be in the near future. Instead of devoting so much space to definition work, the book would probably have benefited from a more sustained and nuanced analysis of the epistemological, sociological, and philosophical implications of the recent widespread enthusiasm for big data. Because ultimately, beyond the six provocations, there is a larger question concerning the legitimacy, continuity, and durability of all scientific research—hence the urgent need

for further reflection, initiated eloquently by Borgman, on the fact that “despite the media hyperbole, having the right data is usually better than having more data” (p. 1).

Thomas Mayer Lemieux, National Institute of Scientific Research (NRS)