The International Journal of Digital Curation
Issue 1, Volume 3 | 2008

Review: *Scholarship in the Digital Age*

Chris Rusbridge,
Digital Curation Centre

June 2008

Summary

A review of *Scholarship in the Digital Age: Information, Infrastructure, and the Internet* by Christine Borgman, MIT Press, 2007, hardback, 336 pp, ISBN-13: 978-0262026192.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. ISSN: 1746-8256 The IJDC is published by UKOLN at the University of Bath and is a publication of the Digital Curation Centre.

Christine Borgman (2007) takes on a massive task in her most recent book in which she examines many of the ways our new digital landscape affects science and scholarship as a whole. In particular this work is about the systems for sharing and communicating science and scholarship. This an important topic, and a significant book. Speaking merely personally, I have already heard it verbally cited in an international discussion session on supporting new scholarship.

This book should be required reading for vice-chancellors, pro-vice-chancellors for research, leaders of research-funding bodies worldwide, librarians and anyone managing a repository whether for data, text or other forms of digital content that represent science and scholarship. In style however, it appears strongly oriented towards social scientists.

The book is organised into 9 chapters, starting with scene-setting in "Scholarship at a Crossroads", and ending with a look forward in "The View from Here". In between we have:

- Building the Scholarly Infrastructure
- Embedded Everywhere
- The Continuity of Scholarly Communication
- The Discontinuity of Scholarly Publishing
- Data: The Input and Output of Scholarship
- Building an Infrastructure for Information
- Disciplines, Documents and Data

I found the two chapters on scholarly communication and publishing particularly hard going; I felt they would have benefitted from being more heavily edited and condensed into a single chapter. They cover ground that has been gone over many times in recent years; while the discussion is valuable, it stands in the way of what seems to me the more valuable part of the book. This is a rapidly changing area, and some emerging developments are ignored: for instance, the role of blogs and wikis in scholarship is not mentioned (blogs get one index entry, to page 1; wikis are not in the index). Social networking appears in the index linking to invisible colleges rather than the tools for shared construction of knowledge. The emerging role of sites like Wikipedia (which has progressed from being denigrated to that of the valued first cautious check for unfamiliar material in a very few years) is not mentioned. But this is a book, the product of long gestation, rather than an up-to-the-minute report. Somewhere here, Borgman does discuss the different role of books, but indexes and tables of contents are poor search mechanisms, and despite several attempts to find it, the reference eluded me.

The real value of this book for me were the chapters on data. The role of data in science and scholarship is a recent emerging theme, but one that has rarely been treated in books, and never as comprehensively as this. Borgman has worked on several science project teams herself, and talked with many other researchers, as well as reading extraordinarily widely. There is a reasonable discussion on data sources and data levels (not all data are raw data, although one person's information product may be another person's raw data). There is a good discussion of how disciplines (or even sub-disciplines, not to mention individual projects) treat data differently. Borgman understands that the information user and information creator have different views on

data sharing, even when they are the same person. She points out the weakness of current policy drivers. “There are [few] incentives, and many dis-incentives, to share … data or make them publicly available.” (Borgman, 2007)

To highlight a few particular points among the many that struck me: the first was a discussion on tacit knowledge, the second on language, and third the social situation of data. All are serious issues in the long term. “Tacit knowledge is most easily explained by the epistemological distinction between knowing how and knowing that… To learn *how* requires applying knowledge of *that* in practice.” She later describes a chemist reporting an experiment; other chemists “can infer much of [the unreported tacit] knowledge through their own background and experience”. Yet even they cannot infer everything, and scholars in areas further afield will have even greater difficulties.

Related to this is language. Scholarly information is not perfectly translatable; language terms are overloaded, and neologisms abound. I do not think Borgman mentions it, but in some genomics areas genes have names like “sleepy” and “hedgehog”. Controlled vocabularies and ontologies are part of the struggle to maintain comprehensibility but she reports that for some, “forfeiting the richness of local language is too high a price to pay for interoperability”.

Elsewhere she writes “Information, whether documents, data, or composite objects, and whether in documentary form such as publications or physical artefacts such as plant specimens, is situated in a social context. When others wish to use the information later, they need to know the context in which it originated to interpret it correctly.” I should confess I had an email discussion with Christine around this topic in 2007; my suggestion was “data are not neutral with respect to the hypothesis”. She referred me to Bruno Latour’s book (Latour and Woolgar, 1986), and mentioned this forthcoming volume. I devoured the Latour account with relish!

OAIS (CCSDS, 2002) attempts to capture some of these ideas within the concepts of the Designated Community and its Knowledge Base. It is a worthy effort, and the best we have at this stage, but the definitions are not yet precise enough to be free from problems in practical data archives.

What are Borgman’s conclusions? She realises that “attempts to predict or invent the future [are] based on the knowledge available today”. She outlines three promising directions: first, invest in content. Second, balance local and global approaches to design. Third, take architectural approaches to infrastructure, separating content, services and tools. Finally, Borgman cautions us to remember that many critical issues (trust, identity, location and intellectual property) “are social issues first and technical issues second”.

With around 800 references, this book is daunting. The social science “build on the past” style is very much in evidence. The author liberally sprinkles these references through the often dense text; but too frequently I failed to see quite how they actually supported the argument. I found myself flicking to the back of the book to get more clues from the full citation - and finding myself wanting to click on it and read the key sections. But this was a book and I was on an aeroplane; not allowed! (There is a

companion Web site[1], which sounded exciting, but in fact only contains the bibliography… but if you read with the book on your lap and the computer by your side, you might achieve what I actually wanted to do!)

I cannot resist a meta-comment, on the process of deep reading for this review. Although by no means a Generation X person, most of my technical reading is online, or near online. As may be obvious, the passive nature of the book was frustrating; I wanted to search and to link, and could not do so. Moreover, the physical construction was annoying; this well-built hardback would not stay open at my page, and most of the time the dust-cover flap was inside the book to keep my place, often obscuring part of the content. On the other hand, once I had managed to give myself permission to make marginal notes (it is a review copy, and only in pencil…) I found the ability to annotate really liberating. I have yet to find an online annotation substitute that works well.

But, my quibbles aside, this is an important and valuable contribution to understanding the world of scholarship in which we live. Those who shape or contribute to this world should buy and read it!

References

Borgman, C.L. (2007). *Scholarship in the digital age: Information, infrastructure, and the Internet*. Cambridge, MA: MIT Press.

Consultative Committee for Space Data Systems. (2002). Reference model for an open archival information system (OAIS). Retrieved June 23, 2008, from http://public.ccsds.org/publications/archive/650x0b1.pdf

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton, NJ, Princeton University Press.

[1] *Scholarship in a Digital Age* http://polaris.gseis.ucla.edu/cborgman/PubsDA.html