Richard Akerman
book review for *Nature*
author manuscript Nov 5, 2007

Open science enhanced by the Internet is making the invisible college visible. Many constraints on communication have been removed, providing opportunities for much broader and richer dissemination and discussion of scientific theories, data and results. The academy and society are still adapting to the rapid changes in the technology and capabilities of the "scholarly infrastructure", and exploring new business models and reward systems.

In *Scholarship in the Digital Age*, Christine Borgman provides a comprehensive overview of these transformations. Her main goal is to ensure that the terminology is well-defined, and the historical perspective well-documented. She states "While the new technologies receive the most attention, it is the underlying social and policy changes that are most profound... This is an opportune moment to think about what [future scholarly environment] we *should* be building."

The book is therefore more about describing current scholarly communications practices and issues set in a historical context than about the technology details. It is a comprehensively referenced study with an ambit encompassing all of science. The scholarly information system is presented as embedded within and constrained by a complex sociopolitical environment, which provides a useful balance to the often ambitious and sometimes utopian goals of major e-research projects.

In order to frame the problem space, the author first presents an examination of the current scholarly communications system, describing the mechanics of the peer-reviewed journal publishing system, the types of scientific conferences, and the roles of authors. This will be well-worn ground for anyone involved in the dissemination of science, but provides a good foundational summary that is necessary if one is to understand the disruptive impacts of new technologies.

It is to this discontinuity that Borgman turns next, exploring the fundamental challenges presented by openly circulating digital information, with a focus on scientific publication. While some of the closed social networks that formed invisible colleges of researchers have now opened up, their erudite discussions, carefully-gathered data and trusted research are now sharing space with, and competing for search ranking with, materials of wide-ranging quality and legitimacy.

As Nature explored in its Peer Review Debate Web Focus

http://www.nature.com/nature/peerreview/debate/

there are challenging issues such as how to select, register and certify scholarly documents in a networked digital environment. As well, how can one locate quality articles, particularly when much of the scholarly record is only available to licensed users

and may be weakly represented in general search results? This brings the author to a detailed discussion of open access, including intellectual property and copyright issues, as well as commons approach (such as Creative Commons).

Moving beyond the fairly well-understood area of scholarly publishing, the next area examined is data in all of its many aspects: from a basic definition of what data are, to the new types and increasing volume of data and its role in scholarly communication. As well, the same certification, curation and intellectual property issues previously examined for publications are examined in the context of data, which has different traditions and conventions.

The author uses all of these issues as the background to her concept of an infrastructure for information, an "advanced scholarly information infrastructure". However those looking for detailed technical information or guidance may come away a bit disappointed, as the focus is entirely on the various behaviours and patterns of different scholarly disciplines, the nature of scholarly artifacts, and collaboration traditions and characteristics. To give one example, in the entire discussion of collaboration in various sections of the book, there is no mention of "wikis", particular types of websites that provide the ability to collaboratively edit and share text, of which Wikipedia is the best-known example. This is by intent, not omission.

The book concludes by providing a compelling look at the many research problems ahead, problems which a comprehensive scientific cyberinfrastructure will need to recognise and address. The author asserts that a focus around the "content layer" is essential. She suggests four promising areas of research:

1) taking a long-term view on content
2) balancing local and global requirements
3) separating content from tools and services
4) identifying tools, technologies and institutional mechanisms that balance the coherence and control of the content layer

Science policymakers would do well to use this book in framing their discussions of the future directions for scholarly infrastructure. Students of information science as well as those seeking a comprehensive understanding of science in our digital age will find it an valuable reference and starting point. In a world where scientific networks and communication are now increasingly visible and open, Christine Borgman has illuminated the discussion of the scholarly communication system itself.