# Preface

Scholars in all fields are taking advantage of the wealth of online information, tools, and services to ask new questions, create new kinds of scholarly products, and reach new audiences. The Internet lies at the core of an advanced scholarly information infrastructure to facilitate distributed, data- and information-intensive collaborative research. These developments exist within a rapidly evolving social and policy environment, as relationships shift among scholars, publishers, librarians, universities, funding agencies, businesses, and other stakeholders. Scholarship in the sciences, social sciences, and humanities is evolving, but at different rates and in different ways. While the new technologies receive the most attention, it is the underlying social and policy changes that are most profound and that will have the most lasting effects on the future scholarly environment. This is an opportune moment to think about what we *should* be building.

This book is grounded in developments of the twenty-first century, set in both a social and historical context. Today's initiatives in cyberinfrastructure, e-Science, e-Social Science, e-Humanities, e-Research, and e-Learning emerged from a tumultuous period in scholarly communication in which technological advances converged with economic and institutional restructuring. Every stage in the life cycle of a research project now can be facilitated—or complicated—by information technologies. Scholars in the developed world have 24/7 access to the literature of their fields, a growing amount of research data, and sophisticated research tools and services. They can collaborate with other individuals and teams around the world, forming virtual organizations. Data have become an important form of research capital, enabling new questions to be asked

by leveraging extant resources. With the mass digitization of books now under way, previously unforeseen possibilities arise to compare literary themes, extract details of events, improve machine translation, and compile extensive indexes and directories. Text and data mining promise everything from drug discovery to cultural enlightenment.

These wondrous capabilities must be compared to the remarkably stable scholarly communication system in which they exist. The reward system continues to be based on publishing journal articles, books, and conference papers. Peer review legitimizes scholarly work. Competition and cooperation are carefully balanced. The means by which scholarly publishing occurs is in an unstable state, but the basic functions remain relatively unchanged. Research data are another matter entirely. While capturing and managing the "data deluge" is a major driver for scholarly infrastructure developments, no social framework for data exists that is comparable to that for publishing. Data can be discrete digital objects, but their use is embedded deeply in the day-to-day practices of research. Scholarly infrastructure also must be understood in the context of legal, policy, and economic arrangements. The "open-access movement" to expand the availability of scholarly publications, data, and other information resources is grounded in several centuries of Western thought about "open science." Open science, in turn, is based on economic principles of public goods. The ethos of sharing that is fundamental to open science and scholarship is threatened by the expansion in scope and duration of copyright protection and patents. These tensions, in turn, are reflected in new forms of publishing and licensing, such as the "information commons" or "knowledge commons." The many stakeholders in scholarly information infrastructure are addressing their own territories, whether technical, legal, economic, social, or political, or in individual research domains, but few are taking a big-picture view of the interaction of these factors. The integrative and interdisciplinary analysis presented here is intended to provoke a conversation among the many parties whose interests depend on a rich and robust scholarly environment.

The argument is laid out in nine chapters, grouped thematically into three parts. The first three chapters frame the issues. The next five chapters identify the problems to be solved if the vision of an ideal scholarly

information infrastructure is to be achieved. Chapters 4 and 5 compare the stability of the scholarly communication system to the instability of the current scholarly publishing system. Chapter 6 compares the role of data in scholarship to that of publications, explicating notions of "data." Chapter 7 explores the social, behavioral, and policy context of data and documents in more depth, set in the context of information infrastructure. Chapter 8 provides an extensive analysis of the scholarly artifacts, practices, and incentive structures of the sciences, social sciences, and humanities, assessing how infrastructure issues are playing out in each discipline. Chapter 9 concludes the book by synthesizing the issues and laying out a research agenda.

My previous book in this series, *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World* (MIT Press, 2000), examined social, technical, and policy issues of information infrastructure with a particular emphasis on digital libraries and the role of libraries as institutions. That book provides a point of departure for this one. It is not assumed that readers will be familiar with the earlier book, although it offers useful background on developments through the 1990s and information infrastructure issues outside the realm of scholarship.

This book draws on literature from many different disciplines and specialties. Each reader will be familiar with some of these sources, but familiarity will vary by community. In the interest of expanding the conversation beyond individual specialties and stakeholders, I have provided an extensive set of references. A permanent Web page associated with this book is located at <http://snipurl.com/BorgmanDigitalAge>.

Let the conversation begin.

Christine L. Borgman
*Los Angeles, California*
*June 2007*