

METHODS AND DATA

This literature review examined a heterogeneous collection of materials. These materials were found by conducting searches of the Web as well as more traditional bibliographic sources. Extensive research has been published by King, McDonald, and Roderer (1980) and by Tenopir and King (2000). This literature review built upon their work.

OVERVIEW OF LITERATURE EXAMINED

To examine the implications of information technology for scientific journal publishing, the study identified and reviewed 382 specific items,⁸ in addition to various bibliographies, bibliographic essays, bibliographic utilities, and websites. The items reviewed are listed in appendix A; they were drawn primarily from U.S. sources and are written in English. Since the focus of this effort was the formal, refereed professional literature, most of the items reviewed are journal articles. Books, conference papers, magazine articles, white papers, and reports were also reviewed. The study did not examine product reviews, whose principal purpose is to advertise or evaluate commercial products and systems for purchases; scientific databases such as those for protein sequences, genomic data, and measurements that support seismic, climate, and meteorological studies, imagery, and mapping; and computational technologies that support data collection and analysis through complex instrumentation, simulation, modeling, and visualization. These materials are discussed in other studies and/or were considered to have a marginal relationship with the present research.⁹

The material covered in the literature reviewed ranges from observation and reportage to analysis and theory, among many other research methodologies. Of the 382 studies examined, 58 were quantitative. Summary information about the characteristics of these quantitative

studies is provided in appendix B.¹⁰ In some cases, quantitative data may be forthcoming, but the results have not yet been fully reported.

STUDY METHODOLOGY

The research conducted by King, McDonald, and Roderer in their 1980 seminal study of the production, use, and economics of scientific journals in the United States, together with work published in 2000 by Tenopir and King, provides the starting point for this study. To build on and supplement these important investigations, a literature search was undertaken. In conducting this search, the following bibliographic utilities (databases) were consulted:

- the Association for Computing Machinery (ACM) Digital Library, which covers all of the society's publications;
- INSPEC, which is produced by the Institution of Electrical Engineers and covers over 6 million articles published worldwide in physics, electrical engineering, electronics, communications, control engineering, computers and computing, and information technology;
- LexisNexis™, which includes the Lexis database for case law and the Nexis database for business and news;
- OCLC (Online Computer Library Center) FirstSearch, Dissertation Abstracts International database;
- OCLC FirstSearch, Library Literature database, which covers materials on libraries and information science;
- PsycINFO®, which is produced by the American Psychological Association and covers the psychological journal literature from 1887 to the present;
- Socio File, which is produced by SilverPlatter Information and covers sociological abstracts from approximately 2,000 journals;

⁸There exists a debate in the technical community over notions of documents, objects, works, and content. For purposes of this study, these distinctions are not pursued since the core artifact is typically an article or document as it is conventionally understood.

⁹For example, the implications of scientific databases, including access by researchers in emerging nations, are discussed in *Bits of Power: Issues in Global Access to Scientific Data* (Committee on Issues in the Transborder Flow of Scientific Data, U.S. National Committee for CODATA, Commission on Physical Sciences, Mathematics, and Applications, National Research Council, Washington, DC: National Academy Press, 1997); this research was supported by the National Science Foundation and other Federal agencies.

¹⁰A study by Hahn (1998) used interviews as a data collection method but did not subject the responses to quantitative analysis; therefore, this study is not included in appendix B.

- Social Sciences Citation Index, which is produced by the Institute for Scientific Information (ISI), and contains citations to articles from 1,400 international social science journals as well as social science articles from journals in the natural, physical, and biomedical sciences; and
- Web of Science, produced by ISI, which is the Web interface for access to ISI citation products covering over 8,000 international journals in the natural sciences, social sciences, and arts and humanities.

The searches were generally confined to the period 1990 to the present, although the period of interest was pushed back to the late 1970s and 1980s in some cases where results alluded to prior studies.

The bibliographic searches were supplemented by detailed examination of key journals and magazines: *D-Lib Magazine*, *Journal of the American Society for Information Science*, *Journal of Documentation*, *Journal of Electronic Publishing*, *Learned Publishing*, *Journal of Scholarly Publishing*, and the annual reviews of the literature supported by the American Society for Information Science.

Web searches were also conducted, and online bibliographies and lists of relevant sources were reviewed. These included C.J. Armstrong, "Collection Management and Scholarly Electronic Publishing Resource," http://www.i-a-l.co.uk/CM_SEP1.htm (2000); Charles W. Bailey, "Scholarly Electronic Publishing," Version 32, <http://info.lib.uh.edu/sepb/sepb.html> (2000); the PEAK (Pricing Electronic Access to Knowledge) project; Stevan Harnad E-Prints on Interactive Publication; and Hal R. Varian, "The Information Economy; The Economics of the Internet, Information Goods, Intellectual Property and Related Issues," <http://www.sims.berkeley.edu/resources/infoecon/> (1998). Research into two publicly sponsored efforts, the Digital Libraries Initiative in the United States (<http://www.dli2.nsf.gov/>) and eLib: The Electronic Libraries Programme in the United Kingdom (<http://www.ukoln.ac.uk/services/elib/>), resulted in the discovery of substantial studies of electronic publishing. Particularly noteworthy were the SuperJournal (<http://www.superjournal.ac.uk/sj/index.htm>) and Open Journal (<http://journals.ecs.soton.ac.uk/>) projects within the eLib effort, which involved collaboration with commercial and learned society publishers.

These results were extended through serendipitous discoveries in the literature reviewed of further relevant books, articles, research reports, and bibliographies. Particularly helpful in this regard were Arms 2000, especially chapter 2 (although summaries of nearly all early projects in this area of study are distributed in sidebars throughout the book); Computer Science and Telecommunications Board 1998, pp. 240–49; Peek and Pomerantz 1998, pp. 345–56, who synopsise numerous early projects in their tables 1 and 2; Schauder 1994, pp. 96–100; and Tenopir and King 2000, pp. 403–63.

ISSUES OF GENERALIZABILITY AND COMPARABILITY

From a methodological perspective, the literature on the implications of information technology for scientific journal publishing is particularly interesting because it is interdisciplinary. Drawing generalizations from and about this material, however—particularly about beliefs and behavior—is challenging. For one thing, rapid technological change complicates the studies and the analysis. In some cases, such as the 1996 TULIP study by Borghuis et al. and Trolley's 1998 ISI electronic library project, the technology overtook the research design before the research and analysis had been concluded. How the combination of technological experimentation and change fused with studies of scientific communication and scientific journal publishing is discussed in the next section.

Looking specifically at the 58 quantitative studies (see appendix B), other analytic and data issues arise. The studies differ widely in terms of sample size (ranging from 3 to 14,368), unit of analysis (individuals, articles, journals, institutions), and research design (e.g., variables studied, questions asked, definitions used). For example, some studies looked at field specialties as defined at the professional or departmental level (physics, computer science, etc.). Others discriminated within fields (experimental high energy particle physics, molecular biology, etc.); still others aggregated specialties (social sciences, life sciences, physical sciences). This disparity across studies leads to a lack of comparability: results from a study that defines "physical scientists" as including the range from theoretical physics through mechanical engineering may be difficult to compare with those from one that looks at theoretical physicists only—particularly when differences in sampling strategies, definitions, and research questions or hypotheses also exist.

Only a handful of the studies examined represent formal surveys based on relatively broad mailings and achieving response rates of generally better than 10 percent. In the main, these survey studies report the results of surveys of relatively small and select populations. Although such small-scale studies may not always meet the requirements of rigorous statistical sampling methodologies, they do capture the reactions of a community to evolving technologies and opportunities. On the other hand, they raise questions as to how far their results can be generalized, and whether they can be replicated in other comparable environments to confirm findings or to elicit variation by designing the study to isolate one or more variables.

The structure for understanding the problem is itself an area of study. Savolainen (1998) reviews the basic approaches to studying users of electronic networks and concludes that, like an earlier generation of studies in the 1980s, studies in the 1990s are still dominated by

system-centered considerations and by university and business contexts: “non-work use[s] have not been given equal attention” (p. 333).¹¹ Computer and information scientists employ user studies to assess user interface design and, more generally, to investigate how a system performs or is likely to perform. Methods include observation, interviews, and analyses of log files. Combined with social science techniques such as questionnaires, surveys, and focus group interviews, these kinds of studies answer three broad questions: (1) what do users want from the system?, (2) what do users do?, and (3) how is the material used?

In sum, research about the implications of scientific journal publishing is in its early stages. Ongoing, projected, and as-yet-unforeseen technological advances and institutional adaptations complicate the field, and their impact is not yet at all clear. Emerging trends and issues, as highlighted in the next sections, can be identified, but definitive empirical results are not yet available.

¹¹The University of California at Los Angeles Internet Project, a longitudinal study with multiple corporate and international partners, which is partially funded by the National Science Foundation, is in the process of collecting and analyzing this type of information. The project’s earliest results were released in November 2000. See <http://www.ccp.ucla.edu>.