

# **ELECTRONIC SCIENTIFIC, TECHNICAL, AND MEDICAL JOURNAL PUBLISHING AND ITS IMPLICATIONS**

REPORT OF A SYMPOSIUM

Committee on Electronic Scientific, Technical, and  
Medical Journal Publishing

Committee on Science, Engineering, and Public Policy  
Policy and Global Affairs Division

**THE NATIONAL ACADEMIES**

THE NATIONAL ACADEMIES PRESS

Washington, DC

**[www.nap.edu](http://www.nap.edu)**

**THE NATIONAL ACADEMIES PRESS • 500 Fifth Street, NW • Washington, DC 20001**

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

Support for this project was provided by the National Academy of Sciences through an unnumbered internal grant. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and symposium speakers and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number 0-309-09161-6 (Book)

International Standard Book Number 0-309-53061-X (PDF)

Committee on Science, Engineering, and Public Policy, 500 Fifth Street, NW, Washington, DC 20001; 202-334-2424; Internet, <http://www7.nationalacademies.org/cosepup/>

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, <http://www.nap.edu>

Copyright 2004 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

# THE NATIONAL ACADEMIES

## *Advisers to the Nation on Science, Engineering, and Medicine*

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

**[www.national-academies.org](http://www.national-academies.org)**

**STEERING COMMITTEE ON ELECTRONIC SCIENTIFIC,  
TECHNICAL, AND MEDICAL JOURNAL PUBLISHING AND  
ITS IMPLICATIONS**

EDWARD H. SHORTLIFFE (*Chair*), Columbia University Medical  
Center, Columbia University

DANIEL E. ATKINS, University of Michigan

FLOYD BLOOM, The Scripps Research Institute

JANE GINSBURG, Columbia University School of Law

CLIFFORD LYNCH, The Coalition for Networked Information

JEFFREY MACKIE-MASON, University of Michigan

ANN OKERSON, Yale University

MARY WALTHAM, Publishing Consultant

*Principal Project Staff*

Paul Uhler, Project Director

Alan Inouye, Senior Program Officer

Julie Esanu, Program Officer

Robin Schoen, Program Officer

Kevin Rowan, Project Associate

Amy Franklin, Senior Program Assistant

**COMMITTEE ON SCIENCE, ENGINEERING, AND PUBLIC  
POLICY**

MAXINE F. SINGER (*Chair*), Carnegie Institution of Washington  
R. JAMES COOK, Washington State University  
HAILE T. DEBAS, University of California, San Francisco  
MARYE ANNE FOX, North Carolina State University  
ELSA M. GARMIRE, Dartmouth College  
MARY-CLAIRE KING, University of Washington  
W. CARL LINEBERGER, University of Colorado  
ANNE C. PETERSEN, W. K. Kellogg Foundation  
CECIL B. PICKETT, Schering-Plough Research Institute  
GERALD M. RUBIN, Howard Hughes Medical Institute  
EDWARD H. SHORTLIFFE, Columbia University Medical Center,  
Columbia University  
HUGO F. SONNENSCHN, The University of Chicago  
IRVING L. WEISSMAN, Stanford University, School of Medicine  
SHEILA WIDNALL, Massachusetts Institute of Technology  
MARY LOU ZOBACK, U.S. Geological Survey

*Staff*

Richard E. Bissell, Executive Director  
Deborah D. Stine, Associate Director  
Marion E. Ramsey, Administrative Associate



## Preface

The use of the Internet and other digital information technologies by the scientific, technical, and medical (STM) research community in the United States and most other countries has transformed many aspects of the research and publishing process. The new technologies have created fundamental changes in the production, management, dissemination, and use of all types of information. It is now possible to communicate research results much more quickly, broadly, and openly than was possible through traditional print publications in the past. Researchers are now able to make available independently their data and articles online, where the information may be easily found, browsed, annotated, critiqued, downloaded, and freely shared. This is resulting in significant changes to the linear path of writing, refereeing, and reviewing of publications as all these functions can be performed concurrently. Most STM publishers also now publish electronic versions of their journals, some exclusively so. The technological developments and resulting changes to the sociology of science are creating both opportunities and challenges for the effective management of scientific communication generally, and STM publishing more specifically.

Because of the far-reaching implications of these developments, the National Academy of Sciences Council's Committee on Publications recommended that the Council commission a study of the factors involved in the changing mechanisms for access to STM information in the scholarly publications and the various technical, legal, policy, and economic issues that they raise. The committee indicated that it is imperative for the National Academies to address, in particular, the increasing concerns about the

implications of various models for access to STM publications for the scientific community.

As a result, the Committee on Science, Engineering, and Public Policy was asked to appoint a committee to oversee the planning for the *Symposium on Electronic Scientific, Technical, and Medical Journal Publishing and Its Implications*, which was held May 19-20, 2003, at the National Academy of Sciences in Washington, D.C. The symposium brought together experts in STM publishing, both producers and users of these publications, to: (1) identify the recent technical changes in publishing, and other factors, that influence the decisions of journal publishers to produce journals electronically; (2) identify the needs of the scientific, engineering, and medical community as users of journals, whether electronic or printed; (3) discuss the responses of not-for-profit and commercial STM publishers and of other stakeholders in the STM community to the opportunities and challenges posed by the shift to electronic publishing; and (4) examine the spectrum of proposals that has been put forth to respond to the needs of users as the publishing industry shifts to electronic information production and dissemination.

The symposium was divided into six sessions, each introduced by opening comments from a moderator, followed by several invited presentations. Session 1 examined the costs involved with the publication of STM journals while Session 2 looked at the related publication business models. Session 3 explored the legal issues in the production and dissemination of these journals. Sessions 4 and 5 looked toward the future and examined, respectively, what is publication in the future and what constitutes a publication in the digital environment. The final session provided several commentaries on the presentations and discussions that took place during the symposium.

The proceedings of the symposium were taped and transcribed, and served as the basis for this symposium report. The formal *Proceedings* of the symposium are available online via the National Academies Press. The National Academies hosted a live audio webcast of the symposium to reach a broad audience and receive additional input. The webcast, along with the edited *Proceedings* of the symposium, can be found on the symposium Web site at <http://www7.nationalacademies.org/cosepup/E-Publishing.html>.

This report is based on excerpts from the symposium *Proceedings* that the committee found particularly useful to highlight. It summarizes the views of the symposium participants but does not contain any consensus findings, conclusions, or recommendations of the committee itself. A footnote in each major section of the report identifies the individuals responsible for the views presented there. In addition, Chapters 3-7 each have a final section entitled "Is-

sues Raised in the Discussion” that summarizes portions of the general discussion of the expert invited panel speakers among themselves and with the audience. Because of the large number of speakers providing comments in those discussions, we have not attributed each point to specific individuals. However, the source of each point may be found in the edited online *Proceedings*.

The report does not cover all the issues that a more comprehensive study might consider, since it is limited only to the topics discussed during the symposium. Several important issues were identified in subsequent discussions, including: the relationship of journal size (number of “pages,” articles, issues) to total annual cost; the “first copy” costs and the relationship of those costs to income sources such as subscriptions or payments from aggregators; and federal policies on the publication of articles in STM journals by recipients of federal research support.

Nor was it possible to involve representatives of all types of STM journals or of all functions in the broader process of scientific communication. The comments and suggestions made by the various participants cannot be generalized easily across the entire spectrum of publications in the STM journal enterprise. For example, significant differences exist between for-profit and not-for-profit journals, between those that represent professional societies and those that do not, between journals that derive revenues from advertisements and those that do not, and between clinical and basic science journals. With regard to the latter, for instance, editors of clinical journals caution that they have a special responsibility in quality control and review. Whereas readers of basic science journals are generally scientists who can critique the articles they read, those who turn to clinical journals for current knowledge frequently may not be experts in the research disciplines covered and have come to depend on the editorial process to assure the accuracy and validity of the papers that they read. The quality and safety of patient care may thus be tied rather directly to the quality of the editorial process in clinical journals.

Despite these acknowledged limitations of the symposium discussions and the resulting report, the committee believes that the material presented here will help identify specific areas for more in-depth inquiry by all the stakeholders in STM journal publishing and in scientific communication more generally.

Edward Shortliffe  
*Committee Chair*  
Paul Uhlir  
*Project Director*



## Acknowledgments

The Committee on Electronic Scientific, Technical, and Medical Journal Publishing and Its Implications would like to thank the following individuals (in alphabetical order) who made presentations during the symposium (see Appendix A for the final symposium agenda): Hal Abelson, Massachusetts Institute of Technology (MIT); Bruce Alberts, National Academy of Sciences; Kent Anderson, *New England Journal of Medicine*; Malcolm Beasley, Stanford University; Robert Bovenschulte, American Chemical Society; Monica Bradford, *Science*; Patrick Brown, Stanford University; Brian Crawford, John Wiley & Sons; James Duderstadt, University of Michigan; Joseph Esposito, SRI Consulting; Michael Jensen, Harvard Business School; Michael Keller, HighWire Press; David Lipman, National Center for Biotechnology Information; Wendy Lougee, University of Minnesota; Richard Luce, Los Alamos National Laboratory; James O'Donnell, Georgetown University; Paul Resnick, University of Michigan; Bernard Rous, Association for Computing Machinery; Alex Szalay, Johns Hopkins University; Gordon Tibbitts, Blackwell Publishing USA; and Ann Wolpert, MIT.

The committee also would like to express its gratitude to the guidance group for this project, which was formed under the Committee on Science, Engineering, and Public Policy. Members of that group included, James Cook, Washington State University; Paul Torgerson, Virginia Polytechnic Institute and State University (retired); and Edward Shortliffe, Columbia University Medical Center, Columbia University.

This volume has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the NRC's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity and evidence. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We wish to thank the following individuals for their review of this report: Martin Blume, American Psychological Society; Karen Hunter, Elsevier Health Services; Justin Hughes, Cardozo Law School; James Neal, Columbia University; Andrew Odylzko, University of Minnesota; and Carol Tenopir, University of Tennessee at Knoxville.

Although the reviewers listed above have provided constructive comments and suggestions, they were not asked to endorse the content of the report, nor did they see the final draft before its release. The review of this report was carried out in accordance with institutional procedures and all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Finally, the committee would like to recognize the contributions of the following National Research Council staff. Paul Uhlir, director of the Office of International Scientific and Technical Information Programs, was the project director for the symposium and principal editor of the committee's report; Julie Esanu, program officer for the Office of International Scientific and Technical Information Programs, helped organize the symposium and edit the report; Alan Inouye, interim director of the Computer Science and Telecommunications Board, and Robin Schoen, program officer for the Board on Life Sciences, provided advice on the project; Kevin Rowan, project associate for the Committee on Science, Engineering, and Public Policy, provided project support for the May symposium; and Amy Franklin, senior program assistant for the Board on International Scientific Organizations, assisted with the production of this report.

# Contents

Executive Summary	1
1. Introduction	7
2. Costs of Publication	9
3. Publication Business Models and Revenue	20
4. Legal Issues in the Production, Dissemination, and Use of the Journal Literature	40
5. What Is Publishing in the Future?	48
6. What Constitutes a Publication in the Digital Environment?	56
7. Symposium Wrap-Up	67
Appendixes	
A Symposium Agenda	75
B Biographical Information for Speakers and Steering Committee Members	80
C Symposium Participants	94



# Executive Summary

The Symposium on Electronic Scientific, Technical, and Medical (STM) Journals and Its Implications addressed these issues in five key areas. The first two—costs of publication and publication business models and revenue—focused on the STM publishing enterprise as it exists today and, in particular, how it has evolved since the advent of electronic publishing. This was followed by a review of copyright and licensing issues of concern to the authors and to universities. The final two sessions looked toward the future, specifically, at what publishing may be in the future and what constitutes a publication in the digital environment.

## **COSTS OF PUBLICATION**

The initial session identified the main elements of expenses (both print and electronic) for many STM journal publications. It was quite clear from the presentations that understanding all the cost elements in the budgets of different publishers is not straightforward and that a comparative analysis, in particular, would be very difficult to do well. At the same time, some speakers indicated that such a study, if done carefully and authoritatively, could be very useful in adding some rigor to the ongoing debate about the high cost of journal subscriptions and the value that publishers bring to the process of scientific communication.

Other cost-related issues that were discussed included the creation and operation of digital archives and the digital conversion of back sets; the

costs of new technology and related cost-containment strategies and the difficulty of moving from print to electronic-only versions; strengths and weaknesses of the peer-review process; cost issues specific to small and mid-sized societies; and the vulnerability of secondary and tertiary publishers.

## PUBLICATION BUSINESS MODELS AND REVENUE

This, understandably, was the most contentious part of the symposium. A number of trends in commercial STM journal publishing were described at the outset. These include the bundling of publications by major publishers in “big deals”; the consolidation of publishers and the targeting of downstream competitors (secondary publishers and subscription agencies) and of vulnerable competitors; diversification of the customer base to more business clients (and a concomitant emphasis on applied research and engineering journals); and market responses to open-access trends, including the creation of meta-content (e.g., documentation and search engines for the open-content resources) and a shift to Web services (e.g., substitutes for the publication of fixed content in print by providing online software, processing, and services for users).

This discussion of commercial publishing trends was followed by the perspectives of a university librarian, a commercial journal publisher from John Wiley & Sons, Inc., and an open-access representative from the Public Library of Science (PLOS). The library overview included some statistics and anecdotal information about the responses of libraries to rising subscription costs and the bundling efforts of commercial publishers; the implications for libraries of changes in electronic journal formats and content, and of digital scientific communications more broadly; and the changing role and influence of libraries in the digital publication context.

The description of the commercial subscription-based model used at John Wiley & Sons highlighted the benefits to readers from this traditional approach and the reasons why Wiley would not switch to the author-pays, open-access model.<sup>1</sup> This was followed by the description of the PLOS

---

<sup>1</sup>According to the definition presented by the Public Library of Science later in this report, an open-access publication is one that meets two conditions. The first is that the copyright holder (either the author or the publisher, if the copyright has been transferred to the publisher) grants to the public a free, irrevocable, perpetual right of access to, and a license to copy, distribute, perform, and display the work, and to make and distribute deriva-

model, including the rationale for reconceptualizing the STM publishing business model on the Internet, the definition of “open access” used by the PLoS, and the advantages of this approach for science.

A vigorous and informative discussion ensued among the panelists and with the audience about the relative merits of the traditional user-pays publishing model versus the author-pays PLoS model. Other related issues were discussed, including the practical difficulties in transitioning to the open-access publishing model, support of the open-access publishing approach by research sponsors, the effect of different publishing business models on the long-term preservation of digital journals, and advertising revenues in electronic publishing generally.

The results of these discussions appeared to be inconclusive. On the one hand, commercial (and professional society) publishers clearly add considerable value to the process of formal scientific communication, and the viability of the author-pays, PLoS type of open-access model is still untested and its future success uncertain. On the other hand, the restricted, subscription-based model clearly has great inherent social costs in comparison with the immediate, free access by any and all users of the information worldwide that the open-access publishing model makes possible. Many participants believed that we are in a period of important experimentation, in which the open-access approach will be tested and refined and in which the traditional publishers will try new approaches and attempt to add more value to STM products and services. A greater differentiation between the practices of commercial and society publishers also may be expected. For example, there are hybrid approaches being implemented in the traditional subscription-based publishing community, mostly by the not-for-profit STM publishers. Moreover, there are other open-access approaches such as self-archiving by authors and open institutional repositories, as discussed in other sessions of the symposium.

---

tive works in any medium for any purpose. The second condition is providing readers with open access to the work. Authors or publishers achieve open access by making a complete version of the article and all supplemental materials available in some suitable standard electronic format, deposited immediately upon publication in at least one internationally recognized, independent online repository that is committed to open access.

## LEGAL ISSUES IN PRODUCTION, DISSEMINATION, AND USE

The focus in this session was on copyright and on licensing issues in the traditional publishing business model. With regard to copyright, there are divergent practices at universities as to whether the university or the author owns the copyright to publications, and in the various derogations from those rights. The question of transfer of copyrights from the author to the publisher, and the limited rights granted back by the publisher to the author, was discussed as well.

The licensing issues pertain to the terms and conditions that publishers and libraries negotiate for site licenses, and to licenses between authors and publishers. The libraries have continued to experience two significant problems in negotiating site licenses, both related to the overall problem of access restrictions: the various limitations and prohibitions on “interlibrary loans” of electronic copies and the problem of long-term preservation of material that is electronic-only under the restrictive licensing regime. Both of these factors inhibit libraries from switching to electronic-only subscriptions. As to the licenses between authors and publishers, two models were suggested that serve the author’s interests better: either retain copyright, while licensing the publisher to deploy the work in all ways that the publisher needs for effective publication and dissemination, or transfer copyright to the publisher, with more rights reserved to the author, such as permission to redistribute the work.

Issues raised in the general discussion included a description of additional problems with the transfer of copyrights by authors in universities (i.e., the author may not own the copyright under the university policy but may not know that, thereby signing void copyright transfer agreements) and significant problems associated with university work-for-hire approaches to academic publications. The burdens for small publishers in developing countries from licensing practices and from restricting access were also discussed.

## WHAT IS PUBLISHING IN THE FUTURE?

The final two sessions of the symposium looked more toward the future, by identifying some of the technology-enabled trends, processes, and projects that are indicative of what may be possible and what may perhaps become more widely adopted. As was noted in the introductory comments for this session, it is quite clear that the digital revolution is changing the

traditional processes of many knowledge-intensive activities, in particular STM publishing and scholarly communication more generally. The various functions—whether metadata creation, credentialing review, or long-term stewardship—can be separated or disaggregated, and players different from those who traditionally have carried out these tasks can, in theory, perform them. Publications can now exist in many intermediate forms, and we are moving toward more of a continuous-flow model, rather than a discrete-batch model. The raw ingredients—the data, the computational models, the outputs of instruments, the records of deliberation—can be online and accessible by others and can be used to validate or reproduce results at a deeper level than traditionally has been possible. Third parties—particularly in an open-access, open-archives context—can then add value by harvesting, mining, enriching, and linking selected content from such collections.

The presentations in this session of the symposium identified some of the social processes, specific pilot projects, and the challenges and opportunities that may provide the basis for future “publishing processes,” which ultimately may be more holistically integrated into the “knowledge creation process.” For example, there are emerging open recommender and reputation systems that use the online environment to get broader public feedback and to develop new indicators of user behavior. Although there are potential problems, such as “gaming” the system, eliciting early evaluations, and “herding,” whereby later evaluators are unduly influenced by previous evaluators, there are some experiments that could be tried in the STM publishing context. Preprint servers, such as the well-known e-Print arXiv established initially for the high-energy physics community by Paul Ginsparg at the Los Alamos National Laboratory, have now been adopted by some other fields. Open institutional repositories, such as the MIT OpenCourseWare project and the MIT-led DSpace consortium for the deposit of various types of research materials in an openly accessible archive, provide additional examples of innovative projects and models.

Despite the exciting possibilities raised by these different initiatives, the subsequent discussion raised a number of potential difficulties in implementing some of these new approaches. Some of the issues raised included caution about an over-reliance on statistical indicators or metrics in judging the quality of information or of publishing activities; the relative merits of the traditional, confidential peer-review process and a more open review system; the constraints of the discrete-batch, print model on adopting some of the more

open, continuous-flow processes; and the inherent tensions between the creation of various open archives and the traditional publishers.

### WHAT CONSTITUTES A PUBLICATION IN THE DIGITAL ENVIRONMENT?

The final session built on the technology-oriented concepts and processes introduced in the prior session. Publication used to refer to the act of preparing and issuing a document for public distribution. It could also refer to the act of bringing a document to the public's attention. Now, publication means much more. It can refer to a document that is Web-enriched, with links, search capabilities, and potentially other services nested in it. A publication now generates usage data and provides many other functions. This session examined three innovative examples—the Signal Transduction Knowledge Environment of *Science*, the publishing of very large data sets in astronomy on Web sites and through the International Virtual Observatory initiative, and genomic data curation at the National Center for Biotechnology Information and the integration of those data with the scientific literature.

Issues raised in the discussion included the need for federal coordination and investment in the cyberinfrastructure to maximize the opportunities for information integration and knowledge creation, some of which were identified in the final two sessions; the difficulties of quality control and review of data in very large or complex databases, particularly in the biological sciences; restrictions on data mining in proprietary STM information that is based on publicly funded research; the opportunities for knowledge discovery from the open publication of large and complex data sets; the transformation of the archiving function in the knowledge discovery process; lost opportunities associated with insufficient people and resources focused on the avalanche of data in all disciplines; and the positive role of the journal publishers in the successful development of databases in molecular biology.