Order Code RL33023

Open Access Publishing and Citation Archives: Background and Controversy

July 26, 2005

Genevieve J. Knezo
Specialist, Science and Technology
Resources, Science, and Industry Division

Open Access Publishing and Citation Archives: Background and Controversy

Summary

Controversies about open access publishing and archiving in science and technology confront issues of copyright, governmental competition with the private sector, and impacts upon the pace of scientific research. Different from traditional publishing which requires fee-based subscriptions, most open access systems give readers free online access to bibliographic citations or the full text of published and non-published literature. There are also free, online, peer reviewed open access scientific journals, sponsored by government or by nongovernmental publishers. Some systems provide access to information in specific scientific and technical fields, or at academic institutions, academic consortiums' websites, or on author's websites. Open access publishing is estimated to constitute between 1% and 2% of the scientific journal publishing market, which is estimated at $9 billion, annually.

Support for the open access "movement" stems from some librarians' and scientists' objections to rising costs of subscriptions to journals; peer reviewers' objections to providing free reviews for journals rapidly escalating in price; rapid technological developments in computing capabilities; and the belief that scientific collaboration, advancement, and utilization will be hastened by free access to citations and articles. Traditional commercial publishers and some scholarly scientific associations object to most open access efforts on the grounds that they duplicate what publishers sell and erode profits. Some critics seek to limit free government-run, citation repositories to include only federally sponsored research results. Some oppose open access publishers' requirements that charge authors fees in the thousands of dollars to pay the costs of publishing articles. Others say that foundation donations that sustain some open access activities are unreliable.

In 2003, a bill was introduced to encourage federal agencies to provide free access to published results of all federally funded basic scientific research. In 2004, congressional report language mandated that authors funded with National Institutes of Health (NIH) research and development (R&D) support voluntarily submit within 12 months of publication, copies of their journal articles to NIH's free access database, *PubMed Central*. Many publishers opposed this policy. In 2005, congressional report language on H.R. 3010 endorsed NIH's new policy to archive journal articles and mandated NIH to work with commercial publishers in expanding its open access repository, *PubChem*, to avoid duplication with private efforts.

Controversial issues that could draw congressional attention include modifying NIH's *Public Access* policy to require that, instead of allowing readers direct access to a published article, the government provide links to the original journal's website to allow publishers to charge fees; limiting federal systems to information derived from federally funded R&D; monitoring the added costs of expanding *PubMed Central*; determining if other agencies will use governmental nonexclusive licensing to allow access to federally funded, commercially published journal articles regardless of copyright ownership by a publisher; assessing the quality of science published in open access journals; and evaluating the economic impacts of open access publishing on traditional publishing. This report will be updated as needed.

Contents

Open Access Publishing and Citation Archives: Background and Controversy

Introduction

This report begins with an inventory of basic information: definitions and guides to histories of the growth of open access publishing and citation archives and descriptions of selected major open access activities. It moves on to summarize major points of difference between proponents and opponents of nongovernmental open access publishing and databases, and then highlights federal, including National Institutes of Health (NIH), open access activities and contentious issues surrounding these developments. The report also briefly describes open access developments in the United Kingdom (where a number of governmental and nongovernmental initiatives have occurred) and in the international arena. Finally, controversial issues which could receive attention in the 109th Congress are summarized.

Definitions of Open Access Publishing and Database Models

The "open access movement" is said to have begun in 1966.[1] The term describes a variety of activities that includes access to archives of indexed citations of articles, access to separate journal articles that were published in traditional journals, and access to free, online journals.[2]

---

[1] With the inception of Educational Resources Information Center (ERIC), launched by the U.S. Department of Education's Office of Educational Research and Improvement and the National Library of Education. This database contains bibliographic citations for privately published journal articles and allows retrieval of the text of other nonpublished materials. *Medline*, a bibliographic system, was launched by the National Library of Medicine in 1966 (but not free until 1997). Source: "Timeline of the Open Access Movement," by Peter Suber, last revised Apr. 13, 2005, at [http://www.earlham.edu/~peters/fos/timeline.htm]. This is an extensive history since 1966, with hotlinks to different systems and databases.

[2] See the following information about open access publishing: Martin Frank, Margaret Reich, and Alice Ra'anan, "A Not-For-Profit Publisher's Perspective on Open Access, as it was planned to be published in *Serials Review*, vol. 30, no. 4, 2004; "Budapest Open Access Initiative," available at [http://www.soros.org/openaccess/read.shtml]. See also Peter Suber, "What You Can Do to Promote Open Access," Last revised April 5, 2005, 11 p. [http://www.earlham.edu/~peters/fos/do.htm]; "Budapest Open Access Initiative: Frequently Asked Questions," last revised March 27, 2005 [http://www.earlham.edu/~peters/fos/boaifaq. htm]. Several open access online journals and

(continued...)

In *traditional publishing*, the publisher, who holds the copyright to an article, pays most printing and distribution costs and, in order to read an article, the journal subscriber pays subscriber fees, whether for hard-copy or online versions. Sometimes an author is required to pay printing page charges for complex graphics or color presentations.

"Open access" publishing generally means that the author or publisher, who holds the copyright to an article, grants all users unlimited, free access to, and license to copy and distribute, a work published in an open access journal (which may be published initially electronically or in hardcopy). Users can also make copies for their personal use, if authorship is properly attributed. Open access publishing often requires an author to pay for publishing or posting of a paper (fees range from about $1,500 to $4,000). Generally, open access publishers require that a complete version of the work and related materials be deposited electronically in an online database that permits open access, distribution, interoperability (allowing users to extract and use the data in other research), and long-term archiving.[3]

In *"free access" publishing* neither an author nor a reader pays for articles to be published or posted on the Internet,[4] but other open access features may not be mandatory.

A few commercial publishers have adopted some open access features in their business models. However, the fundamental difference is that traditional publishers generally require readers to pay to read or print an article, or to search indexes of abstracts or citations. Open access publishers generally do not require readers to pay for these services. Some traditional publishers say they already provide open access in that they may make papers freely available online — but this is usually a year or two after publication. However, they still hold copyright, and they may or may not allow the author to post his or her published articles in an open access repository or database, or on the author's own website.

The scope of *open access repositories or archives* varies. Some contain published journal articles or nonpublished "grey literature" in all fields of science or in specific scientific disciplines. Some archive a specific university's researchers'

---

[2] (...continued)
indexes of collections of these are available. For instance, see *Open Access Bibliography: Liberating Scholarly Literature with E-prints and Open Access Journals*, information is available at [http://www.escholarlypub.com/oab/oab.htm]; "List Issues: Open Access (Journal) Collections: Electronic Resources in Libraries," [http://www.joanconger.net/ERIL/list_issues_openaccess. html]; *SCIELO*, available at [http://scielo.org]; *HighWire Press* [http://highwire.stanford.edu/lists/freeart.dtl], and PubMed Central, at [http://pubmedcentral.com].

[3] Based on "Definition of Open Access," which uses a modified version of the "Bethesda Meeting on Open Access," [http://www.plos.org/about/openaccess.html]. See "Open-Access Publication of Medical and Scientific Research," a Public Library of Science Background Paper, Dec. 12, 2003.

[4] Joanne S. Hawana, "Multiple Publishing Models Critical To Advancing Science, Journal Publishing Societies Argue," *Washington Fax*, Mar. 17, 2004.

preprints, articles, or research reports; or, as in the case of the National Institutes of Health model, published journal articles or other materials funded by a particular federal agency or discipline. Some open access repositories archive only citations for articles or other materials; some archive both citations and full text materials; some allow free downloading and some do not.

Selected Illustrations of Nongovernmental Open Access Activities

A variety of nongovernmental open access publishing activities is illustrated next with summaries of some current major open access information systems or publishers. These are categorized by general type, including commercial open access systems academic-sponsored systems; and subject or disciplinary systems. NIH's *PubMed Central* system is described in detail in the section of this report that focuses on NIH.

Open Access Journal Publishers

Public Library of Science (PLoS). *PLoS* is a nonprofit group, spearheaded in large part by Dr. Harold Varmus, former NIH director. It provides readers with free access to peer reviewed articles published in *PLoS*'s electronic journals. The activity is supported by author payments starting at $1,500 per article and multi-million dollar philanthropic foundation contributions. *PLoS*'s journals include *PLoS Biology, PLoS Medicine, PLoS Computational Biology, PLoS Genetics, and PLoS Pathogens.*[5] *PLoS* seeks to launch journals in other disciplines. It has the goal of publishing highly selective, top-quality articles competitive with the quality of articles in traditional journals like *Science* and *Nature*. Different from traditional publishing, which requires authors to cede copyright to the publisher, authors who publish in *PLoS* retain copyright to an article, but are required to deposit a copy of the article in an open access, online repository that allows long-term archiving.[6] Reportedly, one of the group's major goals is to make research more accessible by eliminating publishers as copyright holders and by ending the "balkanization" of scientific information in separate databases. Under *PLoS*'s editorial policy, "any data can be integrated into new work as long as the original author is credited appropriately. The model is inspired by *GenBank*, the central repository of DNA sequence whose open access policy has driven much of the progress in genomics and biotechnology of the last decade."[7] *PLoS* has announced that it will assist scientists

---

[5] Janet Coleman, "Public Library of Science to Launch 3-4 New Open-Access Scientific Journals in 2005,"*Washington Fax*, Oct. 29, 2004.

[6] Information about *PLoS* and related archives is available at [http://www.plos.org/about/openacess.html].

[7] Amy Harmon, "New Premise in Science; Get the Word Out Quickly, Online," Dec. 17, 2002, *New York Times*.

in developing countries by providing Internet access for readers of limited bandwidth, and will waive or defray author charges for those who cannot afford to pay.[8]

BioMedCentral. A British-founded, independent, commercial publishing system, which provides free access to peer reviewed biomedical research published online.[9] It publishes its own approximately 120 biomedical journals and says articles are rapidly peer reviewed; peer review policies are determined by each journal's board. Authors retain copyright of their work. *BioMedCentral* charges authors or their institutions for the costs of peer review and publication. "Other sources of revenue include subscription access to commissioned articles, sales of paper copies of our journals to libraries, sales of reprints, advertising and sponsorship, and ... a range of subscription-based value added services such as literature reviews and evaluation, personalized information services delivered electronically, provision of editorially enhanced databases, tools that help scientists collaborate, and other software research aids."[10] It archives materials in *PubMed Central*, NIH's open access database.

Faculty of 1000. *BioMedCentral* has created a fee-based subscription service called *Faculty of 1000*.[11] It originated because the publication of so many articles in online journals (sometimes free to readers) with varying degrees of peer review has spawned a new industry: peer reviewers or experts who evaluate articles after publication and provide a selected list of articles recommended for reading to their paid subscribers.

Illustrations of Academic-Related Systems

Some universities ensure that their scholars' publications are available online in a free open access repository by creating their own archives or participating in networked open access archives. Several examples are outlined next.

EScholarship Program. The *EScholarship Program* of the University of California system was launched in the fall of 2003. It is an electronic, searchable repository that makes freely available an archive of the publications (and other media) and some research databases of University of California researchers. The vehicle is also used to disseminate the university's own open access, peer reviewed published journals.[12] Supporters of systems like this say that indexing materials improves access to them and, if full text is available, widens reader access, and improves utilization of federally financed research and development.[13]

---

[8] "PLoS, Frequently Asked Questions," available at [http://www.plos.org/faq.html].

[9] Available at [http://www.biomedcentral.com/info/].

[10] Available at [http://www.biomedcentral.com/info/about/whatis].

[11] Available at [http://www.facultyof1000.com/about/key].

[12] Available at [http://escholarship.edlib.org].

[13] "UC to Launch Open-access Journals" *The Scientist*, June 16, 2003.

DSpace. A number of research universities[14] are participating in *DSpace,* a networked multi-member electronic repository that indexes and shares some research data, articles, and other media.[15] It was developed by the Massachusetts Institute of Technology (MIT) in collaboration with Hewlett-Packard. Some universities, such as Cornell, reportedly, are using it to provide free access to peer reviewed publications.

Highwire Press. This is an archive run by Stanford University that provides full-text articles for biomedical and other scientific journals. It adheres to the post-publication timing policies of each journal, with most articles archived and made accessible between 6 and 24 months after publication in the original traditionally published journal. Some of these articles, but not all, may be viewed for free.[16]

Illustrations of Dedicated Subject or Disciplinary Archives

Some repositories permit free searching for citations, abstracts, articles, or other materials in specific disciplinary fields or areas of application, or by researchers affiliated with specific academic systems, or by other researchers. A few illustrations are given next.

arXIV.org.[17] Initiated in 1991, this is an archive for free scientific publishing, which allows physical science researchers to make preprints of their papers available before formal publication. Maintained by the Cornell University Library[18] (in cooperation with the National Science Foundation and the Department of Energy), it includes articles in the following subjects: physics, mathematics, nonlinear sciences, computer science, and quantitative biology. According to *PLoS*, "This server expanded from its initial role as a vehicle for sharing preprints in theoretical

---

[14] Including Brigham Young University; Case Western Reserve University; Chapel Hill School of Information and Library Science Electronic Theses and Dissertations; Cornell University; Digital Repository at the University of Maryland; DLEARN at the University of Arizona; Drexel University; DSpace@Cambridge; DSpace at MIT; DSpace at University of Rochester; Edinburgh Research Archive; Embry-Riddle Aeronautical University; WETD of Indian Institute of Science, Bangalore (etd@IISc); George Mason University; Hong Kong University of Science and Technology; IDeA, Indiana University Purdue University Indiana; Dspace at Indiana University Of Pennsylvania; Kansas State Publications Archival Collection, Kansas State Historical Society and Kansas State Library; KU ScholarWorks; Los Alamos National Laboratory, New Mexico; Portfolio@Duke University; RIT Digital Media Library; SMARTech Scholarly Materials and Research at Georgia Tech; Texas A&M University Libraries Institutional Repository; T-Space at The University of Toronto Libraries; University of New Mexico, DSpaceUNM; University of Oregon Scholars' Bank; University of Tennessee in Knoxville; University of Texas at Austin, School of Information; University of Washington, Seattle; University of Wisconsin; Vanderbilt University e-Archive; Washington University, St. Louis; Woods Hole Open Access Server.

[15] Available at [http://www.dspace.org/].

[16] Available at [http://highwire.stanford.edu/lists/freeart.dtl].

[17] Available at [*arXIV.org*].

[18] "Scientific Publishing: Who Will Pay for Open Access?," *Nature,* Oct. 9, 2003. See also [http://arxiv.org/].

high-energy physics to its current role as the principal 'library' for a large fraction of research literature in physics, computer sciences, astronomy, and many mathematical specialities. Today, more than half of all research articles in physics are posted to this server prior to their publication in conventional journals. In many fields, these "eprints" are the *de facto* publications of record."[19]

CogPrints. Some types of foreign open access publishing include access to U.S.-generated research findings. *CogPrints* is a free, British-run, self-archive of full-text, electronically available, published, peer reviewed journal articles as well as preprints of unrefereed articles in the "cognitive sciences, including any area of psychology, neuroscience, and linguistics; many areas of computer science (e.g., artificial intelligence, robotics, vison, learning, speech, neural networks); philosophy (e.g., mind, language, knowledge, science, logic); biology (e.g., ethology, behavioral ecology, sociobiology, behavior genetics, evolutionary theory); medicine (e.g., psychiatry, neurology, human genetics, imaging); anthropology (e.g., primatology, cognitive ethnology, archeology, paleontology), as well as any other portions of the physical, social and mathematical sciences that are pertinent to the study of cognition."[20]

patientINFORM. In spring 2005 *patientINFORM*[21] was launched by the American Cancer Society, the American Heart Association, and the American Diabetes Association, in partnership with more than 20 publishing firms, to provide immediate access to free, selected full-text research articles and materials from the three organizations' websites, which provide links to different types of published materials. "According to the group, the initiative 'is being driven by recent trends indicating that public awareness of clinical research, heightened by media coverage and fueled by the spread of broadband Internet, has led more and more patients to go online to find the latest information about treatment options.' "[22] NIH's decision to launch its system also, reportedly, accelerated the formation of the system. After a period of evaluation, " ... the group will determine whether to expand its focus past the three initial diseases into rarer conditions."[23]

---

[19] "PLoS History," available at [http://www.publiclibraryofscience.org/about/history.html].

[20] See [http://cogprints.org/].

[21] Information is available at [http://www.patientinform.org/]. Participating publishers and associations include the International Association of Science, Technical and Medical Publishers; the Association of American Publishers/Professional and Scholarly Publishers; Johns Hopkins' Welch Medical Library; and the National Library of Medicine's *MedLinePlus*; the American Association for the Advancement of Science; the American Medical Association; the American Physiological Association; Oxford University Press; Blackwell Publishing; Elsevier Publishing; BMJ Publishing Group; Nature Publishing Group; and Springer and Wiley.

[22] Andrew Hawkins, "Journal Publishers, Advocacy Groups Spearhead New Open Access Initiative," *Washington Fax,* Dec. 13, 2004.

[23] Hawkins, op. cit., Dec. 13, 2004.

Major Issues Relating to Open Access Publishing

Controversies arise because developments in open access systems and policies seem to have outpaced society's ability to design equitable and efficient mechanisms and economic reward structures to manage transitions between traditional and open access publishing and archiving.[24]

Major arguments made by supporters of open access publishing (largely scientists, librarians, and some non-profit publishers) are that it rides the new wave of inevitable changes in publishing and electronic dissemination of information due to development of the Internet,[25] hastens scientific progress, gives access to more readers, promotes economic development, and, in the case of federally funded research, provides citizens with ready access to the results of research and development that their taxes funded.

Opponents of open access publishing (primarily traditional publishers and major scientific associations) cite such issues as the doubtful permanence of electronic archives, questions of copyright ownership and reductions to traditional publishers' profits, costs to researchers who have to pay to have their manuscripts published in open access journals, the possibly dubious quality of articles published, questions about peer review processing and quality, perceptions of the academic community and the academic reward system which appear to give more status to articles published in traditional journals, and so forth.[26] **See Appendix 1** for a list of additional issues raised about the impact of open access publishing on the academic community, scholarship, and teaching.

The following sections elaborate on some of these issues.

Journal Publishing Costs and Sources of Revenue

The costs of publishing a journal article include preparing the manuscript for publication (initial sorting and selection of manuscripts to be refereed, peer review,

---

[24] See, for instance, Julie M. Esanau and Paul F. Uhlir, eds., *Open Access and the Public Domain in Digital Data an Information for Science, Proceedings of an International Symposium,* Published by U.S. National Committee for CODATA, National Academies Press, Washington, D.C., 2004.

[25] David Stern, "Archival Issues Regarding Electronic Scientific Literature," Presentation at session on "The Future of Scientific Communication (Formerly Known as Publishing)," American Association for the Advancement of Science (AAAS) Meeting, Apr. 21, 2005.

[26] According to one author, barriers to open access publishing include legal framework issues; differences in IT-infrastructure and technologies; business models and costs; indexing services and standards of materials placed in open-access archives; the academic reward system; and marketing and critical mass issues. The importance of each type of barrier varies with the type of open access repositories, whether open-access journal, subject-specific repositories maintained by disciplinary groups; or institutional repositories, maintained by academic institutions. The author provides a matrix and specific details for each of the 18 cells in his analysis in: Bo-Christer Bjork, "Open Access to Scientific Publications - An Analysis of the Barriers to Change?", *Information Research,* Jan. 2004.

selection, editing, layout, table of contents, overhead, letters to the editor, etc.) and distribution. According to a *Wall Street Journal* story, costs for publishing an article typically range from $3,000 to $4,000.[27] However, these costs can average more than $10,000 for some journals, such as *Science* magazine, which publishes only a small fraction of the articles submitted (about 7%),[28] but have high value-added costs, which include reviewing all articles submitted and selecting those that will be published, layout, graphics, distribution, and so forth. Another author has estimated costs for publishing an article in other journals: *BioScience*, about $7,000 per article; *Nature* and *New England Journal of Medicine*, in excess of $1,500.[29]

The comparative costs of publishing electronic journals versus traditional journals are uncertain. Some observers say that article processing costs are similar for print and electronic[30] publications, yet at least one researcher's data indicate that electronic publishing and distribution are cheaper than hard copy publishing.[31]

Who Pays?: Traditional Journals. Traditional publishers usually incur most of the costs of publishing an article. Revenue comes from subscriptions, advertising, reprints, and, in some cases, from authors who are asked to subsidize the costs of color printing or printing of complex graphics, and, in some cases, page charges for publishing articles in traditional hardcopy journals. Preliminary data from an ongoing study which is surveying sources of revenue for traditional and open access publishers indicates that the three largest sources of revenue for traditional journal publishers are subscriptions, which provided 67% of total revenue; industry support (advertising and sponsorship) at 12%; and author fees and charges, at 9%.[32]

Who Pays?: Open Access Journals. Reportedly, most, but not all, open access journals require authors to pay from $1,500 to $4,000 for publishing costs. Open access journals also receive funds from advertising, corporate sponsorships, government grants, the use of volunteers, and, foundation grants.[33] The ongoing

---

[27] Bernard Wysocki, Jr., "Peer Pressure: Scholarly Journals' Premier Status is Diluted by Web," *Wall Street Journal*, May 23, 2005, p. A1.

[28] Wysocki, May 23, 2005, op. cit.

[29] David Malakoff, "Opening the Books on Open Access," *Science*, Oct. 24, 2003, p. 551.

[30] Donald W. King, "The Economics of Science Publishing," Presentation at Session on "The Future of Scientific Communication (Formerly Known as Publishing)," American Association for the Advancement of Science (AAAS) Meeting, Apr. 21, 2005.

[31] King, op. cit., Apr. 21, 2005.

[32] "Variations on Open Access: A Study of the Impact of Alternative Business Models on Financial and Non-financial Aspects of Scholarly Journals," Preliminary Results Presented 14 March 2005, London Book Fair by Kaufman-Wills Group, LLC. The study called traditional publishers "Delayed OA" publishers.

[33] PLoS's webpage includes the following information: "PLoS is a tax-exempt, 501(c)3, nonprofit corporation headquartered in San Francisco, California (Federal Tax ID 68-0492065). PLoS is governed by an eleven-member Board of Directors. PLoS co-founder Harold Varmus is Chairman of the Board. *PLoS* has received financial support in the form of grants from the Gordon and Betty Moore Foundation, the Sandler Family Supporting

(continued...)

study cited above identified the three largest sources of revenue for open access journal publishers as industry support (advertising and sponsorship) at 37%; author fees and charges at 30%; and grants, at 13%.[34]

This same study showed that "contrary to expectations, author fees were charged by a larger fraction of traditional journals than open-access journals."[35] Author fees include charges for color printing, page layout, page publication charges, and so forth. This finding, in combination with the data on percentage sources of revenue, appears to mean that in relation to the total number of publishers, more traditional publishers than open access publishers charged fees to authors, but the payments (as a percentage of publishers' total revenue) were less to traditional publishers than to open access publishers. The fees traditional publishers charged to authors were primarily for small changes, color views, and related items, rather than the larger fees open access journals charge authors to publish in the open access journal.[36]

Federal Policies For Paying Publication Costs in Relation to the Future of Open Access Publishing. Among the issues related to "author pays," and possibly to the future of open access journals, is whether the federal government will allow part of research grant funding to be used to pay charges levied on authors or institutions for the costs of publishing articles resulting from federally funded research. This may become a more prominent issue if open access publishing becomes a larger part of the market.

Now, pursuant to OMB's guidelines, federal agencies that award funds for scientific research permit investigators at universities, colleges, and nonprofit institutions to charge the costs of publishing a scientific article as an allowable direct cost (usually paid in full) if the funding agency agrees that they are an appropriate part of the project. If the costs of publishing are disallowed as direct costs, the federal governments likely will pay for these costs as part of "facilities and

---

[33] (...continued)
Foundation, the Irving A. Hansen Memorial Foundation, the Open Society Institute (OSI), and the Joint Information Systems Committee (JISC). *PLoS* also receives support through donations, sponsorships, and memberships from private citizens, universities, and other organizations" [http://www.plos.org/about/index.html]. It reported that it received a $9 million grant from the Moore Foundation to start operations for four years [http://www.plos.org/about/index.html].

[34] "Variations on Open Access," op. cit. and interviews with an author of the study, July 2005, who agreed with this interpretation of the data.

[35] Lila Guterman, "New Study Compares Open-Access and Traditional Publishing," *Chronicle of Higher Education,* Mar. 25, 2005. "The survey was conducted by the Kaufman-Wills Group, publishing consultants based in Baltimore. It was financed by groups that are affiliated largely with traditional journals: the American Association for the Advancement of Science, the Association of American Medical Colleges, the Association of Learned and Professional Publishers, and HighWire Press, which produces online versions of journals and is operated by Stanford University." An author of the study, "Variations on Open Access," op. cit., agreed with this interpretation of the data. (Interview held July 2005.)

[36] See also "Variations on Open Access," op. cit.

administrative" (F&A) indirect costs, if the research was federally sponsored and if the journal levies similar charges on all research papers published by the journal.[37] If the cost is covered as an F&A indirect cost, full reimbursement may not occur due to limitations on recoveries of some indirect costs.

Some federal agencies have issued policy guidance about allowing as a direct cost of project support, fees for publication and page charges in order to disseminate reports of the agency's federally funded research results. The National Science Foundation (NSF), for instance, says,

> The proposal budget may request funds for the costs of documenting, preparing, publishing or otherwise making available to others the findings and products of the work conducted under the grant. This generally includes the following types of activities: reports, reprints, page charges or other journal costs (except costs for prior or early publication); necessary illustrations; clean up, documentation, storage and indexing of data and databases; development, documentation and debugging of software; and storage, preservation, documentation, indexing, etc., of physical specimens, collections or fabricated items.[38]

According to NIH, the following publication costs are allowed:

> Page charges for publication in professional journals are allowable if the published paper reports work supported by the grant and the charges are levied impartially on all papers published by the journal, whether or not by government-sponsored authors. The cost of reprints and publishing in another media, such as books, monographs and pamphlets, also are allowable. Publications and journal articles produced under an NIH grant-supported project must bear an acknowledgment and disclaimer as appropriate, as provided in *Administrative Requirements — Availability of Research Results: Publications, Intellectual Property Rights, and Sharing Research Resources.*"[39]

While publication costs, library fees, and journal subscription costs related to a research project may be allowable costs of federally supported research, it is not known if these allowances will extend to charges for institutional subscriptions for publication costs that open access journals sometimes charge to cover all author payment charges for scientists affiliated with a specific institution. At least one

---

[37] Rules for educational institutions are found in Circular A-21, "Cost Principles for Educational Institutions," Revised May 1, 2004. See section D, for information about direct costs. See section J39, for treatment of publication and printing as F&A "facilities and administrative" indirect costs. Circular A-21 also allows costs of subscriptions as "facilities and administrative" indirect costs. Rules governing nonprofit institutions are discussed in OMB Circular A-110, "Cost Principles for Non-profit Organizations." See the section, Attachment B, items 33 and 41.

[38] National Science Foundation, *Grant Proposal Guide*, September 2004, NSF 04-23, Section II. 2.g.vi(b).

[39] "Selected Items of Cost," in Part II, Terms and Conditions of NIH grant Awards, Subpart a: General — File 3 of 5," in *NIH Grants Policy Statement (12/03)*.

report cautions that some federal agencies may not allow this.[40] Harold Varmus, a co-founder of *PLoS*, considers "publishing fees as the final, relatively cheap step of a research project" and contends that the federal government should pay for these costs.[41]

In 2003, the UK Wellcome Trust, a large research charity that supports biomedical research in the United Kingdom, announced its support of online open access journals and said it would allow scientists it funds to use a portion of their grant to pay author charges required by the journals.[42] The U.S.-based Howard Hughes Medical Institute allows grantees to use up to $3,000 to spend for publishing in open access journals.

Supporters of open access sometimes contend that now most publishing costs are borne by research sponsors, such as the federal government, and that allowing these sponsors to shift support to pay for open access publishing will not cost more and will provide more benefits to society. For instance,

> Asking research sponsors to pay for publication of the research they support may seem to impose new financial burdens on the government agencies, foundations, universities and companies that sponsor research. But these organizations already pay most of the costs of scientific publishing — a huge fraction of the US $9 billion annual revenue of scientific, medical, and technology journals comes from subscriptions, site licenses, and publication fees ultimately billed to grants or employers. Much of the rest is borne by society in the form of increments to university tuitions; healthcare costs, including drug prices; and state and federal taxes that subsidize healthcare, libraries, and education. Surely the cost of open-access digital publishing cannot, in total, be more than we are already paying under the subscription and licensing model. By simply changing the way we support the scientific publishing enterprise, the scientific community and public would preserve everything we value in scientific publishing and gain all of the benefits of open access.[43]

In opposition, some say if the government paid such costs, money would be diverted inappropriately from research to publishing. Some universities say their costs will increase if they need to reimburse researchers to pay author fees for open access journals and if they still have to pay high costs for subscriptions to traditional journals.[44] In addition, some young scientists/investigators say that business models that force authors to pay for publication in open access journals could hurt them since

---

[40] Catherine Zandonella, "Economics of Open Access," *The Scientist*, Aug. 22, 2003.

[41] Malakoff, op. cit., Oct. 24, 2003, p. 553.

[42] Declan Butler, "Wellcome to Fund Publication in Open-access Journals," *Nature*, Oct. 2003.

[43] Patrick O. Brown, Michael B. Eisen, and Harold E. Varmus, "Why PLoS became a Publisher," *PLoS Biology*, vol. 1, no. 1, p. 1.

[44] Lila Guterman, "The Promise and Peril of 'Open' Access," *The Chronicle of Higher Education*, Jan. 30, 2004, op. cit.

they often have smaller grants and "... an author-pays model could amount to a 'tax for productivity.'"[45]

Rising Subscription Costs

It has been reported that traditional academic publishing has a $5 billion global market,[46] and that one of the leading publishers, Reed Elsevier journals, "bring[s] in about $1.6 billion in annual revenue with an operating-profit margin of about 30%." This profit, according to the same source, could be cut to between 10% to 15% if open access publishing were expanded.[47] (The total scientific and technical journal market has been estimated at $9 billion.)[48]

Subscription costs vary depending upon the journal and how many journals an institution subscribes to. Prices also vary for individual versus institutional subscriptions. According to one article, in October 2003 two scientists at the University of California at San Francisco were charged $91,000 "from Elsevier's Cell Press unit for one-year's access to six biology journals."[49] The University of California in 2003 was reportedly charged $7.7 million a year for subscriptions to 1,200 Elsevier periodicals, which was a 25% price reduction from the original bill, negotiated after faculty moves to boycott Elsevier journals if the original bill price were not reduced.[50] Reportedly, sometimes sales are increased by publishers forcing libraries to subscribe to more than they want because publishers often "... bundl[e] ... journal subscriptions into large contracts often not well matched with institutional research interests."[51] This includes bundling together journals that are made available electronically in database systems that access current and archived journals. Bundling of this sort can force libraries to pay for access to the same journal several times if it is included in more than one database to which the library subscribes.[52]

Rising journal subscription costs, it is argued, are too expensive, making it difficult for libraries, especially university libraries and the public to afford many

---

[45] Andrew J. Hawkins, "Scientists at NIH Open access Meeting Fear Author-pays Publishing Would Hurt Young Investigators," *Washington Fax*, Aug. 21, 2004.

[46] This is the figure for academic market sales and is less than the $9 billion figure cited by Brown, Eisen and Varmus, op. cit., possibly because it excludes some sales and consumers.

[47] Wysocki, May 23, 2005, op. cit. The profit margin reduction figures, according to Wysocki are from an estimate by "Sami Kassab, analyst at investment house Exane BNP Paribas in London...."

[48] Brown, Eisen and Varmus, op. cit.

[49] Wysocki, May 23, 2005, op. cit.

[50] Wysocki, May 23, 2005, op. cit.

[51] Jocelyn A. Rankin and Sandra G. Franklin,"Open Access Publishing," *Emerging Infectious Diseases*, July 2004, pp. 1352-1353.

[52] Interview with CRS librarian, July 2005.

journals,[53] and forcing them to sacrifice spending on other media. Reportedly, Rick Johnson, Director of the Scholarly Publishing and Academic Resources Coalition (SPARC), said that because of rising costs, library spending on print media is shifting from monograph and other materials to support largely journal subscriptions, with price per journal reportedly having doubled within 15 years. He illustrated this by saying that while the Consumer Price Index increased 64 percent, libraries are paying 227 percent more for journal subscriptions.[54] According to a National Library of Medicine report, *Access to Biomedical Research Information,* prepared for Congress in June 2004, "prices of commercial biomedical titles increased 224% from 1988 to 1998, while the prices of nonprofit titles increased 129%."[55] The report was quoted as saying that " 'These trends have adversely affected the ability (from a cost standpoint) of academic and health science libraries to continue to support the needs of the research and health care provider communities for access to biomedical literature ....'"[56]

The current open access movement has been fueled by actions of academics and librarians located at the University of California campuses, as well as at other academic sites, who, in late 2003 and 2004, mounted strenuous objections to increases in costs for subscriptions to scientific journals. Some demanded a 25% reduction in subscription fees from major scientific publishers, with Reed Elsevier often cited as a major target, and said if fees were not reduced, they would relinquish journal editorial board memberships or stop providing free peer reviews for major scientific publishers.[57]

However, according to a May 23, 2005 *Wall Street Journal* article,"[c]urrently, the open-access movement makes up between 1% and 2% of the market, experts say. While that number seems small, the concept is assuming an important role channeling academic discontent" about the rising costs of journals.[58]

The Role of Foundation Support for Open Access Journals. The question as to whether open access journals can exist without subsidies may still be unanswered. Some major open access journals (such as *PLoS* journals) receive large contributions from foundations and also impose publishing costs on researchers, which some universities may pay for via an institutional subscription. Some critics complain that this money might otherwise have been used better to support the cost of research. Others wonder whether open access journals and archives can be sustained without philanthropic contributions and what will happen if foundation contributions are ever reduced. It has been reported that several journals which

---

[53] Rankin and Franklin, op. cit.

[54] Damon Brown, "Open Access Journals Offer a New Way of Publishing," *Journal of the American Dietetic Association,* 2004, p. 1060.

[55] As cited in Bradie Metheny, "Open Access Publishing Language in House Labor/HHS Bill Stirs Controversy," *Washington Fax,* July 20, 2004.

[56] Cited in Metheny, July 20, 2004, op. cit.

[57] This last point has been made by Wysocki, op. cit., and others.

[58] Wysocki, op. cit.

attempted to provide free access to readers reversed policies due to falling subscription rates and revenues for print journals. These journals reportedly included the *Journal of High Energy Physics,* which published online for free for six years; it originally did not charge authors a fee, but ultimately decided "to impose a subscription fee of about $1000 a year" for readers.[59] Likewise, the *British Medical Journal*, which had once allowed free online access, reportedly also said it would start to charge for online access.[60] There is also a question of whether, if publishing patterns and revenue sources change, publishers will obtain enough revenue to be able to risk starting up niche journals in narrow fields of science and which have a small readership, which many traditional publishers have been able to do given their revenue margins.[61]

Publishing Revenues Support Scientific Societies

The point is often made that scientific societies, which may publish on their own or may use commercial publishers to publish their journals, reap considerable profits from their share of journal revenues. They then use these profits to support societies' activities, which can include advocacy and assistance to new researchers in the field. Critics of this practice say that these professional associations need to find different business models, or alternative ways to raise money, to support their activities instead of using publishing profits, which are based on payments from subscribers, university libraries, and, in many cases, indirect costs of federally funded R&D.

On the other hand, revenues to scientific societies may not decrease since, at least according to one professional association, the rise of online publishing does not reduce subscriptions to print journals. For instance, according to the American Physical Society (APS), which receives journal publishing profits, preprints of articles in physics, computer science, and mathematics are published on *arXIV.org*, an open and publicly accessible archive. The editor-in chief of the American Physical Society, reportedly said that

> there has been no decline in the subscriber base of journals in those disciplines. In fact the 'contrary is true,' he said. He explained APS journals have a very liberal copyright policy that gives back to the author the right to post articles on e-print servers even before journal publication. They also allow authors to update articles on the servers, using the corrected journal form, after publication ....[62]

---

[59] David Malakoff, "Money Woes Force Some to Change Course," *Science*, Oct. 24, 2003, p. 553.

[60] Malakoff, Oct. 24, 2003, p. 553. For additional information about financial issues, see Catherine Zandonella, "Economics of Open Access," *The Scientist,* Aug. 22, 2003 and Martin Frank, Margaret Reich, and Alice Ra'anan, "A Not-For-Profit Publisher's Perspective on Open Access," preprint as forthcoming in *Serials Review*, vol. 30, no. 4, 2004, p. 6.

[61] Wysocki, May 23, 2005, op. cit.

[62] Bradie Metheny, "Public Representatives Call for Egalitarian Access to Published Research," *Washington Fax*, Aug. 10, 2004.

Commercial and Open Access Publisher Practices

Proponents of open access have alleged that some publishers' practices limit equitable access to scientific information. These practices have been cited as "restrictive licensing terms overriding copyright and fair use practices, [controls on] long-term archival access to electronic content, and ... selective deletions of published articles from database and e-publications."[63] Traditional publishers often disagree and say that they are beginning to adopt some features of open access publishing. Some of these features include, but are not limited to, developing multimedia enhancements, allowing authors to self-archive their articles, and improved content search capabilities.

Journal Enhancements. Some traditional publishers (like many open access publishers) have taken steps to enhance the content of journal articles they post online by permitting digital access, permitting access to ancillary databases and related materials, or allowing posting of preprints in author's websites or institutional repositories.[64] However, often traditional publishers charge a fee to view the journal article or enhancements, "... with fees ranging from a few dollars to a few tens of dollars."[65] Open access proponents say that fees should not be charged for access to these kinds of information.

Timing of Free Access to Journal Articles. Some publishers already allow free access to journal articles a year or more after publication But proponents of open access have argued that the public or other users should not have to wait a year or more to have access to research findings, especially for biomedical research findings, that could be used to improve a patient's health outcome. Another view is that "... limited access to the full text of research articles is bad for science. Such restrictions make it difficult for researchers to build on the entirety of what has gone before and for readers to check whether they have done so. The practice might contribute to citation bias since authors will only reference journals they can access."[66] Still others may find that traditional publishers do not allow electronic access to data in a form that other researchers can easily use to verify findings or to compare in other research projects.

Self-Archiving. Open access publishers require or allow authors to self-archive their articles immediately and to make them accessible for free. Some traditional publishers now allow authors to self-archive on the author's own website the preprint of their article, or, after a delay, the published journal article. There are a variety of models for this, sometimes with fees charged. Some traditional publishers allow authors to self-archive the preprint and then link to the printed version after publication (American Meteorological Association); some do not allow

---

[63] Rankin and Franklin, op. cit.

[64] Guterman, Jan. 30, 2004, op. cit..

[65] Guterman, Jan. 30, 2004, op. cit.

[66] Citing others, this quote is from Pritpal S. Tamber, Fiona Godlee, and Peter Newmark, "Open Access to Peer-reviewed Research: Making It Happen," *The Lancet*, Nov. 8, 2003, pp. 1575-5777.

posting of the article until a year or more after publication in the journal (American Association for the Advancement of Science); some allow posting of an author's article only on an institutional or educational server, not the author's personal self-archive, (American Anthropological Association); and so forth. The policies of hundreds of U.S. and foreign journals, associations, and publishers are summarized in an inventory, published by SHERPA, a British open access project.[67]

Critics say that archiving only on the author's website makes it hard to find sets of related articles in particular subjects because articles are more accessible when placed in freely searchable repositories that archive articles in many fields by many authors and which can be searched by index or keyword terms.

Commercial and Open Access Search Engines. Some commercial publishers already make available free search engines that allow readers to search for citations or abstracts in specific fields or types of information. These include *Scirus*, a search engine limited to science literature managed by Elsevier, which provides access to a short abstract or excerpt. However, most full text articles found through these searches are not accessible for free; costs to read or download an article average $30 per article, which users or libraries are required to pay.[68] Open access bibliographic archives generally provide free access not only to abstracts or citations in multiple fields, but also often to full-text articles. Open access proponents say that there are multiple benefits to providing access in online repositories of collections of articles since a reader's search of such archives could identify many related papers on one topic and would bypass the need for a reader to search individual authors' websites or to use commercial indexing databases that typically charge a fee to read an article.

Copyright Issues. Supporters of traditional publishing argue that publishers, as copyright holders, need copyright protection in order to market journals and sell reprints which support the costs of publishing and archiving both hardcopy and electronic materials. Some also say that copyright ownership is required to guarantee a researcher's accuracy and the authenticity of authorship of an article. In open access publishing, the author of the article retains copyright ownership, but access to the article normally remains free to readers. As will be discussed below, a mixed model is used in the case of NIH's *Public Access Policy*, which asks authors to voluntarily make published articles accessible on *PubMed Central* within a year of journal publication or sooner. Publishers retain the exclusive right to disseminate the work for the time before free access is permitted on *PubMed Central*, but authors are encouraged to conclude agreements with publishers that allow them to place the article in the database. According to NIH, regardless of the publisher's decision, the agency has the right to utilize the journal article under the government purpose license doctrine. This view may be challenged. (See the section on NIH, below, for more details.)

---

[67] "Publisher Copyright Policies and Self-Archiving," SHERPA, [http://www.sherpa.ac.uk/romeo.php?all=yes].

[68] From [http://www.scirus.com].

Economic Development

Open access publishing, according to many proponents, helps promote economic, social, and technical development and equitable access to scientific knowledge by researchers in countries unable to afford the costs of scientific journals by hard copy or subscription web access. Many open access systems also say that they will waive publication charges for authors from developing counties who cannot afford to pay to have their articles published.

But some traditional publishers say that scientists in developing countries already have free and ready access to many traditional scientific journals. For example, many traditional publishers "... participate in projects sponsored by the World Health Organization and the Food and Agriculture Organization of the United Nations to provide medical and agricultural journals to readers in developing countries at low or no cost."[69] In addition, more than 2,000 biomedical journals are accessible online to researchers and health workers in developing countries via a philanthropic project called Health InterNework Access to Research Initiative (*HINASRI)* supported in collaboration with the World Health Organization.[70]

Peer Review and Quality of Articles In Open Access Journals

There is a diversity of views about whether the articles that appear in open access journals have been subject to the same kind of rigorous peer review as those published in traditional journals and about whether they are of comparable quality. The peer review process traditionally involves review of quality of the article and selection of articles to be published in a journal. Usually journal editors or editorial boards make an initial selection of articles to be peer reviewed from among those submitted; use a panel of expert scientists who may volunteer their time to review submissions; select articles to be published from among the articles peer reviewers ranked as high quality; and sometimes do some editing.

A long-held principle is that the accumulation of high-quality scientific knowledge rests on a foundation of publication, typically in commercially distributed scientific journals, with the findings and results vetted and validated through a process involving peer review and fee-based journal subscriptions. Critics allege that the open access "author pays" model of paying for publishing costs, including peer review, prevents quality control filtering mechanisms from working correctly and that, in the long run, scientific articles published in open access sources may be less credible than those published in journals which charge subscription fees.[71] A survey published in 2005, funded by traditional journal publishers, is reported to have found that the quality of peer review was lower in open access than in traditionally published journals:

---

[69] Lila Guterman, Jan. 30, 2004, op. cit.

[70] Brian D. Crawford, "Open-access Publishing: Where is the Value?" *The Lancet*, Nov. 8, 2003, pp. 1578-1580.

[71] Crawford, op. cit.

> Open-access journals ... received fewer submissions and were less selective in choosing among submissions. [It continued] essentially all of the journals reported using editorial review to select and edit submissions. But nearly all of the traditional journals used external peer review, while only editorial staff members reviewed submissions of about 30 percent of the open-access journals.[72]

On the other hand, a study published in 2005 by a publishing analysis firm showed that the quality of nearly 200 open access journals was almost as high in specific medical disciplines as the quality of articles in traditionally published journals.[73]

Some analysts say that peer review of open access journals suffers from the difficulty of finding enough scientist peer reviewers for both the growing number of open access journals and traditional journals. There is also the view that editorial boards of open access journals, which appear to use primarily electronic review of manuscripts, may not filter out unacceptable manuscripts as much as traditional boards do. Thus peer reviewers for open access journals, which also report and interact primarily electronically, may be swamped and, ultimately, there may be delay in the system. Publication in peer reviewed journals figures prominently in promotion and tenure processes in academia. Some observers contend that members of the academic and scientific communities may not view publication on the Internet or in an open access journal to be as prestigious as publication in traditional peer reviewed journal.[74]

Others use citation data as a surrogate measure for quality. Some analysts cite data showing that articles posted in open access journals or freely available on the Internet are used and cited more frequently than those published in traditional journals and are, therefore, a better model to ensure the speedy utilization of scientific research. For instance,

- Experience in physics where researchers publish in traditional journals and then self-archive their papers in a free database is conducive to scientific communication and favorable to authors since " ... papers listed in free archives often get more citations ...."[75]

- A recent study showed that in four disciplines, philosophy, political science, electrical and electronic engineering, and mathematics, articles that are freely available via open access publishing have a greater research impact than those not available via open access.

---

[72] Lila Guterman, "New Study Compares Open-Access and Traditional Publishing," *Chronicle of Higher Education,* Mar. 25, 2005.

[73] Alison McCook, "Open-access Journals Rank Well,"*The Scientist,* Apr, 27, 2005.

[74] Points raised in the discussion session of a meeting on "The Future of Scientific Communication (Formerly Known as Publishing)," American Association for the Advancement of Science (AAAS) Meeting, Apr. 21, 2005.

[75] Alison McCook, "Open Access to U.S. Govt. Work Urged," *The Scientist,* July 21, 2004.

Impact is measured by citations made by other researchers to the literature in the *ISI Web of Science* database.[76]

- In computer sciences, " ... a 2001 study in *Nature*, showed that, at least in one set of disciplines, papers that appear free online are more likely to be cited by other researchers than those that do not. A scientist at NEC Research Institute analyzed nearly 120,000 papers in computer science and related titles. Those that were freely available online had been cited more often in other papers than were those not online, he found. The average number of citations of offline papers was 2.74, compared with 7.03 for those freely available online."[77]

One implication of these data should be noted. Ease of access to articles readily available online, as opposed to those that may be accessible only in hard-copy journals, may increase the propensity to cite them. Thus citation data may not so much measure quality as accessibility.

National Institutes of Health "Enhanced Public Access Policy"

Legislative Origins

On June 26, 2003, Representative Martin O. Sabo introduced the "Public Access to Science Act" (H.R. 2613, 108th Congress), which would have denied copyright protection to publications resulting from federally funded basic scientific research in order to encourage free dissemination of research results to the public.[78] No action was taken on this bill.

Subsequently, the House Appropriations Committee's report on the FY2005 Labor/HHS bill, H.R. 5006, July 14, 2004, contained language that led to the NIH's

---

[76] Kristin Antelman, "Do Open-Access Articles Have a Greater Research Impact?," *College and Research Libraries*, vol. 65, no. 5, pp. 372-282 (Available via *E-LIS*).

[77] Lila Guterman, Jan. 30, 2004, op. cit.

[78] It proposed to "Amend ... Federal copyright law to declare copyright protection unavailable to any work produced pursuant to scientific research substantially funded by the Federal Government to the extent provided in the funding agreement entered into by the relevant Federal agency pursuant to this Act; Require ... any Federal department or agency that enters into a funding agreement with any person for the performance of scientific research substantially funded by the Federal Government to include in the agreement a statement that copyright protection is not available for any work produced pursuant to such research under the agreement; and express the sense of Congress that any Federal department or agency that enters into such funding agreements should make every effort to develop and support mechanisms for making the published results of the research conducted pursuant to the agreements freely and easily available to the scientific community, the private sector, physicians, and the public." (CRS Summary).

"Enhanced Public Access Policy" (House Report 108-636, p. 104). This report contained language, reported to have been authored by Representative Ernest J. Istook, Jr.,[79] which said that it "recommends" that NIH permit open access to NIH-funded research by "requiring" researchers to deposit peer reviewed articles accepted for publication and associated supplemental materials in NIH's *PubMed Central (PMC)*, an open access information system, within six months after publication of the article in a scientific journal. If NIH awarded funds for publishing, the research would be made available immediately upon publication. It also instructed NIH to draft a report by December 1, 2004 on how it would implement this policy. Reportedly "librarians and the Scholarly Publishing and Academic Resources Coalition, or SPARC," lobbied "the Appropriations Committee behind the scenes to include the open-access language in the committee's report ...."[80]

The conference report on the FY2005 Consolidated Appropriations Act, P.L. 108-447 (House Report 108-792, p. 1177), which included funds for Labor/HHS, directed NIH to consider input from publishers as it developed its public access policy, directed NIH to continue to work with publishers to insure the integrity of the peer review system, and requested that NIH "... provide the estimated costs of implementing this policy each year in its annual' budget justification ..." in response to concerns from publishers that NIH's database cannot easily handle the new articles it will be required to archive.[81]

NIH's Policy

NIH's draft policy about archiving published articles that resulted from NIH funding was released for public review and comment in September 2004.[82] After holding several meetings with stakeholders and considering numerous comments from traditional publishers and others submitted during the public comment period,[83]

---

[79] See:"Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, 2005," Remarks of Rep. Ernest Istook on the Floor of the House, *Congressional Record*, Sept. 8, 2004, p. H6833; Andrew J. Hawkins, "Istook Will Clarify NIH Open Access Publishing Language Intent on House Floor," *Washington Fax*, Aug. 31, 2004; and Jocelyn Kaiser, "... Congress Puts Similar Heat on NIH," *Science*, July 23, 2004.

[80] Andrea L. Foster, "House Committee Tells NIH to Post Research Results Online and Make Them Free," *Chronicle of Higher Education*, July 19, 2004.

[81] Shirley Haley, "Omnibus Report Language on Open Access Called a Win By Scientific Societies," *Washington Fax*, Dec. 7, 2004.

[82] The proposed "NIH Public Access Policy," which solicited comments, appeared in the *NIH Guide for Grants and Contracts* on September 3, 2004 [http://grants.nih.gov/grants/guide/notice-files/NOT-OD-04-064.html] and in the *Federal Register* on September 17, 2004, [http://a257.g.akamaitech.net/7/257/2422/06jun20041800/edocket.access.gpo.gov/2004/04-21097.htm].

[83] These are described on the NIH website at [http://www.nih.gov/about/publicaccess/] and are summarized in many articles, such as: Jocelyn Kaiser, "Seeking Advice on 'Open Access,' NIH Get an Earful," *Science*, August 6. 2004; Bradie Metheny, "Public Representatives Call for Egalitarian Access to Published Research," *Washington Fax*, Aug.

(continued...)

NIH issued the final policy, which was published in the *Federal Register* on February 3, 2005.[84] Implementation of the policy started on May 2, 2005. It asks authors funded by NIH to voluntarily submit to NIH for inclusion in the NIH *PubMed Central* system, articles accepted for journal publication, with the final version to be submitted within 12 months after publication in a scientific journal (instead of six months as originally proposed), or sooner if the publisher agrees. According to NIH, the requirement is not mandatory and no penalties would be imposed if an author did not submit an article to the free archive.[85] Thus, NIH-funded scientists are asked to

> ...submit an electronic version of the author's final manuscript, upon acceptance for publication, resulting from research supported in whole or in part by NIH. The author's final manuscript is defined as the final version accepted for journal publication, and includes all modifications from the publishing peer review process. The policy gives authors the flexibility to designate a specific time frame for public release — ranging from immediate public access after final publication to a 12 month delay — when they submit their manuscripts to NIH. Authors are strongly encouraged to exercise their right to specify that their articles will be publicly available through PubMed Central (PMC) as soon as possible.[86]

NIH allows researcher/authors to use the submission of the manuscript to meet certain NIH grant reporting requirements.[87] According to NIH, its policy is compatible with existing publishing models. The agency said it,

---

[83] (...continued)
10, 2004; Andrew Hawkins, "Open Access Should Be A 'Cooperative Venture' Between NIH and Journals, NAS Urges," *Washington Fax,* Nov. 15, 2004; Meredith Wadman, "Director Hits back at Critics of Free Archive Plan," *Nature,* Nov. 25, 2004; M.T. Cavanaugh, "Open Doors: All NIH-funded Work Could Be Freely Available," *Nature,* Nov. 25, 2004; Shirley Haley, "Publishing Delegation Offers Advice, Alternatives to NIH Director on Open Access Plan," *Washington Fax,* Nov. 5, 2004; Lila Guterman, "NIH Proceeds With Plan to Provide Open Access to Scientific Papers," *Chronicle of Higher Education,* Sept. 1, 2004; Andrew Hawkins, "Publishers Argue for Public Access Flexibility, Links to Journals," *Washington Fax,* Nov. 19, 2004; Andrew Hawkins, "Public Access Will Harm Journal/NIH Relationship, AAI Charges; Advocates Dispute Legal Analysis," *Washington Fax,* Nov. 22, 2004; Jocelyn Kaiser, " NIH Unveils Public Access Policy," *Science,* Feb. 3, 2005; Andrew J. Hawkins, "NIH Says Public Access Policy will Change How Science Is Understood," *Washington Fax,* May 2, 2005.

[84] "Policy on Enhancing Public Access to Archives Publications Resulting from NIH-Funded Research," *Federal Register,* Feb. 9, 2005, v. 70, no. 26, pp. 6891-6900.

[85] NIH said in section P of the *Federal Register* rule, that while the House Appropriations report proposed requiring submission, the NIH policy requesting rather than requiring submission "is consistent with the final report language found on page 1177 of the Joint Explanatory Statement in H.Rept. 108-792." See also: NIH. "Questions and Answers: NIH Public Access Policy," Feb. 2005.

[86] "NIH Calls on Scientists to Speed Public Release of Research Publications," *NIH News,* Feb. 3, 2005.

[87] NIH, "Final NIH Public Access Policy Implementation,"March 15, 2005. The database is available at [http://www.pubmedcentral.gov/].

> examined the access policies of the top 20 journals based on citation impact for medicine and medical research and of the 50 journals published by members of FASEB [Federation of American Societies for Experimental Biology]. As of October 2004, 80% of the 20 high impact journals allow public access of some sort through HighWire press within 12 months of publication; of the 50 FASEB journals, 78% offer public access within 12 months."[88]

NIH has also created a Public Access Advisory Working Group of the National Library of Medicine (NLM) Board of Regents composed of stakeholders to advise NIH and NLM on policy implementation and evaluation. Modifications are to be made to the system as it becomes operational and is studied by the group.

NIH Director Zerhouni justified the new policy by explaining that it provides electronic access to NIH-funded research, permits formation of a central archive of NIH-funded research publications, advances science by creating an information resource that scientists can mine, and helps NIH "better manage its entire research investment."[89]

NIH's *PubMed Central (PMC)*

PubMed Central is managed by the National Center for Biotechnology Information of the National Library of Medicine (NLM). It is "the NIH digital repository of full-text, peer-reviewed biomedical, behavioral, and clinical research journals. It is a publicly accessible, stable, permanent, and searchable electronic archive."[90] It does not publish articles; it provides a free repository for journals and published articles that are posted to the site immediately or several months after articles have been published. Free access is allowed to readers, but in some cases only bibliographic information and abstracts are posted. NIH is statutorily mandated to preserve biomedical literature[91] and make it available, and does so via activities of NLM.[92]

The NIH Director estimated that the added costs for posting all NIH-funded research studies on *PubMed Central's* digital library at around $2 to $4 million annually.[93] According to NIH, agency-supported research resulted in 60,000 to 65,000 published papers in 2003.[94]

---

[88] NIH Director Zerhouni, "NIH: Advancing Science in the 21st Century," March 24, 2005, at FLICC Forum on Federal Information Policy, Library of Congress.

[89] Zerhouni, op. cit., Mar. 24, 2005.

[90] NIH. "Questions and Answers: NIH Public Access Policy," Feb. 2005.

[91] See 42 U.S.C. 286(b)(1).

[92] See 42 U.S.C. 286.

[93] Elias Zerhouni, "NIH Public Access Policy," *Science*, Dec. 10, 2004. See also: Janet Coleman, "Open Access Would Cost NIH Roughly $2.5 Million, Agency's Lipman Estimates," *WashingtonFax*, Sept. 24, 2004.

[94] Questions and Answers, op. cit.

Government Purpose License and Copyright Issues

NIH documents indicate that its *Public Access* policy upholds the principles of copyright since submission of manuscripts is voluntary and the statutory fair use privilege still applies to public use of the archived articles. The agency issued guidelines for authors on how to include, in a copyright agreement with a publisher, language that acknowledges the author's obligation to provide a copy of the article to *PubMed Central*.[95]

NIH relies on obtaining permission from authors as the basis for its policy even though "NIH does not need to seek permission from journals who may acquire copyrights from authors or institutions because any copyright transfer or assignment is currently subject to the government purpose license pursuant to 45 C.F.R. 74.36."[96] The term "government purpose license" is not used *per se* in the cited regulation, but is implied. The regulation says,

> The recipient may copyright any work that is subject to copyright and was developed, or for which ownership was purchased, under an award. The HHS awarding agency reserves a royalty-free, nonexclusive and irrevocable right to reproduce, publish, or otherwise use the work for Federal purposes, and to authorize others to do so (45 CFR 74.36(a)).

The concept of nonexclusive right to use the work is similar to the concept of "government purpose license" that is used in the *Federal Acquisition Regulation*, which governs federally funded contracts. Government purpose licensing permits agencies to disseminate to the public scientific and technical articles based on, or containing data produced from, research funded by the agency. The government may subsequently use and distribute the scientific and technical articles as submitted to a publisher or as published in a journal if the publisher has not added any original materials, such as publisher-prepared abstracts or peer review comments. However, generally an agency should obtain a publisher's written permission to reuse or republish the article as published in the journal.[97] Use of "government purpose authority"*per se* to disseminate published journal articles to the public may be limited to contracts funded by those agencies whose originating or authorizing legislation

---

[95] Questions and Answers, op. cit.

[96] *Federal Register*, Section P. Legal Issues.

[97] According to the source: "FAR Subpart 27.4 — Rights in Data and Copyrights provides copyright guidance for the civilian agencies and NASA. In addition, agencies may have their own FAR Supplements that should be followed." The authority granted to the government to use the published version of an article resulting from federally funded research support is implied to be applicable to grants also. See section 4, "Works Created Under a Federal Contact or Grant," of *Frequently Asked Questions About Copyright A Template for the Promotion of Awareness Among CENDI Agency Staff*, CENDI/2004-8.Updated August 2004, HTML last modified May 04, 2005, Edited and updated by Bonnie Klein, Defense Technical and Information Service and Gail Hodge, Information International Associates, Inc., Published by CENDI Secretariat, Information International Associates, Inc., Oak Ridge, TN, August 2004. CENDI is a federal interagency committee, the Commerce, Energy, NASA, Defense Information Managers Group. Available at [http://cendi.dtic.mil/publications/04-8copyright.html].

mandates them to preserve and/or disseminate information to the general public about the agencies' activities and research results.[98]

Other agencies that support scientific grants are governed by OMB Circular A110-section 36, which allows copyrighting by the owner of the work produced from the award of federal funds, but gives the government a nonexclusive right to use it. Specifically,

> The recipient may copyright any work that is subject to copyright and was developed, or for which ownership was purchased, under an award. The Federal awarding agency(ies) reserve a royalty-free, nonexclusive and irrevocable right to reproduce, publish, or otherwise use the work for Federal purposes, and to authorize others to do so.[99]

The Circular A-110 language does not appear to require agencies' enabling legislation to mandate dissemination of research findings, although agency regulations generally require grantees to publish or disseminate the findings of their research and to share data generated by such research. See, for instance, the *NSF Grant Policy Manual* which specifies that "Investigators are expected to promptly prepare and submit for publication with authorship that accurately reflects the contributions of all those involved, all significant findings from work conducted under NSF grants." [100]

Other federal research funding agencies likely to invoke government purpose license or nonexclusive right to use policies if they were to participate in open access systems like NIH's that archive published articles include the National Science Foundation, the Department of Energy, and the National Aeronautics and Space Administration.

Legislative Action 109th Congress

On June 21, 2005, the House Appropriations Committee approved House Report 109-143 on the FY2006 appropriations bill that includes appropriations for NIH (H.R. 3010). The House bill was passed on June 24, 2005. The report endorsed NIH's objectives in establishing the "Public Access Policy" and included language requiring NIH to develop an outreach program to ensure full participation by grantees in volunteering to submit their journal articles to the NIH archive. It also requested the NIH Director to report to Congress by March 1, 2006 on the number of articles deposited and the length of the embargo by publishers — that is, the delay between publication and submission of each peer reviewed article to *PubMed Central* — and

---

[98] Gary G. Borda, NSA Headquarters, "Government Data Rights Under the FAR," March 4, 2003, Slides.

[99] "Uniform Administrative Requirements for Grants and Agreements With Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations," OMB Circular A-110 (Revised 11/19/93, As Further Amended 9/30/99), Section 36(a)).

[100] Section 734, Dissemination and Sharing of Research Results.

to estimate the total number of articles available for deposit.[101] Senate Report 109-103 on this bill endorsed the objectives of the policy but also emphasized the need for interaction between NIH and stakeholders. It urged NIH to work with stakeholders as it implements the new policy; and asked NIH to report by February 1, 2006 on the number of peer reviewed articles deposited in the database, on "the extent to which the implemented policy has led to improved public access," on the impact on the peer review system, and on the cost of operating the database.[102]

Criticisms of "NIH's Enhanced Public Access Policy"

Criticisms of the NIH policy have come from traditional publishers as well as proponents of open access.

For instance, a report prepared for the American Physiological Society criticized the NIH policy as limiting technology development and commercial competition, specifically that "the open access plan 'undermines the principle of [Bayh-Dole] that the private sector is the preferable vehicle to move federally-funded research results to the public and the marketplace.' "[103] It should be noted that the Bayh-Dole law applies to technology transfer, not to publishing of research results.

Others focus on the notion that NIH policy may promote the forfeiture of patent rights. For instance, a legal analysis contends that pre-publication "manuscripts placed on the PMC database 'likely' can be considered 'printed publications' for patent purposes, thus 'triggering the one-year time period for filing a U.S. patent application covering research disclosed in the manuscript ....' "[104] "Current practice," it is charged, "relies on the date of journal publication to start the clock."[105]

---

[101] House Report 109-143, op. cit., p. 104. See also Jocelyn Kaiser, "House Approves 0.5% Raise for NIH, Comments on Database," *Science*, June 17, 2005.

[102] Senate Report 109-103, op. cit., p. 159.

[103] Based on a legal analysis of the technology transfer implications of the NIH proposal by a Foley and Larnder law firm analysis for the American Physiological Society, as reported in Haley, Nov. 18, 2004. The Bayh-Dole act, (35 USC 200-212) allows the government to transfer control of a federally funded invention to a university or business to promote commercialization; the government can license the invention to a third party if it believes it is not being made publicly available on a reasonable basis. See also CRS Report RL32076, *The Bayh-Dole Act: Selected Issues in Patent Policy and the Commercialization of Technology*.

[104] Shirley Haley, " Open Access Plan Faces Copyright, Regulatory Compliance Questions, Legal Analysis Finds," *Washington Fax*, Nov. 18, 2004.

[105] Haley, op. cit. For other criticisms, see Jocelyn Kaiser, "Seeking Advice on 'Open Access,' NIH Gets an Earful," *Science*, Aug. 6, 2004; John T. Softcheck, "PubMed Central's Capacity to Host Open Access Articles Concerns ASM [American Society for Microbiology], *Washington Fax*, Sept. 1, 2004; Danielle Belopotosky, "Online Federal Library on Health Research Sparks Outcry," *Government Exec. Com*, Sept. 3, 2004; Jeffrey Young, "Journal Publishers Ask Senate to Intervene Against NIH Open Access Policy," *Washington Fax*, Sept. 10, 2004; Jeffrey Young, " 'Unnecessary' NIH Open Access Proposal Should Be Discarded, FASEB [Federation of American Societies for Experimental

(continued...)

PLoS's supporters have criticized the NIH policy for its voluntary compliance requirement and said "... the agency's language should have been to 'require' or 'expect' rather than 'request' the deposition of NIH-funded articles in the National Library of Medicine's free-to-use Internet repository, PubMed Central."[106] In addition, according to *PLoS* "... the maximum allowable delay before articles' public release should have been at most 6, rather than 12 months — particularly since no publisher has presented evidence that the free availability of a fraction of its journals' articles half a year after publication would adversely affect subscription revenues."[107] Others say that the 12-month delay for public access falls short of achieving goals of congressional intent and is too lengthy "in a field as dynamic as biomedicine," where patients need immediate access.[108]

NIH policy has also been criticized by some advocates of open access policy who say that NIH should utilize free access policies that exist in the not-for-profit publishing community, which are more cost-effective. They suggest that instead of putting articles in *PubMed Central*, NIH should create a search engine that has the capability to crawl the full texts of existing journals, including nonprofit journals, to allow access to articles on the original journal's website and to provide access to other articles on the topic. Among the groups who have commented on this position is the Washington DC Principles for Free Access to Science[109] and the American Physiological Society.[110]

By way of example, *Google Scholar*,[111] which was launched in 2004, is a free Internet search engine that allows readers to search for peer reviewed articles, preprints, abstracts, grey unpublished literature and other scholarly analyses. If it links to a full-text article, the article is likely to have been published at least a year before the date of the search. There is no assurance that the search engine captures all current or archived materials available in a field. Full text of publisher-controlled, copyrighted materials may be indexed with a citation, but a reader may be linked to the publisher's website to obtain full text of the published version for a fee. In

---

[105] (...continued)
Biology] Says," *Washington Fax*, Nov. 5, 2004; Andrew Hawkins, "Public Access Will Harm Journal/NIH Relationship, AAI [American Association of Immunologists] Charges; Advocates Dispute Legal Analysis," *Washington Fax*, Nov. 22, 2004.

[106] Andy Gass and Helen Doyle, "PLoS Position on NIH Public Access Policy," Letter to the Editor, *Science*, Apr. 15, 2005, p. 352.

[107] Gass and Doyle, Apr. 15, 2005, op.cit.

[108] Comments made by the Alliance for Taxpayer Access as cited in Andrew J. Hawkins, "NIH Public Access Policy Unenforceable, Violates Copyrights, Opponents Charge," *Washington Fax*, Feb. 7, 2005.

[109] Washington DC Principles for Free Access to Science, "Nor-for-Profit Publishers Call New NIH Rule a Missed Opportunity," available at [http://www.dcprinciples.org/nih_rule.htm].

[110] Haley, op. cit., Nov. 18, 2004, citing a legal analysis by Foley and Lardner, law firm.

[111] Available at [http://scholar.google.com/].

addition, there may be a direct link to the full text of a preprint or a version posted by an author or university archive website.

Issues Relating to Federal Open Access Database Archives and Publishing

In addition to NIH's *Public Access* policy and *PubMed Central*, other federal agencies have engaged in open access activities. Several federal agencies publish free, open access, peer reviewed, Internet accessible journals. These journals include *Emerging Infectious Diseases*, by the Centers for Disease Control and Prevention; and *Agricultural Research* and the *Journal of Agricultural Research*, maintained by the U.S. Department of Agriculture and the National Agricultural Library. Others have free, searchable, electronically available repositories that include abstracts, links to full-text articles, and other research reports, some of which may be read online. However, some agencies have confronted serious obstacles to maintaining such systems and have been forced to terminate them. Below is an overview of agency activities and a review of some of the general issues raised about federal involvement in open access publishing and databases.

Federal Open Access Scientific and Technical Archival Databases

Some agencies maintain databases or repositories containing citations, articles or reports that resulted from government-funded research or research funded by other sources, and some include preprints of scientific and technical materials. For instance, the *DOE Information Bridge* allows readers to access for free all available Department of Energy (DOE) preprint report literature (preprint reports prepared for the government via grant or contract that are usually longer than articles published in journals). DOE also has a tool called *E-print* that allows the user to search major preprint systems and university sites where articles are posted. *E-print* is a gateway to over 17,208 websites and databases worldwide that hold "... e-prints in basic and applied sciences, primarily in physics but also including subject areas such as chemistry, biology and life sciences, materials science, nuclear sciences and engineering, energy research, computer and information technologies, and other disciplines of interest to DOE."[112] The system permits documents to be "... circulated electronically to facilitate peer exchange and scientific advancement. Included are pre-publication drafts of journal articles (preprints), scholarly papers, technical communications, or similar documents relaying research results among peer groups."[113]

Other federal agency open access systems include:

---

[112] Available at [http://www.osti.gov/eprint].

[113] Available at [http://www.osti.gov/eprint].

- The *GrayLIT Network,*[114] which includes the searchable full text of gray literature from the Defense Technical Information Center, the DOE, the NASA Jet Propulsion Lab, NASA Langley, and the Environmental Protection Agency.
- The *Federal Research and Development Project Summaries*[115] system contains information about research projects from the DOE, the National Institutes of Health and the National Science Foundation."[116]
- The U.S. Department of Agriculture's (USDA) *AGRICOLA* (**AGRIC**ultural **O**n**L**ine **A**ccess) system, an online bibliographic data base which provides citations, abstracts, and links, when they are available, to published and non-published agricultural literature in the National Agricultural Library.[117]
- The *Astrophysics Data System (ADS)* is a National Aeronautics and Space Administration (NASA)-funded project which maintains four bibliographic databases containing more than 4.2 million records, including links to external resources dealing with: Astronomy and Astrophysics, Instrumentation, Physics and Geophysics, and preprints in Astronomy. The system also contains full-text scans of much of the astronomical literature (almost 50 astrophysics journals).[118]

Objections to Government-Controlled Databases: Censorship and Competition in the Free Market

Allegations of censorship and governmental competition with free market mechanisms are often cited in opposition to government-maintained databases of scientific and technical information.

Allegations of Governmental Censorship. Some critics contend that governmental control of databases of abstracts and journal articles resulting from federally funded research or funded by other sources implies government "censorship and encroachment upon scholarly discourse"[119] and that federal officials, rather than private publishers, could end up determining what research gets published and what does not.

---

[114] Available at [http://graylit.osti.gov].

[115] Available at [http://www.osti.gov/fedrnd].

[116] Marydee Ojala, "PubSCIENCE Joins the Endangered Special List," *Information Today,* Oct. 1, 2002.

[117] Available at [http://agricola.nal.usda.gov/].

[118] Available at [http://adswww.harvard.edu/].

[119] See, for instance, statement of the Association of American Publishers' Patricia Schroeder in Danielle Belopotosky, "Online Federal Library on Health Research Sparks Outcry," *GovExec.com,* Sept. 3, 2004.

Curbs on Department of Energy Information Systems. Some publishers have objected to government-run scientific and technical databases containing abstracts or articles, saying these threaten their publishing activities and employees' jobs. This controversy is illustrated by the experiences of at least two DOE systems.

The DOE *E-print* system, described above, has been controversial, and, according to a DOE official, a few years ago several publishers threatened to prohibit publication of articles that authors posted on it. But eventually the publishers relented and now each publisher has different rules regarding the posting of preprints.[120]

PubScience, was a U.S. Department of Energy effort to provide a free multidisciplinary database for physical sciences literature. It contained indexed abstracts or citations for federally funded and other literature published in commercial journals. Readers could access indexed abstracts for free, but were directed to the commercial website link to obtain the full text article, usually for a fee.[121] The system was initiated on October 1, 1999 and closed on November 4, 2002. According to one article:

> ... the effort quickly became the target of intense lobbying, spearheaded by the Washington-based Software & Information Industry Association (SIIA), a coalition of for-profit and nonprofit members including Reed Elsevier, ISI, Chemical Abstracts Services, and Cambridge Scientific Abstracts. The SIIA claimed that such a service competed with its members' services and argued that government initiative should confine themselves to government information only.[122]

DOE's Office of Scientific and Technical Information (OSTI) operated *PubScience*. According to one DOE official, intense lobbying by publishers and their

---

[120] Interview, DOE official April 2005.

[121] According to an article written shortly before the termination of *PubScience*: "PubSCIENCE launched in October 1999 with the mission of providing free Web search capabilities for journal article abstracts and citations in the physical sciences. Reading the abstract is free, but hyperlinking to the full text generally involves paying for the article. The collection contains over 1,200 journal titles from 35 publishers, including both professional associations (American Association for the Advancement of Science, American Meteorological Society, American Physical Society, American Society for Microbiology, Royal Society of Chemistry, and the Society for Industrial and Applied Mathematics) and private publishers (Blackwell Science, Kluwer Academic Publishers, Nature Publishing Group, Springer-Verlag, and Taylor & Francis Publishers, Ltd.). A few university presses also contribute to the database. Clearly modeled after PubMed, PubSCIENCE wanted to attract scientists and the general public to its information. Noting that the U.S. federal government funds 80 to 90 percent of scientific research and development, DOE touts PubSCIENCE as a significant taxpayer benefit." (Source: Ojala, op.cit.).

[122] Andrew Albanese, "PubScience Dies Despite Comments," *Library Journal*, Dec. 15, 2002. See also: Ojala, op. cit., and "SIIA Releases Comments on DOE's PubScience Decision," Nov. 15, 2002.

associations threatened OSTI's budget.[123] The House Appropriations committee report on the DOE FY2002 appropriation bill, H.R. 2311 (House Report 107-112, pp. 108-109), cautioned DOE about duplication with commercial information services and asked DOE to keep its efforts focused appropriately. The existence of the commercial database *Scirus*[124] and another called *Infotrieve*[125] were cited as competing commercial vendors.[126]

Attempts to Curtail the Federal Database: *PubChem*. Efforts were made in 2005 to close or curtail an NIH database initiated to advance science by assisting basic researchers to identify chemicals related to genetics and cellular research. According several articles, the American Chemical Society (ACS) initially sought closure,[127] and then modified its position to seek limitations,[128] on *PubChem*,[129] which, it says, duplicates ACS's commercial, fee-based *Chemical Abstract Service (CAS)*.

Reportedly, NIH launched *PubChem* in fall 2004 to provide data and to index hyperlinks to articles on the chemical structures of small organic molecules and information on their biological activities to support the "molecular libraries and imaging component of the NIH Roadmap Initiative,"[130]which is a strategic planning process initiated by the NIH Director.[131] *PubChem* contains data organized into three databases: *PubChem Substance*, *PubChem Compound*, and *PubChem BioAssay*. According to NIH,

> Links from PubChem's chemical structure records to other *Entrez* databases provide information on biological properties. These include links to PubMed scientific literature and NCBI's protein 3D structure resource. Links to PubChem's bioassay database present the results of biological screening. Links to depositor web sites provide further information.[132]

The system, reportedly, will expand as it includes more data from the Molecular Libraries centers and data from other online open access chemical database repositories.

---

[123] Interview with OSTI official, April 2005.

[124] Available at [http://www.scirus.com].

[125] Available at [http://www4.infotrieve.com/default.asp].

[126] Andrea L. Foster, "Energy Department Seeks to Close Web Site That Searches Scientific Journals," *The Chronicle of Higher Education,* Sept. 6, 2002.

[127] Jocelyn Kaiser, "Science Resources: Chemists Want NIH to Curtail Database," *Science,* May 6, 2005.

[128] Andrew J. Hawkins, "Chemical Society Entreats Congress to Pull Funding For NIH's PubChem," *Washington Fax,* May 26, 2005.

[129] Available at [http://pubchem.ncbi.nlm.nih.gov/].

[130] Source: [http://pubchem.ncbi.nlm.nih.gov/].

[131] Available at [http://nihroadmap.nih.gov/overview.asp].

[132] From [http://pubchem.ncbi.nlm.nih.gov/].

PubChem, operated by the National Center for Biotechnology Information (NCBI), also provides readers with free access to links to other NCBI databases. It is operated by 13 staff members with a budget of about $3 million.

According to the ACS, *PubChem* jeopardizes its own *CAS* service, which is reported to "... employ ... more than 1,200 people in Columbus, Ohio, and makes a significant contribution to the society's $317 million in annual revenue from publications."[133] *CAS* subscribers receive summary data on chemicals and links to about 24 million abstracts from about 9,000 journals, as well as patent abstracts on more than 25 million chemical substances.[134] NIH is reported to have said that its database provides indexes and links only to biological journals that overlap only slightly with the journals linked by CAS and focuses on "biological information such as protein structures and toxicology," which CAS does not deal with, not broader chemical reactions which CAS covers.[135] An NIH official, Christopher Austin, senior advisor at the NIH Chemical Genomics Center at the National Human Genome Research Institute, was reported to have said that limitation of *PubChem* would have profoundly negative effects on medical discoveries.[136] One report said "The overlap between the two databases occurs in the indexes of chemical names. NIH maintains the overlap is 'quite modest' and for the most part is 'complementary' to CAS. ACS disagrees, saying *PubChem* duplicates CAS' platform and replicates its search features and information."[137] Several articles noted that the ACS lobbied Members of Congress, especially Appropriations Committee members, to have *PubChem* terminated[138] or limited to include only compounds derived from federally funded R&D and to avoid overlap with a commercial enterprise.

Both the House Appropriations Committee and the Senate Appropriations Committee addressed this issue in their reports on the FY2006 appropriations bill that includes appropriations for NIH (H.R. 3010). They did not reduce funding for the database. Both reports said essentially the same thing — that they understood that the database will include chemical compound information from the NIH-funded molecular libraries screening center network and from other sources. But they both expressed concern about duplication of effort with the private sector and urged NIH

---

[133] Kaiser, May 6, 2005, op. cit.

[134] Hawkins, May 26, 2005, op.cit..

[135] Kaiser, May 6, 2005, op. cit.

[136] Aliya Sternstein, "Chemical Publisher Goes After NIH," *FCWCom*, May 27, 2005.

[137] Hawkins, May 26, 2005, op. cit.

[138] "AmChem Soc Calling for Shutting Down Govt. Chem. Database," email from Patrice McDermott, American Library Association, May 17, 2005; Hawkins, May 26, 2005, op. cit.

to work with private sector publishers to avoid unnecessary duplication.[139] The House passed H.R. 3010 on June 24, 2005. Senate action is proceeding.[140]

Reportedly, "Supporters of *PubChem* see the House language as a victory for NIH."[141] An ACS official is reported to have said that the language is a " 'tremendous step in the right direction.' "[142]

Speculation About Differences in Federal Agency Policies. There are no unequivocal answers as to why some agencies can maintain open access systems more easily than others. It may be that publishers, despite their misgivings, moderated their opposition to congressional action to put articles on NIH's *PubMed Central* since the posted articles are limited to those that resulted from NIH funding. However, NIH may be in a different position from other federal agencies since it has a mandate to preserve and provide health information to the public; other agencies may not have such clear mandates to distribute information and the results of their research funding to the public. Furthermore, support for NIH's open access activities seems based not only on the need to allow taxpayers access to results of research their taxes funded,[143] but also on the emotional argument about need for rapid access to information to improve health and save lives, a compelling rationale to many Members of Congress.

Reportedly, DOE's Scientific and Technical Information Advisory Board is discussing, at the highest levels, the question of whether it should establish an open access policy like NIH's to make DOE-funded articles available in its own database and is preparing a report on this subject. According to several federal agency staff, it seems that in the absence of guidance from the congressional appropriations committees, agencies, other than NIH, would likely find it difficult to mount a system like NIH's because of publisher opposition.[144]

Interagency Activities

Scientific publishing and communications methods are slowly changing as Internet publishing becomes more prevalent. Some observers say that government-supported researchers and sponsoring agency staff should participate in shaping these

---

[139] U.S. Congress, House, Committee on Appropriations, *Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriation Bill, 2006,* House Report 109-143, 109th Congress 1st session, p. 112 ,and U.S. Congress, Senate, Committee on Appropriations, *Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriation Bill, 2006,* Senate Report 109-103, 109th Congress, 1st session, pp. 159-160.

[140] See also Jocelyn Kaiser, "House Approves 0.5% Raise for NIH, Comments on Database," *Science,* June 17, 2005.

[141] Kaiser, op. cit., June 17, 2005.

[142] Quoted in Andrew J. Hawkins, "NIH Should Reign In PubChem's Duplicative Services, House Appropriators Warn," *Research Policy Alert,* June 21, 2005.

[143] Istook, op. cit.

[144] Interview with CENDI official, May 2005.

new methods of delivering scientific information. CENDI (the Commerce, Energy, NASA, Defense Information Managers Group), an interagency committee composed of senior Scientific and Technical Information (STI) managers from 12 U.S. federal agencies, has working groups that are studying open access publishing, indexing, and archiving and has issued reports on it to help develop uniform standards and methods of international cooperation.[145]

International Activities

At least two international organizations and other countries are examining wider implementation of open access publishing. Following the release in 2003 of the "Berlin Declaration" which called for open access to knowledge and its signing by representatives of selected European universities, research groups, and government sectors,[146] the European Union began a study on changes in markets for scientific and technical publishing in Europe. Among its topics of inquiry is the subject of "open access to research findings for all and the need to reconcile authors' rights and the economic interests of publishers."[147] The study is expected to be released in late 2005.

In 2004, the Organization for Economic Cooperation and Development's (OECD) science ministers endorsed a policy "based on the principle that research data from public funding should be openly available"[148] on the rationale that

---

[145] CENDI's members are: Defense Technical Information Center (Department of Defense); Office of Research and Development & Office of Environmental Information (Environmental Protection Agency); Government Printing Office; NASA Scientific and Technical Information Program; National Agricultural Library (Department of Agriculture); National Archives and Records Administration; National Library of Education (Department of Education); National Library of Medicine (Department of Health and Human Services); National Science Foundation; National Technical Information Service (Department of Commerce); Office of Scientific and Technical Information (Department of Energy); USGS/Biological Resources Discipline (Department of the Interior). These programs represent over 96% of the FY2004 federal research and development budget. Among CENDI's open access-related working groups are those that deal with "Archiving, Preservation, and Permanent Access" and "Content Management and Access."According to CENDI, "In 1999, CENDI and the International Council for Scientific and Technical Information (ICSTI) jointly sponsored a review of the state of the practice of digital archiving. Over 30 organizations were surveyed and 18 were interviewed to collect information Regarding technology, policy, procedures, and metadata in operational or prototype projects. The results of this project were reported in *Digital Electronic Archiving: The State of the Art and the State of the Practice*, a report to ICSTI and CENDI. An update of the report was completed in 2004. The updated version, CENDI 2004-3, is available in PDF." (Source: [http://www.cendi.gov]).

[146] Available at [http://www.zim.mpg.de/openaccess-berlin/signatories.html].

[147] "EU Investigates Open Access Scientific Publication,"*News - Medical. Net* ..., June 15, 2004.

[148] Peter Arzberger, et al., "An International Framework to Promote Access to Data," (continued...)

providing such access promotes long-term economic benefits, more informed governmental decisionmaking, and hastens the advancement of scientific research. The ministers asked OECD to develop guidelines to "facilitate optimal cost-effective access to digital research data from public funding ..."[149] that would be balanced in terms of opening access while recognizing "the need for restriction of access in some instances to protect social, scientific, and economic interests."[150] These decisions were based, in part, on a report that was funded by the U.S. National Science Foundation.[151] According to the report's authors, "The ultimate goal ... is to make data sharing and the principle of open access the rule rather than the exception."[152] Open access activities in other countries and in international organizations are summarized in Julie M. Esanau and Paul F. Uhlir, eds., *Open Access and the Public Domain in Digital Data an Information for Science, Proceedings of an International Symposium,* Published by U.S. National Committee for CODATA, National Academies Press, Washington, D.C., 2004.

As noted above, there has been considerable governmental and nongovernmental activity to promote open access publishing in the United Kingdom. Some scientific and medical researchers in Britain took steps to make research results freely available via the British open access publisher, *BioMedCentral*.[153] Subsequently, the Science and Technology Committee of Britain's House of Commons issued a report endorsing open access to research results by proposing to require authors to deposit their published papers in online archives and journals using an author pays model and eliminating subscription fees. It also recommended that government agencies mandate that government-funded researchers put their articles into the archives[154] and that the government pay some publishing fees.[155] In November 2004 the U.K. government (the Department of Trade and Industry) rejected the proposal, maintaining there is no indication that access to scientific journals is impeded under current publishing methods, and that according to the

---

[148] (...continued)
Science, Mar. 19, 2004.

[149] "Science, Technology, and Innovation for the 21st Century. Meeting of the OECD Committee for Scientific and Technological Policy at Ministerial Level, 29-30 January 2004 - Final Communique." Annex I., Available at [http://www.oecd.org].

[150] Annex 1, Available at [http://www.oecd.org/document/15/0,2340,en_2649_201185_25998799_1_1_1_1,00.html].

[151] "International Access to Research Data Critical to Advancing Science for the Public Good, Report Says," *NSF Press Release,* NSF PR 04-031, Mar. 18, 2004. The report was not named in the press release. The author was reported to be Peter Arzberger, director of life sciences initiatives at the University of California, San Diego.

[152] NSF PR 04-301, op. cit.

[153] John T. Softcheck, "U.K. Publishing Deal Makes Public Research Results Available to All," *Washington Fax,* June 30, 2003.

[154] Lila Guterman, "British Parliamentary Panel Endorses Open Access to Scientific Literature," *Chronicle of Higher Education,* July 20, 2004.

[155] Daniel Clery, "Scientific Publishing: Mixed Week for Open Access in the U.K.," *Science,* Nov. 12, 2004.

government, "the true costs of open-access publishing are still not clear ..."[156] and "it is 'not obvious ... that the 'author pays' business model would give better value for money than the current one' ...."[157] Subsequently, in June 2005, the United Kingdom Research Councils (RCUK), the main British supporter of publicly funded research, "which distribute[s] most government science funding,"[158] promulgated for comment a draft policy which mandates researchers it funds to archive their journal articles and conference papers "in a free public archive 'at the earliest opportunity, wherever possible at or around the time of publication.' "[159] But the rules may allow publishers to continue to embargo archiving articles until many months after publication, since the council says "its mandate is 'subject to copyright and licensing arrangements' that can restrict what authors do."[160] Costs of publishing in "author pays" journals would be covered by the Research Councils' funding grant "subject to justification of cost-effectiveness."[161] Apparently some learned societies object to this policy since they fear libraries will cancel subscriptions to their professional societies' publications.[162] Comments are being received on the RCUK policy until August 30, 2005. The British government said it will review its policy options on this issue taking into consideration the draft RCUK policy and any changes to it, as well as other information.[163] The Wellcome Trust, a large British medical foundation, recently announced that it requires all papers produced with its support "... to be submitted to the NIH archive or to the British equivalent that is being developed."[164]

Summary of Policy Issues and Questions

Policies for open access journals and citation repositories are evolving and contentious issues may be raised during the 109th Congress. Those that have implications for academic institutions are discussed in **Appendix l.** Other policy issues and questions are emerging, including the following.

---

[156] Clery, op.cit.

[157] Clery, op. cit.

[158] Jim Giles, "UK Research Councils Claim Success for Open-access Publishing Plan," *Nature*, June 2, 2005.

[159] Eliot Marshall, "Scientific Publishing: Britain's Research Agencies Endorse Public Access," *Science*, July 8, 2005. For earlier history see Giles, op. cit.

[160] Marshall, July 8, 2005, op. cit.

[161] "RCUK Announces Proposed Position on Access to Research Outputs," News release 28 June 2005, at [http://www.rcuk.ac.uk/press/20050628openaccess.asp].

[162] Marshall, July 8, 2005 and Giles, June 2, 2005, op. cit.

[163] Email communication from a staff member of the Office of Science and Technology, a British Government official, July 21, 2005, who said "The government position will be reviewed in the light of advice from RCUK, results of studies by JISC and the report from the EU study."

[164] Giles, June 2, 2005, op. cit.

Copyright

- Assessment of which federal agencies, in addition to NIH, would seek to archive and provide free public access to published journal articles reporting the results of research that they supported, using their government purpose license or nonexclusive right to use published articles, regardless of copyright ownership.

- Assessment of rates of voluntary participation by NIH-funded authors in the *Public Access Policy* and determination of whether there are any negative impacts — from research sponsors or the scientific community — on NIH-funded authors who may not submit articles for dissemination in *PubMed Central*. Similarly, determining if publishers penalize authors who seek copyright agreement terms allowing them to post published journal articles in *PubMed Central.*

Quality Control

- Comparison of the quality of peer review processes and of peer reviewed articles that are published in traditional and open access journals.

- Monitoring of whether academic reward systems react differently to articles published by traditional publishers or open access publishers and assessing the implications for professional advancement of researchers and teachers in academic promotion and tenure systems.

- Assessing the positive and negative impacts on the speed and quality of scientific research, knowledge synthesis, and knowledge accumulation flowing from open access publishing and open access citation/abstract archives in comparison with traditional publishing and archival methods.

Monitoring of NIH *Public Access* Activities and Other Federal Initiatives, Including *PubChem*

- Determination if federal open access databases and archival repositories should be limited to providing access only to publications that result from federally funded R&D.

- Assessment of proposals for governmental citation archives to link to publisher's websites to read published articles, as opposed to posting articles on a free access government system.

- Follow-up to congressional mandates that NIH monitor the implementation of its *Public Access* policy, that it work with traditional publishers to monitor the impacts and costs of open access publishing on *BioMedCentral* as it posts what is estimated to

be thousands (possibly 60,000) of additional articles on the system, and that it work with publishers to monitor impacts on the integrity of peer review processes.

- With respect to *PubChem,* assessing cooperation between NIH and ACS in clarifying the possible overlap between NIH's archive and that of the American Chemical Society's *Chemical Abstracts Service*. Analysis of the impacts on biomedical research in general and on NIH's research and its strategically planned genomic research initiatives if the scope of *PubChem* were to be limited.

Who Pays?

- Determining whether federal regulations for support of contracts and grants will continue to allow agencies to pay individual authors or academic institutions for the costs of publishing articles in open access journals as part of the research process, especially if open access publishing becomes more widespread and a substantial portion of the scientific and technical publishing market. A related issue is determining the possible effects on research support funding.

- Given that federal research sponsors allow some journal publishing and subscription costs to be counted as part of the costs to conduct federally sponsored research, comparing the actual total costs to the government for publishing and reading of scientific articles published traditionally as opposed to those published using open access models.

Economic Implications

- Economic analysis of the impacts on the commercial publication industry (revenues, employment, sustainability, etc.) if open access publishing and archiving activities continue to expand.

- Examination of the extent to which professional scientific societies utilize the profits from publishing to support their activities and of alternative sources of funding for these activities.

Appendix 1. Open Access Publishing: Selected Questions in Academia

Continuing questions relating to controversial issues about open access publishing were raised by Andy Gass and Helen Doyle, "The Reality of Open-Access Journal Articles," *Chronicle of Higher Education,* February 18, 2005. They conclude that although there are problems, support is growing in academia for open access journals. Remaining questions include:

> What will become of the market for secondary filters of primary research articles, services like BioMed Central's Faculty of 1000, which highlight important papers published in a wide swath of journals? Will fee-for-access ventures that collect open-access articles become a new cash cow for publishers? At present, faculty members offer their recommendations to the filtering services free, and publishers sell their aggregated opinions to institutions — will established professors go on contributing their free labor to such entrepreneurial enterprises?
>
> How will the role of the research library change, as open-access scholarly communication becomes more widely practiced? To what extent will librarians be freed from the burdens of subscription management?
>
> Many university libraries now encourage open access by subsidizing a portion of the publication charges in open-access venues for authors affiliated with the university, through channels like our employer's institutional membership program. Will those subsidies continue? If so, will they continue to be paid from libraries' budgets, or will they come from research budgets — a source that would be more consistent with the view of open-access proponents that costs of publication should be part of the costs of conducting research? Or will external granting agencies, many of which already pay scientists' page charges and color-illustration fees, assume the full costs of their investigators' open-access publications?
>
> Will libraries continue to serve as intermediaries through which researchers find open-access information, as well as that available only through subscription, and how?
>
> Those questions relate not just to academic libraries, but to the mission of colleges and universities. The time has come for a comprehensive review of how best to pay for the dissemination of professors' work.
>
> How will reduced legal barriers to reusing articles — a stipulation of most formal definitions of open access — affect teaching, research, and other scholarly activities? There are, of course, good precedents for having few or no legal restrictions on the reuse of scholarly work: Every article published by an employee of the NIH is in the public domain. Some more-restrictive open-access licenses now available, like the Creative Commons attribution license in use for articles from our employer and from BioMed Central, permit users to reproduce scholarly work in any medium, for any purpose, as long as the author receives proper credit.

What kinds of educational tools will such licenses make possible? For example, will we see a proliferation of online articles enhanced with explanatory links and informational sidebars, which make scientific discoveries more comprehensible to a wide audience? Will such resources be produced by commercial enterprises? By nonprofit organizations? Or by networks of volunteers, as is the case with open-source computer software?

Will open-access articles enable more researchers from less-developed countries to work on the frontiers of science? Given that all credible open-access journals waive publication fees for authors who can't afford to pay them, increased availability — and therefore knowledge — of the literature might well allow scientists in the developing world to increase their output of cutting-edge work. Would that change, in turn, help resolve the "10/90 gap" — the unfortunate reality that less than 10 percent of the global expenditure on medical research goes to study the predominant health needs of 90 percent of the world's population?

Most important, what kinds of discoveries might result from searchable, open archives of peer-reviewed, full-text scientific literature? The aggregation of gene sequences in a single, freely accessible information space (GenBank) has spawned entire fields of research; will open access to journal articles have a similar effect on areas of work that could benefit from "mining" full texts and figures? Clearly, comprehensive collections of open-access literature would make it much easier to systematically review published medical studies.

Will open-access literature lead to frequent discoveries of correlations between phenomena previously thought to be unrelated? Will it spark more open access to data sets and databases of laboriously compiled and annotated information? The potential for open access to lead to new discoveries is its single most compelling asset, though one that is frequently overlooked.[165]

---

[165] Andy Gass and Helen Doyle, "The Reality of Open-Access Journal Articles," *Chronicle of Higher Education*, Feb. 18, 2005, pp. B13.