

OPEN ACCESS BIBLIOGRAPHY

Open Access Bibliography

Liberating Scholarly
Literature with E-Prints
and Open Access Journals

Charles W. Bailey, Jr.

Association of Research Libraries

2005

Association of Research Libraries
21 Dupont Circle, NW, Suite 800
Washington, D.C. 20036

© 2005 Charles W. Bailey, Jr.

This work is licensed under the Creative Commons Attribution-NonCommercial License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/2.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

ISBN 1-59407-670-7
ISBN 978-1-59407-670-1

∞ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (R1997) Permanence of Paper for Publications and Documents in Libraries and Archives.

Design and layout by David S. Noble.

CONTENTS

| | |
|-----------------------------------------------------------------------------------|-----------|
| PREFACE | XI |
| ACKNOWLEDGMENTS | XV |
| KEY OPEN ACCESS CONCEPTS | XVI |
| | |
| I. GENERAL WORKS | 3 |
| 1.1 Overviews | 3 |
| 1.2 Analysis and Critiques | 6 |
| 1.3 Debates and Dialogs | 10 |
| 1.3.1 <i>Nature Web Debate</i> on Future E-Access to the Primary Literature | 10 |
| 1.3.2 <i>Nature Web Focus</i> on Access to the Literature: The Debate Continues | 13 |
| 1.3.3 Other | 15 |
| 1.4 Research Studies | 15 |
| 1.5 Other | 16 |
| | |
| II. OPEN ACCESS STATEMENTS | 20 |
| 2.1 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities | 20 |
| 2.2 Bethesda Statement on Open Access | 20 |
| 2.3 Budapest Open Access Initiative | 20 |
| 2.4 NEAR | 21 |
| 2.5 OECD Final Communique | 22 |
| 2.6 Tempe Principles | 22 |
| 2.7 Washington D.C. Principles for Free Access to Science | 23 |
| 2.8 Wellcome Trust Position Statement and Research Reports | 23 |

| | | |
|---------------------------------------------------------------|----------------------------------------------------------------------------------------|-----------|
| 2.9 | World Summit on the Information Society Declaration of Principles and Plan of Action | 24 |
| 2.10 | Other | 25 |
| III. COPYRIGHT ARRANGEMENTS FOR SELF-ARCHIVING AND USE | | 27 |
| 3.1 | General Works | 27 |
| 3.2 | Copyright Ownership and Rights | 28 |
| 3.3 | Creative Commons | 30 |
| 3.4 | Permissions Crisis | 31 |
| 3.5 | Research Studies | 32 |
| IV. OPEN ACCESS JOURNALS | | 33 |
| 4.1 | General Works | 33 |
| 4.2 | Economic Issues | 35 |
| 4.2.1 | General Works | 35 |
| 4.2.2 | <i>BMJ Rapid Responses</i> about “Author Pays” May Be the New Science Publishing Model | 40 |
| 4.3 | Open Access Journal Change Agents | 40 |
| 4.3.1 | SPARC | 40 |
| 4.4 | Open Access Journal Publishers and Distributors | 42 |
| 4.4.1 | BioMed Central | 42 |
| 4.4.2 | Public Library of Science | 44 |
| 4.4.3 | PubMed Central | 47 |
| 4.4.3.1 | <i>General Works</i> | 47 |
| 4.4.3.2 | <i>Science Magazine dEbate</i> on “Building a GenBank of the Published Literature” | 50 |
| 4.4.3.3 | <i>Science Magazine dEbate</i> on “Is a Government Archive the Best Option?” | 50 |
| 4.4.3.4 | <i>Science Magazine dEbate</i> on “Just a Minute, Please” | 51 |
| 4.4.3.5 | <i>Other</i> | 51 |

| | | |
|-------------|----------------------------------------------------------|-----------|
| 4.5 | Specific Open Access Journals | 53 |
| 4.5.1 | Journals in the <i>Directory of Open Access Journals</i> | 53 |
| 4.5.2 | Pioneering Free E-Journals Not in the <i>DOAJ</i> | 55 |
| 4.5.3 | Other | 56 |
| 4.6 | Research Studies | 57 |
| V. | E-PRINTS | 59 |
| 5.1 | General Works | 59 |
| 5.2 | History | 61 |
| 5.3 | Research Studies | 61 |
| VI. | DISCIPLINARY ARCHIVES | 63 |
| 6.1 | General Works | 63 |
| 6.2 | Specific Archives and Projects | 63 |
| 6.2.1 | arXiv | 63 |
| 6.2.2 | NASA Astrophysics Data System | 64 |
| 6.2.3 | RePEc | 65 |
| 6.2.4 | Other | 66 |
| VII. | INSTITUTIONAL ARCHIVES AND REPOSITORIES | 68 |
| 7.1 | General Works | 68 |
| 7.2 | E-Print Archives | 68 |
| 7.2.1 | General Works | 68 |
| 7.2.2 | Specific Archives and Projects | 69 |
| 7.2.2.1 | <i>ePrints-UK</i> | 69 |
| 7.2.2.2 | <i>SHERPA</i> | 69 |
| 7.2.2.3 | <i>Other</i> | 69 |
| 7.3 | Repositories with Diverse Materials | 70 |

| | |
|------------------------------------------------------------|-----------|
| 7.3.1 General Works | 70 |
| 7.3.2 Specific Repositories and Projects | 72 |
| 7.3.2.1 <i>DAEDALUS</i> | 72 |
| 7.3.2.2 <i>DSpace</i> | 72 |
| 7.3.2.3 <i>eScholarship</i> | 74 |
| 7.3.2.4 <i>Fedora</i> | 75 |
| 7.3.2.5 <i>OSU Knowledge Bank</i> | 76 |
| 7.3.2.6 <i>Other</i> | 76 |
| 7.4 Electronic Theses and Dissertations | 77 |
| VIII. OPEN ARCHIVES INITIATIVE AND OAI-PMH | 78 |
| 8.1 General Works | 78 |
| 8.2 Specific Data or Service Providers and Projects | 81 |
| 8.2.1 AmericanSouth.org | 81 |
| 8.2.2 Arc | 81 |
| 8.2.3 Kepler | 82 |
| 8.2.4 OAlster | 82 |
| 8.2.5 OpCit | 82 |
| 8.2.6 Open Archives Forum | 83 |
| 8.2.7 Open Archives Initiative Metadata Harvesting Project | 85 |
| 8.2.8 Other | 85 |
| 8.3 Research Studies | 90 |
| IX. CONVENTIONAL PUBLISHER PERSPECTIVES | 91 |
| X. GOVERNMENT INQUIRIES AND LEGISLATION | 94 |
| 10.1 European Commission Study | 94 |
| 10.2 Sabo Bill | 94 |

| | | |
|------------|----------------------------------------------------------------|------------|
| 10.3 | U.K. House of Commons Science and Technology Committee Inquiry | 95 |
| 10.4 | U.S. House Appropriations Committee Recommendations | 100 |
| 10.5 | Other | 101 |
| XI. | OPEN ACCESS ARRANGEMENTS FOR DEVELOPING COUNTRIES | 102 |
| 11.1 | General Works | 102 |
| 11.2 | Free or Reduced Cost Access | 103 |
| 11.2.1 | AGORA | 103 |
| 11.2.2 | HINARI | 103 |
| 11.2.3 | Other | 104 |
| 11.3 | SciELO | 106 |
| | ABOUT THE AUTHOR | 107 |

PREFACE

Scope of the Bibliography

The *Open Access Bibliography: Liberating Scholarly Literature with E-Prints and Open Access Journals* presents over 1,300 selected English-language books, conference papers (including some digital video presentations), debates, editorials, e-prints, journal and magazine articles, news articles, technical reports, and other printed and electronic sources that are useful in understanding the open access movement's efforts to provide free access to and unfettered use of scholarly literature. Most sources have been published between 1999 and August 31, 2004; however, a limited number of key sources published prior to 1999 are also included. Where possible, links are provided to sources that are freely available on the Internet (approximately 78 percent of the bibliography's references have such links).

There are various definitions of "open access." The scope of this bibliography is determined by the "Budapest Open Access Initiative" definition:

The literature that should be freely accessible online is that which scholars give to the world without expectation of payment. Primarily, this category encompasses their peer-reviewed journal articles, but it also includes any unreviewed preprints that they might wish to put online for comment or to alert colleagues to important research findings. There are many degrees and kinds of wider and easier access to this literature. By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited. . . .

To achieve open access to scholarly journal literature, we recommend two complementary strategies.

I. Self-Archiving: First, scholars need the tools and assistance to deposit their refereed journal articles in open electronic archives, a practice commonly called, self-archiving. When these archives conform to standards created by the Open Archives Initiative, then search engines and other tools can treat the separate archives as one. Users then need not know which archives exist or where they are located in order to find and make use of their contents.

II. Open-access Journals: Second, scholars need the means to launch a new generation of journals committed to open access, and to help existing journals that elect to make the transition to open access. Because journal articles should be disseminated as widely as possible, these new journals will no longer invoke copyright to restrict access to and use of the material they publish. Instead they will use copyright and other tools to ensure permanent open access to all the articles they publish. Because price is a barrier to access, these new journals will not charge subscription or access fees, and will turn to other methods for covering their expenses.¹

The open access movement exists in the broader context of a complex scholarly publishing system. It is widely believed by academic librarians and others that this system is in a state of crisis due primarily to the increasing cost of scholarly journals far in excess of inflation, the proliferation of new journals that are ever more specialized, the failure of library budgets to keep up with these cost and journal proliferation factors, and the resultant increasing restriction of access to journal literature as libraries cancel existing journals and fail to add new specialized ones. Although the open access movement will clearly have a very significant impact on the library “serials crisis” if it succeeds, many of its primary advocates do not see the resolution of this crisis as its primary mission, but, rather, as a desirable potential side effect. This bibliography does not deal with the serials crisis or the important scholarly publishing reform movements that it has engendered that are not related to open access. When general reform-oriented topics, such as changing copyright laws or understanding their impact on research and instruction, are covered in this bibliography, it is in relation to open access concerns.

Likewise, the bibliography limits its coverage of general electronic publishing topics, such as electronic theses and dissertations, to those works that have direct relevance to open access concerns (e.g., electronic theses and dissertations in the context of institutional repositories).

The reader is referred to the author’s *Scholarly Electronic Publishing Bibliography* for an in-depth treatment of the above topics.²

The bibliography does cover a few topics, such as free or reduced cost access to journal literature for developing countries and pioneering free e-journals, that the author views as being very closely aligned with the open access movement, even though they are not open access per se.

Inevitably, there are gray zones between open access and other closely related reform efforts that, in some cases, are intertwined with it. For example, SPARC fosters both open access and low-cost journals. The bibliography includes general articles about SPARC and articles about its open access efforts, but not specialized articles that are solely about its important support of competitive low-cost journals.

The author has attempted to find the right balance between full coverage of a wide range of issues relevant to the open access movement (e.g., major supporting technologies such as institutional repositories and OAI-PMH) and too much inclusion of interesting and important, but potentially irrelevant, material that is closely related to them. While the bibliography covers some esoteric technical areas in detail, it is not intended to be a complete record of all research efforts in these areas, but, rather, a sampling of key works.

There is no consistency in the literature about the hyphenation of “open access” in compound terms (e.g., “open-access journals” or “open access journals”). In this preface, such compound terms are not hyphenated, which appears to be the prevalent trend among scholars.

Construction of the Bibliography

The author has employed a variety of search strategies to identify works for inclusion in the bibliography. Searches were conducted in major index and abstract databases, Internet search engines, OAI-PMH search services (e.g., Arc, Citebase, and OAIster), open access journals (e.g., BioMed Central journals), open access archives (e.g., PubMed Central), Weblogs, freely available e-serials, mailing lists, author and project Web sites, and licensed e-serials and indexes. Of particular note are Peter Suber’s excellent e-publications (*Open Access News*³ and the *SPARC Open Access Newsletter*⁴ among others), which were rich, extremely useful sources of information. A “pearl growing” approach was used: when relevant articles were identified, their reference lists were checked for new sources, and, in turn, the reference lists of these new sources were checked in an iterative fashion. In electronic resources with “articles by,” “related articles,” and “articles that cite this work” search features (e.g., *BMJ*), these powerful capabilities were also used.

While the bibliography is selective, it errs on the side of inclusiveness in cases of doubt in an attempt to fully capture the rapid, vigorous growth of the open access movement, which is still in an early stage of its development. However, it doesn’t include several types of material that may be of potential interest to readers: (1) PowerPoint or similar digital “overheads” from conference presentations (the bibliography does include complete conference papers or digital videos) and (2) electronic-only articles or other works that require free registration. See *Open Access News* for coverage of these works.

Since the bibliography includes many diverse electronic sources, the creation of references required more creative interpretation than would be the case with print sources. Electronic-only works that appear to have been formally published by an identifiable organization are generally treated like equivalent print works; however, in some cases, it was not possible to determine factual information, such as place of publication. Only author, title, and URL information is given for unpublished e-prints or self-published Web pages. To avoid ambiguity, periods are not placed after references’ URLs (except in preface notes).

For some electronic journals, articles are numbered, and they may or may not have internal pagination intended for citation purposes. The article number is usually presented by the publisher in the page position of the reference. These works have been represented in the bibliography with the article number grouped with the year of publication as the following example illustrates: Shidham, Vinod B., Anthony Cafaro, and Barbara F. Atkinson. “*CytoJournal* Joins ‘Open Access’ Philosophy.” *CytoJournal* 1 (Article 1 2004). <http://www.cytojournal.com/content/1/1/1>.

The author has attempted to provide references that give as much relevant information as possible; however, if a dual-format work was only available to him in electronic format, the reference was based solely on that version and, if pagination information was not included, it was omitted from the reference. Some conference paper references have been based on information from electronic indexes, such as *OCLC PapersFirst*.

If an e-print for an article that was published in a restricted access journal could be located, its URL was included with the reference for the journal article. In some cases, it is clear that this is an e-print; in others, it is not. When an e-print is available in a disciplinary archive, an institutional e-print archive, or an institutional repository, the URL is frequently given for the e-print record, rather than for the e-print itself, in an effort to assist the reader in identifying the work as an e-print. For other e-prints, careful examination of the URL will help the reader determine if the article is an e-print or not (e.g., is the URL to the publisher's Web site?).

Some URLs for works in the bibliography have been constructed by their publishers using special characters, such as commas, pound signs, spaces, or underscores. You may have difficulty accessing these works. If so, try using one of the Mozilla family of browsers.⁵

There has not been an effort to standardize author names to eliminate variations. Interviewers and interviewees are treated as article coauthors.

References and URLs were last checked and corrected on 8/31/04. Given the high degree of inclusion of "gray literature" in the bibliography, the reader should expect URL decay and, to some degree, reference decay as well.

The bibliography was created using the EndNote software. Reference formatting and alphabetization was done by this software using a slightly modified version of the *Chicago Manual of Style* 14th edition output style.

Notes

1. Budapest Open Access Initiative, "Budapest Open Access Initiative," 14 February 2002, <http://www.soros.org/openaccess/read.shtml>.
2. Charles W. Bailey, Jr., *Scholarly Electronic Publishing Bibliography* (Houston: University of Houston Libraries, 1996–2004), <http://info.lib.uh.edu/sepb/sepb.html>.
3. *Open Access News*, <http://www.earlham.edu/~peters/fos/fosblog.html>.
4. <http://www.earlham.edu/~peters/fos/newsletter/archive.htm>.
5. <http://www.mozilla.org/>.

ACKNOWLEDGMENTS

The research for this bibliography was conducted while I was on a development leave from the University of Houston Libraries. I wish to thank the Dean of the University of Houston Libraries, Dana C. Rooks, for her support and encouragement. I also wish to thank Peter Suber, Research Professor of Philosophy at Earlham College (as well as Senior Researcher at the Scholarly Publishing and Academic Resources Coalition and Open Access Project Director at Public Knowledge) for his invaluable assistance in reviewing numerous drafts of the bibliography. Thanks to David Prosser, Director of SPARC Europe, for providing very helpful comments on a nearly final draft; Lee Anne George, Publications Program Officer of the Association of Research Libraries, for publishing and expertly editing the bibliography; and David Noble for his skillful layout and graphic design efforts. Special thanks to my wife, Jane D. Segal, for her patience and support during the lengthy research and writing process.

KEY OPEN ACCESS CONCEPTS

This section provides a brief introduction to some of the major open access concepts needed to fully utilize the bibliography. It is not intended as a complete guide to open access (see articles in the “1.1 Overviews” section of the bibliography for such introductions).

Open Access Defined

The “Preface” provided an excerpt from perhaps the most influential open access statement, the “Budapest Open Access Initiative.” The important things to note in the basic definition are that open access deals with peer-reviewed articles or preprints and that free access to these works is not equivalent to open access. Open access also requires no restrictions on how published material is subsequently used except to require that proper attribution of the work be given to the author and that authors retain control over the integrity of their work.

In practice, what makes open access possible is that it is relatively inexpensive to distribute electronic articles on the Internet. Consequently, the open access movement focuses on various electronic publishing strategies.

While the BOAI provides a landmark definition of open access, other groups’ statements have somewhat different views, and the concept of open access is still evolving. For example, the “Bethesda Statement on Open Access Publishing” also includes a requirement that:

A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in a suitable standard electronic format is deposited immediately upon initial publication in at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving (for the biomedical sciences, PubMed Central is such a repository).¹

In spite of an emphasis on peer-reviewed articles in key statements, open access can be applied to a variety of scholarly works that are produced by scholars without the expectation of payment. Although this is likely to be a later development, Peter Suber has suggested that it may even be applied with greater difficulty to scholarly materials that do involve author payment or profit.²

The BOAI statement suggests two strategies for achieving open access: self-archiving and open access journals.

Self-Archiving

Self-archiving can be achieved in at least three ways: (1) putting articles on author Web sites, (2) depositing articles in disciplinary archives, or (3) depositing articles in institutional archives and repositories.

Copyright Concerns

Self-archiving is contingent on authors having the legal right to electronically distribute their articles. This is challenging because many scholars relinquish their copyrights to publishers and, historically, many journal publishers did not want to consider articles that had been distributed as electronic preprints for publication. Since the late 1980's, there has been a growing trend for authors to want to retain their article copyrights. This is primarily the result of the advent of free scholar-produced journals (which often let scholars retain copyright), an increased awareness of copyright issues as a result of the serials crisis, the vigorous self-archiving advocacy efforts of Stevan Harnad and others, and the rise of the open access movement. Authors who retain their copyrights can then grant publishers the limited rights that they need to effectively distribute their works and/or they can put their works under a license that grants certain rights to all potential users and distributors.

A barrier to author granting of rights has been that framing the proper wording of license agreements is a complex process requiring significant legal knowledge. The Creative Commons has greatly facilitated the use of author license agreements by developing a variety of standard agreements that authors can easily select and utilize.³ The Creative Commons Attribution license meets open access requirements.

In recent years, publishers have been more open to considering preprints for publication, granting authors the right to archive their works, and allowing them to retain copyright if they request it.

E-Prints

Self-archived articles may be preprints (i.e., draft articles that have not been peer-reviewed or edited) or postprints (i.e., final, edited versions of peer-reviewed articles). The term used for both is "e-prints." Certain scientific disciplines, such as physics, have a long history of e-print distribution. Initially, this was done by scholars mailing colleagues preprints (or reprints). Later fax was used, then e-mail.

Author Web Sites

While some authors archived articles on FTP or Gopher sites, it was the widespread utilization of the Web starting roughly in the mid-1990's that resulted in a significant growth of personal e-print archives. A key problem with such archives is that they can be unstable, as authors move from institution to institution, retire, make other life changes, or die. As will be seen later, e-prints from such archives are not made as easily visible to the research community as those in disciplinary archives or institutional archives and repositories because they cannot be easily harvested.

Disciplinary Archives

In the early 1990's, formal "disciplinary archives" began to displace scholar-to-scholar distribution in some scientific disciplines. A disciplinary archive provides access to e-prints for one scholarly discipline or multiple scholarly disciplines. Keep in mind that some disciplines have many subfields, and that specialties that draw on many traditional disciplines are increasingly common. Some disciplinary archives provide access to diverse scholarly works, not just e-prints. Usually, a disciplinary archive can be searched and browsed.

The most famous disciplinary archive is probably arXiv, which covers physics, mathematics, non-linear science, computer science, and quantitative biology.⁴ It was established in 1991.

It is important to keep in mind that some disciplines rely more heavily on articles than others, and that some disciplines that rely heavily on articles do not have a strong tradition of using e-prints. Consequently, there can be significant disciplinary differences in receptiveness to open access.

Institutional Archives and Repositories

Where disciplinary archives provide access to the worldwide literature of one or more fields, institutional archives and repositories focus on the literature produced by a single institution.

An institutional e-print archive may contain e-prints written by scholars from many departments, research centers, or other units. Or, it may only contain the e-prints of a single unit.

An institutional repository includes a variety of materials produced by scholars from many units, such as e-prints, technical reports, theses and dissertations, data sets, and teaching materials. Some institutional repositories are also being used as electronic presses, publishing e-books and e-journals. DSpace at MIT is a notable example of an institutional repository (the DSpace project began in 2000).⁵

Typically, e-print archives and institutional repositories can be searched and browsed.

Freeware is usually used to support these efforts. Popular choices include Eprints⁶ for institutional e-print archives and Dspace⁷ or Fedora⁸ for institutional repositories. (The popular EPrints software has also been used for disciplinary archives and institutional repositories.)

Open Access Journals

After self-archiving, the second major BOAI strategy is open access journals. Open access journals allow authors to retain their copyrights, but may require that they agree to license their articles with the Creative Commons Attribution license or a similar license.

Open access journals are primarily electronic journals (print editions are sometimes offered as an optional fee-based add-on). Once the first electronic copy of a journal has been created, the costs of distributing it on the Internet are negligible compared to the costs of distributing additional print copies of a conventional journal. Open access advocates also note other cost savings implicit in their approach, such as the elimination of the need for access controls. Still, open access journals cost money to produce and distribute, especially since they are peer-reviewed and edited like conventional journals. Various funding strategies are in use, but the most common are direct author fees, institutional memberships to

sponsor all or part of author fees, funding agency payment of author fees, grants to open access publishers, institutional subsidies (such as paying the salaries of journal editorial staff), and priced add-ons (such as recommendation services, current awareness services, or print editions).

Open access journals may be included in index and abstract services. The *Directory of Open Access Journals* is a major finding tool, which permits searching at the article level for some journals.⁹

Preliminary research suggests that the “impact” factors of open access journals can be at least as good as those of conventional journals.¹⁰

Three organizations play a major role in the publication and archiving of open access journals: BioMed Central, the Public Library of Science (PLOS), and PubMed Central.

Established in 2000, BioMed Central is a for-profit publishing company that publishes over 100 open access biomedical journals.¹¹

The Public Library of Science is a nonprofit organization that, as of August 2004, publishes one open access journal (*PLoS Biology*).¹² A second journal (*PLoS Medicine*) is expected to be launched in October 2004. The PLoS started in 2000. Its first activity was to circulate an open letter that was intended to convince biomedical publishers to make their journals freely available within six months of publication. Roughly 34,000 scientists from 180 countries ultimately signed the letter, pledging not to publish in (or otherwise support) journals that did not meet this requirement by September 2001. When this letter did not invoke the desired response, the Public Library of Science began to publish its own open access journals.

PubMed Central is a freely available life sciences journal archive that is run by the National Center for Biotechnology Information of the National Library of Medicine.¹³ Journals must meet certain editorial standards to be included in the voluntary archive. As originally conceived in 1999 by Harold Varmus (who was then the Director of the National Institutes of Health), PubMed Central (then called E-biomed) had a broader mission that included e-prints; however, this original vision was very controversial, and it was significantly modified by PubMed Central’s launch in 2000.

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)

Since open access works are scattered across many disciplinary archives, institutional e-print archives, institutional repositories, and open access journals, it can be difficult for scholars to locate all needed works on a particular subject. It requires scholars to search one system after another in a serial fashion. To deal with this problem, the Open Archives Initiative¹⁴ developed the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) to allow search systems (called service providers) to retrieve metadata about open access works from archives and repositories (called data providers) and aggregate this data so that it can be searched with a single query. The first version of OAI-PMH was released in 2001.

There are several notable service providers that can help users locate relevant e-prints. The University of Michigan Digital Library Production Service’s OAIster service harvests metadata from over 340 data providers and provides unified searching of this metadata.¹⁵

The Digital Library Research group at Old Dominion University maintains a similar, smaller scale experimental service called Arc.¹⁶ The Open Citation Project's experimental Citebase service uses citation ranking to enhance search result displays.¹⁷ The Open Archives Initiative maintains a list of other service providers.¹⁸ Major search engines, such as Google and Yahoo, have begun to index metadata from various data and service providers through cooperative projects. These projects supplement existing efforts by search engines to index e-prints, allowing them to overcome special technical problems associated with indexing some of these items.

Government Inquiries and Legislation

There has been increased government scrutiny in the U.K., the U.S., and the European Union of the conventional scholarly publishing system, with particular attention being paid to the perceived high cost of scientific, technical, and medical journals and the fact that much research in these areas is government funded.

In the U.S., legislation was introduced in 2003 that would put works that are “substantially funded” by government money into the public domain (the “Public Access to Science Act,” informally called the “Sabo bill” after its sponsor, Representative Martin O. Sabo).¹⁹

In 2004, the UK House of Commons Science and Technology Committee concluded a significant inquiry into scientific publishing that resulted in a report (*Scientific Publications: Free for All?*).²⁰ This report recommended that articles resulting from government-funded research be deposited in institutional repositories, which would be established at all UK higher education institutions, and that funds be made available to pay open access journal publication fees for such articles (authors would need to apply for these funds).

In 2004, the U.S. House Appropriations Committee recommended that articles that result from NIH grant-funded research be deposited in PubMed Central upon acceptance for publication.²¹ If NIH funds were used to support any publication costs, the articles would be made immediately available. Otherwise, they would be made available six months after publication. NIH would develop a plan by 12/1/04 to implement the recommendation in FY 2005.

Also in 2004, the European Commission announced it would conduct a major study of the scientific publication markets in Europe, which would be completed in 2005.²²

Developing Countries

While not strictly an “open access” strategy, an important closely related effort by traditional publishers and others has been the creation of special journal access arrangements for developing countries, whose scholars may otherwise have very limited access to frequently expensive journals. These arrangements provide free or reduced cost access to journals. Major initiatives include the Access to Global Online Research in Agriculture (AGORA)²³ and the Health InterNetwork Access to Research Initiative (HINARI),²⁴ which provides access to biomedical journals. On a smaller scale, the novel Ptolemy Project provides medical researchers and clinicians in the developing world with free access to licensed e-journals and other e-resources by making them research affiliates of the University of Toronto.²⁵

Open access is taking root in developing countries, and, in addition to other benefits, it provides a way to increase the visibility of research from these countries. For example, Scientific Electronic Library Online (SciELO) is an innovative scientific electronic publishing cooperative that focuses on providing open access to Latin America and Caribbean journals.²⁶

Notes

1. “Bethesda Statement on Open Access Publishing,” 20 June 2003, <http://www.earlham.edu/~peters/fos/bethesda.htm>.
2. Peter Suber, “Creating an Intellectual Commons through Open Access,” 2004, <http://dlc.dlib.indiana.edu/archive/00001246/>.
3. <http://creativecommons.org/>.
4. <http://arxiv.org/>.
5. <https://dspace.mit.edu/index.jsp>.
6. <http://software.e-prints.org/>.
7. <http://libraries.mit.edu/dspace-mit/index.html>.
8. <http://www.fedora.info/>.
9. Lund University Libraries, *Directory of Open Access Journals*, <http://www.doaj.org/>.
10. James Testa and Marie E. McVeigh, “The Impact of Open Access Journals: A Citation Study from Thomson ISI,” 2004, <http://www.isinet.com/media/presentrep/acropdf/impact-oa-journals.pdf>.
11. <http://www.biomedcentral.com/>.
12. <http://www.publiclibraryofscience.org/>.
13. <http://www.pubmedcentral.nih.gov/>.
14. <http://www.openarchives.org/>.
15. <http://www.oaister.org/o/oaister/>.
16. <http://arc.cs.odu.edu/>.
17. <http://citebase.eprints.org/cgi-bin/search>.
18. Open Archives Initiative, “Registered Service Providers,” <http://www.openarchives.org/service/listproviders.html>.
19. Peter Suber, “Martin Sabo’s Public Access to Science Act,” *SPARC Open Access Newsletter*, no. 63 (2003), <http://www.earlham.edu/~peters/fos/newsletter/07-04-03.htm>.
20. House of Commons Science and Technology Committee, *Scientific Publications: Free for All?* (London: Science and Technology Committee, House of Commons, United Kingdom

Parliament, 2004), <http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/399.pdf>.

21. Peter Suber, "NIH Open-Access Plan: Frequently Asked Questions," 2004, <http://www.earlham.edu/~peters/fos/nihfaq.htm>.

22. <http://www.europa.eu.int/comm/research/press/2004/pr1506en.cfm>.

23. <http://www.aginternetnetwork.org/en/>.

24. <http://www.healthinternetnetwork.org/>.

25. <http://www.utoronto.ca/ois/myweb9/>.

26. <http://www.scielo.org/index.php?lang=en>.