RUTGERS
University Libraries

Vice President for Information Services
and University Librarian
Archibald S. Alexander Library
Rutgers, The State University of New Jersey
169 College Avenue
New Brunswick, NJ 08901-1163

gaunt@rci.rutgers.edu
www.libraries.rutgers.edu

732-932-7505
Fax: 732-932-7637

January 10, 2012

Office of Science and Technology Policy
The White House

Re: Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting from Federally Funded Research

The Rutgers University Libraries write in response to the Request for Information, published in the *Federal Register* on November 4, 2011, by the Office of Scientific and Technology Policy regarding public access to peer-reviewed scholarly publications resulting from federally funded research. We appreciate this opportunity to comment.

(1) Are there steps that agencies could take to grow existing and new markets related to the access and analysis of peer-reviewed publications that results from federally funded scientific research? How can policies for archiving publications and making them publically accessible be used to grow the economy and improve the productivity of the scientific enterprise? What are the relative costs and benefits of such policies? What type of access to these publications is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise?

New scientific information spurs research, development, and innovation. Broad access to scholarly publications enables us to build on the past and is fundamental to the scientific enterprise. A major barrier to the scientific enterprise is lack of access to scholarly journals due to prohibitive costs. As research institutions and libraries are increasingly unable to afford subscriptions to peer-reviewed publications that have resulted from government-funded research, the benefit from research and development is lost and productivity of the scientific enterprise is needlessly impeded.

Facilitating the broad dissemination of peer-reviewed scholarly publications resulting from federally funded scientific research is a means for government, at a minimum, to stem the loss in productivity from scientific results that do not currently reach the intended audience due to extreme pricing models and also to expand economic growth in sectors that benefit from such research – health, agriculture, and the environment, among others.

The most effective way to accomplish this broad dissemination in the current research climate is through a national policy, or a coordinated set of national policies, at the federal agency levels, that ensures public access to research through the archiving of peer-reviewed publications resulting from taxpayer-funded research in open access repositories, the so-called *Green Open Access* model. This model will eliminate losses on taxpayer investment in research and will accelerate research time.

The costs and benefits of such policies have been studied. Here below is evidence from two studies:

1. John Houghton, Bruce Rasmussen, Peter Sheehan, et al., *Report Summary: Economic Implications of Alternative Scholarly Publishing Models: Exploring the Costs and Benefits* (London and Bristol, Joint Information Systems Committee, 2009), http://www.jisc.ac.uk/media/documents/publications/summary-economicoa.pdf.

> At these costs, open access publishing would be around £813 per article cheaper than toll access publishing, and open access self-archiving with overlay services around £1,180 per article cheaper. (p. 9)
>
> For UK higher education, these journal article cost differences would have amounted to savings of around £80 million per annum circa 2007 from a shift from subscription access to open access publishing, and £116 million from a shift from subscription access to open access self-archiving with overlay services. (p.10)

2. Malcolm Getz, *Open Access Scholarly Publishing in Economic Perspective* (Nashville, TN, Department of Economics Vanderbilt University, 2004), http:www.vanderbilt.edu/econ/wparchive/workpaper/vu04-w14.pdf.

   > On the positive side, a number of initiatives are at play that may lead to a shift to quality-assured, open-access publishing by the non-profit sector. By the estimated values in table 1, current costs total $6.48 M for a typical library to subscribe to, process, and store its serials compared to a potential cost of $4.142 M with quality-assured, open-access journals published by non-profit publishers. The average US research library might save $2.3 million per year with perhaps $230 M saved per year among all US research libraries. (p. 39)

As the European Union invests in a super-governmental infrastructure – OpenAIRE,[1] as the UK pursues its recently announced national strategy to require "[f]ree and open access to taxpayer-funded research...",[2] and as governments in developing nations across the globe discuss long-term stewardship and public access to scholarly publications, the U.S. too needs to take action to ensure that education and research continue to enable "the creation, access and use of information, knowledge and culture for human development and the exercise of freedoms [that] is widely acknowledged as critical."[3]

Section 6.6 of the UK report on *Innovation and Research Strategy for Growth* states:

> The Government, in line with our overarching commitment to transparency and open data, is committed to ensuring that publicly-funded research should be accessible free of charge. Free and open access to taxpayer-funded research offers significant social and economic benefits by spreading knowledge, raising the prestige of UK research and encouraging technology transfer. At the moment, such research is often difficult to find and expensive to access. This can defeat the original purpose of taxpayer-funded academic research and limits understanding and innovation…. But we need to go much further if, as a nation, we are to gain the full potential benefits of publicly-funded research.[4]

Agencies can accomplish this by partnering with leading institutions, like research university libraries, that have already invested significantly in developing open repositories for the intellectual property of

---

[1] "OpenAIRE: Open Access Infrastructure for Research in Europe," http://www.openaire.eu/.
[2] Secretary of State for Business, Innovation and Skills, *Innovation and Research Strategy for Growth* (London, Department for Business, Innovation and Skills, 2011), 76, http://www.bis.gov.uk/assets/biscore/innovation/docs/i/11-1387-innovation-and-research-strategy-for-growth.pdf.
[3] Sisule F. Musungu, *Using Copyright to Promote Access to Information and Creative Content. Education and Research (Part I),* WIPO/CR/WK/GE/11/2 (Geneva, World Intellectual Property Organization, 2011), 3, http://www.wipo.int/edocs/mdocs/copyright/en/wipo_cr_wk_ge_11/wipo_cr_wk_ge_11_2.pdf.
[4] *Innovation and Research Strategy for Growth,*76.

their parent organizations. Such repositories capture intellectual property in all disciplines and formats, for large and small science, and without regard to funding source. This comprehensive approach is necessary to provide support for modern interdisciplinary research and would be enriched by timely inclusion of peer-reviewed scholarly publications resulting from federally funded research.

University library systems, furthermore, have a long history of collaboration that could be the basis for development of a national access infrastructure. Agency leadership could leverage individual repository achievements into a comprehensive system that would certainly include, but not be limited to, publications resulting from federally funded research.

We would like to clarify our view by noting that the growing of existing markets and creation of new markets relates to access and analysis of peer-reviewed publications, but not to the publications themselves, which should not constitute the markets. Our work going forward should not be about markets in the peer-reviewed scholarship but about markets in scientific and technological sectors made possible through use of federally funded scientific research that is publically accessible without fee, and not monetized downstream. Included are markets that develop practical applications of research results as well as such new markets as those that organize, mine, and present research results for downstream use. Action Science Explorer,[5] a visualization tool developed at the University of Maryland, is an example of a text mining application that seeks to aid researchers seeking understanding of developments in areas that are complicated by multidisciplinary resources. The usefulness of this or any similar tool is proportional to the access it has to resources.

Public access in the form of *Green Open Access* to peer-reviewed scholarly publications resulting from federally funded research is the best way to maximize U.S. economic growth and improve the productivity of the American scientific enterprise. Public access to federally funded research is fair to all stakeholders in society, it best reflects the democratic values upon which the U.S. is founded, it eliminates economic waste, and it encourages research and development to ensure the nation's competitiveness and the well-being of its citizens into the future.

(2) What specific steps can be taken to protect the intellectual property interests of publishers, scientists, Federal agencies, and other stakeholders involved with the publication and dissemination of peer-reviewed scholarly publications resulting from federally funded scientific research? Conversely, are there policies that should not be adopted with respect to public access to peer-reviewed scholarly publications so as not to undermine any intellectual property rights of publishers, scientists, Federal agencies, and other stakeholders?

U.S. copyright law protects the intellectual property interests of publishers, scientists, Federal agencies, and other stakeholders involved with the publication and dissemination of peer-reviewed scholarly publications. It protects these works whether they are publicly accessible without a fee, inaccessible, or available for purchase. The best way to protect the intellectual property interests of publishers, scientists, Federal agencies, and other stakeholders is to maintain federal policies that enable all stakeholders to benefit from their investment in scholarship and research within the copyright system, and to maintain balanced copyright law that includes a full array of limitations and exceptions to support scholarship and research in the digital environment.

Policies modeled on the NIH Public Access Policy effectively balance IP interests by ensuring that authors, as copyright holders, retain the rights needed to allow public access to their works through a non-

---

[5] "Action Science Explorer: Tools for Rapid Understanding of Scientific Literature," http://www.cs.umd.edu/hcil/ase/.

exclusive license; by leaving authors free to contract with publishers on a non-exclusive basis; by ensuring that conditions imposed by Federal agencies surrounding access to research results are met; and by recognizing the right of the public to access and make use of peer-reviewed scholarly publications, to which they have materially contributed, as permitted by copyright law.

(3) What are the pros and cons of centralized and decentralized approaches to managing public access to peer-reviewed scholarly publications that result from federally funded research in terms of interoperability, search, development of analytic tools, and other scientific and commercial opportunities? Are there reasons why a Federal agency (or agencies) should maintain custody of all published content, and are there ways that the government can ensure long-term stewardship if content is distributed across multiple private sources?

We agree with Stevan Harnad, who pointed out in his submission to this RFI, that it is important to distinguish between the problem of access to peer-reviewed research and the problem of digital preservation and storage of the published intellectual and cultural record. Preservation of the published versions of research output (publishers' versions of record in scholarly journals) may best be accomplished within national digital preservation infrastructures, while retention of peer-reviewed final drafts is best accomplished through institutional repositories that may be harvested by centralized subject-based repositories.

To the issue of preservation, the Library of Congress is leading the National Digital Information Infrastructure and Preservation Program, which recognizes the importance of a collaborative, "multiphased plan to collect and preserve a broad spectrum of high-value digital content, with special attention to the needs of the public policy, education and research, and cultural heritage communities."[6] Initiatives such as LOCKSS[7] and Portico,[8] along with research library repositories, could be leveraged to provide the infrastructure for this system.

A parallel federal initiative is needed for public access to peer-reviewed scholarly publications resulting from federally funded research. Federal agencies do not need to maintain custody of all published content, but the government has an important role to play in ensuring long-term stewardship by setting policy and standards, and by providing sufficient and ongoing funding.

(4) Are there models or new ideas for public-private partnerships that take advantage of existing publisher archives and encourage innovation in accessibility and interoperability, while ensuring long-term stewardship of the results of federally funded research?

Public access to research materials in scientific as well as other disciplines is being provided primarily by individual authors, by research institutions, and by university libraries. As trusted institutions dedicated to permanent preservation of and access to the scholarly and historical record, libraries are uniquely positioned to ensure long-term stewardship of the results of federally funded research.

We see opportunity for public-private partnerships where private foundations dedicated to improving the well-being of our citizenry are willing to support development of public access to scholarly publications.

---

[6] *Preserving Our Digital Heritage: The National Digital Information Infrastructure and Preservation Program 2010 Report* (Washington, DC, Library of Congress, 2011), 2, http://www.digitalpreservation.gov/multimedia/documents/NDIIPP2010Report_Post.pdf.

[7] "LOCKSS," http://www.lockss.org/.

[8] "Portico," http://www.portico.org/digital-preservation/.

Universities and research libraries are well positioned to leverage the essential agency-provided funding and infrastructure support referenced earlier to attract such private funds.

There is some opportunity for public-private partnerships involving publishers. Publishers have never been primarily involved in long-term preservation, and it is unlikely that they could or should assume that role now. To bridge the gap from publisher to public access, PubMedCentral, the primary effort of the U.S. government to provide public access to publications resulting from federally funded research, required new law. However, partnerships with publishers could be developed to preserve publishers' archives as part of the infrastructure for federally funded scholarly publications.

Accessibility will best be accomplished through the parallel effort of open access institutional and subject repositories linked in a national open access infrastructure designed for interoperability.

(5) What steps can be taken by Federal agencies, publishers, and/or scholarly and professional societies to encourage interoperable search, discovery, and analysis capacity across disciplines and archives? What are the minimum core metadata for scholarly publications that must be made available to the public to allow such capabilities? How should Federal agencies make certain that such minimum core metadata associated with peer-reviewed publications resulting from federally funded scientific research are publicly available to ensure that these publications can be easily found and linked to Federal science funding?

As a first priority, Federal agencies need to mandate that peer-reviewed scholarly publications resulting from federally funded research be publicly accessible. This would best be done through new national policy like that recently announced by the UK and being pursued by European Union governments, among others, as cited above. The second priority would be for collaborative selection or development, with government leading as convenor of recognized experts, of the tools that will allow search, discovery, and analysis. Government funding will be key.

Core metadata needs to include information about the object itself (descriptive metadata), the physical piece from which a digital file is created (source metadata), the digital file that is created from the source resource (technical metadata), and copyright information and any rights restrictions that may exist that pertain to the collection (rights metadata). It must enable users to find, identify, select, obtain, and understand the appropriate resource. A core metadata schema needs to be flexible, extensible, and interoperable with other schema.

We offer as a model the metadata schema developed by the Rutgers University Libraries as part of the workflow management system for our institutional repository. The metadata schema is available to everyone as part of OpenWMS: Workflow Management System for Digital Objects. The WMS is a platform-independent, open source, web-accessible system that can be used as a standalone application or integrated with other repository architectures by a wide range of organizations. It provides a complete metadata creation system for analog and digital materials, with services to ingest objects and metadata into a Fedora repository and to export these objects and metadata, individually and in bulk in METS/XML Wrapper. Features include:

- Event-based data model for management and rights documentation
- Capability to customize the look and feel of the metadata input and to add default values to data elements
- Ability to customize and add vocabularies to data elements
- Unicode and CJK vernacular character support
- Mapping and import metadata and digital files from standard and in-house formats

- Export digital object in METS/XML Wrapper

Further information is available here: http://rucore.libraries.rutgers.edu/open/projects/openwms/.

(6) How can Federal agencies that fund science maximize the benefit of public access policies to U.S. taxpayers, and their investment in the peer-reviewed literature, while minimizing burden and costs for stakeholders, including awardee institutions, scientists, publishers, Federal agencies, and libraries?

Federal agencies that fund science can maximize the benefit to taxpayers by broadly promoting a national research infrastructure, such as EU governments are demonstrating in their commitment to OpenAIRE, the open access infrastructure for research in Europe, referenced earlier. Federal agencies can educate taxpayers about the availability of research, and on results of taxpayer investment in scientific research. Libraries can also contribute to this effort.

Federal agencies that fund science can minimize the burden and costs for stakeholders by providing a technical and policy infrastructure, with a common platform, and consistent policies and requirements, mandated by federal law. Integration of grants management systems would also be beneficial.

(7) Besides scholarly journal articles, should other types of peer-reviewed publications resulting from federally funded research, such as book chapters and conference proceedings, be covered by these public access policies?

Any taxpayer-funded research should be covered by these public access policies. From the standpoint of accountability to taxpayers for their investment in research, there is no need to distinguish between types of publication. From the standpoint of researchers' requirements, all categories would be beneficial.

Since a researcher may continue to build on the results of his or her federally funded research throughout his or her career, only those peer-reviewed publications that directly result from that funding should be mandated for inclusion. Downstream publications should be requested but not mandated.

(8) What is the appropriate embargo period after publication before the public is granted free access to the full content of peer-reviewed scholarly publications resulting from federally funded research? Please describe the empirical basis for the recommended embargo period. Analyses that weigh public and private benefits and account for external market factors, such as competition, price changes, library budgets, and other factors, will be particularly useful. Are there evidence-based arguments that can be made that the delay period should be different for specific disciplines or types of publications?

Studies have shown that access to research results in various intellectual property regimes offers the greatest opportunity for the greatest good.

In "Minerva Unbound: Knowledge Stocks, Knowledge Flows and New Knowledge Production," Lynne Zucker and her colleagues demonstrate that:

> Regional growth of new knowledge in nanotechnology, as measured by counts of articles and patents in the open-access digital library NanoBank, is shown to be positively affected both by the size of existing regional stocks of recorded knowledge in all scientific fields, and the extent to which tacit knowledge in all fields flows between institutions of different organizational types. The level of federal funding has a large, robust impact on both publication and patenting. The data provide support for the cumulative advantage model of knowledge production, and for

> ongoing efforts to institutionalize channels through which cross-organizational collaboration may be achieved.[9]

In *Open-Access Scholarly Publishing in Economic Perspective*, Malcolm Getz concludes:

> … that substantial cost savings to universities are possible with open-access distribution of quality-assured journals by not-for-profit publishers whose rates reflect cost rather than each university's ability to pay. Open-access to quality-assured materials via the Internet will increase the use of the materials and expand the influence of scholarship worldwide.[10]

A primary goal of public access is to speed up and enlarge our capacity for discovery and application of new knowledge. Logically, that goal is best enhanced by access as early as possible. At the same time, there will always be legitimate reasons for embargoes. These restrictions should be respected. The most common embargo periods are 6 months and 12 months. Shorter embargoes should be encouraged.

Submitted by:

Marianne I. Gaunt
Vice President for Information Services and University Librarian

Jeanne E. Boyle
Associate University Librarian for Planning and Organizational Research

Janice T. Pilch
Copyright and Licensing Librarian

---

9 Lynne G. Zucker, Michael R. Darby, Jonathan Furner, et al., "Minerva Unbound: Knowledge Stocks, Knowledge Flows and New Knowledge Production, Research Policy 36 (2007): 850.
10 Malcolm Getz, Open-Access Scholarly Publishing in Economic Perspective, Vanderbilt University Department of Economics Working Paper No. 04-W14, http://www.vanderbilt.edu/Econ/wparchive/workpaper/vu04-w14.pdf, 2.