Response to the White House Office of Science and Technology Policy's *Request for Information: Public Access to Peer-Reviewed Scholarly Publications Resulting From Federally Funded Research* submitted by the Canadian Association of Research Libraries.

20th of December, 2011

CARL is the leadership organization for the Canadian research library community. The Association's members include Canada's major academic research libraries.

CARL ABRC

The Canadian Association of Research Libraries (CARL) welcomes the opportunity to submit a response to the White House Office for Science and Technology Policy's Request for Information: Public Access to Peer-reviewed Scholarly Publications Resulting from Federally Funded Research. CARL is the leadership organization for the Canadian research library community. The Association's members are the 29 major academic research libraries across Canada together with Library and Archives Canada, the Canada Institute for Scientific and Technical Information (CISTI), and the Library of Parliament.

Across the globe, demands are growing to make publicly funded research more readily available to academia, civil society and industry. Any further steps the White House takes to ensure public access to the results of publicly funded research will lead to more efficient linkages between science and innovation and are highly commendable. While there are already open access policies in place at funding agencies in Canada and across the world, a US policy requiring public access to federally funded research results would likely have a significant global impact and pave the way other countries to follow. CARL's response to the Request for Information is as follows:

(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 result from federally funded scientific research?***

A key step that agencies can take to stimulate new markets related to peer-reviewed publications is to implement policies requiring free and unrestricted online access to the results of federally funded research. Such policies would enable the development of numerous value-added services to be built on top of this content. While it is impossible to envision all of the types of services that may be created, we can anticipate a variety of tools that will evolve as a result-such as search engines, disciplinary portals, data mining, as well as repository services.

In addition, small and medium sized businesses represent a significant portion of library users in Canada and elsewhere, reflecting the importance of scholarly literature to business development. Access to the full corpus of scholarly articles would contribute to the development of new products and services based on the results of 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?

Open access policies will result in greater usage of the research results and will contribute to maximizing the government's investment in research. This statement is validated through a large number of studies that have found that open access increases usage and citation of peer-reviewed articles by researchers.[1] [2] Statistics from PubMed Central, the open access repository for biomedical literature, shows impressively high levels of usage of its open access papers.

---

[1] "The effect of open access and downloads ('hits') on citation impact: a bibliography of studies". (2011) http://opcit.eprints.org/oacitation-biblio.html

[2] Philip M Davis, Bruce V Lewenstein, Daniel H Simon, James G Booth and Mathew J L Connolly. (2008) Open access publishing, article downloads, and citations: randomised controlled trial. BMJ 2008;337: a568. http://www.bmj.com/content/337/bmj.a568

Between March 2008 and March 2010, for example, the monthly number of articles retrieved from PubMed Central doubled from 10 million to 20 million.[3]

The benefits of open access policies are not limited to the research community; industry, public policy and the general public all stand to benefit and use the information contained within the peer-reviewed literature. The majority of users of PubMed Central are from outside the research community. According to David J. Lipman, Director, National Center for Biotechnology Information the users of PubMed Central **"**represent a mix of people from the education and business sectors, as well as private citizens." He estimates that "approximately 25% of our users are from universities, 40% are private citizens or those using personal Internet accounts, and 17% are from companies (the remainder consists of government users or others). These kinds of numbers support the notion that PubMed Central has become a broad-based repository for researchers, students, clinicians, entrepreneurs, patients and their families."[4]

In addition, economic analyses have also shown that the open access model will cost less at the national level than the existing subscription-based system, leading to a more efficient and productive system for disseminating and exploiting the results of publicly funded research.[5] [6] In particular, these studies have found that the benefits of open access via repository self-archiving would result in the greatest cost savings and accrue significant benefits.

What are the relative costs and benefits of such policies?

There are start-up and ongoing costs associated with open access policies, which include repository development and management, monitoring compliance, and so on. However, these costs represent a relatively small percentage of the large public investment already being made in research. The start-up costs for PubMed Central "were about $500,000. Annual operating costs for the system, including ingest of articles, refinement of the submission system and search tools, staffing of a help desk and a central coordinating office for NIH, are approximately $3.5-$4.0 million per year. This represents a small fraction of NIH's budget authority of more than $30 billion per year."[7]

---

[3] Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives. United States House of Representatives. http://www.hhs.gov/asl/testify/2010/07/t20100729c.html

[4] Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives. United States House of Representatives. http://www.hhs.gov/asl/testify/2010/07/t20100729c.html

[5] John Houghton, Bruce Rasmussen and Peter Sheehan. (2010) "Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs". http://www.arl.org/sparc/publications/papers/vuFRPAA/index.shtml

[6] John Houghton. (2009) "Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark". http://www.knowledge-exchange.info/Default.aspx?ID=316

[7] Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives.

In addition, many universities in the US (as elsewhere) already have institutional repositories.[8] These repositories are usually managed by the university library and have dedicated staff to support the deposit of research articles.

As mentioned earlier, economic analyses have found that there are significant overall savings in an open access model. Conservative estimates suggest such policies would result in an estimated five-fold increase on the return on public funds invested in research. Specifically, the benefits of an open access policy, such as has been implemented by the NIH, are estimated to be at least eight times larger than the costs. Extending this type of policy to all US scientific agencies has been estimated to result in a net value of $1.5 billion, of which an estimated 60% would accrue directly to the U.S. economy.[9]

Open access policies for research articles are directly aligned with the Open Government initiative which is seeking to redress a culture " where information is locked up, taxpayer dollars disappear without a trace"[10]. Other key benefits to these policies are that they will contribute to a more transparent government, the development of scientifically sound policy decision-making, as well as serving to increase federal agency accountability.

What type of access to these publications is required to maximize U.S. economic growth and improve the productivity of the American scientific enterprise?

Full and open access to the results of research as published in peer-reviewed scholarly/scientific journals is the ideal type of access. Any restrictions that place limits on the usage of the results of scientific research will dilute the impact of the policy. In order to derive additional benefits via computational analyses and value added tools, open access articles should be free to be crawled and indexed by search engines.

(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?***

It is possible to implement public access policies for federally funded research while fully respecting and working within current copyright frameworks. Creative Commons licenses could

---

United States House of Representatives. http://www.hhs.gov/asl/testify/2010/07/t20100729c.html

[8] 2 Clifford A. Lynch and Joan K. Lippincott, "Institutional Repository Deployment in the United States as of Early 2005," D-Lib Magazine 11, 9 (2005), http://www.dlib.org/dlib/september05/lynch/09lynch.html

[9] Lipman, D. J., M.D. (2011). *Testimony on Public Access to Federally-Funded Research* Committee on Oversight and Governmental Reform, Subcommittee on Information Policy, Census and National Archives. United States House of Representatives. http://www.hhs.gov/asl/testify/2010/07/t20100729c.html

[10] About Open Government. (2011). http://www.whitehouse.gov/open/about

be used in this context which would allow authors and publishers to retain their copyright but mark their work with the re-use rights they feel are most appropriate, while still receiving credit for their work.

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?

CARL strongly advocates for the broadest possible access to knowledge in all fields, and would recommend against the establishment of policies that place unnecessary restrictions on access or reuse of content, such as policies which include lengthy embargo times, or allow "read only" access to scholarly articles.

(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?***

The most effective approach (centralized or decentralized) will vary depending on current practices and infrastructure in any given sector. For physics and biomedical literature, for example, which already have established traditions for depositing into discipline-based repositories, the centralized approach may be most appropriate. For other fields, there are significant benefits to adopting a decentralized approach, making use of the existing institutional repository network based at the universities. University-based repositories already have the capacity and mandate to undertake the long-term preservation and management of scholarly knowledge. The most efficient approach would be to extend the current distributed network environment, which is composed of both institution and discipline based repositories. While it may be easier to monitor compliance via centralized repositories, there are ways to accommodate funders' need to monitor compliance via metadata for specific research grant numbers.

Regardless of approach, it is critical that open access repositories adhere to common standards and metadata that will ensure article level interoperability across repository and across national boundaries. The research enterprise is truly global and it will be critical that US-funded papers are interoperable with the global corpus of research knowledge. The Confederation of Open Access Repositories (COAR), a large international organization representing repository organizations (of which CARL is a member) states, "The real value of repositories lies in the potential to interconnect them to create a network of repositories, a network that can provide unified access to research outputs and be (re-) used by machines and researchers."[11]

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?

[11] Confederation of Open Access Repositories. (2011). "The Case for Interoperability for Open Access Repositories". http://www.coar-repositories.org/files/COAR_Interoperability_Briefing.pdf

The federal government has an important role to play in ensuring that the nation's research outputs are preserved and available into the future. Private industry, which is motivated by profits, has no incentive to preserve content which it can no longer "sell" for access. In addition, there is a risk that commercial enterprises engaged in preserving published content would change business models and go out of business.

A distributed approach that includes a broad range of players including research libraries and other long-term stable institutions may be most likely to succeed. In addition, some redundancy is desired from a preservation point of view – especially with born digital content – by way central, subject and institutional open digital repositories as to where content is archived and accessed.

(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?***

In terms of storage, organizations such as Portico, LOCKSS, and DuraCloud are working with publishers and libraries to assist with the preservation of digital content. Indexing and abstracting services already provide value added services as do journal hosting and repository development services. Indeed, there are numerous models for public private partnerships that should be encouraged with the proviso that these services do not inhibit the easy and open access to research articles.

(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?***

Interoperability at the article level is crucial for the discovery and development of value added services, such as harvesters, which are used to aggregate scholarly papers. Metadata should facilitate both reuse and analysis of published research including machine-readable and machine-interoperable analyses. While it is likely that metadata standards will evolve, we recommend for now the adoption of Open Archives Institute Protocol for Metadata Harvesting (OAI-PMH), which is the international standard currently being applied across repositories. OAI-PMH will ensures a minimum level of interoperability across the global corpus of research literature.

Repository interoperability could also be further enabled and enhanced by the adoption of a few additional protocols such as the Open Archives Initiative Object Reuse and Exchange (OAI-ORE), which serves to facilitate content sharing among systems and the Simple Web-service Offering Repository Deposit (SWORD) that permits researchers to deposit an article through a single interface while also routing the same item to multiple repositories. As well, the Confederation of Open Access Repositories (COAR) recommends the use of the Resource Description Framework (RDF) that expresses digital objects relationships rendered in a machine-understandable way "allowing machines to create sophisticated services over the global representation of knowledge

distributed across repositories and other systems, to make cross-discipline connections, and to combine disparate findings to arrive at new insights."[12]

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?

It is in the interest of all stakeholders to ensure that there is widespread use of standards and metadata. Agencies should work with institutions, universities and, publishers and repository communities to ensure that the appropriate metadata standards are adhered to.

(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?***

The system as it is now is simply not sustainable. Journal prices continue to increase at rates far above the standard of living, to the point that even the most well-endowed research library cannot afford to provide access to all of the content requested by its faculty and students. The implementation of an open access mandate would significantly reduce the systemic costs of access (as discussed earlier in response to question 1).

One way of minimizing the burden for all would be to ensure that there is consistency in the policy language across all agencies. Uniformity in article deposit requirements and procedures across all funding agencies will reduce costs and confusion and ensure a high rate of compliance with open access mandates. In addition, there are a growing number of tools being developed internationally to minimize efforts for depositing articles into repositories, such as the SWORD protocol (discussed above). Agencies can further encourage the development of additional tools and services for depositing and act as facilitators for publishers and repository organizations to work together.

(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?***

The results of publicly funded research in all forms including monographs, conference proceedings and research data should be covered by public access policies. Making all of these products of research available would contribute greatly to maximizing the many billions of dollars that the United States invests each year in research. However, given the different infrastructure and requirements necessary for collecting and providing access to these different types of research output, it may not be appropriate to include all of these under a single policy.

[12] Confederation of Open Access Repositories. (2011). "The Case for Interoperability for Open Access Repositories". http://www.coar-repositories.org/files/COAR_Interoperability_Briefing.pdf

(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?***

In order to maximize the benefit of open access policies, CARL supports the shortest possible embargo periods from 0 to 12 months, maximum. Publishers call for embargo periods as a way of protecting their subscription base. However, there is no evidence that archiving in open access repositories has a negative effect on journal subscription rates. For example, arXiv, the online repository for papers in physics and other fields, has been providing open access to a very high percentage of papers in those fields, but has not resulted in journal cancellations. In fact, the established journals in those fields have continued to thrive.

Canadian Association of Research Libraries (CARL)
600-350 Albert Street
Ottawa, Ontario, Canada K1R 1B1
Telephone: 613.482.9344
E-mail: info@carl-abrc.ca
Website: www.carl-abrc.ca

University Library Member Institutions

University of Alberta
University of British Columbia
Brock University
University of Calgary
Carleton University
Concordia University
Dalhousie University
University of Guelph
Université Laval
University of Manitoba
McGill University
McMaster University
Memorial University of Newfoundland
Université de Montréal
University of New Brunswick
University of Ottawa
Université du Québec à Montréal
Queen's University
University of Regina
Ryerson University
University of Saskatchewan
Université de Sherbrooke
Simon Fraser University
University of Toronto
University of Victoria
University of Waterloo
University of Western Ontario
University of Windsor
York University