

[Assigned ID #]
[Assigned Entry date]
Nick Shockey; nick@arl.org
The Right to Research Coalition
Washington, DC

I am pleased to offer the following comments on behalf of the Right to Research Coalition. Founded by students in the summer of 2009, the Right to Research Coalition is an international alliance of undergraduate and graduate student organizations, representing nearly 7 million students, that promotes Open Access to scholarship. The Right to Research Coalition believes no student should be denied access to the published articles they need because they or their institution cannot afford access. The coalition works to educate the next generation of scholars and researchers about Open Access and to advocate for policies at the campus, national, and international levels that expand access to the results of research.

A full list of the Right to Research Coalition's members is available at the end of this document.

[Question 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? 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?

[Comment 1]

All peer-reviewed articles resulting from federally funded research should be open-access. That is, these articles should be made freely available immediately upon publication with full reuse rights, so users can text mine, data mine, compute on, and create derivative works – including further research – from the articles without commercial restriction.

Open access to federally funded research would greatly improve the resources available to students, at both the undergraduate and graduate levels, to achieve a complete, up-to-date education. Students' educations depend on access to the research literature. These articles are quite literally the building blocks of an education in any discipline; from its core, all the way out to the cutting edge.

Unfortunately, because of the often-high price of journal subscriptions – 15 academic disciplines have an average price per title in excess of \$1,000 per year¹ – students and the professors who teach them are often left without access to the research they need for a complete, up-to-date education. Not only do students routinely run into access barriers when researching for a paper, for a class, or simply to get a better understanding of a given issue, but students’ professors also run into those same barriers and cannot bring the most cutting-edge research into the classroom. Speaking to this point, Dr. Gary Ward, former PubMed Central Advisory Committee Chair and a researcher at the University of Vermont, has said, “In my role as educator, I often find myself teaching my graduate and medical students what I have access to rather than what they most need to know. Just as one example, in a recent lecture I was preparing for our medical students on how drugs can get across the barrier between the blood and the brain to treat neurological disease, I was only able to access about two thirds of the articles that I needed in order to make sure that I was providing these budding young doctors with everything they needed to know about the subject. I can tell you that’s extremely frustrating to me as an educator and it’s clearly not in the best interests of my students.”²

Because the federal government funds a large portion of all published research, a strong open access policy at the federal level would vastly expand the resources available for students to get a complete, up-to-date education. This improvement in education would translate into immediate and persistent economic benefits to the United States economy. As our economy continues to shift toward innovative, research-based sectors like biotechnology and alternative energies, the companies driving our economy will rely on a highly educated and trained workforce. By providing students with immediate access to cutting-edge research, federal agencies can help ensure students are ready to hit the ground running after graduation and put their education to use immediately, rather than having to play catch up.

Similarly, students’ library cards expire at graduation; however, with the current weak economy, it is increasingly common for students to take months or even years to find jobs in their chosen fields. With no institution to pay for journal subscriptions on their behalf, recent graduates lose all access to the subscription-based academic literature and have a limited ability to stay current in their discipline. A strong federal open access policy would open a wealth of cutting-edge research, enabling graduates to maintain an up-to-date understanding of their field and contribute more quickly once hired.

¹ Bosch, et al., *Periodicals Price Survey 2011: Under Pressure, Times Are Changing*. Library Journal. April 2011. Available at http://www.libraryjournal.com/lj/newslettersnewsletterbucketljxpress/890009-441/periodicals_price_survey_2011_under.html.csp.

² Dr. Ward’s full quote can be found at http://www.taxpayeraccess.org/issues/frpaa/frpaa_resources/press-conference-congressman-doyle-to-address-new-.shtml.

Difficulties in accessing the research literature disproportionately impact students at smaller and less wealthy institutions – especially community colleges – which cannot afford the multi-million dollar library budgets required to access large portions of the scientific and scholarly record.³ As our 21st century economy increasingly requires highly skilled workers, community colleges will become more and more essential in providing American businesses with the advanced workforce required for economic competitiveness. With strong open access policies, federal agencies could provide these institutions, which would otherwise have very limited access to cutting-edge research, the ability to incorporate the most up-to-date information into their students' educations. This would help level the playing field between students at less wealthy and wealthier institutions, and have a persistent positive effect on the skill level of the American workforce.

Beyond students, a federal open access policy would pay real dividends to the United States economy and the advancement of scientific research. A useful analogy can be found in the Human Genome Project (HGP), which sequenced the entire human genome and, critically, made the data immediately, openly available for anyone to use without commercial restriction. By any measure, the HGP was an incredible success in providing a return on taxpayer investment, with a \$5.6 billion federal investment yielding \$796 billion in economic output, over \$6 billion in federal, state, and local taxes, and over 3.8 million job-years of employment to date.⁴ Research has shown that the immediate, open availability of HGP data played a significant role in boosting this economic return. One study comparing the use of similar, but closed data from a parallel sequencing project run by the Celera Corporation found “robust evidence that the package of short-term IP used by Celera has been associated with reductions on the order of 30 percent in subsequent gene-level scientific research and product development outcomes.”⁵ There are strong reasons to believe a federal open access policy would lead to a similar increase in return on taxpayer investment in research.

Following the Human Genome Project's example, making articles resulting from federally funded research immediately and openly available would allow them to be utilized and built upon more quickly and by a larger, more diverse group of researchers and corporations. Immediate availability would shorten research cycles by providing researchers with faster access to breakthroughs, and would accelerate the advancement of science, decreasing the amount of time taken for businesses to

³ To get a sense of the size and variation in library journal subscription budgets, see the Association of Research Libraries' Statistics Report from 2008-2009, p. 40-46. Available at <http://www.arl.org/bm~doc/arlstat09.pdf>.

⁴ *Economic Impact of the Human Genome Project*. Battelle Technology Partnership Practice. May 2011. Available at <http://www.battelle.org/publications/humangenomeproject.pdf>.

⁵ Williams, Heidi., *Intellectual property rights and innovation: Evidence from the human genome*. National Bureau of Economic Research Working Paper Series. July 2010. p. 27. Available at http://www.nber.org/~heidiw/papers/5_12_10a_hlw.pdf.

translate theoretical breakthroughs into new products and services. Faster commercialization will, in turn, boost American economic growth and ultimately create new jobs across the economy as innovation can happen more quickly and with less restriction. Similarly, by making the full body of federally funded science openly available to all, federal agencies can greatly expand the number and diversity of those engaged in follow-on research. The expensive nature of journal subscriptions artificially and arbitrarily limits researchers' access to the journals they can afford rather than what they actually need. An open access policy would not only increase readership among an article's intended audience, but it would also lead to an increased likelihood the article would reach unintended readers in adjacent or seemingly unrelated disciplines. This increased diversity promotes additional paths of follow-on research across scientific domains, leading to breakthrough that would not have occurred without an article's availability to unintended readers.⁶

In addition to unintended readers, full open access allows machines as an entire new class of reader to use the literature to its fullest extent. With approximately 1,350,000 papers published annually,⁷ no single person can hope to read even a tiny fraction of all published articles. We will increasingly rely on computational text and data mining to get an overall picture of the state of a discipline and uncover trends, connections, and new research pathways that would otherwise remain hidden. These computational processes can identify relevant articles and enable scientists to work more efficiently, improving scientific productivity. These services also represent a new layer of potential commercialization on top of public databases, like PubMed Central, that is only possible with open licensing and full reuse rights. To be computed on to their fullest extent, articles must be available in a machine-readable format – XML, not proprietary PDFs – and come coupled with the reuse rights necessary to be crawled by computers and for businesses to sell services based on such computation.

One illustration of the value that can only be created from an open repository is the winner of the recent Binary Battle contest hosted by the Public Library of Science (PLOS) and Mendeley, a reference manager and social network for researchers. The winning application, OpenSNP, takes genomic data – either yours or other data that you upload – and "find[s] the latest relevant research and let[s] scientists discover

⁶ An analogous case of openness promoting the volume and diversity of follow-on research in the area of research materials can be found in Murray, et al., *Of Mice and Academics: Examining the Effect of Openness on Innovation*. National Bureau of Economic Research Working Papers. October 2008. Available at <http://www.hbs.edu/units/tom/seminars/2007/docs/Of%20Mice%20and%20Academics%20Ster n.pdf>.

⁷ Björk, et al., *Global annual volume of peer reviewed scholarly articles and the share available via different Open Access options*. Proceedings ELPUB2008 Conference on Electronic Publishing. June 2008. Available at <http://oacs.shh.fi/publications/elpub-2008.pdf>.

new genetic associations."⁸ This application is a great example of how text and data mining can uncover new connections in a way that is only possible when research is open.

It is important to point out that the aforementioned economic benefits of a federal open access policy only represent uses we can currently imagine. Opening this vast literature – not only to a larger audience of readers, but also for unrestricted use – will undoubtedly pay dividends in ways currently unimaginable.

The benefits of a federal open access policy would far exceed the costs. According to a study done last year by the Center for Strategic Economic Studies, an expansion of the NIH public access policy to cover all federally funded research with a six-month embargo period would provide a 500% return on investment to the United States government.⁹ Such a policy would also generate benefits eight times greater than costs, resulting in a net present value gain worth approximately \$1.5 billion.¹⁰ The impact could be even greater with a shorter embargo period or immediate open access. Furthermore, the NIH policy has a proven track record of cost-effectiveness over the past three years. The NIH spends approximately \$4 million per year to make the articles covered by its policy, approximately 90,000 annually, available through PubMed Central – a total of roughly 1/100th of 1% of the NIH's \$30 billion per year operating budget.¹¹

[Question 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?

[Comment 2]

Federal agencies should require articles resulting from federally funded research to be made available under a fully open license that allows the public to freely use, remix, revise, and redistribute the research without commercial restriction, such as the Creative Commons Attribution-Only ("CC BY") license.¹² Only by adopting fully open licensing will we maximize our collective investment in research and allow it to be used, reused, and built upon to its fullest possible extent. Such an approach

⁸ *Winners of the first Binary Battle Apps for Science Contest*. Mendeley Blog, November 2011. Available at <http://www.mendeley.com/blog/design-research-tools/winners-of-the-first-binary-battle-apps-for-science-contest>.

⁹ Houghton, et al., *Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs*. July 2010. p. 7-8. Available at <http://www.arl.org/sparc/bm~doc/vufrpaa.pdf>.

¹⁰ Ibid.

¹¹ Letter from Dr. Francis Collins, Director of the NIH, to Representative Joseph Pitts. December 2011. Available at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

¹² Creative Commons Attribution 3.0 Unported (CC BY 3.0) License summary available at <http://creativecommons.org/licenses/by/3.0>.

would adequately protect authors' interests by requiring citation – the primary mechanism by which researchers build reputation within their field – while allowing the widest possible distribution and use.

The government can implement a policy requiring open licensing fully within the current system of copyright. Using the same mechanism employed by the NIH policy, in which researchers consent at the time of grant acceptance to make their work freely accessible in PubMed Central, agencies can require articles resulting from their funding to be made available under an open license, such as CC BY. Open licenses, such as those offered by Creative Commons, operate within the current system of copyright and have been upheld as legally enforceable by the US Court of Appeals for the Federal Circuit.¹³ Furthermore, the CC BY license is already in use by a federal grant program, namely the Department of Labor's \$2 billion Trade Adjustment Assistance Community College and Career Training grant program (TAACCCT).¹⁴

While the NIH policy has been successful by all accounts, federal public access policies should now go beyond read-only access and include full reuse rights without commercial restriction. When taxpayers fund research, they deserve the full use of the results – to distribute, reuse, data or text mine, and build business on top of – rather than solely the permission to read resulting articles. As mentioned above, open licensing is crucial to maximizing the potential scientific and commercial benefit that can be realized from federally funded research. Opening this vast literature – not only to a larger audience of readers, but also for unrestricted use – will encourage the creation of innovative new tools, such as the OpenSNP application mentioned in comment 1, and pay dividends in ways we cannot presently imagine in the current closed system.

While an immediate open license maximizes the return on taxpayer investment in research, one compromise that could be considered to balance the interest of all stakeholders would be a stepped approach. Initially, articles would be under a period of embargoed access in which usage is restricted to only those uses allowed under copyright and fair use. Then, after the expiration of an embargo period of perhaps three to six months, the articles would be subject to an open license that would allow full reuse rights without commercial restriction, such as CC BY. This approach would allow publishers a sufficient period to recoup their investment, and would still give the public the full reuse rights they deserve for underwriting the research. Much of the additional economic benefit only gained when articles are made openly available would also be captured under this approach.

¹³ Case law supporting the legal enforceability of Creative Commons licenses can be found at http://wiki.creativecommons.org/Case_Law.

¹⁴ See TAACCCT's Notice of Solicitation for Grant Applications, p. 21: <http://www.doleta.gov/grants/pdf/SGA-DFA-PY-10-03.pdf>.

[Question 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?

[Comment 3]

Federal agencies are the appropriate entity to maintain a centralized repository to ensure permanent, public access to publicly funded research. At a minimum, agencies should maintain a mirrored, publicly accessible copy of all articles resulting from federal funding and ensure they retain the rights necessary to do so, as the NIH has done through its public access policy since 2008.

Centralized repositories like PubMed Central (PMC) provide students, researchers, and others with a single point of access to a vast portion of the relevant research literature. This single interface provides students superior ease of use compared to collections of articles scattered across the websites of thousands of individual journals. This ease of use, in turn, enhances discoverability and scientific productivity. NIH's PubMed Central has convincingly demonstrated the excellent return on investment of such a repository. PubMed Central sees 500,000 unique users every day,¹⁵ three-quarters of whom are from outside of the academy.¹⁶

Federal custody is necessary to protect our investment in research by ensuring long-term stewardship over the course of decades or longer. Publishers' incentives and limitations necessitate such an approach. As with any business, publishers can and will fail, and without a properly maintained backup, large numbers of federally funded articles could be erased permanently when a publisher goes out of business. Publishers may also wish to prevent competitors from building products and services on top of their content by stipulating that any centralized repository be a "dark archive." However, public access to such a centralized repository is crucial to maintain archival veracity and maximize the return on our federal research investment.

Furthermore, establishing centralized repositories for other agencies (or groups of agencies) can be accomplished with relatively minor expense or effort. PubMed

¹⁵ Letter from Dr. Francis Collins, Director of the NIH, to Representative Joseph Pitts. December 2011. Available at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

¹⁶ Statement by David J. Lipman, MD, Director, National Center for Biotechnology Information, Public Access to Federally-Funded Research before the Committee on Oversight and Governmental Reform Subcommittee on Information Policy, Census and National Archives, United States House of Representatives. July 2010. Available at <http://www.hhs.gov/asl/testify/2010/07/t20100729c.html>.

Central's existing platform can be customized to meet the needs of other agencies at a fraction of the cost of starting from scratch. Alternatively, NIH's PubMed Central could be expanded to house all federally funded research in one central, cross-agency repository.

[Question 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?

[Comment 6]

For the reasons mentioned in previous comments, a policy requiring immediate, open access to articles through a centralized, PMC-like repository would maximize the benefit to the public and create the highest return on our federal investment in research.

To minimize the burden on all stakeholders, agencies should standardize the language, requirements, and procedures of their policies, being as consistent as possible. As institutions and researchers are often awarded grants by multiple federal agencies, such consistency will be essential to reduce complexity for grantees and increase policy compliance. Researchers should only need to learn one process, not be forced to navigate a web of different, conflicting requirements across federal agencies.

[Question 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?

[Comment 7]

Any peer-reviewed publications resulting from federally funded research and created without the expectation of compensation should be made freely accessible to the public. Free access to these publications would provide significant value to students, researchers and others. For example, conference proceeding papers can provide additional or unique information not present in final publications, include preliminary results that allow insight into future publications, or contain comprehensive reviews of published research to date that can keep others informed of the current state of a given field. However, policies for making these other types of peer-review publications available may differ from those that apply to journal articles; thus, they should be considered separately.

[Question 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?

[Comment 8]

The public should be granted open access to the results of federally funded research immediately upon publication.

American students, in particular, would benefit significantly from immediate, rather than embargoed, access. It is unacceptable to ask students to make do with outdated information. Instead, federal agencies should get cutting-edge research into students' hands immediately. Immediate access to these resources is especially crucial in rapidly evolving fields, such as biotechnology and alternative energy, that form a significant portion of the United States economy and represent some of its most innovative and high-growth sectors. By providing students with improved access to a cutting-edge education, public access policies that provide immediate access can boost American economic competitiveness by helping students hit the ground running after graduation and put their education to use immediately. Furthermore, courses only last three to four months. With an embargo period, a course may be taught many times before the newest research becomes available and thus can be integrated into the class.

If an embargo period is deemed necessary, it should be as short as possible, and the full opportunity cost of slowing the pace of research and delaying students' access to the most up-to-date research should be taken into account when considering the embargo's length. An embargo period should not exceed twelve months and would preferably be six months or less, as is the norm among research funders around the world with such policies.¹⁷ Similarly, hundreds of subscription-based journals voluntarily make their content freely available after embargo periods, typically of six to twelve months.¹⁸ This list includes publishers that have previously expressed concern over the potential negative impact of opening up access to their content. One such example is the Royal Society, the world's oldest scientific publisher, which earlier this year announced it would make its entire historical journal archive available online for free. Finally, the NIH's public access policy provides strong empirical proof that such measures do not harm subscription-based publishers. To date, no publisher has presented any evidence that the NIH policy has harmed its business. In fact, the largest commercial publisher, Elsevier, which owns a large

¹⁷ A complete list of funder access policies, including details and embargo periods, can be found at <http://roarmap.eprints.org/view/type/funder=5Fmandate.html>.

¹⁸ A complete list of subscription journals which allow embargoed access to their content can be found at <http://highwire.stanford.edu/lists/freart.dtl>.

number of journals affected by the NIH policy, has seen its profit margin and revenues increase every year since 2008 when the NIH policy took effect.¹⁹

Embargo periods have a cumulative impact, as they delay new research by their duration at each research cycle. For example, a paper under a twelve-month embargo will not be available to a large portion of researchers until a year after it is published, delaying follow-on research. If papers from that follow-on research are also subject to a twelve-month embargo, then the availability of those results is delayed a full two years. This delay will continue to accumulate with each cycle of research until it far exceeds the original embargo period.

¹⁹ Elsevier's most recent annual financial reports can be found at:
2010: http://reports.reedelsevier.com/documents/pdfs/reed_ar_2010.pdf; relevant figures: p. 134
2009: http://reports.reedelsevier.com/PDFFiles/ReedElsevier_AR09.pdf; relevant figures: p. 91

The Right to Research Coalition includes 48 member student organizations:

American:

- The American Medical Student Association
- The American University Washington College of Law Student Bar Association
- California Institute of Technology Graduate Student Council
- Columbia University Graduate Student Advisory Council
- Cornell University Graduate and Professional Student Assembly
- Dartmouth College Graduate Student Council
- Harvard Extension Pre-Health Society
- Library and Information Science Student Association, Simmons College
- Massachusetts Institute of Technology Graduate Student Council
- Massachusetts Institute of Technology Undergraduate Association
- National Association of Graduate-Professional Students
- Oberlin College Student Senate
- Oklahoma State University Graduate and Professional Student Government Association
- St. Olaf College Student Government Association
- Student Advocates for Graduate Education
- The Student Public Interest Research Groups
- Students for Free Culture
- Trinity University Association of Student Representatives
- Tufts Graduate Student Council
- Tufts University Friedman School of Nutrition Science and Policy Student Council
- The United States Student Association
- Universities Allied for Essential Medicines
- University of California, San Diego Graduate Student Association
- University of Minnesota Graduate and Professional Student Assembly
- University of Nebraska - Lincoln Graduate Student Association
- University of Tennessee - Knoxville Student Government Association

International:

- The Association of Medical Students in Bulgaria
- Athabasca University Graduate Students' Association
- The Canadian Federation of Students
- The Croatian Pharmacy and Medical Biochemistry Students' Association
- Direção Executiva Nacional dos Estudantes de Medicina (Brazil)
- The European Federation of Psychology Students' Associations
- The European Medical Students' Association
- The European Medical Students' Association - Turkey
- The European Pharmaceutical Students' Association
- The Indian Medical Student Association
- The International Association for Political Science Students
- The International Association of Students in Agricultural and Related Sciences
- The International Federation of Medical Students' Associations
- The International Federation of Medical Students' Associations - The Netherlands
- The Lebanese Medical Students' International Committee
- The Macedonian Medical Student's Association
- The Malta Medical Students' Association
- The Medical Students' Association of Kenya
- Medsin-UK
- National Graduate Caucus of the Canadian Federation of Students
- Udruga Studenata Dentalne Medicine (Croatia)
- University of Calgary Students' Academic Assembly