To: publicaccess@ostp.gov

Comments from Lee C. Van Orsdel, Dean of University Libraries

Grand Valley State University
Allendale, Michigan

Summary statement

Thank you for the opportunity to comment on the critical issues associated with providing public access to and preservation of the broad range of research funded by taxpayers in the U.S. I am Dean of Libraries at Grand Valley State University, a Midwestern university with a Carnegie Classification of Masters Large. While we are not a research university, per se, our faculty, staff and students conduct research on a regular basis, often with the help of government and foundation grants, as do a number of distinguished research centers funded by GVSU.

Our University, located on the west side of Michigan, is deeply committed to helping our state recover from the severe downturn in the economy. We are, no doubt, reflective of the country as a whole—trying to leverage the opportunities emerging from an increasingly globalized, hyper-connected and hyper-competitive world into a stronger economy for our region. Innovation, creativity and discovery are the surest paths to success. Timely access to information facilitates all three, making your inquiry into federal information policy of huge importance to our University and to our region.

We strongly support the expansion of the NIH Open Access Mandate to all federal agencies that fund research with taxpayer dollars. Gaining access to all non-classified, federally funded research would unlock countless valuable resources that are not currently available to our university or to the citizens of our region because they are too costly for us to purchase. Then NIH mandate is a case in point. It has unlocked tens of thousands of articles and made them available. Our University has strong baccalaureate, masters and doctoral level programs in nursing and the health sciences, and our researchers have already benefited from immediate, free access to the invaluable collection of cutting-edge research emerging each year from NIH grants.

The NIH mandate has been successful on every level—providing critical, free information to citizens across the whole spectrum of expertise, from doctors and researchers to patients, parents, advocacy groups, medical supply and pharmaceutical companies, inventors, and lawmakers. Its value extends well beyond institutions of higher education like my own. The cost has been minimal—about one-hundredth of 1%--of the NIH annual budget. The NIH policy could be improved upon with a shorter embargo period (6 months as opposed to the current 12 months) and looser restrictions on re-use, but it is fundamentally a viable model that can be easily expanded to accommodate research from other government agencies.

In sum, there are compelling reasons for the United States to adopt an Open Access policy for non-classified, federally funded research. Taxpayers provide the billions of dollars the government invests in research each year. We should have access to the outputs of that

research. Not just on principle, which is important enough, but because it can make a real and immediate difference in their lives at a personal level and across their communities.

You have only to look at the usage data and at anecdotal reports about the effects of the NIH mandate to see how directly the benefits of research can flow to citizens and to communities when access is opened to the public.

In addition, a broad Open Access policy for taxpayer-funded research would help economic growth. The European Union, the UK, China, India and other emerging nations are embracing open access to state-funded research and data. They recognize that information drives discovery and innovation and they want, for their countries, a more robust return on their investment in research. Our global economic competitors will have a competitive advantage if their research and data are open and ours are still locked away behind unnecessary tollgates in the form of subscription fees.

(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?

<u>Encouraging commercialization</u>

The single most important step to grow existing and new markets using results of federally funded research is to require public access to the resulting research results as quickly after publication as possible (typically this means deposit of the final version of the author's manuscript after peer-review and before final formatting as a published journal article within 12 months of publication).

A collateral issue is to increase the speed with which access is provided. The NIH mandate uses a 12-month window; research has to be made publically accessible no more than 12 months after publication of the article. A shorter window—6 months—would still provide a reasonable window of protection to the publishers who have invested in the formal presentation of the article, and it would double the speed with which the returns on our investment in research begin to manifest themselves in the market.

For-profit publishers have pushed hard to protect their economic interest in the existing publishing model. They will say that earlier access to journal articles will hurt their business. Yet no publisher has been able to provide hard evidence to suggest a 6-month embargo results in cancellations of subscriptions to publisher's journals. Even before the recent economic cuts in higher education, which certainly have caused deep cuts in journal subscriptions in most academic libraries, there was no viable evidence that librarians would cut a subscription to a scientific or medical journal because the articles would be free in six

months. Six months is too long to wait for the article if you are teaching and supporting research in those fields. Publishers are using their *fear* of such an outcome to try and block needed reforms in the dissemination of research findings. When all of the factors that lead libraries to cancel journal subscriptions are considered, there is simply no compelling evidence to suggest that the publishers' concerns about the market should outweigh even more compelling evidence about the positive benefits to the populace and the economy when information is free and accessible to all. I believe that six months is a reasonable compromise on this issue. Twelve months is too long.

Increasing scientific productivity

Research productivity is one of the concerns of this RFI. There is no question that Open Access improves research productivity in a number of ways. It improves the visibility of existing research, making prior research more discoverable and less likely to be needlessly duplicated. It speeds the research process, in that it lessens the amount of time a researcher has to wait to get access to material that is not owned by the library in the research facility. It makes research available to users who are not in the academy, such as those in non-profit organizations, for-profit enterprises, hospitals, independent research labs, and museums. Full open access makes possible the crawling and mining of text and data by machines. And it increases the possibility that the benefits of research will be fully exploited by a host of interested users who will no longer be barred from access.

Cost/Benefits

The most extensive research on the economic impact of open access on economies has been done by John Houghton from the Centre for Strategic Economic Studies at Victoria University in Australia. For a number of years now, Houghton's studies have consistently found positive economic effects from open access policies. For example, Houghton conducted a study in 2006 using conservative measures to assess the impact of an open access public policy on the US economy. The report, "The Economic Impact of Enhanced Access to Research Findings" (http://www.cfses.com/documents/wp23.pdf) concluded that, "With the United State's GERD [Gross Expenditure on Research and Development] at UD 312.5 billion and assuming social returns to R&D of 50%, a 5% increase in access and efficiency would have been worth USD 16 billion."

Since the early reports, Houghton has consulted on a number of similar studies, all seeking the connection between the cost/benefit relationship of open access to the scientific literature. A few examples are included here. He was commissioned by SPARC (Scholarly Publishing and Academic Resources Coalition) in 2010 to conduct a study in the United States. The report, "Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs [in the United States]," (http://www.arl.org/sparc/publications/papers/vuFRPAA/index.shtml) was published in August, 2010 and contained the following assessment:

> Preliminary modeling suggests that over a transitional period of 30 years from implementation, the potential incremental benefits of the proposed FRPAA [Federal

> Research Public Access Act] archiving mandate might be worth around 8 times the costs. Perhaps two-thirds of these benefits would accrue within the US, with the remainder spilling over to other countries. Hence, the US national benefits arising from the proposed FRPAA archiving mandate might be of the order of 5 times the costs.

In a study commissioned by Denmark's Agency for Science, Technology, and Information and released in April 2011 (http://goo.gl/pfAf6), Houghton assigned costs to the delays associated with trying to get access to information in the current system. He was studying the effect of open access on small and medium-sized businesses. His finding (emphases mine):

> Access barriers and delays involve costs. It would have taken an average of 2.2 years longer to develop or introduce the new products or processes in the absence of contributing academic research. For new products, a 2.2 years delay would cost around DKK 36 million [EUR 4.8 million] *per firm* in lost sales, and for new processes it would cost around DKK 211 000 *per firm*.

Houghton's research was also cited in an October 2011 report commissioned by the UK Joint Information Systems Committee (JISC), "Benefits to the Private Sector of Open Access to Higher Education and Scholarly Research" (http://open-access.org.uk/wpcontent/uploads/2011/10/OAIG_Benefits_OA_PrivateSector.pdf):

> This study confirms the importance placed by businesses on access to scholarly research and its broad impact in terms of project, service and process innovation….Open Access publishing provides a way of opening much more university and scholarly research to the business sector….[M]ost businesses spend considerable amounts of time working around paywalls….The review suggests that, at a time of accelerating pressure on SME [small and medium-sized enterprises] competitiveness, a shift to Open Access would create significant cost savings by enabling businesses to review more quickly the relevance of individual papers and act accordingly. By boosting discoverability OA may also add value directly to levels and speed of knowledge transfer in this part of the economy.

Data for the positive relationship between open access and scientific literature continues to mount. Most recently, a study was commissioned by the Australian National Data Service to "examine the costs and benefits of public sector organisations making their data….freely available." A quote from their website (http://www.ands.org.au/resource/cost-benefit.html) (emphases mine):

> He was also able to model the savings to users and estimate the wider benefits to the economy, which include factors like multiple and sequential use (and re-use) of data….The study involves a number of agencies, each suggesting different benefit:cost ratios (*but all positive*). Professor Houghton stresses that the ratios themselves should not be compared, because they reflect more the nature of the underlying data rather than the agencies involved, a caveat which is most important.

That said, the study does demonstrate that the benefits outweigh the costs by some considerable margins, meaning that the case for making PSI [Public Sector Information] freely available is strong. Importantly, these are not one-off figures, they are annual and ongoing, and so the benefits accrue.

Houghton's research has been criticized by some publishers. No doubt those criticisms will be part of their response to this RFI. There are formal responses to the publishers' criticisms at http://www.jisc.ac.uk/media/documents/publications/responseoneiaspmreport.pdf and at http://www.cfses.com/EI-ASPM/Comments-on-Hall(Houghton&Oppenheim).pdf. Even better, experts at the White House should review Houghton's models and findings and determine for themselves the credibility of the various reports. The address of his Centre at the University of Victoria is www.cfses.com/.

Type of access needed

There are levels of open access. At a minimum, open access can guarantee online, free access to readers. The NIH mandate exemplifies the basic level of open access. Of far more value is the second level, which allows both free access *and* permission to use and reuse the research with attribution to the author/s. The Creative Commons Attribution (CC-BY) license is the emerging standard among open access licenses that allow these broader uses (http://creativecommons.org/licenses/by/3.0/).

Taxpayers get more for their money when more scientists can both read and re-use the results of the research they supported. Broad re-use allows researchers and businesses to exploit the value of research investment more quickly and for years into the future. *Access without re-use delivers only a fraction of the value.*

(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?

The NIH mandate provides exactly these protections—public access without copyright infringement. When a researcher receives a grant from the NIH, it is made clear to the grantee that when s/he is ready to publish results of the research, the author is required to send the peer-reviewed, final manuscript to the NIH so that it can be posted in an open access repository (PubMedCentral). Publishers are aware from the get-go that manuscripts reporting NIH research are going to carry this stipulation, and they are free to decline to publish the manuscript if they do not like the restriction.

In no way does the NIH mandate infringe on the rights of any publisher, agency or other stakeholder, despite claims to the contrary by publishers. If their copyrights were truly infringed during the four or so years the NIH mandate has been in place, publishers would have filed suit in court. They have not done so because they know there is no infringement. Instead, they are lobbying for legislation that would overturn the NIH mandate and bar all government agencies from issuing open access mandates. Their strategy is to ensure that taxpayer rights to access publicly funded research never come into play. This legislation is not about copyright protection; it is about economic protection for a distribution model that was essential in a paper world but obsolete in a digital one. The latest example of such legislation is The Research Works Act (H.R.3699) that was introduced into the U.S. House of Representatives on December 16, 2011.

It would be economically foolish and ridiculously self-defeating to put the interests of a relatively small segment of the US economy (the scientific publishing industry) ahead of the immense economic and social potential of a federal open access policy for taxpayer-funded research, particularly on the basis of a spurious claim to violated copyrights. The research and its findings are funded by the nation, belong to the nation and must be exploited for the benefit of the nation. Publishers can continue to play a valuable role in scholarly publishing and research dissemination. They can rightfully expect to negotiate with authors for right of first publication, a right that is already protected in the NIH mandate. But they have no right to expect the government to decide federal information policy based on publishers' desires to satisfy their stockholders and their stakeholders. No one has a greater stake in federal information policy than the American taxpayer, and no corporation deserves to limit or control the return on that investment by controlling access to research it had no hand in creating.

(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?

<u>Centralized versus de-centralized approaches</u>

The NIH uses a centralized model, a repository called PubMedCentral, to house the research resulting from NIH research grants. FRPAA, the proposed legislation that would extend the NIH mandate to other federal agencies, proposes a more flexible system, allowing agencies to host their own and/or allow deposit in any public-access repository (at universities, for example) that meets established standards for openness, interoperability and long-term archiving and preservation. In all cases, multiple venues for access and preservation are needed to safeguard the literature.

That said, any public-access policies that are developed must give the federal government adequate rights to archive and distribute publicly funded articles. Simply providing the government with a copy to put in a "dark archive" (accessible only under certain conditions) is not a viable solution. Without regular access and use, the archival integrity of "dark archive" content cannot be ensured.

A caveat. Some publishers have lobbied to be the repository for deposit and archiving of the papers they publish (often in "dark archives"). Private sector repositories have value, certainly. But there are dangers in allowing publishers to house the *only* archives. Publishers go out of business. They have different ideas about who should be able to access "their" articles. Federal agencies have no regulating authority over private sector publishers. These uncertainties suggest that multiple repositories provide more secure, accessible archiving of research over time.

(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/private partnerships already exist. Universities and academic libraries have created archival repositories and are building the infrastructure to house, preserve and provide access to scholarly literature. We have gained extensive experience in the last decade. These libraries should be encouraged to partner with federal agencies.

Under no condition, however, should a single site—library or otherwise—be the only point of access for federally funded articles. There are over fifty research foundations that currently have open access mandates (similar to the NIH mandate) for their research findings. In no case are they using proprietary (typically publisher) sites as their final archives. However, there are many examples of funders partnering successfully with academic and research institutions in this role.

(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?

This is not my area of expertise and I defer to the suggestions of my colleagues who are supplying detailed suggestions regarding metadata, interoperability and other desirable standards.

(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?

Consistent requirements across federal agencies regarding deposit of peer-reviewed articles is necessary. This will not only minimize the burden, it will also increase the rate of compliance and reduce potential costs among stakeholders. Academic libraries, with the help of their professional associations (particularly the Association of Research Libraries), are already working with NIH grantees on our campuses to help with submissions to PubMedCentral. This could continue and expand as part of the public/private partnerships described in question (5) above.

(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?

It makes sense to consider expanding a federal policy for open access to taxpayer supported research to extend to any other type of publication besides peer-reviewed journal articles. Differences in the inherent nature of different types of material—textbooks, for example, usually pay royalties to the author—will need to be taken into account as those policies are developed.

(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?

Immediate access is the ideal time to optimize scientific and commercial advantages of publicly funded research. As economic research has shown (reference the Houghton studies in the Cost/benefit portion of question one above), delays have costs associated with them at the level of small and medium sized businesses, as well as society as a whole. Worldwide, virtually every biomedical research funding agency outside of the NIH has a 6-month embargo for peer-reviewed journal articles, including those in biomedical fields. The US government should work towards a 6-month embargo as expeditiously as possible.

The existing 12-month embargo in the NIH mandate has been a reasonable compromise between the ideals and the concern of publishers. No doubt they would oppose a shorter embargo period. However, *there has been no data provided by publishers that a 12-month embargo has harmed their business.* In fact, embargos of 12 months or less have been adopted voluntarily by hundreds of journals (http://highwire.stanford.edu/lists/freeart.dtl).

It is worth pointing out that journal cancellations are driven by any number of factors, including price and pricing history, the impact of large bundles (so-called big deals) on the cancellation of individual titles from smaller publishers, library funding, pressure to add new and emerging titles, and other factors. If it is useful to know, my library, with an acquisitions budget of over $4M has never cancelled a journal subscription because the articles were going to be available for free in 6 to 12 months.