From: Heather Etchevers
Subject: **Response to Request for Information - FR Doc. 2011-28623**
Date: January 11, 2012 11:48:20 AM EST
To: publicaccess@ostp.gov

Submitted by Heather Etchevers, Ph.D. – Biomedical clinical research scientist writing in a personal capacity, employed by the French National Institutes of Health (Institut National de la Santé et de la Recherche Biomédicale). U.S. citizen.

I am unashamedly basing my response on that of my colleague Dr. Cameron Neylon, from the U.K., but have changed a number of items to reflect my own convictions and experience.

Preamble

Thank you for the opportunity to respond to this request for information. The intellectual property that is generated through publicly funded research takes many forms. It includes patents, the scholarly communications of researchers (including peer reviewed papers), as well as trade secrets, and expertise. The funder of this IP is the taxpayer, through the action of government. Federal funders pay in large part for the direct costs of research, as well as the indirect costs including, but not limited to, investigator salaries, subscription to scholarly journals, and the provision of infrastructure. That the original ownership of this IP is vested in the government is recognised in the Bayh-Doyle act which explicitly transfers those rights to the research institutions and in response places an obligation on the institutions to maximise the benefits arising from that research.

The government chooses to invest in the generation of this intellectual property for a variety of reasons, including wealth generation, the support of innovation, the creation of a skilled workforce, evidence to support policy making, and improved health outcomes. That is, the government invests in research to support outcomes, not to generate IP *per se*. Thus the appropriate debate is not to argue about the final disposition of the IP itself, but how to best support the services that take that IP and generate the outcomes desired by government and the wider community.

By the current focus on the final disposition of IP, a situation currently exists in which different stakeholders argue about who made what contribution. The IP is either divided to the point where it is useless, or concentrated in places where it never actually gets exploited. If instead we focus on the delivery of services that support the generation of outcomes, we will have a framework that recognises the full range of contributions to the scholarly communications process, allows us to optimise that process on a case by case basis, and ultimately enables the U.S.

tax dollar to make the U.S. a more economically successful and a better place to live.

Response

(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?

Over the last decade, there has been an increase in the massive and demonstrated demand from the general public for access to peer-reviewed papers, particularly for access to validated biomedical research. A second crucial market for traditional papers is small and medium enterprise. Estimates of the loss to the U.S. economy from the current lack of comprehensive access to peer reviewed papers by SMEs are around US$16 B (http://osc.hul.harvard.edu/stp-rfi-response-january-2012).

Education at levels from primary through the postgraduate can also benefit from access to current research. In the hands of the right teachers, availability of primary research in domains of interest to our youth and their educators will enable critical thinking on current issues (as opposed to the decades-old ones that make it into scientific textbooks) and the potential blooming of innovative and relevant solutions by this future skilled workforce as it comes of age.

The question is how to demonstrate the sustainability of markets in which the costs of the services required to produce peer reviewed papers can be supported. The incremental cost of providing immediate access upon publication to peer reviewed research communications is at worst zero. The incremental cost of making a publication more widely available once the sunk costs involved in its preparation and peer review have been covered is zero. The infrastructure exists, both in the form of journal websites, and other repositories, to serve this content.

Open Access publishers, such as the Public Library of Science and BioMedCentral, have demonstrated that it is financially viable to make peer reviewed research freely available **via charging for the service of publication up front**. The charges levied by PLoS and BMC are in fact less than those charged by subscription based publishers for vastly inferior “public access” services. For instance, the American Chemical Society charges up to $3500 for authors to obtain the right to place a copy of the paper in an institutional or

disciplinary repository, but limits the rights to commercial use (including for instance use in research by a biotechnology startup or for teaching in an institution which charges fees). I published a few years ago in an Oxford University Press journal using their Open Access option (which I was glad they offered) - but that option alone cost me $3000 on top of the page charges (see http://www.oxfordjournals.org/oxfordopen/charges.html). By contrast the charge made by PLoS for publication in PLoS ONE is $1350. This provides the service of peer review, publication, archival, and places the final, peer reviewed and typset, version of the paper on the web for the use of any person or organisation for any purposes, thus maximising the potential for that research to reach the people who can use it to generate specific outcomes.

The debate over where the IP is finally located, in which **a publicly funded author has to purchase a limited right to use their own work, having donated their copyright to the publisher**, is ultimately sterile. The debate should be focussed on the provision of publication services, the best mechanisms for paying for those services and ensuring a competitive market, and the value for money that is provided for the public investment. It is noteworthy in this context that a number of new entrants to this market, who have essentially copied the PLoS ONE model, are charging exactly the same fee, suggesting that there is still not a fully functional market and that there is a significant margin for costs to be reduced further.

Nevertheless, once the charge for the services attached to publication have been levied, it is important both for the public and the scientific enterprise to have access to the published work as easily and readily accessible.

(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?

It is crucial that **all** service providers, including publishers, research institutions, and researchers themselves receive appropriate recompense for their contributions, intellectual and otherwise. We must enable markets that support sustainable business models for the provision of these services as well as providing competition that ensures **a fair price is being paid by the taxpayer for these services** and encourages innovation. Many of the critical contributions to the process do not generate any intellectual property as I understand the term.

I have gone through the final submitted version, after peer review, of the ten most recent peer-reviewed papers on which I was the corresponding (and therefore most responsible) author. I have examined the text and illustrations of these, which were subsequently accepted for publication in this form, for any intellectual property that was contributed by the publishers during the peer review process. I have found none.

In my view the only relevant intellectual property here is copyright. The expression of our results and ideas in text or in images was not contributed to these documents by the publishers. The final published versions of these papers do have a small contribution of intellectual property from the publishers: the typesetting and layout in some cases, but these are not relevant to the substance of the research itself. Indeed, many for-profit publishers explicitly permit the "self-archiving" of manuscripts in the pre-typeset format, that can be made freely and openly available, for example on the PubMedCentral website, or on other institutional websites. However, none of these publishers to my knowledge are willing to place a link to the free, self-archived version on the same webpage as the version to which they assert they have made a crucial contribution, which requires payment on a per-article basis.

The publishers for each of these papers have indeed provided a range of helpful services, without which the paper would not have been published or would not be visible, including the infrastructure, management of the peer review process, archiving, and deposition with appropriate indexing services. These important services are clearly ones for which **a fair price should be paid to the service provider**. It is therefore the services that we require to purchase and the most effective and appropriate mechanism by which to purchase them, that should be the point of discussion, not the disposition of intellectual property.

Once the costs of preparing and reviewing a research output and making that output available online have been met, there is no economic benefit or reduced cost achieved by reducing access to that output. There is also no gain in paying the full costs for a service that places an output online but then limits access to that output.

(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?

The preference for a distributed or a centralized approach is likely to differ between disciplines, types of research output, and indeed across national borders. Government and federal agencies can best ensure the infrastructure that can support the fullest range of publication, discovery, archival, and integration services by a mix of services, and technical and human infrastructure, provided not only by government but also by commercial entities, and not-for-profits, some of which are centralised, some of which are distributed. Economies of scale mean that it will be more cost-effective for some elements of this to be centralised and done up-front by federal agencies (e.g. long-term preservation and archiving as undertaken by the Library of Congress), whereas in other cases a patchwork of private service providers will be appropriate (specialist discovery services for specific communities or interest groups).

If a service-based model is adopted, in which a fair price for the costs of providing review and publication services is paid up front, guaranteeing that any interested party can access and re-use the published research output, then government will be free to archive and manage such outputs where appropriate. This will not interfere with the freedom to act of any other interested public or private stakeholder, providing greatest flexibility for all. I believe that **long-term stewardship can be guaranteed by contracts that private publishers would pass with the U.S. government in order to publish federally funded research**: to make their archives directly available to a Federal agency, and a reversion of their IP to the government, in case of financial insolvability.

Concretely, this would require government-supported researchers to continue to disseminate all the results of their work with publishers that have made an engagement to provide said open-access service in exchange for a fee-up-front, which cost can be taken into budgetary account in the research funding that is distributed by said government. Such a requirement can be enforced at the level of the research funding contract.

(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?

There are a range of such models ranging from ArXiv, through relatively traditional publishers like PLoS and BMC, to new and emerging forms of low cost publication that disaggregate the traditional role of the scholarly publisher into a menu of services which can be selected from as desired. The role of government and federal funding agencies is to **make a clear statement of expectations as to the service level expected of the researcher and their institution as a condition of funding**.

I am personally very pleased with the evolution of MEDLINE over the years and the increasing use and interoperability of PubMed as it links to publishers' websites as well as to the PubMedCentral repository when possible. Other countries have started to set up similar institutional repositories - links could be added to these. ResearchGATE (http://www.self-archiving.me/), which is promoted by the American scientific honors society Sigma Xi, is another example. It might be harder to make links to finer-grained self-archiving methods, not to mention difficult to police broken links over time, but perhaps registration with MyNCBI could accord the right to the registrant to make a comment on any individual article, in which they submit such a link to a personal website or other repositories. I do not know how things work in other scientific disciplines, with the exception of arXiv in theoretical physics and mathematics, but I think PubMed could be a model to follow for chemistry or ecology, for example.

(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?

Federal agencies should take advice on the adoption of standards only when and where they have widespread adoption and traction (e.g. the PubMed model). The funding of specific targeted developments to support standards and interoperability development is appropriate. Consideration should be given at all times to aligning research standards with standards of wider relevance (e.g. consumer web standards) where appropriate and possible, as these are likely to be better funded. There are however risks that the development of such standards can take directions not well suited to the research community.

Standards adopted by federal agencies should be open in the sense of having:

a) Clear documentation that enables third parties to adhere to and interoperate with the standard.
b) Working implementations of the standard that can be examined and reverse engineered by interested parties.
c) Defined and accessible processes for the development and ongoing support of the standard.

I think in particular of the tag fields that are associated with references in

MEDLINE for importing into bibliographic managers such as Mendeley, EndNote and Zotero. The standard is clear, and are used by the private sector to everyone's benefit.

(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, consistent with the Paperwork Reduction Act and guidance from the Office of Management and Budgets should adopt a "write once - use many" approach. Where possible, the reporting burden for federally funded research should be discharged once by researchers for the communication of each research output. This means in turn that services purchased in the communication of that research should be sufficient to provide for any downstream use of that communication that does not involve a marginal cost.

Thus, for instance, researchers should **not** be expected to write two independent documents, the peer-reviewed paper, and a further public report, to support public access policies. Reporting on the outcomes of federally funded research should depend, as far as possible, on existing previous communications.

(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?

Yes. All research outputs should be covered by coherent federal policies that focus on ensuring that global outcomes of the public investment in research are maximised. The focus purely on research articles is damaging and limiting to the development of effective communication and thus exploitation. I have written a number of peer-reviewed book chapters, some of which are not visible because not listed in PubMed, and that could be valuable to the public and the research community. Self-archiving is a partial solution, but visibility is still an issue. If public access policies covered other peer-reviewed productions such as these, then that would level the playing field.

Still, scholarly journal articles are the true currency of scientific communication, and if a priority list must be made, they are at the top.

(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?

Once the misleading focus on intellectual property is discarded in favour of a service based analysis it is clear that there is no justification for any length of embargo. Embargoes seek to **ensure a private gain through creating an artificial scarcity** by reducing access for a limited period of time. If a fair price is paid for the service of publication then the publisher has received full recompense in advance of publication and no further artificial monopoly rights are required. As noted above the costs of providing such services are no higher than is currently paid through subscription costs. With appropriate competition the costs might indeed become lower.

From the perspective of exploiting the public investment in research embargoes are also not justifiable. Technical exploitation, commercial development, and the saving of lives all depend on having the best and most up-to-date information to hand. Once a decision has been taken to publish a specific research result it is crucial that all of those who could benefit have access, whether they are private citizens with sick family members, small business owners and entrepreneurs, not-for-profit community support organisations, or major businesses.

Given the current environment of intellectual property law, it may be appropriate under some circumstances for the researcher or their institution to **delay publication** to ensure that the research will be fully exploited. However there is no benefit to either the researcher, their institution, or the federal funding agency in reducing access once the research is published. Further it is clear that reducing access, whether to specific domains, communities, or for specific times, cannot improve the opportunities for exploitation of the research. It can only reduce them.

--
Heather C. Etchevers, Ph.D.
INSERM UMR910
Faculté de Médecine – La Timone
27 boulevard Jean Moulin
13005 Marseille, France
tel : +33 4 91 32 49 37
fax: +33 4 91 80 43 19