Committee for Economic Development
The Best of Business Thinking

The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?

The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?

Includes bibliographic references

First printing in bound-book form: 2012

Printed in the United States of America

COMMITTEE FOR ECONOMIC DEVELOPMENT

2000 L Street, N.W., Suite 700

Washington, D.C., 20036

202-296-5860

www.ced.org

Contents

continued

The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?

Preface

In a series of reports since 2004, CED's Digital Connections Council (DCC) has explored how the development of digital technologies and accompanying changes in communications, computation, and the storage of information have affected commerce and public policies. The DCC has been particularly interested in how manifestations of greater openness—seen in open access to research, open standards, open-source software design, and open innovation—add value to America's scientific, technological, and commercial activities and can be harnessed to improve domains such as healthcare and higher education.

This report builds upon that earlier work and delves deeper into the relationship between the traditional means of providing access to federally funded scientific research and the benefits that can be derived from providing greater public access to it. As with virtually any public policy, the benefits and costs of providing public access to federally funded research fall unevenly on different members of society. We find, however, that because public-access policies that make research more open result in accelerated progress in science and faster economic growth, the net societal benefits far outweigh their limited costs.

We wish to thank the DCC's project director, Elliot Maxwell, Chief Strategist of eMaxwell & Associates, and his research assistant, Kevin Bryan, for their outstanding work in researching and writing this report. We also wish to thank the Ewing Marion Kauffman Foundation for their generous support of this project.

Charles E.M. Kolb
President
Committee for Economic Development

Paul M. Horn
Chair, Digital Connections Council
Senior Vice Provost for Research
New York University
Senior Vice President and Executive Director of Research (Retired)
IBM Corporation

Executive Summary

Progress in science is built upon the work of those who came before. Demonstrations of such progress in research can be found in published scientific journals that for several hundred years have been among the most important vehicles for the dissemination of new scientific knowledge.

Until very recently modern scientific journals were funded primarily by institutional subscriptions; subscribers such as libraries and their users had access to new knowledge while others without subscription access had to wait for other means of knowledge diffusion.

The U.S. National Institutes of Health is the largest single funder of biomedical research in the world with a budget of $31 billion that, through its grants making process, generates 90,000 articles each year. Since 2008, NIH's public-access policy has required that its grantees place a copy of their peer-reviewed manuscripts accepted for publication by a scientific journal in PubMed Central (PMC), an online digital repository open to all, no later than 12 months after the version of record is published. Initially, NIH asked grantees to deposit their work voluntarily; when only a small percentage of grantees did so, deposit became mandatory.

Even after less than four years there has been a marked increase in public access to the results of research funded by NIH. PMC now includes 2.3 million full-text articles (not all based on NIH-funded research), and close to a thousand journals now deposit all of their articles, whether or not they deal with NIH-funded research. Roughly 500,000 unique visitors access PMC on a typical workday.

Advocates of greater openness have supported the NIH public-access policy and proposals to extend such policies to research funded by other major federal funders of unclassified scientific, technical, and medical (STM) research. Opponents, primarily publishers of STM journals, have argued that the policy will damage their subscription-supported publishing businesses and, by so doing, will undercut the peer-review system they use to choose articles by subjecting them to examination by subject-matter experts; the financial pressure could also force publishers to close, and even reduce the amount and quality of research, while undercutting publishers' copyrights. Publishers also argue that the manuscripts authored by researchers are not the direct result of the funding of research.

This report examines the costs and benefits of increased public access, and proposals to either extend or overturn the NIH policy. It looks at increased public access to research results through the lens of "openness," with a particular interest in how greater openness affects the progress of science, the productivity of the research enterprise, the process of innovation, the commercialization of research, and economic growth.

CED has issued several reports on how greater "openness," made possible by the digitization of information and the growth of the Internet, can lead to increased benefits to the society, and how it can improve specific domains such as healthcare and higher education. For CED, openness has two aspects—accessibility and responsiveness. To the degree that information or processes are accessible—e.g. are available without need to pay a subscription or be at a particular place—they are more open. And to the extent that what is accessible is responsive—e.g. can be changed, repurposed, and reused—they are more open.

This report finds that:

Public-access policies should be judged by their impact on the society and the development and dissemination of high-quality scientific research and not by their impact on proprietary publishers, open-access publishers (publishers that rely on author payments rather than subscriptions) digital repositories or any particular means of disseminating knowledge.

The NIH public-access policy has substantially increased public access to research results with benefits as described below that far outweigh the costs. Similar benefits can be expected from extending such a public-access policy to other major federal funders.

Increased public access accelerates progress in science by speeding up and broadening diffusion of knowledge not only to researchers in the field of a particular journal but also to others who have not had easy access to research results, such as researchers in other fields, those in the private sector developing new goods and services that rely on scientific research, clinicians and patients, and many others who can contribute to scientific and technological development.

Greater diversity among researchers and the exploration of a larger variety of research paths result from increased public access which leads to faster movement from basic research to the commercialization of new products and services. Faster commercialization increases economic growth and creates new jobs; a whole new segment of the STM publishing industry is growing up focusing on adding value to newly accessible research results.

Research results which are made more publicly available generate more follow-on research and more citations in future articles to the benefit of the researcher. The processes for academic advancement—e.g. tenure and promotion decisions—need to be rethought so as to reward researchers who support greater openness by early disclosure of their findings or by sharing new tools and processes.

Increasing the ability of researchers to locate research and avoid duplicative or dead-end lines of inquiry promotes the maximum return on the government's investment in research and prevents taxpayers from having to pay twice to support research—once through government grants and then again to obtain access to the results through subscriptions.

Making the results of research more available facilitates the continuing evaluation of research and helps promote accountability for funders and better administration of the research enterprise, allowing a sharper focus on priorities.

No persuasive evidence exists that greater public access as provided by the NIH policy has substantially harmed subscription-supported STM publishers over the last four years or threatens the sustainability of their journals or their ability to fund peer review, where experts voluntarily provide evaluations of manuscripts that are submitted by their authors without any compensation from the publishers. No evidence exists of a significant reduction in traditional publishing outlets (since open-access journals have increased to 7300 in the last decade) or that there will be any shortage of outlets for high-quality research.

The benefits of increased access are so great than any delay in availability of research should be minimized. A maximum six month delay, now employed by other government and private research funders has not been shown to have any negative impact; those who seek delay should bear the burden of proof that the benefits of delay to the development and dissemination of high-quality research outweigh the costs.

The NIH policy focuses on being able to read manuscripts that set out the results of government-funded research. But the manuscript is not the only measure by which to judge increased access nor is reading the article the only goal of most researchers; researchers should be able to access the manuscript and its subparts—underlying data, protocols, tools utilized for analysis etc. The return on investment in government research will be increased to the extent the manuscript and its subparts are machine readable, subject to text and data mining and computable, capable of being copied, distributed, displayed, linked and translated into other languages, and subject to analysis with tools chosen by the reader. Greater openness, as seen in both having access and being able to use what is available, should be the starting point for policy.

Major issues remain concerning unlimited use and reuse, particularly with access to data, but such issues should be addressed by the various stakeholders working together. There have been many initiatives in this area that should be supported in an effort to reach consensus which will enable greater openness.

Digital depositories and other mechanisms for dissemination of knowledge provide high returns on investment; a solely private system would be unlikely to realize these returns. Government should work with stakeholders on standards for metadata to enable search and discovery, standards to ensure interoperability, and rules to guarantee access. Government should also minimize differences among public-access rules for federal agencies to promote access and decrease the cost of compliance for both public and private-sector entities.

The Future of Taxpayer-Funded Research: Who Will Control Access to the Results?

"If I have seen further it is by standing on the shoulders of Giants."—Sir Isaac Newton

"Overall the ability of a society to stand on the shoulders of giants depends not only on generating knowledge but also on the quality of mechanisms for storing, certifying and accessing that knowledge." Jeffrey L. Furman and Scott Stern

"Will the model of science magazines be the same 10 years from now? I highly doubt it… I believe in evolution." Dr. Alan Leshner, CEO American Association for the Advancement of Science and Executive Publisher of the journal *Science*[1]

1. Introduction

In April of 2008 the National Institutes of Health (NIH) made a significant shift in its policies regarding public access to NIH-funded research. Since 2005 NIH policy had encouraged researchers to voluntarily deposit articles based on NIH-funded research in PubMed Central (PMC), an Internet-accessible digital archive or repository which had been created in 2000 to advance science by making peer-reviewed journal articles available for free.[2] But by the end of 2007, only 7 percent of covered articles had been deposited.[3] Congress, in response, required NIH to have all NIH-funded researchers "submit or have submitted for them to the National Library of Medicine's PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication, to be made publicly available no later than 12 months after the official date of publication."[4] Congress instructed NIH to implement this new public-access policy "in a manner consistent with copyright law."[5]

This was a really big deal.

Academic researchers generally obtain funding from their institutions, private companies, not-for-profit organizations and governments to support their research. Once the results are written up the manuscripts are submitted for possible publication to a scholarly journal;* for most academic fields, and particularly in the sciences, scholarly journals are the primary medium for reporting the results of new research.†

Historically journal revenue has come from institutional subscribers such as universities and laboratories. In order to promote subscriptions, most traditional (often called "gated" or "toll-access") journals place substantial limits on access to, and the dissemination of, articles that they publish, as well as restrictions on what subscribers can do with articles to which they have purchased access.††

NIH is the largest funder of biomedical research in the world. Its FY 2011 budget was $31 billion. It is estimated that in 2010 NIH research funds generated over 90,000 peer-reviewed journal articles reporting on the results of research across a wide spectrum of scientific activity. Under the mandatory deposit policy the percentage of covered research placed in PMC rose to 73 percent between July 2008 and December 2010.[6]

---

* Scholarly journals are run by for-profit publishers (Wiley, Elsevier etc.) and not-for-profit groups, particularly academic scholarly societies.

† After an article is submitted by the researcher (without any expectation of being compensated by the publisher) and he or she pays a fee—usually though not always nominal — the journal editor sends the article to experts in the field for their peer-review comments as to the quality of the manuscript, suggested edits etc. Journal editors tend to be full-time researchers who are often paid a small fee by the publisher. Peer reviewers occasionally receive honoraria, though most often they are not paid and participate voluntarily as part of their scholarly obligations. If an article is accepted, the publisher sets the type appropriately for the journal, performs grammatical and sometimes copy editing, and prints paper copies of the final article (although more and more journals are becoming completely electronic in order to cut costs and simplify access). Most journals now archive their articles online.

†† Individuals may subscribe, though high annual subscription rates makes that less usual. Subscribers generally receive access to journal archives, which increasingly are being kept online.

Through its public-access policy NIH has provided interested readers with free access to hundreds of thousands of articles on taxpayer-funded, cutting-edge research which, in the past, were restricted to journal subscribers and those whom they served. Supporters of increased public access have argued that the NIH policy should be extended in some fashion to other Federal government agencies that provide large amounts of funding for unclassified extramural research. Opponents, primarily (but not all) publishers (both proprietary and not for profit) have argued that the policy is damaging or will damage their publishing businesses and, by so doing, will eventually undercut the peer-review process, reduce the number of outlets disseminating research, and even reduce the amount and quality of the research that is made available, while infringing or undercutting the value of the publishers' copyrights.[7]

This paper examines the origins and impact of the NIH policy and the principal claims made for and against it. The paper also considers proposals to narrow, or even to reverse the policy, as well as proposals to extend increased public-access policies to other federally funded research. The paper focuses on the costs and benefits of increased public access to the results of taxpayer-funded research; of particular interest is how greater openness affects the progress of science, innovation, the commercialization of research, the productivity of the research enterprise as a whole, and the growth of the economy.

2. A Brief Note on Terminology

Over the past several years the Digital Connections Council of the Committee for Economic Development (CED) has examined how the growth of the Internet and the digitization of information have dramatically increased the "openness" of information, processes, and institutions.[8] The degree of openness of information, for example, can vary dramatically. To the extent that people have access to information without restrictions, that information is more open than information to which people have access only if they are subscribers, or have security clearances, or have to go to a particular location to get it. But accessibility is only one of two attributes of openness. The other is responsiveness. Can one change the information, repurpose it, remix and redistribute it? Information is, by CED's terminology, more open when there are fewer restrictions on what can be done with it. CED has come to view openness as a continuum running from closed to open, with the degree of openness depending on the limitations, if any, on accessibility and responsiveness.

Supporters of the goals of the Open Access movement often use the term "open access" somewhat differently, to encompass more than the common understanding of "access." They understand open access to include not only free, immediate, and unrestricted access to digital information online, but also full rights to use the information without restrictions other than requiring attribution for the information's creator.[9] Open access so defined would fall at the openness end of CED's continuum of openness.

Under CED's taxonomy, the NIH policy would be considered to be somewhat open because it provides increased access to research but not fully or even nearly fully open because the research is not necessarily available immediately—there may be an embargo of up to 12 months after publication—and because restrictions remain on how a reader can use the manuscript beyond simply reading it e.g. copying and distributing it or conducting data and text mining or incorporating it in a new work. Open Access advocates, while welcoming the greater access under the NIH policy, likewise would not characterize the policy as producing true open access.

To avoid confusion we will refer to the NIH policy as a public-access policy—rather than as an open-access policy.

This leads to another set of definitions that might help in navigating through the debates on open access. "Green Open Access" (Green OA) or self archiving involves researchers placing copies of their own manuscripts on their websites or in digital repositories such as PMC. The present NIH policy is Green OA. "Gold Open Access" (Gold OA) consists of publishers making articles in their journals freely available on their websites. "Hybrid" OA journals allow access on an article-by-article basis. "Gratis" Open Access describes those situations in which access is immediate and free from fees or other charges. "Libre" Open Access removes price barriers and at least some licensing or copyright restrictions—what some have called "permission" barriers regarding the use of the information.

3. The Trend Toward Greater Public Access to Research Results

Scientists and other researchers, particularly academics, have a long history of sharing their findings, often through journals and scientific societies. In recent years, with the rise of the digitization of information and the growth of the Internet, this sharing has become far easier and cheaper; costs for processing, distributing, transporting, and storing information have become negligible.

A new era in sharing began in 1991 when Paul Ginsparg created "arXiv" as an open digital archive or repository for preprints—research papers not yet published in "version–of-record" form by a journal—in physics, a field with a long history of exchanging research findings. Submissions far exceeded expectations and the number of preprints and the fields covered have continued to grow with estimates that arXiv now contains pre-prints of 95 percent of the peer-reviewed articles in high-energy physics.[10] In 1994, Steven Harnad called for researchers to place the results of their research in publicly accessible digital archives in order to make them more widely and quickly available rather than relying on intermediaries such as journal publishers; this would replace "publish or perish with self-archive and flourish."[11]

Digital repositories or archives now come in many different flavors, e.g. large public repositories like PMC, discipline-based repositories such as RePEc in economics, and institutionally based repositories such as Dspace at MIT. The coverage of repositories is not universal. Not all institutions have them and policies differ among those that do as to whether deposit of articles is mandatory, how they treat their accessions, and what users may do with them.

During the same period, the intellectual framework for increasing access to knowledge was being reinforced both in the United States and elsewhere. The 2002 Declaration of the Budapest Open Access Initiative captured the spirit of the movement:

> An old tradition and a new technology have converged to make possible an unprecedented public good. The old tradition is the willingness of scientists and scholars to publish the fruits of their research in scholarly journals without payment for the sake of inquiry and knowledge. The new technology is the Internet. The public good that they make possible is the world-wide electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all scientists, scholars, teachers, students, and other curious minds. Removing access barriers to this literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge.[12]

In the succeeding years other meetings of advocates, in Bethesda[13] and Berlin,[14] echoed the call for greater access to knowledge.

At the same time, and at least partly in response, a growing number of Open-Access (OA) publishers, such as the Public Library of Science and BioMed Central, emerged. These published peer-reviewed articles in OA journals that were funded not by subscription revenues but by charging a fee to authors under a system known as "author pays," in essence shifting payment indirectly to those who fund the research. (Some OA journals received outside funding including to pay start up costs.) While there is some diversity in the access and permission policies among OA journals, basically all of their articles are available for free to anyone on the Internet. Many are deposited as a matter of course into repositories like PMC.

4. Increased Access to Government Funded Research

Even before the emergence of the commercial internet in the mid 1990s governments around the world had begun to focus on the transition away from economies built on industrial might to a new Information Age whose raw materials were intangible bits processed and distributed by information and communications technologies. A growing body of economic literature was showing the substantial public returns of scientific research, the cumulative nature of discovery, the importance of research results and their dissemination for innovation, and the importance of innovation for economic growth. No economically advanced country wanted to be left behind in the Information Age. Each wanted to build a Knowledge Economy which could successfully compete internationally. Accordingly, governments began increasing their financial support for research, particularly in the biomedical arena.

In 2005, the Organization for Economic Cooperation and Development (OECD), made up of the leading developed nations, issued a report on scientific publishing that addressed this special category of research:

> Governments would boost innovation and get a better return on their investment in publicly funded research by making research findings more widely available…And by doing so, they would maximize social returns on public investments.[15]

During this period there was increasing support for greater access to research results within the European Union. As the EU Commissioner for Science and Research noted at the launch of the EU's 7th Research Framework Programme in 2007:

> Easy and free access to the latest knowledge in strategic areas is crucial for EU research competitiveness. This open access pilot is an important step toward achieving the 'fifth freedom', the free movement of knowledge amongst Member States, researchers, industry and the public at large. Beyond, it is a fair return to the public of research that is funded by EU money.[16]

To further the goal of increasing access, the EU has committed millions of euros to support the creation and maintenance of an open-access research infrastructure known as OpenAire—the Open Access Infrastructure for Research in Europe.[17]

The NIH adopted its "Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research," in 2005. It encouraged the *voluntary* deposit of research findings it had funded into PMC.[18] Based on the limited success of this voluntary policy Congress instructed NIH to develop a *mandatory* deposit policy which took effect in April of 2008.[19] This policy required that NIH-funded researchers make an electronic version of their final peer-reviewed manuscripts accepted for publication—not the version of record published in a journal—available to the public by depositing them in PMC within 12 months of the journal's publication.

The rationale underlying NIH's public-access policies, both voluntary and mandatory, was to:

> [E]stablish a "central archive of NIH-funded research publications—for now and in the future—preserving vital medical research results and information for years to come.
>
> Provide electronic access to NIH-funded research publications for patients, families, health professionals, scientists, teachers, and students.
>
> Create an information resource to make it easier for scientists to mine medical knowledge and NIH to manage better its entire research investment."[20]

As one NIH official explained it, because of the "cumulative nature of science" a "stable repository" that provided improved access to "cutting edge research" would "advance basic science and accelerate its application to solving today's problems…."[21*]

Numerous other governments have implemented or are considering similar policies to increase public-access to government-funded research. In the United Kingdom, for example, over 90 percent of articles resulting from government biomedical funding through the UK Medical Research Councils are available under their public-access policy.[22] The Wellcome Trust, the UK's largest private biomedical research funder, has also adopted an increased public-access policy. Both the Medical Research Councils and the Wellcome Trust

* "The "America Competes Act" of 2007 also addressed access to NSF research. The Act required that "All final project reports and citations of published research documents resulting from research funded, in whole or in part, by NSF, should be made available to the public in a timely manner and in electronic form through NSF's website."

This mandate was limited to project reports rather than the final author's manuscript detailing the research findings and analyzing them or the published article. The follow-on "America Competes Reauthorization Act" suggested the need for further steps by requiring the White House Office of Science and Technology Policy (OSTP) to establish a working group to coordinate public-access policies to journal articles and data across the federal agencies.

have established policies requiring that articles based on research they fund be made publicly available no later than six months after publication, rather than the twelve months allowed under the NIH policy.

Major private funders such as the Howard Hughes Medical Institute, the largest private funder of biomedical research in the United States, have adopted similar policies to increase access to the research they fund.

The European Commission is now funding Publishing and the Ecology of European Research (PEER) to "investigate the effects of the large-scale, systematic depositing of authors' final peer-reviewed manuscripts… on reader access, author visibility, and journal viability as well as on the broader ecology of European research."[23] PEER is "a collaboration between publishers, repositories and researchers" examining how to maximize the use and impact of peer-reviewed research including such issues as the length of embargos before articles must be made publicly available. In 2012 the Commission "will propose a European Research Area Framework and supporting measures […]They will notably seek to ensure through a common approach to […} dissemination, transfer and use of research results, including through open access to publications and data from publicly funded research." The Commission is planning to promote open access to the results of publicly funded research. It will aim to make open access to publications the general principle for projects funded by the EU research Framework Programmes.[24]

In a major development in December 2011 the British Government announced its intention to require that all publicly funded scientific research be published in OA journals in what the *Guardian* newspaper described as "a direct challenge to the business models of the big academic publishing companies which are gatekeepers for the majority of high-quality scientific research."[25] The British Government explained its reasoning in its report, "Innovation and Research Strategy for Growth,"

> Free and open access to taxpayer-funded research offers significant social and economic benefits by spreading knowledge, raising the prestige of UK research and encouraging technology transfer. At the moment, such research is often difficult to find and expensive to access. This can defeat the original purpose of taxpayer-funded research and limits understanding and innovation.
>
> We seek "to ensure that government policies stimulate rather than hinder, UK innovation through…[i]ncreasing access to public data or to knowledge created as a result of publicly funded research."[26]

5. Maintaining, Extending or Limiting Public Access to Government Funded Research

Even before NIH adopted its mandatory public-access policy in 2008, legislation had been introduced to increase public access to federally funded research. The Federal Research Public Access Act (FRPAA), introduced by Senators Cornyn and Lieberman in 2006, would have required that researchers funded by 11 federal agencies with budgets for unclassified extramural research greater than $100 million develop policies to provide access to manuscripts no later than 12 months after publication. This legislation, reintroduced in 2010 as S. 1373, would have set the embargo period at six months; the agencies that would be covered have annual research budgets, including that of NIH, of over $60 billion.[27]

A very different piece of proposed legislation would have ended NIH's mandatory deposit policy. The Fair Copyright in Research Works Act, introduced by Representative Conyers in 2008, also would have prevented any other federal agency from adopting a similar policy.[28] The bill was reintroduced as H.R. 6845 in the 110th Congress and H.R. 801 in the 111th Congress. A bill recently introduced by Representatives Issa and Mahoney, the Research Works Act (H.R. 3699), and supported by the Association of American Publishers, also would reverse the NIH policy and prevent its adoption by other agencies. It would accomplish this by eliminating the government's ability to condition research funding on the researcher's agreement to provide a publicly accessible copy of the researcher's peer-reviewed manuscript accepted for publication by requiring the approval of the publisher for any public dissemination if the publisher had engaged in a peer-review process (or any other value-added activity) prior to accepting the article.*

This new legislative proposal came as the US Office of Science and Technology Policy (OSTP) began to receive comments in a public proceeding addressing the issues surrounding increased public access. In late 2011, OSTP issued two requests for information seeking responses to a number of questions—one concerning the NIH policy and its possible extension to other federal funders, while the other raised issues surrounding increased access to data.[29]

The wisdom of maintaining or limiting the NIH public-access policy, or extending it or some variant to other federal research funders, depends ultimately on whether the policy's benefits exceed its costs. As one member of the House Judiciary Committee noted in hearings on the Fair Copyright in Research Works Act, ultimately one "must decide whether the perpetuation of the NIH policy will promote or inhibit the development and dissemination of medical knowledge."[30]

This report generally adopts this perspective and will evaluate the impact of the NIH policy and proposals to extend it to other federal funders based on the effect on the long-term development and dissemination of high-quality scientific, technical, and medical research and scientific progress. Happily, based on an analysis of congressional testimony and filings before NIH and OSTP, both critics and supporters of the policies promoting increased public access appear to share these goals. Whether the policies have positive or negative effects on proprietary publishers, open-access publishers, institutional repositories, or other institutions is a secondary concern, except insofar as their existence is required for the dissemination of knowledge.

6. The Impact of The NIH Public-Access Policy

Although the mandatory deposit policy has been in effect for less than four years it has already had an impact on public access to NIH-funded research. As of November 2011, roughly 26,000 articles were being deposited into PubMed Central every month; about 10,000 of the items are being deposited pursuant to the mandatory deposit policy—a dramatic increase in the percentage of covered articles from the days when deposit was voluntary.

In addition, the policy is positively correlated with an increase in access to research not directly covered by the policy. PMC now receives far more articles not covered by the policy from far more journals than before the adoption of mandatory deposit. PMC now includes a total of 2.3 million full-text articles including both manuscripts and final published articles. Nine-hundred-ninety one journals now provide their full content up from 330 in April 2008. Another 300 journals systematically provide their final, published NIH-funded articles, while 1,620 provide individual articles.[31]

In 2011, roughly 500,000 unique visitors were accessing PMC on a typical weekday, downloading over one million articles; the number of articles retrieved has doubled in the past three years.[32] An NIH analysis identified some 25 percent of the visitors as from academic institutions, 40 percent from the general public and 17 percent from companies.[33]

If the issue is simply whether the policy had increased public access to high-quality research, the answer is simply yes. But critics argue that the question is not so simple. They argue that the NIH policy has important negative consequences. The next section addresses the principal arguments raised against the present policy.

7. The Arguments Against the NIH Public-Access Policy and Other Policies Designed to Increase Public Access

Based upon the public record the opponents' principal claims may be summarized as follows:

A. The NIH policy threatens the sustainability of scholarly publishing. Even with a 12-month

* Under current copyright law the author can retain the right to grant the government a non-exclusive license to the work while still providing the publisher with the copyrights necessary to publish it. This proposed legislation would, In effect, provide the publisher with a new right beyond those necessary for publication and available from the author—a right to prevent a funding agency from making a manuscript accessible through deposit in a repository without the publisher's approval. Given this new right, if the funder wanted to make a copy of the manuscript publicly available, it would have to enter into a negotiation with the publisher, potentially being locked into an obsolete format or other restrictions.

embargo, the policy will likely cause many customers to cancel their subscriptions and wait for the articles to be available for free. By reducing or eliminating this revenue stream the policy will undermine the financial viability of existing publishers.

B. The financial impact will be so damaging that publishers will be unable to afford to perform crucial tasks such as conducting peer review. This threatens the quality of the available research and would allow "junk science" to flood the marketplace.[34]

C. The financial impact will force layoffs and reduce the availability of high-wage, high-skill jobs in scholarly publishing in the United States.

D. Some publishers will be driven out of business thereby reducing the number of outlets for research and discouraging researchers from even conducting research.

E. The present system of publishing has produced a robust and vibrant market for publishing and disseminating research results that is serving users well. Users are satisfied with the present state of access and publishers are already making changes to increase access even more.

F. The cost of operating the government archive will significantly increase NIH costs and reduce the amount of funding available for research.[35] At a time when government budgets are being slashed there is no reason to spend public funds on a function being performed effectively by the publishers who already archive the research they publish.

G. The mandatory deposit policy appropriates the value that the publishers bring to the publishing process without compensation, infringes or undercuts the value of their copyrights, and signals internationally that the U.S. government is not supportive of intellectual property protections.[36]

A. Do the NIH public-access policies threaten the financial viability of the subscription based STM publishers and, in doing so, threaten the quality of research available?

According to critics, government mandates "threaten the sustainability of scholarly publishing."[37] Even with an embargo, the policy will likely cause many customers to "forego continuing their subscriptions and to simply wait for the articles to be freely available on PMC."[38] By reducing or eliminating subscription revenue the policy will so weaken existing publishers that it will place "the peer review process…at risk."[39] The ultimate impact of the policy would be to threaten the amount and quality of the research which is available.[40] In testimony before Congress, a former Register of Copyrights forecast a dire future for proprietary publishers: the policy "will sooner rather than later destroy the commercial market for scientific, technical, and medical journals." He went on: "With plummeting sales how could the [scientific, technical, and medical] publishers stay in business?"[41]

Have these predictions been borne out over the last three and a half years?

First a caveat. It is may be too early to determine the *full* impact of the NIH policy. The mandatory deposit policy only applies to NIH-funded research accepted for publication after April of 2008. Even assuming that the manuscript was accepted for publication within the following six months, the accepted manuscript would not be required to be made available for public access until November of 2009. Thus, the *full* impact of the policy is probably not yet visible today although there is already substantial evidence of an increase in public access to government-funded research.

But if the critics are correct, some negative impact should be discernable—some shadow should be visible over the subscription-supported publishing sector. Such a shadow should have deepened given the existence of additional threats to traditional publishers: other mandates requiring the deposit of funded research in digital repositories and the rise of new OA publishers such as BioMed Central and the Public Library of Science which were creating multiple OA journals to compete for manuscripts with existing outlets. (By the end of 2011 there were 7,304 OA journals listed in the Directory of Open Access Journals.[42])

Even facing these challenges, the scientific, technical, and medical (STM) publishing sector experienced average revenue growth (not counting mergers and acquisitions) between 2003 and 2008 of roughly 7 percent annually.[43] Profit margins remained enviable in comparison with most industry sectors at between 20-30 percent, sometimes even higher.[44]

Starting in 2008, however, revenue growth decreased and profit margins faced heightened pressure. Were these changes the result of the NIH public-access policy and other mandatory deposit mandates? Subscribers could have determined what research of interest was covered by the NIH policy and begun to cancel their subscriptions and watch PMC and other repositories for the freely available articles.

In a filing to the White House's Office of Science and Technology Policy[45] concerning public access policies the Association of American Publishers (AAP) pointed to a 2006 study commissioned by the Publishing Research Consortium (PRC) which found that a large proportion (44 percent) of librarians would cancel their subscriptions if more than 40 percent of a journal's content would be available for free 12 months after publication. AAP also noted that "its members report declining sales and usage since the mandate went into effect," although they provided no further detail.

The PRC study was completed before the NIH public-access policy went into effect and was based on responses to hypothetical situations that might affect subscription behavior. It was also done before the financial crisis hit the United States, which, one might argue, created a situation that increased the likelihood that the study's prediction of a large number of cancellations would be borne out.

Financial analysts following the STM publishing sector have not concluded that the NIH policy caused a substantial number of cancellations—or even that the number of cancellations was anywhere close to the study's predictions. While recognizing that the NIH policy, other deposit mandates, and OA journals might have long-term impacts on the gated journals, they point to another cause for the slowing growth: the stagnation or even reduction in the United States of budgets of academic libraries,[46] which in some cases constitute 80 percent of the subscription revenues of a particular journal.[47]

The financial crisis beginning in 2008 hit the finances of U.S. academic institutions, both private and public, very hard; surveys by the U.S. Association of Research Libraries over the last several years showed a large percentage of libraries had their budgets reduced—and then reduced again, and then again.[48]

Over the last twenty years, STM publishers have raised their subscription prices—justifying this to their customers as "largely reflect[ing] the growth in the volume of scientific research which doubles every 13 years"[49] —by 300 percent, far outstripping the general rate of inflation. Some journal subscriptions now cost more than $20,000 per year. Continuing to raise the price of subscriptions—which in the past has helped contribute to the sterling financial performance of STM publishers—is now much more difficult, as institutional subscribers have less and less money to spend.

Even facing these challenges, proprietary publishers have been generally successful in retaining their customers and keeping their earnings growing sufficiently to please investors. They have entered into so-called "Big Deals" with subscribers, bundling "must-have" journals which have high impact in the scholarly world with less desirable journals, and offering these bundles at prices far less than the total that would have to be paid if subscriptions were bought separately. They have extended subscription contract terms and built in cancellation charges that delay or make cancellations less likely. In some cases, they have waived or reduced the price hikes built into these longer contracts.

Although there have been cancellations, including some notable cancellations of "Big Deal" bundles by major research libraries, there have not been widespread cancellations—certainly not in the 40 percent range described in the PRC study.

There are a number of factors that work to limit cancellations. Beyond the longer contract terms and cancellation fees, research librarians strive to meet the needs of their readers—in the case of research libraries the academic researchers who rely on journals to keep them abreast of research in their field. The research librarian's ultimate goal, unobtainable even for the most well-endowed institutions, would be to provide coverage of all serial publications. Cutting subscriptions means potentially disappointing some number of readers. Moreover, if they were to cancel a subscription the institution's researchers might not have access to the articles based on government-funded research until the embargo period—now up to twelve months long—had passed; few researchers would be pleased to wait twelve months for access to cutting-edge research in their field. Shorter embargo periods, such as the six months utilized by the UK research councils, Wellcome Trust and the Howard Hughes Institute, in theory might have a marginal effect on the behavior of subscribers, but a number of publications have chosen shorter embargo periods without reporting substantial negative effects.[50]

Because most journals are not exclusively reliant on publication of government-funded research, cancellation would mean foregoing access to those articles not based on NIH-funding. (NIH generates roughly 90,000 articles annually—about 13 percent of the approximately 700,000 articles indexed for PMC—covering approximately 5,500 NIH-selected health and biomedical sciences journals. There are a larger number of health and biomedical sciences journals and a far larger number of STM journals and articles.)[51] In addition, cancellation of a subscription may lead to their researchers losing access to archives of already-published articles to which they previously had access.

Research for the proprietary publishing industry, including by publishing trade groups, confirms the analyes of financial analysts on the central role of budget pressures on subscription cancellations. Mike Peine of the Allen Group, in a 2010 study,[52] noting that journal subscription price increases have risen an average of more than 6 percent for the past decade, observed that during the current recession institutional cancellations seemed to be driven mainly by careful examination by purchasers of "duplicate purchases." Libraries, for example, may have noticed that they were subscribing to a journal both individually in print and online as part of an aggregation. A 2010 study by the Publishers Communication Group (PCG) found a similar result.[53] In a study that included 45,000 subscriptions from 2005 to 2010, PCG called libraries that had recently cancelled subscriptions to at least one of the journals in their study and asked for the reason for the cancellation. Reasons cited by more than 5 percent of respondents were: duplicate subscription to online and print versions, library budget cuts, journal price increases, low usage, or the cancelled journal was for a faculty or research group no longer with the subscribing institution. Increased public-access policies, open-access journals, or digital repositories were either not mentioned or were mentioned by less than 5 percent of the respondents.[54] This finding is consistent with the statement by medical libraries that PMC access has not led to subscription cancellations.[55]

The experience of traditional publishers in physics regarding the effect of public access is also instructive: "Physics is the field with the highest level and longest history of OA archiving, and in physics [subscription-funded] journals have publicly acknowledged that they've seen no cancellations attributable to OA archiving. In fact two publishers [of subscription supported journals in physics] have launched their own mirrors of versions of arXiv."[56]

While arguing to policymakers that the NIH policy and its extension to other federal agencies would "cripple the business prospects of peer-reviewed journals,"[57] STM publishers have painted a more optimistic future to financial analysts, who predicted that growth rates would return to the 4-5 percent level as the economy improves.[58] In fact, according to several recent surveys, the last four years have been marked by an *increase* in both the number and subscription prices of STM journals.[59] In their communications with investors, these analysts do not appear to be describing an industry at a "tipping point" or spiraling downward with increasing subscription cancellations and with journal after journal ceasing publication.

B. Is the NIH policy a threat to the continued practice of peer review and the quality of science?

Critics argue that the NIH policy is such a threat to the viability of subscription-supported journals that it jeopardizes their ability to conduct peer reviews. The collapse of the peer-review process would call into question the quality of research that is available, potentially opening the floodgates to "junk science," and "imped[ing] medical and scientific innovation."[60] Publishers argue that they "invest hundreds of millions of dollars every year managing and coordinating the work of millions of authors, editors and reviewers, and vetting millions of submissions through an independent review process," something they would not do if they were unable to afford these sums.[61] Further, "[p]ublishers then also assume the responsibility and costs associated with bringing peer reviewed articles to the attention of other scientists and the news media, including the editorial staff that coordinate multiple revisions, extensive proofreading, layout, design, publishing, distributing and archiving of articles."[62]

Peer review is a crucial part of the process that publishers utilize to certify the quality and importance of the scholarly work that they choose to publish, from among a large number of submissions. When functioning properly, it allows experts to identify for publication the very best research being done and to provide invaluable insights on how to improve the manuscript's reporting and analysis.

The most important value-added component of peer review is provided via the time and effort supplied by these experts, time and effort which are generally voluntarily donated—just as the author provides the article without compensation by the publisher. The donated services must be organized and coordinated, but the more costly activities of the publishers—copy editing, formatting, publication/distribution/marketing/advertising etc.—take place after an article is accepted. It is hard to imagine that the coordination function is a major cost factor when compared to these costs particularly as the labor is donated and the Internet is available to reduce coordination costs. And the NIH policy does not require submission of the published "version of record" but only the final, peer-reviewed author's manuscript accepted for publication.

Critics have also argued that the NIH policy would impede "medical and scientific innovation."[63] In their view, "[t]he nature of peer-reviewed journals makes them attractive to the nation's top research experts and institutions."[64] By risking the future of the peer-review process, the NIH policy and its extension "would discourage future research efforts by America's best and brightest research minds."[65] (It should be remembered that peer-review processes are being utilized by successful OA journals published by BioMed Central—purchased by Springer, a for-profit publisher—the Public Library of Science and many other OA publishers. Since 2010 the PLoS, a pure OA publisher, has completely covered costs with revenue, while employing high-quality peer-review processes.[66])

A letter from 33 Nobel laureates in support of the NIH policy provides a response to charges that it would discourage research efforts:

> This is one of the most important public access initiatives ever undertaken. Finally, scientists, physicians, health care workers, libraries, students, researchers and thousands of academic institutions and companies will have access to the published work of scientists who have been supported by NIH.
>
> For scientists working at the cutting edge of knowledge, it is essential that they have unhindered access to the world's scientific literature. Increasingly, scientists and researchers at all but the most well-financed universities are finding it difficult to pay the escalating costs of subscriptions to the journals that provide their life blood. A major result of the NIH public access initiative is that increasing amounts of scientific knowledge are being made freely available to those who need to use it and through the internet the dissemination of that knowledge is now facile.
>
> The clientele for this knowledge are not just an esoteric group of university scientists and researchers who are pushing forward the frontiers of knowledge. Increasingly, high school students preparing for their science fairs need access to this material so that they too can feel the thrill of research. Teachers preparing courses also need access to the most up-to-date science to augment the inevitably out-of-date textbooks. Most importantly, the lay public wants to know about research findings that may be pertinent to their own health diagnoses and treatment modalities….[67]

A further argument of critics is that mandating the submission of the author's manuscript before it has been edited leaves open the possibility that errors in that version will not be corrected, posing yet another threat to the quality of science. But NIH has addressed this possibility. Its policy is to accept corrections from both authors and publishers that will be reflected in the PMC version. NIH has also made clear that publishers are always welcome to deposit the final published copy of the article in PMC; in those situations PMC will link to the publisher's web site.

C. Will the cost of implementing the NIH policy significantly reduce the funds available to support research?

According to opponents the cost of operating the government archive will "significantly increase NIH costs" and "siphon dwindling resources away from research funding to accomplish a task already well done and paid for by private sector publishers."[68] Moreover, at a time when government budgets are being slashed, there is no reason to spend public funds on a function being performed *effectively by the private sector; and, given the growing pressures on the federal budget there is some question about the "sustainability" of any governmental archives established to implement the new access policies.* [69]

The 2011 budget for NIH was approximately $31 billion. The budgets of the 11 largest funders of non-classified research in the federal government including NIH totaled roughly $60 billion in 2011. NIH estimates its cost of implementing its policy at roughly $4,000,000 or about 1/100 of 1 percent of its budget. (Even if NIH were to pay $3000 in author-pays fees for each article generated by its research it would still be spending less than 1 percent of its budget not taking into account the $100 million it already pays to publishers in publishing fees.)[70] As to the sustainability of any federal repositories, it should be noted that there appear to be considerable economies of scale and scope for large databases such as PMC. Each individual publisher's own archives lack such scale and would obviously be under some pressure if the financial future of scholarly publishing under the NIH policy were as bleak as its critics contend.

It is clearly possible, even likely, that government archival mechanisms will face pressure on funding. But the argument cuts both ways. Is there sufficient reason to believe in the superior sustainability of private as opposed to governmental archives? Is there significant tension between paying the costs of maintaining an archive of articles that lose some of their market value each year with the obligation to maximize shareholder returns? Is there good evidence to conclude that materials archived by a proprietary publisher will remain accessible if the publisher ceases to do business or if the journal is shuttered for any reason? Will gated publishers continue to make available open-access repositories that they may acquire or, as recently occurred, will they put them behind subscription gates?[71] Are profit-maximizing publishers likely to cooperate with their competitors to maximize the availability of all their work for researchers, as opposed to trying to increase the "stickiness" of their own websites? Given these questions, it would not be prudent to rely exclusively on private archiving efforts to maintain access to NIH-funded research.

D. Will the financial impact of the NIH public-access policy, and/or its extension, force layoffs and reduce the availability of high-wage, high-skill jobs in scholarly publishing?

We have not been able to locate sufficiently disaggregated data to make any reasonable judgments about the impact of the NIH policy on employment in the STM publishing sector. Publishers had the opportunity to submit evidence of the employment impact during congressional hearings on the NIH policy and during OSTP's public-comment process on the policy, but no detailed information regarding jobs has been made available and disaggregated data are not available from traditional sources. Even if data were available, it would be difficult to disentangle the effects, if any, of the NIH policy from the effects of the financial crisis, which has had such a negative impact on employment in the United States over the last several years.

There are, however, several observations that can be made about employment in the STM publishing sector. Mergers and acquisitions have increased consolidation in the sector, allowing the remaining publishers to spread costs over a larger number of publications, but in some cases resulting in redundancies in personnel. Publishers have reduced costs and head counts through layoffs and by actively outsourcing pre-publication work from high-wage countries such as the United States, to India and other low-wage countries (a practice also followed by OA publishers.) Given these activities, and the absence of authoritative employment data, it is not possible to make any meaningful judgment regarding the effect of the NIH policy on employment by proprietary publishers. Moreover, any job losses attributed to the policy would need to be weighed against job growth from the thousands of new OA journals and from products and services built around PMC. As one knowledgeable observer of the STM sector wrote:

> [Public access] stimulates the growth of a new industry adding value to the newly accessible research itself. This new industry includes search, current awareness, impact measurement, data integration, citation linking, text and data mining, translation, indexing, organizing, recommending, and summarizing. These new services not only create new jobs and pay taxes, but they make the underlying research more useful….[S}cholarly publishers are themselves in a good position to provide

> many of these value-added services which could provide an additional revenue source for the industry.[72]

Perhaps even more important, any job losses would have to be weighed against job gains from the faster commercialization of research in an environment of greater public access as described in section 9.

E. Will the financial impact of the NIH policy drive some publishers out of business, thereby reducing the number of outlets for research and discouraging researchers from even conducting research?

As of 2011, there was no evidence that any prominent STM journal had ceased publishing or reduced the number of articles published based on the NIH policy adopted in 2008. No prominent STM journal has ceased publication in the past half decade for any reason. The number of STM journals has actually increased over the last 4 years.

There is an argument that high-impact journals will derive some benefit from the increased access. They are likely to gain readers who can cite them in future work. A number of high-impact journals have begun to deposit all their articles in PMC even if the articles are not covered by the NIH public-access policy, some even before twelve months have passed after publication—perhaps for this very reason.[73]

Lower-quality/lower-impact journals may face greater threats, particularly as librarians review the "Big Deals" that bundled these journals with those considered to be "must have." But the number of journals in actual jeopardy in this category is likely to be far lower than the thousands of open-access journals that have been created since 2001. There may also be increased financial pressures on learned societies that publish scholarly journals. Traditionally, the relatively high profit margins available from scholarly publishing have provided support for a wide range of society activities so changes here too are likely; but increased access has created opportunities for new services and new revenue streams for the traditional publishers.

All in all, there is no reason to believe, as a recent economic survey of the literature surrounding open access pointed out, that there is now or will be in the future any difficulty in finding outlets for high-quality research.[74]

In the face of apocalyptic predictions it should be remembered that subscription-based publishers, both for profit and not for profit, are sophisticated business entities. They are not helpless giants unable to navigate and survive changes in the marketplace. They have been taking numerous steps to retain customers, increase revenues, and cut costs including beginning or expanding lines of business not as dependent on the direct publication of research results and providing new value-added tools and services to make the research they publish more useful including investing in data-analysis firms. In particular, they retain important advantages in their existing connections with customers, their publishing experience, their economies of scale, and, perhaps above all, their reputations for excellence at identifying and publishing high-quality research. Even today, with all the digital tools available, when the impact of an article is measured, its importance is often determined by the impact of the journal in which it appears, not necessarily how often it is cited or by whom.[75]

Regardless of the policies aimed at increasing public access, traditional STM publishers are not surprisingly facing a trend that has upended many media markets—the disruptive technologies of the digitization of information and the growth of the Internet. As one group of economic scholars knowledgeable about the Internet has written:

> "Digitization initiated significant shifts in market structure and significant revisions in longstanding competitive behavior in newspapers, music, movies, and other media markets. In each case established firms have faced financial stress from declines in revenues. In each market new entrants offer users new services, and new ways to gain familiar services at much lower cost. In other words, digitization has been closely associated with ushering in Schumpeterian creative destruction in many knowledge-based industries."[76]

While any predictions about the development of markets roiled by disruptive technologies are suspect, one plausible scenario is for an increasingly heterogeneous and competitive marketplace. The STM market is already heterogeneous with profit and not for profit, OA and not OA, and hybrid publishers; even today the major proprietary publishers include OA journals in their own publishing portfolios, an acknowledgement of the viability of the OA publishing model.[77] There are also widely varying norms across disciplines on the access to and use of journals, working papers, and preprints.

The increased competition may prove challenging for lower-impact journals but the higher-impact journals have advantages that are likely to allow them to continue to play a major role while changing in ways that would contribute to their long-term viability. Such a scenario would not be a cause for regret, but a positive development that would have a wide variety of firms competing to better serve the research enterprise and increase the development and dissemination of high-quality research to the benefit of all.

Given our limited ability to predict the future the Executive Publisher of the journal *Science* may have summed it up well: "Will the model of science magazines be the same 10 years from now? I highly doubt it… I believe in evolution."[78]

F. Are users satisfied with their access to results of research so there is no need for government intervention?

Critics of the NIH policy argue [79] that subscription-based journals are the foundation of a publishing system that has made "more scientific, technical and medical (STM) information available to more researchers at lower cost than ever before." They claim that, "90 percent of researchers say they have sufficient access to the STM journals they need. Access to journal content ranks very low on their list of concerns," and publishers are continuing to work to "make information even more widely available online." The NIH public-access policy, to these critics, is a solution "in search of a problem that has not been proven to exist."

More information is available to more researchers–and to everyone–at lower cost than ever before—there is little disagreement about that. And it is not particularly surprising that academics at large U.S. institutions are generally satisfied with their level of access—the survey cited above reflects their views. These researchers are, in fact, the target market for the journals and one would expect them to be satisfied. Subscription rates for key journals are so high that only those whose strong needs to stay current in the field are being satisfied would be willing to pay.

But on deeper examination, the portrait of customer attitudes becomes more complex. While access to new research is generally sufficient for academics at large institutions in the United States and Europe, there are several surveys documenting access challenges in the academic setting including one publisher survey that found "15% of USA and Canadian scholars from all disciplines reported their level of access to the journal literature 'varies' or was 'poor or very poor'."[80] Surveys find much less satisfaction among researchers in poorer countries and among private-sector researchers. Ware and Monkman queried scientists worldwide about their ability to access new research. While 85 percent of American respondents reported good or excellent access, only 66 percent in the Europe/Middle East region did, with even lower numbers (53-56 percent) outside of those regions.[81]

Drilling down further, how does the present structure serve the needs of those who place a lower value on staying current and who are therefore less likely to subscribe? These potential readers may be researchers in the field but with fewer resources, researchers who have left academia but are still active, researchers in a related field, private-sector professionals in large and small enterprises scanning across a broad range of fields for discoveries that might be commercialized or relevant to their own activities, medical practitioners interested in the underlying science, patients, teachers, or the simply curious. Anecdotal evidence in a number of surveys of private-sector researchers[82] and medical clinicians and patients,[83] as well as the experience of academic fields, such as high-energy physics, that offer free preprints online,[84] suggests the presence of a "long tail" of potential readers who place a value on access less than the cost of subscription (and greater than zero dollars). The number of those potential readers is substantial—in aggregate perhaps even larger than the number of present subscribers.

To give a feel for the size of this group that is most likely to benefit from expanded access of the sort provided by the NIH policy: According to the Bureau of Labor Statistics, there are 621,700 industrial R&D workers at research labs in the United States—not including those in other establishments who do research part of the time.[85] The OECD tries to account for these other researchers and counts slightly more than one million industrial R&D workers.[86] NSF tracks all employed science and engineering PhDs separately; there were 649,800 nationally as of the latest data (2008) employed in research professions.[87] The total number of researchers, academic and in industry, is higher.

Such individuals could search for relevant articles at an individual publisher's website and purchase a copy of any article that they find of interest. From an economic standpoint, once research is conducted, written up, peer reviewed and edited for publication, the cost of distributing an additional electronic copy to an interested reader is close to zero. As has been noted, subscription prices have increased far faster than inflation but the prices charged for individual articles—unrelated to any increases in the amount of research being performed—have increased even more dramatically, so that buying individual articles is not likely to be perceived as an attractive substitute for a subscription. For example, a single article in *Applied Microbiology and Biotechnology* published by Springer costs $34.95. Potential readers who value the article at more than zero dollars but less that $35 will be unlikely to purchase it.

There is what economists call "a deadweight loss" to society when buyers who would place a positive value on something, such as research, do not purchase it. The loss to society is even larger when the cumulative nature of scientific inquiry is considered.

Academic research is generally an input into more refined research or into new products. These creations, in turn, inspire further innovative activity. Impediments to access do not merely harm the original potential reader who does not find or read an article, but also those who would learn from the applications to which that reader might have put the research.[88]

Beyond the direct cost of the article, these potential readers face additional problems in obtaining access to research of interest to them. A 2009 study by Ware, finds that about half of English small-and-medium-sized enterprises that believe access to new research is important for their business suffered at least occasional access difficulties that caused them to avoid reading ten to twenty percent of research articles they considered important.[89] A study by Houghton commissioned by the government of Denmark finds that, among private-sector respondents who value access to new research, 68 percent of articles they tried to read presented access difficulties, with only about half of that access difficulty a result of having to pay for research.[90] And their problems with access had a substantial impact on how long it took them to develop and introduce their new products.

Lacking a centralized site to which to turn, individuals must search for research on individual publishers' websites. The challenge of navigating through a large number of publisher's websites should not be underestimated. Forty-four percent of respondents were simply not sure how to find an article they needed.[91] Because publishers control the details of their own websites they may employ, for example, proprietary formats that cause difficulties for some users or devices.*

If a researcher finds an article of interest she must set up an account and make individual payments; one estimate suggests that 60 minutes of staff time were spent for each academic article located and accessed via this route.[92] In addition, there is a built-in impediment to purchasing an article, as it is often difficult to determine whether it is worth the price because the available abstracts are often not particularly informative;[93] at the same time, it is commonplace to need to read multiple articles to find the smaller set that is genuinely useful. Finally, each journal may set different rules on the functionality available to the purchaser of an article. For those who need to scan in an interdisciplinary fashion across two or more fields— a common situation for private sector firms working across multiple academic disciplines—the costs and difficulties mount up quickly.

These potential readers are not generally the targets of subscription-funded publishers. But the loss to society of impeding these potential consumers of research may be substantial. Section 10 will address the important

* Even in the academic world, where many potential readers have access to research through their institution's subscription, the minor annoyances of a clumsy user interface or the availability of a superior one can be enough to change reader behavior: a study of a bibliometric service in physics, which offered links both to the publisher's website copy and a copy on arXiv, found 82 percent of users going to arXiv when both links were available.

contribution that these "non-experts" within a particular field can make to innovation and economic growth. And, as previously noted, given the cumulative nature of scientific progress, society loses the ability to build on contributions that they might have made.

Over the last decade many publishers have made changes to increase access to their published works. Some have made some or all of their back issues available for free on their websites; *Science*, for example, provides free access on its website for all its articles after 12 months. *Nature* encourages the self-archiving of the author's accepted manuscript in PMC six months after publication. Others have chosen to accept some "author-pays" articles, which are made more accessible than traditional articles.

Publishers have pointed to these actions as additional evidence that there is no need for a mandatory deposit policy—that the NIH policy was a solution "in search of a problem that has not been proven to exist." Voluntary efforts by the publishers to increase access are certainly helpful and should be applauded. But the mandatory deposit policy was put into place because NIH's voluntary deposit policy had only increased access to a very small percentage of NIH-funded research.

G. Is the policy, by forcing "scholarly journals to surrender their peer reviewed work to the federal government…without compensation…, tantamount to an unconstitutional taking of intellectual property raising…serious 'due process' and other legal questions"?[94] Are the researcher's manuscripts that are covered by the policy improperly targeted by the policy as they do not "result" from government funding?

According to critics,[95] the NIH policy "could adversely impact the rights of U.S. copyright holders" by effectively stripping them of their value and signals internationally that the U.S. government is not supportive of intellectual property protections. Publishers also dispute what the government has funded, arguing in the AAP's filing to OSTP that the government has a rightful claim only to the reports required by the Federal Acquisitions Regulations; the government does not have a claim on the manuscripts created by the researcher that "describe, explain, or report" on the NIH-funded research but which are not the "direct result" of federal funding. According to the publishers, "[i]n some ways these [federal acquisition reports] are superior to the journal article in achieving the goal of public access to federally funded research results"…as they are "available earlier than any journal article", …"contain significantly more detail and information than the peer-reviewed literature"… and "report on both positive and negative results." (This comparison is somewhat problematic as the NIH policy has no requirements regarding the journal article itself but only affects the accepted manuscript; in other portions of the filing the publishers repeat the point that the government has not funded the journal article as if the policy were directed to the article itself.)

This paper focuses primarily on the economic impact of the NIH policy and its possible extension on the development and dissemination of high-quality research. An analysis of the "takings" argument or the copyright-related arguments—infringement, devaluation of copyrights, the effect on international copyright enforcement—or on what the government has funded—requires a legal analysis and is beyond the paper's scope. But because these arguments were raised during the consideration of the policy and during testimony before Congress, the endnotes of this report include pointers to relevant documents by both proponents and opponents of the policy. Having access to these source documents should allow readers to draw their own conclusions.[96]

In summary, there has been no persuasive evidence of substantial harm to subscription-supported publishers resulting from the NIH policy, as of today. There is, however, substantial evidence of increased access to NIH-funded research. There is also evidence that a large number of potential readers of research results believe that they have less than satisfactory access to research that they consider important and that the lack of access is costly in terms of their potential contributions to scientific progress and to the development of new products and services. There is also substantial evidence of new entrants in the STM market creating new jobs through starting new journals and offering value-added services that increase the returns on the research results made available by the NIH policy.

The next section will describe research on the effects of increased openness.

8. Research on the Effects of Increased Openness

The impact of greater access on academic work is fairly well studied, generally using citation analysis; research on the broader impact of increased access to research, particularly with regard to nonacademic users and its potential impact on the commercialization of research is much less developed. Some important recent research in these areas is discussed below in section 9.

The impact of increased access on citations in follow-on research

One much-debated question within the academic world is whether greater access leads to more citations. The importance attached to the answer is understandable. Citations are the currency of reputation in the academic world, and to the extent open access leads to more citations, academics who benefit will support it—and be more likely to publish in journals that provide greater access, deposit their work in digital repositories or make it available in work papers or preprints. To the extent that citations represent real transfers of knowledge—and they do to some extent, although scholars of bibliography regularly warn that citations are not a perfect measure[97]—increases in academic citations mean increases in diffusion of ideas and real world "impact," from which society as a whole benefits.

Davis (2011) reports on an experiment where publishers made some articles on journal websites open access; the articles were chosen at random. The "open" articles had nearly double the readership of non-open-access articles.[98] A series of academic articles finds an 8 to 250 percent increase in citations to articles openly available.[99] This is not merely driven by authors electing to make their "best work" available openly: some universities have open-access mandates for their faculty, and the citation advantage was roughly equal for articles written by authors who had no ability to self-select.

That increased access results in more citations is clear. But how big is the citation advantage? Although there are methodological issues with much of the research, a citation advantage for open access at the lower end of the range appears the most plausible, as shown by McCabe and Snyder (2011) in their recent review of this literature: studies that explicitly control for selection, whereby authors may make their best research available online but not their marginal research, tend to find the most limited change in citations from being made openly available.[100] Even though the citation advantage is relatively limited, it is likely to be enlarged if coupled with the use of an easily searchable digital repository; making an article openly available can only have a positive effect on citations if potential readers can actually find it. The magnitude of the combined effect of open access and use of a digital repository remains conjecture at this point; there is no convincing research on the topic, perhaps due to the brief amount of time that well-publicized repositories have been operating and the lack of ubiquity in their coverage of various disciplines.

The impact of increased access on the speed of scientific progress

For researchers, developments that increase the speed and breadth of dissemination of cutting-edge research accelerate their own research production. This acceleration, in turn, is likely to be a key determinant of economic growth in the future. This argument has been well articulated by Benjamin Jones, an economist at Northwestern University who specializes in the development of technology. In a series of papers he demonstrates that science appears to be "getting harder" over the past few decades, mainly because it takes a researcher more time to reach the frontier in a given field.[101] Thus, if public-access policies can increase the speed at which frontier scientific results are disseminated the rate at which cutting-edge science is performed should also increase.[102] These papers suggest that maintaining high levels of economic growth in the future will require policymakers to embrace policies and technologies that allow individual scientists to reach the frontier in their specialty more quickly, as well as technologies which allow specialists separated across time and space to connect with each other to form research teams.[103]

The best research on how increased access may affect the speed at which science progresses by allowing cutting edge research to diffuse more quickly concerns preprints rather than open-access articles *per se*. For example, physicists generally place preprints and working papers in arXiv before publication. Gentil-Beccot, et. al, find that approximately 20 percent of all follow-up research cite these deposited works *before* the original article is ever "published by a journal."[104]

Open-access mandates do not require availability of preprints, but they do increase manuscript access and decrease the time which users must wait for access; such policies are likely to have similar effects to those seen in the pre-print research by Gentil-Beccot et. al.

Public-access policies have different effects on different groups. OA journals have little impact in fields like physics where nearly everything is already available online through digital repositories.[105] On the other hand, in fields like medicine, where traditional publishers do not authorize preprints or self archiving *before* the version of record is published, the impact is likely to be greater. The impact of free online access, versus online availability behind a subscription gate, is particularly pronounced for citations coming from researchers in poorer nations.[106]

Different kinds of gain from open access can be seen through the actions of practitioners such as clinicians. In the *Journal of Clinical Psychology*, Hardisty and Haaga (2008) sent mental-health practitioners a link to articles which were variously open access and gated.[107] A week later, the open-access article was found to have been read twice as often. The practical effect was even more impressive: reading the open-access article was associated with the practitioner recommending a more cutting-edge treatment. Another survey showed six out of ten physicians changed their initial diagnosis based on information accessed on line.[108]

The benefits from open access may also be visible with respect to the success of journals themselves. Some journals have switched from a subscription-supported business model to OA or have become hybrids. For example, the journal *Food and Nutrition Research* became OA. As a subscription supported journal, its predecessor had about 800 subscribers. After becoming OA, the journal found substantial increases in unique readers of articles on its website, many more downloads, and increased rates of citations and submissions. Readers came from over 170 countries; previously only 15 countries had even a single subscriber.[109]

As this section has shown, increasing public access to research results increases the speed and breadth of their dissemination and their use by others as demonstrated by follow-on citations.[110] Increased public access and the resulting faster diffusion of cutting edge research via preprints also increase the speed of scientific progress.

The next section examines recent works that illustrate the benefits of increased access to research results in two critical areas: applied research and the commercialization of research results.

9. Increased Public Access and Its Impact on Follow-On Research, Innovation and Commercialization

Greater access to research[*] increases the efficiency and productivity of the research enterprise as a whole and makes it more likely that research will lead, however indirectly, to new products and services which result in increased economic activity and the creation of new jobs. This is true regardless of the state of the economy but is even more critical given the continuing high rate of unemployment.

Four relatively recent papers greatly enhance our understanding of the impact of increased public access and how it may encourage cumulative innovation. In an article in the August 2011 issue of the *American Economic Review*, "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," Jeffrey L. Furman and Scott Stern examine a set of institutions called biological resource centers (BRCs).[111] These centers are generally federally funded; they archive and certify biomaterial used in research studies, and forward archived samples to researchers who want to follow up on earlier work. Famously, Kary Mullis, a private-sector researcher, used a strange organism that lived in geysers at Yellowstone, and had been archived a decade before at a BRC, to develop a technique allowing quick replication of genetic material. When the organism was archived, there was little reason to expect that it would have any practical use. Mullis won a Nobel Prize for this work.[112]

Furman and Stern looked at whether the use of freely accessible materials in the BRCs led to articles that are cited more than other works in similar journals on similar topics using less accessible materials as well

[*] In this case increased access to research includes not only increased access to research articles but also increased access to data and materials, fewer limitations on how information that is being accessed is used, and a lack of restrictions on downstream research and other activities. This broader view of how to increase openness will be discussed more in section 14.

as whether the citation profiles looked different over time. They found that articles based on openly accessible BRC materials got 220 percent more citations.[113] This in itself is not particularly striking given that higher-quality articles are both more likely to be cited and more likely to have their materials archived. Their more-telling finding is that when materials that had been previously archived in private, not-freely-accessible archives, were deposited in a BRC, the citation rates for the articles increased by 50-125 percent. In addition there was a roughly 100 percent increase in unique labs and universities citing the materials[*]

Furman and Stern hypothesized that with private, not-freely-accessible archiving, follow-up studies generally came from those who knew of and had access to the archived materials, particularly friends and associates of the original author. In contrast, materials in the BRC were open to everybody and were used by a far broader circle.[114] A back-of-the-envelope calculation suggests that, given the positive marginal impact of placing materials in a publicly accessible BRC, it would be 3 to 10 times more cost effective for NIH to increase funding of BRCs and other means of dissemination than to fund the lowest ranked but still funded proposals for original research. (Such complementarities between funding of original research and funding of methods for its dissemination and certification also apply to federally funded repositories like PMC.)

"Of Mice and Academics: Examining the Effects of Openness on Innovation," Fiona Murray, Philippe Aghion, Mathias Dewatripont, Julian Kolev and Scott Stern, used a natural experiment in the 1990's to demonstrate the impact of greater openness on both increasing the rate at which follow-on research was taking place, and on broadening the scope of the follow-on research into additional fields (what they called "horizontal extension" and "vertical extension" or "intensity" of research), which is the movement from basic research toward applied research, a step toward commercialization.[115] The natural experiment existed for most of the 1990s when there were three strains of transgenic mice optimized for disease-related research. One of these strains was "open" in that while subject to intellectual property protections these were not enforced; they were thus available without limits on the kind of research being done or on what could be done with the fruits of the research. There were also two strains of "closed" mice. Dupont had patents that covered one of the closed strains, the "Cre-lox mice," and had licensed Harvard's "Oncomouse." Dupont licensed their use by researchers and imposed stringent restrictions on how researchers could breed such mice, how they could report results of their studies, and how any revenues ultimately generated would be shared between the researcher and Dupont. These restrictions, not surprisingly, made researchers wary of using either closed strain.

Murray and her collaborators were able to trace the citations to articles resulting from the use of the various strains of mice. They found that the research based on open mice generated substantially more follow-on research and greater "horizontal extension" of the follow-on research; this had an important advantage—"a substantial increase in the rate of exploration of more diverse research paths."[116]

They also found that citations in follow-on research to research based on open mice were more likely to be found in applied research journals than citations to the research based on closed mice. This suggests that use of the open mice might well lead to faster commercialization.

Their research identified three main channels whereby openness can influence the level and nature of scientific research:

> First, by reducing the costs of accessing key research inputs openness encourages new researchers to enter, thus increasing the diversity

[*] The finding was based on three occasions in their sample, when a private sector "special collection" was forwarded to a major public BRC – for example, one private sector lab shut down and deposited many of their collected samples. Papers using biomaterial from that special collection had an expected lifetime citation profile given the citations they received during the years they were not in a BRC. The positive marginal impact of being added to a BRC was 50 to 125 percent. The impact is largest for articles originally published in lower-ranked journals—what would appear to be a certification effect.

[†] The natural experiment involving "open" and "closed" mice last almost a decade, until NIH reached agreements with the patent holders and licensees so that researchers could use any of the strains without restrictions.

> of academic research participants. Second, relative to what would happen in the case of industrial research, openness makes [academic] researchers more likely to engage in experiments that broaden the number and diversity of research lines…[because] their research can itself have subsequent impact. Finally, there is, of course a direct expropriation effect—an increase in the level of openness of an upstream research tool should encourage the exploitation of that tool in research which is already well down the research line and in the more applied phase. [117]

Openness, they concluded, particularly in academic research, "may increase the overall flow of research output, and in particular it is closely associated with the establishment and exploration of entirely new research lines."[118] They further pointed out: "Overall, our findings highlight a neglected cost of IP: reductions in the diversity of experimentation that follows from a single idea." They stressed the importance of such diversity: there is "value in experiments… that can lead to an entirely new research line, consistent with the idea that scientific discoveries do not follow a purely 'linear' model."[119]

These findings by Murray et. al set the stage for perhaps the most interesting and provocative recent work, a 2010 paper "Intellectual Property Rights and Innovation: Evidence from the Human Genome," by Heidi L. Williams.[120] Like Murray et. al, Williams used a natural experiment, the widely watched effort to decode the human genome. For several years at the turn of the century two groups competed intensively to determine gene sequences. Unlike other research in this area, Williams was able to go beyond the effect of openness on citations in follow-on research and track both publications resulting from the gene sequencing and, even more striking, the commercial product development based upon the resulting articles. As she described it, she was seeking to determine "whether differences in scientific publication translate into differences in the availability of commercial products."[121]

From the early 1990s to 2003, the Human Genome Project (HGP) was a publicly-funded project to map the human genome for future use in medical testing and medicine. From 1999 to 2001, a private effort, by Celera Corporation, was competing in the sequencing. The results of Celera's sequencing were available only to those who paid for them and were protected by "click-wrap licenses," a relatively weak form of protection. Once the publicly funded HGP sequenced a gene the HGP made the sequence freely available to all. In 2003, two years after Celera's data was sequenced, the HGP finished its sequencing and the entire genome was publicly available.

Many research labs and private companies paid Celera a significant amount of money to access Celera's data as soon as Celera completed a sequence, rather than waiting for the HGP to make the data publicly available. One can assume that Celera's customers believed such licensing would allow them to obtain a head start in commercializing the data.

Williams collected data on follow-on research papers about genotype-phenotype links—papers reporting research that linked a specific gene or genes with a specific medical condition, a necessary first step in creating a diagnostic test for a gene-related disease. She also tracked the diagnostic tests that were brought to market. She found that Celera-sequenced genes led to roughly 30 percent fewer articles about genotype-phenotype links, and a similar reduced availability of diagnostic tests based upon these articles. The advantage in publications and, particularly in commercialization, enjoyed by the open-HGP research results was notable. It stands in contrast to what might have been expected to take place based on intellectual property theory, which would have predicted faster commercialization by Celera's customers who had IP protection, a head start, and strong economic incentives to recoup their payments to Celera.

Economists refer to areas where follow-up research is restricted as an "anticommons" since such policies limit the ability of third parties to exploit earlier research in follow-up studies and new product development. Anticommons research such as the Oncomouse and BRC papers would have successfully predicted that increased public access to the HGP data would result in a larger and more diverse group of qualified researchers entering the race to find links between diseases and genes and to develop diagnostic tests based on those links.

Williams' findings are clear and noteworthy: "Celera IP has had negative impacts on economically meaningful science on both scientific research and product development outcomes."[122] In essence, "IP reduces the diversity of scientific experimentation."[123*]

Williams also found that the advantages from openness were long-lasting. Even today, genes initially sequenced by Celera are the subject of fewer publications per year, suggesting "increasing returns to R&D" by those who chose to work on HGP's openly available genetic sequences.[124 †]

As Williams summarized her findings:

> This paper provides empirical evidence on how intellectual property (IP) on a given technology affects subsequent innovation… [A]nalyses suggest Celera's IP led to reductions in subsequent scientific research and product development outcomes on the order of 30 percent. Celera's short-term IP thus appears to have had persistent negative effects on subsequent innovation relative to a counterfactual of Celera genes having always been in the public domain."[125]

This section has described research that demonstrates the advantages of greater access in stimulating follow-on research, and, in particular, the advantages gained by facilitating greater diversity of research paths. The research also demonstrates the advantages of greater openness in moving research toward commercialization and, in the case of the Human Genome Project, in allowing the faster discovery and production of commercial products; these benefits were shown to be long lasting and the negative impacts of restrictions on access to be persistent.

A forthcoming paper by Joshua Gans and Fiona Murray shows that disclosure of the results of funded research is particularly important for government-funded work. As Nelson (1959) and Arrow (1962) pointed out, the private sector tends to devote too little effort to basic research as compared to applied work, a mismatch often solved by government funding of basic research. But applied research funded by government is not identical to private sector applied research; the private sector often has reason to keep results secret so disclosure mandates are critical for maximizing the societal benefits from government-funded science. Gans and Murray provide a useful model of this type of disclosure, as well as a survey of theoretical and empirical evidence.[126]

All of these papers raise important questions about the effects of intellectual property protections on scientific progress; they may suggest a closer look at the effects of the Bayh-Dole Act as currently interpreted by colleges and universities. (In its report on harnessing openness for the benefit of higher education, CED raised this issue.)[127] These papers also suggest that even minor restrictions on openness in science can have large effects, and that since progress in science is cumulative, the results can be long-lasting. A Celera click-wrap license, relying on a private archive rather than an open BRC for biospecimens, and restrictions on use of a transgenic mouse might seem, in the context of a major scientific project, minor hurdles. Yet they are hurdles with real effects. Restrictive access policies for published scientific work may likewise be a relatively minor hurdle, but the preponderance of research on

---

* If the genes sequenced by Celera and the HGP are not randomly selected, it might not be possible to distinguish the effect of IP from the effect of HGP having chosen better genes to sequence earlier. Williams addresses selection in three ways. First, she restricts her sample to genes sequenced late in the HGP study—chromosomes that were known to be very important, as based on earlier scientific research on the chromosome, were, in fact, sequenced first. But in the final years of the HGP, when Celera was active, these particularly important genes had already been sequenced, and what was left was roughly a collection of genes of random known importance. Second, she looks at individual genes, examining follow-up research and commercialization depending on the time that HGP ended up sequencing a gene already discovered by Celera. Third, she looks at the full panel of future research and commercialization as a function of how long a gene is held by Celera. The qualitative conclusion remains the same in each specification: follow-up research on a gene is initially harmed by IP restrictions, this impacts future applied products using those genes, and the cumulative nature of science means this impact persists for many years after the full human genome is sequenced by HGP in 2003.

† One could argue that Celera's competitive engagement in gene sequencing had a secondary benefit in spurring on HGP's sequencing activities, or that Celera used the subscription funds for other cutting-edge research, but these questions were not the subjects of Williams' research. One might also argue that intellectual property protections such as patents promote commercialization—a notion underlying the Bayh-Dole legislation; Williams' research shows that commercial products built on the IP-protected research by Celera were less frequent, however, than those based upon HGP's open research. But this was not the subject of the research.

open science more broadly defined suggests they may have major impacts on both future scientific research and the creation of products and services based on that research. Or to state it more directly, increased access to research results appears to have clear and demonstrable benefits for the advancement of science and the interests of society.

The next section will build upon the importance of encouraging and facilitating the development of diverse research paths by examining the impact of those who have, in the past, had the greatest difficulty in obtaining access to research results.

10. Increased Public Access and the Role of Unforeseen Contributors

Given the non-linear nature of scientific progress and the importance to progress of what Murray et al. describe as "the diversity of experimentation that follows from a single idea," it is noteworthy that the benefits of increased access to research results are likely to be greatest for those who have had the greatest difficulty in accessing them—those who might be "unforeseen contributors to scientific progress." These are not the experts in the field from whom one might expect breakthroughs. Rather they may be former academics, academics in related fields, clinicians, patients, those whom Malcolm Gladwell has described as "tweakers"[128] (and whom others call "tinkerers"), even the simply serendipitous reader.

The role of these unforeseen contributors to innovation is not yet well understood, though there is some anecdotal and historical evidence on the topic. The famed historian Elizabeth Eisenstein argued that the printing press was critical to the Enlightenment because of its ability to "bridge over the gap between town and gown."[129] Economic historian Joel Mokyr, a leading scholar of the Industrial Revolution, likewise argues in a series of books,[130] that tools and customs which led to the dissemination of knowledge, rather than merely the production of knowledge, were essential to jump-starting the period of phenomenal economic growth in the West since the early 1800s.

It is crucial that we add computational devices to this list of unforeseen contributors. These machines are playing an ever-increasing role in the progress of science as they are used to gather materials to support the work of individual researchers or research networks. If research results are not machine readable and computable they will be closed to the most diligent—if not the most creative—of researchers regardless of their putatively open status. Scientists are just beginning to make use of computational abilities now employed in other fields; if Amazon can be searched by pictures, scientists should be able to search for tumors, bacteria, etc.

Identifying the unforeseen contributor as an important beneficiary of increased public access is not to downplay the role of established experts in a field. These experts benefit from increased access, but they are, in general, more likely to have access already; they are more likely to benefit from increased responsiveness—more freedom to use the information in a variety of ways, with fewer restrictions on what they can do with it (as described in sections 13 and 14).

There are many anecdotal examples of serendipitous readers who scan scientific work and end up making important breakthroughs; for example, Hugh Rienhoff, at the time a biotech venture capitalist, made the cover of *Nature* as a result of his independent work in genetics inspired by a sick daughter.[131] MIT Professor Eric von Hippel's work on the role of the knowledgeable customer on innovation has identified another source of commercially significant advances.[132] The computer industry was, both metaphorically and in a handful of famous cases quite literally, nurtured in garages by non-academics whose knowledge about computers was shared and tested among amateurs and experts alike in networks sometimes called "invisible colleges" by sociologists and historians.

The cost of equipment necessary for gene-related research is inexpensive enough today that biotech firms may also grow out of "garage-based" firms that do not have the resources to purchase subscriptions to the latest journals. For such firms, or similarly situated individuals, increased public access is critical.[133*]

* The political scientist Donald Stokes divided science into purely basic, purely applied, and 'Pasteur's Quadrant' at the intersection of the two. This style of research is typified by Louis Pasteur, whose research investigated basic scientific knowledge and directly developed important commercial products. Open-access policies are likely to be most important for research in this quadrant, since the development of new inventions which rely on basic science obviously benefits from access to new basic science results.

The vehicles for dissemination of information are also changing. The journal, OA or gated, will continue to play an important role. It may be that social media will take the place of venues like the "Home Brewed Computer Club" that nurtured Steve Jobs, Steve Wozniak, and many others. In a 2006 article entitled "The Effect of Open Access on Citation Impact," Yanjun Zhang examined two journals in the same field with roughly equal academic importance as measured by academic citations; one of the journals was completely open access and one was not. He then used a computer robot to scan the Internet for "informal" references, such as links from Wikipedia articles. The open-access journal's articles were referenced by informal sources almost twice as often as articles in the gated journal.[134]

David McKenzie and Berk Ozler, both of the World Bank, recently ran an experiment concerning latent interest in cutting-edge research among a group of economic development practitioners that included non academics.[135] They found that weblog links to a subset of open-access papers were able to increase readership of those papers by orders of magnitude—a single link from a popular economics weblog led to the equivalent of two years worth of readers. They also found, using a randomized sample, that they were able to induce development practitioners into reading new research by pointing out a new development weblog via email, and then linking to particular papers. Many of these practitioners rarely read new research on their own before this experiment. This result accords well with other anecdotal reports from private industry, NGO workers, clinicians, etc., concerning their interest in seeing relevant new research if only they could find what was interesting and access it without significant trouble.[136]

Two other studies point to the importance of the unforeseen contributor.

Karim Lakhani's study of InnoCentive.com, "Scientific Problem Solving through Broadcast Search: InnoCentive.com," describes the operations of a research network established by the pharmaceutical giant Eli Lilly to obtain solutions for problems that Lilly hoped non-affiliated experts, rather than their in-house R&D staff, would solve.[137] Both before and after InnoCentive was spun off by Lilly, it allowed "solvers" to join the network voluntarily, have access to the posted problems, and be able to propose solutions for a monetary reward—in effect initially outsourcing some of Lilly's R&D.

Solvers were not necessarily known beforehand by anyone at Lilly nor did they have to be pre-qualified as experts in the field of the problem or in any other field. They were, as a group, quite varied, including many people from different backgrounds, countries, personal and professional situations, and experiences."[138] They self selected for some reason—the monetary reward, whatever prestige might result from solving the problem, or perhaps curiosity. They could choose to attempt to solve one or many of the problems; together they were engaging in what Lakhani called "massively parallel processing." A substantial number of the winning solvers did not come from the field primarily associated with the problem. Lakhani sees this as confirmation of the "local-search phenomenon:" "the tendency to identify, define, and grapple with problems using local expertise, experience, knowledge and tools that had proven effective."[139] Because of the local-search phenomenon, if the problem was not solved by local resources it would lay dormant; even if it were solved, the solution might be "suboptimal due to the limited space from which the solution was sought." "Other possible alternative approaches were ignored—or simply not perceived."[140] Innocentive's experience confirms the findings of Murray et. al. and Williams on the value of a diversity of approaches and the creation of new avenues of exploration that greater access facilitates.

One other recent article illustrates the role of the unforeseen contributor in speeding up the process of innovation; this analysis explicitly draws on the open-source software model where the unforeseen contributor also figures prominently. In a commentary in *Nature Chemistry*, entitled "Open Science is a Research Accelerator," Michael Woelfle, Pierro Olliaro, and Matthew H. Todd describe a case study of an effort to produce an off-patent drug, "where open-source methodologies were employed to accelerate the process of discovery. The acceleration occurred because the project was open: relevant experts could identify themselves."[141] Woelfle et. al. set the stage by describing what they do when they face a problem they can't solve—in effect describing what Lakhani called the "local-search" phenomenon:

> [We] go to see our colleagues and ask for their advice. Our professional network is valuable. It is also limited. Perhaps there are people who are well placed to help us in another university or company in a different country, but we unfortunately don't know them. Surely science would proceed faster if we could reach these people. Or, better, if they could find us.[142]

In attempting to find a less expensive way to produce a lower-cost and more palatable drug for a ravaging parasitic infection called schistosomiasis (and commonly called Bilharzia), Woeefle and colleagues employed open-source methods including making publicly available their earlier and partially successful results, as a kind of "open lab notebook." Many of those who contributed new approaches were not previously known to them.[143] The open-source software axiom, "with enough eyes all bugs are shallow," seems applicable to the processes they employed; if enough people have access to a problem, there will be someone in the group with the right experience and insight and incentives to find a solution.[144]

The crucial message of their work is that the research was accelerated by being open: "Experts identified themselves and spontaneously contributed….The research therefore inevitably proceeded faster than if we had attempted to contact people…Perhaps this is not surprising but if it is the case that 'none of us is as smart as all of us' and if we wish to reach scientific goals quickly, why is so much science not practiced this way?"[145]

The authors cited other advantages of openness beyond greater speed. The "process is transparent and the public can be assured that funding arising from their taxes is being used responsibly and there is no suggestion of political interference."[146] Moreover, the research funding is leveraged because it was supplemented by valuable inputs from those who were not funded. Because everything is open and available the project "need not cease with the graduation of students or the demise of a principal investigator…" and the work is "subject to the most rigorous peer review because the review process never ends…."[147 *]

This section has made clear the importance of the unforeseen contributors who would benefit—along with society—from increases in public access to research results.

The next section will look at the potential economic impact of greater public access.

11. The Relationship Between Increased Public Access and Economic Growth

How might increased public access affect economic growth? Any estimate will necessarily be quite rough, but John Houghton (2010) has developed a useful framework, and has calculated the net present value gains of extending NIH-style policies to all federally funded research. Houghton provides a range of estimates with the low end being on the order of magnitude of 1.5 billion dollars (net of the costs of running a digital repository like PMC).[148] Of that figure, approximately 60 percent accrues directly to the U.S. economy; the rest benefits locales where researchers make use of the research. The benefits of the public-access policy are approximately eight times larger than the costs; this accords well with Furman and Stern's (2011) research on the value of BRCs and provides additional evidence on the need to fund the means of disseminating taxpayer-supported research, such as digital depositories.[149]

What does this estimate entail? The Solow model of economic growth, for which economist Robert Solow won a Nobel Prize, provides a method of separating the elements of economic growth into growth in the labor supply, the national capital stock, and general technical knowledge. One of the most robust findings

* While the focus of the commentary was the value of the open-source methodology in accelerating research the authors include a critique of traditional publishing: "Academia is associated with the free transmission of data and resources but in many ways this no longer is how it operates." The scientific community "communicates research results through publications relying on pre-publication peer review. Papers frequently omit some experimental information, or ignore negative results. The delays involved in publication of papers or reviewing of grants are significant. Many of us still publish papers in journals where comments on papers are not permitted, meaning that technical errors can remain uncorrected because rebuttals are usually required to be substantial works in their own right. Improvements in the existing state of the art are made through subsequent substantial and stand-alone articles where there can be significant delays arising from the peer review process of both the papers and the grant proposals required to fund the work."

of the Solow model is that increases in the base of knowledge are by far the dominant explanation for increased American prosperity over the twentieth century.

The standard version of Solow's equation used to study R&D allows the estimation of a parameter which shows how much a given change in the stock of technical knowledge contributes to the broader economy. The stock of research knowledge is assumed to depreciate like any other asset every year. (The U.S. Bureau of Labor Statistics uses the working assumption that basic knowledge does not depreciate in usefulness, but that the worth of an average piece of applied science depreciates at 10 percent per year as its applications become less economically important.) Since 20 percent of federal funding goes to basic research, this means that all federally funded research depreciates at approximately 8 percent per year. Houghton modifies the standard Solow equation by adjusting the stock of R&D for accessibility (economic benefits from research can only be gained when practitioners can access the research).[150] After that adjustment, the increase in economic output due to expanded public access can be found by estimating: the marginal rate of return of scientific research, the relative importance of R&D to the overall economy, and the increase in accessibility caused by the application of the NIH policy across the federal research enterprise. These estimates can be placed in a dynamic model which values the economic gains resulting from the production of, and access to, a body of scientific knowledge at any given time in the future.

The marginal rate of return of basic research has generally been estimated to be quite high. Salter and Martin (2001) review the literature on social rates of return from R&D spending, and find estimates ranging from 20 to 160 percent.[151] The leading researcher on this topic, Berkeley's Bronwyn Hall, noted in a 2009 review of these estimates that "on the whole, although the studies are not fully comparable, it may be concluded that R&D rates of return in developed economies during the past half century have been strongly positive and may be as high as 75 percent or so, although they are more likely to be in the 20 to 30 percent range."[152] Houghton uses the conservative lower bound of 20 percent. Of this return, Adam Jaffe and others have estimated about two-thirds to three-quarters of the social gains of research in the United States accrue to American firms and consumers.[153*]

While the social return from basic research is well-studied quantitatively, estimates of the economic impact of problems in obtaining access to research are less precisely estimated. Ware (2009) surveyed British small and medium enterprises concerning access difficulties, and found that gated-access policies caused them to avoid reading ten to twenty percent of research articles they considered important.[154] In the absence of any federal policy, approximately 20 percent of all research articles are already open-access.[155] Open-access policies increase readership and citations in future research. Based on similar data, Houghton uses 5 percent as the potential increase in accessibility to federally funded research for applied scientists as a result of the extension of the NIH mandate to all federally funded unclassified research—a relatively conservative estimate.

Running a digital repository is relatively inexpensive—arXiv spends about $7 per article; NIH costs per entry are somewhat higher because they provide some additional functionality.[156] NIH has estimated that the annual operating costs for PMC including the ingestion of articles, refinement of submissions, creation and use of search tools, staffing of the help desk and the staffing and operation of a central coordinating office for NIH at $3.5 million to $4.0 million. Extending the NIH policy to other federally funded extramural research would roughly double the number of articles covered, so costs would definitely increase; the amount of the increase would depend on how an extended public-access policy would be implemented, e.g. through centralized versus decentralized repositories etc.

[*] Publishers have criticized the NIH policy because its benefits do not accrue solely to the United States and because manuscript deposited in PMC may be subject to piracy. As the director of NIH has pointed out, "all the articles that are publicly available in PMC are already accessible in electronic form from outside the United States (via subscription-based internet access and from libraries) at the time they become available in PMC." Letter to Honorable Joseph R. Pitts from NIH director Francis Collins MD Ph.D dated December 16, 2011. We know of no publisher who voluntarily restricts subscriptions to publications containing reports on NIH funded research to those in the United States or who voluntarily limit access by non US citizens who subscribe. Given the increasing percentage of research being conducted outside the U.S. we should be encouraging greater openness worldwide.

As to piracy the NIH Director noted in the same letter that NIH has "established sophisticated monitoring systems to protect publisher interests and prevent piracy." http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf

Translating these increases in accessibility into GDP gains requires making assumptions about the lag between basic science and its application in new products; in general, Houghton errs on the side of long lags and low social returns to R&D. The $1.5 billion net gain would be more than five times higher, yet still reasonable, if higher estimates of the social return of R&D were used. Houghton's sensitivity analysis of the possible gains shows that a positive net welfare gain from extending the public-access mandate occurs under even the most conservative estimates of the parameters used in previous economic research on social returns to R&D.

Section 8 described evidence that in addition to increasing the diffusion of scientific knowledge, increased public access may also increase the speed at which such knowledge diffuses. To the extent that researchers can use easily searchable digital repositories to have both quicker and broader access to cutting edge research, Houghton's estimates may be too low because his model assumes that the lag between basic research and applied products is unaffected by policies that increase public access.

12. Are Mandates Necessary Given the Benefits of Increased Access?

Increasing public access has been shown to have substantial benefits ranging from increased citations, to more broad-based gains in the progress of science and in a society's greater economic growth. But will the existence—and recognition—of these gains lead to free and ubiquitous access to the results of federally funded research (either through publication in OA journals or by immediate or briefly delayed deposit in a digital repository) in the absence of a governmental mandate? The prospects are unclear.

One major obstacle is the nature of the STM publishing market which economists describe as "two-sided." In this marketplace the burden of the costs of increased access are borne by the researcher but the benefits accrue mainly to the society at large.

OA journals charge authors a non-trivial fee to publish their articles. That fee is borne either directly by the author or indirectly by his or her funder.[157] The gains from having an openly accessible article include increased citations, which benefit the academic researchers and enhances their personal reputations. But most of the gains accrue to society. Since the total benefits (public plus private) cannot be captured by the authors, their incentive to pay a fee to publish in an OA journal or to take the steps necessary to voluntarily deposit copies of their work in an open digital repository (particularly if traditional publishers discourage it) is less than would produce the maximum benefit for society. According to economic principles an under-supply of articles for OA journals or digital repositories is the likely result of a voluntary policy.

The incentives for researchers to support increased access would be greater if their contributions to enhanced openness were more directly rewarded by the traditional means of recognition in the academy. As the Berlin Declaration pointed out, there will need to be new mechanisms to recognize accomplishments in scholarship in a more open research environment. For example, the immediate posting of research results may accelerate scientific progress as was seen in the Human Genome Project, but it may lead to lost publishing opportunities for the researcher as the public posting is considered by some publishers to constitute a "publication" and they will not publish work that has already been "published." If greater openness is to be encouraged, then the criteria for granting academic recognition—e.g. via tenure and advancement decisions—will need to be rethought.

In addition to this economic disincentive, Kling and McKim (2000) have pointed out that customs and not simply cost-benefit analyses cause many of the differences in publication practices across academic disciplines; this can be seen in the contrasting practices in high-energy physics where almost all research is available through preprints and chemistry and medicine where access is much more restricted. Such customs can take a long time to change in the absence of an outside force.[158]

The societal benefits of increased public access described above are substantial. They justify efforts by funders, particularly governments, to increase public access. Voluntary efforts—such as NIH's voluntary deposit policy—did not succeed in the past and do not appear to be particularly successful now as demonstrated by the results of voluntary efforts to preserve university holdings.[159] A growing awareness of the benefits of greater openness—and a recognition that the value of easily searchable and interoperable digital repositories increases as their holdings increase—has led to an increasing number of mandates to increase access around the world.[160]

13. Would a Simple Extension Of The NIH Policy to Other Federal Funders Capture the Potential Benefits of Greater Openness?

Simply increasing public access by extending the present NIH policy to other federal funders would not deliver anywhere close to the full benefits of greater openness. The NIH policy ensures that those interested will be able to locate and read peer-reviewed manuscripts accepted for publication based on research funded by NIH. This policy has led to important benefits even in the limited time since its adoption. But the policy does not deal with restrictions on how the research results can be used and built upon, particularly by other researchers. Merely extending the policy would not provide the greatest possible return on the public's investment in research either in terms of scientific progress or in benefits to the society such as faster commercialization of research or greater economic growth.

Openness, according to CED, includes not only accessibility but also responsiveness—being able to do something with that which is accessible. The next section addresses the future of the research enterprise and the benefits to be gained by researchers from the removal of restrictions on how they can use taxpayer funded research to which they now have improved access.

14. The Future of the Research Enterprise: Overcoming Use Restrictions and Other Impediments to Openness

The scholarly journal has a long and distinguished history having provided perhaps the most valuable means of diffusing scientific research results over the last several hundred years. The NIH policy has certainly increased access to journal manuscripts (and articles) but only provides a starting point for developing other means for increasing responsiveness and the ability to use what is available, dramatically increasing its value. This section looks at some of these further steps.

Moving Beyond the NIH Public-Access Policy

Even today, NIH recognizes that read-only versions of manuscripts can be made more valuable. PMC is converting deposited articles into XML, allowing them to be more easily searched and linked. It is providing a link from the article to the publisher's website if the article is deposited by the publisher. It is linking articles to other sources, including openly available data such as NIH databases on clinical trials, gene sequences, protein structures and chemical compounds. To facilitate searches, particularly by non technically trained users, NIH is using natural language processing to direct users to other articles of possible interest.

For the active researcher at the heart of the research enterprise, even this access is not enough. Researchers traditionally have been trained and certified in a particular discipline and were expected to have command of the key articles in the field, sometimes three or four hundred. But today's explosion of research and new tools makes it possible to imagine that the same researcher may be able to extend his or her scholarly reach, having command of two or three or more times that many articles and processing a greater percentage of the flood of new research. And a 21st Century definition of mastery will necessarily extend beyond articles to encompass critical data sets and emerging new tools for search, discovery, and analysis.

One challenge for a 21st Century open research infrastructure is to get the researcher the right information at the right time. Researchers should be able to create, maintain, and utilize their own personalized collections. This will require not only access but also the discovery tools needed to search the broadest possible range of potentially valuable literature and to identify items of interest, whether they be articles or manuscripts or what until now were thought of as sub parts of an article—the text, the underlying data, the protocols followed, the tools developed, the computer programs employed etc. Discovery might involve use of scholarly reviews, recommendations from colleagues, reputational or ranking systems, comments in social media, or the results produced by machine readers that hunt relentlessly, using their algorithmic senses to discover, digest, and contextualize data from machine-readable sources. Accordingly, the focus of what needs to be accessible needs to be expanded and the notion of the article as the unit of measure of accessibility reconsidered. In a major change, debates about public policy may well shift from the manuscript to an "accessible object" which might consist of any or all the sub parts that make up the manuscript; the debates will not be about the ability to read a manuscript but rather about being able to more fully use it or one or more of its component parts.

A challenge for a 21st Century open research infrastructure is, in CED's terms, to increase the responsiveness of what is now accessible. In essence, the challenge is to provide the researcher with the opportunity to use the materials to extract as much value as possible by whatever means. This goes far beyond reading; it would be as if all the published literature was available in one searchable database. But what are the limits, if any, on use and reuse?

The Berlin Declaration defines use and reuse very broadly for the purposes of defining open access: for an object to qualify, the reader must have the ability to "copy, distribute, transmit, and display the work publicly, make and distribute derivative works in any digital medium for any responsible purpose," subject to proper attribution to the author. Such open access goes far beyond the public access embodied in the present NIH policy.

There are many incremental steps that can be relatively easily imagined and implemented between what is called for by the NIH policy and what is described in the Berlin Declaration. A mandate could require that the manuscript and its sub parts must be machine readable and capable of being copied, distributed, displayed, and linked to other sources. It could specify that the manuscript and its subparts be available for both data and text mining and be subject to computation and translation into other languages. Requirements regarding the availability of specified metadata could be put into place to facilitate search, discovery, indexing, and contextualization; competition in the development of new tools could be enhanced by mandating that a researcher be able to extract the manuscript and its subparts to be analyzed using tools chosen by the researcher.

Each step in the broadening of access and the elimination of restrictions on use is likely to raise new issues. The issues may be different for each of the groups who would benefit from greater openness, for example, the individual researchers, the networked collaborators, the infrastructure providers such as digital repositories and the citizen scientists. The issues are also different for an article, its different sub parts—particularly the underlying data—and for large public data sets such as those of de-identified patient records.

The issue of the accessibility and responsiveness of data is likely to be at the heart of the debates about greater openness. The research enterprise of the 21st century thrives on the accumulation and manipulation of data, from images of the cosmos produced by robotic telescopes scanning the skies to genetic information spewed from thousands of gene-sequencing machines to the accumulated medical records of hundreds of millions of patients particularly through advanced text and data mining techniques. This change is fundamental. In the past the researcher formulated a hypothesis and then gathered data to test it while today, awash in data, the researcher looks at the data in order to formulate the hypothesis.

The OECD's 2004 Ministerial meeting focused on the issue of access to government-funded research and led to a 2007 OECD Recommendation Concerning Public Access to Research Data from Public Funding, which noted the benefits of increased access to research results and the underlying data:

- Reinforces open scientific inquiry,
- Encourages diversity of analysis and opinion,
- Promotes new research,
- Makes possible the testing of new or alternative hypotheses and methods of analysis,
- Supports studies on data-collection methods and measurement,
- Facilitates the education of new researchers,
- Enables the exploration of topics not envisioned by the original investigators, and
- Permits the creation of data sets when data from multiple sources are combined.[161]

The OECD Recommendation, citing a U.S. National Research Council Report, stated: "The value of data lies in their use. Full and open access to scientific data should be adopted as the international norm for the exchange of scientific data derived from publicly funded research." Access should be "easy, timely, user-friendly, and preferably Internet based." Of particular interest to governments were that such access "helps to maximize the research potential of new digital technologies and networks, but provides greater returns from the public investment in research."[162]

Although greater openness for data can provide significant benefits it also raises difficult issues regarding privacy, security, data integrity—and many others. These are issues that are not easily resolved. This suggests that a relatively simple mandate such as the NIH policy directed at manuscripts is not sufficiently nuanced to be easily applicable. One example: the study of Alzheimer's disease is likely to benefit from the gathering of the broadest array of Alzheimer's patient records. The Institute of Aging is doing just that. But large data sets funded by the Institute as recently as six years ago could not be connected with those now being gathered; consent for further sharing of the data was not obtained in the earlier studies because no one imagined that the data might be made available beyond the researchers then engaged. Simply saying that this data should be made available within "x" months would not be adequate for the complex issues that need to be resolved. (NIH has made progress with its data policies, so it can resolve this and other issues.)

Another current example: There are difficult technical and policy issues surrounding the use of even anonymized or de-identified patient data because it is reasonable to believe that, with enough time and effort, any data can be linked with other data resulting in a non-trivial chance of re-identification of the patient. In an earlier report the CED recommended that the National Academies of Science and Engineering be tasked with recommending the technological steps necessary to show a good faith effort to protect data privacy through anonymization and de-identification.[163]

Huge collections of data pose a different problem —there may simply be too much information for the relatively small number of scientists focused on the data to make sense of it. But the Internet provides an opportunity to use the collective brainpower of unforeseen contributors to solve problems through what is being called crowd science or crowd-sourced science, or citizen science.[164]

Researchers have used crowd science to classify galaxies and solar explosions and in more than 400 other projects.[165] Other researchers have taken crowd science a step further, for example inviting citizen scientists to use video game-like tools to attempt to fold a protein molecule—how a protein is folded is a problem that must be solved before research on certain diseases can progress. In September 2011, one of the teams of players in the video game FoldIt successfully "deciphered the folding of a protein important in AIDS research."[166]

Yet another issue with data is that it is often treated as proprietary:

> A scientist works long and hard to generate original data, and then expects to reap the reward in the form of publishing the first research paper to describe some new phenomenon. She is not going to want share this data with others, particularly strangers, any more than say, an investigative reporter would want to share his notes before a story has been written. Harnessing 1,000 people requires sending your data out into the world —something that science is loath to do. The scientist's interest in keeping things private and getting credit, in other words, is directly opposed to society's interest in tackling some problems with a hive of the best minds.[167]

Any policy that seeks to improve the productivity and effectiveness of the research enterprise of the 21st Century will need to find ways to provide incentives for actions that increase the openness of scientific research, such as the timely sharing of data or tools. This is an area where the government can play a leading role through its own practices such as recognizing actions by researchers in supporting greater openness when evaluating their grant applications.

Other issues will need to be addressed if restrictions on use are to be reduced and openness expanded. The need for attribution has been part of the vision of the Open Access movement from the earliest days. Most researchers are aware of materials that have been used without proper attribution; given the role that published materials have traditionally played in tenure and advancement decisions it is not surprising that many researchers tend to treat their work products as proprietary. If the practice of open or networked science becomes customary these inclinations may be reduced but greater attention will need to be paid to how to ensure attribution, encourage sharing, and discourage scholarly "piracy."

Many licenses governing copyrighted materials are restrictive and not machine readable. The development of licenses such as Creative Commons' CC-BY license, which allows greater use, requires attribution, and is machine readable, is helpful. But in the world of digitized information, license enforcement may not be easy. (Some Open Access advocates have supported the placement of materials into the public domain to increase openness, but recent cases suggesting that materials in the public domain can be "restored to copyright" and open-access archives made proprietary raise concerns about a public-domain strategy.)[168]

Other issues arise concerning unlimited reuse and derivative works. Unlimited reuse could, for example, involve the creation of mash-ups incorporating another's research into one's own. Copyright is protective of authors' rights to control the use of their own work in someone else's creation—which may sometimes be problematic given the cumulative nature of progress in science. And the law has not settled on a sustainable course in the new digital world; for example it has struggled with "sampling," whereby someone uses small bits of a creative work such as a song in the creation of a new work. How should the law treat sampling if the object sampled was not a song but a data set critical to a study of the safety and efficacy of a new drug? The stakes would be far higher than for music. A whole range of interests must be taken into account including society's interest in the creativity and experimentation critical to scientific discovery.

These are not brand new issues. And the list above is hardly exhaustive—just dealing with the impact of intellectual property issues or the administration by colleges and universities of the Bayh-Dole Act would discourage the faint-hearted. But a good deal of creative work has already been done by NIH, other agencies, the Institute of Medicine among others, and there are many initiatives aimed at finding acceptable solutions taking into account the very wide range of interests. There is, however, not yet a consensus on these issues. We would encourage others to help identify solutions which will best serve society's interest in creating the most robust research enterprise possible.

Digital Repositories in a 21st Century Open Research Infrastructure

Digital repositories are likely to play a key role in enhancing openness in a 21st Century Open Research Infrastructure. At present, the coverage of digital repositories is hardly ubiquitous. They are often silos, open but isolated from each other. If the research enterprise is to thrive there must be agreed-upon standards that facilitate search and discovery and ensure interoperability. There must also be agreed-upon rules for access, when factors such as competition or local custom may lead to discrimination in access, the favoring of particular proprietary tools, or the isolation of key materials etc. Such standards and rules should build upon existing efforts. Whether repositories are centralized or decentralized the measures of their value for the research enterprise will be the scope and scale of the resources that can be easily and efficiently shared and their utilization for scientific discovery; if a decentralized route is taken there must be means of assuring that all participants understand and follow the same rules to ensure access and interoperability. Because of the issues discussed earlier, it would not be prudent to rely on a system of solely private archives.

Care must also be taken to dovetail efforts to broaden access and remove restrictions on use with the efforts directed at the preservation of materials. While reducing barriers to access and use increases the likelihood of preservation by allowing copies of materials to be present in many locations, there are new issues of data management and versioning which arise.

15. Findings and Conclusions

The central question for this paper has been the impact of the NIH public-access policy and proposals to extend policies to increase public access to research supported by other federal funders based on their effects on the long-term development and dissemination of high-quality scientific, technical, and medical research and scientific progress.

The available evidence, underlying the findings set out below, demonstrates the value of the current NIH public access policy and supports the extension of public policies to increase public access to research supported by those federal agencies that provide significant amounts of funding for unclassified extramural research. In order to obtain even greater benefits from increased openness, restrictions on the use of the accessible research results need to be reexamined and reduced.

This research shows the following:

- The NIH mandatory deposit policy has led to a substantial increase in access to the results of NIH-funded research.
- Policies that increase public access to research results:
  - accelerate progress in science. They speed up and broaden the diffusion of knowledge and allows researchers to get to the frontiers of knowledge in their fields more quickly. Progress is enhanced by the use of digital repositories and other means of improved dissemination such as biological resource centers and gene banks that are open, easily accessible and searchable, and interoperable expanding the amount and the scope of the materials available.
  - have differing effects on differing groups. Researchers within a given field at major research institutions in the U.S. and Europe generally have access to journals within their field; researchers in fields such as high-energy physics where there is a culture of ubiquitous access to preprints even before articles are published in journals are also able to stay abreast of relevant research relatively easily. But there are substantial numbers of potential readers who have difficulty in obtaining access to research results that are of interest to them, including academic researchers in unrelated fields or at institutions with fewer resources, researchers who have left academia, professionals in large and small enterprises attempting to find research useful for their businesses, clinicians, and patients, among many others. Their access difficulties result in their wasting time and effort in locating and obtaining access. Access problems for businesses raise costs and cause delays in the development of new products and services. The contributions that those with access difficulties can make to scientific progress or economic growth are delayed and diminished. And as progress in science is cumulative, society loses the benefit not only of their potential contributions but the contributions which would subsequently have been built upon their work.
  - lead to greater diversity in researchers, the exploration of more varied research paths, and more experimentation with more solutions to a given problem making it more likely that the solution will be optimal. In so doing it ameliorates the "local-search phenomenon" where the solution set to a given problem is limited by the experiences and knowledge of a more restricted group of potential problem solvers.
  - generate more follow-on research, more academic citations, greater visibility for authors, and heightened prestige for the research, and are particularly valuable to academic researchers. Actions that support greater openness would likely increase if there were effective incentives for them. For academic researchers the most effective incentives lie in academic processes such as those leading to tenure or advancement decisions. These processes for providing recognition need to be rethought in order to foster acts that support greater openness, even those that do not result in publications, including the early release of research results and the sharing of new research tools. Because of its importance to academic recognition, the value of a scholarly work should be judged on its own impact not only on where it is published or made accessible. Government recognition of contributions made in support of greater openness in the government grant making process would increase the incentives for greater openness and would reflect the same values that led to the adoption of increased public access policies.
  - provide greater transparency in the research enterprise and enable improvement in the administration and evaluation of the overall research effort promoting greater accountability. Enhanced openness facilitates the continuing review and critique of research results, assists in error detection, discourages the reporting of selective results, and helps researchers avoid

duplicative research and the pursuit of dead-end lines of research.

- allow for the fullest possible use of the research results and heighten the return on government investment in research. They also prevent taxpayers from having to pay twice to support research—once through government grants and then again to obtain access to the results through subscription payments.

- promote greater economic activity and economic growth. Gains in GDP associated with increased public access far exceed the costs of providing it.

- stimulate enhanced "intensity" in follow-on research, moving it more quickly toward applied research and commercialization into new products and services. By speeding commercialization greater openness accelerates job creation. A whole new industry sector is now emerging based on adding value to newly available research results.

- merit other forms of funder support. Research funders should be prepared to consider applications for funds for "author-pays" forms of publication. In making funding decisions, the importance of (and the high return on investment in) the mechanisms for dissemination and diffusion of knowledge, such as digital repositories, biological resource centers, and shared data banks, should be carefully weighed. Grant applications should require plans to increase the dissemination of results and those plans should be considered as part of the evaluation of the grant application.

- have not substantially harmed the traditional subscription-based STM publishing industry over the last four years, based on the available evidence at this time. There is no persuasive evidence that increased access threatens the sustainability of traditional subscription-supported journals, or their ability to fund rigorous peer review (which is also utilized by OA journals such as those published by PLoS and BioMedCentral). There is no evidence of a significant reduction in the number of high-impact journals, or that a significant reduction of traditional publishing outlets is likely, or that there will be any shortage of outlets for high quality research that would threaten the amount or quality of research because of policies increasing public access. The most important factor affecting proprietary STM publishers since 2008 has been the economic downturn in the United States that hit universities particularly hard and left the publishers' most important customers with less and less money to spend on scholarly publications.

- have such an important set of benefits that any delays in its application should be minimized. There is no persuasive evidence supporting the need for a twelve-month delay before an author's accepted manuscript must be deposited in PMC; a maximum six-month delay, as adopted by many funders, both governmental and private, would allow the speedier harvesting of the benefits of greater openness including the acceleration of scientific progress. There is evidence that even shorter delays could be adopted without risk. The length of delays, if any, might differ depending on the rate of change within a field, but the burden of proof should be on those who advocate greater delay.

- no longer should focus on the journal article as the only possible unit of measure for increased access. More attention should be given to considering the accessibility and responsiveness of the article's subparts. The article and its subparts can provide the most value when machine readable, subject to downloading, data and text mining, copying, display, distribution, linkage, computation, translation into other languages, and extraction for analysis with tools of the researcher's choice. Issues surrounding the use of data and the creation of new works using accessible materials ranging from privacy and security to appropriate data management practices and intellectual property rights require further analysis and discussion in order to build greater consensus on the appropriate steps to foster greater openness.

- must address issues of attribution which have been integral to calls for greater access to knowledge since the early days of the Open Access movement. More permissive licenses, such as the Creative Commons CC-BY license, which are machine readable and require attribution, allow realization of greater societal benefits.

A system of interconnected and interoperable digital depositories lies at the heart of a 21st Century Open Research Infrastructure. The goal should be the development of a publicly accessible shared corpus of research results as close to all inclusive as possible. There are several steps the federal government can take to maximize the benefits of all scientific research while minimizing both public and private costs. Among these are:

- Convening the appropriate stakeholders e.g. researchers, funders, research institutions, depository operators, and publishers, to recommend metadata standards, based on existing efforts, so that the various accessible objects can be located, described, indexed and contextualized.
- Convening the appropriate stakeholders regarding depositories to recommend (based to the extent possible on existing work) standards for access and interoperability and rules to prevent unreasonable discrimination in accessibility and responsiveness, the isolation of key materials, the continuation of research "silos," or other actions that would inhibit the emergence of new tools or otherwise thwart innovation. Whether a centralized or decentralized system of digital depositories emerges all depositories will need to follow the same rules regarding accessibility and responsiveness in order to maximize interoperability and the value of the shared research corpus. A solely private archival system is unlikely to realize the potential benefits of ensuring access to NIH-funded research.
- Ensuring coordination between policy efforts to increase accessibility and responsiveness of government-funded research and efforts to promote the preservation of digital objects.
- Attempting to achieve as much uniformity and simplicity in requirements for increased public access across federal funding agencies as possible, given the fact that researchers and institutions may be supported by multiple agencies and conflicting or inconsistent or overly burdensome policies increase confusion and the cost of compliance.
- The National Academies of Science and Engineering should recommend what technological steps should be taken to have acceptable, good-faith privacy and security standards for personally identifiable information, including anonymization and de-identification of patient records.

Endnotes:

1. Thomas Lin, "Cracking Open the Scientific Process," New York Times, January 16, 2012.

2. House Report 108-636 was adopted in 2004, suggesting a mandatory access policy for NIH funded research. See http://thomas.loc.gov/cgi-bin/cpquery/?&db_id=cp108&r_n=hr636.108&sel=TOC_338641&.

3. At the time, there were other articles in PubMed Central that came from journals that were making all of their articles—NIH funded or not—available by depositing them in PMC .

4. PL 110-161, Appropriations Act of 2008, Section 218, Division G, Title II. Available at http://thomas.loc.gov/cgi-bin/query/z?c110:H.R.2764.enr:

5. Ibid.

6. NIH Public Access Policy: Overview and Recent Developments, Jerry Sheehan, Assistant Director for Policy Development National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services September 28, 2011.

7. For example, the Association of American Publishers letter to the Director of the National Institutes of Health < http://www.pspcentral.org/publications/Letter%20to%20Zerhouni%20NIH%20-%20AAP-DC%20Principles%204-16-08.pdf> lists many of these complaints. Similar white papers listing the concerns of the publishing industry can be found in the Congressional Record of hearings on the Fair Copyright in Research Works Act and the Federal Research Public Access Act.

8. Committee for Economic Development, "Harnessing Openness to Improve Research, Teaching and Learning in Higher Education," Committee for Economic Development, 2009. "Harnessing Openness to Transform American Health Care," Committee for Economic Development, 2008. "Open Standards, Open Source and Open Innovation: Harnessing the Benefits of Openness," Committee for Economic Development. These reports can be found at http://www.ced.org.

9. See, for example, the Berlin Declaration on Open Access, available at http://oa.mpg.de/berlin-prozess/berliner-erklarung/.

10. Anne Gentil-Beccot, Salvador Mele and Travis C. Brooks, "Citing and Reading Behavior in High-Energy Physics." Scientometrics 82(4), 2010.

11. Steven Harnad. "Subversive Proposal" Presented at the Network Services Conference, November 28-30, 1994. Transcribed online at https://groups.google.com/group/bit.listserv.vpiej-l/browse_thread/thread/682843612b4ff4d/62b3bd13407dc1d8.

12. "Declaration of the Budapest Open Access Initiative," available at http://www.soros.org/openaccess/read.

13. "Bethesda Statement on Open Access Publishing," 20 June 2003, available at http://www.earlham.edu/~peters/fos/bethesda.htm.

14. Berlin Declaration on Open Access, available at http://oa.mpg.de/berlin-prozess/berliner-erklarung/.

15. OECD, "Working Paper on the Information Economy," Sept. 2, 2005, available at http://www.oecd.org/dataoecd/42/12/35393145.pdf.

16. European Commission Open Access Initiative, available at http://cordis.europa.eu/fetch?CALLER=EN_NEWS&ACTION=D&RCN=29774.

17. See OpenAIRE: Open Access Infrastructure for Research in Europe, available at www.openaire.eu

18. "Policy on Enhancing Public Access to Archived Publications Resulting from NIH-Funded Research," available at http://grants.nih.gov/grants/guide/notice-files/not-od-05-022.html.

19. PL 110-161, Appropriations Act of 2008, Section 218, Division G, Title II. Available at http://thomas.loc.gov/cgi-bin/query/z?c110:H.R.2764.enr:

20. "Policy on Enhancing Public Access to Archived Publications Resulting from NIH Funded Research," available at era.nih.gov/docs/ePUG_Public_Access.ppt.

21. Prepared testimony of Davd J. Lipman, M.D, Director, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services before the Subcommittee on

Information Policy, Census, and National Archives, Committee on Oversight and Government Reform, United States House of Representative, July 29, 2010 p.1.

22. See the UK MRC Position on Open Access at http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Openaccesspublishing/Position-statement/index.htm.

23. http://www.peerproject.eu/.

24. Cited in Open Access Update, Caroline Sutton, Publisher, Co-Action Publishing, President, Open Access Scholarly Publishers Association (OASPA), ISMTE Meeting, August 9, 2011, Washington, D.C.

25. Alok Jha, "Results of Publicly Funded Research Will be Open Access," The Guardian, December 8, 2011, available at http://www.guardian.co.uk/science/2011/dec/08/publicly-funded-research-open-access.

26. BIS, "Innovation and Research Strategy for Growth," available at http://www.bis.gov.uk/assets/biscore/innovation/docs/i/11-1387-innovation-and-research-strategy-for-growth.

27. Breakdown of Congressionally Funded R&D by type and agency available at http://www.aaas.org/spp/rd/fy2012/total12c.pdf.

28. Hearing on the Fair Copyright in Research Works Act, available at http://judiciary.house.gov/hearings/printers/110th/44326.PDF.

29. See http://www.federalregister.gov/articles/2011/11/04/2011-28621/request-for-information-public-access-to-digital-data-resulting-from-federally-funded-scientific and http://69.175.53.20/federal_register/2011/dec/23/2011-32943.pdf.

30. Hearing on the Fair Copyright in Research Works Act, available at http://judiciary.house.gov/hearings/printers/110th/44326.PDF.

31. See the counter at the PubMed Central website, http://www.ncbi.nlm.nih.gov/pmc/.

32. Letter to the Honorable Joseph R. Pitts, Chairman, Subcommittee on Health, House Energy and Commerce Committee, United States House of Representatives from Francis S. Collins, M.D., Ph.D., Director, National Institutes of Health at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

33. NIH Public Access Policy: Overview and Recent Developments, Jerry Sheehan, Assistant Director for Policy Development National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services, September 28, 2011.

34. See comments by the Partnership for Research Integrity in Science & Medicine (PRISM) at http://www.prismcoalition.org/problems_threats.htm. See also http://www.publishers.org.

35. Ibid.

36. Ibid.

37. Ibid.

38. See comments of Representative Berman during the Congressional Hearing on the Fair Copyright in Research Works Act, available at http://judiciary.house.gov/hearings/printers/110th/44326.PDF.

39. Ibid.

40. Ibid.

41. Comments of former Register of Copyrights Ralph Oman during the Congressional Hearing on the Fair Copyright in Research Works Act, available at http://judiciary.house.gov/hearings/printers/110th/44326.PDF.

42. See http://www.doaj.org.

43. Claudio Aspesi, "Reed Elsevier: The Inevitable Crunch Point," Bernstein Research, March 10, 2011. The same analyst has written a number of other proprietary reports on STM publishing for Bernstein from a equity investor perspective.

44. Ibid.

45. Association of American Publishers, Response to White House OSTP Request for Information Regarding Public Access, January 12, 2012, available at http://publishers.org/_attachments/docs/library/01-10-2012%20ostp%20publications%20rfi%20aap-psp%20response%20-%20final.pdf.

Endnotes:

46. Claudio Aspesi, "Reed Elsevier: The Inevitable Crunch Point," Bernstein Research, March 10, 2011.

47. Ibid.

48. Association of Research Libraries, ARL Statistics 2009-2010, available at http://publications.arl.org/ARL-Statistics-2009-2010/.

49. PRISM, Myth vs. Fact, available at http://www.prismcoalition.org/myth.htm.

50. SPARC Executive Director Heather Joseph during the Congressional Hearing on the Fair Copyright in Research Works Act, available at http://judiciary.house.gov/hearings/printers/110th/44326.PDF. The PEER activity in Europe may provide additional useful data.

51. Letter of Francis S. Collins at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf

52. Mike Peine, "2010 Study of Subscription Prices for Scholarly Society Journals," Allen Press, available at http://allenpress.com/system/files/pdfs/library/ap_journal_pricing_study_2010.pdf.

53. Publisher's Communication Group, "Journal Subscription Renewal Trends, 2006-2011," available at http://www.pcgplus.com/pdfs/Journal%20Subscription%20Renewal%20Trends.pdf.

54. Ibid.

55. Letter to The Honorable Darrell E. Issa from Gerald Perry, AHIP, President, Medical Library Association, and Gary Freiburger, AHIP, President, the Association of Academic Health Sciences Libraries, January 30, 2012.

56. Peter Suber, "Publishers launch an anti-OA lobbying organization," Open Access News, August 23, 2007, available at http://www.earlham.edu/~peters/fos/2007/08/publishers-launch-anti-oa-lobbying.html.

57. PRISM, Comments on Federal Research Public Access Act, available at http://www.prismcoalition.org/legislation_2695.htm.

58. Claudio Aspesi, "Reed Elsevier: Bringing Down the House," Bernstein Research, March 29, 2011.

59. "From 2007 to 2011, the number of biological sciences and agricultural journals and medicine and health journals grew by 15 percent and 19 percent, respectively. In the same time span, the average price of biology journals and health science journals have increased 26 percent and 23 percent respectively." Letter of Francis S. Collins at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

60. PRISM, Comments on Federal Research Public Access Act, available at http://www.prismcoalition.org/legislation_2695.htm.

61. PRISM, Myth vs. Fact, available at http://www.prismcoalition.org/myth.htm.

62. Ibid.

63. PRISM, Comments on Federal Research Public Access Act, available at http://www.prismcoalition.org/legislation_2695.htm.

64. Ibid.

65. Ibid.

66. PLoS 2010 Progress Update, available at http://www.plos.org/media/downloads/2011/2010_PLoS_Progress_Update_lo.pdf.

67. Sir Richard Roberts et. al, Open Letter to the US Congress concerning the NIH Public Access Policy, September 9, 2008, available at http://www.taxpayeraccess.org/supporters/scientists/an-open-letter-to-the-us-congress-signed-by-33-nob.shtml.

68. PRISM, "Problems with Government Intervention – Government Spending," available at http://www.prismcoalition.org/problems_costs.htm.

69. Ibid.

70. Letter of Francis S. Collins at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

71. Joshua Gans, "Berkeley Electronic Press Closes Up Journals," Digitopoly, January 26, 2012, available at http://www.digitopoly.org/2012/01/26/berkeley-electronic-press-closes-up-journals/.

72. Alan Garber, "Harvard Response to the White House RFI on OA Publications," January 4, 2012, available at http://osc.hul.harvard.edu/stp-rfi-response-january-2012.

Endnotes:

73. For example, the New England Journal of Medicine makes articles open access after a 6 month embargo, whether those articles are NIH-funded or not.

74. Mark J. McCabe and Christopher Snyder, "The Economics of Open Access Journals," July 2010, available at http://mccabe.people.si.umich.edu/EOAJ_2010.pdf.

75. The UK and Australia, among others, implicitly use some estimation of journal quality when judging the overall quality of research produced by a university. This has large impacts on funding. Tenure reviews in many fields also implicitly use the prestige of the journal that a tenure candidate's work has been published in as a proxy for the quality of the underlying article. See, for instance, http://www.timeshighereducation.co.uk/story.asp?storycode=403502.

76. Shane Greenstein, Josh Lerner and Scott Stern, "The Economics of Digitization: An Agenda for NSF," NSF Grand Challenges 2010, available at http://www.vanderbilt.edu/AEA/econwhitepapers/white_papers/Shane_Greenstein.pdf

77. The Public Library of Science has had a positive cash flow since 2010 despite being completely open access. See the 2010 PLoS Progress Report, available at http://www.plos.org/media/downloads/2011/2010_PLoS_Progress_Update_lo.pdf. BioMed Central, previously an independent OA publisher, was purchased by Springer, a traditional STM publisher, seems to validate the business model.

78. Thomas Lin, "Cracking Open the Scientific Process," New York Times, January 16, 2012.

79. PRISM, Myth vs. Fact, available at http://www.prismcoalition.org/myth.htm.

80. Letter of Francis S. Collins at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

81. Mark Ware and Mike Monkman, "Peer Review in the Scholarly Journals," Publishers Research Consortium, January 2008, available at http://www.publishingresearch.net/PeerReview.htm.

82. Ibid.

83. Comments of NIH Director Elias Zerhouni during the Congressional Hearing on the Fair Copyright in Research Works Act, available at http://judiciary.house.gov/hearings/printers/110th/44326.PDF.

84. Anne Gentil-Beccot, Salvador Mele and Travis C. Brooks, "Citing and Reading Behavior in High-Energy Physics." Scientometrics 82(4), 2010.

85. Bureau of Labor Statistics, "Career Guide to Industries 2010-11," available at http://www.bls.gov/oco/cg/cgs053.htm.

86. OECD, "Science, Technology and Industry Outlook 2010," available at http://www.oecd.org/dataoecd/40/27/46666029.pdf.

87. National Science Foundation, "Science and Engineering State Profiles," available at http://www.nsf.gov/statistics/states/.

88. Nathan Rosenberg, Inside the Black Box, Cambridge University Press, 1982.

89. Mark Ware, "Access by UK small and medium-sized enterprises to professional and academic information," Publishing Research Consortium, 2009, available at http://www.publishingresearch.net/documents/SMEAccessResearchReport.pdf.

90. John Houghton, Alma Swan & Sheridan Brown, "Access to Research and Technical Information in Denmark," Danish Agency for Science, Technology and Innovation, 2011. Available at http://www.fi.dk/publikationer/2011/adgang-til-forskningsresultater-og-teknisk-information-i-danmark.

91. Ibid.

92. Ibid.

93. Ibid.

94. PRISM, Myth vs. Fact, available at http://www.prismcoalition.org/myth.htm.

95. Ibid.

96. Hearing on the Fair Copyright in Research Works Act, available at http://judiciary.house.gov/hearings/printers/110th/44326.PDF; And House Hearing on Public Access to Federally-Funded Research, available at http://republicans.oversight.house.gov/index.php?option=com_content&view=article&id=922%3A07-29-2010-information-policy-qpublic-access-to-federally-funded-researchq&catid=14&Itemid. Ultimately,

the copyright issue can only be definitively resolved by the courts, but the publishers have not initiated litigation. As to the federal acquisition reports, we have not seen any evidence of widespread interest in substituting these reports for the researcher's manuscripts.

97. This field is generally called "Bibilometrics;" see http://www.nihlibrary.nih.gov/researchtools/pages/bibliometrics.aspx.

98. Philip M. Davis, "Open Access, Readership, Citations: A Randomized Controlled Trial of Scientific Journal Publishing," The FASEB Journal 25(7), 2011.

99. Among many papers in this literature, see: Yassine Gargouri et. al, "Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research," PLoS One 5(10), 2010; Yanjun Zhang, "The Effect of Open Access on Citation Impact," Libri 56, 2006; James A. Evans & Jacob Reimer, "Open Access and Global Participation in Science" Science 20, 2009.

100. Mark J. McCabe and Christopher Snyder, "The Economics of Open Access Journals," July 2010, available at http://mccabe.people.si.umich.edu/EOAJ_2010.pdf.

101. Benjamin F. Jones, "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?" Review of Economic Studies 76(1), 2009. See also Benjamin F. Jones, "Age and Great Invention," Review of Economic Statistics 99(1), 2010.

102. Ibid.

103. Ibid.

104. Anne Gentil-Beccot, Salvador Mele and Travis C. Brooks, "Citing and Reading Behavior in High-Energy Physics," Scientometrics 82(4), 2010.

105. James A. Evans & Jacob Reimer, "Open Access and Global Participation in Science," Science 20, 2009.

106. Ibid.

107. David J. Hardisty & David A. F. Haaga, "Diffusion of treatment research: Does open access matter?" Journal of Clinical Psychology 64(7), 2008.

108. Letter of Francis S. Collins at http://publicaccess.nih.gov/Collins_reply_to_Pitts121611.pdf.

109. Cited in Open Access Update, Caroline Sutton, Publisher, Co-Action Publishing, President, Open Access Scholarly Publishers Association (OASPA), ISMTE Meeting, August 9, 2011, Washington, D.C., and private communication on February 2, 2012.

110. By and large the academic research has focused on the link between open-access policies and dissemination of research by using future academic citations, but there is controversy about the use of citations as a key measure; for example, it is well-known in the bibliometric data that citation patterns vary across fields, that some citations are simply "honorary" rather than representing real transfers of knowledge, and that citation studies will neglect diffusion of knowledge outside of academia.

111. Jeffrey L. Furman & Scott Stern, "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," The American Economic Review 101(5), 2011.

112. Ibid.

113. Ibid.

114. Ibid.

115. Fiona Murray et. al, "Of Mice and Academics: The Role of Openness in Science," MIT Working Paper, 2008, available at http://www.hbs.edu/units/tom/seminars/2007/docs/Of%20Mice%20and%20Academics%20Stern.pdf.

116. Ibid.

117. Ibid.

118. Ibid.

119. Ibid. Two other papers by Murray and her coauthors explore the "Oncomouse" in greater detail. See Fiona Murray and Scott Stern, "Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge?" Journal of Economic Behavior and Organization 63(4), 2007; and Fiona Murray, "The Oncomouse that Roared: Hybrid Exchange Strategies as a Source of Productive Tension at The Boundary Of Overlapping Institutions," American Journal of Sociology 116(2), 2010.

120. Heidi Williams, "Intellectual Property Rights and Innovation: Evidence from the Human Genome," NBER Working Paper #16123.

121. Ibid.

122. Ibid.

123. Ibid.

124. Ibid.

125. Ibid.

126. Joshua Gans and Fiona Murray, "Funding Scientific Knowledge: Selection, Disclosure and the Public-Private Portfolio," in The Rate and Direction of Inventive Activity Revisited, University of Chicago Press, forthcoming 2012; Kenneth Arrow, "Economic Welfare and the Allocation of Resources for Invention," in The Rate and Direction of Inventive Activity, National Bureau Committee for Economic Research and the Committee on Economic Growth of the Social Science Research Council, Princeton University Press, 1962; and Richard Nelson, "The Simple Economics of Basic Scientific Research," Journal of Political Economy 67 (3), 1959.

127. "Harnessing Openness to Improve Research, Teaching and Learning in Higher Education," Committee for Economic Development, 2009.

128. Malcolm Gladwell, "The Tweaker," The New Yorker, November 14, 2011, available at http://www.newyorker.com/reporting/2011/11/14/111114fa_fact_gladwell

129. Elizabeth Eisenstein, The Printing Press as an Agent of Change, Cambridge University Press, 1980.

130. For instance, Joel Mokyr, The Lever of Riches, Oxford University Press, 1992; and Joel Mokyr, The Gifts of Athena, Princeton University Press, 2004.

131. Brendan Maher, "Personal Genomics: His Daughter's DNA," Nature 449, 2007, available at http://www.nature.com/news/2007/071017/full/449773a.html.

132. Eric von Hippel, The Sources of Innovation, Oxford University Press, 1988. Available online at http://web.mit.edu/evhippel/www/sources.htm.

133. For instance, Dr. Rob Carlson has written about this trend extensively on his website. See http://www.synthesis.cc/bio.html.

134. Yanjun Zhang, "The Effect of Open Access on Citation Impact," Libri 56, 2006.

135. David McKenzie and Berk Ozler, "The Impact of Economics Blogs," World Bank Working Paper, 2011, available at https://blogs.worldbank.org/impactevaluations/files/impactevaluations/dispatchsection2.pdf.

136. Mark Ware, "Access by UK small and medium-sized enterprises to professional and academic information," Publishing Research Consortium, 2009, available at http://www.publishingresearch.net/documents/SMEAccessResearchReport.pdf.

137. Karim Lakhani, "InnoCentive.com (A)." Harvard Business School Case 608-170.

138. Ibid.

139. Ibid.

140. Ibid.

141. Michael Woelfle, Pierro Olliaro, and Matthew H. Todd, "Open Source is a Research Accelerator," Nature Chemistry 3, 2011.

142. Ibid.

143. Ibid.

144. "With enough eyes, all bugs are shallow" is a well-known paraphrase based on a property noted by early open source advocate Eric Raymond in his 1999 O'Reilly book "The Cathedral and the Bazaar."

145. Michael Woelfle, Pierro Olliaro, and Matthew H. Todd, "Open Source is a Research Accelerator," Nature Chemistry 3, 2011.

146. Ibid.

147. Ibid.

148. John Houghton, Bruce Rasmussen & Peter Sheehan, "Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs," SPARC Report, 2010. Available at http://www.arl.org/sparc/bm~doc/vufrpaa.pdf.

149. Jeffrey L. Furman & Scott Stern, "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research," The American Economic Review 101(5), 2011.

150. John Houghton, Bruce Rasmussen & Peter Sheehan, "Economic and Social Returns on Investment in Open Archiving Publicly Funded Research Outputs," SPARC Report, 2010. Available at http://www.arl.org/sparc/bm~doc/vufrpaa.pdf

151. Ammon J. Salter and Ben R. Martin, "The economic benefits of publicly funded basic research: a critical review," Research Policy 30, *2001.*

152. Bronwyn Hall, Jacques Mairesse and Pierre Mohnen, "Measuring the Returns to R&D," NBER Working Paper No. 15622, 2009, available at http://www.nber.org/papers/w15622.

153. There are a number of useful summaries on the economics of returns to R&D. For example, "Measuring the Returns to R&D," cited above, looks at how economists statistically estimate the return to research spending accounting for spillovers and value created that is not captured by the inventing firm. Another line of research investigates the value of research based on the impact of new research results on stock market value of the inventing company; a good summary is Alessando Grandi, Bronwyn Hall and Rafaele Oriani, "R&D and Financial Investors," in Chiesa & Frattini (eds.), Evaluation and Performance Measurement of Research and Development: Techniques and Perspectives for Multi-Level Analysis, Edward Elgar, 2009. The economist Adam Jaffe has also written extensively on this topic.

154. Mark Ware, "Access by UK small and medium-sized enterprises to professional and academic information," Publishing Research Consortium, 2009, available at http://www.publishingresearch.net/documents/SMEAccessResearchReport.pdf.

155. See Philip M. Davis, "Open Access, Readership, Citations: A Randomized Controlled Trial of Scientific Journal Publishing," The FASEB Journal 25(7), 2011, or Mikael Laasko et. al, "The Development of Open Access Journal Publishing from 1993 to 2009," PLoS One 6(6), 2011.

156. Private communication from Neil M. Thakur to the author, February 2, 2012.

157. Not all researchers are supported financially by funders. Many private sector researchers, as well as a nontrivial number of academics, rely primarily on institutional support.

158. Rob Kling & Geoffrey McKim, "Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication," Journal of the American Society for Information Science, 51(14), 2000.

159. Oberlin Group of Libraries, Response to Office of Science and Technology RFI on Public Access at www. whitehouse.gov/sites/default/files/.../ostp/papf-email-3.pdf.

160. At the end of 2011 there were 310 open-access mandates tracked by ROARMAP. 51 of these were from funders but the majority were thesis or faculty repositories run by universities. A number of prominent research universities currently have faculty mandates. See http://roarmap.eprints.org.

161. OECD, "Principles and Guidelines for Access to Research Data from Public Funding," 2007, available at www.oecd.org/dataoecd/9/61/38500813.pdf.

162. Ibid.

163. Committee for Economic Development, "Harnessing Openness to Transform American Health Care," 2008, available at http://www.ced.org/images/library/reports/health_care/report_healthcare07dcc.pdf

164. Gareth Cook, "How Crowdsourcing is Changing Science," Boston Globe, November 11, 2011.

165. scienceforcitizens.net.

166. Gareth Cook, "How Crowdsourcing is Changing Science," Boston Globe, November 11, 2011.

167. Ibid.

168. For example, the open access Berkeley Electronic Press journals were sold in 2011 to a proprietary publisher, and previously open access articles were "re-gated" by the new owner. Authors of archived articles were not informed of this change. See Joshua Gans, "Berkeley Electronic Press Closes Up Journals," Digitopoly, January 26, 2012, available at http://www.digitopoly.org/2012/01/26/berkeley-electronic-press-closes-up-journals/.

About The Author

Elliot E. Maxwell advises public and private sector clients on strategic issues involving the intersection of business, technology, and public policy in the Internet and E-commerce domains. He is a Fellow of the Communications Program at Johns Hopkins University, and Distinguished Research Fellow at the eBusiness Research Center of the Pennsylvania State University. He also advises the Office of the National Coordinator for Health Information Technology in the U.S. Department of Health and Human Services, EPCglobal/GS1, and the Committee for Economic Development.

From 1998 until 2001, Maxwell served as Special Advisor for the Digital Economy to U.S. Secretary of Commerce William Daley and U. S. Secretary of Commerce Norm Mineta. In this position he was the principal advisor to the Secretary on the Internet and E-commerce. He coordinated the Commerce Department's efforts to establish a legal framework for electronic commerce, ensure privacy, protect intellectual property, increase Internet security, encourage broadband deployment, expand Internet participation, and analyze the impact of electronic commerce on all aspects of the economy. He was deeply involved in the development of e-government activities and was a founding member of the Federal Interagency Working Group on Electronic Commerce.

After leaving the federal government he was Senior Fellow for the Digital Economy and Director of the Internet Policy Project for the Aspen Institute's Communications and Society Program. Previously, Maxwell worked for a number of years as a consultant and as Assistant Vice President for Corporate Strategy at Pacific Telesis Group where he combined business, technology, and public policy planning. He served at the Federal Communications Commission as Special Assistant to the Chairman, Deputy Chief of the Office of Plans and Policy, and Deputy Chief of the Office of Science and Technology. Maxwell also worked for the U.S. Senate as Senior Counsel to the U.S. Senate Select Committee on Intelligence Activities.

Maxwell went to Brown, where he co-authored, with Ira Magaziner, the report which led to the introduction of the "New Curriculum," and Yale Law School, where he served on the Board of Editors of the Yale Law Journal. He has written and spoken widely on issues involving the Internet, electronic commerce, telecommunications, and technology policy. His most recent work, "Harnessing Openness to Improve Research, Teaching and Learning in Higher Education," was issued by the Committee for Economic Development (CED). The predecessors of that work, "Harnessing Openness to Transform American Health Care," "Open Standards, Open Source, and Open Innovation: Harnessing the Benefits of Openness," and "Promoting Innovation and Economic Growth: The Special Problem of Digital Intellectual Property" were also published by CED. The RFID Journal recently published his views on "Rethinking Privacy."

In 2005, Maxwell presented "Some Reflections on the Future: Dipping a Toe in the Datastream" to the Organization for Economic Co-operation and Development (OECD) in Paris. That same year, the Economic Policy Institute published his report, "A New Future for Telecommunications Policy: Learning from Past Mistakes." His study of Internet Governance "Rethinking Boundaries in Cyberspace," written with Erez Kalir, was published by the Aspen Institute in 2002.

More information is available at www.emaxwell.net

Statements On National Policy Issued By The Committee for Economic Development

Selected Recent Publications:

After Citizens United; Improving Accountability in Political Finance (2011)

Hidden Money: The Need for Transparency in Political Finance (2011)

Partial Justice: The Peril of Judicial Elections (2011)

Are Institutional Investors Part of the Problem or Part of the Solution? (2011)

To Reform Medicare, Reform Incentives and Organization (2011)

This Way Down – To A Debt Crisis (2011)

Restoring Trust in Corporate Governance: The Six Essential Tasks of Boards of Directors and Business Leaders (2010)

Harnessing Openness to Improve Research, Teaching and Learning in Higher Education (2009)

Teacher Compensation and Teacher Quality (2009)

Rebuilding Corporate Leadership: How Directors Can Link Long-Term Performance with Public Goals (2009)

Leadership and Shared Purpose for America's Future (2008)

Built to Last: Focusing Corporations on Long-Term Performance (2007)

The Employer-based Health-Insurance System (EBI) Is At Risk: What We Must Do About It (2007)

The Economic Promise of Investing in High-Quality Preschool: Using Early Education to Improve Economic Growth and the Fiscal Sustainability of States and the Nation (2006)

Open Standards, Open Source, and Open Innovation: Harnessing the Benefits of Openness (2006)

Private Enterprise, Public Trust: The State of Corporate America After Sarbanes-Oxley (2006)

The Economic Benefits of High-Quality Early Childhood Programs: What Makes the Difference? (2006)

Education for Global Leadership: The Importance of International Studies and foreign Language Education for U.S. Economic and National Security (2006)

A New Tax Framework: A Blueprint for Averting a Fiscal Crisis (2005)

Cracks in the Education Pipeline: A Business Leader's Guide to Higher Education Reform (2005)

The Emerging Budget Crisis: Urgent Fiscal Choices (2005)

Making Trade Work: Straight Talk on Jobs, Trade, and Adjustments (2005)

Building on Reform: A Business Proposal to Strengthen Election Finance (2005)

Developmental Education: The Value of High Quality Preschool Investments as Economic Tools (2004)

A New Framework for Assessing the Benefits of Early Education (2004)

Promoting Innovation and Economic Growth: The Special Problem of Digital Intellectual Property (2004)

Investing in Learning: School Funding Policies to Foster High Performance (2004)

Promoting U.S. Economic Growth and Security Through Expanding World Trade: A Call for Bold American Leadership (2003)

Reducing Global Poverty: Engaging the Global Enterprise (2003)

Reducing Global Poverty: The Role of Women in Development (2003)

How Economies Grow: The CED Perspective on Raising the Long-Term Standard of Living (2003)

Learning for the Future: Changing the Culture of Math and Science Education to Ensure a Competitive Workforce (2003)

Exploding Deficits, Declining Growth: The Federal Budget and the Aging of America (2003)

Justice for Hire: Improving Judicial Selection (2002)

A Shared Future: Reducing Global Poverty (2002)

A New Vision for Health Care: A Leadership Role for Business (2002)

Preschool for All: Investing In a Productive and Just Society (2002)

From Protest to Progress: Addressing Labor and Environmental Conditions Through Freer Trade (2001)

The Digital Economy: Promoting Competition, Innovation, and Opportunity (2001)

Reforming Immigration: Helping Meet America's Need for a Skilled Workforce (2001)

Measuring What Matters: Using Assessment and Accountability to Improve Student Learning (2001)

Improving Global Financial Stability (2000)

The Case for Permanent Normal Trade Relations with China (2000)

Welfare Reform and Beyond: Making Work Work (2000)

Breaking the Litigation Habit: Economic Incentives for Legal Reform (2000)

CED Counterpart Organizations

Close relations exist between the Committee for Economic Development and independent, nonpolitical research organizations in other countries. This International Network of Private Business Organizations includes groups of business executives and scholars with objectives similar to those of CED, which they pursue by similarly objective methods. CED cooperates with these organizations on research and study projects of common interest to the various countries concerned. This program has resulted in a number of joint policy statements involving such international matters as energy, assistance to developing countries, and the reduction of nontariff barriers to trade.

CE	Circulo de Empresarios Madrid, Spain
CEAL	Consejo Empresario da America Latina São Paulo, Brazil
CEDA	Committee for Economic Development of Australia Melbourne, Australia
CIRD	China Institute for Reform and Development Hainan, People's Republic of China
EVA	The Finnish Business and Policy Forum Helsinki, Finland
FAE	Forum de Administradores de Empresas Lisbon, Portugal
IDEP	Institut de l'Entreprise Paris, France
経済同友会	Keizai Doyukai Tokyo, Japan
NBI	National Business Initiative Johannesburg, South Africa
SMO	Stichting Maatschappij en Onderneming The Netherlands

Committee for Economic Development

2000 L Street N.W.
Suite 700
Washington, D.C. 20036
202-296-5860 Main Number
202-223-0776 Fax
1-800-676-7353

www.ced.org