



# Mapping the Future of Scholarly Publishing

The Open Science Initiative (OSI) is a working group convened by the National Science Communication Institute (nSCI) in October 2014 to discuss the issues regarding improving open access for the betterment of science and to recommend possible solutions. The following document summarizes the wide range of issues, perspectives and recommendations from this group's online conversation during November and December 2014 and January 2015.

The 112 participants who signed up to participate in this conversation were drawn mostly from the academic, research, and library communities. Most of these 112 were not active in this conversation, but a healthy diversity of key perspectives was still represented.

Individual participants may not agree with all of the viewpoints described herein, but participants agree that this document reflects the spirit and content of the conversation.

This main body of this document was written by Glenn Hampson and edited by Joyce Ogburn and Laura Ada Emmett. Additional editorial input was provided by many members of the OSI working group. Kathleen Shearer is the author of Annex 5, with editing by Dominique Bambini and Richard Poynder.

CC-BY 2015

National Science Communication Institute (nSCI)

[www.nationalscience.org](http://www.nationalscience.org)

[info@nationalscience.org](mailto:info@nationalscience.org)

nSCI is a US-based 501(c)(3) nonprofit organization

First edition, January 2015

Final version, April 2015

Recommended citation: Open Science Initiative Working Group, *Mapping the Future of Scholarly Publishing* (Seattle: National Science Communication Institute, April 2015).

# Mapping the Future of Scholarly Publishing

Executive Summary	i
Introduction	1
Recommendations	2
Goals, Objectives & Ideals	5
Current Confusion	7
A Crossroads	9
Why Act Now?	14
Questions that Need Answering	16
Possible Solutions	21
Conclusion	26
Annex 1: Journal Statistics	28
Annex 2: Current Issues in Journal Publishing	30
Annex 3: Current Issues in Open Access	33
Annex 4: The All-Scholarship Repository	35
Annex 5: Challenges for the Global Periphery	41
Annex 6: Recommended Reading	44
Annex 7: OSI Working Group Members	51

THE OPEN SCIENCE INITIATIVE WORKING GROUP

## ACRONYMS

APC	Article processing charge
arXiv	"Archive," the IR for physics/astronomy
ASR	All-Scholarship Repository
CC	Creative Commons licensing
CC-BY	A particular type of CC license
DOE	Department of Energy
IR	Institutional repository
NIH	National Institutes of Health
nSCI	National Science Communication Institute
NSF	National Science Foundation
OA	Open access
OSI	Open Science Initiative
OSTP	White House Office of Science & Technology Policy
SPARC	Scholarly Publishing and Academic Resources Coalition



The Open Science Initiative (OSI) is a global effort to improve the future of science, discovery and society by opening vast new troves of research information to researchers, educators, policymakers and the public. The main initial focus of OSI is to reform the scholarly publishing system. Two important side tracks are to find answers to key questions in scholarly publishing, and also investigate building the world's first all-scholarship repository.

Why journals? Scholarly journals are the backbone of science communication and discovery, and have been for centuries. However, for the past 20 years or so—roughly coinciding with the growth of the Internet—the scholarly publishing system has been under a tremendous and increasing amount of stress due to rapidly increasing subscription prices, rapid proliferation in the number of journals being published, distorted publishing incentives in academia, lax editorial oversight, massive escalation in the global rate of knowledge production, changing communication patterns and expectations in our society, the emergence of open access as a compelling model of free and open information access, and a wide array of other important factors. This stress is particularly affecting access to medical research information today, and particularly in the developing world.

The National Science Communication Institute (nSCI) hosted a conference in late 2013 to explore the broad outlines of this issue. The proceedings of this conference are available online at [bit.ly/1zkk6PJ](http://bit.ly/1zkk6PJ).

In early September of 2014, nSCI recruited and organized over 100 thought-leaders from around the world into a three month long online conversation—named the Open Science Initiative (OSI) working

group—to begin looking into viable ways to reform the scholarly publishing system. The transcripts of this conversation have been preserved and are summarized herein.

What are the problems with the current system of scholarly publishing? What are the different perspectives on these problems? What are some possible solutions? What should our goals and our guiding objectives be regarding improving access to research information? Should we even bother worrying about this issue (is the current state of affairs adequate)? What would a future with more open science look like? What might a future without more open science look like? How do we get from where we are now to where we need to be, considering there are so many competing interests and entrenched positions? Why might it be important to act now?

The OSI working group discussed these issues and many others at length. The group also made these three important recommendations (the first two being majority viewpoints):

1. **Convene an annual series of high-level conferences** between all key stakeholders over the next 10 years to discuss, implement, adjust, and track

major reforms to the scholarly publishing system. The first conference is currently being planned for early 2016. The delegate list will be an invited group of 200 decision-makers representing every major stakeholder group in scholarly publishing, participating with the understanding that they will try to reach an agreement on the future of scholarly publishing and will then work to help implement this agreement. The United Nations will be backing these conferences (through UNESCO) and will help mobilize broad and ongoing international support, participation, and funding. Very broad participation from US stakeholders—publishers, authors, federal agencies, companies who use research, institutions that produce research, and more—is critical to getting this effort up and running. While scientific research is certainly a global interest and enterprise, the US is the largest single producer and consumer of this research information, so without strong US participation, global adoption will be difficult to achieve.

2. **Find answers to key questions** related to reform, as detailed in the summary document. What do we really mean by “publishing” today? Are self-archiving mandates practical? Are impact factors accurate? Do embargoes serve the public interest? Are there better ways to conduct peer review? Why isn’t open access growing faster? These and many other questions have been identified in this report as starting points for discussion.
3. **Investigate the possibility of constructing the world’s first all-scholarship repository (ASR).** Our initial discussion regarding this repository is included in Annex 4. Conversations are currently ongoing on this matter. The Department of Energy’s Los Alamos National Laboratory (LANL) will explore building the prototype ASR (LANL also created arXiv). We are currently preparing a briefing paper for the White House Office of Science and Technology Policy so they can align upcoming federal compliance efforts with this repository. A number of OSI working group members

feel that creating the world’s first all-scholarship repository will need to be a precursor to truly comprehensive journal reform, and creating it the right way may end up having a greater impact on science discovery than anything ever attempted to date.

As we push forward with this initiative, the OSI group will need the following kinds of help: Broad buy-in and participation from research agencies, companies and institutions; more input and perspective from publishers, research institutions, government agencies, the public, and other stakeholders; subject matter expertise (such as programming, database construction, user interface design, customer experience, and so on), hardware/hosting support, data integration support, conference support (facility support, logistics, etc.); outreach/PR expertise; and finally, backing by policymakers and major funders. Building this support base will be the only way to achieve effective and long-term sustainable reform.

The budget for the first conference will range between \$150k and \$500k depending on how many of the costs we can cover for participants (more coverage is better—we don’t want people declining our invite on account of budget reasons). The repository effort can begin modestly but will eventually require millions of dollars annually, although much of the eventual operating cost can be recouped through sponsor support, advertising, and value-added services. A start-up budget of \$10 million would help get a critical mass of experts working full-time on this project right away.

This initiative already has a broad range of stakeholder support, but as we move forward we want to make sure that everyone has a seat at the table and also make it clear that we’re not just spinning our wheels to produce another white paper for discussion. OSI, nSCI, UNESCO, LANL, and others have committed to undertake an effort to actually shape the future of how we as a society value, share and use science. Care to join us?



For over 300 years now, scholarly journals have served as a vital cornerstone of learning, teaching and discovery. Even today, scientists value and rely upon journals and conferences far more than any other knowledge-sharing tools.

But with the explosive growth of information available through computers and the Internet in recent decades, our rapidly changing societal expectations about having free and rapid access to information, the continued emergence of many new research specializations, and the explosive growth of knowledge creation, the scholarly publishing system has reached perhaps the most significant crossroads in its history. Considering the speed of knowledge creation alone, we are losing our ability to keep track of information, let alone manage and make fair and proper use of it. But well-known problems with this system are also affecting timely access to critical research—particularly in global periphery and other lower resource regions (see Annex 5)—as well as library budgets, public education and awareness, public policy formulation, and more. Indeed, distorted publishing incentives and lax oversight may even be affecting science itself in terms of quality and replicability.

A group of experts from around the globe recently convened online to discuss scholarly publishing and develop a proposal for how to improve it. Coined the Open Science Initiative, or OSI, this group's proposal, along with a brief summary of the details and various perspectives discussed by participants is included herein. For the purposes of this conversation, the group defined "scholarly publishing" as the research output that originates from public and private universities, research centers, businesses, institutions and individuals, and ends

## The scholarly publishing system has reached perhaps the most significant crossroads in its history

up in journals. The group noted that journal publishing is a largely university-centric exercise: Most journal articles are written by academics, even though most science researchers work outside of academia (see item 1 in the first section of Annex 6). The group recognized that this view of scholarly publishing is limited; that it is far more applicable to STEM fields than most humanities; that there is a growing awareness among scholars of the need to create different types of scholarly products (such as open notebooks); that there is a wealth of scholarship that, to varying degrees by field, can be even more important than journal articles and can be similarly difficult to find and use (such as government research reports, books, patent literature and databases); and that as time goes by, it will be important to understand how the complete set of scholarly resources and not just journal articles are produced, stored, used, preserved and linked together.

However, the group chose to focus on journals for now because this is what science relies on the most, this is where the most urgent reform is needed, and because changing this system will create the greatest impact for science and society and will transition us from where we are today into a more efficient and effective yet still familiar and proven system. Whether this transition will work well for the humanities and social sciences is a perspective that needs to be included in this process. We leave the discussion of new modes of communicating science and scholarship for a future activity.

What can be done? Action is needed, and reasonable first steps are within our immediate grasp.



The OSI working group is recommending three immediate, parallel tracks of action: (1) establish a series of conferences between key stakeholders to reach a broad, inclusive, and sustainable agreement on the future direction of scholarly publishing, (2) investigate the issues related to and the prospects of establishing a new, global, all-scholarship repository (ASR), and (3) conduct, commission, or otherwise locate research to find answers to the key outstanding questions listed in this report.

With regard to inclusiveness, the conferences recommended by this group will include not just US and UK stakeholders but stakeholders from throughout the world. While it is true that the largest proportion

of scholarly articles are published in US- and UK-based journals (see Annex 6, Overview section, item 6), and about one-quarter of all research published annually comes from US institutions (same Annex section, item 7), the network of research and publishing is a vast and global one. Changes made for US and UK audiences will need to address the concerns and perspectives of global audiences in order to be truly workable. Annex 5 describes some of the unique challenges faced by scholarly communication reform efforts in the global periphery and elsewhere.

Getting this kind of broad input hasn't been attempted before in this matter, and yet it is clearly vital...

scholarly publishing. This first meeting in 2015 should be followed by a series of annual meetings until 2025 to continue focusing on the details of reform, and refinements to newly-adopted scholarly publishing processes and outcomes.

Given the range of issues and interests in scholarly publishing and the often-ingrained policies and practices of stakeholders, we don't anticipate that a single meeting will immediately resolve all issues for everyone. However, the frameworks of a solution are reachable, as this document describes.

In order for these reforms to be accountable and inclusive, a broad alliance of stakeholders should participate in and take ownership of these meetings, including faculty, publishers, libraries, grant agencies, policymakers, administrators, information schools, science communication groups, and scholarly associations and societies. And to the extent practicable, a wide array of groups advocating for scholarly publishing reform should collaborate on creating and managing these meetings. Getting this kind of broad input hasn't been attempted before in this matter, but it is clearly vital, particularly in order to create an effective and long-term sustainable solution to this issue.

In order for these changes to be ultimately workable, the OSI working group is recommending to:

1

**ORGANIZE A SERIES OF 10 ANNUAL DECISION-MAKING CONFERENCES, STARTING THIS YEAR**

About three-fourths of OSI members advocated convening a series of high-level meetings, beginning in 2015, to reach an agreement between stakeholders on what the path forward will look like in

1. Hold annual meetings on this issue for the next 10 years, beginning in 2015. Each annual meeting will focus on updating stakeholders regarding progress, necessary refinements, and challenges for the coming year. Reforming aca-

demographic publishing will take longer than one year, so accepting the need for a sustained multi-year effort is key.

2. Remain flexible for now on the best format for these conversations. One possible format might be a combination of congress and summit, where the congress consists of invited participants from across the stakeholder spectrum, and the summit consists of selected leaders from each stakeholder community who have decision-making authority for their institutions and organizations. The congress would help inform the summit, and the summit would give an equal voice and vote to stakeholder groups.
3. Remain flexible for now on costs. The conference organizing committee has started searching for sponsors, and at the time of this writing already has some interesting possibilities lined up. However, in order to ensure that this event stays on track regardless of major sponsors, every institution should tentatively plan on covering its own costs for participating.

The main focus of this first conference will be scholarly publishing reform since focusing just on one sub-issue, such as open access, will divert attention away from some of the larger context that also needs to be discussed. The objective of this first meeting will be to reach an agreement between stakeholders to: (1) improve scholarly publishing practices and outputs over both the short-term and the longer-term, (2) create a framework and timetable for reform, (3) commit to work together to implement these reforms in a timely manner and spread the new processes beyond the borders of those in attendance, and (4) monitor and fine-tune these efforts to continue improving academic publishing outcomes for all stakeholders worldwide.

Specific actions may also come out of this first meeting, but the first and most critical outcome will be to agree to work together toward our common goals. It is important to note in this regard that these proposed conferences

will involve university faculty members as one of the key stakeholder groups. Many open access publishing policies currently in place in the US have come from faculty and not from university administrations. Says Rick Anderson, Associate Dean for Scholarly Resources & Collections at the University of Utah's Marriott Library, "In an environment where academic freedom is paramount, changes in faculty behavior will mostly have to be initiated by the faculty."



## START SEARCHING FOR ANSWERS TO IMPORTANT QUESTIONS

A number of research questions have been posed in this document. These questions should be addressed before (or maybe as a part of) the initial conference—through original research, or locating existing research information:

1. What are the prevailing opinions among stakeholders in scholarly publishing about what "open access" really means? Is it realistic or even necessary that all stakeholders should have the same goals in mind?
2. What are the economic impacts of "traditional" journal publishing versus open access publishing? Are there enough data available to compare these modes to other modes of information dissemination such as books, blog posts, and media appearances? Deloitte and McKinsey each produced solid reports on the economic impacts of open data in the public sector (each showing substantial benefits; see the Economic Benefits section of Annex 6), but these reports did not include specific mention of open access publishing.
3. What is the actual rather than predicted impact of institutional repositories on journal subscription rates? Embargo policies to-date have been constructed on the assumption that access to the materials in these repositories needs to be restricted (embargoed) for a minimum amount of time in order to protect journal revenues.

4. Do researchers participate in the current system of scholarly publishing because they believe in it, or do they participate because it's required and they have no viable alternatives? What are their true preferences when it comes to publishing?
5. Are researchers getting what they need from the current publishing systems? Where are the satisfaction gaps in publishing across disciplines—with issues such as access, peer review, embargo periods, replicability, and so on? Some of the surveys noted in Annex 6 (the Current State of OA, item 16) hint at the answers. A meta-analysis of these and other similar surveys would help paint a clearer picture.
6. What factors are affecting the current growth rate of open access (supply, demand, available solutions, incentives, and so on)? Here again, a meta-analysis of current survey data would help.
7. What concerns do researchers have about Creative Commons licensing, which many see as being integral to the success of open access? Surveys have suggested low adoption rates, especially in the humanities. Is there existing survey data that can help answer this question?
8. What is the real extent of the information “underload” problem around the world (a term coined by Stewart Lyman to describe information paucity caused by access restrictions)?
9. How useful are impact factors? Conceptually, they drive faculty to publish in particular journals (those with higher impacts), which in turn skews the perceptions of these journals and therefore raises their impact value. We need a statistical analysis on the whole system of feedback loops that go into impact calculations—

journal visibility, press, perceived reputation, and so on, and not just citations and usage comparisons.

10. What is the current state of tenure evaluation across the US with regard to how journal publishing is weighed?



### INVESTIGATE CREATING A NEW, ALL-SCHOLARSHIP REPOSITORY

Institutional repositories of research information exist everywhere, but they are not widely used, and they do not communicate effectively with one another. Still, the most prominent US, EU and global efforts to integrate scholarship rely on connecting these repositories. This may be the right approach, but it may not be. The OSI working group has taken a look at this fundamental issue. Why aren't we building a single warehouse of global scholarship that researchers can search, integrate and mine? Is it unrealistic to consider this approach? The group is neither endorsing nor dismissing this approach, simply noting it as an idea worth more (a splinter group of OSI members has already started this task). A summary of the detailed conversation held by the group on the all-scholarship repository issue is contained in Annex 4 .

It is important to note here that scholars already have the right to deposit pre-publication versions of their research work in institutional repositories—more on this practice is discussed later. Most often, though—and for a wide variety of reasons including inertia, habit, complexity, a lack of time, a lack of interest, a lack of awareness, concerns about permanence, concerns that pre-print versions differ significantly from the actual versions of record, and more—these institutional repositories generally aren't used much (although participation rates varying widely by institution and discipline). So improving repository use (whether existing or new) might go a long way toward improving the current state of information accessibility.



## What are the main goals of this discussion? What are the objectives and ideals we should keep in mind?

Liam Earney is the head of library support services for the UK's influential Joint Information Systems Committee (JISC). As Dr. Earney said last year in a report examining the future of so-called open access publishing (see The Current State of OA section in Annex 6, item 1), "There has rarely been such high-level institutional and political interest globally in the publication and dissemination of scholarly research."

But where do we start? *Science* editor-in-chief Marcia McNutt suggests that in terms of identifying the ideals to which any reform effort should aspire, an effective scholarly publishing system is "widely accessible, filters the flood of information, and has indicators of quality."

Within this framework, University of Kansas economics professor Joshua Rosenbloom suggests these four goals for reform, with a fifth goal offered by Rick Anderson:

1. share our own research as widely as possible (the benefits of sharing scale with access),
2. ensure that we can find the relevant information produced by others,
3. provide for the persistence of this information in a stable form into the future (print has a good track record here, whereas other media are more problematic than we sometimes want to acknowledge),
4. identify and weed out spurious, incorrect or false claims, and
5. provide for review and certification.

**Our objective? As a starting point, it's to improve the discoverability, timeliness, usability, affordability, and intelligibility of scholarly and research information.**

What about ultimate objectives? A starting point is to improve the discoverability, timeliness, usability, affordability, and intelligibility of scholarly and research information. The overwhelming majority of scholars already agree with these points (see the author surveys in Annex 6 under The Current State of OA), and feel that improving access is vital to the future of scholarship and discovery.

Are there any other ideals to which we can aspire to along the way, setting aside the question of whether or not they are immediately achievable? For most people who care about the open access side of this question, the answer is yes, although not everyone agrees whether these are desirable or achievable.

PLoS co-founder and leading OA authority Michael Eisen describes the ideals this way:

1. Works of science should be made available freely as soon as the authors feel it is ready to share,
2. Assessment of works should happen from the moment a work is made available and continue for as long as it is useful,
3. The process of assessment should be fully open, and, while there is certainly room for organized entities in this process, everybody should be able to participate,
4. Assessment of a work is multidimensional, and involves issues of validity, importance, audience, ethics and many other factors, and whatever systems we have for evaluating works should be multidimensional,

5. The costs of supporting the science communication system should be borne by the entities that fund and otherwise sponsor science, and the economic system should be designed so that funds are never an obstacle to participation, and
6. The form of publication and the infrastructure should be optimized to facilitate all uses of published works, either on their own or in bulk.
6. Exposes the research for examination and discussion at early stages of development, explanation, and dissemination,
7. Expands peer review and assessment processes and possibilities,
8. Stimulates development of affordable tools and systems for analysis, interpretation, translation, remixing, sharing, archiving, etc.,

Joyce Ogburn, the dean of libraries at Appalachian State University, suggested that one approach to take toward these ideals would be to create an environment with legal, ethical, resource, service, policy, reward and incentive dimensions that:

1. Recognizes the value that scholarly research and inquiry offers to an audience beyond fellow scholars and researchers,
2. Places authoritative information in the hands of anyone who may benefit,
3. Allows for creative reuse and repurposing in the name of education and the public good,
4. Encourages experimentation and reduces risk in order to test different paths of inquiry and dissemination,
5. Fosters sustainability and scalability both for the traditional and the new,
9. Spawns new organizations and avenues for wide engagement with the creation, discovery, use, and preservation of knowledge,
10. Provides routes for more inclusive and participatory scholarship, much like citizen science welcomes amateurs and enthusiasts to collect data and engage in problem solving, and
11. Appreciates a multimodal system of scholarship where many different methods, forms, products, and channels emerge and flourish. Examples include extensive use of images and visualizations, interactive knowledge products, collaborative collections of structured information like the Encyclopedia of Life, multimedia, blended scholarship that incorporates both the physical and the digital, and infrastructure development that makes new kinds or methods of scholarship possible (tools, instruments, software, methodologies, algorithms, and more).

## So how do we support the goals of scholarly publishing as we move forward toward our objectives (with these ideals or other optimums in mind)?



And how do we evaluate that the changes we make to the current system are helping and not hurting? To answer these questions, and given the current state of scholarly publishing, some say we should start over at square one—that the system we have now is built on publishing realities that existed 300 years ago and that we would all be best-served by just rethinking this system in its entirety. Others argue that given enough time and encouragement, the incremental reforms happening now will eventually create a sustainable reality better than the one we have now but not necessarily much different.

Given the current state of scholarly publishing, some say we should start over at square one...

For the past decade or so, open access has been viewed by many as the golden chalice of publishing reform—as the most widely accepted and adopted approach that holds out the best promise for improving access to scholarly works. OA is best described by Peter Suber, the director of Harvard's Office of Scholarly Communications and global leader of the open access movement, as "a kind of access, not a kind of business model, license, or content."

OA comes in two main varieties—green and gold. Green OA refers to free, self-archived versions of scholarly works—the research articles that researchers and/or institutions can put on their own servers. Depending on a particular journal's policy, green archives might include either pre and/or post-peer reviewed versions of papers, or even the pdf versions of final published articles (including journal editing and formatting). Gold refers to OA peer-reviewed journals—the research articles that get published (as opposed to simply made available online).

The publishing costs of journals of all kinds—OA and non-OA—can be supported in part by author fees (called processing charges or APCs in OA, and page charges in non-OA journals). Exact estimates of how much support comes from pages charges and APCs varies widely, but the majority of OA journals do not charge author-side fees while the majority of non-OA journals do charge (there are, however, important economic dif-

ferences between journal publishing in STEM fields and in the humanities, as explained in Peter Suber's article listed in the Current State of OA, Annex 6, "Promoting Open Access in the Humanities").

Also, these fees are higher for non-OA journals and higher still for high impact journals (although some have suggested that the charges of some OAs are actually loss-leaders—below-cost rates). Both kinds of journals can also be supported by society or institution fees and by advertising.

Only non-OA journals and "hybrid-OA" journals charge subscription or access fees. In hybrid OA, the publisher maintains a subscription model but an author may pay an article charge for his/her individual paper to be "open" in an otherwise "closed" access journal. The ownership of copyright may also be different—held by the author instead of the publisher, for instance.

Both green and gold OA may be immediately available to the public (the latter upon publication). In most cases, the published versions of articles cannot be posted back to an author's or institution's green archive. Gold OA journal articles with Creative Commons licensing are an exception here.

Against this colorful backdrop of green, gold, open, closed, hybrid and so on, the number of journals in the global knowledge marketplace continues to grow at dizzying speed (see the Overview articles in Annex 6), but not all disciplines have been able to keep pace in a manner that ensures readability, affordability and access. Medical research in particular (in part due to increasing specialization and high demand) has found itself in a precarious position of having much of its knowledge rendered unavailable or unaffordable for discovery and collaboration. As described later in this document, the number of subscriptions—many of them very expensive—that science researchers need to keep current in their fields has grown explosively in recent years. Not only is this information voluminous, it often suffers from a number of usage impediments, including being disconnected from related research or its underlying data and analyses; inaccessible to non-subscribers for a period of time due to embargoes; very expensive to access at the article level; copyright-protected by entities other than the author, and often-times, very poorly-written, and unintelligible even to experts in the same field.

The downstream impacts of these impediments affects more than just science discovery and collaboration. Effective public policy on issues like climate change is also a casualty, as may be potential societal benefits such as technology transfer and science education.

Different perspectives and tensions are driving both the publishing reform and the open access discussions, but what we must remember is that research and educational institutions (and their researchers and scientists) own the work everyone is fighting over. These particular stakeholders need an important seat at the table—indeed, they need to be at the head of the table, and they aren't right now. In a recent report by K|N Consulting (Appendix 5, Novel Ideas, item 3), it is estimated that academic institutions pay about 84% of the global publishing and distribution costs of journals—by way of subscription fees, labor, editing, author charges, and other associated activities. The content and existence of journals as critical communication instruments be-

gin and end with these institutions—they are both the primary producers and the primary consumers of scholarly knowledge—so they need to take an active role in deciding whether and how the current system should change.

This scholar-centric reality is complicated by the fact that scholarly publishing takes place upon an uneven playing field. Some disciplines, particularly in the humanities, are experiencing less of an access crisis than science since their journals are generally much less expensive. Still, many subscription journals in the humanities—as well as monographs—are struggling for sustainability, and most of them are reluctant to make their content OA. Also, as pointed out by Colleen Lyon, Scholarly Communications Librarian at the University of Texas at Austin, journals in the humanities may be priced lower “but library monograph budgets are being cut in order to pay for the large increases in cost for STEM journals. Since the humanities depend heavily on books, this ends up impacting the scholarship they have access to—not to mention the reduction in the amount of material that university presses are publishing, meaning it's harder than ever to get a book published.”

Questions of access are also more pressing in some fields and institutions than others. Reforming journal practices for one field may end up impacting practices in another field by fiat, but not by necessity. We may not find a one-size-fits-all solution or model. Also complicating this discussion of the reform of publishing is the fact that there are tens of thousands of publishers in the world with a wide array of business models and constituencies. Some provide better services and produce better products than others. Some fields are being better served than others. Some publishers are tied to science societies that scientists want to support, some are tied to industry groups that value secrecy over transparency, and others are devoted to open access. Some publishing groups produce hundreds of journals, some handle just one.



With the proliferation of ideas, models, opinions and needs, scientific publishing—and by extension, open access—is at a crossroads today.

This crossroads can be described as a nexus of four different perceptions about what open access is and is not.

1. The moral imperative. Not everyone agrees that society has a moral imperative to share knowledge, or at least to share it freely, immediately, and without copyright restrictions. There are those who contend that research paid for by governments belongs to the people, and others who contend that the marketplace of ideas and innovation simply wouldn't function without secrecy and the right to protect ownership and discovery.
2. Public versus open. The essential difference between "public" access and "open" access is that in the public access model, authors or publishers retain copyright, which means that the liberal reuse of content can be limited (or at least not as rich and instantaneous as in open access). Some advocates for freer information are seemingly content to simply have more public access to information available regardless of price, license type, or timing.
3. Who are the stakeholders in this conversation? Where are the disagreements about OA occurring? A general observation from our working group (not backed up by survey data of broader opinions) is that there is little disagreement inside the core of the OA movement. Peter Suber notes that "the OA movement has been at this crossroads, or has contained these intramural disagreements, for at least 10 years," and William Gunn, the Head of Academic Outreach for Mendeley, agrees that "there's also a general consen-

sus view that is held by pretty much all the major thought leaders and that serves the majority."

However, it is the position of most in the OSI working group that OA perspectives and policies among the broad array of OA stakeholders are fragmenting now more than ever, and that a widespread understanding and agreement between the stakeholders in scholarly publishing is needed as soon as possible before this fragmentation produces undesirable outcomes for OA (or even regression due to confusion and disagreement), before more opportunities for discovery are missed, more articles get locked away behind paywalls, and new access models take shape that could ultimately deepen the information access rift between upper-tier economies and the rest of the world.

4. Where are we going? Different stakeholders in the publishing reform conversation have clearly different goals and perceptions, and these differences are sowing confusion, inaction, and even hostility toward more cooperation and collaboration between researchers and publishers to develop workable and mutually beneficial solutions. The differences of opinion about the goals of OA are numerous, and range from questions about whether Creative Commons licensing is required, to what the pace of reform should be, to whether the end goals of open access include the elimination of subscription journals—the scorched earth model—or the fertile garden model of creating a world of more information that will provide vast new opportunities for many.

With regard to the public versus private disagreements, David Wojick, formerly a senior consultant with the Department of Energy's Office of Scientific and Technical Information, describes that from a policy perspective, a "bifurcation...has certainly occurred at the national policy level, with the UK (and EU) going for open access and the US choosing public access. Moreover the OA movement seems to be largely silent on this pending policy schism. So...some urgency is called for, lest public access become the default solution."

Joyce Ogburn agrees with Suber that this crossroads isn't necessarily an impasse—that there are various avenues and options available regarding licenses, formats, and so on—but also agrees with Wojick that public access may be emerging as the settling point for the moment—that "this is as far as OA can be pushed right now under the current circumstances." Ogburn also describes the efforts to pass the US medical research public access law of 2008 that ended up creating the NIH PubMed repository. "I can attest to how hard it was to get public access," says Ogburn. "It took a lot of time, alliances, and compromises to achieve this step. Pushing for total OA was not feasible at that time." Adds Wojick, "There is a very real danger that US public access is not just a small step, rather it is the last step. Once established, US policy will be very hard to change and it is a potential model for other countries."

Not everyone agrees whether this bifurcation is real or imagined, permanent or temporary, acceptable or not acceptable. William Gunn and Jean-Claude Guédon, a long-time scholarly publishing expert and professor at the University of Montreal, suggest that this split, if it exists, exists only among those who are trying to implement open access policies and not within the OA advocacy community itself. Further, Guédon is confident that the emerging public access model is only a way station on the path to full open access, and that in the meantime, having some public access is at least better than having no access.

Rick Anderson agrees that it would be an exaggeration to say that there is a "split" or "bifurcation" in the OA community, but that there is certainly a diversity of beliefs in that community as to what constitute suitable goals for reform of scholarly communication. Some groups and individuals see public access as an acceptable end goal, while others see it only as a step in the process towards OA; some see some role in the future for nonprofit scholarly toll-access publishing, while others believe that anything less than universal OA would constitute failure. Some are willing to accept embargoes as a permanent feature of the OA landscape, while others are not; some hold strongly to the view that OA is not OA unless it includes CC BY licensing (or the functional equivalent thereof), while others strongly oppose mandatory CC BY while still considering themselves supporters of OA.

Wojick expresses reservations about whether the 2013 White House Office of Science and Technology Policy directive mandating public access is also a stopping point on the way to open access or simply an end in itself. "The OSTP public access program just extends the long existing NIH model to the rest of the funding agencies, so in a way it is a step sideways not forward. Once a program like this is established it is very hard to change, plus it can become a model for others. I have been surprised at the lack of objection to the OSTP public access program from the OA community." Wojick, who was part of the interagency work group that led to this OSTP directive, concludes that there is "no Federal sentiment that this public access policy is just a stopping point on the way to full OA."

Anderson notes that this public access stopping point has not been limited to government agencies. "Every library that provides an institutional repository that does not require CC-BY licensing is also offering what amounts to a public access solution, rather than a fully OA solution. I'm not aware of any library that has plans eventually to require BOAI-compliant licensing of the

papers placed in its repository.” Guédon suggests that library policies are being shaped more by the fact that they have archives filled with copyrighted materials than by a reluctance to embrace OA.

This CC-BY licensing issue is one of the larger issues in the “Where are we going?” split. Some feel that open access doesn’t necessarily need to be tied to Creative Commons licensing to function well. Others feel that Creative

## THE CREATIVE COMMONS QUESTION

The Creative Commons (CC) licensing model is not an attempt to abolish copyright. Rather, it is a model that in combination with free access can maximize the utility of information and opportunities for discovery. Some see the CC-BY license (the most accommodating Creative Commons license, which lets others modify and redistribute work, even commercially, as long as credit for the original creation is noted) as an essential component of open access. Others posit that open access can still function without Creative Commons licensing, but that it works best when research can be both freely accessed and used without restriction.

However, the basic philosophy of open access has a strong orientation toward sharing, so CC-BY licensing is seen by many advocates as a natural and even necessary component of OA. This orientation is based in part on the fact that, as Michael Eisen says, “the standard forms of reuse allowed under fair use are clearly inadequate, and prevent many forms of desirable reuse.” Indeed, as Jean-Claude Guédon points out, “fair use” doesn’t even exist in some countries and/or it is considerably more restrictive than in the US.

In general, Creative Commons licenses are far more liberal than copyright with regard to these fair use limitations. Under US fair use rules, copyrighted works of commentary, criticism, research, teaching, or news reporting can be cited without the copyright owner’s permission. This use is restricted, however, when it comes to creating derivative uses of the owner’s work, using large portions of this work, or developing works with commercial potential. Therefore, in science research in particular, there is a hope and expectation among many that more liberalized licensing approaches will lead to more reuse and discovery. Others argue that since CC-BY licensing rates are currently relatively low (varying by field), we may also be witnessing a reluctance by researchers to “let go” of scholarly work with commercialization potential. Therefore, they posit that attaching CC-BY to OA may actual lower participation rates in OA.

At this juncture, an important question among many in academia and research seems to be whether it is better to encourage more public access regardless of whether it is CC-BY licensed, or to hold out for more sharing under CC-BY licenses. Perhaps the sticking point in this conversation is a lack of understanding in academia about Creative Commons. These licenses do not replace copyright. Rather, they work alongside copyright, allowing users to authorize the levels of usage and protection that best suit their needs. Instead of accepting a default “all rights reserved” copyright license for journal papers, selecting a CC license allows users to hold some rights and let go of others that may increase the reuse of their research (such as allowing free sharing)—to choose a “some rights reserved” option if you will (of which there are six available, including CC-BY). Notes Paul Groth, Disruptive Technology Director at Elsevier Labs, “There needs to be better education about the licenses available to researchers for all the content they produce. Many researchers are not aware of what CC-BY or other licenses mean. For example, CC-BY can be hindrance for data reuse by data integrators (CC-0 being a better option for this use-case).”

Setting the CC-BY education and acceptance issue aside for the moment, the case for more CC-BY may be easy to establish: open materials are used more often in the here and now, and the wholesale removal of copyright from millions of papers in the future to enable greater sharing would be an essentially impossible task. Therefore, allowing greater access now (without necessarily giving away all rights) at least allows for the possibility of greater future use and reuse.

One other important advantage of Creative Commons licensing is its potential to solve the “version of record problem” in green repositories. In general, only preprint versions of copyrighted articles can be deposited into institutional archives. The final version resides in the journal’s website (and often behind a paywall). There may be only minor differences between these two versions, but the differences may also be significant. With Creative Commons licensing, the versions in all repositories can be the same, ensuring that the large, interoperable repositories we construct will be more robust and contain the most accurate information available for use and comparison.

Public access repositories that contain a mix of copyrighted and Creative Commons licensed materials may contain a mix of these “final” and “pre-print” versions of papers. In this sense, the public access model may not be nearly as valuable as the OA-CC model in creating reliable, authoritative repositories. OA is also not delayed—it is immediate—whereas PA articles are often embargoed for around 12 months to allow the publisher an opportunity to recoup costs.

Commons licensing is a core requirement of open access and that real OA can't happen without it.

The CC-BY question is only one point of confusion and disagreement. Another is that different institutions, fields and organizations who are advocating for more open access—such as SPARC, ACRL, the Alliance for Taxpayer Access, Create Change, the IFLA Open Access Taskforce, OASIS, the OASPA, the Open Data Foundation, Public Knowledge, PLoS, and the Right to Research Coalition, and others—have different end goals in mind, which therefore makes working together toward solutions problematic.

On the one side are those who firmly believe the goal of open access is to eliminate subscription journals and that subscriptions are intrinsically incompatible with universal access. On the other side are those who say the goal of open access is simply to make information more accessible to researchers, not eliminate subscription journals. And somewhere in the middle are those who say this isn't about open access at all but how journals and scholarly societies (who publish many journals) will adapt to change and whether we will end up seeing a net gain for science as a result.

"If there is/was a consensus" between these groups, says scholarly publishing expert and journalist Richard Poynder, "I suspect it is beginning to weaken as the practicalities of implementing open access come more sharply into focus." Furthermore, he says, "many of the views espoused by OA advocates may not be representative of the larger research community. As such, the larger research community might reasonably ask: 'Why should we accept that you know best and do what you say?'"

This disagreement has created more acrimony than necessary, with those who see open access as being inherently hostile to pay models of access either rallying to

the defense of subscription journals (or requiring more proof that open access works), or hastening to bid journals adieu. The reality is that at some point this issue became owned by the public, and in doing so, perceptions splintered and this splintering has meant that not everyone who wants more open access feels the same way about the goals of OA, or even the meaning. There are now many "owners" of this issue and they are speaking with different voices, so there is a lot of misunderstanding (or more accurately, different understandings), mistrust, confusion, and sub-optimal efforts.

Indeed, notes Anderson, "all of these entities are not, in fact, getting behind the same thing. The NIH and NSF have gotten behind public access, not OA. SPARC sees embargoes as something that should be allowed for now (as a compromise measure during what it be-

lieves is a period of transition to universal OA), but says that embargoes are not acceptable in the long run. By contrast, the latest revision of the RCUK mandate provides structurally for embargoes of various

The larger research community might reasonably ask: "Why should we accept that you know best and do what you say?"

lengths. Wellcome allows embargoes as well, and does not require CC-BY (though it encourages it), whereas RCUK does require CC-BY. And then, of course, there are thought leaders like Robert Darnton who explicitly disagree with other thought leaders who take the position that universal OA is the only acceptable future outcome for scholarly communication. None of this is to mention the wide diversity of thought and opinion that exists within the global community of scholars, whose work is the lifeblood of the scholarly communication system."

At least part of what the OSI working group is hoping to accomplish by publishing this paper and coordinating future conferences on this issue is to air these different perspectives and lay the groundwork for a better, common understanding so we can all come together and move forward toward more effective, more workable solutions.

Finally, with regard to the timing of reforms, Anderson suggests that three broad categories of perspective have emerged in OA circles:

1. **Universal OA Now:** All scholarship should be available on an OA basis and without embargoes, and we need to achieve this reality immediately or as soon as possible.
2. **Universal OA Eventually:** All scholarship should be available on an OA basis and without embargoes, but it's okay if we get there incrementally over time.
3. **Blended Solution Indefinitely:** We should always work to expand the public's access to scholarship, but it's okay if some embargoes, some traditional copyright restrictions, and/or some varieties of toll access remain a feature of the landscape indefinitely.

These categories seem to exist even where all are in agreement that OA means free public access plus the equivalent of CC-BY licensing. However, it's also true that despite the fact that the OA definitions offered earlier have been widely accepted, not everyone who thinks of

him- or herself as working for OA is working from the same understanding of what "open access" means. So in fact there are also multiple subcategories of perspective on this issue.

A fourth broad category of perspective—and not an insignificant one—is "No Significant Change Is Needed." For many, the current system seems to be serving their needs just fine, or they are skeptical of or disinterested in OA. Those who hold this view aren't included in the schema above, which is intended to address categories of orientation within the OA movement. The OA attitude surveys included in Annex 6 suggest that about half of all researchers fall into this "neutral" category, including economist Joshua Rosenbloom. "Philosophically," says Rosenbloom, "I think I can see the appeal of full OA, but practically I am not convinced that it is in fact feasible or that imposing it would indeed be desirable. Can we do better than the current system given the massive changes in technology that have taken place in the last two decades? Certainly. But I would prefer to look for a solution that explicitly articulates and seeks to promote all the goals of scholarly communication, and recognizes that there may in fact be trade-offs across them, rather than pursuing a single-minded focus on open access."

## What's wrong with the current pace of change? Do our different perspectives necessarily need to converge?



In particular, is it acceptable for the OA movement to contain multiple perspectives? Some say yes, others say no. Recognizing as we do today that there are multiple perspectives on this issue, “organizations working for OA have an obligation,” says Rick Anderson, “especially to their members, to be up-front about what their goals are.” Some also suggest that the current lack of clarity is sowing confusion and impeding progress toward more OA—and going back to David Woyack’s concerns, that this lack of clarity may make the public access option look more attractive by comparison since its end goals are clearer or at least comparatively easier to achieve. But could seeking convergence on this issue actually end up hurting OA? To most in this working group, the answer is no. More clarity, transparency, and unified action can only help every interest in this effort.

But to some, including Jean-Claude Guéron, there remains a concern that if trying to find common ground in this debate steers us away from an idealized version of OA—as a tactical, inter-

mediate move—it will result in a sub-optimal communications outcome for science (albeit temporary) “Perfection must not become enemy of the good,” says Guéron, “but the good is not good enough.”

Are there any other compelling reasons why we should drive for a resolution to the OA situation sooner rather than later? Of course. This initiative isn’t just about clarifying OA or preventing public access from taking hold as the default model. It’s about organizing our world’s ever-increasing flood of information, storing and preserving it in ways the global public can find and use, building

new and better analytical and discovery tools, and using this information to enrich knowledge, find answers, save lives, and even multiply the economic impacts that research and development is already so good at doing. Now is the time to reach an agreement and begin moving forward swiftly—not ten years from now after even more of this critical information has become siloed, pay-wall-protected, and otherwise hidden from researchers and the public for another generation.

And there’s the efficiency and effectiveness argument to consider as well, says William Gunn. “In a time of funding crunches, we have a duty to...make the case to funders that we’re using their funds as effectively as possible.” When knowledge is created and then not made discoverable and accessible, research and discovery obviously suffer, along with our funding efficiency and effective-

ness. We risk traveling down the same roads again and again without clear knowledge and understanding of what has been tried and learned in previous research.

When knowledge is created and then not made discoverable and accessible, research and discovery obviously suffer...

Weighing these many different perspectives and concerns, the OSI working group feels that the best option is to get in front of the issue of publishing reform and to help define the direction and solutions that will otherwise evolve without the considered input of all stakeholders. To characterize this approach as crafting a “grand compromise” is inaccurate. Rather, the approach needed here is to move forward together on the basis of advancing the interests that all stakeholders have in common—which at very minimum includes the effective production, communication and application of knowledge.

What if differences remain at the end of the day? While a frank and forward-looking discussion of our interests and options will go a long way toward bridging the gaps in understanding, Michael Eisen cautions that there are those in this conversation whose positions are nearly immovable, but that these positions shouldn't keep us from moving forward as a group. "As a professor," says Eisen, "I am routinely asked to do things that are not necessarily in my direct personal interest in the name of some greater good. For example, I can not keep all of

my data hidden away for me to reuse in publication after publication. I can not refuse to share research reagents with my competitors. I can not commercialize ideas that come out of my lab without sharing them with the university and/or public. And I think this is another one of those cases," he says, "where we have to accept" that as we create a new publishing paradigm with new demands on scientists and institutions, some people will not get on board. But the group's understanding and effort should still move forward.

## WHY?

### 1. Why should we do this?

To lay the groundwork for a future where we can maximize contributions to, and use of (or at least the opportunity for use) scholarly resources

### 2. Why now—what's the hurry?

The rate of information creation in our world is accelerating at breakneck speed. The more of this information that gets locked behind paywalls and copyright restrictions, the more opportunities for discovery are lost and the harder it will become to unlock these materials for future generations. This situation is especially critical in STEM, where inquiry is increasing becoming segregated into silos and separate streams of inquiry, and where one year embargoes can quickly dry up leads. Public access alone is not sufficient in science to allow for maximum information sharing and use, although it may be adequate in the humanities and some social sciences. Also, there is also concern in the global periphery that the current direction of reforms will do more harm than good.

### 3. What's so bad about taking more of a wait-and-see approach to publishing reform?

The publishing industry is changing, but it's important to get ahead of reforms and ensure that the changes include input from all key stakeholders and maximize benefit to the public and ensure equality of access around the world.

### 4. Where is the proof this approach will work?

Open access adoption continues to grow. However, there is confusion among the consumers of OA materials (primarily academia). Many don't understand what OA entails and how it benefits them. There is also genuine, fully-informed disagreement in academia about these questions, and these include disagreements in principle rather than just confusion or miscommunication. Our feeling is that by injecting more clarity and unity into this conversation, academia can move forward together toward a mutually-acceptable solution, adoption rates will markedly improve, and systems for use and reuse will improve along with this increased demand.

### 5. How does the scholarly publishing audience feel about these reforms?

Most of the producers and consumers of OA (mostly academia) have positive feelings about OA. There is definitely more enthusiasm for OA among the STM community, however, than in the humanities and social science communities. A one-size-fits-all approach to OA may be neither possible nor desirable.

This conversation has many high-level questions hanging over it.



There are also many specific publishing-related questions to answer (as noted in Annex 2), but these overarching questions need to be examined first (note that some of the research questions mentioned in the first part of the document were drawn from this section):

### PUBLISHING

1. What goals should scholarly publishing have? Do we as a society have a moral imperative to ensure the timely, accurate, and—to the extent practicable—free dissemination of research information, or do the structural and economic constraints to the publishing process need to define what we can and should do with this information?
2. Do researchers and scientists participate in the current system of scholarly publishing because they believe in it, or do they participate because it's the only game in town? What are their true preferences when it comes to publishing? Are they getting what they need from the current system? Where are the satisfaction gaps in publishing across disciplines—with issues such as access, peer review, embargo periods, replicability, and so on? Some of the surveys noted in the Annex hint at the answers. A meta-analysis of these and other similar surveys would help paint a clearer picture.
3. NIH data scientist Bruce Hutchins and US Cellular statistician Gordon Reikard point out that in some fields the total financial outlay required to publish can be a driving consideration when considering publishing options—including the decision to not publish. Both subscription and open access publications can charge to publish (with

charges levied per submission, article, page, and/or figure). The financial barriers to publishing can be significant, and the financial distinction between open access and subscription publishing can be large, modest, or non-existent. This is not considering, of course, the cost of access after publication—just the cost of getting published. How much are these investment considerations skewing publishing decisions (in different fields and research environments)?

4. What do we mean by “publishing?” Do we mean archiving? Or the value-added services provided by publishers (such as editing and managing peer review)? When we talk about publishing reform, are we talking about reforming the polished end products, the process, the profits, the mechanisms, the archives, or something else?
5. Where should publishing reform efforts be focused? Everywhere, or just science?
6. Who has the power to make changes to scholarly publishing practices? Do these powers flow from publishers, institutions, tenure committees, funding agencies, authors, or all of the above? Or none of the above? Is federal intervention needed?

### PUBLISHERS

7. Jean-Claude Guédon posits the question of how the commercial concerns of publishers may or may not be affecting the contents of journals and the orientation of research programs. Is this an issue?
8. What is the most appropriate role for publishers? As Rick Anderson notes, “This is a tough

issue, because while we might all agree that knowledge shouldn't be treated as a commodity, we do need a mechanism for people to be paid for their work. Much of the work that gets done in producing scholarship is either institutionally funded (by the academy) or grant-funded, but much of it is done by scholarly publishers who function outside of those spheres. A few of these are for-profit commercial entities; the great majority of them are nonprofit professional or scholarly societies. Is there a way for publishers to be paid, reasonably, for the work that they do that won't constitute the inappropriate commodification of knowledge? Of course, not everyone will agree that the work publishers do is necessary, and not everyone who feels it's necessary will agree that publishers (as traditionally understood) should be the ones to do it. If we remove publishers from the picture, then we need to figure out either how the work they now do will be done, or how we will move forward without that work being done." There is also the issue of how publishers like the AAAS—scholarly societies that publish major journals like *Science*—will fare with publishing reforms. Whether we end up seeing a net gain for science in this case, says Anderson, will depend on quite a few additional factors, "including the degree to which science benefits from the activities and services that are currently underwritten by subscription fees. Bear in mind that most scholarly journal publishers don't put their surpluses (if any) into the pockets of shareholders; they plow them back into the work of their societies."

## FINANCING

9. Guédon also argues that publishing is an integral and important part of the research process, and that "most research in most countries is supported by either public money or foundations (charities in the UK). In other words, scientific research has been subsidized from the very beginning (say the creation of the Royal Society in Britain in the 17th century). Why the publishing phase of research came to be treated separately and differently in terms of money is an interesting historical problem." Should we be focusing on merging these two phases back together again? What would this merger cost (or are we just talking about shifting buckets of money around)?
10. Is our current funding mechanism for scholarly journals fundamentally unjust? Many have complained that research funded by the federal government should not rightfully be repackaged by publishers and resold back to organizations from which the research originated, once again at the taxpayer's expense. This is not to diminish the important role that many publishers play in the process of bringing science out of institutions and to the public (at least nominally; the primary customer of journals is still academia). But are there other ways to fund this enterprise that wouldn't involve double-dipping? For instance, the K|N report (as noted earlier) describes a plan for supporting the publishing process that involves creating collaborations between scholarly societies, universities, libraries, and other stakeholders. Is this kind of approach too sane, rational or optimistic for the real marketplace?
11. In a related question and stepping back even further from the issue of who pays, Guédon asks whether our current perspective of commodifying of science research is itself the main financial issue. "There has been an insertion of a market approach to a category of entities—namely research results—that are not commodities, but rather elements of an informational infrastructure. If we define the task of designing a research problem as following a road made up of various papers distributed across many journals, we can

see the problem easily: imagine travelling from A to B with the requirement that you must buy your way at every step of the trip in order to reach your goal." Is the correct question to ask, then, whether our perspective itself is wrong? Instead of trying to justify new approaches to financing journal publishing, should we instead be questioning why research results should be a private commodity to begin with (at least in the case of government-funded research)?

## OPEN ACCESS

12. What has "open access" come to mean as the practices and broad perceptions (and misperceptions) about OA have developed over the years? Even if publishers and research institutions decided to work together to increase open access, what would they be creating—or more accurately, what do they think they would be creating? A system that is completely free, immediate, and Creative Commons licensed (with the most liberal sharing form of license)? Or something else? And what about the open sharing of data, source codes, protocols, and other research materials? Does talk about open access encompass these myriad concerns for some, or is a separate discussion needed?
13. What amount of annual growth and progress in OA is satisfactory to institutions and publishers who need to gauge how quickly they should move and how much of a commitment (including time and money) they should invest? Is the currently slow growth of OA due to a lack of supply, sub-par solutions, a lack of demand, a lack of clarity about what OA means, or all of the above? Are more personal motives involved, like academics wanting to support (through toll-access) the societies to which they belong, or a genuine fear that green OA (depositing pre-published versions of papers in institutional repositories) means giving away publishing rights

or diminishing the value of scholarly work to a publisher? Or are fears about unauthorized commercial reuse, plagiarism or fraudulent use the main roadblocks to more rapid and widespread application, adoption and acceptance of OA? Here again, a meta-analysis of current survey data might help.

14. As OA publishing continues to grow, how do we ensure an open access publishing system without financial obstacles for readers or for authors? Financial obstacles to publishing can be significant and can keep scholars from underfunded institutions from publishing their research (particularly scholars from the global periphery and the humanities).
15. As discussed, many OA advocates see Creative Commons licensing (specifically CC-BY) as an important part of achieving full OA, but acceptance and adoption of CC-BY has not been widespread to-date, particularly in the humanities. How do authors actually feel about Creative Commons licensing? Some author surveys (see Annex 6) suggest that attitudes are mixed. What are the underlying reasons? Confusion, a lack of concern, institutional influence, concerns that aren't being addressed, perhaps even ignorance? And once we identify these reasons, what solutions should we develop to help improve CC-BY adoption rates, if this is in fact the right tool for the job? Will more education alone solve the problem? Or if CC-BY isn't the right tool, what is? Would creating a new Creative Commons licensing category help or hurt?

## INFORMATION ACCESS

16. What is the true extent of the information "underload" (paucity) problem around the world? Some have suggested that the price of knowledge is too high, and that these high prices are creating knowledge scarcity, particularly in medical research and practice where access to

timely information is critical and the price of this information is the highest. Access prices may be too expensive for small research groups in the US, unaffiliated researchers, and researchers in the global periphery.

17. How do we avoid replicating the usual patterns of exclusion in making sure that scholars, publishers, administrators, policymakers, and other stakeholders from the global periphery are part of any serious future attempts to reshape a knowledge system that impacts all of us?
18. Are we satisfied with the current state of knowledge preservation? There appear to be two main issues here: methods, and actors. With regard to methods, Joshua Rosenbloom says that “In addition to ensuring wide current access to scientific information, the scientific communication system must ensure that this information is preserved in a fashion that is accessible to future generations of scholars. Print journals are widely held in a distributed network of libraries making it unlikely that they would be lost. Transition to online publication poses challenges of designing a robust and sustainable storage architecture that is not dependent on continuing infusions of funding to preserve access to historical information.” Organizations like Portico and LOCKSS are doing important work in this area. With regard to actors, some have noted that we should avoid leaving preservation in the hands of private companies that could cease to exist.

## PEER REVIEW

19. What is really needed from peer review, and is the established review process meeting these needs? The arguments against the current process are numerous: reviewers often know less about the subject matter than authors, value is not always added, the review process is fundamentally unfair in a commercial sense that reviewers are not paid for their work (and the finished product is sold, with no royalties accruing back to authors or reviewers), anonymity is not always preserved, peer review ends up marginalizing creative work (which, in science in particular, is not what we should be doing as opposed to merely ensuring that work is legitimate), and more. Is there a better way, such as open peer review, or pre-publication peer review?

## EMBARGOES

20. What public or scholarly interests are being served by placing embargoes on federally-funded research? What is the impact of this delay on discovery and the public interest? What other more nuanced options might we consider to address the wide variety of concerns on this issue? See Annex 2 (Issues in journal publishing) for a more detailed description of embargoes.
21. Is there actually evidence that journals have experienced subscription cancellations in response to manuscripts being made available in open access repositories? According to Shan Sutton’s research (see Annex 6, Embargoes section), the answer is no.

## MANDATES

22. How can self-archiving be improved in the US? Are mandates the answer? A variety of concerns about mandates are mentioned in this summary, but some are more salient than others. And what is the likelihood of crafting and enforcing a federal mandate on self-archiving—one uniform, irrevocable requirement that all research be archived in a free and open location?
23. Do we need to tighten and clarify our use of the term “mandate,” such that it will be applied only to those policies that actually require authors to do certain things (see Peter Suber’s book on this subject, listed in the Mandates section of Annex 6)? As Rick Anderson states, there is a meaningful difference “between policies that require

certain behaviors and policies that amount to statement of institutional preference. The landscape is filled with the latter, some of which are kind of disguised as the former...that say, in essence, 'You must deposit your paper in our repository unless you say that you don't want to, in which case you don't have to.'"

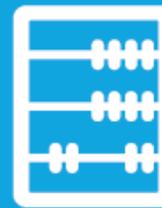
## IMPACT FACTORS

24. How useful are impact factors? Conceptually, they drive faculty to publish in particular journals (those with higher impacts), which in turn may skew the perceptions of these journals and therefore raises their impact value. Has anyone conducted a truly rigorous statistical analysis on the whole "ecosystem" of what goes into impact calculations—journal visibility, press, perceived

reputation, and so on, and not just citations? More on this is discussed in the next section. Are there meaningful alternatives to evaluating impact, or do we even need to bother (after all, we can't possibly hope to quantify the future impact of science). Would faculty be satisfied with an alternative system as long as it is recognized as reflecting meaningfully on the quality of their scholarship?

25. Altmetrics are often cited by open access advocates as an alternative method to measuring scholarly impact. How much do the aims and goals of researchers in the field of altmetrics align with the aims and goals of those in the open access community (see <http://altmetrics.org/manifesto/>)?

When it comes to developing workable solutions, there can often seem to be as many entrenched positions as there are stakeholders.



However, the needs and interests of these different groups do overlap, and it's clear that workable solutions exist at these areas of overlap if the various stakeholders begin talking on a regular basis about what to fix and how. The OSI working group is proposing holding a series of annual meetings between high-level stakeholders to begin working toward our common goals, and adjusting our plans along the way as technology and the publishing landscape continue to evolve.

Pre-judging solutions will not be helpful for these meetings. However, identifying some possible solutions might be helpful if only to note that a variety of possible solutions exist, and some of these may be quickly achievable.

### THE ALL-SCHOLARSHIP REPOSITORY

For instance, what if we looked at the current process of scholarly publishing as really consisting of two distinct parts—sharing and publishing? As Paul Groth notes, “The decoupling of the journal I think will help us have the open content we want while highlighting the value of peer review and quality.”

Sharing research is simple in some disciplines. Most physics research gets immediately posted to the free, open access server arXiv.

Astronomers, computer scientists, mathematicians, and increasingly evolutionary biologists make heavy use of this system, depositing 7,000 papers every month. Impact factor is irrelevant here (although not later when it comes to getting published), as is peer review since

arXiv is just an open access repository and not a journal; a moderated system helps flag papers that might be of dubious quality. Licensing formats vary (all rights can be retained by the author or permissions openly granted through a Creative Commons license—see <http://arXiv.org/help/license>). The common thread is that arXiv requires sufficient rights to allow submitted articles to be non-exclusively distributed by them in perpetuity. At present, the total budget for running this system is approximately \$850,000 per year.

With a new global system like arXiv for all disciplines (or by expanding arXiv to include all disciplines), all research would be freely and immediately accessible. That goal might be immediately within reach and immediately accepted by stakeholders on all sides of the OA issue. The advantage of creating a single repository is that it would solve some of the barriers that currently exist with institutional self-archiving, and it would also allow scalable solutions for metadata, search, visibility, and so

on to be implemented across all research papers at once. The details can be worked out later, but the technology certainly exists to create this kind of system from scratch or leverage one of the existing systems for this purpose, and the operations budget is surmountable. Modest fees (perhaps based on student

population) from every research institution should be more than enough to support the operation of such a facility (e.g., 200 institutions contributing \$5,000/yr provides a \$1 million annual budget), with any excess going toward value-added services. Notes William Gunn, “This

Identifying some possible solutions might be helpful if only to note that a variety of possible solutions exist, and some of these may be quickly achievable.

is, in essence, what the OSTP intended to accomplish. They wanted to expand PubMed Central to PubFed Central. This broke apart due to interagency politics and now each agency is to come up with their own plan. They're overdue to submit their plans and so far only NIH and DOE are close, with DOE leaning towards CHORUS and various other agencies working on their own solutions. A university-centric solution is being developed by the Association of Research Libraries (ARL), AAU, and APLU. This solution is called SHARE and it is working with the Center for Open Science as development partner."

How would the world be convinced to use a new all-scholarship repository? And would it necessarily need to be entirely OA to be effective? Would authors still maintain copyright restrictions, for instance? An agreement on how to improve open access publishing might include these elements as identified by Michael Eisen: convince, compel, and create. Moving toward such a plan would mean convincing academia about the merits of OA, compelling them (through real and enforceable mandates, not voluntary OA policies) to publish using this model, and creating an OA publishing system that is attractive.

How would academia react to this effort? Would a large new mandate fly, or are there other mechanisms that would be just as effective (or even more so), like tying funding scores to the percentage of green papers authors have deposited (as suggested by biotechnology consultant and author Tom Hopp), or creating new processes to handle deposits on behalf of busy faculty (as suggested by Joyce Ogburn)?

A more detailed discussion on the all-scholarship repository idea is contained in Annex 4 of this document.

## **EXPANDING THE PIE**

What role would publishers have in this kind of a world where one vibrant pre-print repository existed instead of hundreds of small, disjointed ones? A big one. Someone still needs to comb through this mountain of in-

formation. Even the more promising works would still undoubtedly benefit from editing and peer review. An existing audience of institutions, libraries, researchers and policymakers would still need science societies and publishers to sift through and identify the information and research of importance, and to package this information (including through translation and simplification) into articles that are clear and readable. And these steps are not even addressing what lies ahead—tasks like making connections between research and research data, and creating other new information products like these for the benefit of scientists, educators, policymakers and the general public. In other words, the benefits pie can greatly expand for everyone.

What would such an expanded pie system look like? Here's where the real possibilities and options come into play. One scenario to propose for consideration is this: What if the journal world started resembling the book world with regard to competing for authors? Journal publishers may not have to compete as fiercely as book publishers for the business of authors, but competition still exists—authors seek out and often have choices about where they should publish. What if more market forces were brought to bear in this relationship? In the book world, for instance, publishers compete aggressively for the business of the best authors and for control of the best manuscripts, using approaches that include royalties, advances, limited-duration contracts, derivative product development and more. Could this marketplace model be applied to journal publishing as well? The best pieces (of the most interest to their respective audiences) could be pulled out of this all-scholarship repository and polished up for publication. Copyright would remain with the authors, not with the publishers, and royalties would be paid accordingly (to authors or their institutions). The most glamorous journal "imprints" would, of course, remain the most coveted publishers, but having a paper selected for publication by a coveted publisher would be less a proxy of pure merit and more accurately a reflection of some merit-marketability-interest metric (although some would argue that this is already the case).

## PEER REVIEW

What would happen to peer review in such a scenario? Publishers would continue to offer peer review as they do today, but some form of peer review could also happen in the archive. A crowd-sourced peer review system has been proposed (by Marcia McNutt and others), wherein the papers deposited in an arXiv-type of repository would then be subjected to crowd filtering, review and comment, which would push the better papers to the top where these could be subjected to more rigorous review and editing before becoming formally published works. A journal of astrophysics, overlaid on arXiv, has in fact already been constructed as an example of how a system connecting repositories directly to peer review and publishing might work (see <http://theo.j.org> for details). Biophysicist Daniel Mietchen notes that the Electronic Proceedings in Theoretical Computer Science (see <http://www.eptcs.org/> for details) has already been doing this for years.

## MARKETPLACE REFORMS

What other scenarios or factors might lead to the marketplace reform of journal publishing? Sales is certainly one area to consider. Journals aren't really a big open-market commodity right now—most sales are made directly between journal publishers and academic libraries. Would more market competition help bring down prices? Since December of 2013, nSCI has been working with Amazon to investigate the groundwork of selling journals through the Amazon website and bringing this retailer's considerable market power to bear in journal pricing, availability and distribution. Amazon is still investigating this matter, and over the past year has interviewed many of the key players around the world in open access and journal publishing. A decision by the company is expected soon.

Scenarios that make publishers more of a client than a manager are also being tried. In the field of particle physics, for instance, where papers are already published in arXiv, a three-year deal (through SCOAP3, a broad alliance of libraries, researchers, funders and publishers)

was recently reached wherein select journal publishers will provide peer-review, editing, and publishing for most of the high-energy physics papers deposited in arXiv. Winning publishers were selected on a variety of factors, including article processing charges—a technique that froze out the field's previously most prestigious journal. This arrangement was funded by up-front payments from libraries.

A straight-to-market approach for journal articles has also been proposed, using off-the-shelf tools like Smashwords, or a yet-to-be-invented iTunes-like system that allows for article-level browsing and downloads at \$0.99/article. Existing resources like DeepDyve have made interesting inroads to this approach, allowing users to read millions of articles online for a \$40/month subscription fee (although downloading articles is still a separate and often hefty charge). According to William Gunn, "Pubget, acquired by the Copyright Clearance Center, was pursuing a pure pay-per-view model and found that the market really best supports institutional sales at the moment."

Finally, any business with high profit margins (as is currently the case in science journal publishing as a whole) is ripe for entry, even consolidation. So it's entirely possible that a force like Amazon or Google could enter this market in a disruptive manner, either seeking to buy hundreds of smaller publishers and roll them into one, buy one or more large publishers, or start from scratch with their own label. Entities like this don't need high margins and can further reduce costs by leveraging their existing technology frameworks.

## CLINICAL TRIALS

One could also posit that a one-size fits-all system for all scholarly publishing isn't necessarily ideal or the only way to think about the future. Clinical research, for instance, is an area where reform is critically needed, and the arXiv solution would not necessarily be a good off-the-shelf fit. What is needed in clinical trials is to disclose nearly everything, including protocols, modifications, recruitment figures and datasets (see the Open Sci-

ence Framework for an example of how this might look, <https://osf.io>; see also the Immune Tolerance Network's TrialShare site, built using LabKey, at [www.itntrialshare.org](http://www.itntrialshare.org)).

In the case of clinical trials publishing, we may want to think in terms of creating a brand new archive for academic knowledge (or at least a whole new format)—one that is flexible enough to hold everything associated with research and not just final papers. Creating such a system would likely be a tremendous benefit to science, medicine and discovery. Considering medical research alone, clinical trial results are currently scattered and often difficult to access; providing a summary in journal article format isn't enough for scientists to really cross-compare comprehensibly and with confidence. This approach also results in publication bias, where only favorable results are being published. Delays are a factor, where time is of the essence, so a catch-all repository would dramatically speed time to access. Upstream from this, publishing could still proceed as with other disciplines, with the exception that traditional peer-review systems may not be needed here, given that clinical trials undergo rigorous review by institution-based, industry/sponsor-based, and/or network/consortium-based scientific review committees (most often multiple groups with different areas of focus). This review process examines issues such as scientific worthiness, methodology, radiation safety, data safety monitoring, and investigator credentials. Further, trials are reviewed later from a patient protection perspective by institutional review boards. Peer review may be superfluous in this case.

Building a new pre-print archive for all scholarly works (perhaps with links to published works) not only meets the needs of clinical research, but it solves the problems of having hundreds of different institutional repositories with differing standards and versions, and communication difficulties between repositories. At the same time,

it addresses the concerns commonly expressed by authors about current green archives—concerns such as visibility, sustainability, complexity, and so on. Issues of longevity could be managed across the spectrum by adhering to accepted and emerging standards for digital archiving and preservation.

## **PUBLISHER AND PUBLISHING REFORMS**

The consequences of a single repository system might be good for knowledge and library budgets, but would it also be good for publishers? Would it mean a reduction in subscription revenues for publishers? This is an issue that will require careful thought and further discussion. One would hope that a new marketplace would abound with new opportunity. For instance, publishers will still have their work cut out for them in terms of identifying, editing, reviewing, publishing, and curating quality works, and they will no doubt also recognize the potential for developing an array of new value-added information products (such as data mining and other tools that work with the repository to analyze and work creatively with the content)—which would diversify their publishing portfolio and also provide a much-needed and much-valued services back to science.

What other publishing reforms might be on the horizon? To start, should we think in terms of whether journal articles are the right way to go at all? They're difficult to produce, and as in the case of clinical research information, they aren't necessarily the best-suited formats for capturing information and showing how it is connected. Data files are sometimes provided but they aren't connected to each other or to the analyses. For the most part, papers today look just like they did 100 years ago. A number of interesting alternatives for showcasing the interrelatedness of information exist, such as David Wojick's issue-tree idea (<http://scholarlykitchen.sspnet.org/2013/07/10/the-issue-tree-structure-of-expressed-thought/>.)

Many other models exist, including these two suggested in the K|N report (as cited earlier):

1. Having scholarly societies form publishing cooperatives to benefit from economies of scale
2. Creating collaborations between publishers and institutions where all institutions pay in a certain prorated amount to help support an entirely open access publishing system.

## **TENURE REFORM**

As for reforming the tenure system, technology consultant and former CalTech researcher Eric Van de Velde suggests a gradual approach that lets tenure committees try new approaches in a risk-free manner. Specifically, he suggests this: “Suppose, for example, that every senior appointment requires at least one scholarly work in a non-traditional form. Such a requirement would hardly be a hardship on candidates. It is difficult to imagine that such a requirement would harm the institution. Yet, it would guarantee that every senior faculty member has personally experimented with at least one nontraditional form of scholarly communication. Even a candidate who complained about that requirement and succeeded in obtaining an exception would be forced to think seriously about scholarly-communication issues. I have no idea where such an approach would lead us. I cannot guarantee it would even lead to open access. However, in my experience, most people tend to become open-access converts after studying the issues. Mostly, it is just a mechanism to suspend the autopilot approach to scholarly publishing for at least a brief moment in every researcher’s career. In addition, it injects the issue into the discussions of the committees that really matter.”

## There is a significant diversity of attitudes among stakeholders in the scholarly communication system.



And these stakeholders all have an important role in shaping the future of scholarly publishing. As such, they are all charged with interpreting this system and turning reform concepts into reality.

Are they all acting with the same understanding about issues like open access, and moving forward together toward solutions that will work for everyone? The answer is a resounding “no.”

We have noted that the majority of authors are enthusiastic about open access but at the same time have serious concerns about Creative Commons licensing (see the Taylor & Francis survey in The Current State of OA section of Annex 6, item 16). We have noted that there are open access advocates who are willing to accept a limited role for toll-access publishing or for public access models or embargoes, institutions that embrace the self-archiving of research materials while others eschew it due to complexity and dubious benefit, pricing models that work for some large libraries but cause information scarcity for smaller libraries and research institutes, some publishers who make open access part of their business and others who try unsuccessfully to create OA models their customers embrace, authors and institutions who have embraced the “author pays” mode of open access publishing, and those in the global periphery for whom all pricing models—on the author side and the consumer side—have been harmful.

Without coordinated movement and action, we have confusion, dissension, and outcomes that are neither optimal nor timely. Therefore, more concerted and continual communication between the key stakeholder groups in scholarly communications is a vital first step toward improving outcomes, and this is the key recommendation of the OSI working group.

Whether compromise ends up being the order of the day or simply realizing where we can work together toward our common interests, and perhaps figuring out new and creative ways to reach our goals, reforming the scholarly publishing system is an important task that may end up having more impact on the dissemination and sharing of knowledge than almost any other task we can imagine. The path we need to follow may even be reasonably simple and straightforward, and the solutions may be rooted in simply reflecting thoughtfully on our practices, expectations and cultural norms and recognizing our common goals.

This change may be gradual and predictable, or it may end up being disruptive. Either way, the efforts and outcomes may create new worlds of opportunity and discovery.

Can we demonstrate conclusively one way or the other that reforming the current scholarly publishing system is important and that open access is the best tool for the job? No. Sound opinions exist on all sides of this issue. But the evidence suggests that without working together to clarify this issue now, the information access situation in scholarly publishing is only going to deteriorate.

Therefore, the OSI working group is recommending that a series of decision-oriented conferences be organized over the next ten years between all key stakeholders in the scholarly communication system. The purpose of these conferences will be to discuss the issues at hand, find common ground, and plan a roadmap forward—and importantly, to revisit and revise these plans annually. Every key group in the scholarly communications system will be heard, and every group will need to take ownership of the process and outcome in order for this agreement to be effective and sustainable. Effective and

sustainable solutions to complex issues like this cannot be imposed from the outside—they need to be created by the affected stakeholders—and this is what we’re proposing happen, for the first time on this issue.

Ultimately, these conference deliberations may come down to a leap of faith—to the realization that we don’t know what we don’t know. Immediate and unfettered access to science data on a large scale has never been attempted, so there’s no expectation of such use, no traditions of such use, and no evidence. Even if immediate and unfettered access happened tomorrow, it might take years before science can start capitalizing on the potential. Pockets of science already do, of course, but these are the exceptions. “We can’t predict,” says Michael Eisen, “what creative people and organizations will do with the bounty of information. What we do know—from countless examples within science and elsewhere—is that when people are given the ability to use data like this, they do amazing and unpredictable things, and that once they do we can’t understand how people lived without them.” Of course, absolutely nothing may happen as a result of vastly improved OA or the development of an all-scholarship repository. But then again, we may see collaborations, innovations and creations that will change the world.

David Wojick suggests that “A little computer modeling might go a long way here. Modeling is not prediction but it does show what is possible given plausible assumptions.” Research by Wojick and others has shown that increasing the contact rate might speed up science significantly, since increasing the contact rate is arguably just what OA wants to do (see Wojick’s paper in The Future of Open Science section of Annex 6).

Even without accelerated discovery, it’s hard to imagine that nothing at all will come of this effort. The fact remains that there has never been a systematic, rigorous cataloguing and indexing of science research over the years that enables researchers on a granular level to easily examine what’s been tried, what’s been discovered,

who is working in their field around the world, what the outcomes have been from various experiments, or how their study data relates to previous similar experiments—the data headers are different, units are different, methods of analysis are different, and more.

So gathering everything together in one place—like we’ve speculated about doing with the all-scholarship repository—might be an important first step in what may eventually develop into a monumental long-term project with significant impacts.

Practically speaking, of course, there are huge challenges ahead. Data collection and standards aside, for instance, companies and industries (like big pharma, who sponsor a lot of important drug research) and institutions (think technology transfer units) will want to play their cards close to their vests and not release information that compromises their future intellectual property positions. And public policy may become more challenged, not less, since everyone with a political axe to grind will be able to take this same data and misrepresent it (not necessarily maliciously, but just through inexpert analyses) to support their own viewpoints.

Still, on balance, the push for more-better-faster has become part of our modern hi-tech society’s mantra. The digital natives entering science today won’t settle for anything less than broad, rapid, easy access. The prevailing culture of science and academia will corrupt them for a while, but eventually the walls that prevent greater sharing are going to come down. Having a process in place to facilitate this dismantling, along with a roadmap to ensure a continuity of quality and access, is a vital task.

The OSI working group is pleased to present this working paper, and invites your feedback.

Sincerely,

The OSI working group

Exactly how many journals are there anyway? According to Michael Mabe in his 2003 study of journal publishing trends (see Annex 6, Overview section, item 9) there are wildly differing estimates of the number of academic journals because there are widely differing definitions of what constitutes academic publishing. “The main reason there have been so many varying estimates of the number of learned periodicals in the world,” says Mabe, “is almost entirely down to the simple matter of definition. What exactly constitutes a journal? More importantly, what constitutes a learned journal? The most important characteristics of a learned journal are the nature of its content, whether the title is actively publishing at the present point in time, and whether the content has been through a peer review system to ensure its quality. Unless these key distinctions are taken into account when estimating learned journal numbers from directories such as Ulrich’s, erroneously high values will be obtained.”

nSCI researcher Marcus Lambert conducted a new count from Ulrich’s in November 2013 and came up with a figure of 30,796 journals, of which 26,529 were print and 4,267 were online-only (using the search criteria: (“Active”) Serial Type:(“Journal”) Content Type:(“Academic / Scholarly”) Key Feature:(+“Refereed / Peer-reviewed”).

Mabe concluded that the growth of journals has been relatively constant over the past 300 years, with an average annual growth rate of about 3.5% since 1800. This rate results in a doubling of the number of journals about every 20 years. This growth rate appears to correlate closely with (and be caused primarily by) the growth in the number of researchers. For every additional 100 refereed papers these new researchers add to the journal universe, a new journal will be born.

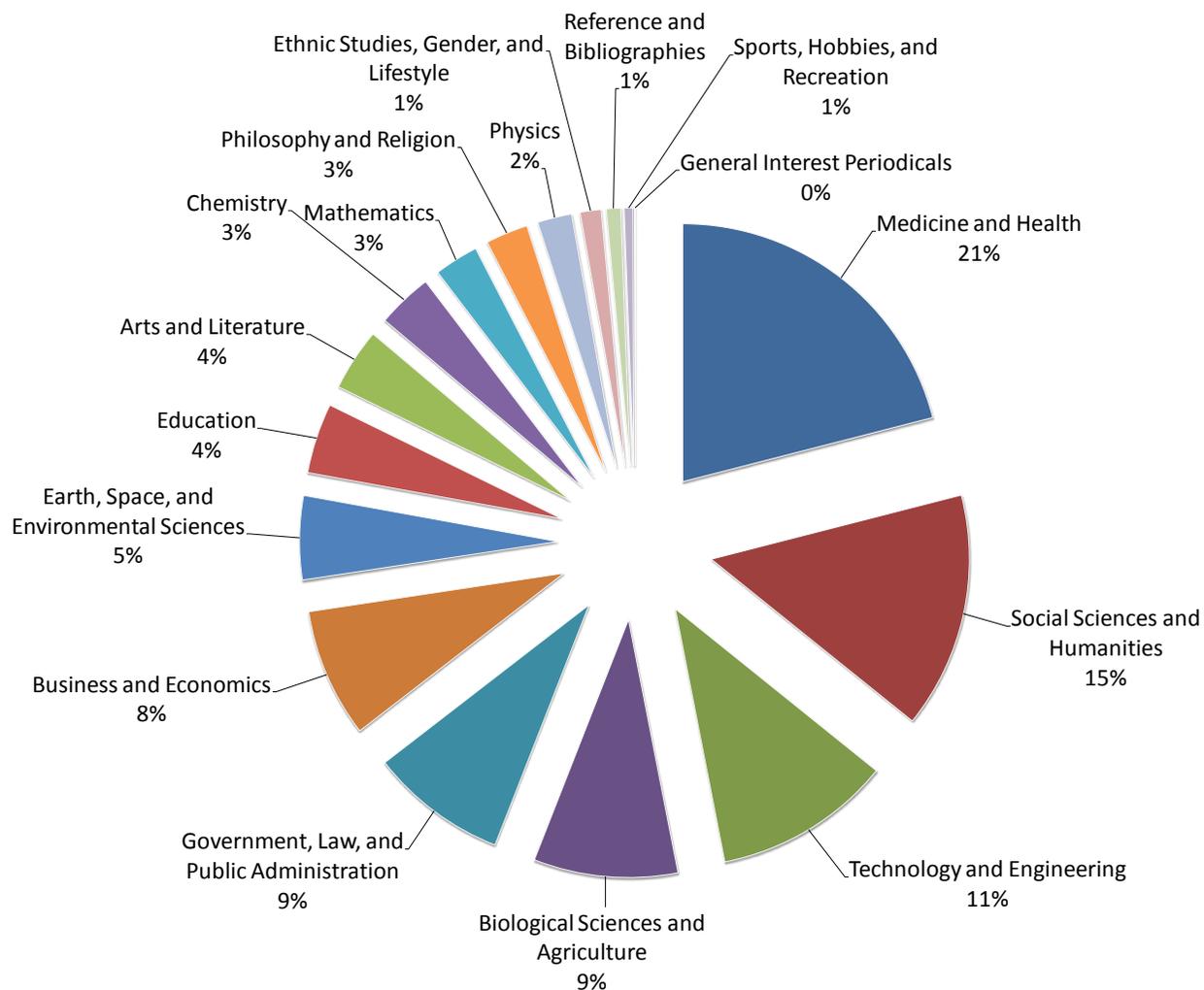
Who publishes these journals and what are they about? As with the problem of getting an accurate count of journals, getting accurate breakdowns of journals de-

pends on what’s being counted. The subject breakdown in the pie chart to the right is based on what’s covered in the 30,796 active and refereed academic journals listed in Ulrich’s as of November 2013 (according to Lambert). The breakdown in the table the right is based on a much broader listing of journals and publishers that appears in journalseek.net—some 100,000+ journals and 5500+ publishers. The figures provided by this website aren’t entirely accurate and are helpful for approximation purposes only.

Taking into account only the top group of 42 publishers from this list, who publish 100 or more journals each and together account for 16,825 journals between them, the top four publishers (in order, Elsevier, Springer, Wiley, and Taylor Francis) account for 44% of the market. The generally-accepted figure is that these top four account for 42 percent. According to this source, the Chinese journal conglomerate Chinese Electronic Periodical Services is the fifth largest publisher, with Sage coming in sixth. Both print and online journals are included in these charts, and some journals are listed in more than one category of the pie chart.

Language profiles are also interesting. Some have observed that the rapid growth of publishing in China suggests that Chinese will become the lingua franca of science. Others note that English continues to grow as the primary language of science when measured in terms of the science citation index. When simply calculating the percent of the English language journals appearing in the latest count from Ulrich’s, the result is 25,531 out of the 30,796 active and refereed scholarly journals counted by Lambert, or 83 percent.

nSCI hosted a conference in the Fall of 2013 covering a wide range of journal-related issues. More statistics about journal readership, language, search patterns, impact factors, prices, distribution and more are posted online as part of the conference summary at <http://bit.ly/1zkk6PJ>.



PUBLISHERS WITH 100 OR MORE JOURNALS	JOURNALS	% OF MARKET
Other (24 publishers)	3,853	22.9
Elsevier Science	2,655	15.8
Springer (before 2015 merger with Nature Publishing Group)	1,759	10.5
John Wiley and Sons	1,484	8.8
Taylor and Francis Group	1,446	8.6
Chinese Electronic Periodical Services	1,307	7.8
Sage Publications	580	3.4
Emerald Group Publishing Limited	414	2.5
Bentham Science Publishers	375	2.2
IEEE	357	2.1
Inderscience Publishers	346	2.1
SciELO	330	2.0
Lippincott Williams and Wilkins	298	1.8
BioMed Central	293	1.7
Walter de Gruyter GmbH & Co. KG	288	1.7
Portal of Scientific Journals of Croatia	281	1.7
Japan Science and Technology Information Aggregator, Electronic	258	1.5
Cambridge University Press	256	1.5
Oxford University Press	245	1.5
<b>TOTAL</b>	<b>16,825</b>	<b>54.6</b>

A constellation of issues in journal publishing is affecting access to the knowledge produced by academia, particularly in science.

Here is a summary of a few of these issues:

1. Rapid growth of journals. Since around the time of Isaac Newton—indeed, 2015 marks the 350th anniversary of the publishing of the first science journals in 1665—journals have served as the official repositories of and vehicles for sharing research, and inasmuch they have been extremely important to both science and society. But journals have grown increasingly specialized (driven at least in part by the increasing volume of and specialization in science research) and their numbers have doubled every 20 years to exceed 36,000 today. So while it was possible for a biologist in the 1990s to keep current by reading a handful of major journals like *Science*, *Nature* and *Cell*, a biologist's required reading list in 2014 has fragmented and multiplied into dozens of journals. This multiplication and segmentation has had consequences. In some fields scientists now have to read 400 articles a year to stay current (see the nSCI 2013 conference proceedings for more details). Editors must sort through and interpret this ever-growing pile of science information to identify worthy research news for scholars (through their journals), and journalists—who funnel this information to the public—face the same challenge of sorting through a growing mountain of scientific research while their own newsrooms shrink.
2. Rapid growth of research outputs. According to a 2012 study by the National Science Foundation, global spending on research and development more than doubled between 1996 and 2012, leading to a concomitant increase in the number of articles produced, particularly from Asia. This increased output puts serious pressure on traditional journals, who not only have more articles to shepherd through the publication process, but also more articles to reject (a costly process that creates no direct revenue). If these articles are unable to get published, there is a risk of publication bias and research waste, especially if novel or statistically significant research is preferred for publication.
3. Quality. The amount of growth and segmentation of research has affected quality. The editorial and scientific integrity process (usually called peer review) is not uniform from one journal to the next. The complete and transparent reporting necessary to assess the quality and validity of research is lacking in most publications. Also, increasing pressures on postdocs and declining grant funds may be contributing to higher rates of retraction and issues of reproducibility.
4. Fraud. Outright fraud, while still exceedingly rare, has nonetheless also become more evident as the number of journals continues to proliferate. The possibility that unchecked fraud exists in some journals (especially to the extent that it may be driven in part by pressures to publish) is an alarming prospect since knowledge is built upon knowledge in science. If some of this latticework is corrupt then many research studies may need to be retraced and rebuilt.
5. Costs. Increased costs are yet another consequence. Journal prices have been increasing rapidly along with the number of journals that

libraries need to acquire, even as academic budgets are squeezed. Libraries have responded by bundling and cutting subscriptions and monograph budgets while smaller research organizations and researchers in less-developed regions of the world have been pushed into information underload. Even single article download costs from publishers can be constraining for individual researchers and smaller firms, and are certainly not sustainable at scale for larger institutions—usually around \$32 for a research article only a few pages in length.

6. Access. Most journals have copyrighted content just like any other book or periodical, which means research gets locked up for reuse after publication. In many cases, authors simply sign over copyright to journal publishers; they often don't know they may be able to negotiate to retain copyright, and therefore don't try. In effect, research institutions must pay to gain usage-restricted access to the very information that they paid to develop in the first place. Publisher policies continue to be highly restrictive over the longer-term as well, including disallowing the posting of published versions of articles even a decade and more after publication.
7. Literacy. Information literacy is an issue that extends beyond the borders of this discussion. While journal articles are written primarily for academicians, there is also a broad audience of educators, policymakers, professionals, and interested citizens who would benefit from having access to visible, accessible, and understandable journal articles. Widespread and barrier-free access to high-quality research is not the case at present.
8. Readability. Closely connected to the issue of literacy and access is the manner in which journals are written. Increasingly over the last 30 years—in part due to specialization, in part to the growing internationalization of English (as establishment of English as the lingua franca of science), and also in part due to changing expectations about how journal articles should be written—the language of journal articles has become increasingly impenetrable. If greater cross-discipline and even public access is a desired outcome of increased open access availability (and not all are agreed on the latter), one wonders whether the pressure to share more effectively will encourage a reform in journal writing standards.
9. Tenure. The link between tenure and evidence of published scholarship is old and needs to be updated to reflect new possibilities and evidence of quality and impact. Career motives (pressured by more intense postdoc competition than ever before) incentivizes researchers to publish in the highest impact factor journals available. The scientific community needs to ensure that assessments don't work against improving openness. To be clear, this isn't just about tenure—it's about academic evaluation in general, and as Jean-Claude Guédon points out, funding agencies are equally guilty of the same bias. Quite typically, for instance, grant proposals will ask for a list of publications “in major international peer-reviewed multi-disciplinary scientific journals and/or in the leading international peer-reviewed journals... the number of citations (excluding self-citations) they have attracted (if applicable);” and so on. This emphasis on “leading and major” outlets is pervasive in science in general—not just tenure—and it needs to change before the OA discussion can change (and there are, of course, quite a few assessment alternatives available nowadays). William Gunn notes that NIH and NSF have both recently amended their biosketch instructions to call for “research products” or other non-journal contributions.

10. Embargoes. In scholarly publishing, an embargo is a prescribed period of time during which access to research results is restricted to subscribers. Afterward, the research may be made publicly available, at least for US federally-funded research. The purpose of this restriction is to protect the revenue of the publisher. The established timing of the embargo is a balance between what constitutes a “fair” waiting period for the publisher before an article becomes free to access, and what makes sense in terms of so-called article “half-lives”—how long it takes for about half of the downloads of an article to occur. In 2013, the US Office of Science and Technology Policy, or OSTP, initiated a public access program to make federally-funded research available after an embargo period of 12 months. Many in science feel this figure creates an unnecessary delay and puts a brake on discovery and innovation. Some in the humanities generally feel that this wait can be even longer (and in fact, the data suggest that a 24 month embargo may be fairer in the humanities—see the Chris Wickham interview in Annex 6, The Current State of OA, item 10). One academic association in the humanities has actually adopted a policy urging universities to allow dissertations to be embargoed for up to six years; in the UK, the allowable embargo periods for government-funded work are six months for STM research and 12 months for the humanities. Are embargoes an unjust and unjustified burden on science? Putting a mandatory brake on the entire system of research scholarship to redress the needs of publishers—a need which has not been established to exist (see research questions posed earlier)—may not even be justified from a publishing perspective. See Shan Sutton’s research article on recent embargo practices (listed in the annex) for more details.
11. Peer review: As described earlier, peer review has come under fire in recent years for not adequately “protecting” science from fraud, for not adding value in all cases or with all journals, and for not promoting the most innovative science. Is the peer review system broken as currently constructed?
12. Impact factors: Impact factors today are a hot topic. While they may serve a useful function in scholarship, insofar as trying to visualize the flow, importance, and interconnectedness of research, they also incentivize academics to publish in certain journals over others, and this incentive affects choices that may need to change in order to improve open access. In addition, most of our attempts to measure impact lack accuracy (some newer measures are getting better, but traditional measures are about as accurate as a coin-flip, showing no significant correlation between measured and actual impacts; see the arXiv paper in the impact factor section of recommended reading). So if we need to measure impacts, we need to find better ways of doing so—at minimum, making sure that the impacts we’re measuring have more to do with articles than journals. “Fundamentally,” says William Gunn, “reputation should accrue at the level of the author and the paper, and journal brand should be a reflection of this, not the other way around.” But we may also want to ask ourselves why we’re focusing on evaluation to begin with. Is it to improve science, improve publishing, provide ROI data to funders (which may also foster the temptation to fudge data), give benchmarks to tenure committees, or exactly what? As with the publishing question itself, there doesn’t need to be a one-size-fits-all answer here, and what works best for evaluating interconnectedness might not be at all suitable for evaluating success toward earning tenure.
13. Open access. See Annex 3.

## There are also a few important issues in open access itself affecting the ability of this movement to expand.

These issues include:

1. Defining open access. The idea and practice of providing free access to journal articles has been around for decades, but it was clarified and formalized as “open access” through a series of initiatives in the early 2000s (such as the Budapest Initiative of 2002 and the Berlin declaration of 2003). As has been discussed earlier in this paper, although OA has seen widespread acceptance and adoption since then, open access today is encountering growing pains as proponents have developed different interpretations of what “open” really means. Does it mean only gratis (free to read) or must it also mean libre (free to reproduce and adapt as long as the creators are attributed)? Are embargoes allowed and if so under what conditions? What about author fees and subscription fees? How do these variations factor into the conversation? Is the line between public access and open access getting blurred or conflated in some quarters?
2. Perceptions. Publishers and their audiences are still struggling with perceptions (many completely unfounded, and some less so) of the “open” model. Lots of confusion and doubt still exists among faculty—about the quality of OA journals and peer review, impact factors, growth rates (i.e., whether this is just a fad), costs (article processing charges, or APCs), preservation (is print a more permanent medium than digital?), fraud (concerns about publishing in a questionable or disreputable journal), and more.
3. Slow growth. At present, approximately 15% of journals are open access; approximately 5% are both open access and Creative Commons licensed (essentially meaning free to share with attribution, but with a number of possible variations). Institutions have embraced the OA movement with varying degrees of enthusiasm. Some institutions encourage OA publishing by paying for their faculty’s article processing charges (APCs) that help make OA publishing possible, some do not.
4. Mandates and OA self-archiving. Self-archiving mandates were supposed to solve at least the bulk of the problem caused by journal articles disappearing behind paywalls. These mandates, which have been around for about ten years now, were designed and intended to allow the open-access archiving of all scholarly works in their pre-print format—that is, usually, in an unedited and non peer-reviewed format. As Michael Eisen states, they are about “adjusting the system so that the interests of individuals and the interests of funders, universities and the public are aligned,” but they are also about making people do what they otherwise would not do. Many journal publishers agreed to this arrangement, with the caveat that published versions of works (if a work was accepted for publication in a journal, which is not always the case)—which might differ a lot from the pre-print versions or maybe just a little—could not be self-archived, although the peer-reviewed and accepted manuscript version could be shared. These “published” versions would remain under the control of the publisher, sold for a profit (called “paywall-protected”), and often be copyright-restricted as well. A quick review of the current state of self-archiving policies around the world shows mixed results.

For instance, hundreds of mandates have been attempted (about 350 total, one-third of which are institutional, and most of which have been voluntary, calling into question whether these voluntary measures can actually be called “mandates” at all), but the results of these efforts have been mixed at best—inconsistent both in scope and in implementation from one institution to the next, and from one type of mandate to the next (institutional vs. thesis, program, thesis, or funder). Some mandates specify time periods, others mention specific repositories, still others require Creative Commons licensing, and the list goes on.

Second, publishing analyst Steven Harnad estimates that globally, the current level of green OA is about 12% but that this figure could jump to 80% if effective green mandates were implemented. But what’s an “effective mandate?” Maybe not voluntary? A growing number of US universities are adopting OA policies, but virtually all of these allow authors to “opt-out” on a paper-by-paper basis. Also, even if new federal “mandatory mandates” were introduced, most US institutions of higher education are private, and most public ones are governed at the state level (and also at the corporate level in the case of for-profits). Therefore, it is unlikely that Congress would pass a new law saying that “all universities and research institutions

shall do things this way.” However, new federal mandates regarding information sharing policies can work if they are attached as a condition of receiving federal research funding (for instance, all institutions receiving federal funding must comply with an array of provisions regarding equal rights, non-discrimination, equal access, and more—Title IX, ADA, etc.).

Finally, there are a number of problems with the current system of institutional repositories. For one, institutional repositories don’t always contain the final version of record (and locating the final version isn’t always easy). Also, different repositories may use different licensing rights. There are formatting differences: Some prefer XML or HTML-formatted articles, while others rely on pdfs. Metadata is also an issue: All try to use a common metadata structure (data that describes the author, subject, content, etc.), but this data is usually incomplete or inadequate, including missing data on copyright status (the copyright holder, type of Creative Commons license, etc.). The most common metadata structure is “Dublin Core”—a set of 15 metadata elements that are not specific enough to be useful across a huge variety of disciplines. More subject or discipline-specific metadata constructs might be helpful, as well as better format standards and conversion tools.

A variety of important initiatives are happening worldwide to integrate existing warehouses of academic knowledge (most commonly known as institutional repositories, or IRs).

The main issue with integration is that research repositories are built to a wide variety of specifications, utilize a wide range of technologies, and are generally incomplete (in terms of the information they contain, the recency of their information, and the degree to which this information is accurately and completely notated). Tied to the issue of integration, the ability to “hook” into different repositories and pull out information is limited and mostly results in portals that simply provide links to articles instead of full-text resources that are truly integrated.

The prospects of improving repositories are questionable: University repositories are generally not well funded and therefore may not offer a wide range of services, including metadata support or review. There also aren’t adequate incentives for depositing into repositories (and/or, there is apathy, a lack of awareness, confusion, and even disincentives in some cases). Goodwill appears to be the primary motivator at present (as the survey data bears out; see Annex 6). Therefore, the success of integrative efforts depends on the use and development of institutional efforts, as well as the ability to integrate these disparate systems into a single, usable system. It may be worth considering whether a better approach is to simply work toward building a single repository, perhaps with robust distributed capabilities to allow for continued institutional, community and discipline-level “ownership” and management of information.

The current state of institutional repositories is tangled, and yet worldwide there are a number of worthy efforts trying to come to grips with the issues and solutions. Quoted here is a portion of what Richard Poynder wrote about the effort to network research repositories in May 2014 interview of Kathleen Shearer, executive director of the Confederation of Open Access Repositories (source: <http://bit.ly/164kjDI>):

“In October 1999 a group of people met in New Mexico to discuss ways in which the growing number of “eprint archives” could cooperate.

Dubbed the Santa Fe Convention, the meeting was a response to a new trend: researchers had begun to create subject-based electronic archives so that they could share their research papers with one another over the Internet. Early examples were arXiv, CogPrints and RePEc.

The thinking behind the meeting was that if these distributed archives were made interoperable they would not only be more useful to the communities that created them, but they could “contribute to the creation of a more effective scholarly communication mechanism.”

With this end in mind it was decided to launch the Open Archives Initiative (OAI) and to develop a new machine-based protocol for sharing metadata. This would enable third party providers to harvest the metadata in scholarly archives and build new services on top of them. Critically, by aggregating the metadata these services would be able to provide a single search interface to enable scholars interrogate the complete universe of eprint archives as if a single archive. Thus was born the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). An early example of a metadata harvester was OAIster.

Explaining the logic of what they were doing in D-Lib Magazine in 2000, Santa Fe meeting organizers Herbert Van de Sompel and Carl Lagoze wrote, “The reason for launching the Open Ar-

chives initiative is the belief that interoperability among archives is key to increasing their impact and establishing them as viable alternatives to the existing scholarly communication model.”

As an example of the kind of alternative model they had in mind Van de Sompel and Lagoze cited a recent proposal that had been made by three Caltech researchers.

Today eprint archives are more commonly known as open access repositories, and while OAI-PMH remains the standard for exposing repository metadata, the nature, scope and function of scholarly archives has broadened somewhat. As well as subject repositories like arXiv and PubMed Central, for instance, there are now thousands of institutional repositories. Importantly, these repositories have become the primary mechanism for providing green open access—i.e., making publicly-funded research papers freely available on the Internet. Currently OpenDOAR lists over 3,600 OA repositories.

### **WORK IN PROGRESS**

Fifteen years later, however, the task embarked upon at Santa Fe still remains a work in progress. Not only has it proved hugely difficult to persuade many researchers to make use of repositories, but the full potential of networking them has yet to be realised, not least because many repositories do not attach complete and consistent metadata to the items posted in them, or they only provide the metadata for a document, not the document itself. As a consequence, locating and accessing content in OA repositories remains a hit and miss affair, and while many researchers now turn to Google and Google Scholar when looking for research papers, Google Scholar has not been as receptive to indexing repository collections as OA advocates had hoped.”

### **GLOBAL PROGRESS**

According to Shearer, “There are numerous national and thematic repository networks around the world, which link repositories with each other. These have evolved based on unique requirements and mandates; are at different stages of development; and reflect varying levels of integration. Some national networks, such as in the UK, Portugal, Argentina, and Spain, are very cohesive and have a number of robust services supporting their repositories. Others are less developed and unified, and revolve more around a community of practice for repository managers. Broadly speaking, repository networks can be characterized as having one or more of the following aspects: community of practice, adoption of common standards for metadata and vocabularies, centralized harvester, catch-all repositories for orphan publications, and other value added services. In addition to national and thematic networks, regional repository networks are being developed to connect repositories across national boundaries.”

The La Referencia repository network in nine countries at present (Argentina, Brazil, Chile, Colombia, Ecuador, and Mexico, Peru, Venezuela and El Salvador) harvests from national nodes that in turn are harvesting from institutional repositories. “The initiative,” says Shearer, “began as a project funded by the Inter-American Development Bank (IDB) and is now managed by CLARA, the organization that manages the high-speed network in Latin America.” OpenAIRE is a project funded by the EC to develop repository infrastructures in the EU. OpenAIRE aggregates the research output of EC-funded projects and makes them available through a centralized portal. Shearer notes that “OpenAIRE-compliant repositories adopt common guidelines so that content can be aggregated into the central portal. OpenAIRE, with renewed funding from the EC, will soon begin to develop other value added services such as text mining and reporting tools, which enable users to better use the content and track funded research outputs.” China is also investing in the development of institutional repositories and sees these as the main route toward open access, as opposed to gold OA.

## US PROGRESS

The US is taking several different approaches to this challenge:

1. CHORUS—the Clearinghouse for the Open Research of the United States—is a private enterprise run by CHOR, Inc, a 501c3 nonprofit. According to their website at [chorusaccess.org](http://chorusaccess.org), CHORUS “leverages widely used technology to facilitate a simple compliance process, optimized search and dashboard services, and multi-party archiving and preservation capabilities.” It is intended to be “interoperable with other public-access delivery solutions, agency search portals, publisher platforms, and scholarly repositories and archives.” CHORUS points users to the best available version of articles on the publishers’ publication sites. The open programming and interface is intended to encourage innovators to develop new tools and functionality “that further support public access and facilitates text/data mining on funded research.”
2. PAGES—the Department of Energy’s Public Access Gateway for Energy and Science—is the DOE portal that makes scholarly scientific publications resulting from DOE research funding publicly accessible and searchable at no charge to users. Launched in August 2014 in response to the February 2013 OSTP memorandum, PAGES offers free public access to the best available full-text version of DOE-affiliated scholarly publications after an “administrative interval” of 12 months. According to their website at [osti.gov/pages/search.jsp](http://osti.gov/pages/search.jsp), “DOE PAGES<sup>Beta</sup> contains an initial collection of journal articles and accepted manuscripts as a demonstration of its functionality and eventual expanded content. Over the next year, additional metadata and links to articles and accepted manuscripts will be added as they are submitted to OSTI, with anticipated
- annual growth of 20,000-30,000 publicly-accessible articles and manuscripts. When DOE PAGES moves beyond the “beta” period, it will offer distributed full-text access to all DOE-affiliated accepted manuscripts or articles after an administrative interval of 12 months.”
3. SHARE—the “Shared Access Resource Ecosystem”—is a university-centric solution being developed by the Association of American Universities, the Association of Public and Land-grant Universities, and the Association of Research Libraries, working with the Center for Open Science as development partner. According to a June 2013 briefing paper (<http://www.arl.org/storage/documents/publications/share-proposal-07june13.pdf>), “SHARE envisions that universities will collaborate with the Federal Government and others to host cross-institutional digital repositories of public access research publications that meet federal requirements for public availability and preservation. SHARE will have a central metadata aggregator and will also release event metadata. According to the July 2014 SHARE newsletter, in June “the team began evaluation of content, API, and harvesting options from ClinicalTrials.gov, the Department of Energy’s (DOE) SciTech, ImpactStory, Cross-Ref, the National Library of Medicine’s PubMed Central, and the Public Library of Science (PLOS). Development of initial code libraries and parsing of content began on ClinicalTrials.gov, DOE SciTech, and ImpactStory. In addition to these initial content streams, the team built a scraping integration API called “scrap\_i,” which will be used to manage the incoming content from many sources.”
4. DPLA: Although not focused on scholarly publishing, the Digital Public Library of America has put together a collaborative model that might work well for combining institutional repositories of scholarly works. This model combines a

central website with distributed ownership, allowing libraries across the country to update and maintain materials while also making these materials accessible both through the DPLA website (dp.la) and through regional service hubs.

## ISSUES WITH CURRENT EFFORTS

All of these systems—PAGES, OpenAIRE, and others—are just portals and aren't designed to include the full text of articles (or additional research documents). And the success of all of these efforts will depend on the continued development and integration of institutional and commercial repositories. PAGES, for instance, will incorporate publisher-supplied metadata from CHORUS and link to publicly-accessible content.

The reason PAGES chose this route instead of the ASR route, according to Mel DeSart, head of the engineering library at the University of Washington, dates back to the February 2013 OSTP directive encouraging the development of "a strategy for leveraging existing archives, where appropriate, and fostering public/private partnerships with scientific journals relevant to the agency's research." OSTP was also looking for solutions that encourage public-private collaboration to:

1. maximize the potential for interoperability between public and private platforms and creative reuse to enhance value to all stakeholders,
2. avoid unnecessary duplication of existing mechanisms,
3. maximize the impact of the Federal research investment, and
4. otherwise assist with implementation of the agency plan.

A single, integrated platform is, in essence, what the OSTP intended to accomplish, agrees William Gunn. "They wanted to expand PubMed Central to PubFed Central. This broke apart due to interagency politics and

now each agency is to come up with their own plan. They're overdue to submit their plans and so far only NIH and DOE are close, with DOE leaning towards CHORUS and various other agencies working on their own solutions."

What are the keys to making these systems work as hoped? According to David Wojick, "The key as far as PAGES is concerned will be providing good metadata that accurately identifies the funders of the research being reported on. This metadata is turning out to be surprisingly difficult to get from the authors, even for the publishers. PAGES also needs the publication date in order to apply the mandated embargo period. Acceptance per se is not enough. Linking to publishers and IRs on a grand scale is pretty complex."

The biggest problem with SHARE, writes Wojick, is that they are simply going to harvest whatever metadata each repository provides, "so it is not clear how much true integration we will get. They are just beginning to address that issue."

## TO BUILD...

So what might a single, connected, all-scholarship repository accomplish? One way to approach this question is to explain what the currently disconnected system cannot accomplish. Without a common platform or interoperability framework, digital fragmentation will continue to make information access worse and worse. Imagine the Internet without common standards—using computers that don't talk to each other, don't use a common language, and yet are supposed to work together to create a robust, dynamic, and usable pool of global knowledge.

This communications dysfunction is real and immediate in science. It demands not just creating a pie-in-the-sky future, but getting a handle on the information we have right now. Take HIV/AIDS research. No single database of research exists—nothing that ties together 30 years of research studies, data, protocols, and ideas. There are si-

los of databases, each with their own unique data headers, but there is no standard between them (one is under development), no effort that ties all of this work together (a small trial effort is struggling for funding), and no agreement among research institutions and networks to do this on a grand scale—not even an archive of all HIV/AIDS research papers that ties everything together and helps modern researchers look back and see what’s been done already and what was learned.

Even if the sole purpose of an ASR effort was to improve how we spend our increasingly scarce research dollars, we can posit that building a single, interoperable repository will save billions of dollars and vastly improve research efficiency and effectiveness.

And that’s not even considering what it will do for discovery. Or considering fields outside of HIV/AIDS research.

What would happen to institutional repositories and efforts like SHARE, CHORUS, PAGES, and other if a new, massive ASR was funded? Probably nothing, at least for years and years. It will take a village to pull this off, from institutions to current repository players to publishers and government agencies. This enterprise will take at least a decade to build but support will grow over time (and with it, funding and discovery). As it grows, the incentives will become the elephant in the room. For reasons of visibility, discoverability, access, sustainability, reputation, ease of use, and so on, the reasons for using this system will be obvious and the reasons for not using it will sound irresponsible.

Also, a hundred new actors will enter this space—far more than there are now—and the competition this creates will drive innovation. Specifically, there doesn’t need to be just one ASR. In fact, says Glenn Hampson, the executive director of nSCI, “for reasons of sustainability, access, innovation, security, preservation, and more, there should be many, and these should be shared, linked, archived, and replicated. The first ASR should be shared with Elsevier, so Elsevier can put its own spin on

the information architecture and attract users based on how well it does in this regard. The same goes for AAAS and anyone else who wants to take this data and run with it. The first ASR would be the primary, common repository, but not the sole interface. Every institution and research organization should be able to get into this system and clean up its files, add links, add profiles, add new metadata, create colloquial summaries of the research that can be understood by outsiders, add databases, create connections to other research and researchers, do real-time test-comparisons of data (already being done through systems designed by LabKey in Seattle), and much more.”

The right balance of carrots and sticks is needed moving forward. Writes Dee Magnoni, research library director at the Los Alamos National Laboratory, a new all-scholarship system built from the ground up can include all the incentives needed to encourage participation—“the emerging ability to feed profiles with repository content, altmetrics, ORCID and VIVO info, etc. can go a long way to building incentives.”

### **...OR NOT TO BUILD?**

On the other hand, do we need to design a new system in order to have true interoperability between repositories? That’s the big question right now. Some say yes, others say no—at least not yet. It’s possible, for instance, that we can accomplish the same end-goals of an ASR by simply improving participation in existing institutional repositories (maybe through a massive sustained outreach campaign, suggests William Gunn). Or by continuing to build systems to interface with the repositories that researchers already use. “There are millions of researchers on Mendeley,” for instance, says Gunn, and “many hundreds of thousands ... have up to date publication records on their profiles, suggesting one option is to leverage this work already done (and the hundreds of thousands in ORCID, for example), to fetch the publications automatically. Mendeley and ORCID have APIs & repos have deposit interfaces.”

Therefore, not everyone is convinced that re-inventing the wheel (yet again) is the best way to approach this issue. Writes Richard Poynder, “While I like the idea of an ‘all-scholarship repository,’ I do wonder how practical it is. The OA movement has been promising a distributed network offering something like that for over a decade. Yet thirteen years after the OAI-PMH protocol was created to facilitate interoperability between repositories I do not believe an effective network has yet to be created, not least because repositories do not implement metadata properly. So I suspect the task is far from trivial, both in terms of the technical and metadata work needed to build the infrastructure and because of the consequent cost, both the cost of setting it up and of maintaining it. And for so long as researchers continue to go to traditional publishers to publish their work any such network is going to be in competition with those publishers, who have a number of advantages.”

Indeed, Eric Van de Velde suggests that the objections to making any attempt at some version of an all-scholarship repository (either centralized or distributed) can be grouped like this:

1. In the absence of strong OA pressures from the scholarly discipline (as for HEP and ArXiv), local repositories are likely to be best to obtain the maximum content. Moreover, they distribute the costs and efforts.

2. Thus far, standardization of local-repository practices and metadata has not succeeded.
3. The problems with metadata together with the limitations of OAI-PMH have made the federated approach a complete mess in terms of discovering/searching/managing/measuring Green OA scholarship.
4. The metadata approach is fundamentally flawed for the long term.
5. Repositories completely open up full text and all data. A more complete argument on this here: The Metadata Bubble (<http://scitechociety.blogspot.com/2014/10/the-metadata-bubble.html>)
6. Institutions should be more open to having complete collections harvested by other institutions/services and collected into super-repositories.

## CONCLUSION

There are many inside and outside the current OSI group who are interested in continuing this initiative to explore a single repository approach. We all recognize that institutional repositories are an important—perhaps even central—issue in the OA conversation, and therefore, are calling it out here for further investigation and effort.

The current system has presented two major challenges for those in the developing world (often referred to by economists as the global periphery, or those nations outside the global “core” economically): a lack of access to research information, and a lack of visibility (and therefore perceived value) of local research and science.

The well-discussed problems of access to scholarly journals are even more acute in places where research and university (and library) budgets can be extremely low. At first, researchers in the global periphery took the view that open access would be a win-win situation, since it would allow them to read all research articles and the assumption was that they would be able to continue to publish (especially since most open access has not been an author-pays model)

However, there has been a gradual realization that new forms of scholarly communication may increase inequalities rather than bridge them. It has been observed that as “sweeping changes in the global north will see more northern research freely available to all online, the danger for locals is twofold: firstly, that they may be limited in their opportunities to publish (especially by expensive APCs) and, secondly, that their own research drowns in the worsening invisibility of the online discoverability sphere.”[15].

Indeed, as the author-pays model has become increasingly prevalent, there is a concern that open access could further marginalize researchers from these regions, creating “new enclosures to knowledge”[1], [14]. With average fees of US\$2,097 to US\$2,727 per article to publish in open access[2], APCs are completely out of reach for these researchers and make it very difficult to publish research generated in the south and by scholars from the ‘global north/center’ (as opposed to the global periphery), certainly not in the international journals produced in the north that would provide the kind of prestige measured by academic assessments systems all

over the world. In a recent commentary, Indian scholars assert, “The great hopes raised by the rise of OA seem to be receding, despite the fact that the share of papers available on OA is increasing and more and more funders and governments are coming up with OA policies. The different constituencies—librarians, scientists, economists, activists in advanced countries and emerging countries—expect OA to solve widely different problems.”[3]

It was initially espoused that APCs would not be an issue for researchers in the global periphery because publishers promised to grant those in the developing world a fee waiver (although those researchers would have to go hat in hand to publishers to give them charity publishing privileges, an unfortunate and power-laden relationship in itself). However, there is some indication that waiver rules are gradually tightening, and one can envisage a time when they are eroded away to practically nothing. For example, the most liberal waiver policy has always been operated by PLoS. Over time this policy has degraded. Initially, PLoS had a no-questions asked policy, but this ended in 2010 and in 2014 the rules were tightened further.[4] BioMed Central began to tighten their rules on fee waivers in 2006.[5] Obtaining a waiver is even harder with hybrid OA. Elsevier’s policy reads, “If an author would like their article to be published open access, but cannot afford these fees, then individual waiver requests are considered on a case-by-case basis and may be granted in cases of genuine need”.[6] Even with waivers, researchers in the developing world do not want handouts, but rather to be equal participants in the scholarly communication system. In a letter to *Na-*

ture, Raghavendra Gadagkar writes, “Page charges may be waived for authors who cannot afford to pay, but a model that depends on payment by authors can afford only a few such waivers. And why should anyone want to survive on charity?”[7] In addition, the costs for a researcher from a non-English speaking country may well be larger due to the need to translate his/her paper into English. Even when articles and books are written directly in English and other major languages (non-native), the costs for language checks may be very high (i.e., prohibitive).

Intricately linked to these problems is the current value system for research. Scientific quality and visibility can often be tied (by grant agencies and tenure boards in particular) to journal rankings, as determined by impact factors. This situation has effectively left research evaluation in the hands of international publishers and large scientific societies. Through their editorial policies, publishers decide what questions are important, and thus create a kind of “collective, flexible, largely unplanned, yet narrowly controlled, form of science policy for the developing world.”[8] Research published outside of these international journals, hence, is perceived to be of lesser quality. These journals have little incentive to change the current evaluation system, as it is directly related to their capacity to generate revenue.

As with all researchers, there are external pressures for researchers in the global periphery to be visible and therefore publish in international journals (in effect understood to be those published in the global north/center). However, serious consequences arise from this. For example, scientists from the global south must contribute to problems affecting mostly rich countries rather than their own in order to be published in these journals, and important scientific questions are being slighted or ignored because of the need to publish in prestigious journals.[9][16] This issue has been referred to as the 10/90 problem. The phenomenon in which 90% of the world’s R&D money is spent on the 10% of diseases that primarily affect people in developed countries, while only 10% is spent on diseases that mainly affect the 90% of people who live in the developing world. The problem

is described by Leslie Chan, Associate Director of Bioline International, in an interview with Richard Poynder. “[W]e still don’t have a good handle on malaria, on sleeping sickness and many other very common diseases that are found in the developing world,” says Chan. “And we don’t have a good handle on them because there just hasn’t been enough R&D money spent on them. They are neglected diseases.”[10]

In order to address these challenges, Latin American and African countries have sought to make local research more visible through local services like AJOL, Redalyc and SciELO. However, this could further isolate research in various regions. Research should be a global endeavour. Having two systems would surely be inefficient and perhaps counter-productive. It may also be worth noting that the SciELO model in Brazil is being challenged. On October 29, 2014, the “Coordination of Improvement of Personnel in Higher Education” (CAPES) announced that it will launch two tenders to “internationalize” 100 Brazilian journals through an agreement with a non-Brazilian publisher.[11] In protest to this move, the community of science editors in Brazil published an open letter directed to CAPES asking to be consulted and that the process be more transparent.[12]

It is essential that the decisions regarding a scholarly communication system which benefits global knowledge should be “determined through a genuinely global conversation to create scholarship shaped by academic rigour and quality, disciplinary frameworks and research imperatives, not determined by geographical borders, technical and other inequalities or commercial gains”. [15]

Participation by commercial publishers in scholarly publishing (with their need to generate revenue) will, inexorably, have implications for the global periphery in terms of their ability to fully participate in the system. For that reason, there are those who contend that scholarly publishing is not an area where for-profit companies should be involved. Others in this group, however, contend that the marketplace is better suited to assess demand and create sustainable solutions. The question for either approach, then, is how can we start to build institutional

capacity and mechanisms for regional and global collaboration so that we can capture, curate and share the knowledge that is being produced in our universities, and do so in a way that is inclusive and accessible?[13] After all, we have not yet done so despite over a decade of institutional repositories and OAI-PMH.

---

1. Bambini, Dominique: <http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/982/1856>
2. Björk and Solomon, *Developing an Effective Market for Open Access Article Processing Charges* (London: Wellcome Trust, 2014). <http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Open-access/Guides/WTP054773.htm>
3. A, Subbiah et.al., "Heads I Win, Tails You Lose: The Intransigence of STM Publishers," *Proc Indian Natn Sci Acad* 80:5 (December 2014): 919-929. [http://insa.nic.in/writereaddata/UploadedFiles/PINSA/Vol80\\_2014\\_5\\_Art03.pdf?](http://insa.nic.in/writereaddata/UploadedFiles/PINSA/Vol80_2014_5_Art03.pdf?)
4. <http://www.PLoS.org/publications/publication-fees/PLoS-publication-fee-assistance/> and <http://www.PLoS.org/PLoS-updates-fee-waiver-program/>
5. <http://www.the-scientist.com/?articles.view/article-No/23940/title/BioMedCentral-faces-angry-editors/>
6. <http://www.elsevier.com/about/policies/pricing-policy#apc-policies>
7. <http://www.nature.com/nature/journal/v453/n7194/full/453450c.html>
8. Vessuri, Guédon and Cetto, "Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development," *Current Sociology*, 2013. <http://eprints.rclis.org/23682>
9. Ibid.
10. Comments by Leslie Chan in interview with Richard Poynder. [http://poynder.blogspot.ca/2008\\_06\\_01\\_archive.html](http://poynder.blogspot.ca/2008_06_01_archive.html)
11. <http://blog.scienceopen.com/2014/11/brazilian-science-editors-open-letter>
12. Ibid.
13. Comments by Michelle Willmers in interview with Richard Poynder. <http://goo.gl/PL5TaZ>
14. E. Bonaccorso et al, "Bottlenecks in the Open-Access System: Voices from Around the Globe. *Journal of Librarianship and Scholarly Communication* 2:2 (2014):eP1126. <http://dx.doi.org/10.7710/2162-3309.1126>
15. Czerniewicz & Goodier, "Open Access in South Africa: A Case Study," *South African Journal of Science* 110:9-10 (2014): 97-105. <http://www.sajs.co.za/open-access-south-africa-case-study-and-reflections/laura-czerniewicz-sarah-goodier>
16. R. Hamman, "Balancing the academic terms of trade: The paradox of publishing in top-tier journals from the periphery," 2012 (unpublished). As referenced in Czerniewicz and Wiens, "African Research Visibility Online: The Poverty Alleviation Case," *African Journal of Information and Communication*, 13 (2013): 1-11.

Unless otherwise cited, all author quotes in this report were taken from the OSI online conversations held between October of 2014 and January of 2015. The exact dates and timestamps of particular quotes are available on request. Recommended citation: [Author], as quoted from the Open Science Initiative working group conversation, National Science Communication Institute, October 2014-January 15.

### OVERVIEW OF SCIENCE PUBLISHING

1. Ware and Mabe. *The STM Report: An Overview of Scientific and Scholarly Journal Publishing* (The Hauge: International Association of Scientific, Technical and Medical Publishers, 2012). [http://www.stm-assoc.org/2012\\_12\\_11\\_STM\\_Report\\_2012.pdf](http://www.stm-assoc.org/2012_12_11_STM_Report_2012.pdf)
2. "Free-for-all." *The Economist*. May 04, 2013. [http://www.economist.com/news/science-and-technology/21577035-open-access-scientific-publishing-gaining-ground-free-all?goback=.gde\\_27230\\_member\\_237917550](http://www.economist.com/news/science-and-technology/21577035-open-access-scientific-publishing-gaining-ground-free-all?goback=.gde_27230_member_237917550)
3. "Scientific Discourse: Buckling at the Seams." *Science*, October 4, 2013. <http://www.sciencemag.org/content/342/6154/56.short>
4. "The Future of Publishing." *Nature*, March 28, 2013. <http://www.nature.com/news/specials/scipublishing/index.html>
5. "Journals & Science Conference Wrap-up," National Science Communication Institute, Seattle, Washington, November 20, 2013. <http://nationalscience.org/news/nsci-news/2013/journals-science-conference-wrap-up/>
6. *Geographies of the World's Knowledge* (Oxford: Oxford Internet Institute and the Convoco Foundation, 2011). [http://www.academia.edu/2992610/Geographies\\_of\\_the\\_Worlds\\_Knowledge](http://www.academia.edu/2992610/Geographies_of_the_Worlds_Knowledge)
7. *Global Publishing: Changes in submission trends and the impact on scholarly publishers* (Thomson-Reuters, April 2012). [http://scholarone.com/media/pdf/GlobalPublishing\\_WP.pdf](http://scholarone.com/media/pdf/GlobalPublishing_WP.pdf)

8. "The true cost of science publishing," *Nature*. <http://www.nature.com/news/open-access-the-true-cost-of-science-publishing-1.12676>
9. M. Mabe, "The growth and number of journals," *Serials* 16:2 (2003), 191-197. <http://uksg.meta-press.com/link.asp?id=f195g8ak0eu21muh>
10. 2013 *Library Journal*/Periodical Price Survey. [http://lj.libraryjournal.com/2013/04/publishing/the-winds-of-change-periodicals-price-survey-2013/#\\_](http://lj.libraryjournal.com/2013/04/publishing/the-winds-of-change-periodicals-price-survey-2013/#_)

### PUBLIC FRUSTRATIONS WITH SCIENCE JOURNALS

1. "Kristof's good start," National Science Communication Institute. <http://nationalscience.org/blog/nsci-focus-areas/science-writing/2014/kristofs-good-start/>
2. "The Economist broadsides science," National Science Communication Institute. <http://nationalscience.org/blog/nsci-focus-areas/science-design/2013/the-economist-broadsides-science/>
3. "Jumping off the retractions bandwagon," National Science Communication Institute. <http://nationalscience.org/nsci-focus-areas/science-writing/2013/jumping-off-the-retractions-bandwagon/>

### THE CURRENT STATE OF OA

1. Liam Earney, "2013: The Year Open Access Broke," *Information Standards Quarterly*, Summer 2014. [http://www.niso.org/apps/group\\_public/download.php/13692/isqv26no2\\_crctd.pdf](http://www.niso.org/apps/group_public/download.php/13692/isqv26no2_crctd.pdf)
2. Hilda Bastain, "Open access 2013: A year of gaining momentum," *Scientific American*, December 26, 2013. <http://blogs.scientificamerican.com/absolutely-maybe/2013/12/26/open-access-2013-a-year-of-gaining-momentum/>
3. Steven Harnad, "Is OA growing? It depends," National Science Communication Institute (reprinted from an email from Steven to the Global Open Access list), December 6, 2013. <http://nationalscience.org/blog/nsci-focus-areas/science-writing/2013/is-oa-growing/>

4. Sue Gardner, "Is OA growing?," National Science Communication Institute, Sept 3, 2014. <http://nationalscience.org/blog/nsci-focus-areas/science-writing/2014/open-access-crossroads/>
5. Richard Poynder, "Robin Osborne on the state of Open Access: Where are we, what still needs to be done?," Open and Shut, December 23, 2013. <http://poynder.blogspot.co.uk/2013/12/robin-osborne-on-state-of-open-access.html>
6. Peter Suber's, "Ten challenges for open-access journals," SPARC, 2009. <http://www.sparc.arl.org/resources/articles/ten-challenges>
7. Peter Suber, "Open Access Overview." <http://legacy.earlham.edu/~peters/fos/overview.htm>
8. Peter Suber, *Open Access* (Cambridge: MIT Press, 2012). <http://bit.ly/oa-book>
9. Peter Suber, "Promoting Open Access in the Humanities." <http://legacy.earlham.edu/~peters/writing/apa.htm>
10. David Crotty, "Open Access in the Humanities and Social Sciences: An Interview with Chris Wickham," Scholarly Kitchen, April 27, 2014. <http://scholarlykitchen.sspnet.org/2014/04/17/open-access-in-the-humanities-and-social-sciences-an-interview-with-chris-wickham/>
11. "arXiv financial projections for 2013-2017," Cornell University Library, March 28, 2012. <https://confluence.cornell.edu/download/attachments/127116484/arXiv%20Business%20Model.pdf?version=2&modificationDate=1349107159000&api=v2>
12. Jørgen Burchardt, "Researchers Outside APC-Financed Open Access: Implications for Scholars Without a Paying Institution," Sage Open, October-December 2014. <http://sgo.sagepub.com/content/4/3/2158244014551714>
13. "The effect of open access and downloads ('hits') on citation impact: a bibliography of studies," OpCit Project, last updated June 25, 2013. <http://opcit.eprints.org/oacitation-biblio.html>
14. "The Open Access Citation Advantage," SPARC Europe. <http://sparceurope.org/oaca/>
15. Prudence Adler, "ARL Response to White House RFI on Public Access to Scholarly Publications," Association of Research Libraries, January 28, 2012. <http://www.arl.org/focus-areas/public-access-policies/federally-funded-research/2431-arl-response-to-white-house-rfi-on-public-access-to-scholarly-publications#.VlpOvcnddWy>
16. Surveys of faculty attitudes toward OA:
  1. J. Xia, "A longitudinal study of scholars attitudes and behaviors toward open-access journal publishing," *J. Am. Soc. Inf. Sci.*, 61 (2010): 615–624. <http://onlinelibrary.wiley.com/doi/10.1002/asi.21283/abstract>
  2. *Taylor & Francis Open Access Survey, June 2014* (Oxford: Taylor & Francis Group, 2014). <http://www.tandfonline.com/page/openaccess/opensurvey/2014>
  3. *Author Attitudes Towards Open Access Publishing* (Oxford: InTech, April 2011). <http://www.intechopen.com/authors-favor-open-access.html>
  4. *Learned Society attitudes towards Open Access* (Oxford: TBI Communications, May 20, 2014). [http://www.edp-open.org/images/stories/doc/EDP\\_Society\\_Survey\\_May\\_2014\\_FINAL.pdf](http://www.edp-open.org/images/stories/doc/EDP_Society_Survey_May_2014_FINAL.pdf)
  5. *Canadian Researchers' Publishing Attitudes and Behaviours* (Toronto: Phase 5 Research, March 7, 2014). <http://www.cdnsience-pub.com/learning-centre/impact-and-discovery/Researcher-Survey-Results.aspx>
  6. Daniëlle van Geresteyn, "What scholars think about open access: An overview of the developments in scholars' awareness and attitude towards open access publishing between 1992 and 2014" (MA thesis., Leiden University, October 9, 2014).

<https://openaccess.leidenuniv.nl/handle/1887/29761>

7. DJ Solomon, "A survey of authors publishing in four megajournals," *PeerJ* 2 (2014): e365. <https://peerj.com/articles/365/>
8. A variety of studies by Diane Harley on faculty attitudes and behaviors: <http://www.cshe.berkeley.edu/diane-harley>

### JOURNAL PUBLISHING BEST PRACTICES

1. Claire Redhead, "Principles of Transparency and Best Practice in Scholarly Publishing," OASPA, December 19, 2013. <http://oaspa.org/principles-of-transparency-and-best-practice-in-scholarly-publishing/>
2. *Trustworth Repositories Audit & Certification: Criteria and Checklist* (Chicago: Center for Research Libraries, 2007). [http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf)

### COPYRIGHT

1. Richard Poynder, "The Open Access Interviews: Paul Royster, Coordinator of Scholarly Communications, University of Nebraska-Lincoln," Open and Shut, August 31, 2014. <http://poynder.blogspot.co.uk/2014/08/the-open-access-interviews-paul-royster.html>
2. John Wilbanks, "Licence restrictions: A fool's errand," *Nature*, March 28, 2013. <http://www.nature.com/nature/journal/v495/n7442/full/495440a.html>

### PEER REVIEW

1. C.G. Begley and L.M. Ellis, "Drug development: Raise standards for preclinical cancer research," *Nature*, March 29, 2012. <http://www.nature.com/nature/journal/v483/n7391/full/483531a.html>
2. Diane Harley and Sophia Krzys Acord, "Peer Review in Academic Promotion and Publishing: Its Meaning, Locus, and Future," UC Berkeley: Center for Studies in Higher Education, 2011.

<http://www.cshe.berkeley.edu/publications/peer-review-academic-promotion-and-publishing-its-meaning-locus-and-future>

### PRICING

1. Richard Poynder, "Open access: What price affordability?," *eCancer*, 2014. <http://ecancer.org/journal/editorial/41-open-access-what-price-affordability.php>
2. Danny Kingsley, "Managing Article Processing Charges," Australian Open Access Support Group, updated July 30, 2014. <http://aoasg.org.au/managing-article-processing-charges/>

### EMBARGOES

1. Shan Sutton, "Open access, publisher embargoes, and the voluntary nature of scholarship," Association of College and Research Libraries, October 2013. <http://crln.acrl.org/content/74/9/468.full>

### IMPACT FACTORS

1. J. Bollen et al, "A Principal Component Analysis of 39 Scientific Impact Measures," *PLoS One*, June 29, 2009. <http://journals.PLoS.org/PLoSone/article?id=10.1371/journal.pone.0006022>
2. G. Lozano et al, "The weakening relationship between the Impact Factor and papers' citations in the digital age," *arXiv*, May 19, 2012. <http://arXiv.org/ftp/arXiv/papers/1205/1205.4328.pdf>
3. Jeroen Bosman, "Nine reasons why Impact Factors fail and using them may be harmful to science," March 11, 2013. <https://im2punt0.wordpress.com/2013/11/03/nine-reasons-why-impact-factors-fail-and-using-them-may-harm-science/>
4. <http://altmetrics.org/manifesto/>

### TENURE REFORM

1. "Guidelines for Evaluating Work in Digital Humanities and Digital Media," *Modern Language*

Association, last revised January 2012. [http://www.mla.org/guidelines\\_evaluation\\_digital](http://www.mla.org/guidelines_evaluation_digital)

## NOVEL IDEAS

1. Richard van Noorden, "Open-access deal for particle physics," *Nature*, September 24, 2012. <http://www.nature.com/news/open-access-deal-for-particle-physics-1.11468>
2. Elizabeth Wager, "Publishing Clinical Trial Results: The Future Beckons," *PLoS One*, October 27, 2006. <http://www.PLoSclinicaltrials.org/article/fetchArticle.action?articleURI=info:doi/10.1371/journal.pctr.0010031>
3. Kennison and Norberg, *A Scalable and Sustainable Approach to Open Access Publishing and Archiving for Humanities and Social Sciences* (New York: K|N Consultants, 2014). <http://knconsultants.org/toward-a-sustainable-approach-to-open-access-publishing-and-archiving/>
4. "Not-so-modest Proposals: What Do We Want Our System of Scholarly communication to Look Like in 2010?" (paper presented at the CIC Summit on Scholarly Communications, Chicago, Dec. 2, 2003). <http://people.brandeis.edu/~unsworth/CICsummit.htm>
5. Herbert Van de Sompel et al., "Rethinking Scholarly Communication: Building the System that Scholars Deserve," *D-Lib Magazine*, September 2004. [www.dlib.org/dlib/september04/vandesompel/09vandesompel.html](http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html)
6. "Peer Review as a Service: It's not about the journal," May 9, 2014. <http://theojo.org>

## MANDATES

1. Peter Suber, section entitled "Digression on the Word 'Mandate,'" in Peter Suber, *Open Access* (Cambridge: MIT Press, 2012). <http://bit.ly/oa-book>: <http://goo.gl/19dHyW>
2. Steven Harnad, "Regulating the BioRxiv Stampede," National Science Communication Institute (reprinted from an email from Steven

to the Global Open Access list), November 13, 2013. <http://nationalscience.org/blog/nsci-focus-areas/science-collaboration/2013/regulating-the-biorxiv-stampede/>

3. J. Xia et al, "A Review of Open Access Self-Archiving Mandate Policies," *Libraries and the Academy* 12 (2012): 85–102. [https://www.press.jhu.edu/journals/portal\\_libraries\\_and\\_the\\_academy/portal\\_pre\\_print/current/articles/12.1xia.pdf](https://www.press.jhu.edu/journals/portal_libraries_and_the_academy/portal_pre_print/current/articles/12.1xia.pdf)
4. David Wojick, "US Federal Open Access Rule-makings To Come," *Scholarly Kitchen*, October 21, 2013. <http://scholarlykitchen.sspnet.org/2013/10/21/us-federal-open-access-rule-makings-to-come/>
5. Stuart Shieber and Peter Suber, "Good practices for university open-access policies," Harvard Open Access Project, last modified January 22, 2015. <http://bit.ly/goodoa>

## ECONOMIC BENEFITS OF OPEN INFORMATION

1. *Open data: Unlocking innovation and performance with liquid information* (New York: McKinsey & Company, October 2013). [http://www.mckinsey.com/insights/business\\_technology/open\\_data\\_unlocking\\_innovation\\_and\\_performance\\_with\\_liquid\\_information?cid=other-eml-alt-mgi-mck-oth-2910/](http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information?cid=other-eml-alt-mgi-mck-oth-2910/)
2. *Market assessment of public sector information* (London: Deloitte, May 2013). [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/198905/bis-13-743-market-assessment-of-public-sector-information.pdf)
3. New research in-progress focusing specifically on impacts of open access: [http://www.nsf.gov/awardsearch/showAward?AWD\\_ID=1360205&HistoricalAward](http://www.nsf.gov/awardsearch/showAward?AWD_ID=1360205&HistoricalAward)

## THE FUTURE OF OPEN SCIENCE

1. "Recent advances in massive, online, open science research," National Science Communication Institute, Feb 24, 2014. <http://nationalscience.org>

org/blog/uncategorized/2014/recent-advances-massive-open-online-science-research/

2. Another good example of open science in action—not what-if stuff: <http://sagebase.org/>
3. “University of Chicago Launches Bionimbus Protected Data Cloud to Analyze Cancer Data Using the cloud to access science data,” NSF, May 15, 2013. [http://www.nsf.gov/news/news\\_summ.jsp?cntn\\_id=127935&WT.mc\\_id=USNSF\\_51&WT.mc\\_ev=click](http://www.nsf.gov/news/news_summ.jsp?cntn_id=127935&WT.mc_id=USNSF_51&WT.mc_ev=click)
4. L. Bettencourt et al, “Report for the Office of Scientific and Technical Information: Population Modeling of the Emergence and Development of Scientific Fields,” SciTech Connect, October 4, 2006. <http://www.osti.gov/scitech/servlets/purl/990671>
5. *Nature* now publishing datasets: <http://www.nature.com/sdata/>
6. Bowser and Shanley, *New Visions in Citizen Science* (Washington, DC: Woodrow Wilson Center, 2013). <http://www.wilsoncenter.org/sites/default/files/NewVisionsInCitizenScience.pdf>
7. Peter Murray Rust, “Unlocking chemistry: it’s time to make the subject as open as bioscience,” *The Guardian*, January 6, 2014. <http://www.theguardian.com/higher-education-network/blog/2014/jan/06/open-chemistry-patents-research>
8. Steven Harnad, “Will Google start dabbling in science journals” National Science Communication Institute (reprinted from an email from Steven to the Global Open Access list), August 22, 2014. <http://nationalscience.org/nsci-focus-areas/science-writing/2014/will-google-start-dabbling-science-journals/>
9. *2014 Global R&D Funding Forecast* (Seattle: Battelle, December 2013). <http://nationalscience.org/nsci-focus-areas/science-policy/2013/battelle-2014-report/>
10. “List of countries by research and development

spending,” Wikipedia. [http://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_research\\_and\\_development\\_spending](http://en.wikipedia.org/wiki/List_of_countries_by_research_and_development_spending)

11. R. Celeste et al, *Furthering America’s Research Enterprise* (Washington, DC: National Academies Press, 2014). <http://www.nap.edu/catalog/18804/furthering-americas-research-enterprise>
12. M. Nielsen, *Reinventing Discovery: The New Era of Networked Science* (Princeton: Princeton University Press, 2011). <http://www.amazon.com/gp/product/B005OQGZ54>

## ISSUES AFFECTING THE GLOBAL SOUTH

1. UNESCO Global Open Access Portal (GOAP), snapshot of OA in each country: <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/access-by-region/>
2. Barbara Kirsop, Leslie Chan and Subbiah Arunachalam, “Access to Scientific Knowledge for Sustainable Development: Options for Developing Countries,” *Adriane*, July 29, 2007. <http://www.ariadne.ac.uk/issue52/kirsop-et-al>
3. E. Bonaccorso et al, “Bottlenecks in the Open Access System: Voices from Around the Globe,” *Journal of Librarianship and Scholarly Communication* 2 (2014): eP1126. <http://dx.doi.org/10.7710/2162-3309.1126>
4. Richard Poynder Interviews:
  1. Richard Poynder, “Michelle Willmers on the state of Open Access: Where are we, what still needs to be done?,” *Open and Shut*, October 30, 2013. <http://poynder.blogspot.co.uk/2013/10/michelle-willmers-on-state-of-open.html>
  2. Richard Poynder, “Open Access in India: Q&A with Subbiah Arunachalam” *Open and Shut*, June 9, 2014. <http://poynder.blogspot.co.uk/2014/06/open-access-in-india-q-with-subbiah.html>

3. Richard Poynder, "Dominique Babini on the state of Open Access: Where are we, what still needs to be done?," *Open and Shut*, July 25, 2013. <http://poynder.blogspot.co.uk/2013/07/dominique-babini-on-state-of-open.html>.
5. Leslie Chan and Eve Gray, "Centering the Knowledge Peripheries through Open Access: Implications for Future Research and Discourse on Knowledge for Development," in *Open Development: Networked Innovations in International Development* (Ottawa: IDRC, 2013). [https://tspace.library.utoronto.ca/bitstream/1807/44099/1/Chan%26Gray\\_Open%20Development.pdf](https://tspace.library.utoronto.ca/bitstream/1807/44099/1/Chan%26Gray_Open%20Development.pdf)
6. Natalia T. Bowdoin, "Cultures of Access: Differences in Rhetoric Around Open Access Repositories in Africa and the United States and Their Implications for the Open Access Movement," University of South Carolina, 2013. [http://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1005&context=aiken\\_gregg\\_graniteville\\_library\\_facpub](http://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1005&context=aiken_gregg_graniteville_library_facpub)
7. Xiaolin Zhang, "Development of open access in China: strategies, practices, challenges," *Insights* 27(1), March 2014. <http://uksg.metapress.com/content/p4t2157824453387/fulltext.pdf>
8. Laura Czerniewicz, "Inequitable power dynamics of global knowledge production and exchange must be confronted head on," London School of Economics and Political Science, April 29, 2013. <http://blogs.lse.ac.uk/impactofsocialsciences/2013/04/29/redrawing-the-map-from-access-to-participation/>
9. Jorge L. Contreras, "Open Access Scientific Publishing and the Developing World," *SSRN*, May 24, 2012. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2065887](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2065887)
10. Leslie Chan, Barbara Kirsop and Subbiah Arunachalam, "Open Access Archiving: the fast track to building research capacity in developing countries," *SciDev Network*, Nov 2005. [https://tspace.library.utoronto.ca/bitstream/1807/4415/1/Open\\_Access\\_Archiving.pdf](https://tspace.library.utoronto.ca/bitstream/1807/4415/1/Open_Access_Archiving.pdf)
11. C.P. Chandrasekhar, "Open Access vs academic power," *Frontline*, October 4, 2013. [http://www.frontline.in/columns/C\\_P\\_Chandrasekhar/open-access-vs-academic-power/article5134150.ece](http://www.frontline.in/columns/C_P_Chandrasekhar/open-access-vs-academic-power/article5134150.ece)
12. Leslie Chan, Barbara Kirsop, and Subbiah Arunachalam, "Towards open and equitable access to research and knowledge for development," *PLoS Medicine*, March 29, 2011. <http://journals.PLoS.org/PLoSmedicine/article?id=10.1371/journal.pmed.1001016>
13. Laura Czerniewicz, Sarah Goodier, "Open access in South Africa: A case study and reflections," *South African Journal of Science* 110, Sept/Oct 2014. [http://www.sajs.co.za/sites/default/files/publications/pdf/Czerniewicz\\_Research%20Article\\_0.pdf](http://www.sajs.co.za/sites/default/files/publications/pdf/Czerniewicz_Research%20Article_0.pdf)
14. Eve Gray, "Supersize it!," *Research Professional*, January 9, 2014. [http://www.researchresearch.com/index.php?option=com\\_news&template=rr\\_2col&view=article&articleId=1346155](http://www.researchresearch.com/index.php?option=com_news&template=rr_2col&view=article&articleId=1346155)
15. Gajendra Sharma, "Transforming Open Access Scholarly Publishing and Scientific Delivery: Challenges and Opportunities in Asian Regions," *Research Journal of Information Technology*, 2014. <http://docsdrive.com/pdfs/academicjournals/rjit/0000/67086-67086.pdf>
16. H. Vessuri, J.-C. Guedon, and A. M. Cetto, "Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development," *Current Sociology*, 2013. <http://eprints.rclis.org/23682/>
17. *Report of the Regional Latin American and Caribbean Consultation on Open Access to Scientific Information and Research* (Paris: UNESCO, March 2013). [http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/news/report\\_open\\_access\\_en.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/news/report_open_access_en.pdf)

18. Juan Pablo Alperin and Gustavo Fischman, "Quality in the Open Scholarly Communication of Latin America," IDRC, October 1, 2012. <http://flacso.org.br/oa/wp-content/uploads/2013/11/first.interim.report.20120928.pdf>
19. Dominique Babini, "The risk of open access becoming integrated into existing commercial publishing - the need of a global system of non-commercial open access scholarly communications," *RECIIS* 8(4), 2014. <http://www.reciis.icict.fiocruz.br/index.php/reciis/article/viewFile/982/1856>
20. Jean-Claude Guédon, "Open Access and the divide between 'mainstream' and 'peripheral' science," 2008. <http://eprints.rclis.org/10778/>
21. Susan Murray, Abby Clobridge, "Scholarly Publishing in Africa," October 2013. [http://pubs-for-dev.info/wp-content/uploads/2013/10/PfD-2013\\_05\\_Abby-Clobridge-and-Susan-Murray.pdf](http://pubs-for-dev.info/wp-content/uploads/2013/10/PfD-2013_05_Abby-Clobridge-and-Susan-Murray.pdf)
22. Juan Pablo Alperin, Dominique Babini and Gustavo Fischman (eds.), *Open Access indicators and scholarly communications in Latin America* (Buenos Aires: CLACSO, August 2014). <http://biblioteca.clacso.edu.ar/clacso/se/20140917054406/OpenAccess.pdf>
23. Scholarly communication in Africa Programme (various resouces): [http://www.scaprogramme.org.za/media\\_presentations/](http://www.scaprogramme.org.za/media_presentations/)
24. Mohamed Hanif Nashipudi and B. Ravi, "Indian research going global: A study on the status of open access publishing," *International Journal of Information Research*, 2014. <http://www.spoars.org/journal/v3n4p4>

The 112 participants who signed up to participate in this conversation were drawn mostly from the academic, research, and library communities.

Most of these 112 were not active in this conversation, but a healthy diversity of key perspectives was still represented. Individual participants may not agree with all of the facts and viewpoints described herein, but participants agree that this document reflects the spirit and content of the conversation. Those who chose to sign this document (listed in order of signing) are:

Joe Kraus  
 Ada Emmett  
 Rick Anderson  
 Paul Groth  
 Colleen Lyon  
 Wm. Joseph Thomas  
 Glenn Hampson  
 April Hathcock  
 William Gunn  
 Nina Collins  
 Denise Troll Covey  
 Eric F. Van de Velde  
 David Wojick  
 K. Jane Burpee  
 Maria Bonn  
 Debbie Chaves  
 Joyce Ogburn  
 Richard Poynder  
 Grisell Rodriguez  
 Barbara Harvey  
 Eric Olson  
 Joann C. Delenick  
 Joshua Rosenbloom  
 Dominique Babini

**And 90 others (some participants cannot list their names due to conflict of interest with their employers)**

