Open Scholarship Initiative Proceedings, Volume 1, 2016
doi: http://dx.doi.org/10.13021/G8CS33

Report from the What is Publishing? (2) Workgroup

Melanie Dolechek, Sharon E. Farb, Steven Hall, Harriette Hemmasi, John Inglis, Lisa Macklin, Ivan Oransky, Mark Parsons, Nancy Rodnan, Andrew Tein

Abstract / OSI2016 Workgroup Question

What do we mean by publishing in today's world? What should be the goals of scholarly publishing? What are the ideals to which scholarly publishing should aspire? What roles might scholarly publishers have in the future? What scenarios exist where publishers continue to play a vital role but information moves more freely? What impact might these reforms have on the health of publishers? Scholarly societies? Science research? Why?

Introduction

At the inaugural Open Scholarship Initiative (OSI) in 2016, held in Fairfax, VA, April 19-22, our workgroup discussed the question "What is Publishing?" This larger question included additional aspects for us to consider (see Abstract/Workgroup Question, above). After two days of discussion, review, and refinement of ideas, we have endeavored to point a way forward that understands and recognizes our past and our current state of publishing.

Publishing Today

Henry Oldenburg, who founded Philosophical Transactions in 1665, defined the four functions of scientific publishing as registration, certification, dissemination, and archiving.

Registration (establishing who first made a discovery) and certification (review of the findings by one's peers) remain fundamental functions of scholarly publishing today. In fast-moving areas of science, in which competition for grants is fierce, days matter, and the pressure on publishers to publish quickly, to establish a researcher's precedence, is intense. In certification, we are seeing many experiments in post-publication review and commentary, but pre-publication peer review remains the norm in the vast majority of disciplines.

Dissemination remains a critical function. In a digital environment, discovery requires high-quality, reliable metadata and new tools and services such as CrossRef, Open Funder Registry, XML-coding, semantic indexing, and indexing by search services. Archiving is a cooperative undertaking, between publishers, libraries, and third-party services.

Oldenburg might not recognize many of the digital elements of today's scholarly publishing, and would likely be astonished at the scale—two million published journal articles per year, compared with fewer than 130 in the first volume of *Philosophical Transactions*—but he would probably still recognize the two principal forms of output of scholarly research, the journal article and the monograph. In each case, publication is essentially defined by a single final output, which is largely still text-based and consumed in printed, or PDF, form.

It is also worth considering the extent to which open access has changed the fundamentals of journals publishing. The most recent study of the global growth in open access publication, *Monitoring the Transition to Open Access*, published in August 2015 by the Research Information Network for the Universities UK Open Access Co-ordination Group, found that 16.6 percent of all articles in peer-reviewed journals were published under the gold open access model in 2014.[1] That figure is likely to have grown to 18 percent or more in 2015. With green open access mandates also affecting many articles published on a subscription basis, and with the growth of subject and institutional repositories, dissemination has certainly changed: primarily digital, multi-channel, and multi-format, with preprints available alongside accepted manuscripts alongside the final published versions. However, open access has not changed the two core functions of registration and certification. Traditional and open access publishers alike develop and curate journals and monographs for specific scholarly communities, manage peer review, produce and index articles/monographs and their metadata, and distribute them digitally. While open access has seen the launch of a small number of so-called 'mega-journals,' many researchers choose to publish in specialist journals which directly address their communities, and driven by systems of academic recognition and reward, prefer to publish with the strongest brand that will accept their submission.

While journal publishing is now largely a digital, and digital first, enterprise, even if the final output would be recognized by Oldenburg, monograph publishing is still to a large extent a 'print first, digital second' undertaking, with only a few publishers experimenting with open access models. All too often the digital book is little more than a PDF of the print edition, with the occasional ancillary supplement of data and multimedia. We are beginning to see some publishers, especially in the sciences, producing digital monographs which take advantage of the potential of the format—full html, embedded multimedia, interactive charting, and so on—and some university presses are experimenting with open access and digital monographs, but these experiments should be conducted far more widely, and across more areas in the humanities and social sciences.

Publishing in the Future

Scholarly research is an increasingly diverse, complex, and interdisciplinary endeavor, with growing importance assigned to incremental stages of evidence and argument. We therefore envision a future publishing paradigm that is networked, open, and significantly more dynamic than the traditional model. Emerging signs point to a shift from the current "event-driven" model that focuses primarily on the publication of a print (or PDF) article or monograph to an ongoing process-driven, digital model that reflects

more of the research lifecycle through a scholarly record that is comprised of much more than text. The boundaries of what constitutes a book or article are becoming blurred, and we are increasingly recognizing the value of other content. Over time, we anticipate that publishing will regularly encompass a richer and more interconnected range of scholarly content, including data, lab and field notes, software, preprints, social media posts, video—performative and experiential—and multimedia, with possibilities for interactivity among the various elements. We will likely see additional forms of content and new methodologies introduced through the use of virtual reality, gamification, and other innovative technologies—all of which are intended to help construct and enhance meaning and have the potential to make scholarship more relevant to society. By sharing, collectively assessing, and interconnecting this diversity of scholarly output, the research community will recognize these results as valued, first-class objects of research. The ability to continually access, assess, and interact with these many assets will be key markers of successful research as well as successful publishing.

As more of the research lifecycle is shared, we anticipate that researchers will experiment widely with formal and informal channels of disseminating their scholarly results. Through this experimentation, researchers and others will learn what is most valued and needed from the established publishing process versus what is best suited for alternative methods. In contrast to today's practices, tomorrow's myriad scholarly artifacts are likely to be released incrementally through diverse global systems. They may represent smaller (or larger) unbundled components, distributed at a much more timely pace with varying degrees of openness. Linkages among the discrete parts would be established in standardized ways to retain, expose, and build on the inherent intellectual relationships. Assessment and certification of all elements would be considered essential, and would go well beyond today's more limited approach to peer-review. Disciplinary experts and data curators, for example, would be collaboratively assess data in order to ensure methods are appropriate, uncertainties are well described, documentation is complete, and standards are followed. Software assessment would include expert reviews as well as test-driven methods such as continuous integration. There would be mechanisms for ongoing community assessment of all published artifacts, and the major assessment criteria (beyond those of originality and research merit) would be reproducibility and reusability. The published "book" or "article" would not be a final event, but part of an ongoing scholarly conversation.

This model assumes a general trend toward openness, with restrictions based on ethical norms rather than proprietary considerations, as they are currently defined. Given the differences in disciplines, we would expect that the timing of openness might vary widely among different communities. At present, humanists are much less inclined to share their in-progress work, whereas many scientists are accustomed to working in teams and distributing repetitive drafts for quick review and feedback. In addition, there may be variations in the degree of openness with which scholarship is released. For example, a virtual lab notebook may initially be shared only among a small team of researchers and curators and later released to others in the research community. Even then, the entire notebook

might not be fully open, but instead, be released selectively, on as-needed basis, to provide evidence or provenance for another publication. The timing of release may also be dictated by intellectual property concerns, such as patent application filing.

This dynamic and diverse publishing ecosystem will involve changing roles for today's key partners in the world of publishing, including researchers, publishers, universities, and funders. In addition, new partners will be required if we are to successfully deliver, disseminate, and preserve future research. Data curators, data scientists, software developers, and designers would work with researchers, librarians, editors, and publishers to develop a network of interconnected scholarly work supported by a variety of institutions including: data repositories, research labs, libraries, publishers, network providers, standards organizations, software providers, and professional societies from public, private, as well as nonprofit sectors. These new, shared roles and the resulting scholarly output will require different types of reward and recognition that go well beyond the traditional measures of scholarly impact.

Finally, it is important to note that the future of publishing is likely to be a mix of today's formal and informal methods of scholarly dissemination. Questions remain whether these two approaches will continue to develop in parallel; become combined, complementary systems that incorporate the best of both models; or evolve into entirely new alternatives. Regardless, it seems clear that the open, global network will increasingly be the primary means and method of distributing scholarly communication. The network will also serve as the dominant forum in which scholars interact with their peers, and will increasingly become the preeminent platform for building audiences (and collaborators) that extend scholarship beyond academia.

How to Get There

Conversations about where publishing is today, and what it might become in the future, inevitably turn to the purpose of scholarly works. At present, these outputs form a large part of how scholars and institutions are evaluated. That means any significant change will require thought and action on the part of those who conduct this evaluation—funding agencies, tenure committees, scholarly societies, accrediting bodies, and others.

While no one solution or player will effect systemic change, each can contribute with its own experiments. Instead of rewarding only the peer-reviewed and published paper, evaluators may assign value to post-publication contributions by others, data sharing, collaboration, and other important steps in the scientific process. In short, we want evaluation metrics that reflect how we want science to work.

Funders of scholarly activity, particularly scientific research, are eager to support innovative work that pushes the boundaries of basic knowledge and maximizes translational possibilities. In the past decade, initiatives such as NIH's Pioneer Awards, the HHMI Investigator Program, Wellcome Investigator Awards, and the joint HHMI/Gates/Simons Faculty Scholar Program have encouraged the unfettered imagination of early career investigators, the spark of cross-disciplinary collaboration, or the risk-taking of established scientists willing to abandon the familiar for untried approaches. Dur-

ing the same period, compelled to evaluate the impact of the supported work on scholars and the general public, funders discovered systems of publication out of step with their goals, characterized by restricted access to output, restrictions on redistribution, peer review ill-disposed to novelty, and community journals closed to cross-disciplinary work. And they found few means of measuring the consequences of their funding decisions. Their responses, from open access mandates and the funding of new breeds of journals, to the instigation of new metrics, are at the heart of the present turmoil in scholarly communication and their voices will surely continue to be heard.

Collecting, evaluating, and disseminating a wide range of research artifacts will affect scholars, librarians, and publishers in a variety of ways. Scholars will have to adapt to changing expectations from funders regarding distribution of scholarship resulting from the funding. If universities change promotion and tenure evaluation requirements, scholars at those universities will shift their publishing and dissemination practices accordingly. It is logical that all stakeholders will focus their efforts on what is rewarded. As publishing becomes a more continual process, as described above, scholars will be expected to follow standards and provide complete documentation, including ensuring that uncertainties are well described for data sets and other works that are not currently a routine part of published research results. Documenting and communicating the related outputs of the research process will also fall to scholars, just as we have citation norms in the current system.

Librarians have already expanded their role from assisting scholars with the research process, to assisting scholars with research outputs. For example, librarians help scholars comply with funder requirements, such as the NIH Public Access Policy, and this role will expand with new funder requirements. Librarians also continue to play a role in assisting scholars organize, apply metadata, and preserve their research outputs, such as data sets. The collection of these materials will need to be standardized. The growth of digital humanities centers in libraries has brought new tools to bear on the research process, with librarians serving as guides to scholars on applying the most appropriate technology to their research question. On a broad scale, services like the Research Center of HathiTrust, which provides text and data mining across the corpus of works in the HathiTrust digital archive, will continue to leverage library collections and librarian expertise to provide new research opportunities for scholars. These new roles will require new skills and new ways of collaborating with scholars and third parties at other universities or in the private sector.

Publishers also assist authors with meeting funder requirements, such as depositing manuscripts in PubMed Central. Publishers continue to play a vital role in promoting standards and best practices that facilitate the ecosystem of available research results. In addition, publishers will need to establish new products and services, partner with new service providers, and invest in talent and training as the industry shifts away from the manuscript-centered, broadcast approach and toward an approach that reflects the way users both consume and interact with Web content and engage with other users. All of this will require visionary proactivity, continuous user monitoring and analysis, and swift reactions when the unexpected inevitably emerges.

The new, dynamic publishing ecosystem we envision will require new types of business models and expanded partnerships and collaborations to succeed and be sustainable. It is important to note that this new ecosystem relies on a complex infrastructure of many different organizations, technologies, and professionals. When infrastructure works well it is transparent and often taken for granted. As such, it is often difficult to garnish the attention and funding necessary for ongoing support. As we see with the evolution of traditional infrastructures like the power grid, many different funding models will need to work together across multiple scales in a decentralized network. Governments will continue to play a foundational role in establishing appropriate regulatory frameworks and support of public goods, but private funders, universities, commercial and nonprofit labs, technology companies, scholarly societies, libraries, and publishers will also play critical roles. Governments and institutions need to take a lead in establishing new public-private partnerships, private finance initiatives, and local-level partnerships that coordinate and sustain necessary publishing services. Again, specific approaches are likely to vary across disciplines.

Ensuring that the needs of the researcher remain at the heart of these multi-stakeholder partnerships will be fundamental to the success of this future vision of publishing. Researchers are increasingly confronting time and resource challenges that distract from the core mission of discovery and scientific advancement. In response, governments, institutions, and the private sector must work jointly to establish a friction-free environment for researchers to share scholarship. This includes both infrastructure and policy elements. Infrastructure solutions require a seamless user experience based on platforms and tools that are interoperable and in constant communication. Policy solutions require coordinated efforts that will drive and incentivize behavioral change, including short-term efforts to recognize and measure individual contributions in the research process and the longer-term systemic changes to academic evaluation mentioned previously. The future of publishing is enabling researchers to accelerate pre- and post-publication discovery and increasing the discoverability and impact of their scholarly work.

Recommendations for OSI

Our team identified several recommendations to inform future OSI efforts. The three primary recommendations focus on aspects of open scholarship and open access publishing that require more data, evidence, and discussion. There were several themes mentioned in many, if not most, of the teams including peer review, business models, and role and needs of academic authors that were not adequately addressed and would benefit from additional research and discussion. The recommendations below have been identified as priorities:

- identify existing studies and initiatives relevant to open scholarship, including systems of academic recognition and reward; and identify gaps in evidence and knowledge
- define unmet publishing and dissemination needs of scholars
- develop disciplinary approaches, use cases, and experiments or pilots rather than one-size-fits-all approaches

- invite additional stakeholders to next OSI meeting, including research communities, domain repositories, research software providers, and academic authors.

The OSI "What is Publishing (2)" Workgroup

- **Melanie Dolechek**, Executive Director, Society for Scholarly Publishing (SSP)
- **Sharon E. Farb**, Associate University Librarian for Collections and Scholarly Communication, UCLA
- **Steven Hall**, Managing Director, IOP Publishing
- **Harriette Hemmasi**, University Librarian, Brown University
- **John Inglis**, Executive Director and Publisher, Cold Spring Harbor Laboratory Press and Co-founder, bioRxiv
- **Lisa Macklin**, Director, Scholarly Communications Office, Emory University
- **Ivan Oransky**, Distinguished Writer In Residence, New York University Arthur Carter Journalism Institute, and Co-Founder, Retraction Watch
- **Mark A. Parsons**, Secretary General, Research Data Alliance
- **Nancy Rodnan**, Senior Director of Publications, American Society for Biochemistry and Molecular Biology (ASBMB)
- **Andrew Tein**, Vice President, Global Government Affairs, Wiley

Notes:

[1] "Monitoring the Transition to Open Access: A report for the Universities UK Open Access Co-ordination Group," Research Information Network, August 2015, as of May 25, 2016: http://www.researchinfonet.org/wp-content/uploads/2015/09/Full-report-FINAL-AS-PUBLISHED.pdf