Report from the "What is Publishing?" (1) Workgroup

Amy Brand, James Butcher, Meg Buzzi, Kathleen Fitzpatrick, Ann Gabriel, Rikk Mulligan, Vivian Siegel, Matt Spitzer, Jamie Vernon

Abstract

The mechanisms used for scholarly publishing have remained largely unchanged over time, even as we've moved from a print-based world to a digital world. The scholarly communication ecosystem, however, is now undergoing a period of rapid transformation, including the introduction of new actors, new services, and increased pressure to improve the means of scholarly communication in order to meet the growing expectations of an information-rich world. Where to begin? The first question is to ask how scholarly publishing can provide the greatest benefit to global society in a sustainable way. Our two-day conversation about this question led us to the conclusion that the "black box," monolithic model of scholarly publishing no longer serves most researchers. The most sustainable approach that best responds to the needs of authors and researchers today, and that may also pose the least amount of risk in completely disrupting the system, is disaggregated services—unbundling the products and services that publishers currently provide and letting market forces drive the development of, and demand for, a new and improved à la carte world of knowledge artifacts and knowledge management tools.

OSI2016 Workshop 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?

The last 350 years

In all aspects of academic scholarship, research builds on the work of others. Historically, scholars have shared their work in the form of scholarly journal articles and monographs—static "versions of record" that capture a snapshot in time of a research project, crafted into a story that seems worth reading and therefore worth telling.

In this context, print publishing has been the predominant method by which these stories are shared. More than simply printing (or, these days, posting) articles and monographs, publishing also includes filtering (peer review to determine whether a piece of scholarship is "publishable"), editing (cleaning up the language so that other people can understand the work), making the work "discoverable" (getting it into libraries, indexes, newspaper articles, and so

on), and the registering and quality-stamping of ideas (establishing priority for the authors, which can lead to rewards like promotion and tenure, or funding, in particular in light of the hard-won reputation of a publisher or publication for quality discernment).

We use largely the same mechanisms for publishing today that we did not merely 30 years ago, but 350 years ago. Many of these mechanisms have moved online, but they have nonetheless retained their prior characteristics. This traditionalism and resistance to change gets in the way of genuine transformation. This is not to say that publishers as a group are not innovative. In fact, the academic publishing community as a whole has collaborated on several transformative infrastructure initiatives like Crossref,[1] ORCID,[2] and CRediT.[3] And not all publishers are created equal: large commercial publishers, university presses, societies, and other non-profit players each have different missions and offer different value propositions to their customers, readers and authors.

Still, publishing as we've known it for the last several centuries has been something of a "black box" from the perspective of both author and reader: The raw manuscript goes in and the published work comes out, while the mechanism of this transformation remains largely hidden from view.

While aspects of this process have been updated for the digital age—including mechanisms for online submission, peer review management, editing, typesetting, and dissemination—the fundamentals remain the same: Content of a certain kind goes in one end, and publications of a few traditional categories emerge. If we are to

effect real, transformational, sustainable change, we are going to have to break open the black box.

None of these ideas are new, of course. Theodore "Robbie" Fox, who was editor of *The Lancet* from 1944-1964 predicted on his retirement that "A day will come when journals will be superseded as a means of publishing new research."[4] However, despite the disruptive technologies of the internet, and the vast data stores and APIs it makes possible, we have not yet wholly changed our vision of what publishing itself can and should be.

Framing the question

At the outset we formulated the following question to guide to our discussions about desired change: "How can scholarly publishing provide the greatest benefit to global society in a sustainable way?" We returned to this question throughout two days of discussion, and it helped remind us as to what the ultimate model should aspire.

We were wary of restricting our discussion to traditional forms of academic publishing (namely journals and monographs). Rather, we considered a future form of publishing that might be defined as follows: *A process that captures (or creates) and makes discoverable artifacts of knowledge in order to facilitate the use and reuse of scholarship on a global scale, and that enables research communities to build upon the work of others and provides a venue for evolving discourse.*

Part of our wariness was informed by the reality that scholarly publishing already extends well beyond journal articles and monographs, and includes blogs, white papers, information briefs, lab notes, conference presentations, videos, movies,

static and interactive datasets, information visualizations and more. Expanding this ecosystem even further, many of these products remain in flux, either temporary or deliberately: databases might be updated with more current information, or articles updated to include or connect to the most recent findings. Many products are "dark" (not visible to the public for whatever reason, including being unavailable electronically), while others exist in a variety of languages and formats, and so on.

Our approach was also informed by an awareness that, as the volume of information in our world continues to increase at a staggering rate, the current structure of our scholarly publishing system is not equipped to manage the flow. Unleashing this flow would entail both risks and benefits. We returned at several points to the analogy of scholarly publishing being like a hydroelectric dam: Research flows downstream, through many tributaries, and collects in a reservoir. The dam, a man made structure, restricts the flow of water (research), but in a way that allows power (return on investment) to be generated or crops (knowledge/education) irrigated. While many people within academia and scholarly publishing are rightly frustrated by the slow pace of change and the restrictions that the scholarly dam places on academic progress, destroying the dam would have negative consequences for everything that sits downstream and would also reduce the ability to generate power from the structure. Our conclusion was that we need to find a way of redirecting the flow, while maintaining as many benefits of the structure as possible.

In this regard, we felt that it is important to analyze the system as a whole and consider how individual changes might impact the scholarly endeavor. The scholarly black box is a highly complex system, with many moving parts and dependencies. As tempting as it may be to break it apart, a more considered approach may be advisable. After all, scholarly publishing has undoubtedly contributed to the ultimate goal of providing benefit to global society. Balancing the desire for rapid change, while also maintaining the benefits, will be difficult, but crucial.

The path forward

Given that scholarly publishing involves a rich ecosystem of products, and given that our current system of scholarly publishing is ill-equipped to capture the current and growing volume of inputs and outputs, but recognizing as well that the current system does provide important and established benefits to society and to the research community, the question for us came down to this: *How can scholarly publishing provide the greatest benefit to global society in a sustainable way?* Our guidepost, our ideal, was the future form of publishing described earlier— a process that captures (or creates) and makes discoverable artifacts of knowledge in order to facilitate the use and reuse of scholarship on a global scale.

We agreed that to find the answer to this question would first require:

1. **A systems analysis**, looking at how individual changes might impact the publishing system. The current reality of publishing might be characterized as a complex network of actors and processes that excel at creating discrete research products but not at working together to facilitate research and scholarship on a global scale. What happens when we start changing different parts of this system?

2. **Identifying the obstacles** to change (such as infrastructure limitations, attribution practices, tenure and promotion systems, business models, peer review, and the pace of information processing).

3. **Clearly specifying the reasons** for change (such as providing access to everyone who needs it, providing choices to authors, and encouraging more innovation in the publishing business model).

4. **Being prepared to peel away generations of expectations** surrounding our culture and practices of research communication. What would happen if there were no longer a "version of record" for a research breakthrough, for instance, and documents could "breathe" and change as new information comes along? What if more sharing happened throughout the research process (as occurred with the human genome project, for example), enabling global collaboration and rapid correction and translation? Are we comfortable with the current pace of change in publishing? Our group acknowledged that, while considerable efforts have been made to transform the scholarly record through content innovation and new metadata standards, we need faster change but in a way that is practical and sustainable.

Our group focused mainly on the change that could happen at point 1, not for a lack of interest about the other points but simply for lack of time. Still, we were encouraged by how much improvement might be possible by making a few changes at this system level. Specifically, we concluded that the simplest path forward might begin by viewing publishing not as a monolithic enterprise, but as a bundle of inputs, outputs, processes, partners, standards, products, services, and so on. The next step would to determine which of these facets should be unbundled. The primary goal of an unbundled world of publishing services would be to meet individual researchers "where they are" in terms of their preferred modes of authoring and sharing, and to allow those who contribute knowledge artifacts to have access to a wider variety of the à la carte bits and pieces they need to advance discovery—more knowledge, and more robust services (like editing, peer review, dissemination, storage, processing, data standardization and integration, analysis, and so on).

What would happen if, by breaking down publishing services into "app-sized" pieces, new players could easily enter the market, creating publishing solutions that current publishing behemoths can't provide? Can this happen? Should it? Are the vested interests of some commercial entities preventing this from happening? How would this approach expand the places where publishing can serve scholarship, extend sharing from infrequent punctuated static events to include multiple steps in the research process? How can we experiment in this new environment while watching the effect at a systems level on the sharing and growth of knowledge?

The bottom line is that the pace of scholarly innovation should not be throttled by a centralized system that can no longer keep up with the rate of input. If these limitations can be mitigated by shifts in how we view scholarly artifacts and how we view publishing as a monolithic enterprise, then changes must be made. A new era of open, interconnected, and standardized services needs to be created, supported, extended, and adopted by key stakeholders in

order to enable the truly effective use and reuse of scholarship.

This concept of publishing extends well beyond the remit of publishers as we have traditionally known them. It is possible, even likely, that this change may mean the dissolution of some of the services that publishing has traditionally provided, while at the same time giving rise to a host of other new services that have not yet been conceived.

Encouragingly, this approach need not only be motivated by competition between publishers; there is room—and given the interconnectedness of publishing, there are good reasons—for collaboration as well, especially with regard to core infrastructure. Crossref, ORCID, and DOIs are all examples of where stakeholders—whether publishers, funders, institutions, or individuals—have collaborated in efforts to create systems that benefit everyone. Creating these shared infrastructures has not always been easy, but they have shown that collaboration is possible with the right governance in place.

So how do we get there from here? We can't just announce tomorrow that henceforth, all publishers will offer unbundled products and services. Here is one proposal:

1. As a first step, explore how current scholarly publishing stakeholders might be incentivized to embrace these changes. Begin by looking at where innovations in publishing are emerging, including organizations outside of the traditional publisher model that operate on principles of openness and cross-compatible data and use. Look at the challenges these services have

faced, the successes they have gained, and how differences in their funding, adoption models, and operating structures have contributed.

2. Following this report, a working group of OSI participants and other representatives should explore how the work of these free-standing services could be extended to cover the traditional publishing components like filtering, editing, dissemination, registration, and so on. All services in these areas will not be equally adaptable to new models, but those with highest potential for impact, or most easily shifted, should be highlighted by the working group.

3. Ongoing work on this model should seek to describe an open source, federated infrastructure that can serve as the underlying architecture, the primary goal of which will be to allow those contributing scholarly artifacts to have access to a wider variety of standardized services offered a-la-carte or in bundles. Integration of existing publisher infrastructures and allowing low cost and easy participation by new services should also be addressed.

Conclusions

It is important not to fall prey to false dichotomies and oversimplifications, such as the belief that the totality of scholarly publishing is broken and that what was once all paid print must now be all digital. Improving the world of scholarly publishing where it needs improving will be challenging, as this is a centuries old system with entrenched expectations and cultural norms. In today's society, there is an increasing demand for (and expectation of) more and freer access to information. In addition,

without planned change, the scholarly publishing system may simply fall too far behind what researchers and technologists working in collaboration are able to create on their own, outside of traditional publishing.

The challenge for reformers to this system will be to figure out how to improve it quickly, realistically, and sustainably without losing the myriad benefits it provides. To this end, unbundling publisher services is an idea worth exploring. A careful examination of the current scholarly publishing ecosystem needs to be made first, followed by discussions about developing a shared infrastructure for publishers.

No one can deny the appeal of a world in which research knowledge moves more freely. The future possibilities for more efficient discovery alone are tremendously exciting, to say nothing of the new products and services that might arise in a rapidly growing and rejuvenated publishing ecosystem. Getting to the starting gate will take perseverance, goodwill, and significant cooperation, but working together, we think this is a challenge that can be realistically and sustainably met in the near future, and that once met, will provide our global society with vast benefits.

OSI2016 "What is Publishing?" (1) Workgroup

Amy Brand, Director, The MIT Press

James Butcher, Publishing Director, Nature Journals

Meg Buzzi, Director, Opus Program, UCLA

Kathleen Fitzpatrick, Director of Scholarly Communication, Modern Language Association (MLA)

Ann Gabriel, Vice President, Academic & Research Relations, Elsevier

Rikk Mulligan, Program Officer for Scholarly Communication, Association of Research Libraries (ARL)

Vivian Siegel, Director of Education and Training, Global Biological Standards Institute

Matt Spitzer, Community Manager, Center for Open Science (COS)

Jamie Vernon, Director of Science Communications and Publications at Sigma Xi and Editor-in-Chief, American Scientist

Notes:

[1] Crossref, as of June 13, 2016: http://crossref.org/

[2] ORCID, as of June 13, 2016: http://orcid.org/

---

[3] CRediT, as of June 13, 2016: http://casrai.org/credit

[4] T. Fox. Crisis in Communication: The Functions and Future of Medical Journals. London: The Anthlone Press. 1964.