Making the future of scholarly communications

Steven A. Hill

INTRODUCTION

The scholarly communications system is a central feature of the way in which research works. In fact, it could even be argued that this system is the primary driver of the culture that shapes research practice within the academic sphere. After all, the insights from research are of little, if any, value until they are shared, and scholarly communications are the currency that allows the reputation economy of academia to function. The future of scholarly communication is, therefore, central to the future of research itself. What are the drivers for change? How will, and should, the scholarly communications system respond? And, perhaps most importantly, which of the possible futures do we want to create? These are important questions for both researchers and research policy makers.

The questions are also important for the scholarly publishing business – commercial publishers, learned societies, and university presses. Developments in the world of university presses are especially interesting in this regard. While there are, of course, many long-established presses, there is currently a renaissance and reconfiguration of the university press, especially in the UK (Cond, 2016). This renaissance is surely a response to the changing landscape of scholarly communications, and how university presses respond has real potential to shape the future of the system.

My aim in this essay is to consider the future of scholarly communication. I will argue that the primary drivers for the future of scholarly communications are the transformations brought about by digital technology and the cultural and societal push towards openness. While these drivers are indeed leading to sometimes disruptive change in the scholarly communication system, I will suggest that there are opportunities being missed. Conservatism and lock in are leading to only a digital representation of the physical world, with limited opportunities for real innovative practice. In the final part of the essay, I will reflect on the policy response that might facilitate innovation in the scholarly communications system.
best known of these is the Public Library of Science (PLOS) collection of journals. The uptake and adoption of these new platforms has been considerable; for example, although only founded in 2001, in 2015, more than 31,000 articles were published in the journal PLOS ONE (Denker, 2016). In the UK Research Excellence Framework (REF) in 2014, PLOS ONE journal articles were the third most commonly submitted research outputs, accounting for 0.9% of all journal articles (Research Excellence Framework, 2014).

Alongside these changes in the journal publication process, the ubiquity of digital communication and storage has also seen the development of a new route of communication, the preprint. The first preprint server, arXiv, was initiated in 1991 and has become an established part of the publishing landscape in physics, computer science, and mathematics, hosting one million articles by the end of 2014 (Van Noorden, 2014). Preprint servers allow the publication of manuscripts that have been submitted for publication prior to peer review. This provides a route for rapid dissemination of findings, albeit through versions that have not been modified following peer review. Following the success of arXiv, equivalent facilities have been created for biology (bioRxiv) and social science (the Social Science Research Archive, SSRN, and following the acquisition of SSRN by Elsevier, SocArXiv). Further innovation has developed around the idea of preprint servers, with journals, such as F1000 Research, where all articles are published in preprint form, prior to peer review.

The digital revolution has also had an impact on long-form scholarly communication: monographs, edited collections, and scholarly editions. Such long-form publications are increasingly available in digital formats, alongside print editions, although this is still by no means always the case. The usage of digital versions for long-form scholarly work has remained relatively low, however, and certainly does not approach the near-ubiquity found with journals.

As well as changing the formats in which scholarly material is made available, the digital revolution, especially the Internet, has had a dramatic impact on the ways that scholars find scholarly material, both for the location of already-identified sources and for the search for new content. For example, a recent survey of UK academics revealed the central importance of general purpose search engines and Google Scholar (Wolff, Rod, & Schonfeld, 2016).

On one level then, the scholarly communication system looks very different as a result of this digital transformation, and academic practices have changed significantly. However, in many ways, the changes are of a very limited nature. Henry Ford, speaking about the development of the model-T, famously said ‘If I had asked people what they wanted, they would have said faster horses’, and to me, the digital scholarly communications system looks to have gone in this direction.

Journal articles are primarily consumed in the form of PDF files that electronically replicate the appearance of pages in a printed journal. While many journals make versions available in HTML, researchers continue to prefer to use the PDF version (Aalbersberg, 2013). Alongside this emulation of the print format, many journals also continue to limit the length of articles and require the inclusion of page numbers in citations, both anachronistic hangovers from the print era.

Similarly, digital versions of long-form publications essentially mirror physical books. Again, the PDF format dominates, although other formats more suitable for e-Readers are sometimes available (Cox, 2010). Also the case is that currently available e-Readers replicate the experience of the physical book. From the perspective of scholarly use, e-Readers only offer a limited set of functionality compared to the physical book (Schonfeld, 2013), with the focus being on the linear reading of the whole volume.

Perhaps most striking is the retention of the dichotomy of short-form, journal article, and long-form communication, which split along essentially arbitrary disciplinary lines (Crossick, 2015; Hill, 2015). These differences persist, even in the face of supposedly considerable policy pressures from research assessment (Digital Science, 2016). There is some innovation in this area from some publishers (e.g. the Palgrave Pivot imprint, Palgrave), but the focus and uptake is largely in disciplinary areas where monograph publishing is commonplace.

THE INNOVATION GAP

The digital transformation of scholarly communication is only partial, and that begs the question of whether a system based on the digital replication of the physical world is sufficient. While the current system has operated for centuries, I will argue that a wider transformation of the scholarly communication system has the potential to revolutionize the process of research in a positive way. There is an innovation gap that is preventing progress.

In the digital world, the length of a communication does not need to be constrained, so scholarly communications could be simply the size they need to be to communicate ideas and findings clearly. Indeed, even the idea of a fixed communication point is unnecessary. Scholarship is by its nature incremental and developing, so it is possible to imagine scholarly outputs that are living documents and that are updated as new data, analysis, or insights become available.

The availability of longer-form material in digital format allows authors to move away from the linear presentation of ideas. Instead, it is possible to present a range of paths through material or even for readers to develop their own paths. There can also be layers of content of increasing detail or sophistication that allow readers of differing specialities or levels of expertise to navigate entry into content. In addition to opening scholarship up to audiences beyond the research community, there is also the potential to facilitate working across academic disciplines. There are already experiments on breaking down the linear form in creative writing (see, e.g. Arcadia by Iain Pears; http://arcadiatheapp.com/). Developing technologies that allow this non-linear access to scholarly content is important as they mirror the ways in which scholars actually use long-form material (Crossick, 2015).

While there are opportunities for readers to navigate their own paths through digital scholarly works, perhaps greater...
opportunities exist for researchers to link portions of content from existing material to create new contributions. While this already occurs through citations, the potential for linking and ‘remixing’ scholarly content is currently untapped. Alongside such remixing, there is also the potential for annotation of content. Annotation in various forms currently shapes part of the workflow of most scholars, and in certain disciplines, ‘scholarly editions’ are recognized research outputs. However, digital representations offer the potential for these annotations to be shared and to become contributions to scholarship in their own right. This is currently limited by the lack of interoperability between annotations in different formats. While it is currently possible to annotate PDF files, web content, or content for e-Readers, there are limited opportunities for blending annotations together from different sources. The recent developments towards the adoption of hypothesis (see https://hypothes.is/) as a standard for annotation offers considerable hope to address this, but the effectiveness is limited both by the range of digital formats that will be available and the extent to which scholarly reading occurs through physical formats.

Another potential benefit offered by a more digital-only scholarly communication is the seamless inclusion of content beyond the text and image currently associated with scholarly works. In addition to alternative formats for communication, such as video and audio, it is also possible to include access to primary data and sources directly within the digital work.

A further area that is only just beginning to be exploited is the use of text mining and semantic analysis. With the development of appropriate algorithms and the availability of sufficient processing power, it is now possible to generate real insights from large bodies of text. Much of the pioneering work in this area is being driven by research questions in the humanities, as evidenced by the rise of the new discipline of Digital Humanities. In addition to using automated text analysis as a research tool, it is also time to turn that power onto the products of scholarly communication themselves. An interesting and promising example is the emerging approach of ‘semantometrics’, which uses full-text analysis to measure the cognitive distance between articles that have a citation relationship (Herrmannova & Knoth, 2015; Knoth & Herrmannova, 2014). While many conventional bibliometric approaches are of only limited value (Wilsson et al., 2015), this new technology offers the potential to develop truly meaningful measures of research progress.

For these innovations to come to pass, there is also a need for real progress on two areas of openness: first, open access to content and, secondly, content made available in open standards that facilitate reuse. All of the innovations described in this section require access to the full text of scholarly works, in formats that are both human- and machine-readable and with permissions that allow easy reuse. For these innovations to happen, scholarly works need to be both in the right formats and also exist outside of ‘walled gardens’ that prevent the merging of content (Johnson, 2010). As discussed by Neilson (2012), the confluence of digital and open has the potential to transform scholarship.

**CLOSING THE GAP**

In the preceding sections, I have argued that the scholarly communication system has only been partially disrupted by the transformation to digital and that, as a result, we are missing opportunities to transform not just scholarly communications but scholarship itself. Building on these arguments, I will now consider the barriers that are preventing this transformation and discuss the role of the policy environment in removing those barriers and bringing about change.

There are technological constraints that need to be overcome. Current e-Readers, for example, make dealing with texts in a flexible way difficult, and the handling of PDF files in various software readers has limited flexibility too. There are also infrastructure and standards gaps. While standards are emerging for annotation, for example, they are far from universal and need to cope better with the complex multi-device online/offline world that is the reality for most scholars. While HTML has many advantages over proprietary formats, it also has limitations, so advances are needed here too (see, e.g. http://scholarlyhtml.org/).

While these technological barriers are important, their solution is, at least in part, likely to be demand-driven. Until authors and readers ask for innovation, the impetus to produce solutions will be limited, and there is evidence to suggest that demand is currently limited by deep-seated cultural norms within the academic community.

Looking at the scholarly monograph, for example, there is evidence of resistance to the use of the digital format within the academic community (OAPEN, 2015; Wolff et al., 2016). The digital versions of scholarly books appear to be primarily used to filter content, with deep reading being reserved for the print copy (Faherty, 2016). This reflects a view that the format of the book is somehow optimal (Crossick, 2015), but there is little evidence either way to inform this position. Similarly, there is little pressure from the users of journal articles to move away from the status quo.

Alongside the general acceptance of current scholarly communications norms, there is the tight linkage of those norms into the mechanisms of reputation and reward within the research system. In many disciplines, there are implicit (sometimes explicit) requirements to publish in particular forms, or in particular venues, to demonstrate performance, which serves only to reinforce current behaviours and act as a deterrent to innovation.

The position poses a dilemma for policy makers. If we are to see the additional benefits to research and scholarship that will flow from innovation in scholarly communication, we must either impose change on a resistant (or at least ambivalent) academic community or construct a policy environment that encourages incremental change. The former approach is sometimes argued as going against the principles of academic freedom, and the imposition of change risks stimulating innovation for its own sake, without generating real benefits for research.

But evidence suggests that only focusing on removing barriers to innovative practice is unlikely to produce change at
significant pace. For example, in the UK’s REF, and the Research Assessment Exercises before it, the policy position has been clearly permissive of a diversity of research outputs submitted for assessment. For example, the guidance on submissions for REF 2014 (Research Excellence Framework, 2011) states:

‘An underpinning principle of the REF is that all forms of research output will be assessed on a fair and equal basis. Subpanels will not regard any particular form of output as of greater or lesser quality than another per se.’

Despite this, and the undoubted strength of research assessment as a policy lever, the outputs submitted to REF are, in the vast majority, journal articles and long-form, text-based scholarly works. The creation of a permissive policy environment on its own appears insufficient to bring about change.

In my view, the permissive policy environment needs to be accompanied by carefully thought-through policy interventions that actively encourage innovation. Funding to support the production of new forms of scholarly output needs to be part of the mix. There is also an important role for projects like the AHRC/ British Library-funded Academic Book of the Future (see https://academicbookfuture.org/) that both apply rigorous research to the process of scholarly communication and engage in innovation themselves (see, e.g. https://www.ucl.ac.uk/ucl-press/ucl-press-news/call-for-content-booc). I also think we should be doing more to identify and celebrate innovative practice where it occurs. A national prize for innovative scholarly communication might be a method for doing this.

University presses have a central role to play in shaping this new world of scholarly communications. Their position close to researchers gives them access to leading edge thinking, in terms of the possibilities that digital publishing offers, how those possibilities can be deployed to enhance scholarship, and the technological innovation needed to make the possibilities practical. Working in partnership with scholars, university presses can and should be the leaders in innovation. An example, already mentioned, is the Book as Open Online Content (BOOC) project, through which the outputs from the Academic Book of the Future will be disseminated. This innovative approach is being developed and delivered with UCL Press.

Policy makers have perhaps not focussed enough on university presses as potential agents of change and innovation. Although not always easy for research funders, direct funding of innovative activity in university presses is a potentially powerful way to influence change.

CONCLUDING COMMENTS

The world of scholarly communication stands at a crossroads. On the one hand, there is much potential and actual innovation; on the other hand, cultural, commercial, and technological barriers are limiting this potential. A policy environment that is simply permissive of innovation has been insufficient. Instead, we need a policy environment that will shape a positive future for scholarly communications and so the research system itself.

REFERENCES

Aalbersberg, I. J. (2013, July 9). PDF versus HTML – Which do researchers prefer? [Web log post]. Retrieved from https://www.elsevier.com/connect/pdf-versus-html-which-do-researchers-prefer

Cond, A. (2016, March 16). The University Press Redux: Balancing traditional university values with a culture of digital innovation [Web log post]. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2016/03/16/the-university-press-redux/

Cox, L. (2010). Scholarly book publishing practice: The ALPSP survey findings. Learned Publishing, 23, 347–356. doi:10.1087/20100412

Crossick, G. (2015). Monographs and open access: A report to HEFCE. Retrieved from http://www.hefce.ac.uk/media/hefce/content/pubs/indirreports/2015/Monographs,and,open,access/2014_monographs.pdf

Denker, S. P. (2016, February 24). PLOS 2015 reviewer thank you [Web log post]. Retrieved from http://blogs.plos.org/plos/2016/02/plos-2015-reviewer-thank-you/

Digital Science. (2016). Publication patterns in research underpinning impact in REF2014. Retrieved from http://www.hefce.ac.uk/media/HEFCE.2014/Content/Pubs/Independentresearch/2016/Publication,patterns,in,research,underpinning,impact,in,REF2014/2016_refimpact.pdf

Faherty, A. (2016). Academic book discovery, evaluation and access. Retrieved from https://academicbookfuture.files.wordpress.com/2016/06/faherty_academic-book-discovery-full-report.pdf

Herrmannova, D., & Knoth, P. (2015). Semantometrics in coauthorship networks: Fulltext-based approach for analysing patterns of research collaboration. D-Lib Magazine, 21(11/12). doi:10.1045/november2015-herrmannova

Hill, S. A. (2015). On scholarly monographs and the sciences. Retrieved from http://steinenhill.org/on-scholarly-monographs-and-the-sciences/

Johnson, S. (2010, April 23). The glass box and the commonplace book [Web log post]. Retrieved from http://www.stevenberlinjohnson.com/2010/04/the-glass-box-and-the-commonplace-book.html

Knoth, P., & Herrmannova, D. (2014). Towards semantometrics: A new semantic similarity based measure for assessing a research publication’s contribution. D-Lib Magazine, 20(11). doi:10.1045/november2014-knoth

Neilson, M. (2012). Reinventing discovery. Princeton, NJ: Princeton University Press.

OAPEN. (2015). Researcher survey 2014. Retrieved from http://oapen-uk.jiscebooks.org/files/2012/02/OAPEN-UK-researcher-survey-final.pdf

Palgrave. (n.d.) Palgrave pivot. Retrieved from https://www.palgrave.com/gp/palgrave-pivot

Research Excellence Framework. (2011). Assessment framework and guidance on submissions. Retrieved from http://www.ref.ac.uk/media/ref/content/pub/assessmentframeworkandguidanceon submissions/GOS%20including%20addendum.pdf
Research Excellence Framework. (2014). Analysis of REF 2014 outputs data. Retrieved from http://results.ref.ac.uk/DownloadSubmissions/ByForm/REF2

Research Information Network. (2009). E-journals: Their use, value and impact. Retrieved from http://www.webarchive.org.uk/wayback/archive/20120107234518/http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/e-journals-their-use-value-and-impact

Schonfeld, R. C. (2013, December 10). Stop the presses: Is the monograph headed towards an e-only future? [Web log post]. Retrieved from http://www.sr.ithaka.org/wp-content/uploads/2013/12/SR_BriefingPaper_Presses_120913.pdf

Van Noorden, R. (2014, December 30). The arXiv preprint server hits 1 million articles. Nature. doi:10.1038/nature.2014.16643. Retrieved from http://www.nature.com/news/the-arxiv-preprint-server-hits-1-million-articles-1.16643

Wilsdon, J., Allen, E., Belfiore, E., Campbell, P., Curry, S., Hill, S., …, Johnson, B. (2015). The metric tide. doi:10.13140/RG.2.1.4929.1363. Retrieved from http://www.hefce.ac.uk/media/HEFCE2014/Content/Pubs/IndependentResearch/2015/TheMetricTide/2015_metric_tide.pdf

Wolff, C., Rod, A. B., & Schonfeld, R. C. (2016, June 15). UK survey of academics 2015. Retrieved from https://repository.jisc.ac.uk/6437/1/ithaka-survey-of-academics-2015.pdf