

Knowledge infrastructure and the role of the university

Amy Brand*

The MIT Press, Cambridge, MA, USA

Abstract. As open access to research information grows and publisher business models adapt accordingly, knowledge infrastructure has become the new frontier for advocates of open science. This paper argues that the time has come for universities and other knowledge institutions to assume a larger role in mitigating the risks that arise from ongoing consolidation in research infrastructure, including the privatization of community platforms, commercial control of analytics solutions, and other market-driven trends in scientific and scholarly publishing.

Keywords: Open science, knowledge infrastructure, institutional responsibility

1. Introduction

The world has changed enormously in the few months since I gave the opening keynote address at the first NISO+ conference on which this paper is based, back in February of 2020. My presentation concerned the need for more diversity and competition in knowledge infrastructure, and in particular the imperative for universities to step into a larger research dissemination and technology role [1]. My argument started from the premise that the steep rise in openly accessible research content will not turn out to be the panacea that many in the research community were hoping for, as long as the platforms for delivering and analyzing that content are themselves locked down [2]. Not only has the coronavirus pandemic brought into even sharper focus how important ready-access to credible research information is for scientists and scholars, for policy makers, and for the global public, but it will also no doubt accelerate the demise of legacy business models across the research information arena. It may be too soon to say with any specificity what the longer term impact of this crisis will be on publishers, libraries, aggregators, and other stakeholders in the information ecosystem, but it is already clear that the impact will be momentous.

2. The pandemic, research, and changes in communication

As I write this in May of 2020, several major research institutions have begun to share information about the staggering financial losses they anticipate in relation to COVID-19. It is clear that the pain will be felt across the academy; indeed, it already is [3]. Many heads of university libraries are reporting considerably reduced budgets in the year ahead, and their intentions to move away from “big deals” and

*Corresponding author: Amy Brand, PhD, Director, The MIT Press, 1 Rogers Street, Cambridge, MA 02141, USA. Tel.: +1 (617) 955 6864; E-mail: amybrand@mit.edu.

de-prioritize further the collection of print materials [4]. As the director of a large university press with prosocial values and no endowment funds, producing books and journals in a fast-changing information marketplace, I was already spending a good deal of time focused on sustainability in publishing academic books and journals. Now, I think about it nonstop. With physical bookstores closed around the world for months and other less visible hits to the supply chain, and with the attention of potential book buyers understandably elsewhere during this crisis, sales of adult nonfiction books are down significantly across the industry in the last quarter of 2020. For many small University Presses who were already struggling [5], the future is looking even more precarious.

At the same time, the pace of research efforts to address the pandemic is unprecedented. As a result, we are witnessing an explosion of research communications related to the pandemic, much of it on preprint platforms like bioRxiv and medRxiv, reflecting an accelerated shift towards what is now referred to as the publish-review-curate (PRC) model. Developments that speed discovery in the face of a crisis such as this should be welcome, yet there are significant risks to public health when research outputs, especially those with clinical applications, are circulated unvetted and misused or misinterpreted in public media. This growing concern has been reported widely in the mainstream media [6].

3. Changes in peer review

The preprint model emerged in the early 1990s with the launch of arXiv for pre-refereed papers in physics. arXiv has since grown to support the deposit of works in a wider set of quantitative fields – mathematics, computer science, quantitative biology, quantitative finance, statistics, etc. In many ways, the preprint model is more “organic” to those fields and publishing ecosystems than it is for biomedical and clinical fields, where the vetting of protocols, human subject practices, data, statistical methods, and potential conflicts of interest figures centrally in the publishing process, in order to protect individual patients, public health, and the overall credibility of the research publishing enterprise. Hence, numerous initiatives to layer peer review on top of the pre-print ecosystem have been launched in recent months, including a forthcoming journal from the MIT Press called *Rapid Reviews Covid-19*. To meet the demand for trustworthy information about the pandemic, we have also responded with a series of digital-first, rapidly produced books called *MIT Press First Reads*.

If the truth wars being waged in recent years in news media and on social platforms have taught us nothing else, they have shown us just how inextricable information is from the systems that host and deliver it. This is true for everyday journalism and scientific publishing alike. Indeed, the battle for control over information and knowledge is a defining struggle of our time, and it is a battle over infrastructure broadly defined. For publishers and service providers in the research information sector, the basic challenges in delivering peer-reviewed research content to a target readership for its intended purpose and impact are in some ways more complex and interesting than ever before, as we work to harness the power of multiple platforms, metadata tags, standards, licenses, identifiers, business models, and metrics to accelerate discovery and promote trust in scientific and scholarly information.

4. Building a university-centric publishing ecosystem

Research universities are in the business of creating and transmitting knowledge. They are the locus of empirical research and critical inquiry; they contain libraries and some even have university presses.

Yet they have historically under invested in capturing and communicating the knowledge being created on their campuses, having outsourced much of the relevant functions and infrastructure, including not only publishing, but also the design and tracking of metrics upon which academic reputations are judged. As we come to terms with the fact that established models and practices are increasingly unfit for purpose in a world of shrinking library budgets, overpriced journals, unpurchased monographs, and oligopolistic analytics platforms (not to mention demand for increased speed and access), the imperative grows for universities themselves to assume a greater knowledge dissemination and infrastructure role. How might we build a more sustainable university-centric publishing ecosystem? How do we foster collaboration and collective investment within and across universities? At the same time, how do we avoid “reinventing the wheel” of academic publishing through new efforts that overlook or devalue the expertise and skill sets of publishing professionals?

I am far from alone in believing that it behooves institutional leaders today to fully explore the implications of commercial control of research data, analytics, and publishing infrastructure, along with the positive potential for community-based alternatives [7,8]. MIT took the unusual step in 2018 of “manifesting” the future we would like to help create, by launching the Knowledge Futures Group (KFG) with the twofold goal to incubate home-grown technologies and to spark a movement towards greater institutional investment in knowledge infrastructure.

5. Conclusion

The research community is rightfully celebrating more open access and open data, yet there is growing recognition in the academic community that pay-to-publish open access is not the panacea people were hoping for when it comes to affordable, sustainable scholarly and scientific publishing. Publication is, after all, only one step in a flow of research communication activities that starts with the collection and analysis of research data and ends with assessment of research impact. Open science is the movement towards open methods, data, and software, to enhance reproducibility, fairness, and distributed collaboration in science. The construct covers such diverse elements as the use of open source software, the sharing of data sets, open and transparent peer review processes, open repositories for the long-term storage and availability of both data and articles, as well as the availability of open protocols and methodologies that ensure the reproducibility and overall quality of research. How these trends can be reconciled with the economic interests of the publishing industry as it is currently organized remains to be seen, but the time is ripe for greater multi-stakeholder coordination and institutional investment in building and maintaining a diversified open infrastructure pipeline.

References

- [1] A. Brand, The other I-Word: Infrastructure and the future of knowledge. Posted to the NISO Plus Repository in April of 2020, https://nisoplus.figshare.com/articles/NP2020_The_Other_I-Word_Infrastructure_and_the_Future_of_Knowledge/12093690, last accessed on June 6, 2020.
- [2] C. Aspesi and A. Brand, In pursuit of open science, open access is not enough, *Science* **368**(6491) (2020), 574–577. doi:10.1126/science.aba3763, last accessed June 6, 2020.
- [3] F. Furstenberg, University leaders are failing, *The Chronicle of Higher Education* (2020), <https://www.chronicle.com/article/University-Leaders-Are-Failing/248809>, last accessed June 6, 2020.
- [4] L. McKenzie, Libraries brace for budget cuts, *Inside Higher Ed* (2020), <https://www.insidehighered.com/news/2020/04/17/college-librarians-prepare-looming-budget-cuts-and-journal-subscriptions-could-be>, last accessed June 6, 2020.

- [5] P. Givler, University press publishing in the United States. in: *Scholarly Publishing: Books, Journals, Publishers, and Libraries in the Twentieth Century*, R.E. Abel and L.W. Newman (eds), Wiley, New York, 2002.
- [6] W. Yan, Coronavirus tests science's need for speed limits, *The New York Times* (2020), updated April 21, 2020, <https://www.nytimes.com/2020/04/14/science/coronavirus-disinformation.html>, last accessed June 6, 2020.
- [7] S. Kember, Distributed open collaborative scholarship, *The Knowledge Futures Commonplace* (2020). doi:[10.21428/6ffd8432.fdf89d72](https://doi.org/10.21428/6ffd8432.fdf89d72), last accessed June 6, 2020.
- [8] D. Greenstein, Next-generation university publishing: A perspective from California, *Journal of Electronic Publishing* **13**(2) (2010). doi:[10.3998/3336451.0013.205](https://doi.org/10.3998/3336451.0013.205), last accessed June 6, 2020.