Big Deal

Should Universities Outsource More Core Research Infrastructure?

January 4, 2018

Roger C. Schonfeld

ITHAKA S+R

Ithaka S+R provides research and strategic guidance to help the academic and cultural communities serve the public good and navigate economic, demographic, and technological change. Ithaka S+R is part of ITHAKA, a not-for-profit organization that works to advance and preserve knowledge and to improve teaching and learning through the use of digital technologies. Artstor, JSTOR, and Portico are also part of ITHAKA.

Research universities have developed in symbiosis with a robust set of commercial providers that serve their needs. From food service providers to run dining halls to private equity firms to manage parts of the endowment, outsourcing has allowed universities to remain more focused on their core educational and research functions. But universities have also at times elected to outsource academic infrastructure. Commercial firms have developed a major role in several significant university functions, including scientific publishing, library management systems, and course management systems. And in all three cases, the commercial priorities of vendors have at times left academia frustrated. While outsourcing is not uniformly good or bad, services with a principally academic market seem to be especially susceptible to monopoly or oligopoly dynamics among commercial providers.
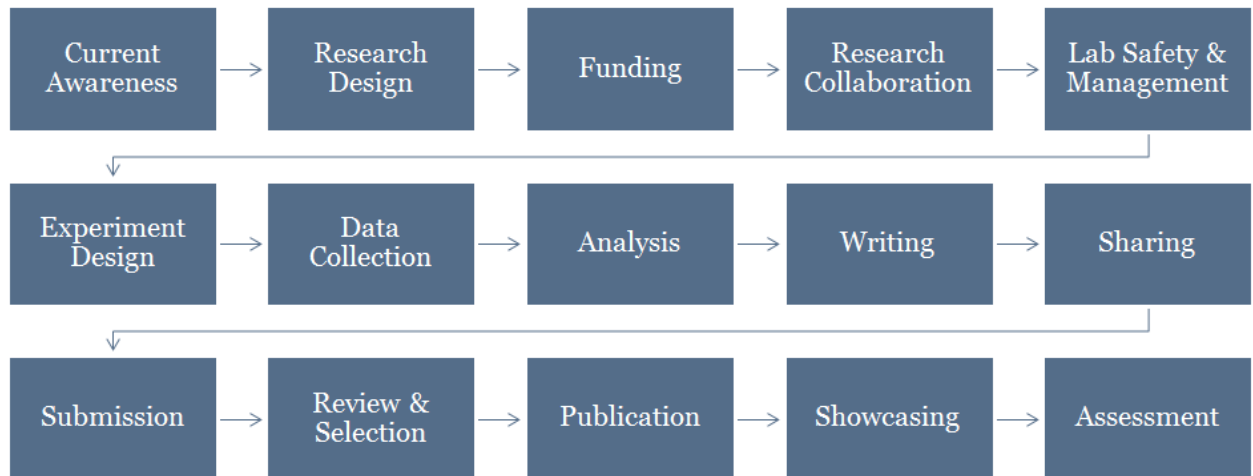
Today, an entirely new class of tools and platforms is developing that stands to revolutionize scientific research in academia. These research workflow tools will variously improve transparency, reproducibility, accountability, and efficiency of the scientific research process, an enormous boon to scientists themselves and their universities that have so much invested in academic science. But in assessing and adopting these valuable new research workflow tools, research universities once again face key questions about whether and under what circumstances to outsource core scholarly infrastructure with a principally academic market.[1]

Research workflow

Scholars and in particular the laboratory scientists who work as teams can dramatically increase their productivity and effectiveness with tools that support them systematically from research design and funding through data collection and analysis to assessment and showcasing, as can be seen in this very rough schematic of their overall workflow:

---

[1] For discussion and guidance that has helped me frame my thinking in these areas, I thank Emily Drabinski, Joseph Esposito, Lisa Hinchliffe, Kimberly Lutz, and Donald J. Waters. Any weaknesses or limitations are my responsibility alone.

Current Awareness	→	Research Design	→	Funding	→	Research Collaboration	→	Lab Safety & Management
Experiment Design	→	Data Collection	→	Analysis	→	Writing	→	Sharing
Submission	→	Review & Selection	→	Publication	→	Showcasing	→	Assessment

It isn't hard for a scientist to imagine moving seamlessly across the various steps of the research lifecycle: from programing a scientific instrument, to analyzing the dataset it created, to collaborating on the draft article derived from the analysis, to working through peer reviewer comments, to examining the impact of the research once published, to including those statistics in the grant application for the next phase of research, to keeping up with the newest scholarship in the field.

Traditionally, each of these steps along the workflow utilizes different tools and systems. In part, this is because the overall workflow is supported by, or engages with, a variety of different functions, including grant writers, research funders, notebook providers, instrument providers, analysis and writing software, publishers, and libraries.

Individual digital tools and services can bring great improvements to the researcher experience at almost every point along the workflow. Bringing greater seamlessness across the full research lifecycle through integrations across these discrete steps is a path towards yet greater improvement.

Providers

In recent years, scientists have been developing a variety of tools and services to support many of these discrete stages of their research process, with the primary objective of increasing efficiency, dissemination, and impact. Today, investments by private equity and publishers alike are helping these tools and services develop and carrying them in new directions.

I've written extensively about this turn to workflow over the past two years, as several companies build out their support for scientific workflow.[2] Elsevier has done so by acquiring, among a variety of startups, the preprint services SSRN and bepress, in addition to laboratory notebook and research collaboration tools. And Elsevier is steadily integrating these tools and services with one another and with its publishing empire. Digital Science has built a rival stable of individual firms in a similar space, though speculation that it will be integrated with its half-sibling Springer Nature is being vociferously denied.[3] Clarivate has important strengths in review and assessment. The not-for-profit Center for Open Science is trying to determine its competitive position in this rapidly changing market. The Chan Zuckerberg Initiative acquired Meta and is rumored to have additional interests in scientific research. And there are many independent players as well. Ownership changes can be expected in the future, and the academy should examine all these arrangements with equal care and forethought.

And these providers are pursuing strategies that differ to some degree. Some are more focused on integrating with publishers, others are more focused on maintaining their independence. In some cases, links are being established to tie individual services into more coherent packages if not actual product bundles, and there are opportunities to create various forms of lock-in.[4]

These investments, accompanying integrations, and the risks of lock-in, should cause us to wonder what objectives and strategies are being pursued and what the implications might be. I have elsewhere addressed these questions for scholarly publishers, which face

[2] For more on these acquisitions, see Roger C. Schonfeld, "Elsevier Acquires SSRN," *Scholarly Kitchen*, May 17, 2016, available at https://scholarlykitchen.sspnet.org/2016/05/17/elsevier-acquires-ssrn/ and Roger C. Schonfeld, "Elsevier Acquires bepress," *Scholarly Kitchen*, available at https://scholarlykitchen.sspnet.org/2017/08/02/elsevier-acquires-bepress/. With respect to the Center for Open Science, see Roger C. Schonfeld, "The Center for Open Science, Alternative to Elsevier, Announces New Preprint Services Today," *Ithaka S+R Blog*, August 29, 2017, available at http://www.sr.ithaka.org/blog/the-center-for-open-science-alternative-to-elsevier-announces-new-preprint-services-today/.

[3] For the speculation, see Roger C. Schonfeld, "Who Owns Digital Science?" *Scholarly Kitchen*, October 23, 2017, available at https://scholarlykitchen.sspnet.org/2017/10/23/ownership-digital-science/. For the denial, see Daniel Hook, "'Who owns Digital Science' – That is the Question…," November 8, 2017, available at https://www.digital-science.com/blog/news/owns-digital-science-question/.

[4] On provider strategies, see Roger C. Schonfeld, "Strategy & Integration Among Workflow Providers," *Scholarly Kitchen*, November 7, 2017, available at https://scholarlykitchen.sspnet.org/2017/11/07/strategy-integration-workflow-providers/. On lock-in, see Roger C. Schonfeld, "Workflow Lock-In: A Taxonomy," *Scholarly Kitchen*, January 2, 2018, available at https://scholarlykitchen.sspnet.org/2018/01/02/workflow-lock-taxonomy/.

important questions of their own.[5] In the remainder of this piece, I will examine some of the implications for universities and for science.

Structure

Today almost no university is positioned to address its core interests here in any truly coherent way. The reason is essentially structural. There is no individual or organization within any university that I am aware of that is responsible for the full suite of research workflow services. All agree, broadly speaking, about the importance of the new tools and services that touch on their areas of responsibility. But coordination across the responsible parties to address the integrations that are beginning to emerge across services, and therefore across responsible university divisions, is limited at best. Fragmented decision-making cannot address issues of collective strategy.

Fragmented decision-making cannot address issues of collective strategy.

No campus office or organization has responsibility for anything other than a subset of the system. Some examples help to illustrate the fragmentation. The library is principally responsible for discovery and access, and to some extent for preprint dissemination. The CIO or the senior executive for research may be responsible for certain services, for example research computing and research data. The provost or the senior research executive may be responsible for tools that support funding and research productivity assessment. A safety office may have compliance responsibilities for certain aspects of laboratory management. And in many cases the principal investigator running a laboratory has wide discretion as well. Notwithstanding growing efforts by the VP for Research and/or the research library to engage with many of these other parties on campus, and find ways to collaborate, these efforts are typically broad and informal or focused on a specific subset of the overall issues, for example just on research data management. There is rarely a working group, or senior executive, responsible for the full array of research workflow services and support.

---

[5] See Roger C. Schonfeld, "Workflow Strategy for Those Left Behind: Strategic Context," *Scholarly Kitchen*, December 18, 2017, available at https://scholarlykitchen.sspnet.org/2017/12/18/workflow-strategy-left-behind-context/ and Roger C. Schonfeld, "Workflow Strategy for Those Left Behind: Strategic Options," *Scholarly Kitchen*, December 19, 2017, available at https://scholarlykitchen.sspnet.org/2017/12/19/workflow-strategy-left-behind-options/.

The high-level implications of this disorganization are straightforward. As individual university divisions acquire new categories of core scholarly infrastructure, it is virtually impossible for the university as a whole to have a coherent strategic position on key issues. Among the most important questions that they are unable to address strategically is whether core scholarly infrastructure should be outsourced and if so to what kind of organization(s).

Restructure

It is therefore imperative that universities -- and the systems and consortia through which they collaborate -- position themselves accordingly. They need to ensure that the appeal of the new category of workflow products is addressed not just at the level of existing university divisions but in a way that recognizes how the paradigms for research support is shifting under a workflow mindset.

At a minimum, each university should be positioned to serve as a strong interlocutor and negotiator, not a disorganized jumble, to companies when they come calling with these new products to offer. This means that a company cannot go around one "no" or set of requested license revisions to another office for a "yes" -- just as a publisher cannot, in a well-functioning organization, sell a licensed e-resource around the library's licensing unit. Addressing the need for a strong interlocutor is about communication and coordination, at a minimum, but can also involve changes to organizational structure.

In considering this issue, universities should to the extent possible pursue greater ambitions than simply seeking to coordinate their negotiating posture when a vendor comes calling. Beyond this, a university should act with single purpose to establish strategy, set objectives, coordinate budgets, and act operationally. To do so, individual units might need to give up some autonomy to ensure that they can act with other units together in the university's overall best interests.

How can individual universities develop visions that are sufficiently shared that they can realize the economic and technical advantages in shared infrastructure that scales across institutional boundaries?

One interesting example is the institutional repository, typically run by a library and playing increasingly important roles in compliance and showcasing scholarly works. With the Elsevier acquisition of the Digital Commons repository service, a number of libraries are considering their alternatives.[6] While there is an interest in finding alternative repositories, the arguments in this piece about workflow generally may suggest that, as I have argued elsewhere, "libraries adopting standalone institutional repositories are moving in exactly the wrong direction strategically."[7] How will library repository planning efforts -- or for that matter the VP of Research's research information management system planning efforts -- fit into a larger university vision? And, how can individual universities develop visions that are sufficiently shared that they can realize the economic and technical advantages in shared infrastructure that scales across institutional boundaries?

Strategy and Outsourcing

The very basic strategic question facing universities is the extent to which it makes sense to continue down the path of outsourcing basic scholarly infrastructure. Surely it is undesirable to outsource core scholarly infrastructure such that switching costs are high enough to limit the effective functioning of a free market. Balancing the desire for low switching costs with the need for a seamless end-to-end experience has been one of the substantial challenges with other outsourced academic infrastructure.

The switching costs for platform infrastructure are often enormous. This is likely to be even more so the case for research infrastructure that reaches into the laboratory than it is for other categories of university business systems. Providers will pursue a lock-in strategy to maximize switching costs, if not outright bundling, for individual research workflow tools. It is imperative that universities understand the nature of switching costs, not only for the university-facing assessment and showcasing tools but also for the increasingly integrated laboratory management systems. If universities cannot hold down these switching costs low enough to ensure they can take advantage of market competition, they will quickly find themselves locked into a single provider or set of providers.

One approach is to insist on loosely coupled discrete services rather than a single integrated workflow. The university can thus adopt a strategy of working with an array of

---

[6] See for example the University of Pennsylvania's "Bexprexit" initiative, https://beprexit.wordpress.com/.

[7] Roger C. Schonfeld, "Reflections on 'Elsevier Acquires bepress': Implications for Library Leaders," *Ithaka S+R Blog*, August 7, 2017, available at http://www.sr.ithaka.org/blog/reflections-on-elsevier-acquires-bepress/.

different providers -- even if pricing makes it advantageous to go "all in" with a single provider. This approach may put the university in a stronger negotiating position in the long run, but it can be expected to do so at the expense of the seamless experience and robust feature set that these platforms make possible.

> ## If individual researchers determine that seamlessness is valuable to them, will they in turn license access to a complete end-to-end service for themselves or on behalf of their lab?

And, indeed, whatever model the university may select, if individual researchers determine that seamlessness is valuable to them, will they in turn license access to a complete end-to-end service for themselves or on behalf of their lab? So, the university's efforts to ensure a more competitive overall marketplace through componentization may ultimately serve only to marginalize it.

If componentization may not ultimately be a good choice, some universities might conclude that research workflow infrastructure should be community controlled rather than commercially outsourced. The Center for Open Science appears to be positioning itself to serve this role, albeit with resource and strategic constraints that deserve scrutiny.

Ultimately, these questions of architecture, strategy, seamlessness, and outsourcing are ones that require sustained attention in a coordinated fashion, both within a university and across research universities.

Licensing

Even in a scenario where research workflow infrastructure is provided on a community controlled basis, it will likely still be provided on some kind of shared basis, for example by a collaboration or consortium. In those cases also, but especially in cases where research infrastructure is provided on an outsourced basis, universities should give substantial attention to licensing terms. Because, in either scenario, universities will wish to ensure that key academic values are protected. The contracts and policies that would govern such infrastructure are therefore essential. Will universities insist on contractual terms that strongly protect their interests and those of their researchers?

Will universities insist on contractual terms that strongly protect their interests and those of their researchers?

There are important questions about who owns which data placed into or generated through the use of research workflow services. What is owned by the university itself, what by its researchers, and what by their funders? In some contexts, addressing this question not only legally but also technically may be no small challenge. Of course data security and privacy considerations are also vital.

And what of this data are reusable by others, including the platform provider itself, in what some might see as "non-consumptive" forms of analysis? One way that research workflow tools are already being monetized is on the basis of selling aggregated anonymized analyzed data to third parties. The appeal of monetizing university data outside of academia is one of the factors driving at least some of the investment flowing to these tools. These are issues that should be subject to negotiated license agreements where standard terms do not adequately protect the university.

Looking Ahead

This is new terrain for academia, but the ground is shifting rapidly. If academia can organize its work and develop a strategic vision for research workflow, there is yet an opportunity to avoid the negative consequences of outsourcing core scholarly infrastructure.