The Second Digital Transformation of Scholarly Publishing

Strategic Context and Shared Infrastructure

Tracy Bergstrom
Oya Y. Rieger
Roger C. Schonfeld
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 with a mission to improve access to knowledge and education for people around the world. We believe education is key to the wellbeing of individuals and society, and we work to make it more effective and affordable.

Copyright 2023 ITHAKA. This work is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of the license, please see https://creativecommons.org/licenses/by/4.0/.

We invite you to comment on this draft report. Please send your suggestions and questions to sr-infrastructre@ithaka.org by August 31, 2023. The final report will be published on the Ithaka S+R website in October 2023.
# Table of Contents

**Introduction** .................................................................................................................................................. 3

**Definitions and Methodology** .................................................................................................................. 4

**Strategic Context** ...................................................................................................................................... 5
- Transitioning to Service Provision ............................................................................................................ 5
- Consolidation and Competition ................................................................................................................... 7
- Humans and Machines ................................................................................................................................. 9
- Trusted Global Public Good ....................................................................................................................... 10

**Some Notes on the Landscape** .............................................................................................................. 11
- Publishing Organizations ........................................................................................................................... 11
- Perspectives .................................................................................................................................................. 12

**Shared Infrastructure Today** ................................................................................................................ 14
- Identifiers and Standards ............................................................................................................................ 15
- Enterprise Publishing Systems .................................................................................................................. 16
- Discovery, Collaboration, and Trust .......................................................................................................... 18
- Preservation ................................................................................................................................................ 19
- Reflections ................................................................................................................................................... 20

**New Opportunities for the Shared Infrastructure** .................................................................................. 21
- Spine of the Scholarly Record ..................................................................................................................... 21
- Research Integrity ....................................................................................................................................... 22
- Making Meaning ........................................................................................................................................ 22
- Supporting New Business Models ........................................................................................................... 23
- Interoperability ......................................................................................................................................... 24

**Recommendations** ................................................................................................................................. 24

**Appendix: List of Interviewees** ............................................................................................................... 24

**Appendix: Interview Questions** .............................................................................................................. 26
Introduction

Today, the scholarly publishing sector is undergoing its second digital transformation. Rather than being recognized primarily as the producer and disseminator of copyrighted materials, the sector is increasingly identified as a service provider to various groups that produce and consume scholarly information. In the face of alternative publishing models and competition from newer entrants, there is substantial consolidation among traditional publishers. The sector’s core goal of facilitating human authorship and readership is giving way to the enablement of machine-to-machine communication. Advancing scholarship and science as a trusted global public good is becoming increasingly complicated due to a polarized political and information environment, enhanced attention to academic fraud and misconduct, and growing geopolitical divergences that amplify variations in scholarly communication policies and practices in different geographies. Scholarly publishing has grown ever more complicated, even as researchers expect a seamless experience.

A robust and nimble infrastructure is imperative to support the vital work of scholarly communication and effectively and efficiently meet the emerging service needs of different stakeholders. Publishers and other scholarly communication services and providers (hereafter collectively, “publishing organizations”) rely on this shared infrastructure in many key parts of their work, and it forms a foundational part of their technology stack and service framework.

Many publishing organizations find that substantial components of the shared infrastructure are either no longer fit for purpose or do not yet exist to support emergent needs. There is widespread consensus that new investment in shared infrastructure is required and even some degree of agreement on the broad purposes that this shared infrastructure should serve.

Yet these gaps exist, and rarely because of raw technical challenges. Rather, they are the result of stubborn strategic, governance, and business model impediments. At the working levels necessary to develop and sustain a thriving shared infrastructure, publishing organizations face real challenges in generating strategic alignment with each other. In many key categories, governance of the shared infrastructure extends beyond well-aligned publishing organizations, adding a further layer of complexity. And there is the ever-present issue of the business model and investment case—who pays, who will pay, and for what—which in turn provide incentives for innovation or inaction.

Through this report, we hope to stimulate a discussion about the future of the shared infrastructure for scholarly communication with its key stakeholders. The draft you are reading now is issued in July 2023 for feedback, which can be shared with us directly at sr-infrastructure@ithaka.org through August 31, 2023. We will publish a final version of this report in October 2023.
Definitions and Methodology

Scholarly communication is the process through which research products and outputs (such as articles, audiovisual materials, data, code, and research methods) are created, assessed, improved, shared, disseminated, and preserved in a variety of modes including through formal and informal publications, conferences, and other academic networking methods. In the digital environment, shared infrastructure has emerged as the key enabler for delivering the services that authors and readers need. It is composed of standards, platforms, technologies, policies, and the communities that enable and support them.¹ Services like reference linking, repositories, identifiers, single sign-on, and digital preservation have supported the digital transformation of scholarly publishing, achieving real efficiencies for all stakeholder communities. The ultimate goal of publishing organizations and their shared infrastructure is to support the global community of researchers to discover, access, and use relevant and trustworthy materials as effortlessly as possible.

Developing, maintaining, and sustaining fit-for-purpose community infrastructure is a challenge particularly when the technology, policy, and business environments are in flux and user behavior and needs are evolving. It is necessary to sustain and, in some cases, improve existing shared infrastructure as some elements of it become more critical while others may decline in value. Given the proliferation of providers, the sustainability of some shared infrastructure elements will be dependent on the competitive differentiation among them. Infrastructures are known to be embedded in ways that only become visible upon breakdown or glitches. It is natural for outside observers to take an infrastructure component for granted if it is working well and has become well-established and embedded, such as metadata standards that facilitate discovery and access. This makes it more difficult to focus on the seamless aspects of shared infrastructure than it is to describe its gaps and challenges.

The purpose of this report is to examine the strategic context behind the development of shared infrastructure and to assess what aspects of that infrastructure are working well and where improvements are needed. We draw on a high-level overview of the shared scholarly communication infrastructure that we published in April 2023 to provide scoping for the array of shared infrastructure examined here.²

In spring and early summer 2023, we conducted interviews with 49 infrastructure service providers, publishers, librarians, advocates, analysts, funders, and policy makers. Given the scope of this project, we focused our interviews on individuals from the European Union and Anglophone countries, with only limited outreach to individuals from, or with expertise on, Africa, Asia, and Latin America. Appendix A includes a list of the individuals we interviewed. We

---


will conduct several additional interviews prior to publishing the final version of this report in October 2023.

The interviews were framed around several key questions:

- What factors are driving change in scholarly communication and its infrastructure?
- What elements of infrastructure provision are working well?
- Where are there gaps or new opportunities for shared infrastructure?3

In addition to our formal interviews, we have benefited from project-specific advice from and conversations with a number of individuals, including IJsbrand Jan Aalbersberg, Laird Barrett, Steven Heffner, Rose L’Huillier, Hylke Koers, Clifford Lynch, Katherine Skinner, Eefke Smit, Todd Toler, Paul Tuten, Craig Van Dyck, Ralph Youngen, and Charles Watkinson. We thank Joanna Dressel for her substantial contributions in scheduling and coordinating our interviews and Kimberly Lutz and Mark McBride for their comments on an earlier draft of this paper.

STM Solutions provided sponsorship support that made this project possible, for which we express our gratitude. The research and analysis are solely the work of the project team members, and we accept all responsibility for it.

**Strategic Context**

In this section, we provide a high-level outline of some of the key strategic contexts faced by publishers and infrastructure providers. We acknowledge here, but do not further examine below, some of the broader factors at play, such as the macroeconomic uncertainty that was growing as we conducted this project. Our focus is on the sense of great opportunity—and substantial uncertainty, grave challenges, and diverging perspectives—amidst this second digital transformation.4

**Transitioning to Service Provision**

Scholarly publishing is in the midst of a substantial transition away from more traditional models of publishing and towards a service-provision model. Publishers fall into a variety of places along this spectrum, and the transition is occurring at a different pace in different parts of the sector. There is a rich complexity of factors driving this transition, including scholarly incentives and the rise of open access, as well as broader changes in the environment for digital content distribution and protection.

The strategic context includes the transformation of digital business models broadly. These include ad-supported free access (for example through YouTube), traditional subscriptions (for

---

3 See Appendix B for the full list of questions explored during the interviews.

4 As one important source for strategic context, we recommend the STM Trends reports, including the recent STM Trends 2026 report, “The Beauty of Open at Scale”: https://www.stm-assoc.org/standards-technology/stm-trends-26/.
example through Substack), and various kinds of hybrid models (for example through Spotify). Copyright enforcement can remain a tactic for content producers and distributors, particularly for piracy at scale. For customers, though, the winning strategy has proved to be business model innovation that responds to particular customer categories and market sectors. Scholarly publishers have been coming around to this mode of thinking, particularly as they consider how to respond to open access policy dynamics.

Open access is now a factor for a growing share of scholarly publications. In the United Kingdom and European Union, government and funder policy initiatives have driven change largely through gold-based models which have matured into transformative agreements that now seem likely to be challenged through rising interest in diamond and repository-based green open access models. In the United States, the policy landscape that will result from the Nelson Memo will also drive an increase in immediate open access, certainly for federally funded materials, although there is some contention about which models will be utilized for compliance. Other geographies include a range of models as well.

Business models that provide fairly direct rewards for growing article volumes have changed the marketplace in a variety of ways. Perhaps most importantly, the interest in sourcing articles has translated into a growing strategic focus on improving the author experience. This is a critical element of the “service provider” mindset, which several major publishers see as vital strategic priorities in terms of competitive differentiation. Additionally, transactional processes underpinning new business models introduce new kinds of risks and require an array of new platforms, skill sets, and organizational structures. Finally, the transition to these new business models has provided an opportunity for new entrants into the marketplace, bringing new forms of competition including, at least in some cases, on price. While new entrants enter the market, transformative agreements, in particular, are seen by many observers to continue to drive consolidation among incumbents, as described in greater detail in the following section.

This strategic context has also raised questions about the nature of the publishing services that are needed and alternative models for how to organize them. There is a long-standing effort to build repositories, at institutional, disciplinary, funder, and in some cases national or supranational, levels, to reduce costs and improve speed relative to Gold models. In some fields and geographies, these have been transformative, while in others they have thus far been additive.

---

5 There are several models for open access. Green is a version of the publication archived online not through the publisher. Gold denotes immediate open access by the publisher usually based on a publishing charge. Diamond refers to a publication model in which journals and platforms do not charge fees to either authors or readers. Transformative agreements are negotiated between institutions (e.g., libraries, national and regional consortia) and publishers to bundle together subscription and open access publishing fees with a goal to transition to greater levels of open access. More expansive definitions of these and other terms, and some discussion of them, may be found at: https://scholarlykitchen.sspnet.org/2020/04/07/seeking-sustainability-publishing-models-for-an-open-access-age/.


Most of these services offer more streamlined review compared with traditional editorial and peer review processes, but both traditional and repository-based models have a wide range of standards and processes. Low-cost publishing models and alternatives, including some diamond models and repository models, may yet prove to be a source of disruptive competition for the sector. Some community members contemplate a future in which publishers are not meaningfully distinguished from one another by brand and other factors but rather come to serve as contracted suppliers for open access publishing services, believing that this second digital transformation cannot be complete without this essential shift.

Notwithstanding these broad transitions, there are key exceptions as well. While the decline of print as a format for publications in many STEM fields is a foundational element of many of these transitions, print remains important for monograph publishing in the humanities and social sciences. The challenge of monographs going “out of print” in response to low levels of demand has eroded in response to new printing technologies. There nevertheless persist substantial questions about the extent to which monograph publishing will transition to open access, let alone some of these second order transitions that digital has enabled for other fields.

Consolidation and Competition
The strategic directions and marketplace dynamics of scholarly publishing are essential to understanding the landscape for shared infrastructure. For traditional publishing activities, this includes two major forms of consolidation that have swept the marketplace, even as there are several important new entrants that have brought renewed competitive vigor. Additionally, particularly for some of the largest publishers, strategy is as much around the development of a platform, analytics, or services business, as it is around scholarly publishing.

During the first digital transformation of scholarly publishing, there were a number of notable acquisitions and mergers, including Elsevier’s acquisitions of Academic Press and Cell Press, Wiley’s acquisition of Blackwell, and the deal that brought together Springer and Nature. This activity created the small set of largest houses that currently populate the landscape. Today, further acquisitions are most likely either to bring scale to open activities (witness Wiley’s purchase of Hindawi) or among other publishers.8

But another form of consolidation is very much still in play. Many independent societies have housed their publishing with some of the largest publishers. During the first digital transformation, this was particularly important because it allowed them to gain access to services such as journal hosting and other infrastructure as well as to a global sales team. During this second digital transformation, societies are seeking to access the umbrella of global transformative agreements that large publishing houses have negotiated. Experts believe it is

---

difficult for a society to leave their publisher partnerships, meaning that this form of consolidation may be no less permanent than an actual acquisition.9

Consolidation of both types leads some to forecast that only a dozen or so independent publishers will remain. Should this scenario play out, each could be expected to have an essentially separate technology stack that it develops itself. This could in turn lead to questions about the implications for certain types of shared infrastructure.

At the same time, there are new categories of competition. The transition to gold open access models created an opportunity for a new group of pure OA publishers, several of which have experienced skyrocketing growth rates through a combination of mega journals and more traditional offerings, almost all driven through article processing charges (APCs). While these new entrants use a variety of shared infrastructure, they are probably more likely to build aspects of their own infrastructure than similarly sized traditional publishers, given their distinctive OA-only needs.

Another new category of competition has been developing for over a decade, as several of the largest publishers, in particular, have been building platform, analytics, and services businesses to complement their primary publishing business. As a result, these publishers now also compete with pure-play shared infrastructure providers (such as Clarivate and Digital Science among others).

That said, the largest publishing houses are not by any means identical. Each has a strategy that is distinguishable from its peers, both because of the different nature of their publishing lists and even more importantly because of their different forays into platform, analytics, and services businesses. The result is that their interests, while sometimes aligned, in other cases are not. This has tremendous implications for where they can align as users of shared infrastructure.

University presses are an important type of publishing organization, especially for monographs. They have different kinds of organizational dynamics, as their university-based identity has tended to reduce otherwise natural opportunities to seek scale through consolidation. While there have been efforts made at cross-university press models, these have been exceptions that have not always proved sustainable. Several university presses have developed solutions businesses to serve one another in areas such as fulfillment, warehousing, and digital distribution. Others have been integrated into their universities’ libraries with the goal of aligning on values and strategy.

---

Humans and Machines

The longstanding purpose of scholarly communication—to support human readership and authorship—is itself in transition as we move to a future where more and more creation and consumption of the scholarly record involves machine-to-machine communication. Over time, infrastructure will be a key enabler of this transition.

Over the last couple of decades, substantial transformations have taken place in the ways knowledge is generated, discovered, accessed, shared, utilized, and repurposed. We will not reiterate those here beyond observing that, with the proliferation of research outputs generated from the investigation to dissemination stages, the scholarly record has become much more heterogeneous, variable, dynamic, and distributed.

Publishing is largely digital, but still to a large degree reliant on print-era structures and workflows. That said, the scholarly record may be atomizing into component elements such as data, code, and research methods. A publication may be understood to incorporate the article identified as the version of record, related to a preprint in a disciplinary server, and multiple datasets in various repositories, as well as code, methods, protocols, and so forth. Increasingly, medical and other fields are assessing the viability of “article extenders,” meant to recap the contents of research publications for non-specialist audiences via videos, infographics, or other formats. Ideally, all these individual components will be connected together to represent the complete set of outputs of a given research project. It should be noted that the atomization of the scholarly record in these component parts raises questions about assessing the quality, integrity, and impact of research at various levels beyond the version of record alone. Large publishers in particular believe that remaking the scholarly record for computational analysis and reproducibility is an important priority.

Another approach, which some see as complementary and others as an alternative, is text and data mining (TDM) access to the literature. Rather than decomposing it into a series of linked objects, a number of services and tools are developed to support non-consumptive analysis of the text by providing secure access that respects rights and aggregates at scale. While aggregators provide one set of services for such work by scholars, there is growing interest from the commercial sector, for example from pharmaceutical companies, to access such materials at scale for their own analytical purposes.

Paralleling the well-known surge in human-based production of knowledge has been the increasing practice of computational forms of content production and consumption. For example, the current wave of generative AI has numerous implications for scholarly communication. With respect to scientific research, generative AI will accelerate the transition away from human authorship of a version of record as the key output of a given research project. This has led to questions among publishers about how to monetize their publications as training data for commercial AI services, as well as how to integrate AI-based tools into their services. Generative AI will also raise an array of questions on ethics and integrity for scientists and publishers alike.
Publishers have been relying on AI-powered applications (and machine-learning) in publishing for a number of years to make scholarly communications better, faster, and more trustworthy (e.g., copyediting and proofreading). They have recently developed more sophisticated practices to bring efficiencies to the peer-review process and improve the quality of published research.

**Trusted Global Public Good**

Scholarship is, at least in its ideal form, a trusted global public good. It is a public good insofar as increasing access to knowledge does not diminish the value of knowledge to anyone else. It is available to all given our globally shared interests in pursuing societal goals and human understanding. It is high quality and trustworthy, generating broad public support for the expertise that it reflects. All this to say, it has become clear in recent years how far away we are from the ideal.

Sadly, some of the challenge comes from within. It has become clear that the incentives to commit academic fraud and misconduct far outstrip the consequences, and the result is a rash of manipulated images, fabricated data, paper mills, and other threats to the integrity of the scholarly record. Today, responsibility for addressing these threat vectors and investigating suspicious activity is often all too diffuse, with universities, publishers, and funders, among others, all having a role but rarely collaborating effectively. Maleficient actors using generative AI will introduce a new set of challenges. Quality publishers are investing in expertise, tools, and processes designed to block the vectors through which misconduct can otherwise enter the scholarly record, but as recent examples illustrate new business models operating at scale globally can nevertheless pose real challenges. Today, large publishers are existentially concerned about research integrity because they operate at scale, but other publishing organizations even if less exposed have a shared interest in ensuring trust in and authority of the scholarly record and the work of publishers and scientists.¹⁰

Although serious, fraud and misconduct are probably not the primary drivers of growing public mistrust in science and scholarship. In the polarized political environment in the United States and other countries, higher education and scientific institutions are among the organizations that are no longer uniformly trusted. This has manifested itself in conspiracy theories, fake news, and politicization of science, among other unfortunate outcomes. In the long run, taxpayer support for science, through public funding agencies, underpins a great deal of research activity and publishing. It is in the common interests of universities and publishing organizations to bolster public trust in these institutions and this expertise.

The era of globalization has been winding down for several years, particularly given splits that have taken place with China and Russia. These obvious divergences have placed limits on certain forms of scientific collaborations, either outright because of new regulatory frameworks

---

or indirectly through severe chilling effects. The extent to which global science and scholarly communication will hold up in the face of an increasingly fractious international environment remains uncertain.

More broadly, though, there are several different “geographies” for scholarly communication, for example the US, Latin America, Europe, and India, as well as China. Many providers and much scholarship readily cross these geographies but they do so operating within decidedly different policy environments. Beyond the policy environment specific to scholarly collaboration and scholarly communication, there are other types of regulatory divergences, for example on data, privacy, and artificial intelligence, which can serve as impediments to global solutions for publishing services and/or shared infrastructure. Also, local systems of knowing and knowledge-making are another related consideration, for example with respect to Indigenous knowledge, where the entire research and communication lifecycle needs to adhere to the principles of knowledge sovereignty. These can sometimes diverge from other practices and overall policy directions. Divergent collaboration and communication geographies and practices place real limits not only on the global reach of publishing organizations but also the global distribution of certain forms of knowledge.

Ultimately, at least some observers believe that there is a real tension between, on the one hand, providing a validated and trustworthy scholarly record and, on the other hand, providing an open system in which anyone globally can participate.

Some Notes on the Landscape

Publishing organizations and infrastructure providers alike are grappling with, and attempting to provide leadership for, this strategic context. In this section, we provide some observations both about the landscape of publishing organizations as well as the array of perspectives on infrastructure provision that we encountered.

Publishing Organizations

The publishing landscape remains rich with divergent interests despite the compact nature of the field. The field is dominated by several large publishers, many of whom have built their own underlying technological infrastructure to manage internal processes. Midsize and university presses can’t utilize the functionality of these systems, and the market isn’t robust enough to warrant extensive commercial investment and innovation for the needs of smaller publishers. Because serial publications command more market share than single titles, functionality for journals drives innovation more so than monographic publishing. The same is true for STEM publications, the needs of which are currently driving change more so than for humanities or social science publishing.

Large publishers also have the ability to mix commercial, in-house, and open infrastructure to create a best-in-breed approach as the landscape continuously shifts. We heard from several publishers that the adoption of open tools is often advantageous to their workflows to fill
specific needs or gaps in processes. Some are also evaluating open infrastructure components for the benefit of the field at large to increase transparency and functionality between systems. But it is most often that business needs, rather than altruism, drive adoption of open tools at this level.

Smaller, university, or mission-driven publishers have various constraints that often force a competitive rather than collaborative environment. Branding, especially for university presses, is critical to their success; this, however, impedes mergers that might help stabilize multiple small presses. These publishing houses are also often thinly staffed, which makes innovation difficult. In our interviews, several small publishers spoke of their challenges in managing their back-office functions, including title management and payments. Outsourcing some functions such as customer service may be an option, but this inevitably adds costs at a time when revenue is likely diminishing.

The rise of open access publishing has disrupted the landscape with new considerations. A number of individuals representing multifaceted perspectives spoke about the specific challenge of APCs, both from a payment and receival perspective. While larger and more agile publishing houses are able to respond to new challenges like collecting APC payments by establishing new business models and building responsive systems, small and university presses often struggle. Better infrastructure is needed to ensure open access is viable and sustainable. Several interviewees also commented that the current dominance of open access and the exponential need for its support have pushed other essential discussions to the background, including on the collective need for stronger linked data and the industry’s continued dependence on the PDF as a format.

Overall, there is broadly articulated alignment about critical directions within the publishing landscape. But on a practical level, a variety of divergent incentives and strategies are at play. This problem will only proliferate as research publishing becomes more international, especially around open access publishing and dissemination.

Perspectives

In scoping this project around “scholarly communication,” we recognize the array of publishing models and indeed communities at play. We have tried to include in our interviews, and listen sympathetically to, voices representing different models and communities. Based on these interviews, we believe it is no longer valid to frame the current fissures around open versus closed models. Ultimately, we did not find so much a clear set of coherently articulated schools of thought about shared infrastructure as we did a series of spectra on which perspectives varied significantly.

Everyone we interviewed expressed a strong belief in the importance of creating and sustaining a shared infrastructure for scholarly communication, but the commercial and strategic rationale for doing so varies, as does the actual sense of what elements of a shared infrastructure are most important. We spoke with employees of commercial firms, academic and research institutions,
not-for-profit organizations, and funding organizations, and encountered a range of perspectives.

Some interviewees were highly critical of capitalism and/or commercial organizations, at least with respect to scholarly communication and its infrastructure. Several are concerned about the profitability of commercial providers and worry that commercial interests diverge from those of the academy, of researchers, or of science.

As part of these conversations, some interviewees advocated for the creation of new modes of publishing that disrupt the traditional publishing ecosystem. They see these new modes—for example, a certain kind of repository ecosystem—as a key form of shared infrastructure. Others believe outright disruption to existing businesses is infeasible. They are looking instead for specific interventions that might dilute the long-term control of the scholarly record by commercial firms while working with these firms today.

We also heard about the importance of funding innovation within the academic and research sectors, even as interviewees acknowledged that these sectors on their own may not be able to provide the resources necessary to launch and sustain new infrastructure. Some of the people we interviewed are advocating for new workflows for research and communication—and platforms for facilitating rapid sharing of research and data to facilitate communication before publication—that would dilute the primacy of formal publishing and publication branding.

A number of our interviewees work for commercial, in some cases publicly traded, firms involved in publishing or infrastructure provision. They tended to focus on the importance of creating systems and processes at scale, marketplace dynamics, and the strategic directions of science and the academy. Interviewees from traditional publishers spoke about their own commercial and strategic rationale for creating and maintaining a common infrastructure. Many commercial providers are comfortable thinking as community members interested in governance, representation, and business model issues for the shared infrastructure, at least for the portions of it that benefit them.

Most interviewees from disruptive and academy-owned publishers also described their own commercial and strategic rationale for creating and maintaining shared infrastructure. While some noted preferring certain forms of governance or business models for this infrastructure, they focused mostly on that strategic and/or commercial rationale given their own goals in the marketplace.

In considering the shared infrastructure, some interviewees, particularly those from the academy or other not-for-profit organizations, express a particular focus on the values that underpin this infrastructure (beyond business requirements). We heard from them about the importance of values and principles such as transparency, open systems, and a form of academy or non-commercial community control.

“Quality” as a principle in scholarly publishing has profound implications for shared infrastructure. In some cases, interviewees believe strongly in quality as a core value of certain
forms of publishing. Others express a strong view that quality cannot be determined objectively or that when invoked it serves too often as a mechanism of inappropriate exclusion, with some feeling that as a result nearly any contributions should be welcomed into the scholarly record.

We heard several different perspectives on privacy and anonymity in terms of data practices. Some feel very strongly that absolute anonymity is the only position that adequately protects users. Others feel that appropriate safeguards can be put into place to protect user privacy while adding new services that they believe users expect from their broader digital experiences.

Several interviewees take a decidedly globalist perspective, believing that science cannot respect borders. Other interviewees expressed concerns that, whatever their own preferences, global science and scholarly communication may be fracturing.

Several interviewees used the term “open science” but hold different definitions for the term. For some, it means making research products openly available and reusable for everyone and making the research process transparent early on in the process. For others, the emphasis is on increasing international and cross-institutional scientific collaboration to address societal problems and communicating scientific outcomes to the public beyond the traditional scientific community. For others, it is about extending free and open principles from open access and open source into other parts of the scientific process.

Finally, some interviewees believe that identifying the common layer of shared interests is the best starting point for creating shared infrastructure, even if that results in somewhat of a least common denominator outcome. Other interviewees are interested in building “coalitions of the willing,” in some cases to challenge incumbents and drive change or other cases to carry forward what is seen to be a pressing community priority.

While there is some clustering together of groups of these views, we were not able to identify two or three or four coherent overall schools of thought on shared infrastructure. Later, we will suggest our view that different models make more or less sense for different parts of the shared infrastructure.

Shared Infrastructure Today

The foregoing strategic context and landscape informs our understanding of the needs for shared infrastructure. Next, we turn to an examination of the current state of shared infrastructure. We make no claim to an exhaustive treatment here. Instead, we examine key strategic issues in several important categories of shared infrastructure. We studiously avoid assessing individual infrastructure elements or providers, as that is outside the scope of our project, but rather attempt to identify key themes. Our objective is to inform an understanding of what is working well and where there may be limitations.

At a high level, shared infrastructure for scholarly communications is extensive and impressive. It addresses many foundational needs of publishing organizations and leverages collective action, while its providers serve as vehicles for common (or in some cases outside) investment.
Identifiers and Standards

As scholarly communication increasingly becomes more complex, the need for common identifiers and standards continues to grow in order to facilitate discovery, access, linking, rights management, and assessment of scholarly content. Starting with object identifiers, and more recently identifiers for researchers and research organizations, the research community has benefited tremendously from the availability of what are generally termed persistent identifiers (PIDs).

PIDs are provided by an array of organizations. For objects, the basic standard has become the digital object identifier (DOI), which is stewarded by the DOI Foundation. Several community-governed not-for-profits such as CrossRef and DataCite, as well as several national consortia, issue DOIs and provide a variety of enabling and related workflow and metadata services. Somewhat more recently, proprietary efforts to provide interoperable researcher identity gave way to the ORCID, which is stewarded and issued by the ORCID organization. Today, there is widespread agreement about the need for an identifier for research organizations, but no single model has achieved consensus. A well-developed commercial provider was acquired by the Copyright Clearance Center, while other actors are developing standards that are promoted as being more open, but any long-term convergence remains unclear.

Membership organizations like CrossRef, DataCite, and ORCID are prominent in identifier work, and they are joined by somewhat similar kinds of organizations that develop and steward both formal standards and community-wide recommended practices. We group them together because they have some key dynamics in common.

First, several of these organizations were founded with a strategic purpose—or a strategic constraint—for example to drive interoperability across publishing platforms in ways that would take away strong network effects advantages for the largest publishing houses or to avoid interfering with current or future business directions of key sector players. Baked into their early history is that they are more than just neutral stewards of PIDs but rather enabled or constrained with a strategic purpose. Some of the founding members may have forgotten these dynamics but they remain just under the surface in understanding contemporary dynamics and opportunities.

Further, whoever the original founders of these membership organizations, they were developed with the expectation, or at least the ability, to serve the widest possible array of members and communities. To take just one example: despite the challenges posed by different geographies, to say nothing of geopolitical tensions, there remains a dream of a single identifier encompassing every researcher across the globe. Several PID organizations have been characterized by growth not only in the number but also the type of members, enabling the PIDs they issue to become widely accepted standards while at the same time leading to some foundational questions about organizational purpose, governance, and business models. For example, if members pay widely divergent fees, should each receive the same weight in voting? Or, as member-controlled not-for-profits, to what extent do PID organizations exercise power to
keep members on board or, alternatively, should they ensure that it is easy for members to depart?

Some large publishers express concern about “losing control” of governance for organizations that they had originally established, when their developing strategic needs are, as they see it, insufficiently prioritized by these organizations. Several interviewees wondered whether it remained in their organization’s interest to maintain membership in community identifier organizations.

Specifically, as these PIDs have become more widely adopted and their governance has broadened, are these organizations evolving rapidly enough to meet publishing organizations’ emergent strategic needs? For example, publishing house representatives expressed that identifier organizations should provide services that advance research integrity, for example a trusted digital identity for researchers.

Since each identifier organization operates and sets direction independent of one another, there is also no shared, common direction for organizations to move forward collectively. There are also some concerns about the governance bandwidth and overhead burdens imposed by multiplying membership organizations. These considerations help to explain why several recent identifier and infrastructure initiatives have been set up inside of one or more existing organizations. Even so, several individuals questioned whether more was needed, specifically whether a model of shared governance or services might act as an umbrella over multiple projects with similar purviews to mitigate overhead and provide common objectives.

**Enterprise Publishing Systems**

Publishing systems offer integrated workflows for authors, editors, and designers to facilitate content management and distribution. Some systems specifically focus on facilitating the submission and processing of manuscripts and managing the review process while others provide extended functionality to support delivery, distribution, analytics, and e-commerce. In recent years, some publishing systems have been expanded to allow for newer features such as preprint deposit and dataset submission and review. Platforms that couple hosting, publishing, and delivery are deployed in many scholarly societies, university presses, and other publishers. Nevertheless, there is no turnkey solution that can handle all the requirements of every publisher as the technical infrastructure, priorities, and resources vary.

Although there are a number of enterprise publishing systems in the marketplace, large publishers tend to develop and maintain customized in-house systems to support their large operation. Such internal investment enables them to position services such as manuscript submission and editorial management for competitive differentiation. On the other hand, the majority of mid-size or small publishing operations rely on off-the-self publishing platforms in order to control in-house maintenance and development costs.

Since the early 1990s, with funding from foundations or governmental agencies, several open-source publishing and hosting platforms have been developed with the goal of reforming
scholarly publishing. This category of publishing systems aims to promote and enable community-based, collaborative, and academically-driven publishing and introduces new workflows for content creation and review. The sustainability of these systems is based on adoption rates and successful fundraising efforts.

On the commercial side, most of the innovation has been introduced by start-up initiatives that attract investors, who provide the capital for development. If these start-ups are successful, they mature into service providers with the potential to achieve substantial market share and revenue. As the market share of an enterprise software provider develops, their publisher clients often find that they become “monolithic,” in the sense that the dependencies of a vast array of clients impede strategic innovation either by the enterprise provider or, as a result, among the clients. Typically, strategic innovation is enabled by a new enterprise platform—either from an existing provider willing to disrupt its own product line or a new entrant. But before this happens, the general trend is for large publishers or other commercial organizations with a strategic interest in the business to acquire mature enterprise providers, raising unavoidable questions from clients about their independence and neutrality.

Some small or not-for-profit publishers feel that the overall industry suffers from a lack of innovation because of its small size. While major players can develop and maintain systems for advanced features, they feel the pressure to utilize cost-efficient out-of-the-box solutions, leaving a potential gap in innovation. Regardless of the publishing system innovation coming from the commercial or not-for-profit sector, existing legacy systems, editorial workflows, and hosting platforms often can slow down adoption and transformation at local level.

Many publishing organizations are agnostic about the innovation of any one particular provider. Indeed, we heard directly from some interviewees that their goal is decidedly not to sustain individual infrastructure providers but rather to ensure that they have access to needed infrastructure, a pointed distinction. The key factor for them is to ensure that market forces are adequate to enable a disruptive innovator to compete. Unfortunately, today there is some concern that such market forces are not working all that effectively, and some publishers feel at least somewhat trapped by existing providers. Additionally, as they adopt new systems, smaller independent and society publishers tend to lack economies of scale to afford and sustain new system deployment.

For smaller publishers, supporting back-office functions is a big challenge as there are different platforms and title management systems for different audiences (e.g., libraries). For instance, for some university publishers, insufficient integration between their payment systems and a university’s financial framework makes it complicated to manage the revenue generation process. For smaller or mission-driven publishers, there are opportunities to outsource administrative support, but each new organizational layer is likely to add new costs. Some have to operate multiple publishing platforms as no single system meets all their needs, especially for publications that are paywalled or open. For instance, publishers that need to fulfill the funder OA mandates need to use a system that will support not only chapter-level but also volume-level creation of PDF version. Publishers in the textbook sector may need a supplemental system that allows rentals and low-level DRM use. In terms of publishing workflow support systems, there is
no perfect solution yet, but small publishers are looking for predictable hosting fees, good customer service, and a system that can be managed internally using existing resources/expertise.

The strategic context of consolidation is a particularly important factor for enterprise systems. On the one hand, a strong marketplace for shared enterprise systems makes it possible for smaller publishers to thrive, and new publishers to enter the marketplace, without needing to create their own publishing platforms. In this sense, enterprise systems are a vital element of ensuring a competitive marketplace. But a downward spiral is also a real risk: a more consolidated marketplace reduces demand for shared enterprise systems and in turn the ability for their providers to make the investments necessary to enable the remaining independent publishers to resist consolidation.

**Discovery, Collaboration, and Trust**

In this section we explore a set of categories of the shared infrastructure where we see particular opportunity for innovation and disruption alike, including the discovery of scholarly materials, scholarly collaboration, efforts to establish trust and broadcast the trustworthiness of scholarship, and methods for analyzing the impact of scholarship. The strategic importance of this infrastructure is that researchers—scholars, students, and indeed machines as well—need to be able to discover the content they need, gauge its trustworthiness before utilizing it, and then collaborate in incorporating it into future scholarship. And publishing organizations have a strong incentive to maximize the discovery—and, they therefore hope, measures of usage and impact—of their publications.

In the discovery category, an array of commercial providers provide both “library discovery” and “research discovery,” along with several important non-commercial offerings as well. Research discovery is often integrated into or offered alongside various impact metrics, which are also in some cases integrated into library discovery. Some services offer tools that are intended to provide discovery of “everything” a user might reasonably seek, while others are bounded based on analyses of quality, openness, or other characteristics.

Content discovery and assessment are critical functions in the researcher journey and there is tremendous value in providing these services. Many organizations also express serious concern when a competitor, or a party with divergent interests, controls these functions. Some of the recent efforts to ensure the integrity of the scientific record have exposed the market tensions that ordinarily lay further under the surface. Shared infrastructure, and the interconnectedness of research content through a network of PIDs, has tremendous benefits for researchers but also introduces complexity. For instance, if error exists in metadata, it is difficult to rectify once the error has proliferated across systems and unclear whose responsibility this may be.

A newer model has grown through what are sometimes called scholarly collaboration networks. The current offerings in this category have built what can be understood as researcher-centric platforms, with a core focus on the researcher experience that generates traffic and engagement that can ultimately be monetized not only through advertising but also by providing services to
publishers (or competing with them), as well as other parties interested in research and researchers.

In recent years, it has become clear that the fractured nature of the publisher-specific websites has been an impediment to a strong user experience. Instead, there is a growing effort to place the content where the users are, to drive a more seamless discovery to access experience. Syndication models and some of the standards that enable them are a major driver. While it is still early days in how this overall shift will play out, it is possible to imagine substantial opportunities to bring together discovery, trust markers, access, and potentially even collaboration through the same platform experience.

The existing infrastructure services described above are provided through strikingly different service and business models. Some are provided as or closely connected to enterprise software bundles sold to academic libraries. Others rely strongly on brand, particularly for impact measurement, and are sold as databases to academic libraries. And scholarly collaboration network platforms are monetized through publisher charges for what is essentially content marketing or through advertising models.

As mentioned, we see a particular opportunity in this space for innovation and, perhaps, disruption. New forms of AI will lead to new ways of understanding the purpose and nature of search, discovery, and academic collaboration. And, it will be essential to address integrity, trust, and authority in ways that respond not only to changes within science and scholarly communication but also external threats and risks as well. Existing infrastructure providers in these categories have access to content, which is an absolutely foundational requirement but not one that has, in recent years, impeded new competitors from entering the marketplace.

Preservation

Preservation relies on trusted, sustainable environments to provide critical functionality for the long-term stewardship of research outputs. Although enduring access has a fundamental primacy in scholarly communications, the concept of preservation as necessary and shared infrastructure is often overlooked or taken for granted. The interviewees rarely mentioned the role of infrastructure in supporting long-term access to digital content. Yet concerns persist about the stability of digital preservation across the landscape and preparedness for accommodating different types and formats of publications. More commonly adopted, shared infrastructure could both mitigate some concerns and be more agile for future development as scholarly communication continues to evolve. Some disciplines are also clearly better prepared and organized in this regard.

The libraries and archives took the lead in ensuring the preservation of print cultural heritage through selection, acquisition, and long-term storage—in some cases with redundancy by making multiple copies of materials available at different institutions. The digital sphere requires different models as maintaining a centralized infrastructure is not feasible due to the characteristics of digital content, which is both pervasive and ephemeral at the same time. The challenge is not only preserving the bits of digital objects but also being able to transition over
time their affordances, software environment, and the context required for interpretation and consumption.

There are several successful initiatives and services to support the preservation of scholarly journals and monographs, including at an article and chapter level. That said, there continue to be gaps in the service infrastructure for addressing the requirements of other content types, including those that are emerging or will derive from the atomization of the scholarly record. Digital preservation remains a critical challenge for many institutions, but especially for cultural heritage organizations from low- and middle-income countries. To this end, international, national, and regional preservation services, advocacy organizations, and technology initiatives continue to play a critical role as needed infrastructure.

Reflections

There is an enormously rich terrain of shared infrastructure, as we documented in our landscape review and analyzed in the sections above. The providers of this infrastructure, as well as those who fund and govern it, can take real pride in this achievement. We have studiously avoided addressing individual providers and their particular strengths and weaknesses as such an assessment is outside our scope. Instead, in this section, we offer some overall reflections on the current state of the shared infrastructure.

Different categories of shared infrastructure seem to accrete providers with discrete blends of organization types and governance models. In some cases, we believe there is a basic logic to the current models, for example the membership organizations that provide identifiers and standards, as well as the commercial start-up ecosystem for enterprise systems which periodically disrupts the status quo. In other cases, such as in the areas of discovery, trust, and impact, the infrastructure is unstable or still being formed, and the appropriate governance and business models have not yet necessarily been accepted across the community.

At the individual provider level, we note that few infrastructure providers have achieved some of the key characteristics that stakeholders would like them to entail:

- **Trustworthiness** among clients and other stakeholders,
- **Financial returns** that yield breakeven or sufficient profit margin, depending on the model, to ensure sustainability;
- **Access to capital**, i.e. working capital for maintenance and growth capital for additional development;
- Creative **innovation** in response to emergent needs; and
- **Agility and flexibility** as the market changes.

To be sure, not all providers may need to possess all five of these characteristics in equal measures. For example, a start-up enterprise systems provider may optimize differently than a community governed PID provider.
In some geographies, including North America and Europe, most—not all—of the infrastructure we have examined is paid for, in whole or in part, by publishing organizations. On some level, it makes good sense for publishing organizations to pay for their own infrastructure. On the other hand, like any other cost, this is a kind of barrier to entry to the marketplace. We note this because in some countries, the government pays for a substantial portion of this infrastructure, enabling domestic publishers to focus on adding value on top of it.

New Opportunities for the Shared Infrastructure

In the above section, we examined some of the challenges and opportunities for key categories of the existing shared infrastructure. In this section, we build upon the strategic context discussed above to examine opportunities to create new categories of shared infrastructure. In some categories, shared infrastructure will be highly contested as there will be meaningful strategic consequences from the choices that are made.

Spine of the Scholarly Record

Many observers expect to see traditional articles atomizing into their component parts in ways that will enhance machine-to-machine communication. We are already seeing the growing importance of dataset deposit, as well as indications that research protocols, code, and other elements may themselves become first-class research artifacts. The vision is that ultimately these various research artifacts can be re-used and recombined to advance not only replicability but also new research and interdisciplinary collaborations.

Up to now, we are mostly seeing different services used for the deposit of these different types of artifacts. For example, datasets end up in various repositories rather than “in” the journal publishing the article itself. There is a need for a set of standards, and possibly a technology layer, to help link all of these elements together. It is probably not just a matter of a link or DOI from an article to a dataset as it is. Instead, we may see the need for the elements of the scholarly record to be persistently connected with one another and actively managed through versions, corrections, retractions, and the like. Ultimately, we may see the need for a new way of understanding and visualizing the scholarly record.

Publishing organizations tend to envision that the version of record will link out to each of the constituent elements. Even if the version of record grows shorter as its component parts end up elsewhere, in this vision it still maintains its value as the element of the scholarly record that undergoes peer review, that counts for academic career advancement, and that ultimately serves as the spine of the scholarly record. It is understandable that publishing organizations would envision this model, given that it maintains the centrality of their brands and businesses.

Others might see the world differently. For a funder, for example, the key unit of measure is the funded research project. The funded project typically will have several different articles resulting from it, as well as datasets, various coded programs, a number of methods, and so forth. Each of these has a many-to-many relationship with one another. And of course the researcher may
instead see themself as the organizing unit of measure, no matter how collaborative their research may be.

If the foreseen atomization of the research article comes to pass, there could be opportunities to rethink and possibly even disrupt the nature of scholarly record and its leading custodians. The shared infrastructure that would support this atomization would have an embedded business logic that could favor one type of party or another.

This is ultimately a tremendous innovation opportunity and vehicle for disruption that no publishing organization has the ability to address on its own.

**Research Integrity**
The publishing community has woken up to concerns about research integrity. Shared infrastructure in support of research integrity to provide confidence and trust in the methods and the findings of the research is being developed both by enterprise software providers, particularly those that work on manuscript submission and editorial review, as well as on a shared basis through community organizations. Existing infrastructure and that currently in development will address some but not all threat vectors to research integrity.

The issue that we identified as the biggest gap today is the perceived need for a secure digital identity for legitimate scholars, to help editors triage submissions into more and less trusted categories. We see opportunities for researcher identifiers to be used as the hub for much greater information about digital identity, in part by allowing publishers and other parties to submit markers of identity into identifier records. As examples, publishers that have processed APC transactions using credit cards have substantial signs of verified identity, as do universities that have securely linked an email address.

The boundaries of the scholarly record represent another aspect of research integrity that requires new forms of infrastructure. Of course the record has never had absolute boundaries. But in a subscription landscape, libraries played an important role in establishing the metes and bounds of the scholarly record (and what would be preserved over time) based on their selection decision-making. In a gold or diamond open access environment, libraries may have a reduced role and so other forms of boundary-setting may be required. Journal rankings may increasingly serve to set the boundaries of the scholarly record, although whether that is the right form of shared infrastructure, or whether it has the right governance and business model to allow it to serve this role without fear or favor, is not yet settled.

**Making Meaning**
Research publishing shares scholarship within a community of peers. Scholars are trained to read the literature with a critical eye and to recognize that any single new study is part of a broader tapestry. Basic literature reviews, systematic reviews, and meta-analyses are among the formal techniques that are used to draw meaning from the literature. There are vast
opportunities to make meaning more readily from scholarly literature, both for scholars themselves and for a broader public.

This developing infrastructure category has connections with several existing services and features and is in some sense an extension of existing discovery techniques. It could be compared with the clinical decision support services that are widely used in medicine but extended to other fields and/or audiences. And it surely could build in some cases on the new generation of generative AI. One critical element of this category is that it must bring together publications in a single field or set of fields, across all publishers active in that field, in order to be successful.

While we believe this category will be valuable for scholars themselves, we think it may be of greatest importance for a broader public. In an information ecosystem that is driven by troubling degrees of political polarization, trusted interpretations and translations of science and scholarship to inform current awareness and policymaking is absolutely vital. While some scholarly associations see scientific communication as an important element of their work, the research publishing sector as a whole tends to see this work as outside its ambit. And yet the environment of misinformation and declining trust in science means that scientific communication is not just important for its own sake but as a key element in the political process that ultimately leads to public funding of science itself.

Supporting New Business Models
The transition to open access has introduced new business models for scholarly publishing. As publishing organizations have developed and implemented these new business models, they have discovered that certain key elements of their shared infrastructure are incapable of supporting them. Shared infrastructure was, unwittingly, designed to support the subscription model only. As a result, publishing organizations have been faced with a dilemma about how to provide the infrastructure necessary to support open access business models, and in some cases they have been forced to utilize manual processes as a stopgap. There remain substantial opportunities for the development of shared infrastructure that would support and enable open access business models.

One major category of infrastructure would support the transfer of money necessary to pay for publishing services. This includes systems necessary for individual universities to bundle together their resources to pay for consortial transformative agreements. It also includes the systems necessary to tie together individual institutional customers of transformative agreements with the workflows for manuscript management, where authors unaffiliated with such institutions are otherwise charged an APC. And it also includes the related systems necessary to support the multi-payer model that bundles divergent resource budgets together within a single institution to pay for transformative agreements. Major publishers (and in some cases universities) have invested in building their own systems to enable such services, some of which are mere stopgaps and others of which serve as a point of competitive differentiation.
There is also infrastructure enabling libraries to analyze and manage their collections budgets to help them optimize acquisitions, holdings, impact, and spending, all of which supports the transition to open access (as well as other institutional priorities). There are some existing products in this category but there are unmet needs as well.

**Interoperability**

Finally, while not a single category of infrastructure, we wish to note the need to think in terms of greater interoperability for existing and new types of infrastructure.

Different elements of the infrastructure can be somewhat siloed as a result of the specific needs of each. There are opportunities for greater attention to interoperability across competing infrastructure providers and infrastructure categories.

In addition, thinking in global terms, we note that global infrastructure will not always be possible due to geopolitical dynamics and values mismatches. Opportunities may exist for increasingly sophisticated forms of interoperability across geographies.

**Recommendations**

[We expect to provide recommendations in our final report when it is published. In the meantime, we invite you to comment on this draft report. Please send your suggestions and questions to sr-infrastructure@ithaka.org by August 31, 2023. The final report will be published on the Ithaka S+R website in October 2023.]

**Appendix: List of Interviewees**

- Clare Appavoo, Executive Director, Canadian Research Knowledge Network
- Allison Belan, Director for Strategic Innovation and Services, Duke University Press
- Laird Barrett, Head of Product, Springer Nature
- Amy Brand, Director, MIT Press
- Rachel Bruce, Head of Open Research, UK Research and Innovation
- Johannes Buchmann, Chief Operating Officer, De Gruyter
- Adrian Burton, Director, Data, Policy, and Services, Australian Research Data Commons
- Matt Buys, Executive Director, DataCite
- Ana María Cetto, Research Professor, Universidad Nacional Autónoma de México
- Angela Cochran, Vice President, Publishing, American Society of Clinical Oncology
- Raym Crow, Senior Consultant, SPARC
- Chris Freeland, Director of Open Libraries, Internet Archive
- Nicko Goncharoff, Managing Director, Osmanthus Consulting Ltd
- Joshua M. Greenberg, Program Director, Digital Information Technology, Sloan Foundation
Appendix: Interview Questions

1. What do you see as the most important drivers of change in scholarly communication?
   ▪ What are the key trends in [specialty area of interviewee]?
   ▪ For example, trends like
     ◦ Growth of new form of content such as datasets, code, methods, and implications for DOIs
     ◦ Protecting the integrity of the scientific record
     ◦ Widespread shift to open
     ◦ Increasing importance of computational forms of content consumption and content production
     ◦ Divergent demands from users for such features as providing seamless discovery and access; broadening participation through movements like “citizen science; producing translational outputs that meet the needs of the general public
   ▪ Successes registered and roadblocks and impediments

2. What kinds of shared infrastructure do you rely upon in your role? What kinds of shared infrastructure do you see as most important in the scholarly communications sector?
   ▪ Prompt with categories of shared infrastructure drawn from the landscape review (ie Assessment, Metrics, and Reporting; Authentication and Authorization; Discovery, Syndication, Hosting, Delivery, and Aggregation; Licensing and Rights Management; Manuscript Submission and Editorial Management; Metadata; Peer Review, Annotations, and Commenting; Persistent Identifiers; Preservation; Publishing Platforms and Repositories; Research Information Management; Researcher Identity and Marketing; Research Data Curation and Management; Usage Data)

3. What is working especially well in this shared infrastructure?

4. What are some of the challenges and pain points that you see with this infrastructure provision?
   ▪ Are you satisfied with technical abilities; product strategy, governance, business model, etc.

5. What kinds of infrastructure doesn’t yet exist but would be valuable to your work or goals?
• You mentioned XXX strategic direction. Do you see infrastructure needs associated with that?

6. Do you have any other comments or suggestions to inform our study?