| Method | Evidence | Discussion |

Outcomes

PROCESS

AFTERMATH

| Discussion | Revision | Re-use |

The Evolving Scholarly Record

By Brian Lavoie, Eric Childress, Ricky Erway,
Ixchel Faniel, Constance Malpas, Jennifer Schaffner
& Titia van der Werf

OCLC® Research

The Evolving Scholarly Record

Brian Lavoie, Eric Childress, Ricky Erway, Ixchel Faniel, Constance Malpas, Jennifer Schaffner & Titia van der Werf

#scholrec

OCLC Research

Contents

Figures

Acknowledgments

Introduction

The ways and means of conducting scholarly inquiry are experiencing fundamental change, with consequences for scholarly communication and ultimately, the scholarly record—the curated account of past scholarly endeavor. The scholarly record is evolving into a corpus of material vastly different from its previous print-based version. While in the past the scholarly record was largely defined by the formally published monographic and journal literatures, its boundaries are now both expanding and blurring, driven by changes in research practices, as well as changing perceptions of the long-term value of certain forms of scholarly materials.

Understanding the nature, scope, and evolutionary trends of the scholarly record is an important concern in many quarters—for libraries, for publishers, for funders, and of course for scholars themselves. Many issues are intrinsic to the scholarly record, such as preservation, citation, replicability, provenance, and data curation. Often these issues must be discussed and resolved across a range of stakeholder groups. With this in mind, OCLC Research has developed a conceptual framework that will help organize and drive discussions about the evolving scholarly record, by providing a high-level view of the categories of materials the scholarly record potentially may encompass, as well as the key stakeholder roles—and configurations of those roles—associated with the scholarly record.

A framework of this kind can serve as a common point of reference in discussions surrounding the scholarly record, by introducing shared concepts and terminology that promote mutual understanding and consensus around the broad features of the scholarly record.[1] This can help support discussions about the scholarly record within domains, and also—and perhaps more importantly—across domains. A framework can help knit together the fragmented strands of work addressing various aspects of the scholarly record into a cohesive whole. It can also equip libraries, publishers, funders, scholars, and other stakeholders with a resource to support strategic planning around issues associated with the scholarly record.

Shared understanding and collaborative relationships are especially important in regard to the scholarly record, because the transition from print to a digital, networked environment likely means that decision-making around the scholarly record will have to become more consciously coordinated. The broader range and greater volume of materials now perceived to be relevant to the scholarly record means that no single institution can hope to gather and

manage it all—or even a significant share of it. Therefore, perpetuation of the scholarly record is likely to become a much more collective and deliberate enterprise, with more tightly integrated and explicit roles and responsibilities. A common view, in the form of a framework that conceptualizes the directions in which the scholarly record is evolving, would be a valuable starting point for shaping the contours of these new decision-making arrangements.

Defining the Scholarly Record

The Nobel-prize winning scientist Francis Crick, when asked to define molecular biology, responded that molecular biology is whatever interests molecular biologists (Teitelman 1994, 183). While the boundaries of a discipline are, in a sense, set by the directions its researchers take them, a definition of this kind is nevertheless not very helpful from the practical standpoint of providing an understanding of the basic ideas of molecular biology. In the same way, when we ask "what is the scholarly record?" we might be told that the scholarly record is whatever interests scholars. Here, as with molecular biology, we might accept that this response has some validity, but it does not take us far in understanding the kinds of materials that form the scholarly record.

Ross Atkinson (1990, 356) provides a more helpful definition of the scholarly record, defining it as "that which has already been written in all disciplines . . . that stable body of graphic information, upon which each discipline bases its discussions, and against which each discipline measures its progress". This definition offers an eloquent conceptualization of the scholarly record, but is nevertheless resistant to practical application, in that it does little to establish boundaries around the specific kinds of materials the scholarly record might encompass.

Instead of a "top-down" conceptual view of the scholarly record, we could instead take a "bottom-up" approach and enumerate the specific types of materials the scholarly record might include. But this quickly becomes a challenging task, as the range of candidate materials seems to stretch on endlessly: print books and journals (and their electronic equivalents), data sets, computer models, blog postings, e-prints, interactive programs, complex visualizations—and many more too numerous to mention. To complicate matters further, the boundaries of the scholarly record also depend on the perspective that particular groups of stakeholders bring to bear on it. For example, a young faculty member interested in establishing their credentials might view the scholarly record as the portion of scholarly materials relevant for a tenure review. That same faculty member in the role of researcher might view the scholarly record as any material that is useful in furthering their research interests. A publisher may view the scholarly record as those materials that have been made

available through a formal publication process, including peer review and professional editing, as well as dissemination via an established communications channel like a journal or book. A library, on the other hand, might view the scholarly record as those scholarly materials that have been systematically gathered and organized into collections for long term use.

Even the question of what is scholarly is an open one, and reasonable people can disagree: for example, the International Studies Association recently announced that editors of their journals should not engage in blogging, because the blogging environment often falls short of the professional standards the ISA wished to promote, coupled with the risk that personal blogs might be confused with professional blogs (Straumsheim 2014). The proposal was later tabled.

In the end, precise definition of the scholarly record is a difficult if not impossible task. Therefore, a framework that proposes to conceptualize the content of the scholarly record must find a suitable middle ground between one that is too stratospherically abstract to be of practical use, and one that embodies such specificity that it provokes argument rather than consensus, and is too rigidly defined to permit cross-disciplinary application and future evolution.

The Scholarly Record: Evolutionary Trends

In addition to basic definitional issues, a range of evolutionary trends are also shaping the scholarly record. First and perhaps most obvious, we are witnessing a shift from what was traditionally a print-centric scholarly record to one that is increasingly manifested in digital form and resides on the network. Second, the boundaries of the scholarly record are shifting and blurring. While in the past we might have thought of the scholarly record as consisting primarily of text-based materials like journals and monographs, today the cohort of materials over which the scholarly record can potentially extend has expanded dramatically, to include research data sets, computer models, interactive programs, complex visualizations, lab notebooks, and a host of other materials.

Third, some of the fundamental characteristics of the scholarly record are changing. At the risk of oversimplifying an admittedly nuanced point, one might characterize the scholarly record of the past as largely static, in that much of it was manifested in fixed formats like print; it was made available primarily through formal publication channels like books or journals; and its focus was on the documentation of final outcomes, rather than the entire process of scholarly inquiry. Currently, however, these characteristics are being turned on their heads: the scholarly record, by virtue of its transition to digital formats, is now much more mutable and dynamic than in the past; it is made available through a blend of both

formal and informal publication channels; and the scholarly record's boundaries are expanding to include a much wider context surrounding the publication of a scholarly outcome. This last point is driven in part by an increased emphasis on replicability of scholarly outcomes, as well as higher expectations around what might be termed "leverageability": that is, the ability to take previously published work and integrate it seamlessly into new work.

Finally, another trend of note is the reconfiguration of the stakeholder roles associated with the scholarly record. The pathways by which materials comprising the scholarly record are created, managed, and consumed are changing in a variety of ways, with traditional stakeholders taking on new roles, and new stakeholders taking on traditional roles. The scholarly communication "supply chain" is evolving in concert with the scholarly record itself.

These trends, as well as others,[2] are hastening the scholarly record along an evolutionary path that promises to transform our view of the nature and scope of the scholarly record, as well as the configuration of stakeholders roles associated with it. However, this is not to suggest that the scholarly record was previously in some kind of stationary equilibrium, and has only recently been subject to transformative forces. The scholarly record is always evolving, and there always have been, and always will be, issues and debate over its definition and scope.[3]

While we are not experiencing a sudden revolution in the development of the scholarly record, we are experiencing a particularly emphatic confluence of trends that are accelerating the evolutionary process that is changing the boundaries of the scholarly record, as well as the configurations of the stakeholder roles associated with it.

Conceptualizing the Scholarly Record

Given current trends in the evolution of the scholarly record, OCLC Research developed a conceptual framework that is intended to promote understanding of what the scholarly record looks like today, and may look like in the future, as well as characterize changes in the configurations of key stakeholder roles associated with the scholarly record. The framework is divided into two parts: the first addresses the content of the scholarly record, and the second focuses on the stakeholder ecosystem associated with the scholarly record.

Content of the Scholarly Record

Figure 1 presents a conceptual view of the evolving scholarly record. At the center of the picture are published outcomes[4]—the traditional coin of the realm for scholarly inquiry. These are the materials that constituted the scholarly record as it was traditionally understood,

through which scholars reported their ideas and findings, and on the basis of which received credit and attribution; that were formally published and disseminated through established scholarly communication channels; and that libraries collected and curated for long-term access and use. Even as the scholarly record evolves in terms of scope and content, published outcomes are still of primary importance, and therefore are appropriately positioned at the center of the picture. Traditionally, most of these outcomes were text-based—for example, journal articles and books—and indeed these materials are still important today. However, it is not uncommon to see them augmented—and in some cases displaced—by other forms of outcome, including video, data sets, interactive programs or complex visualizations. But as figure 1 illustrates, the boundaries of the scholarly record are stretching even more, beyond the traditional focus on published outcomes, to encompass materials generated in the process
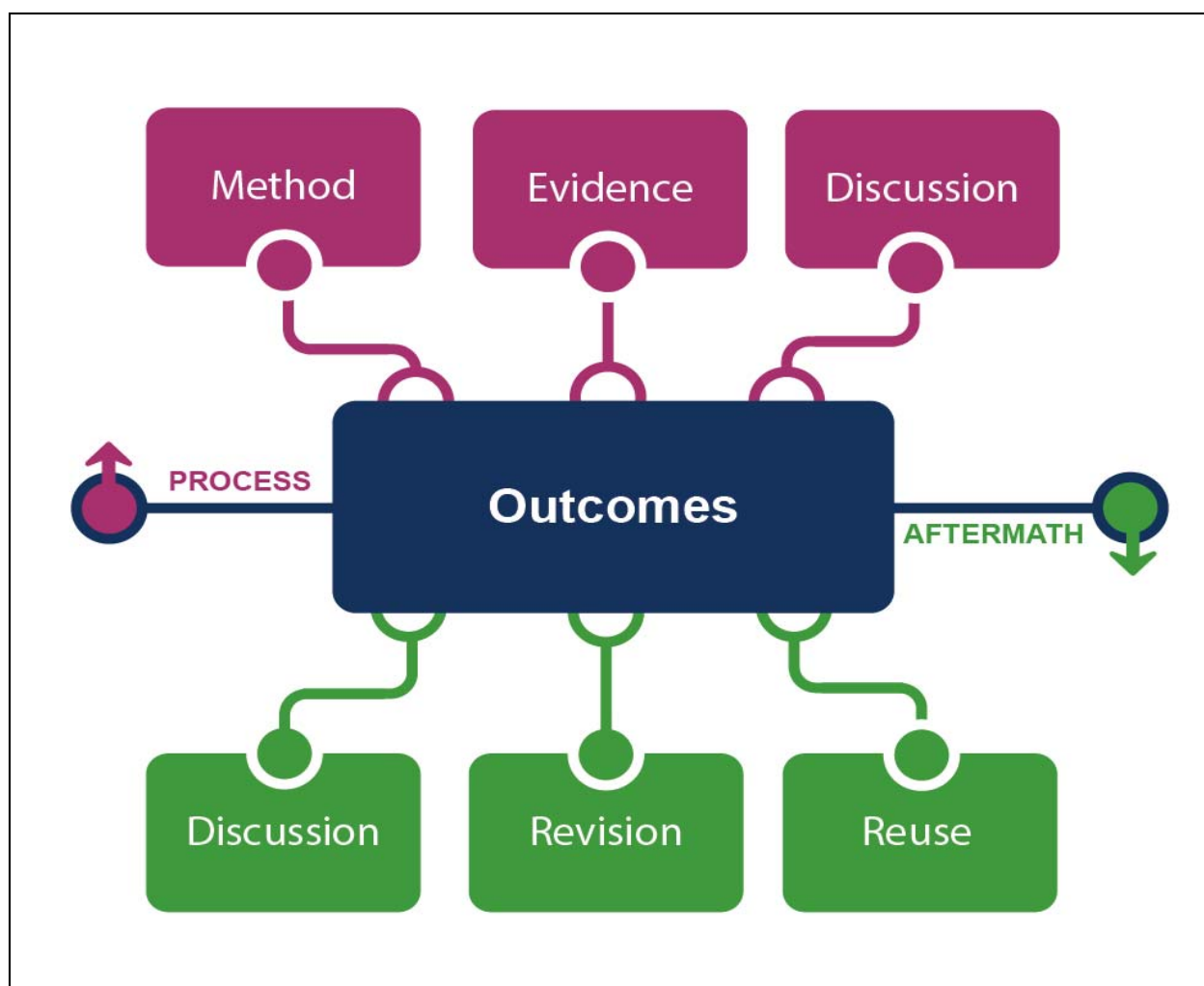
and aftermath phases of scholarly inquiry.

Figure 1: The content of the scholarly record

Process

The process phase of scholarly inquiry refers to the research activities leading up to the production of published outcomes. These activities generate materials that potentially could migrate to the permanent scholarly record. As figure 1 shows, these materials can be divided into three classes:

- *Method*: materials representing or documenting methodological techniques or innovations (e.g., software, computer models, digital lab notebooks, sampling frames, experimental protocols, instrument calibrations).

- *Evidence*: the "raw materials" of, or inputs to, the scholarly process from which outcomes are derived (e.g., data sets, survey results, new or enhanced primary source documents, links to findings in other scholarly works).

- *Discussion*: materials capturing formative discussions and other interactions with colleagues, experts and other interested parties that coalesce around a particular scholarly endeavor while it is underway and help shape or refine the research process and its outcomes (e.g., preprints, listserv/blog discussions, conference presentations, annotated commentary, grant proposals/reviews).

Anchoring outcomes directly to the methods employed, evidence used, and formative discussions conducted during the process of scholarly inquiry, helps contextualize and deepen our understanding of these outcomes, facilitates replicability, and eases the task of leveraging published outcomes into new research. As we will see in the next section, efforts have already sprung up to capture these materials, and make them part of the permanent scholarly record. As interest continues to grow in extending the scholarly record to include materials related to method, evidence and discussion, the "paper trail" of science will be captured in ways it never has been before, when published results were often the only record of research activity.

Aftermath

Once the outcomes from a research project have been formally published or otherwise made available, scholarly activities surrounding that piece of work may still continue in the aftermath phase. As figure 1 illustrates, three classes of materials can potentially be created in this phase:

- *Discussion*: materials capturing discussions and other interactions between scholars and other interested parties pertaining to the ideas or findings presented in a

published outcome (e.g., listserv/blog discussions, conference presentations, annotated commentary, post-publication formal reviews).

- *Revision*: materials representing alterations to the substance of a published outcome (e.g., the outcome may be enhanced with additional findings; errors may be corrected or clarifications made).

- *Reuse*: materials produced by editing or repackaging a published outcome for a new venue or audience (e.g., conference presentations, summaries, blog posts, versions for popular audiences).

Capturing the materials generated as a result of these aftermath activities, and connecting them back to their antecedents in the process and outcome phases of scholarly inquiry, underscores the notion that the fruits of scholarly inquiry, and indeed the scholarly record itself, are dynamic in nature: as they enter the stream of scholarly discussion and use, they are subject to critique, refinement, and repurposing.

It is important to emphasize that the conceptualization of the scholarly record depicted in figure 1 does not attempt to present a model of the research process itself. Such a model is beyond the scope of this paper. Instead, a very simple chronological context has been layered on top of the categories of material we identify as potentially constituting the permanent scholarly record. This chronological layer is anchored around the publication of some type of scholarly outcome, and divides categories of material into those generated before the outcome is made available, and those generated afterward. In this sense, the focus of figure 1 is not on the end-to-end process of scholarly inquiry, but instead on the materials created as a result of scholarly inquiry. Put another way, the framework distinguishes between materials that are byproducts of research activities (process phase), and those that are byproducts of the publication of the scholarly outcome (aftermath phase).

Additionally, we do not mean to suggest that the categories of materials surrounding published outcomes have only now sprung into existence, or have suddenly become important. In fact, these materials have always existed in one form or another, and they have always represented an integral part of the context surrounding scholarly inquiry. Our point, rather, is that traditionally these materials were not formalized as part of the permanent scholarly record through systematic collection, consistent referenceability, and persistent accessibility.

New scholarly work is usually built on the foundations of the existing scholarly record, producing new materials that in turn become part of the scholarly record and inform future work. In this sense, the entire existing scholarly record is available as input to any new

research process, a point we have not captured in figure 1, but that is at the heart of the value of collecting and preserving the scholarly record.

The specific nature of the materials that might be considered part of the scholarly record in each of the categories of material surrounding published outcomes will change from context to context, from discipline to discipline, and from stakeholder to stakeholder. The point is that whatever form these materials may take, they can be classified according to one of these categories. Moreover, we are not suggesting that everything depicted in figure 1 will necessarily end up in the permanent scholarly record. Rather, the picture represents the maximal scope and depth of the materials in which there is increasing interest in systematic collection and curation.

Finally, materials in the categories surrounding scholarly outcomes might become outcomes in their own right. Data sets are a good example: in some disciplines, the publication of an important data set is now considered a first-class scientific outcome, ranking equally beside findings derived from the data.

In summary, figure 1 envisions a scholarly record that is evolving into a greater emphasis on collecting and preserving context, by incorporating materials generated during the process and aftermath phases that proceed and follow the release of published outcomes. This in turns suggests a scholarly record that is becoming a deeper and more complete record of scholarly inquiry.

Examples

In order to make the concepts presented in figure 1 more concrete, figure 2 provides examples of specific materials for the components of the scholarly record surrounding published outcomes. In these examples, it is interesting to note not just the content itself, but also the services or publication channels through which the materials are made available to the scholarly community.
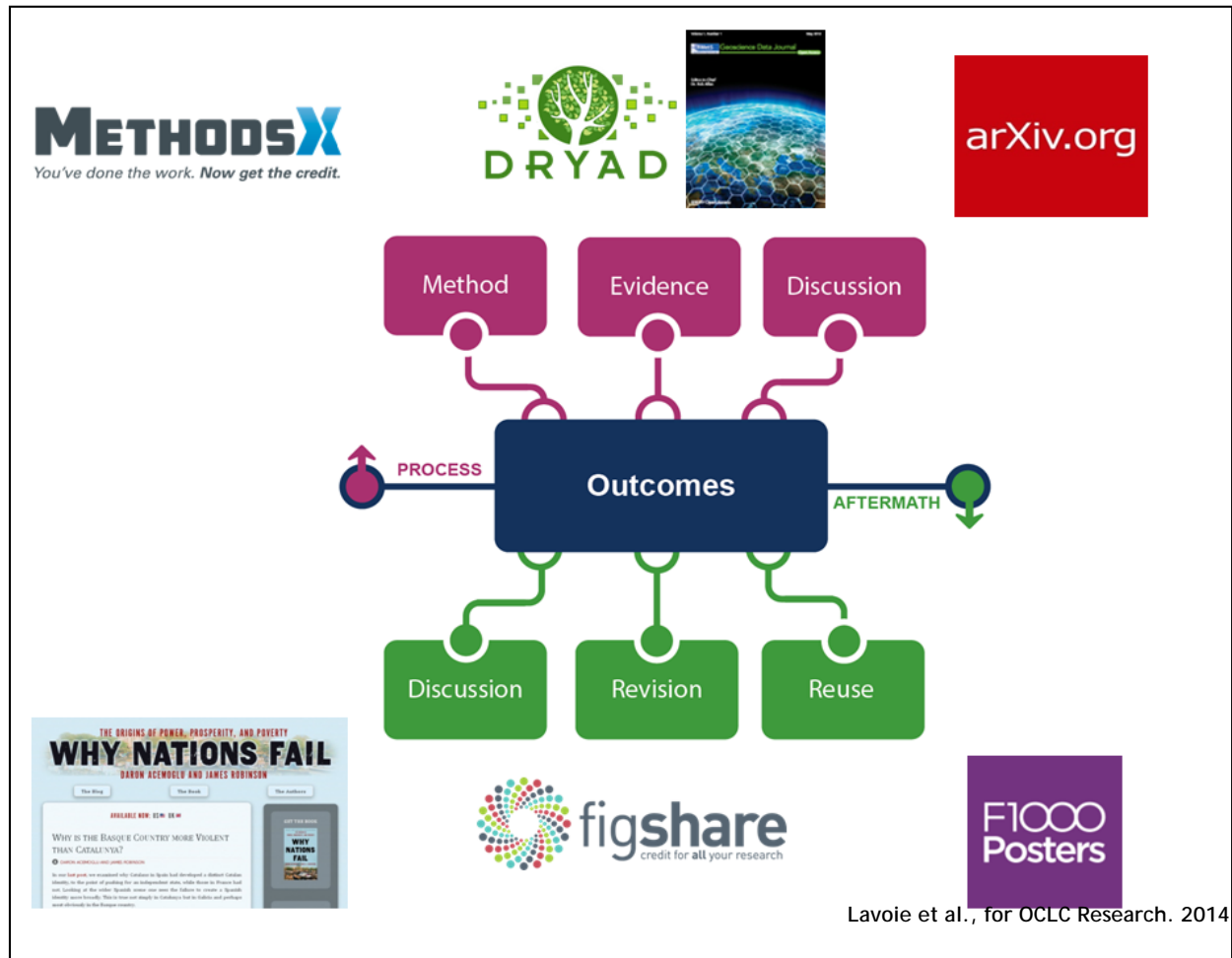
Figure 2: Examples of materials migrating to the scholarly record

Process:

- *Method*: *MethodsX*[5] is a new journal launched by Elsevier that allows researchers to publish details about methodological techniques and innovations that may be of broader interest. On the journal's website, researchers are encouraged to "releas[e] the hidden gems from your lab book."

- *Evidence*: Dryad[6] is a repository of data sets associated with published articles in the life sciences literature. *Geoscience Data Journal*[7] also makes research data available to the scholarly community, albeit through a different channel: the journal, published by Wiley, provides a platform for publishing research data through a peer-reviewed process.

- *Discussion*: ArXiv[8] is a preprint repository that provides an opportunity for scholars to expose their work to peers for discussion and commentary prior to formal publication.

Aftermath:

- *Discussion*: *Why Nations Fail*[9] is a blog launched in parallel with the publication of a book (written by two economists) of the same title. The blog provides a venue for the authors and other interested parties to discuss the ideas presented in the book, as well as relate them to current events.

- *Revision*: Figshare[10] is a service that allows researchers to enhance their published work by uploading ancillary materials and making them accessible and citable. Materials uploaded to Figshare may be revised (or deleted) as needed, with version control support for publicly available data.

- *Reuse*: F1000 Posters[11] is a permanent repository for posters and presentations. It is a channel for making materials available that repackage published outcomes into new forms for different venues—in this case, conference posters and presentations.

The examples presented in figure 2 represent an interesting split between materials that are made available through what we might term "publications"—e.g., *MethodsX, Geoscience Data Journal*, even the *Why Nations Fail* blog—and those that are made available through what would more appropriately be labeled services—e.g., Dryad, arXiv, or Figshare. While the former can be gathered and organized into collections in ways at least somewhat analogous to the traditional print-centric scholarly record, the latter may pose greater challenges for systematic collection and curation.

MethodsX and *Geoscience Data Journal* are examples of new wine in old bottles. The new wine is the material that is entering the permanent scholarly record—methodological techniques and research data sets. The old bottles are the channels through which these materials are being made available: the traditional scholarly journal, complete with formal peer-review and editorial processes.

Finally, it is interesting to note that both *MethodsX* and Figshare emphasize private credit to the researcher as a key incentive for making materials available through these venues. *MethodsX* advises "You've done the work. **Now get the credit.**" Similarly, Figshare promises "credit for **all** your work." An interesting aspect of the evolution of the scholarly record is that its extension to a much wider range of materials is fueled by the dual (and complementary) public and private incentives to deepen and expand documentation of

scholarly activities, and to credit the researcher for the full range of materials produced during the process of scholarly inquiry and its aftermath.

Stakeholder Ecosystem

The second part of the scholarly record framework addresses the stakeholder ecosystem attached to the scholarly record (figure 3). In depicting this ecosystem we focus on the key roles filled by stakeholders, rather than the specific identities of the stakeholders themselves. By moving up to this layer of abstraction, the ecosystem can be more flexibly applied across many contexts, where different entities (e.g., publishers, libraries, proprietary information services, etc.) can be configured in a variety of ways. In this sense, it is important to emphasize that in our view of the ecosystem, multiple roles could be subsumed within a single entity.
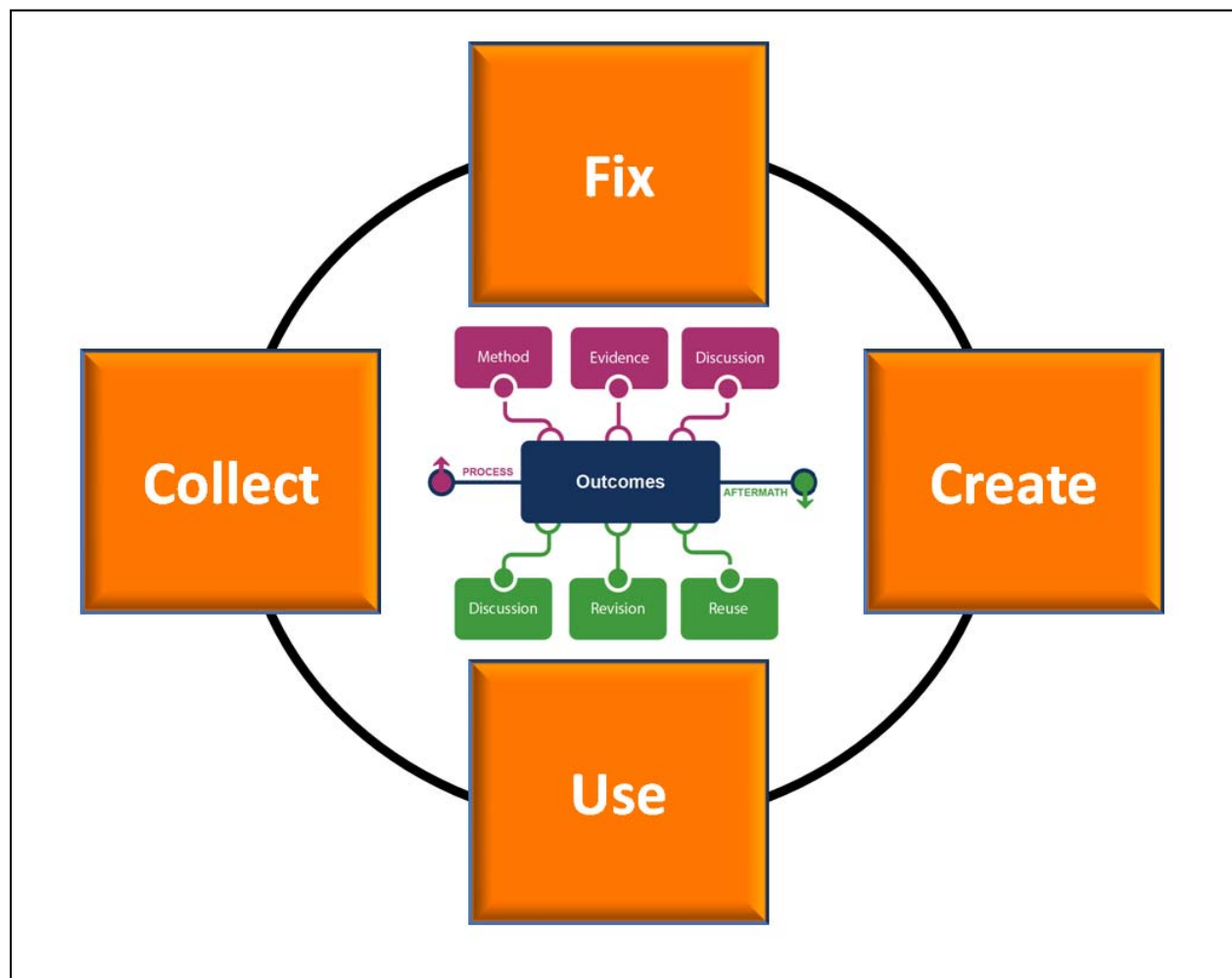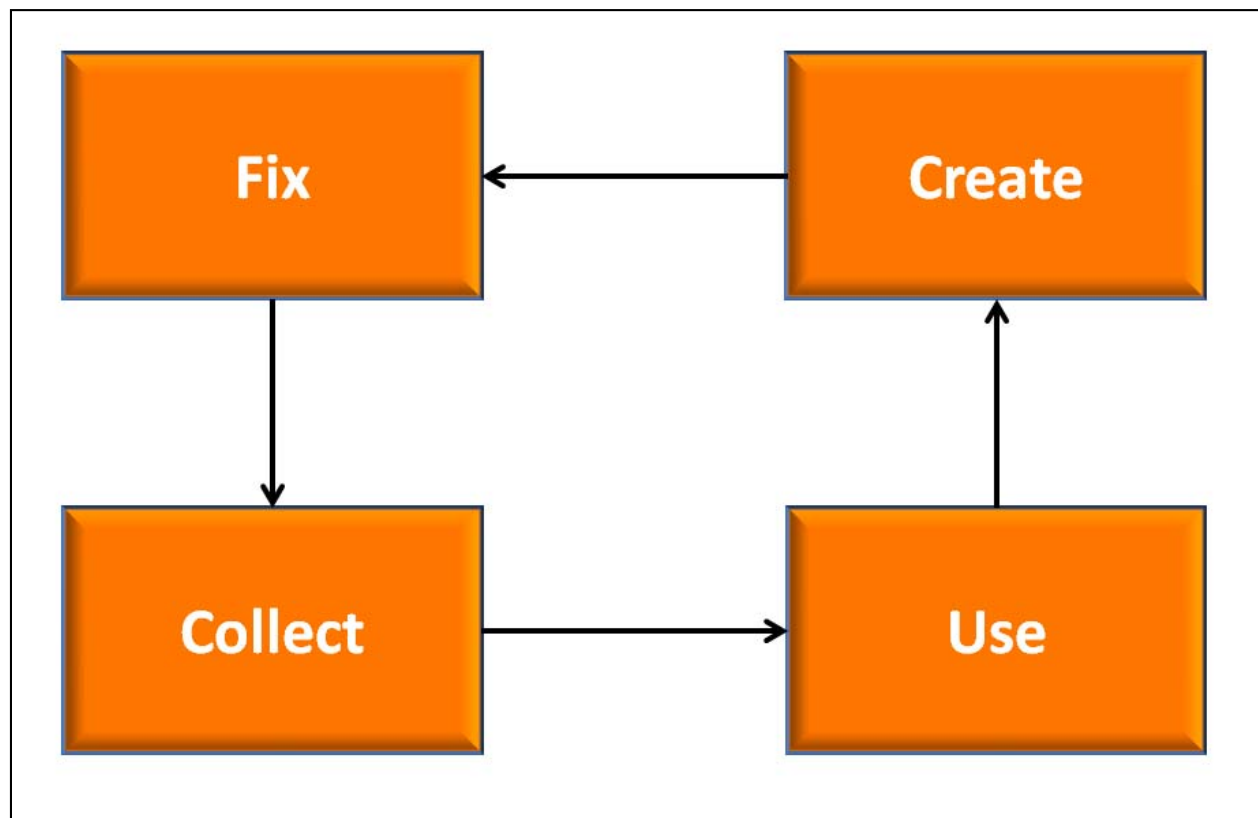
Figure 3: Roles in the scholarly record's stakeholder ecosystem[12]

In the ecosystem depicted in figure 3, four broad categories of stakeholders are distinguished by the nature of their activities. The *Create* role is filled by authors or other agents[13] who create the materials that may eventually migrate to the scholarly record; referring back to figure 1, these materials include published outcomes, as well as the categories of materials in the process and aftermath phases related to these outcomes. The *Fix* role is performed by publishers and other organizations that fix the materials created by authors in the recognized literature, by transitioning them into an authoritative (often peer-reviewed) version that is in a persistent, citable, accessible form—often through an established scholarly communications channel such as a journal. The *Collect* role is filled by libraries and other organizations that gather scholarly materials and organize them for long-term use and preservation as part of the permanent scholarly record. Finally, the *Use* role refers to the multiplicity of uses that researchers, students, and others make of the scholarly record, which in many cases will result in the creation of new materials, thus restarting the cycle.

We acknowledge that there is a lot of nuance buried in these four very broad characterizations of stakeholder roles.[14] Moreover, the boundaries or distinctions between the roles are not as sharp in practice as the picture seems to suggest. The goal here is to provide a high-level conceptualization of the stakeholder ecosystem that can be used to support thinking about how these roles are changing or being reconfigured in conjunction with the evolution of the scholarly record itself. We present a few examples to make this idea more concrete.

Lavoie et al., for OCLC Research. 2014.

Figure 4 illustrates what might be considered the "traditional" or print-centric configuration of the stakeholder roles. In this setting, materials comprising the scholarly record followed essentially a cyclical path through the various stakeholder roles: materials were created by scholars; fixed in the literature through formal publication; collected by libraries and made available for local use; and then used by faculty and students to support research and learning. Use of the scholarly record often involves the creation of new materials, which begins the cycle anew.
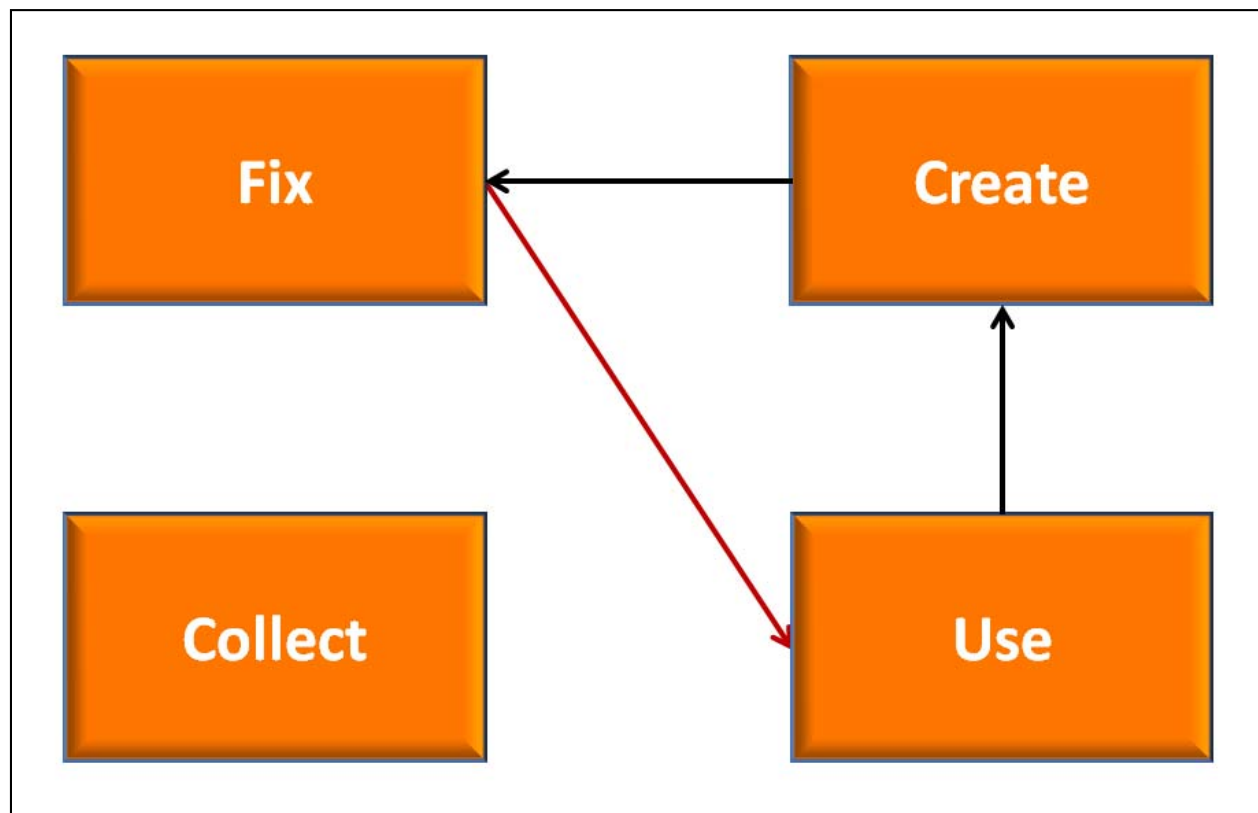
Lavoie et al., for OCLC Research. 2014.

Figure 4: Traditional or print-centric configuration of stakeholder roles

In some quarters of the evolving scholarly record, however, this traditional configuration of roles is changing. Some important contributions to the scholarly record may not be formally "fixed" in the literature through traditional agents like publishers. Instead, they may be disseminated outside the formal publication process through venues like institutional repositories or faculty web pages, thus disintermediating—or at least redefining—the *Fix* role in the configuration. Similarly, some portions of the scholarly record may be accessible directly from the scholars who created them, rather than through collections built and maintained by libraries and other institutions. This disintermediates, or at least redefines, the *Collect* role. Reconfigurations of stakeholder roles pose new challenges for securing important segments of the evolving scholarly record.

Figure 5 illustrates how we might depict the configuration of the stakeholder ecosystem in regard to the e-journal literature.
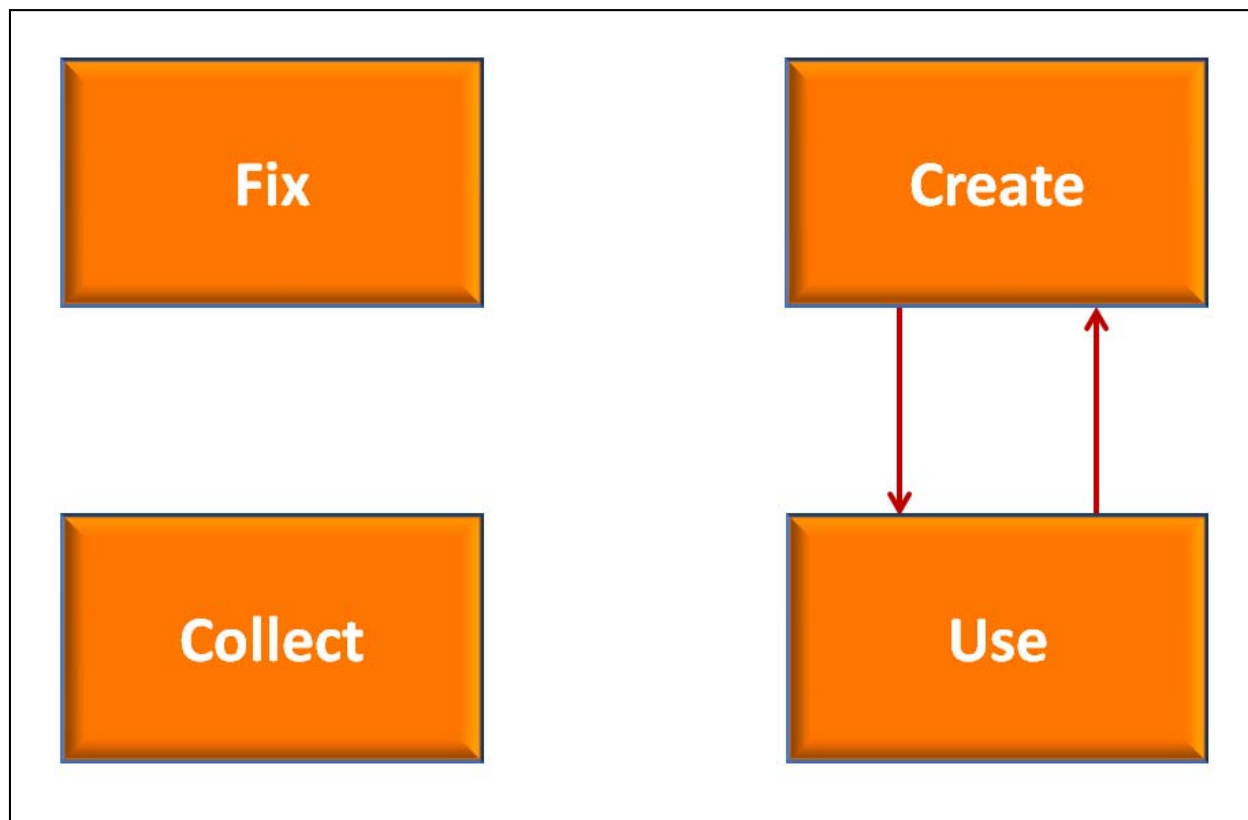
Lavoie et al., for OCLC Research. 2014.

Figure 5: Stakeholder configuration: e-literature

In this configuration, the Collect role has largely been bypassed. Materials remain in the custody of publishers, and are accessed through the publisher's proprietary platform. The academic library, which traditionally performed the functions associated with the Collect role for the print journal literature, finds itself now performing a "brokerage" role, by licensing access to the e-journal literature on behalf of its constituent faculty and students. However, as the environment continues to evolve in this segment of the scholarly record, the Collect role has worked its way back into the configuration—sometimes through relative newcomers on the scene: for example, organizations like JSTOR and Portico have obtained custody of portions of the e-journal literature and are committed to sustaining long-term access to them.

Figure 6 presents the configuration of stakeholder roles associated with the portions of the scholarly record that are manifested in social media (e.g., blogs, Twitter) and social storage (e.g., SlideShare, YouTube, Flickr).

Lavoie et al., for OCLC Research. 2014.

Figure 6: Stakeholder configuration: social media and social storage

In this setting, the traditional Fix and Collect roles are both disintermediated from the ecosystem. Materials of this kind are often made available directly from creators to users, through an intermediary (often proprietary) service that perhaps performs some of the functions associated with the Fix and Collect roles, but not all. Again, as time has passed we have seen efforts to resurrect the Fix or Collect roles in this configuration: for example, the transfer of the Twitter archive into the custody of the Library of Congress, or various initiatives aimed at collecting and preserving academic blogs.

The key point about the stakeholder ecosystem depicted in figure 3 is that the traditional roles inherited from the print-centric version of the scholarly record—Create, Fix, Collect, and Use—are still relevant in today's environment. Although one or more of these roles may initially be neglected as new pathways of scholarly communication emerge and develop (with accompanying shifts in the configuration of stakeholder roles), gaps in performance of the functions associated with these roles must eventually be filled. In some cases, this may involve new stakeholders adopting roles traditionally associated with other institutions or organizations—as we saw with the e-journal literature example.

Using the Framework: Example

The scholarly record framework is intended to, among other things, help organize and support discussions and consensus-building efforts around issues having to do with the scholarly record. To illustrate this, we suggest an example of a context in which the framework might serve as a useful resource.

In 2010, the National Science Foundation's Blue Ribbon Task Force on Sustainable Digital Preservation and Access published its final report *Sustainable Economics for a Digital Planet: Ensuring Long-term Access to Digital Preservation*.[15] The report dealt with issues and solutions for achieving economically sustainable digital preservation activities. Among the report's recommendations was the following:

> . . . Libraries, scholars, and professional societies should develop selection criteria for emerging genres in scholarly discourse, and prototype preservation and access strategies to support them. (55)

The scholarly record framework could be used to support an effort to address this recommendation. The first part of the framework, which conceptualizes the potential content of the scholarly record (figure 1), could serve as a reference point for the first part of the recommendation: in particular, one could imagine a discussion working through each of the components in the picture, in order to identify specific materials that are deemed important to secure as part of the permanent scholarly record within a particular disciplinary context. In this way, priorities can be established for collection and long-term curation.

The stakeholder ecosystem (figure 3) could help address the second part of the recommendation, by elucidating the pathways by which the selected materials are currently being created, managed, and used, and translating these pathways into specific configurations of stakeholder roles. These configurations would help identify key relationships among stakeholders that need to be forged, as well as possible gaps in fulfillment of responsibilities or functions embedded within the various stakeholder roles that might impact prospects for achieving long-term preservation and access goals.

In short, the scholarly record framework supports efforts to address this recommendation by mapping out the conceptual spaces in which the discussions must take place—in the first instance, the potential scope of the content of the scholarly record (within which certain discipline-specific priorities need to be established), and in the second instance, the stakeholder ecosystem attached to the scholarly record (whose configuration in regard to the selected materials impacts the formulation of preservation and access strategies).

Issues to Think About

Identification, collection, and long-term curation of the evolving scholarly record leads to a host of attendant issues and challenges. Some of these are quite new, rooted in the increasingly digital and networked character of the scholarly record; others are better described as new manifestations or amplifications of familiar challenges from the traditional print-centric scholarly record. As we think about the implications of an evolving scholarly record similar to that which we describe in this paper, a number of issues are readily apparent—for example:

- *Drawing a distinction between the scholarly record and the cultural record*: The boundaries of the scholarly record must be distinct enough to avoid drawing everything into it. How can we distinguish between the scholarly record on the one hand, and the broader cultural record on the other? Often, this question must be addressed on a discipline-by-discipline basis, as perceptions of what is "scholarly" can shift from context to context.

- *The dynamics of the scholarly record*: The digital age has introduced a higher degree of dynamics into the scholarly record. Digital materials are more mutable than those in fixed form, and because of this, materials in the new scholarly record may not necessarily be static. They can be altered, supplemented, or even retracted (as suggested by some of the material identified in the "aftermath" phase of figure 1). In the digital environment, versioning can be a much more complex issue than in the print world. This mutability has important implications for scholarly citation and referencing practices.[16]

- *Manifesting a "scholarly work" for discovery, access, and use*: Preserving not just the published outcome of a research project, but also the various ancillary materials associated with that outcome (e.g., data sets, computer models, lab notebooks, etc.) suggests that a set of relationships will be needed that binds together the various pieces of a "scholarly work," which may be distributed across many locations on the network. Moreover, these relationships need to be instantiated in a data layer upon which services such as discovery and fulfillment can operate.

- *"Selecting" the permanent scholarly record*: As with print resources, selection is still an important issue for a digital, networked scholarly record. Vast amounts of material are being produced that could potentially be included in the permanent scholarly record, but it is unlikely that sufficient capacity exists to gather, organize, and curate all of these materials. Choices will have to be made, and priorities established.

- *Stewardship models for the evolving scholarly record*: The increasing volume and complexity of the content potentially comprising the scholarly record, as well as a widening distribution of custodial responsibility, suggests that "local copies" of the scholarly record are becoming increasingly partial—that is, the portion of the scholarly record that a single institution can hope to collect, store, and offer locally is getting smaller and smaller. This has important implications for the type of stewardship model best suited for securing the long-term persistence of the scholarly record.

Conclusion

A framework that conceptualizes the scope of the evolving scholarly record, as well as its key stakeholder roles, is a useful first step toward working through a host of issues related to gathering and perpetuating a deeper, more comprehensive record of the process, outcomes, and aftermath of scholarly activity. Such a framework helps cultivate the shared understanding, and ultimately, the collaborative relationships needed to effectively identify, collect, and make accessible the wide range of materials the boundaries of the scholarly record are stretching to encompass. Of course, the concepts, generalizations, and abstractions of the framework must eventually give way to the practical details of the specific materials and stakeholder entities around which strategies for curation of, and access to, the permanent scholarly record must revolve. But it is helpful to build toward these practical solutions from a shared understanding of the basic contours of the evolving scholarly record and its stakeholders.

Notes

1. The OAIS reference model is a good example of a framework serving to organize cross-domain discussions within a particular problem space (digital preservation). (See CCSDS 2012)
2. A number of interesting discussions of recent trends in the scholarly record and scholarly communication, and their implications, are available. A survey of these contributions is beyond the scope of this brief paper, but two good examples are: "Full-spectrum Stewardship of the Scholarly Record" (Schottlaender 2010) and "On-line Scholarly Communications: vd Sompel and Treloar Sketch the Future Playing Field of Digital Archives" (Angevaare 2014).
3. To corroborate this point, consider Dewald, Thursby and Anderson's (1986) "Replication in Empirical Economics: The Journal of Money, Credit, and Banking Project" published nearly 30 years ago. The authors collected two years' worth of articles published in the *Journal of Money, Credit, and Banking*, and attempted to collect the data and computer models used to generate the findings reported in the articles. With these in hand, the authors then proposed to try to replicate the findings reported in the articles. Predictably, the authors had a great deal of difficulty collecting the data and computer models, and even in instances where they did, they often had trouble replicating the reported findings. As a result the study concluded that journals needed to make an effort to secure data and computer code prior to publication of the article itself. These are issues that are still debated today.
4. By published, we mean either formally published or otherwise made available.
5. See http://www.journals.elsevier.com/methodsx/.
6. See http://datadryad.org/.
7. See http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%292049-6060.
8. See http://arxiv.org/.
9. See http://whynationsfail.com/.
10. See http://figshare.com/.
11. See http://f1000.com/posters.
12. This framework is based on earlier unpublished material developed by Lorcan Dempsey and Brian Lavoie.
13. Although for brevity's sake our discussion focuses on human authors, we acknowledge that other agents, such as computers, measuring instruments, sensors, etc., can also create materials that could become part of the scholarly record (such as data sets) without explicit human intervention.
14. We acknowledge that stakeholders can influence the scholarly record in ways beyond the roles depicted in figure 3. For example, technology providers facilitate the Create role, and in doing so, shape the materials we might include in the Method category from figure 1, and more broadly, the characteristics of the materials that might be included as part of the permanent scholarly record (e.g., by defining format).
15. One of the authors of this paper served as co-chair of the task force.
16. This issue has been addressed in the broader context of Web archiving by the Memento framework, which introduces protocols for gathering and integrating different versions of Web-based resources that have appeared over time. (See http://mementoweb.org/)

References

Angevaare, Inge. 2014. "On-line Scholarly Communications: vd Sompel and Treloar Sketch the Future Playing Field of Digital Archives" *Research in KB* (blog). National Library of the Netherlands, 22 January. http://researchkb.wordpress.com/2014/01/22/on-line-scholarly-communications-and-the-role-of-digital-archives/.

Atkinson, Ross. (1990). "Text Mutability and Collection Administration" *Library Acquisitions: Practice & Theory.* 14(4):356. doi:10.1016/0364-6408(90)90006-G.

CCSDS (Consultative Committee for Space Data Systems). 2012. Reference Model for an Open Archival Information System (Oais): Recommended Practice CCSDS 650.0-M-2. Washington, DC: CCSDS Secretariat. http://public.ccsds.org/publications/archive/650x0m2.pdf.

Blue Ribbon Task Force on Sustainable Digital Preservation and Access. 2010. *Sustainable Economics for a Digital Planet: Ensuring Long-term Access to Digital Preservation*. La Jolla, California: Blue Ribbon Task Force on Sustainable Digital Preservation and Access. http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf.

Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. 1986. "Replication in Empirical Economics: The Journal of Money, Credit, and Banking Project" *The American Economic Review.* 76 (September): 587-603.

Schottlaender, Brian. 2010. "Full-spectrum Stewardship of the Scholarly Record." Presentation at the 30th Annual Charleston Conference, Charleston, South Carolina, 5 November 2010. http://www.slideshare.net/CharlestonConference/full-spectrum-stewardship-of-the-scholarly-record-by-brian-e-c-schottlaender-university-of-california-san-diego.

Straumsheim, Carl. 2014. "Is Blogging Unscholarly?" *Inside Higher Ed*. 29 January. http://www.insidehighered.com/news/2014/01/29/international-studies-association-proposes-bar-editors-blogging#sthash.RkX99zTI.dpbs.

Teitelman, Robert. 1994. *Profits of Science: The American Marriage of Business and Technology.* New York, NY: BasicBooks.