

Toward Improved Discoverability of Scholarly Content: Cross-Sector Collaboration Essentials

Mary M. Somerville (Mary.Somerville@ucdenver.edu)
(ORCID: <http://orcid.org/0000-0002-4201-8335>)
University of Colorado-Denver

Lettie Y. Conrad (Lettie.Conrad@Sagepub.com)
(ORCID: <http://orcid.org/0000-0003-1580-7037>)
SAGE

Abstract

By way of follow-up to earlier work in understanding and improving discoverability of scholarly content, this article reports on recent data and reflections that led to clearer definitions of discovery and discoverability, as well as deeper cross-sector collaborations on standards, transparency, metadata, and new forms of partnerships. Recent advances in discoverability are also described - from enhanced library-based web-scale searching to serving researcher needs through the Open Researcher and Contributor ID (ORCID) registry. The article points to a 2014 SAGE white paper that presents in greater detail opportunities for wider collaboration among libraries, publishers, service providers, and researchers in the interest of furthering discovery, access, and usage of scholarly writings and creative work.

Keywords: Open Researcher and Contributor ID; ORCID; Scholarly collaboration; Discovery; SAGE White Paper; Scholarly content

Introduction

In January 2014, SAGE released a white paper at the American Library Association Midwinter Conference in Philadelphia examining the current discoverability landscape. It updates the initial SAGE white paper¹ on the topic released two years earlier. The 2012 research report was summarized and extended in a January 2013 article in *Collaborative Librarianship* titled "Discoverability Challenges and Collaboration Opportunities within the Scholarly Communications Ecosystem: A SAGE White Paper Update."²

The 2014 SAGE white paper, *Collaborative Improvements in the Discoverability of Scholarly Content: Accomplishments, Aspirations, and Opportunities*,³ continues a cross-sector research collaboration begun in 2011 when expert interviews were conducted to gather baseline data on the state of discoverability. During the course of analyzing research data, Conrad and Somerville discovered a shared passion for discoverability and continue - to this day - to exchange information and insights on developments in this rapidly evolving

area of scholarly communications. The author byline of the second SAGE white paper formally recognizes this three-year 'thought partner' collaboration.

Given their respective roles in different sectors, co-authors Conrad - in the publishing world - and Somerville - in the academic world - receive dissimilar but complementary information throughout the course of their day-to-day work lives. In addition, Conrad serves as a publisher representative on the National Information Standards Organization (NISO) Open Discovery Initiative (ODI), and is therefore well-informed about the respective viewpoints on web-scale discovery services among primary and secondary (abstracting and indexing/ A&I) publishers, academic libraries, and service providers (vendors). All of this is by way of saying that information resource sharing between publishers and libraries is valuable when developing 'big picture' understanding of discovery and discoverability. Furthermore, as this "From the Field" contribution proposes, full realization of the white paper aspirations to improve discoverability of the scholar-

ly corpus requires vigorous information sharing, action-oriented analysis, and cross-sector collaborations.

Shared Goals

SAGE's sponsorship of two surveys of cross-sector discoverability experts aims to further innovative collaborations that significantly advance researcher experience and, ultimately, academic progress and results. This aspiration assumes that, "Despite increasingly challenging scholarly ecosystem circumstances exacerbated by economic uncertainty and disruptive technologies, the driving missions of academic publishing and librarianship have not changed."⁴ In other words, the SAGE white papers assume that both sectors – and their technology providers – share a continuing commitment to advancing discovery and usage of the scholarly corpus for knowledge creation and, ultimately, societal good.

Toward this end, data collection methods in both studies recognize the exchange of viewpoints and perspectives necessary to further discussion of discoverability accomplishments, aspirations, and opportunities – the particular focus of the 2014 white paper. Therefore, the substance of *Collaborative Improvements in the Discoverability of Scholarly Content* reflects both information exchanged and harvested since 2012 and also insights gleaned from the survey results of 16 experts. Survey contributors shared with the co-authors a commitment to furthering the advancement of scholarly knowledge creation and dissemination, which necessarily requires renegotiated cross-sector relationships within and across publisher, library, and service provider/vendor communities.

Expert opinions corroborated the observations of the white paper's authors that, in the past two years, considerable progress has been made in raising awareness of the importance of scholarly discovery enhancements and establishing baseline codes of practice and rules for engagement. For instance, the National Federation of Abstracting and Information Services (NFAIS) has proposed codes of practices for web scale discovery services and primary and secondary academic content publishers. And the NISO Open Discovery Initiative will soon be publishing recom-

mended guidelines for discovery services on fair linking, metadata workflows, and usage statistics for shared assessment routines. Within this context of considerable refinement in recommended standards from various international bodies, with the aim to advance cross-section cooperation and collaboration, the white paper proposes actionable recommendations for discoverability improvement. Additionally, the paper urges collaborative action in its three-fold purpose:

- Propose actionable recommendations for discoverability improvement,
- Advance heightened cooperation among librarians, publishers, and service providers, and
- Encourage 'new ways of thinking and doing' that improve researcher experience.

Study Findings

In establishing context and generating recommendations, the white paper draws upon peer reviewed published articles, as well as professional blog posts, conference presentations, research projects, and commissioned studies. In addition, the report is enriched with insights from ecosystem experts with national and international reputations in the publishing, consulting, and vendor/service communities, as well as in academic librarianship and user experience. In considering published thought and survey contributions together, new definitions and recommendations emerged.

In recent years, definitions have evolved that distinguish between discovery and discoverability, and recognize the importance of optimizing researcher user experience. They now enable greater precision in cross-sector conversations that intend to further collaboration by differentiating:

- *Discovery* - the process and infrastructure required for a user to find an appropriate item and
- *Discoverability* - the description or measure of an item's level of successful integration into appropriate infrastructure maximizing its likelihood of being found by appropriate users.

Within the context of improved understanding of researcher discovery aided by discoverability infrastructure, the white paper offers four actionable collaborative discoverability recommendations.

- Standards - All sectors should observe ratified 'content optimization' and interoperability standards, as part of business-as-usual routines.
- Transparency - Ratified standards and best practices compliance are critical for successful discovery, and the development, implementation, and enforcement of these standards require open relationships across the ecosystem to achieve common goals.
- Metadata - Quality metadata, observing ratified standards, enable successful discovery of scholarly content, products, and services.
- Partnerships - Opportunities exist for new discovery innovations across the industry that support enriched discoverability to produce integrated pathways for achieving researchers' 'jobs-to-be-done.'

Library Implications

These recommendations have significant implications for libraries and their users. When common standards are not observed or enforced, libraries incur tremendous costs in terms of re-work, including manual correction of 'bad' (non-standards compliant) metadata, and researchers experience annoying and time-wasting inconveniences, such as partial search results or broken links, both evidence of disconnects between discovery and delivery. Libraries can influence publisher adoption of standards for systems interoperability and 'good' metadata through purchasing decisions that reward standards compliance.

Collaborative partnerships can also influence adoption of standards that advance heightened cooperation among librarians, publishers, and service providers (vendors). Such partnerships require inventing new ways of thinking and doing that improve researcher experiences. Systems co-design is among the most fruitful approaches to cross-sector collaboration. As the following example illustrates, participatory design partnerships can produce enhanced functionalities and expanded content within an immersion environ-

ment. In addition, when discovery and delivery are seamlessly integrated, researchers can more swiftly move to finding resources when searching curated library content through a customizable 'single search box' environment that addresses the vexing problem of knowing where, as well as how, to search in order to discover.

In the spirit of systems co-design, since 2010, the Auraria Library in Denver, Colorado has invested significant time and attention across the organization to the selection, implementation, and refinement of a web-scale discovery service. Although there was inconsistent evidence available to confidentially judge adherence to standards, selection criteria privileged web-scale discovery services that ensured vendor neutrality through non-biased indices and algorithms, provided usage reports (both primary content and A&I publishers), enabled user testing for publishers to make sure their content is visible, and clarified what is and what is not indexed. Following extensive dialogue and debate, ProQuest Summon was selected. Then agreement was reached to position the Summon discovery tool as "the place to go for full text, images, and media."⁵

Achieving the discovery tool brand promise required partnerships with both ProQuest and publishers. For instance, ProQuest added a video facet as a search refinement option that now permits efficient search for content in *NBC Learn* and *Films on Demand*. Working with ARTstor, ProQuest also added the 'image spotlighting' feature that highlights the 1.4M art, art history, and architecture images collection. To provide integrated researcher experience through a single search box, Summon next indexed *CREDO Reference* that now provides contextualizing content from multiple publishers for 'getting started' research advice. Refinement and enhancement of Summon continues during preparation across the library for release of Summon 2.0. This next step in the library discovery service 'lifecycle' will continue to require cross-sector collaborations.

These examples illustrate the impact of cross-sector cooperation and collaboration aimed at optimizing the user experience and ensuring library integration in researcher workflows. Another cross-sector example involves a researcher identification system that further illustrates the

potential of the four white paper recommendations of standards, transparency, metadata, and partnerships.

The Open Researcher and Contributor ID (ORCID), which celebrated its first-year launch anniversary in November 2013, assigns unique identifiers to researchers in order to associate scholars or entities with research outputs, whether technical papers, grant awards, patent applications, journal articles, or conference papers, across multiple hosts or databases.⁶ To address the researcher name ambiguity problem, ORCID provides a registry of persistent unique identifiers for published researchers and scholars. Widespread adoption and usage by the research community at key workflow and dissemination points – manuscript submissions, datasets deposit, grant applications, patent applications, and faculty records – support cross-disciplinary, cross-sector linkages across multiple grant awards, clinical trials, scholarly publications, patents, and datasets. Toward that end, 350,000 researchers registered and ninety organizations joined in the first year, and more than 50 organizations have already integrated ORCID identifiers into workflows, including CrossRef, Researcher ID, PubMed Central, and Scopus.

In reflecting on the potential of this cross-sector initiative, SAGE white paper contributor and expert consultant Judy Luther stated,

The potential for ORCID to add a very significant contribution to discovery is huge. Imagine finding all the work of one author readily linked through a CrossRef DOI to the item. From a researcher perspective, their ORCID page can serve as their public vita for publications. A page of ORCID snapshots would provide a quick overview of an academic department. One group in the social sciences wanted to see a map of Ph.D. students linked to the schools where they studied so it would be easier to understand the spread of ideas. That could then serve as a valuable reference point in the greater disciplinary network – a mapping of academic disciplines. ... The use of ORCID would simplify this process and enable other applications.⁷

In considering the conditions needed to realize this ‘ideal’ discoverability, Luther continues, “Full text publishers, A&I publishers and scholarly societies need to employ the ORCID ID and encourage its adoption. Subsequently, librarians could use it as part of their authority records for creating MARC records which would need to be included in processing search results by search engines and ILS systems.”⁸ Widespread adoption of this cross-sector initiative holds great promise for reducing the fragmentation of the current discovery environment that requires navigation of multiple pathways and associated choices, and that unduly complicates user experience and potentially compromises search results.

Ideal Futures

The SAGE white paper details the considerable progress made in recent years to establish standards, practices, and codes with the aim of enhancing discovery and usage of the scholarly corpus. In addition, the white paper encourages ‘new ways of thinking and doing’ that improve researcher experience. A continuum of strategies can advance working better together across sectors. Individuals can productively engage in cooperative information gathering and exchange with thinking partners in their own field as well as in other sectors of the scholarly ecosystem. Collective pursuit of such cross-sector ‘sense making’ opportunities with thoughtful collaborators can occur at conferences, such as the Charleston Conference, as well as through social media or conventional email.

Ultimately, however, heightened collaboration among librarians, publishers, and service providers is essential if we are to successfully create significant advancements that improve researcher experience, researcher workflow integration, and the user experience. Thus, the ultimate aim of this white paper depends on additional cross-sector discussion and, ultimately, action-oriented planning to create nimble, responsive initiatives for co-designing library and publisher websites and product customizations.

Joint ventures would necessarily be supported by endorsed cross-sector standards and practices and fostered by the considerable goodwill among supply chain contributors.⁹ The benefits are

many, including for libraries a shifting of attention from production to analysis.¹⁰ For researchers, this holds promise for moving from searching to finding and using. For the scholarly community, it offers the opportunity to build upon the significant achievements realized in recent years. As the report concludes, “the shared goal remains furthering discovery, access, and usage of scholarly publications and creative work. As library, publisher, service, and researcher roles evolve and as more sophisticated discovery and discoverability strategies emerge, robust and sustainable progress very much depends on continued, heightened cross-sector collaborations.”¹¹

Endnotes

Websites cited:

Auraria Library <http://library.auraria.edu/>
KBART <http://www.uksg.org/KBART>
NFAIS <http://nfais.org/>
NISO <http://www.niso.org/home/>
ORCID <http://orcid.org/>
SAGE <http://sagepublications.com/>

¹ M. M. Somerville, B. J. Schader, and J. R. Sack, *Improving the Discoverability of Scholarly Content into the Twenty-First Century: Collaboration Opportunities for Librarians, Publishers, and Vendors*. A SAGE White Paper. (Los Angeles, CA: SAGE Publications Ltd., 2012) Available: <http://www.sagepub.com/repository/binaries/librarian/DiscoverabilityWhitePaper/>

² M. M. Somerville and L. Y. Conrad, “Discoverability Challenges and Collaboration Opportunities within the Scholarly Communications Ecosystem: A SAGE White Paper Update, *Collaborative Librarianship* 5, no. 1 (2013): 29-41. Available: <http://collaborativelibrarianship.org/index.php/jocl/article/viewFile/240/181>

³ M. M. Somerville and L. Y. Conrad, *Collaborative Improvements in the Discoverability of Scholarly Content: Accomplishments, Aspirations, and Opportunities*. A SAGE White Paper. (Los Angeles, CA: SAGE Publications Ltd, 2014). DOI: 10.4135/wp140116. Available:

<http://www.sagepub.com/repository/binaries/pdf/improvementsindiscoverability.pdf>

⁴ L. Y. Conrad, “Discovering Authoritative Reference Material: It’s all about ‘Location, Location, Location,’” in *E-reference Context and Discoverability in Libraries: Issues and Concepts*, ed. Sue Polanka (Hershey, PA: IGI Global, 2011): 137-147.

⁵ M. M. Somerville, “Digital Age Discoverability: A Collaborative Organizational Approach,” *Serials Review* 39, no.4 (2013): 234-239.

⁶ L. L. Haak, M. Fenner, L. Paglione, E. Pentz, and H. Ratner, “ORCID: A System to Uniquely Identify Researchers,” *Learned Publishing* 25, no. 4 (2012): 259-264.

⁷ J. Luther, SAGE white paper expert survey response (November 2, 2013).

⁸ J. Luther, SAGE white paper expert survey response (November 2, 2013).

⁹ J. Bunnelle, J. Emery, M. Levine-Clark, E. McElroy, A. McKee, and M. Page, “The Charleston Conference Continues – Getting to No: Calling for an End to Contention,” *Against the Grain* (December 2013-January 2014): 77-78.

¹⁰ M. Collins, M. Somerville, N Pelsinsky, and A. Wood, “Working Better Together: Library, Publisher, and Vendor Perspectives,” in *Proceedings of the 32nd Annual Charleston Conference* (Charleston, South Carolina, 2014), in press.

¹¹ M. M. Somerville and L. Y. Conrad, *Collaborative Improvements in the Discoverability of Scholarly Content: Accomplishments, Aspirations, and Opportunities*. A SAGE White Paper (Los Angeles, CA: SAGE Publications Ltd., 2013). DOI: 10.4135/wp140116. Available: <http://www.sagepub.com/repository/binaries/pdf/improvementsindiscoverability.pdf>