

A Brief History of the Open Discovery Initiative

Kenneth J. Varnum

Senior Program Manager at the University of Michigan Library
varnum@umich.edu

Key Points

- The Open Discovery Initiative was created in 2011 to improve participation among content providers in the then-novel web-scale discovery services
- The issue addressed by the ODI included metadata standards, content indexing, and availability of licensed content in discovery services
- After adoption of the recommended practice, June 2014, a new Standing Committee has been formed to address unfinished business and implementation.
- A challenge for the ODI is finding the right balance between the needs of commercial partners and library customers.

Comment [PS1]: Author – please check you are OK with edited key points

Introduction

The NISO Open Discovery Initiative (ODI) was born at the American Library Association Annual Conference in 2011 with a straightforward aspiration: to foster transparency across the participants in the then-novel arena of web-scale discovery tools. While these discovery services -- Primo Central, Summon, EBSCO Discovery Service, Google Scholar, and their like -- are now commonplace, a half-decade ago they were new to the marketplace. The emergence of these one-stop discovery layers promised to integrate a library's more traditional metadata-driven online catalogs of physical items with full-text search of a library's entire licensed content portfolio, which raised many questions. Especially in the early days of these systems, there was little common ground for all participants in these discovery products to clearly understand the inputs and the outputs. The ODI has been working to level the playing field and shed light on the sometimes opaque world of library discovery services.

In September 2016, the author conducted written interviews with the co-chairs of the Open Discovery Initiative's original Working Group, Marshall Breeding and Jenny Walker, and the co-chairs of the current Standing Committee, Laura Morse and Rachel Kessler. The following article draws from these conversations, looking back on the 5-plus years of progress and looking ahead to ongoing challenges and opportunities, as well as the author's own experiences with the ODI.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi:

[10.1002/leap.1078](https://doi.org/10.1002/leap.1078)

Formation of the Working Group

As one of the founding co-chairs of the Open Discovery Initiative (ODI), Marshall Breeding, now an independent library consultant and then a librarian at Vanderbilt University, recalls two main areas of discomfort that inspired the formation of ODI. First, there was a growing concern among discovery service providers that there was an “unevenness of participation of publishers in regard to contributing their metadata or full text to populate discovery indexes.” Likewise, many content providers were concerned that their contributions to discovery -- especially abstracting and indexing work -- would not be represented or credited in discovery services, potentially resulting in a loss of use of their standalone products. A related issue, according to Laura Morse, a participant in the ODI process since its inception, was that “many libraries felt (and perhaps continue to feel) that content coverage and bias concerns across the various options limits choice.”

To begin addressing these concerns, Breeding and Jenny Walker, then a Vice President for Product Management at Ex Libris (maker of Primo Central), convened a conversation at the American Library Association (ALA) Annual Conference in 2011, with invited guests representing discovery services, content providers, and libraries along with NISO, the National Information Standards Organization, which has facilitated and supported ODI since its inception (see box: Open Discovery Initiative Kick-Off Meeting). Participants in the meeting talked about their interest in a framework for better understanding the inputs, mechanisms, and outputs of web-scale discovery tools. As Breeding notes, we “confirmed considerable interest in pursuing an official organization to pursue the issues related to open discovery.” Todd Carpenter of NISO invited the attendees to formalize their interest through a work item proposal to NISO (Walker, 2011). That proposal was, in turn, accepted by the Discovery to Delivery Topic Committee and a new NISO workgroup, the Open Discovery Initiative, was officially formed in October 2011, with the committee created and beginning work in January 2012. The group was co-chaired by Marshall Breeding and Jenny Walker, with members drawn from libraries, discovery providers, and content providers in roughly equal numbers (Open Discovery Initiative Working Group Roster, n.d.).

The new ODI Working Group started by conducting a survey of discovery participants (libraries, content providers, and discovery providers) to help understand the scope and scale of issues the working group would need to address and to prioritize the group’s work (NISO ODI Working Group, 2013). There was clearly a great deal of work to be done, according to the survey. In response, the workgroup determined that recommended practices were needed in five areas, according to Walker:

1. Technical recommendations for data format and data transfer, including methods of delivery and ongoing updates.
2. Recommendations for the communication (automated or through reporting) of libraries’ rights for their users to access specific content (e.g., restricted to users from subscribing libraries versus open to all users).
3. Clear descriptors regarding the level of indexing performed for each item or collection of content and the availability of the content.

4. Definition of fair linking from the discovery service to the published content.
5. Determination of what usage statistics should be collected and for whom, and how these data should be disseminated.

These five broad work areas led to the formation of five subgroups (the subgroups responsible for areas 2 and 3 were subsequently merged, resulting in four subgroups for the bulk of the time). Each subgroup, like the overall ODI, was made up of members from each constituency and some members served on multiple groups. Over the course of the following year, each subgroup worked on its assigned task, with monthly meetings of the whole Initiative to report out on progress and discuss issues that crossed over multiple subgroups' responsibilities. While the range of issues and concerns was quite broad, they can be distilled into a few core areas, as described by Breeding:

Discovery services providers desire content from the broadest range of publishers to increase the effectiveness, and therefore value of their products. Publishers need to know how this content is consumed by library patrons so that they can adjudicate the value of participation with discovery services. Libraries have an interest in discovery services which approach comprehensive representation of their collections, and must have a thorough accounting of the content provided to each of the discovery services.

One of the benefits of serving on this NISO working group was exposure and deeper understanding of the needs and concerns of the other two constituencies. As might be obvious, the internal business needs of one participant group do not necessarily align perfectly with those of another. Finding a balance, having an open dialog about the need for understanding what is included in discovery systems and what is excluded, has mutual benefits that were only discovered through this process.

In fact, as Breeding says, a main challenge of the entire process has been "to foster greater openness and participation in ways that go beyond what the commercial interests of publishers and discovery services providers would have done otherwise." The work group focused its effort on fostering transparency in three broad areas: 1) libraries' desire to understand what specific abstracting, indexing, and full text information from a content provider was available in each of the discovery services; 2) content providers' desire to understand how their metadata was processed in a discovery service to enable library users' to find their materials; and 3) discovery services' desire to understand how library users used their services and what was contributing value to their enterprise.

The ODI Recommended Practice

The ODI Working Group completed a draft recommended practice (Open Discovery Initiative Working Group, 2013) in October 2013 and released it for a public comment period. Comments received were discussed and incorporated into the final recommended practice, NISO RP-19-2014, released in June 2014 (Open Discovery Initiative Working Group, 2014). The recommended practices focused on sharing metadata between content and discovery services and statements of conformance, in which participants would declare, through a published

statement based on a template provided by ODI, the degree to which they followed the recommended practice. A few examples will suffice; the full recommended practice outlines the elements in detail.

- Content providers should provide specific metadata elements to discovery providers when they contribute their content to the index. These elements are broken down into “core” -- a minimum metadata set for each contributed item -- and “enriched” -- additional metadata that, if contributed, should be provided in a standardized way. This reduces complexity and cost in contributing materials to multiple discovery services and, for discovery services, minimizes ingest processes.
- Content providers should provide to libraries a description of the scope and depth of contributed content in a standardized way so that libraries can understand the degree to which the content they license is equally available through the discovery service.
- Discovery service providers should share coverage lists, with specified title-level metadata, with libraries so that libraries could accurately understand what materials from their licensed content providers are included in the discovery service.
- To foster transparency, discovery service providers were asked to declare whether there are any non-disclosure agreements between themselves and individual content providers that affect the indexing, relevance ranking, or linking to a particular vendor’s content.
- Content and discovery providers should follow specific file formats and data exchange protocols to minimize complexity and lower barriers to access for future contributors to this process.
- Content usage reports provided by discovery services to content providers and libraries should contain minimum features.

Current Work

With the release of the Recommended Practice in June 2014, the original Working Group was replaced with a new Standing Committee to address ongoing issues, encourage publication of conformance statements, and recommend any future work for an updated version of the recommended practice. The Standing Committee was originally co-chaired by Laura Morse of Harvard University and Lettie Conrad of Sage; Rachel Kessler of ProQuest replaced Conrad in 2016.

According to participants, the current phase of the ODI process has several challenging areas where more work is needed: First, in terms of growing the number of organizations that have issued conformance statements. Kessler says, “I think there’s a misconception that you need to be perfect in order to declare conformance. While I understand the apprehension behind declaring imperfections in writing, the goal of ODI, at least as I see it, is transparency, to show that your organization is honest and forthcoming. Organizations should, therefore, publish their conformance checklists and make plans to improve upon the areas where they are not yet perfect.” Additionally, Kessler adds, “once organizations do submit their conformance statements, I don’t think there’s enough of an incentive to keep improving.... [I]t’s hard for an

organization to publicize each small improvement. Therefore, once the statement has been made, ODI often no longer figures into the organization's priorities."

Furthermore, there has been a low rate of publication of conformance statements, particularly among content providers. As of September 2016, only five content providers (Credo, Gale, EBSCO, IEEE, and Sage) have published conformance statements. (All of the "big three" discovery providers -- EBSCO, Ex Libris, and ProQuest have done so.) ODI has embarked on a series of presentations at conferences to help encourage vendors to declare their ODI conformance and to encourage librarians to ask vendors to do so.

Future Challenges

While the Recommended Practice formed a baseline for the Standing Committee's work, it left unresolved a number of issues.

One such issue is striking the appropriate balance between the needs of commercial partners and library customers. This has been a continuing challenge. Laura Morse notes, "The heart of ODI lives at the intersection of commercial concerns for DSP and CPs and library efforts to provide valuable, visible services to patrons. Ensuring that the focus remains on cross-stakeholder group collaboration to promote participation and true conformance to ensure benefits are realized by library users is sometimes challenging."

A second issue, along similar lines, continues to center on transparency of indexing and linking. Jenny Walker comments, "I would like to see an environment in which there is transparency regarding the availability of content, fair linking, etc., and that vendors of discovery services can focus predominantly on user experience, addressing the rapidly evolving needs of the users."

Finally, an issue that the original ODI Working Group decided to pass on to the current group was an audit process. As Kessler notes, "Some vendors are not interested in participating because of the lack of an audit process. While I would have thought that an audit would intimidate vendors from participating, these vendors feel that without an audit, the ODI seal of approval loses its value." Defining the framework for such an auditing process, and establishing a method for implementing it, would represent major efforts. At the same time, Kessler says, "I would, perhaps naively, argue that ODI is founded on principles of transparency and honesty, and continuing along those lines I would hope that an audit would not be necessary."

The Open Discovery Initiative has laid out a path for open communication and shared understanding around large-scale discovery services – but that alone is not a solution. It is important that libraries continue to work with their vendors to understand the level of ODI conformance currently in place. All parties must come together to ensure content coverage and increase the level of disclosure by discovery services and content providers. The success of web-scale library discovery services does not rest with one sector alone and ODI is only as productive and beneficial as the diligent, well-intentioned efforts invested by each member of the scholarly communications community.

References

- NISO ODI Working Group (January 2013). Retrieved from
http://www.niso.org/apps/group_public/download.php/9977/NISO%20ODI%20Survey%20Report%20Final.pdf
- Open Discovery Initiative Working Group (October 16, 2013). *Open Discovery Initiative: Promoting Transparency in Discovery*. Retrieved from
http://www.niso.org/apps/group_public/document.php?document_id=11606
- Open Discovery Initiative Working Group (June 25, 2014). *Open Discovery Initiative: Promoting Transparency in Discovery*. Retrieved from
http://www.niso.org/apps/group_public/download.php/14820
- Open Discovery Initiative Working Group Roster. Retrieved from
<http://www.niso.org/workrooms/odi/roster/>
- Walker, Jenny (August 10, 2011). *Proposed NISO Work Item: Open Discovery Initiative: Standards and Best Practices for Library Discovery Services Based on Indexed Search*. Retrieved from http://www.niso.org/apps/group_public/document.php?document_id=7175

Box: Open Discovery Initiative Kick-Off Meeting

Open Discovery Initiative

AGENDA

1. Introductions
2. Short overview of the current landscape and background to this initiative (Marshall Breeding)
3. Benefits and values for stakeholders
 1. The library perspective – Marshall Breeding, Vanderbilt University
 2. The discovery vendors – Oren Beit-Arie, Ex Libris
 3. The information providers – Bonnie Lawlor, NFAIS
4. Suggested areas of focus – policy/business issues as well as technical issues
5. Statement from Todd Carpenter, NISO
6. Next steps

Meeting attendees:

Hosts

- Marshall Breeding, Vanderbilt University and independent consultant
- Oren Beit-Arie, Ex Libris
- Jenny Walker, Ex Libris

In-Person:

- Kristin Antelman, North Carolina State University and OLE
- Laura Krier, CDL
- Joe Luccia, University of Villanova and VuFind

- 
- Judy Luther, Informed Strategies
 - Nettie Lagace, NISO
 - John McDonald, Claremont Colleges (also on COUNTER board)
 - John Meador, University of Binghamton
 - Laura Morse, University of Harvard
 - Michelle Newberry, FCLA and ICOLC
 - Kari Paulson, EBL
 - Roger Schonfeld, Ithaka/JSTOR
 - John Tagler, AAP
 - Mike Teets, OCLC

By telephone:

- Renny Guida, Thomson Reuters
- Bonnie Lawlor, NFAIS