OPEN DIALOG: CHRIS BULOCK, COLUMN EDITOR

Finding Open Content in the Library Is Surprisingly Hard

Chris Bulock
Collection Access and Management Services, Oviatt Library, California State University Northridge, Northridge, CA, USA

ABSTRACT
While library discovery systems offer great promise for delivering open access materials, they often fall short. Discovery tools merge multiple records together, and in the process show a preference for linking options that are less likely to successfully deliver open access full text. Third party tools offer options for bringing users to missing content, but these tools are often not fully supported in library systems. Library systems are so focused on keeping unauthorized users out of restricted content that it compromises their ability to get users into content that is freely available.

KEYWORDS
Open Access; library discovery tools; link resolvers; library service platforms

Introduction
Open access (OA) publishing has the promise to bring simplicity and flexibility, but this is far from the reality of accessing OA resources in library search platforms. Library systems include a great deal of complexity, and they are inherently systems of barriers. Link resolvers are simply not up to the task of navigating hybrid journals and multiple versions of articles (Jones, 2014). Library discovery tools showed much more promise for successfully delivering OA results, as they include article level metadata and may include direct links to articles as well. However, libraries have been using these systems for several years now, and they have proven to have their own pitfalls. While some discovery tools allow for the incorporation of outside tools and plugins, these come with their own drawbacks and limitations as well.

Throughout this column, I will draw on my own experience working with library systems. My library implemented the Alma library service platform and Primo discovery service from Ex Libris in 2017. From 2018 to 2020, we also used the oaFindr service, which integrated with Primo to deliver OA articles. Given that each library’s implementation of these services may function somewhat differently, I will attempt to outline these examples clearly enough to illustrate my points without dwelling too much on the technical details of our particular implementation. This column also draws on research I did with Heather Cribbs, which we presented at the eCAUG (the Ex Libris Users of North America California User Group) 2020 conference (Bulock and Cribbs, 2020).

Many copies
One of the fundamental difficulties of library discovery tools is sorting out the multiplicity of records and links, especially knowing that many of them are for the same item. Article level metadata for a single article may be available from a journal publisher, from an aggregator, from an institutional repository, from a subject repository, and from additional parties like CrossRef. In some cases, these records may pertain to different versions of the article. For example, there may be records in the search system for a copy of an article that the author deposited in an OA institutional repository as well as for a closed access copy from the publisher of record. In many cases, though, and particularly in the current Central Discovery Index (CDI) implementation in Primo, these multiple records may end up merged into a single record to present to researchers (Ex Libris, 2021b). This certainly makes sense; most researchers would find it frustrating to stumble upon seven records for the same article in every search. However, the merging of records has implications for delivering full text to users.

CONTACT Chris Bulock christopher.bulock@csun.edu Collection Access and Management Services, University Library, California State University Northridge, 18111 Nordhoff St., Northridge, CA 91330, USA. © 2021 The Author(s). Published with license by Taylor & Francis Group, LLC. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
To go with the multiplicity of records is a multiplicity of links to deliver full text access. There are two basic methods to link to full text in a library search tool. Some records may include a direct link to the content on a particular platform, with no connection necessary to other library systems. For example, the Directory of Open Access Journals (DOAJ) provides article level metadata to Ex Libris, complete with direct article level links to full text. Because the collection is comprised of OA materials, they can ensure that any library activating the DOAJ records in Primo would be able to access all content included, making these direct links a good choice.

Another delivery method is to rely on the library’s link resolver to construct a link to a copy available to members of the library. In this method, the metadata provider does not supply a link to full text. Instead, the supplied metadata may include a Digital Object Identifier (DOI) or perhaps just an article title, journal title, page number and date that the link resolver may use to match against the library’s knowledge base. With a match there, the resolver can construct a link to a platform where the library has a subscription or has determined free access is available. Interestingly, DOAJ also maintains a link resolver collection with Ex Libris. So while there are DOAJ article records available in Primo with direct links, DOAJ resolver links may also appear on records from aggregators or other metadata sources that do not include direct links.

Potential for confusion

There is the potential for multiple problems to arise in such a complicated system, from primarily superficial problems to those that hinder discovery. One of the lowest impact problems is the likelihood for OA indicators to appear without any OA links. The metadata that providers send to Ex Libris includes an indication of whether the resource is open or closed access. OA resources include a visual indicator showing they are OA, and a search can be limited to just those resources as well. However, as was previously mentioned, records are often merged together, and there is the possibility for records of OA versions to be merged with records of closed access versions. When this occurs, the OA indicator is retained. This is further complicated by policies regarding which links to display in such situations. An OA repository that relies on authors depositing articles will generally include isolated articles from any given journal rather than complete volumes or issues. This makes them unsuitable for link resolver systems, which can only track entitlements down to the issue level, not beyond to articles. As such, repositories usually supply direct links in the discovery system metadata they provide to Ex Libris. The general rule is that, if a merged record in Primo includes some records with direct links and some that relied on resolver linking, then the direct links are preferred (Ex Libris, 2021a). Heather Cribbs and I did some large-scale link checking and established that direct links tend to be much more reliable than resolver links (Bulock and Cribbs, 2020). However, Ex Libris reverses this policy when the direct links come from OA collections, then preferring the resolver links instead (Ex Libris, 2021a). Returning to our example of a merged record with an OA indicator, such a record will then only feature resolver links, often only those that lead to closed access resources with institutional logins required.

Offering only closed access links on a record labeled as OA is likely to cause confusion. This situation is not likely to materially affect the success of getting to full text for institutional users, but it will stop anyone without a login. On the other hand, the decision to privilege resolver links over OA direct links will likely result in more researchers hitting dead ends. Link resolver links are prone to errors of many kinds, and are especially ill equipped when an article is missing a DOI (Wakimoto, et al, 2006). Direct links, on the other hand, do not have such a difficult task to complete and are more reliable (Bulock and Cribbs, 2020). The decision to prefer resolver links for OA resources will end up preventing researchers from getting to those resources and represents a step back from the promise of discovery tools for providing better access.

Outside services

In order for OA materials to get into the hands of library users through the search tool, they need an entry point into the search tool, and they need a link to full text. We have mentioned some of the issues with delivery, including a preference for resolver links that puts OA content at a disadvantage. With major publishers offering hybrid journals, the records have a route into the system, but they rely on resolver links, and in the case where a library lacks a subscription to a journal with some OA articles, that may mean that those records do not show up as having any full text access. There are also issues with records getting into the discovery service in the first place. While organizations such as the DOAJ offer a way in, it still
requires that publisher provide article metadata to DOAJ (DOAJ, 2020).

Some services offer a way to get around some of these problems. While Primo offers some plugin functionality that can query a DOI against OA data like Unpaywall and insert a link, this is not reflected in searches that limit to full text results. Other services, such as 1findr, work in the typical structure of the discovery service, providing article level metadata for OA materials to Primo. While this is a truly promising approach, Ex Libris, at least, seems unprepared to fully support it. The crux of the problem is that OA items in a discovery service do not get a proxy prefix added, while closed access links do. This approach allows institutional users to access paywalled resources while preventing library staff from adding an endless stream of OA URLs and domains to the proxy server configuration file. Because oaFindr is a subscription service for accessing OA material, it ended up falling into a gray area. Its designation was changed several times during my library’s subscription, with the proxy prefix being added and dropped sporadically. This led to months at a time when our users would be unable to access any of the links in the records.

Conclusion

Library search tools do show a good deal more promise than previous generations of systems for delivering access to OA articles and other resources. However, that promise means little when implementation decisions are made to obscure access and favor older technologies like the link resolver. Many of the problems I have encountered came from library systems’ role as a means of getting only the right people into paywalled resources. Library systems are just as much about preventing people from researching as they are connecting people to resources. This emphasis on paid, protected access is sometimes just too difficult to reconcile with OA resources. As the OA share of scholarship grows ever larger, library systems providers must find more creative ways to bring the two together.

References

Bulock, C., Cribbs, H. (2020, June 30). Open access discovery and delivery in primo [Paper presentation]. ELUNA California User Group Conference online. https://drive.google.com/file/d/1H1s9T_wmAkAJHtjEZWspNnK9kMKAziI/view?usp=sharing.

DOAJ. (2020, August 6). “Myth-busting: Article content will appear in DOAJ once a journal is accepted for indexing.” DOAJ News Service. https://blog.doaj.org/2020/08/06/myth-busting-article-content-will-appear-in-doaj-once-a-journal-is-accepted-for-indexing/

Ex Libris. (2021a). “CDI and linking to electronic full text.” In Primo Central Discovery Index Documentation and Training. Ex Libris Knowledge Center. https://knowledge.exlibrisgroup.com/Primo/Content_Corner/Central_Discovery_Index/Documentation_and_Training/Documentation_and_Training_(English)/CDI__The_Central_Discovery_Index/CDI_and_Linking_to_Electronic_Full_Text

Ex Libris. (2021b). “Match and merge in CDI.” In Primo central discovery index documentation and training. Ex Libris Knowledge Center. https://knowledge.exlibrisgroup.com/Primo/Content_Corner/Central_Discovery_Index/Documentation_and_Training/Documentation_and_Training_(English)/CDI__The_Central_Discovery_Index/Match_and_Merge_in_CDI

Jones, M. A. (2014, June 28). Discovery of open access articles in hybrid journals: What role does the library play? [Paper presentation]. 2014 American Library Association Annual Conference, Las Vegas, NV.

Wakimoto, J. C., Walker, D. S., & Dabbour, K. S. (2006). The myths and realities of SFX in academic libraries. The Journal of Academic Librarianship, 32(2), 127–136. https://doi.org/10.1016/j.acalib.2005.12.008