



A Guide to Evaluating Content Neutrality in Discovery Systems


ExLibris[®]
a ProQuest Company

■ Importance of Content Neutrality

Content neutrality is fundamental to the library's mission. Users expect that whatever search tool they use exposes and delivers access to all pertinent resources in the collection in an unbiased manner. This is deeply concerning to many within the library industry, as evidenced by the work of the Open Discovery Initiative. Two years in the making, it involved more than two dozen individuals representing multiple facets of the library community.¹

Content neutrality means that search results lists are ranked based on objective measures such as their fit to the user's search and intent. It means that full-text access is directed according to the quality and reliability of the service available to an institution, as determined by library staff. In other words, content is discovered and linked according to its relevance and quality, not according to any commercial interest a vendor might have.

Content neutrality, therefore, is deeply rooted in the very definition of librarianship. Librarians, after all, are not just gatekeepers; they are also guides and connectors. They know what their students need, know the institution and its mission relative to the content they acquire, and the content that is most important in their collections.

Librarians and libraries need assurances that the discovery tools they provide fully support their

essential role. They want and need users to be able to discover resources, unencumbered and unimpeded by external motivations. The reputation of the library itself, and the value of it, depends on how well the discovery service fulfills users' expectations.

Content neutrality in a discovery system means that students and researchers are equally exposed to the entire wealth of information from all sources. If

Discovery tools must be unencumbered and unimpeded by external motivations.

the return of results or their presentation is influenced by factors that in turn influence a user's ability to select or access relevant

content, it interferes with the library's mission.

The library's mission is to provide the user with as much information as possible to support research and allow the user to critically evaluate the results. Libraries invest large sums in licensed and purchased content. If search results are based on the dollars at stake, neutrality can take a back seat. Discovery vendors' declarations of neutrality aside, libraries are the stewards of their users' experiences. Therefore, libraries need to make an effort to understand, control, and ensure the content neutrality of their discovery systems.²

How does a library ensure that its discovery tool of choice is content-neutral? We present here a standards-organization viewpoint, some principles for evaluating content neutrality, and finally a set of questions that libraries should ask.

If other factors influence results, they interfere with the library's mission: providing unbiased access to content.

■ Open Discovery Initiative

One useful starting point is the recommendations of NISO's Open Discovery Initiative. The ODI working group's output was primarily focused on recommendations for metadata sharing among discovery providers, content providers and libraries, but also proposed protocols related to "fair linking" by discovery systems.³ Especially relevant to content neutrality, they included:

- Non-discrimination among content providers in how results are generated, relevance-ranked or link-ordered
- Mechanisms that allow libraries to establish preferences as to link targets, order or priority when the same content is available through multiple content platforms
- Library-configurable options in other aspects of link presentation such as number of links, link order, and labeling and branding of full-text links
- Seamless linking to the A&I interface, when the A&I metadata was incorporated in building the index that made the item discoverable
- Disclosure to content providers and libraries when changes could impact result sets, relevance ranking or link ordering
- Disclosure of any business connections with content providers, whether direct or indirect ownership or through negotiated or paid agreements
- Affirmative statement of the neutrality of their algorithms, and especially for content provided by the discovery provider

A fundamental principle of content neutral discovery is ensuring that metadata processes do not affect discovery and delivery options.

The ODI working group in its recommendations stated that while full disclosure of search-relevancy ranking algorithms should be encouraged, discovery vendors may rightfully consider them to be proprietary trade secrets. However, the question as to whether a vendor's methods result in biased results was a key question that figured into their recommendations.

■ Content-Neutral System Design

Libraries should call for discovery providers to commit to content-neutral practices. We present the following principles.

Make content equally discoverable. Fair and equitable treatment of the metadata that the discovery tool provider uses is fundamental to content neutrality. A particular concern is how to manage duplicate records when a resource is available from more than one content provider. There are multiple approaches used today. A match-and-merge-records methodology aggregates all of the available data from all sources into a single record. A grouped-

records approach uses the ISSN/ISBN or a unique set of identifiers as a key to tie duplicate items together, enabling users to expand the 'super record' and view the specific items. A third approach keeps all records separate and distinct, where some records will have "richer" metadata than others, and thus will be ranked higher and linked to. Whichever method is used, a fundamental principle of content neutral discovery is ensuring that discovery and delivery options are not affected due to metadata handling or processes.

Ensure that technical considerations are balanced.

To prevent biased results, it is important that the relevancy-rankings approach does not introduce a bias by favoring one provider over another. The merging and grouping approaches help to improve discovery of the resource without eliminating resources from consideration, or favoring one publisher or database provider at the expense of another. Relevancy ranking algorithms that value one metadata record over another record when both reference the same content create an opening for content discovery and delivery bias. Regardless of the technical details, a fully objective relevance ranking orders results in a neutral way, without a bias toward any given content provider.⁴

Keep platforms separate, in delivery and in visual presentation.

When a content provider also furnishes a discovery system, each should run on a separate platform. The two should be distinct and not built off of or layered upon one another, both in technical implementation and in presentation. Branding elements in the discovery system that reference the content platform promote that platform at the expense of others. Avoiding brand extension through the discovery service is critical to ensuring content neutrality.

Make content-neutrality the default. Libraries shouldn't be burdened with making extensive changes to the system configuration just to achieve content neutrality — for example, in order to link users to the preferred full text. Systems should be content-neutral by default. Librarians should be able to control resource preference, with linking preferences easily customizable

through a simple and straightforward interface. However, discovery services that push all responsibility for content-neutral linking configurations onto the library make it less likely that the goals of equal and fair discovery will be achieved.

Make delivery equitable. Access to full text should be treated the same regardless of source. Applying special type treatments or graphic elements such as a .pdf icon or logo for some providers but not others influences user selection by making those links more attractive. Ideally, clicking on an article title should take the user directly to the article, similar to a Google user experience.⁵ However, when the interface takes the user to a list of links, that presentation should also be unbiased. Link lists that intermix clear and confusing names exert their own influence over user selection.

Provide a neutral user interface.

The user interface should be impartial and not influence the selection of resources. All content that is accessible through a discovery-user interface, whether it is the result of index-based searching or federated searching,

should appear in the same set of results. Otherwise, all content from all providers is not being treated equally. Moreover, if a specific database or content provider is consistently presented as providing more matches across multiple searches, it leads to the supposition that it is a better source in general. This can potentially lead to familiarity bias where one resource becomes the “go to” resource for a user. Resource preference if it exists at all should be in the hands and control of the library.

Relevancy ranking algorithms that value one metadata record over another when both reference the same content create an opening for discovery and delivery bias.

■ Questions to Ask

Librarians should closely evaluate discovery systems for bias. They should also work closely with discovery vendors to ensure that they follow the guidelines and principles at the core of content neutrality. Here are some sample areas of inquiry that librarians should consider.

- How does the system protect a content-neutral approach?
- Are records handled in a way that doesn't filter out data?
- Does the handling of duplicate items systematically favor one content provider over another?
- Does the relevancy-ranking approach weight items disproportionately?
- Does the discovery provider enrich the metadata of some content provider records, but not others, which influences the discovery of those records?
- Are all content providers treated equally in the display of results?
- How do the user interface and experience maintain a provider-agnostic approach?
- Are there any aspects of the interface that bias the user, or foster future bias?
- Is the delivery of data influenced by any factors that the vendor controls?

- Does the vendor do anything from a technical standpoint that overtly favors the vendor's own content?
- If the return of results seems to bias in favor of the vendor's own content, is that circumstantial or deliberate?
- If there are elements such as URLs, return of results, platform elements, results display and interface design that seem to favor a vendor's own content, is that incidental or intentional?

■ The Need for Transparency

As discovery tools become more-popular avenues for accessing content across collections, libraries have a critical stake in how search results are ranked and presented. Furthermore, they not only have the right to control preferences over the priority of resources, but also the right not to be influenced by the system or the platform. After all, beyond the purely academic standpoint, content neutrality is important from the standpoint of collection use and therefore collection development which goes to the very core of what libraries do. Until we achieve greater collaboration among all parties involved — libraries, discovery vendors and content providers — librarians need to scrutinize the neutrality of their systems, ask questions and receive satisfactory answers.

¹ Open Discovery Working Group (2014). Open Discovery Initiative: Promoting Transparency in Discovery. *NISO*.

² Amira, Aaron (2015). Collaboration, Exclusivity, Choice Oh My! (Presentation).

³ Open Discovery Working Group (2014).

⁴ Breeding, M. (2014). Library resource discovery products: context, library perspectives and vendor positions. *Library technology reports*, 50(1), 33-59.

⁵ Shi, X. and Levy, S. (2015) An Empirical Review of Library Discovery Tools. *Journal of Service Science and Management*, 8, 716-725.

About Ex Libris

Ex Libris, a ProQuest company, is a leading global provider of cloud-based solutions for higher education. Offering SaaS solutions for the management and discovery of the full spectrum of library and scholarly materials, as well as mobile campus solutions driving student engagement and success, Ex Libris serves thousands of customers in 90 countries. For more information, visit <http://www.exlibrisgroup.com>