

What's Next for Library Discovery? Insights from North Carolina State University Libraries' Discovery Retreat

Hiva Kadivar and Kristen Wilson

Abstract

This paper discusses a one-day staff retreat held by the North Carolina State University Libraries to explore the future of library discovery, and the outcomes from that retreat. It describes how we defined, planned, and executed the retreat, and summarizes and reflects on the themes of identity, future direction, and challenges that surfaced during retreat discussions. The future directions addressed at the retreat include linked data, machine learning, social media presence, discovery outside the library, and whole library discovery. The paper also includes details of how we have used the outcomes of the retreat to advance new work within the Libraries.

Keywords: library discovery, library catalogs, artificial intelligence, social media, retreat planning

Introduction

For North Carolina (NC) State University Libraries, the two decades between 2000 and 2020 saw rapid innovation in discovery services as books and journals moved online. We adopted popular web technologies and we invested in building our own user-focused tools. More recently, many staff members have remarked on how our pace of

innovation seems to have slowed, and today's trends feel harder to grapple with. How do we adapt to a rapidly changing environment in which our users increasingly turn to Google, ChatGPT, and TikTok to conduct research and complete their assignments?

To help us reconnect to our culture of innovation, the Libraries held an in-person, one-day retreat on September 12, 2023, to consider the future of the Libraries' discovery environment in light of local priorities and new trends in the field. The concept of discovery was considered in a broad sense to encompass all library systems and workflows that facilitate the discovery of the library as a whole, including resources, services, spaces, technology, and people.

We extended that holistic approach to planning the retreat by involving leaders from several different departments that engage with the concept of discovery. The discovery retreat planning team included: Josh Boyer, User Experience; Maria Collins, Acquisitions and Discovery; Adebola Fabiku, Access Services; Hiva Kadivar, Digital Library Initiatives; Andreas Orphanides, User Experience; Jason Ronallo, Digital Library Initiatives; and Kristen Wilson, Digital Library Initiatives.

The retreat focused on surveying the discovery landscape, discussing the background and state of existing systems, and identifying ideas that can be expanded for further investigation, pilot projects, and experimentation. To maximize our time together and facilitate open discussions at the retreat, we considered it out of scope to make decisions about what will be done or when it will be done. With participation from thirty-five Libraries staff and senior leadership across nine different departments, the retreat included presentations, lightning talks on emerging technologies, new products and services, as well as breakout sessions and large group discussions.

Identity

Throughout the retreat, staff reflected on identity in terms of who we are, what we do well, and where we want to excel. Some major themes

that manifested in these reflections were related to our organizational culture, staffing, and discovery tools.

Retreat participants noted innovation, curiosity, and willingness to adapt and pursue new ideas as strong elements in our organizational identity. We quickly and dynamically respond to user needs and trends and adopt an evidence-based approach. We have a culture of valuing and prioritizing accessibility, user experience (UX) and design, eliminating barriers to access by creating usable interfaces that do not sacrifice results.

In terms of our staffing identity, we have excellent staff and strong in-house technical, development, user experience, metadata creation, and collection building skills. We value iterative development. Although we have the willingness and technical expertise to pursue new technical projects, participants noted that we do not necessarily need to excel at the most technically difficult things, but need to excel at the ones that impact users the most. We take pride in our culture of collaboration and want to excel at protecting our staff from burnout and from being under-resourced.

In addition to staffing, many discussions focused on how we have created many high-quality discovery tools with a strong UX focus: QuickSearch, the Drupal website, the Catalog, Special Collections tools, Journals, and Databases. Participants noted that we have been keeping up to date with general discoverability standards and adapting them to the library environment. For instance, our QuickSearch tool enhances discovery by offering autocomplete, categorizing search results by format, and integrating information from many sources such as the Catalog, website, journals, databases, people, technology, and events. When it comes to our discovery tools, we are not just an academic library that makes use of out-of-the-box, vendor-supplied tools, but one that aims to do discovery better than the vendors and enhances it further based on our users' needs.

Challenges

Throughout the retreat, staff described a number of challenges facing the Libraries as we try to advance our discovery tools. These challenges

relate to rapidly changing student expectations and behavior, the siloed and proprietary nature of scholarly content, and the staffing models needed to achieve ambitious goals.

Before the retreat, Josh Boyer asked students to post sticky notes listing their favorite discovery tools and research tips. The responses encompassed a wide array of tools including search engines, AI assistants, social media, and databases of pirated material. Although some users are still turning to the Libraries' search tools, and others no doubt make their way to them eventually, we can no longer assume that students see the library as the natural starting point for their work. We need to focus more on understanding how and where students prefer to work and meeting them where they already are. One of the challenges in doing this will be learning about emerging needs and responding to them quickly enough to make an impact before trends and students move on. We will also need to figure out how to insert ourselves into spaces that we do not control or that we have not typically inhabited.

In addition to looking outward, many of our discussions focused on ways to optimize use of our own data and resources, such as incorporating linked data and machine learning into our discovery tools. But unlike many commercial organizations that have full ownership of and access to their data, the Libraries' data is spread across many environments. This occurs locally, as we store data in the Catalog, Special Collections tools, Drupal, a locally developed electronic resources management system, and other locations. Beyond our own tools, much of the electronic content we provide is owned by publishers and vendors and stored on authenticated third party websites. Getting access to the full spectrum of data needed for these types of projects and figuring out how to make connections across them will be another major discovery challenge.

Retreat participants also noted challenges about the Libraries' current staffing, commitments, and priorities, and how these might negatively affect our ability to innovate. On the retreat's wrap-up questionnaire, a clear majority of responders listed challenges related to

inadequate staffing, and many noted that a lean workforce is not well-positioned to do creative or innovative work. People also listed concerns about the Libraries' current maintenance commitments and level of technical debt. Our long history of creating our own tools means that we have a very large portfolio that requires constant upkeep. Finally, moving into new areas of work may require us to shift the way we think about our effort. If we want to focus more attention outside the Libraries' boundaries, what internal commitments can we give up? No matter what directions we decide to move in, we will need to address the challenges of ensuring we have the staff, time, support, and intellectual and emotional breathing room to succeed.

Future Directions

AI and Machine Learning

Although there is a lot of hype and fear about AI, we focused our retreat discussions on practical ways that AI and machine learning could help us improve the discovery experience. The ideas we came up with included using AI to create new ways of searching, improve search results, and create customized content, rather than on generative AI tools like chatbots or assistants. Specific ideas for AI and machine learning include:

- Implement semantic search within the Libraries' search tools to better identify relevant results using terms that are not explicitly part of the original search.
- Recognize the type of material a user has searched for and provide customized results. For example, if the search appears to be a book title, we would prioritize Catalog results; if it appears to be a question about study rooms, we would prioritize our website results.

- Provide user recommendations based on similar searches, course affiliations, trending topics, and other sources.
- Use machine learning and natural language processing to support staff workflows like metadata creation and data clean up.

Linked Data

At NC State, we have done some experiments with linked data, but we have not yet figured out how we can use it more systematically. Retreat participants focused their discussion on the potential for linked data to enhance our searching and make our data more discoverable outside of library environments. Specific ideas for linked data include:

- Revisit our linked data approach holistically and determine what areas of work are worth investing in, possibly with help from outside experts.
- Create linked data connections between MARC data and Drupal data so that we can connect bibliographic resources to people, events, and spaces.
- Publish linked data about our resources that can be consumed by outside sources.
- Create a knowledge graph that creates links across internal data sources and with outside data sources and vocabularies.

Whole Library Discovery

Jason Ronallo presented a lightning talk about whole library discovery, which is the concept of providing our users with a holistic view of the Libraries' resources including collections, services, spaces, equipment, and people. The goal of whole library discovery is to present the user with this holistic view regardless of where they have started their search. Whole library discovery is deeply entwined with machine learning and linked data, which are different and potentially complementary strategies for creating the connections needed to discover

information from many sources. Specific ideas for whole library discovery include:

- Look more deeply at QuickSearch and Catalog logs to see what users are looking for.
- Embed related Drupal search results in the Catalog.
- Create additional structured data and endpoints within Drupal to support related results.

Discovery Outside the Library Environment

Knowing that many users do not start their search from the Libraries' website, we talked about ways that we can reach them in the spaces where they are most likely to be working already. This is an especially challenging direction, because we do not have control over these spaces and focusing our attention on them takes away from work on our own tools. Retreat discussions focused on getting library data and content into popular services and creating content that is appealing and familiar to students. Specific ideas for discovery outside the library environment include:

- Investigate getting Catalog records indexed by search engines—the goal is that a user can search “NCSU [title]” and get a useful result.
- Investigate browser extensions and tools that can be used across websites—We could improve an existing prototype extension or look at vendor products.
- Tie-in to social media: Use social media to create more useful and appealing search results across non-library websites, i.e., not just text-based results.

Social Media

Beatrice Downey, Libraries Fellow in User Experience and Research Engagement, presented a lightning talk on using social media as a

discovery tool. She conducted informal user research on instances where, rather than going to Google or other discovery platforms, users go to social media platforms first to find information. She also conducted an environmental scan of professional social media websites, within and outside of libraries. Many libraries are utilizing social media as a discovery tool by promoting library services and events, reaching library users and non-users, teaching information and media literacy and how people can use the library services.

This talk inspired many discussions on social media engagement throughout the day. Retreat participants had conversations on how we can make better use of social media to humanize the library in places where our users are already looking for information, like TikTok and Instagram. Specific ideas for social media engagement include:

- Create engaging short-form content, like videos, that highlights the Libraries' useful tools and services, and connects users to our existing long-form content.
- Create short-form content that highlights simple tips and tricks, broadly useful tools, and how to get the most out of popular non-library tools.
- Foster collaboration opportunities with student organizations who already have established social media presence.
- Embed social media content into the Libraries website to create more discovery paths.

Follow-Up Projects

Search Engine Integrations

One of the most popular ideas from the retreat was allowing results from the Libraries' catalog to be indexed by Google and other popular search engines. Anecdotal experience suggests that many users begin their research with a Google search, and user research has confirmed

this assumption. A 2022 internal user research study about streaming video access described a user who first Googled “ncsu” followed by the title of a video.¹ When this failed, she amended the search to “ncsu library video” followed by the title. Neither of these searches were successful because none of the Libraries’ catalog records are indexed by Google, but it would have been clearly beneficial for the user if a search of this kind had succeeded.

There are challenges in allowing our catalog to be harvested by Google’s web crawlers. Frequent visits by web crawlers would increase the amount of the traffic to the catalog, as well as the load on the application programming interfaces (APIs) provided by our integrated library system (ILS) for live availability checks. Concerns that web crawler activity could bring down our catalog or ILS have been the main reason we have not attempted search engine integrations in the past.

In light of the excitement generated by the discovery retreat, staff in the Digital Library Initiatives department decided to revisit the possibility of exposing our catalog to search engines and investigate strategies to prevent excessive traffic in the process. We considered the possibility of creating static versions of our catalog pages that would not include live item availability information from our catalog, which could be visited by crawlers without triggering API calls. We completed a pilot project by creating a small set of static pages, but they were not properly indexed by Google, so we did not continue with that approach. The project was put on hold due to competing priorities, but we hope to return to work in this area soon. Future strategies include making better use of site maps to control what pages are crawled and increasing overall resources to our servers.

Semantic Search

The discussion about AI and machine learning at the retreat focused primarily on ways to use these tools to improve the search experience. One problem that can be addressed by semantic search is the ability

to match the user's search terminology with the vocabulary used in the search results, even if that terminology is different. For example, a search for "rust in metals" may still bring up results related to plant diseases known as rusts. The more relevant results may use a different term, such as "corrosion" or "tarnish" and the name of an alloy and are not found by a traditional lexical search.² Semantic search uses embedding models, a type of AI model that generates vector coordinates from text, then like using latitude and longitude coordinates, those vectors can be used to find which documents are the nearest neighbors to the query vector. This approach supports natural language searching and provides results that match the meaning of the user's search, rather than just the vocabulary.

This focus on semantic search at the retreat dovetailed with work in the Digital Library Initiatives department that had already begun and has since intensified. Jason Ronallo has worked on incorporating semantic search into the Libraries' Citation Index tool to improve the quality of search results, and Kevin Beswick has begun to prototype a semantic search in our QuickSearch tool, using data from Frequently Asked Questions (FAQ) articles.³ These proofs of concept have been valuable in helping Libraries staff understand how to work with large language models (LLMs) and identify the kinds of technical resources needed to support them.

As we learn more, we hope to expand the amount and types of data we can use with semantic searches and incorporate them into our production systems.

Browser Plug-In

One of the biggest challenges with library discovery, as discussed frequently in the retreat, continues to be that users are looking for what they need elsewhere. To that end, one follow-up from the retreat was investigating the offerings from the LibKey service including their Nomad browser extension. Niqui O'Neill was experimenting with a browser plug-in before the discovery retreat and after the retreat, Kevin Beswick extended that work.

One of the common things such browser extensions do is find DOIs and link to full text. Our prototype also matches on other identifiers such as ISBN and we are looking to extend it further. In the prototype, a button next to the identifier connects to the full text when available and allows the user to quickly request the full text when unavailable.

We wanted to go even further by incorporating resources beyond the catalog. The extension also shows a button in the browser toolbar, which users can click on to bring up recommended items and services from the Libraries related to the content on the page they are currently visiting. If you are on a page to learn Python, for instance, clicking on the extension button can show that we have various resources related to Python such as workshops and books. We hope this is another way we can get closer to whole library discovery. A significant challenge to this approach is trying to encourage adoption even if we develop a tool that is useful.

Conclusion

The NC State Libraries' discovery retreat proved to be a successful way to reflect on the Libraries' history of discovery projects, explore new ideas, and spur innovation. The retreat helped staff to explore our identity as a creative and user-focused workplace and affirm that we will continue to uphold those values. It provided us with a chance to talk about the challenges we currently face, including pace of change, loss of control, technical debt, and staffing. It also helped us identify promising future directions that will allow us to maintain our culture of innovation and meet some of our challenges. To be successful, we will need to meet our users where they are, even if that is outside of traditional library spaces, and work to understand new technologies and implement them within a library context.

The discovery retreat has prompted staff to begin new projects, including an investigation into search engine indexing and development of a browser plug-in prototype. It has also affirmed the work

already being done on AI and machine learning and underlined that this is a direction we must continue to pursue. The retreat was the source of a great many new ideas, and there remain some areas where we have yet to make real progress, including social media and linked data. But if nothing else, it raised awareness of these needs, which may lead to future projects as opportunities arise.

Staff feedback on the retreat and its outcomes were overwhelmingly positive. Attendees praised the importance of the subject matter to library users, the chance to be creative, and the collaborative spirit of the event. There is still much work to be done, and we hope that the inspiration provided by the retreat will help us be ready to engage and innovate to improve discovery for our users.

Contributor Notes

Hiva Kadivar is Libraries Fellow, North Carolina State University, Raleigh, North Carolina, USA.

Kristen Wilson is Discovery Systems Manager, North Carolina State University, Raleigh, North Carolina, USA.

Notes

- 1 North Carolina State University Libraries, "Panopto Usability Testing: Mini-Study" (internal user research study, June 14, 2023), accessed December 19, 2024.
- 2 Jason Ronallo and Kevin Bewick, "Semantic Search" (PowerPoint presentation, North Carolina State University, Raleigh, NC, June 10, 2024), accessed June 27, 2024, <https://ronallo.com/presentations/semantic-search-ai-share-out-2024/#1>.
- 3 Ibid.