

10-1-2013

The Zen of Web Discovery

John Hubbard

University of Wisconsin - Milwaukee, hubbardj@uwm.edu

Follow this and additional works at: https://dc.uwm.edu/lib_staffart



Part of the [Library and Information Science Commons](#)

Recommended Citation

Hubbard, John, "The Zen of Web Discovery" (2013). *UWM Libraries Other Staff Publications*. 2.
https://dc.uwm.edu/lib_staffart/2

This Article is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in UWM Libraries Other Staff Publications by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

The Zen of Web Discovery

John Hubbard <hubbardj@uwm.edu>, University of Wisconsin–Milwaukee Libraries

Abstract: Many libraries have revitalized their role by offering an updated resource discovery layer. This article gives an informal overview of design principles and best practices for implementing a web-scale discovery system. Emphasis is placed on the proper service philosophy of supporting a new technology, with open communications and evidence-based customization.

“Be water, my friend.” - Bruce Lee

Introduction

The latest wave of “next-generation” search engines has been around long enough that any library without such a product—and which instead relies on a conventional online catalog and research databases to fulfill its members research needs—is nowadays more akin to an institution that offers an interface as outdated and cumbersome as the card catalog rather than a library devoted to the mission of accommodating modern information seeking behavior. Selecting, implementing, and maintaining a new resource discovery platform (which is often layered on top of an existing integrated library system and electronic resource subscriptions) may seem a daunting task, especially within an institutional environment that has historically been change-averse. This paper presents a few lessons learned from the successful launch of such a system (Ex Libris’ Primo platform) at the University of Wisconsin–Milwaukee.

My favorite library joke poses the question: “How many librarians does it take to change a light bulb?” The answer, of course, is “Change?!” Although libraries have long existed by resisting fads, refusing to embrace the latest technology no longer has the survival value it once did. The Internet has by and large marginalized much of what libraries used to do. Unfettered access to information is not necessarily a bad thing, just as few among us would bemoan the loss of closed stacks, save those who derived a misplaced sense of power from being an information gatekeeper. Web-scale discovery tools, moreover, have been around the better part of a decade, and are far from the latest flavor of the month in trendy innovations to library services.

Catalogs and databases are important tools in the search process, yet they are merely a means to an end, and are but a slice of a library user’s overall research process. Librarians who extensively deal with these search systems sometimes fail to grasp that the average person is not going to devote their career to understanding dozens of interfaces, settings, and advanced features. A typical searcher follows the principle of least effort and looks for results that are “good enough.” They are much happier to use Google than to map out a convoluted query with Boolean operators and nested Dialog-era commands. Rather than setting specific search criteria before doing a search, most sites are presently built on the use of facets to set post-search limits on results. This is a flipped way of thinking about mapping out search terms,

and instead simply trying some searches, and adjusting results as necessary based on what is retrieved.

Libraries need to do a better job at responding to preferences for easy-to-use interfaces. Even the most advanced library user spends more time on a site that is not the library's, meaning that adherence to prevalent standards is a necessary component of any good library website. There is instead often a disconcerting disjoint between how we think we should force people to search and how they chose to search regardless. If even the most stellar subscription content is bundled in a clumsy interface, people will turn elsewhere to obtain information. As much as we shake our heads at this—often with the same sort of bemused condescension that 19th century anthropologists exhibited toward alien cultures—it doesn't change the fact that many people use Wikipedia more than the library.

The challenge for libraries, therefore, is to provide a search experience on par with Google for accessing library content. We must embrace the changes that such discovery systems bring, as web users have already done by using search engines instead of library catalogs. When web-based catalogs first came into being, some librarians advised caution, and there were many sentiments yearning for the telnet interface if not the card catalog. Such resistance to technological progress is no longer warranted. The successful growth of research products marketed directly to individuals (e.g., Mendeley, JPASS, Udini) illustrate how libraries that continue to ignore these changes to information delivery do so at their peril.

Product Selection

The major players in the vendor landscape for web-scale discovery systems are EBSCO, Ex Libris, OCLC, and Serials Solutions. Each company offers a mature, cloud-based product with a useable web interface, extensive content coverage, and integration with existing bibliographic data such as a local catalog, and optionally, federated searching of other subscription databases. Similar to trends in the development of operating systems and web browsers, they are all moving towards a common look and feel with shared underlying content, and although measurable differences definitely remain, preferences for one system over the other may be based more on company reputation than any tangible distinctions.

A formal request for proposal (RFP) process is usually how libraries come to select and purchase a new discovery product. Depending on the levels of bureaucracy involved, the evaluation and procurement process may become protracted across several years. Obtaining funding for a finding tool that may not necessarily add anything original to a library's holdings, given these systems' price tags, may be difficult. Then again, a reference librarian, who also exists to facilitate access, provides no content either. While the features, coverage, support, and price of different systems are being compared, parallel planning discussions can take place to build consensus within the library and its parent institution, as well as across any involved library systems or consortia, regarding implementation.

Internal Preparations

Libraries face an uncertain future. More than technological obsolescence or political and economic threats, our profession is in danger of self-destructing by failing to adapt to meet people's needs. Implementing a new discovery system can be as much an upheaval for a library as it was to open the stacks or retire the card catalog. It is pivotal to address the fear of the unknown that librarians may misinterpret as cause to hold back an otherwise promising project. The planning process must be devoted to allaying such concerns and fostering buy-in amongst all effected library staff.

Sabotage from within an organization can derail almost any successful service. It is easy to see the ridiculousness from an outsider's perspective of a position that, "we can't switch to e-journals because our shelvers would be out of a job." Yet you may hear something along the lines of, "we can't change to a system catalog because we have several people devoted to building local records" stated in earnest by librarians who feel their livelihood threatened due to the evolution of stand-alone catalogs to a more shared environment.

As important as the logistical planning for designing, coding, and testing the system's features are, maintaining frequent and open communications about the process amongst all constituents, for the purpose of informing, educating, and cultivating support for the new service, is vital to a successful product launch. Even those not directly involved in implementation, especially employees whose workflow will be impacted, should be invited to feel ownership by having their voices heard about any upcoming changes. Reference librarians need to instruct people about the characteristics of the new interface; and catalogers must adapt to the new role that local records hold in a shared setting. Most importantly, any librarian will be in a position to laud or disparage a new discovery system and make a substantial impact on the library's marketing efforts. These individuals should all be focal points of a collaborative effort to explicate the major selling points and improvements that a new discovery service will bring.

Gathering such sponsorship and endorsements is a time-intensive process. Similar to how the bulk of most programming code is devoted to error handling, the majority of time taken on developing even the most technologically-oriented service will be spent on formal communications and informal discourse that are primarily designed to build consensus and garner support. As with website development, there may also be no clear delineation of responsibility and authority regarding certain decision processes. It is a testament to the worth of a library's web services that almost anyone working in a library is compelled to declare a final say about any particular component of that library's online presence. A constant stream of feedback and suggestions may at times be useful, but without administrative support to empower designated individuals to finalize design decisions, accommodating a barrage of demands from multiple perspectives—especially those that may not be representative of typical end user requirements—can instead create unnecessary and humbling roadblocks, and have a chilling effect of paralyzing an otherwise productive workflow.

Pragmatic Customization

One of the biggest problems with our profession is exemplified by the results from a web search for this: popular scholarly site:edu. This query will obtain hundreds of variations on what is essentially the same guide to the complementary characteristics of different publication types. The redundancy of this duplicated content—akin to the overlapping, just-in-case print collections that many libraries continue to hoard, ignoring the advent of rapid lending and direct acquisitions—signals a staggering amount of wasted effort spent by librarians. Aside from any lack of awareness over the ability to make use of shared resources, it is a consequence of the characteristic enthusiasm for control over every detail of the environment, and a temperamental disposition that is possessed by many members of the profession. However these traits are ill-suited when it comes to progressing with the pace of current technological trends, and have negative consequences on efforts to rapidly deploy systems which are for their part functional out of the box.

Modifications are not always a bad thing. It's fitting and proper to work towards improving usability. However, customization efforts often go beyond that. Installing extra scripts and plug-ins and add-ons is flirting with the mechanical equivalent of combined drug intoxication (i.e., some kludge crashing the whole system with the slightest update); excessive tinkering eventually creates an impossible workload to maintain, and is not usually necessary. Considering the importance of adhering to a standard interface, put another way, those who would give up uniform design to obtain a little local customization deserve neither.

Not everything needs to be customized. Certain adjustments are necessary to synchronize local metadata, enable patron authentication, and make available subscription resources display properly within the new interface. Yet no doubt due to the meticulousness of their clientele, many library vendors now offer a veritably paradoxical level of configuration options that go beyond basic branding choices into an array of aesthetic options and preferred wordings (many of which are altered, reset, or otherwise rendered dysfunctional with each maintenance patch and version update) that are not only unnecessary choices for each subscribing library to be burdened with, but also detract from the number of available librarian hours that would be better devoted to addressing the smooth operation of the new system's core features. Early OPACs needed to be heavily customized to make them not only usable but merely functional. This is not the case with web-scale discovery systems, any more than it is necessary for a library to make a tailored version of the Google homepage with the logo slightly smaller and the font one pixel larger.

Just as the sophisticated search syntaxes in the age of dial-up information retrieval systems have given way to the faceted and fuzzy discovery interfaces of today, it is necessary to embrace a loss of control over every precise detail of a search engine's form and functionality in order to keep up with its development. This is the same principle behind many individuals' "good enough" mentality, and it would do many a librarian good to learn to let go of the need for rigid overanalysis of every systemic detail when it is done at the risk of losing sight of larger issues.

Think of the analogous efforts that librarians began in cataloging individual webpages whilst search engines were off crawling millions of sites. Rather than being bogged down with design changes by committees and making substantial but optional modifications (arguably with dubious usability benefits—especially those requested solely from internal feedback and not based on testing public users), the best course would be to consider all changes warily. This is doubly true within a consortium environment, where aside from incorporating any necessary campus or branch nomenclature, any and all customizations should be made at the system-wide level, and not repeatedly or inconsistently amongst member libraries.

It is important that someone who fully understands what is technologically feasible be placed in a position undiluted in authority and be made responsible for determining which optional customizations should be implemented. Many information technology workers, librarians included, have shirked their duties by cultivating a “guru” aura (itself a remnant of the obsolete information gatekeeper) beneath from which it may be claimed that virtually any requested change is impossible. Short of breaking the laws of physics, any describable functionality is possible, if not reasonable, to program. Many seemingly easy changes, however, are conversely difficult to enact. Only by weighing the expected usability improvements against the necessary workload to make such possibilities a reality can the project be successfully managed.

Key Considerations

I am listing three characteristics of our new system that the UWM Primo Implementation Team deemed the most important options, based either in terms of the amount of time spent discussing various possibilities, or the significance of impact our adjustments had on a user’s search experience.

- *Prominence and presentation of “legacy” interfaces that were being largely replaced by the new discovery system.* Multiple tabs and search boxes on the library homepage were replaced with a single basic form that is directed at Search@UW (the external name for our Primo instance). From the homepage, diminished links to old and native interfaces are still available. Certain holdings searches are still best accomplished via the local catalog. Individual database interfaces, given the Primo Central Index’s vast coverage as a source for articles, won’t necessarily obtain more results on a particular topic, but may be more efficient to search within their isolated scope. It is critical that the efforts to funnel most traffic through the new discovery layer be congruent with and supported by the user services provided by librarians (i.e., you should never hear a library student being taught, “never use that search box on the homepage”). Documenting the relative strengths and weaknesses of the new and old systems as research starting points on public library guides is useful for this purpose.
- *What is included in a default search, and how different pre-search delimiters (called “scopes” in Primo) are labeled.* For most libraries, this should be all available holdings records and article collections as well as whatever digital repositories can be harvested.

Aside from a few foreign language news sources, which were not full-text and consequentially caused troublesome interlibrary loan requests, we have activated all contractually-allowed sources to populate results from the Primo Central Index. The default “Everything” search scope includes records for books from throughout the UW System. These materials can be obtained by patrons via an inter-campus delivery service. The optional search scopes are titled with relatively basic library terms (“Articles,” “Books & Media,” and “Digital Collections”). Although for example dissertations and book chapters are included under “Articles” while “Books & Media” is a label for all holdings, maps and scores and microfilm, etc. included, we opted for less pedantry in favor of more simplified labels. Lastly, we chose only to display, for article results, records with online full-text availability. This is a post-search setting that may be manually toggled by the end user, but the default configuration was retained this way as it is more representative of what most searchers want.

- *For records derived from local catalogs, which MARC fields are presented in the results display.* Finalizing this arrangement was a lengthy process, consensus for which was ultimately determined by a special task force working in cooperation with the vendor. Although this topic was the source of many fervent discussions amongst system catalogers, analogous to how local holdings records comprise a tiny minority of Search@UW results in comparison to items from other sources, it is arguably one of the more insignificant customizations made in the eyes of the general populace.

These issues highlight the importance of streamlined customization procedures wherein feedback may be solicited and decisions made without delaying the overall implementation process. Unfortunately, as seems to be the case with many library decisions, there is a point in design debates when a basic intuition of some sort is invariably invoked (‘This setting is helpful, yours is not’) and there seems to be no way to decide disagreements rationally, especially within an environment wherein staff may successfully repeatedly raise an addressed and resolved issue *ad nauseum* until whatever they want is implemented. To avoid such situations, we must endeavor to act with the best interests of the average user as a priority. All things being equal, trusting in the default settings provided by the vendor—even more than attempting to recreate the skeuomorphic features of legacy systems just because “we’ve always done it that way”—is a good rule of thumb for those valuing economy of effort.

Unintended Features

No library system is without bugs, and discovery products are certainly evidence of this. Many of the more cumbersome usability impediments in our new interface, however, can be traced to bad source data from erroneous catalog records that is dutifully displayed more prominently by a system that adheres to current resource description principles for bibliographic records, especially at times as a consequence of merging records across several system libraries with disparate cataloging practices. For example: (1) Multiple records with an “English” entry for their ISBN were de-duplicated, and shown as different versions of a single work, due to their matching

identifier field; (2) Different works with a common series title, such as “Great Books of the Western World,” were similarly “FRBRized” and displayed as the same edition; (3) Misleading “View Online” links showed up for records with 856j entries that only led to a table of contents.

Related to this last issue, it is also notable that many catalogs still contain vestigial holdings entries with deep links to subscription resources. These records, even if properly maintained, are duplicated by a dedicated link resolver or some other electronic serials list that is more capable of maintaining an accurate repository of a library’s complete online holdings. With the advent of new discovery systems and their potential abilities to call on electronic resource management systems to populate their database of items, it is past time to rethink the traditional catalog, no longer as something suited to be an exhaustive record of holdings, but rather as an inappropriate tool for maintaining an inventory beyond what is on the shelf in the library.

For records that are found to be malformed enough to cause problems with their display within a discovery system, cleanup efforts to standardize and restore the integrity of such MARC entries should be a priority. Especially within a consortium environment, this is an opportunity to review, ideally prior to a full product launch, decades of differently applied cataloging standards. This may seem a more daunting task than cleaning the Augean stables, however most library systems include powerful tools that can make global changes to data, rather than having to check and update individual records manually.

Since launching Search@UW, it has been widely adopted and accepted by most users, however a few shortcomings have come to light. There are concerns about the weighted results ranking and facets display. It has consequentially proven problematic for basic searches because it yields too many results, creating the opportunity for user education about forming a precise search statement, making use of limits and facets, and when to explore more specialized interfaces. Our hosted server performance has not been ideal. As with all limitations, it is important to provide a forthright appraisal of how and when to use the system and what workarounds are necessary. We have long desired this “Holy Grail” of a combined book and article search, but the enthusiasm for that system needs to be tempered with the realities of how the technology currently works. To this aim, a public library guide has been created for Search@UW that not only describes its features but also lists all known problems and issues for reference.

Ongoing Adjustments

There is no final product when it comes to a library’s web services. Vendors issue a steady stream of updated releases with new features, and librarians must continually adjust interfaces to meet patron expectations. Many of the changes made to our system were not completed before a public debut. After the considerable work to achieve basic functionality and maintain operations (in the three months surrounding launch, the internal UW System mailing list for Primo received over 750 messages), the people involved with Search@UW’s implementation have had the time to program additional usability improvements. Examples include minor

styling and interface adjustments, such as a script that changes patrons' default item retrieval pick-up location to their home campus.

It is important to solicit and acknowledge feedback from all individuals both throughout a live public beta and after the system is put into production. We have published a suggestion and error reporting form linked to at the bottom of every Search@UW page. This form has yielded constructive comments from multiple perspectives, and changes have and continue to be made based on such feedback. As mentioned earlier, making substantial customizations based either on internal suggestions that are not necessarily indicative of non-librarians' search behavior—or even formal usability testing that can carry the weight of enough demand characteristics that its ecological validity can be questioned—is no substitute for basing design decisions on genuine feedback from library users.

Just as change for change's sake is bad, given the size of our user base, there will always be negative reactions to anything different that we present. Some complaints may not even pertain to the updates that were made. With any such communication mechanism, there are invariably contrarian-minded critics, such as those that cling to historical research methods that are not applicable in the present information gathering technology landscape. These messages are opportunities to engage and inform individuals why all suggestions (including, as with previous upgrades, earnest requests to ditch the new system and revert to the old one) are not necessarily actionable. As Bill Cosby once said, "I don't know the key to success but I know the key to failure is trying to please everybody."

Conclusion

Many librarians are loathe to give up the control of being able to construct a detailed search query; reverse engineer results based on known content coverage; and tinker with virtually all aspects of an interface. With today's search engines this is not always possible. Web-scale discovery systems are the best hope for libraries to maintain relevance by saving the time of readers seeking quality results. For some librarians, learning to love these resources, which often bring imprecision and a perceived loss of control, requires the adoption of a different mindset. Practicing transparency in communications, establishing a well-defined decision process, and committing minimal customizations can make the implementation of a new discovery system a smooth process.

*

October, 2013

This work is licensed under a [Creative Commons Attribution-ShareAlike 3.0 Unported License](https://creativecommons.org/licenses/by-sa/3.0/).