Knowledge Creation Platforms

The Next Step after Web-Scale Discovery

Carl Grant

Abstract: Technology has long been used to provide access to the holdings of libraries. Yet, end-users continue to migrate towards web search tools. This article suggests attractive differentiation could be provided using business processes to develop “Knowledge Creation Platforms”, i.e. technology for accessing knowledge and allowing end-users to seamlessly start creating new knowledge using social networking, publishing it using Open Access and moving the resulting work into digital repositories.


1. Introduction

Digital technology has long been transforming the search/find tools used by academic library members and end-users. In libraries we have seen Online Public Access Catalogs (OPACs), federated searching, scholars’ portals, and more recently discovery and Web-scale discovery products. Yet librarians continue to find themselves in a constant and losing competition with other publicly available search tools, most notably Google and Google Scholar. As librarians look to provide value to their organizations through the services they offer their end-users, and at the same time, differentiate those offerings from these other tools. This article suggests one possible new offering: that of a Knowledge Creation Platform as the logical next step for achieving that differentiation.

2. History – How we arrived where we are today

Before we examine where we might go in the near future, we owe it to ourselves to take a brief look at where we have been and where we are today.

OPACs have consistently been a part of library technology for several decades and even today this technology continues to be the primary means for many library users to locate library resources. Not surprisingly, as this technology has aged it produced a large number of complaints due to its limitations. The XC Survey Report summarized it as follows: “Respondents were asked an open-ended question to name their ‘top three gripes’ about their OPAC, and then to list additional issues if they wished to do so in a fol-
low-up question. The top issues expressed in those complaints were:

a. Difficulty of customization (42 instances)
b. Inadequacy of search functions (31 instances)
c. Opacity of results and lack of grouping or faceting (27 instances)
d. Limitations of the user interface (16 instances)
e. Lack of Web 2.0 functionality (9 instances)
f. Backend problems (8 instances)
g. Lack of integration with databases or other systems (8 instances). There were also multiple complaints about:
   • Lack of access to data (7 instances)
   • Difficulty finding journals and the articles in them (6 instances)
   • Lack of updates (7 instances)
   • Lack of an API (Application Programming Interface; 6 instances)
   • Usability problems (6 instances)” (Foster/Bowen/Lindahl/Randall 2007:2-3)

Sadeh identified two problems with OPACs: first, they “are inherently librarian-centric; their design in terms of data structures and workflows is focused on library administration and hence severely limits the possibilities for the end-user interface” (Sadeh 2008:10) and Dempsey wrote: “The catalog emerged at a time when information resources were scarce and attention was abundant. However, the situation is now reversed: information resources are abundant and attention is scarce.” (Dempsey 2013:1)

Attempts to replace the catalog with improved technology first saw federated search tools. However, in the end the numerous issues surrounding that technology including slow response time, failed connectors and issues surrounding the Z39.50 standard upon which federated searching was largely based, ended up limiting this technology’s adoption.

Kress/Del Bosque/Ipri noted that “users expect library systems to behave like Internet searches and are frustrated when they do not. Some participants chose to avoid the complexity of the library web site by using Google. Web-scale Discovery Platforms have the most potential to remove some of the burden from the user and place it on the technology” (Kress/Del Bosque/Ipri 2011:163).

As a result of all these frustrations, accompanied by the emergence of Google as a search tool, academic initiatives such as the Scholars Portal and the North Carolina State University’s efforts with the Endeca software, a new type of interface emerged, which ultimately became known as a “discovery” tool. Per the “Code of Practice: Discovery Services” developed by the National Federation of Advanced Information Services (NFAIS), discovery is defined as one that provides “a ‘single search box’” to “access a central index of pre-indexed metadata and/or full-text” (NFAIS 2012:2).

Commercial discovery products have now been in the market since January 2009, when Serials Solutions first introduced Summon. In the time since discovery products were introduced, they have been successfully implemented in thousands of libraries around the world. Now, in addition, these products exist as open source software solutions (VuFind, Blacklight), collaborative solutions (WorldCAT by OCLC) and proprietary solutions (Summon by Serials Solutions, Primo by Ex Libris and Chamo by VTLS) among the many available choices.

In libraries today discovery has become the de-facto choice for any library wishing to provide single box search systems. Most recently, the continued development of cloud-computing based discovery solutions coupled with discovery has led to the evolution of “Web-scale discovery” as the latest iteration in this technology. Web-scale discovery is different in that it brings discovery tools together with a large aggregated index, containing content harvested from designated sources and it is a multi-tenant, cloud computing based service (i.e. not locally installed).

3. Some problems discovery tools still do not solve

As with any technology, today’s discovery tools are not perfect. While they have solved some problems, they have created new problems and have left some issues unaddressed.

For instance, it is not uncommon to hear concerns expressed about the lack of support for item specific call number searching, known item searches or the inability to search sub-sets. Some systems do not support user browsing very well and as metadata records are increasingly purchased through, and maintained by the vendor of the discovery systems, the quality of metadata is sometimes degraded.

Hofman/Yang (2012:262) state that typical problems not yet solved by discovery tools include:

1. “Significantly, only four of the 75 discovery tool implementations in the sample allowed for any kind of traditional index browsing, which may be one of the reasons why 96 percent of institutions using discovery tools offered access to both their discovery tool and their classic ILS OPAC” and,
2. “Also, as the study by Emde et al. (2009) suggests, faculty and graduate students tend to ignore changes to a library web site (such as a single search box) and continue to follow their tried and true search techniques. The same may be true of librarians and other expert searchers. Thus, providing users with a familiar interface may be another reason libraries keep access to their classic catalogs.”

Dehmlow (2013) recently wrote: “Debates over vendors’ unwillingness to share data have overshadowed other important features and functionality that are important to making discovery systems valuable”. Furthermore, he cites the:

1. Lack of a fully integrated request and delivery service.
2. Need to “build more intuitive connections between primary resources and the bibliography records that relate to those resources” (Dehmlow 2013:2).

He concludes: “These systems are incredibly useful and valuable for support of different research needs at different levels – getting novice users quick access to a handful of scholarly resources and augmenting the deep research process of our expert users.” (Dehmlow 2013:3)

Furthermore, web-scale discovery has not stopped the tide of users turning to Google, or other search engines, as their primary access/search tool. As noted by Camden: “Despite these improvements [discovery tools], most users have moved to Google and other search engines as their tools of choice. Often, the library catalog is used only when absolutely necessary (…).” (Camden 2008:1) This was further confirmed by OCLC’s 2010 Perceptions of Libraries Report, which stated: “Not a single U.S. respondent in 2010, including college students, began an information search on a library web site, down from 1% in 2005.” (OCLC 2010:95) Even more interesting was the announcement by Theo Engleman, University of Utrecht: “Delivery is the core business of libraries. Discovery? Not anymore.” (Kortekaas 2012) He then noted that the University of Utrecht has decided not to buy a discovery interface at all, believing that library data now needs to be delivered via APIs and web services, into other portals, interfaces and software that are being utilized by users.

So, clearly, issues remain in the current versions of discovery tools. Even if the list of issues above were resolved, the reality is that they would provide incremental improvement over today’s products. Would that be enough? Or possibly, as librarians we need to ask ourselves this question: How can we add such compelling value to the life of our users that they will reach out to us first in some aspect of their daily work/life flows centering around knowledge?

4. Differentiating between today’s discovery systems

To answer that question, first we need to look closely at the iterations available from various suppliers or open source initiatives. While local needs will always take precedence in determining which product is best for any one library, within the context of this paper the following differentiating characteristics will be important to note:

1. The discovery system should include an aggregated article index, preferably one developed and maintained by a separate supplier from the developer of the discovery interface so as to maintain the checks and balances in the content supply chain used by the library.
2. The product should be offered as a multi-tenant cloud-computing offering. Multi-tenancy keeps the costs low that are associated with running a product in a cloud-computing environment. In addition, it makes it easier to aggregate click-data about end-users.
3. User behavior and click-data should be dynamically aggregated for all institutions using the product. This is what will ultimately allow analytic tools to offer the ability to start driving pro-active services to end users, a radical change from the more reactive model of library service that is commonly found in libraries today.
4. All of the functionality available through the native interface supplied with the product, should be made available through APIs, or web services (REST calls) so that an organization could entirely replace the interface and deliver selective, but needed functionality into other interfaces where search/find/use of library knowledge would be beneficial.
5. Faceting should be under the total control of the library as delivering library resources into other interfaces may well require customization of this capability.

Even with that set of characteristics, the environment in which today’s discovery systems operate is changing rapidly. Dempsey points out that librarians need to realize “the catalog itself is being reconfigured in ways that may result in its disappearance as individually identifiable component of library service” (Dempsey 2012). Furthermore, he points out that “on the network, we increasingly expect services to be built around our workflows” (Dempsey 2012).
These are key statements and provide us with a logical transition point for considering what libraries need to offer next, either as a logical step beyond or in place of, today’s discovery tools.

5. Applying the Blue Ocean Strategy to web-scale discovery

While we would all like to believe what Larry Page of Google once said when asked: “The search engine of the future? A reference librarian with complete mastery of the entire corpus of human knowledge.” (Battelle 2006:252) We also realize this is not possible. A better strategy for shaping the future is outlined by Kim/Mauborne in “Blue Ocean Strategies”. This work points out: “The only way to beat the competition is to stop trying to beat the competition.” (Chan 2005:4) In other words, we need to simply understand that for the foreseeable future, our users will continue to use Google and similar search engines as a primary starting point for their searches and instead focus on areas where we can innovate and better meet the needs of our end-users by providing real value to their lives.

Matthews/Heron confirm this: “There is an explosion in the competitive market environment that is perhaps the most significant change facing libraries today.” (Heron/Matthews 2013:10) Competitive indeed. Challenging as well. Including the facts that we are facing flat or decreasing budgets, dealing with staffing issues, open data, open source software, open educational materials, open teaching and open courses. The list is long and impressive.

Technology has brought to us many of these challenges and, at the same time, many of the possible solutions. Unfortunately, we let the technology define us, allowing ourselves to be turned into a commodity set of services. Commodity services are where the “red, bloody waters of competition” occur, as outlined in “Blue Ocean Strategies”. When this happens I frequently encourage librarians to see this for the opportunity it is, to step back and perform an overall analysis of their services in relation to all those that are members/users have available to them so as to understand that competitive landscape. Doing that will give us the opportunity to find ways to use technology as a platform on which to develop new and valued services that are not commodity services. “The creators of blue oceans, surprisingly, didn’t use the competition as their benchmark. Instead they followed a different strategic logic that we call value innovation (…) making the competition irrelevant by creating a leap in value (…).” (Chan 2005:12) As librarians, this is what we need to do.

After reading “Blue Ocean Strategies”, one should next read David Lankes’ work “Atlas of New Librarianship”. There one finds a definition of the mission of librarianship, which should be widely adapted and only lightly modified to meet local circumstances. It says the mission of librarianship is “to improve society by facilitating knowledge creation in their communities” (Lankes 2011:15). By applying technology to this mission statement, librarianship could develop both the next step in the evolution of the catalog and, at the same time, substantial differentiation between existing web search systems and web-scale discovery products. In doing this, we must insure that the solution meets this goal: “Innovation for differentiation must be bold enough that, if it wins, it achieves separation.” (Chan 2005:8) Success in achieving that level of differentiation is, as noted by Heron/Matthews: “The result of a proactive approach to planning is that the library understand how it adds value to the life of its customers.” (Heron/Matthews 2013:40) So, that is our goal and if we achieve it we would open up “Blue Ocean” for libraries to sail across, at least the near-term future.

6. What would a Knowledge Creation Platform do?

To answer that question, let us start with an examination of the definition of the word components of the phrase. “Knowledge¹⁰, according to Webster’s Dictionary is “the circumstance or condition of apprehending truth or fact through reasoning”. While “Creation”¹² is “the act of making, inventing, or producing”. Finally, “Platform”¹³ is “a place or opportunity for communicating ideas and information”.

Then let us look at “Knowledge Creation” to see how it has been utilized. According to the Business Dictionary, “Knowledge Creation¹⁴ is the “Formation of new ideas through interactions between explicit and tacit knowledge in individual human minds”. This definition is based on the work of Nonaka. While his work is focused on management practices, much of what he wrote is applicable to today’s world of higher education. He wrote: “Innovation can be better understood as a process in which the organization creates and defines problems and then actively develops new knowledge to solve them. Also, inno-

---

vation produced by one part of the organization in turn creates a stream of related information and knowledge, which might then trigger changes in the organization’s wider knowledge system.” (Nonaka 1994:14-15)

Using those definitions to define where we are trying to go, we would then say: *A knowledge creation platform would be a place where users could apprehend truth or fact and determine how to solve created or defined problems and then openly communicate those ideas with others.*

Applying technology to providing this solution, the components of a Knowledge Creation Platform (KCP) in libraries could possibly include, but not be limited to:

a. **Discovery.** As noted above, knowledge creation stands on the shoulders of existing knowledge. So having the capability to look broadly and deeply into the existing domains of knowledge is an important requirement of a KCP. Yet to leave it there is to only begin to recognize the complexity and extensive nature of the various workflows of academic end-users, be they students, staff or faculty. Plus we know that finding knowledge is an iterative and interactive process. As we find answers to questions we have posed, it changes what we are thinking and thus what we want to ask. So, if we are looking to create differentiation and new value for our end-users, we must ask ourselves how can we best do that? Examining the existing workflows of end-users, one typically finds they are using a variety of existing tools, ranging from Google Scholar, library website/discovery/search tools, apps and numerous others. Overall, this leaves a very disjointed and clumsy workflow for them to utilize. It is out of this that the opportunity emerges to provide value and differentiation.

b. **Social Networking.** The KCP should integrate social networking within the interface. Hernon/Matthew point out: “Collaboration can take place with people seemingly without the limitation of geography” (Hernon/Matthews 2013:9) and Nielsen observes: “A fundamental requirement that must be met if we’re to amplify collective intelligence: participants must share a body of knowledge and techniques.” (Nielsen 2012:75) Furthermore, “online tools create an architecture of attention whose purpose is to help participants find tasks where they have the greatest comparative advantage. Ideally, the architecture of attention will bring the attention of the right expert to the right problem” (Hernon/ Matthews 2013:32). Equally important, this is where the library can leverage its centrality as the hub of knowledge on every campus, the point where ideas can intersect and be leveraged into greater use and impact. Pariser recognized the importance of this need when he said: “Creativity is often sparked by the collision of ideas from different disciplines and cultures. Combine an understanding of cooking and physics and you get the nonstick pan and the induction stovetop.” (Pariser 2011:15) Sunstein further confirmed this by saying: “Information is widely dispersed in society. Most human beings on the planet have bits of information from which others might benefit. But the groups and institutions often fail to obtain the information that individuals have. As a result they end up making avoidable and sometimes disastrous mistakes.” (Sunstein 2006: 7) So the KCP could fill a vital and differentiating need by facilitating users being able to hold a discussion via social networking tools, around a knowledge object, be it a video, sound recording, text or data.

c. **Librarianship.** A KCP will also bring the value of librarians as facilitators to the forefront of such knowledge explorations and developments. “The (...) librarian is not only the keeper of all the information, but also has the knowledge of what data connects to other data. Thus the relevant content is found and linked where appropriate. Such connections may be to other works, but could just as easily be connected to a movie, a song, or a theatre performance. It could be a video interview with the author. It could be a video of a discussion of the work.” (Gould 2011:125) LibGuides have shown that we can add a great deal of value to subject explorations, but we should, through the KCP, enhance our ability to bring face-to-face discussions to our users (and as many libraries have already done). Additionally, the KCP should provide a rating mechanism so that users could rate librarians and over time, create a virtual spectrum of librarians in demand, by subject areas, because of their ranked value as determined by end-users.

d. **Predictive, proactive library services.** As discussed previously, analytics will provide librarianship with the opportunity to understand how users who won research grants, scored the highest grades, matriculated, advanced on to graduate schools, published award winning research, obtained employment soon after matriculation, et cetera in comparison to how they interacted with the library, its resources and at what points in time and what resources they utilized. By analyzing this, we will be able to look at new users and suggest to them, proactively, what they might wish to consider doing and at when in order to replicate that success. The KCP would ideally be the tool to front-end that pro-active function for the end-user.

e. **Integrated tools for creating new knowledge.** Using either open source office productivity tools, or through proprietary office appli-
cations featuring rich APIs, word-processing, spreadsheet, database and presentation tool should cleanly integrate within the interface of the KCP so that again, the end-user does not need to step out of the interface in order to actively work on their research or assignment. The flow should be seamless and instantaneous as they go from discovery to reaching behind the research to reanalyze the data using a spreadsheet tool to writing up their findings. Furthermore, as they cite existing works, the tools should be fully integrated that would permit them to record their citations and develop their bibliographies. As the user’s work progresses, the KCP should also integrate with Open Access publication tools, so that the paper/research could be quickly and seamlessly moved into the peer review publication process. Once published, the final step would be to offer the submission tool to move the work into the organization’s institutional repository where it completes the circle by becoming part of the existing base of knowledge to be discovered by the next research/end-user.

f. Serendipity and Contextual Support. As Pariser famously noted in his work “The Filter Bubble”, search tools that are dependent on advertising are incented to develop profiles of end-users that turn into silos or filtered bubbles of knowledge wherein what the end-user is exposed to becomes a reflection of their previous expressions of interest. Since librarians are not incented with this same model, but in fact, quite the opposite, one of the goals of the KCP should be to leverage serendipity. Pariser also offered that: “They can solve for serendipity by designing filtering systems to expose people to topics outside their normal experience.” (Pariser 2011:235) Specifically he suggested: “(…) place a slider bar running from ‘only stuff I like’ to ‘stuff other people like that I’ll probably hate’ at the top of search results and the News Feed, allowing users to set their own balance between tight personalization and a more diverse information flow.” (Pariser 2011:235) With regard to contextual support, the goal would be to facilitate end-users, when they are working with existing knowledge, to easily understand the environment in which research was created, the funding sources and the owners of those funding organizations behind the research being considered. Then, as librarians, we will truly have met the need of our end-users, to set context around a piece of knowledge they might want to utilize.

Through these techniques, we make sure we do not let our users get placed in such a filter bubble, while at the same time offering clear differentiation for librarianship and the tools we develop and offer our end-users. We then use technology to place our users in a “learning bubble” not a “filter bubble” – a place above biases and above unspecified and unmodifiable filtering.

7. Summary

It is important for us to realize that as librarians a lot of our future rests in our ability to provide substantial value and differentiation through unique services. Those services must be what our users want, need and will value because they so clearly integrate into and facilitate their day-to-day workflows.

The Knowledge Creation Platform could be one way we can do this and in a highly scalable fashion. It would also support our efforts to demonstrate our drive to go from being inwardly to outwardly focused. We can do so while being consistent and supporting our core mission (as outlined by Lankes). The KCP would be a next generation of library discovery technology that would be a powerful enabler of new, or greatly enhanced, library services that are proactive, facilitating and purposeful. It would facilitate our ability to help our end-users to create new knowledge while, at the same time, allowing librarianship to leverage existing knowledge and culture. This technology would amplify our profession and make its value more widely understandable, measurable and documentable.

Nielsen said it best when he said: “We’re at a unique moment in history; for the first time we have an open-ended ability to build powerful new tools for thought. We have an opportunity to change the way knowledge is constructed.” (Nielsen 2012:206)
References


Carl Grant is Associate Dean for Knowledge Services and Chief Technology Officer at the University of Oklahoma Libraries, 401 West Brooks Street, Norman, OK 73019, Phone: +1 405 325 2611, E-Mail: carl.grant@ou.edu