

Planning and Implementing Resource Discovery Tools in Academic Libraries

Mary Pagliero Popp
Indiana University, USA

Diane Dallis
Indiana University, USA



Managing Director: Lindsay Johnston
Senior Editorial Director: Heather A. Probst
Book Production Manager: Sean Woznicki
Development Manager: Joel Gamon
Development Editor: Hannah Abelbeck
Assistant Acquisitions Editor: Kayla Wolfe
Typesetter: Milan Vracarich, Jr.
Cover Design: Nick Newcomer

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: <http://www.igi-global.com>

Copyright © 2012 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Planning and implementing resource discovery tools in academic libraries / Mary Pagliero Popp and Diane Dallis, editors.
pages cm

Summary: "This book addresses the many new resource discovery tools and products in existence as well as their potential uses and applications"--Provided by publisher.

Includes bibliographical references and index.

ISBN 978-1-4666-1821-3 (hardcover) -- ISBN 978-1-4666-1822-0 (ebook) (print) -- ISBN 978-1-4666-1823-7 (print & perpetual access) (print) 1. Online library catalogs. 2. Federated searching. 3. Library web sites. 4. Library catalogs and users. 5. Information behavior. 6. Academic libraries--United States--Case studies. 7. Academic libraries--Canada--Case studies. I. Popp, Mary Pagliero, 1949- II. Dallis, Diane, 1971-

Z699.35.C38P58 2012

025.5'24--dc23

2012003667

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Chapter 11

Setting a Direction for Discovery: A Phased Approach

Janet Fransen

University of Minnesota, USA

Lara Friedman-Shedlov

University of Minnesota, USA

Nicole Theis-Mahon

University of Minnesota, USA

Stacie Traill

University of Minnesota, USA

Deborah Boudewyns

University of Minnesota, USA

ABSTRACT

While many other academic libraries are currently or have recently faced the challenge of setting a new direction for their discovery platforms, the University of Minnesota is perhaps unique in its phased approach to the process. In the spring of 2011, the University of Minnesota Libraries appointed a Discoverability task force to identify a Web-scale discovery solution, the third phase in the Discoverability research process. Discoverability 3 Task Force members are now synthesizing the work of two previous phases and other relevant internal and external analyses to develop requirements and selection criteria for the solution. Some of these requirements and criteria are standard for any large-scale system implementation. Others were derived from the findings of the previous two phases of the Discoverability project. The authors discuss the Libraries' phased approach to developing a vision for discovery and selecting a solution that puts the Libraries on a path to fulfilling that vision.

DOI: 10.4018/978-1-4666-1821-3.ch011

INTRODUCTION

As library users become more and more comfortable with each new innovation in the Web-based world, libraries have found themselves struggling to provide an experience that seems as effortless as Google yet directs users to resources owned or licensed by the library. The path to the perfect search tool is further complicated by the range of needs an academic library must meet. If users are to rely on a single search tool, that tool must provide a balance of ease of use and breadth/depth of results that is appropriate for:

- The freshman who needs five credible sources he or she can read, understand, and synthesize.
- The upperclassman who is capable of reading and synthesizing academic work, but still may want to find articles and books from popular sources to provide context for a topic.
- The graduate student who must find and review any and all literature related to his or her topic.
- The researcher who needs to know about new publications in his or her field, but also understand the basics of other disciplines as interdisciplinary work grows.

University of Minnesota Libraries has approached this challenge by charging successive groups with exploring aspects of user expectations and needs, inventorying data sources both within and external to the Libraries, developing a vision for an ideal discovery tool and, finally, issuing a request for proposal (RFP) for the Libraries' next discovery platform. The groups have each operated under the umbrella term "Discoverability," and each has built on the work of the previous Discoverability groups.

In this chapter, members of the three Discoverability groups describe what they found in each phase of this approach and how theoretical

discussions and explorations of the current state of the art in discovery led to a series of requirements for the Libraries' next discovery solution.

BACKGROUND

The University of Minnesota is a large public research university, with a flagship campus located in the Twin Cities of Minneapolis and St. Paul, and four coordinate campuses across Minnesota. About 52,000 of the 69,000 enrolled students are on the Twin Cities campus. The University community includes over 4,000 faculty members among its 25,000 employees. The University offers Bachelor's, Master's, and Doctoral degrees in a wide variety of majors, with strong programs across disciplines.

A core value of the University of Minnesota Libraries is offering appropriate services and tools to meet the Libraries' diverse user population's needs. Over the last decade, Libraries staff members have expressed that value through projects such as:

- The Undergraduate Virtual Library, a library-novice-friendly portal for key resources, implemented in 2004 as part of a University-wide focus on undergraduate learning and retention (Prescott & Veldof, 2010).
- The Multi-dimensional Framework for Academic Support project, which analyzed user behavior and resulted in numerous initiatives to create a more productive research support environment (University of Minnesota Libraries, 2006).
- The myLibrary service, a site within the University's institutional portal with content customized according to the user's *affinity*: their role, college, department, and degree program (Hanson, Nackerud, & Jensen, 2008).

In mid-2008, the Libraries began promoting Ex Libris's™¹ Primo^{®2} product, locally branded as MNCAT Plus, as the default search interface for the catalog. Although this early discovery layer was a step forward for undergraduates seeking a friendlier face than the librarian- and researcher-oriented Ex Libris™ Aleph^{®3} OPAC, more experienced researchers often found it frustrating. Clearly, it would—and will—take some time to design and implement a single discovery tool that serves all users well.

Although library roles and user needs are evolving quickly, the Libraries administration felt it was safe to assume that the Libraries will be connecting people to the information they need for many years to come. With that in mind, the time was right to launch a long-term project to determine how libraries, users, and the technology environment are evolving and what those changes mean for library-mediated discovery.

LITERATURE REVIEW

Each of the first two phases of the Discoverability project involved extensive literature reviews. Interested readers can find bibliographies in the groups' final reports (Hanson, et al., 2009 and Hanson, et al., 2011). Many resources used in those studies are included in the Additional Readings at the end of this chapter.

The Discoverability reports are part of a large volume of literature written between 2008 and 2011 that focuses or touches upon subjects such as the evolving meaning of the library collection and role of the catalog, information-seeking behavior, and user preferences. The discovery product landscape also changed dramatically, spurring librarians to consider just how to evaluate products in light of their own institution's needs.

The shift to online content and collections in libraries presents a new opportunity and challenge for discoverability in libraries. Today's library collections consist of print and non-print items,

as well as data and a variety of other information formats. Pradhan, Trivedi, and Arora (2011) discuss how Web-scale discovery allows users to search, refine, and discover all content comprising the library's collections seamlessly and cohesively. Although not entirely perfect, discovery tools allow users to perform a broadcast search across information sources and refine their search as needed. However, as Chew (2010) mentioned, these new tools are using MARC in new ways to produce a contextualized and faceted discovery environment for the user.

At the same time, libraries must contend with the fact that users' information needs and seeking behaviors cannot be generalized to one set of user requirements. Connaway, Dickey, and Radford (2011) use several research methods to show the importance of convenience as a factor in information seeking, but concede that its importance relative to other factors and even its definition depend on the user's situation. Fisher and Julien (2009) review the extensive research over the last few years on the role of context, as well as the "human factor:" academic role and discipline, occupational group, or people in everyday settings.

As discovery tools evolve, generalized assumptions ("They want it to work like Google") are becoming more nuanced. In a recent Library Journal Reviews post, librarians from four institutions discussed their experiences with the newly-adopted discovery tools. Amanda Clay Powers of Mississippi State University praises EBSCO Discovery Service™⁴ for solving common problems for both lower-level undergraduates and advanced researchers. At the same time, Powers notes that students such as upper-level undergraduates still have more success with subject-specific databases (Discovering What Works, 2011).

A small but growing body of literature evaluates the next generation discovery tools that have been adopted by other libraries facing the same set of challenges. Most of these articles focus on the selection process (Cai, Dou and Jiang, 2011; Brubaker, Leach-Murray, and Parker, 2011; Becher

and Schmidt, 2011). Missing from the literature are thorough studies or evaluations of the tools once deployed; while many go into significant detail on the features and concerns regarding the discovery tools as well as the criteria used to make a decision on which to adopt, there is as yet little research on how the latest Web-scale tools perform “in the wild.” Yang and Wegner’s (2010) article notably relies strictly on real-life examples and demonstrations rather than vendor claims, but the study was limited to verifying the presence of specific features, rather than a thorough evaluation of their performance and/or usability. A few such studies (e.g. Emmanuel, 2011) are now available for the “next generation” catalogs that began appearing in the mid- to late-2000s, but the next wave of innovation represented by Web-scale discovery tools is still too new to have been the subject of this type of investigation. As Marshall Breeding notes in “State of the Art in Library Discovery” (2010), adoption cycles are sluggish, despite the wealth of new tools that are now available, so relatively few institutions have much experience with any of them.

PHASE 1: SURVEYING THE LANDSCAPE AND GATHERING DATA

In 2008, the Libraries convened the Discoverability Exploratory group to explore discovery behavior and corresponding tools or systems that might make relevant resources more visible and easier to find. The group’s investigation involved identifying trends about user search and discovery behavior, gathering and analyzing systems data, and determining if the trends identified in the current literature aligned with that of the local data collected on user behavior at the University of Minnesota.

Trends

To identify user behavior drawn from user studies, the group compiled and analyzed available literature (as of December 2008) describing user discovery activity. The literature review included case studies such as the task force report on resource discovery from the University of Wisconsin and usability testing by the University of Washington, the EDUCAUSE study of undergraduate use of technology, and the 2008 Horizon Report. The group also consulted articles written by leading practitioners such as Peter Brantley, Lorcan Dempsey, Derek Law, and Clifford Lynch. For a complete list of articles used for the trend review, consult the suggested reading section at the end of this chapter.

A number of common themes emerged from the group’s study and discussion of these articles and studies. The group consolidated these themes into five major trends:

1. Users discover relevant resources outside traditional library systems
2. Users expect discovery and delivery to coincide
3. Users increasingly rely on portable Internet-capable devices
4. Discovery increasingly happens through recommending
5. Users increasingly rely on nontraditional information objects (Hanson et al., 2009)

Although none of the individual trends identified were surprising, even in 2008, as a group the trends made a compelling case for rethinking the Libraries’ discovery systems and priorities. The trends raised questions about how much, given the degree to which users rely on Google and other outside portals, the Libraries should invest in developing a search interface without a very specific or specialized audience in mind.

The trends also highlighted the importance of making local resources discoverable in those

Table 1. The group investigated any likely data source in an effort to prove or disprove the trends

Question	Data Sources investigated	Findings
How much are the Libraries' current discovery tools used? Has that number increased or decreased over time?	Statistics from the OPAC (Aleph®) and current discovery tool (Primo®)	Statistics show that users searched the catalog through either interface more often than any other library resource. However, comparisons over time proved impossible due to changes in how searches were counted.
	OpenURL link resolver (SFX®) source and target logs	The OpenURL link resolver moves users from discovery to delivery, and the "Find It" page that presents the options for fulfillment is the most-viewed page on the Libraries website. However, logs show that users can and do discover articles through tools not controlled by the Libraries. Google Scholar™ is the most striking example: Requests to the link resolver from Google Scholar™ have climbed steadily since its release in late 2004.
	Individual database statistics (five of the top ten most-used databases were sampled)	Not useful for answering this question due to the number of ways a user might come in contact with a database, including through proxy links, federated search, OpenURL resolver, or directly from a bookmark.
	Server logs for Archives and Special Collections finding aids	Although finding aid statistics didn't provide evidence of more or less use of library discovery tools, the statistics did underscore the fact that users approach finding aids from a very different perspective than other library web pages. Users of the main site viewed an average of four pages per visit, while finding aid users viewed an average of almost twelve pages. Twenty-five percent of finding aid users spent more than fifteen minutes per visit viewing content.
	Server logs for repositories hosted by the Libraries	Statistics for all three repositories hosted by the Libraries showed extensive traffic directed to individual items through search engines.
The Libraries currently expose the catalog to users throughout the state through a discovery layer called the MnLINK Gateway. Do users discover the Libraries' resources through that tool?	Statistics from the OPAC on how many searches were done via the Z39.50 protocol	Users of the MnLINK Gateway search University Libraries content by default, whether they intend to or not. In addition, the MnLINK Gateway "polls" multiple times for one search. Therefore, even though the top-ranked method of searching the catalog during the time period investigated was the Z39.50 protocol, the group could not conclude that users intended to search the Libraries catalog more often through that interface than any other.
What percentage of visits to the Libraries' Web site comes from mobile devices? Is that percentage increasing over time?	Server logs	Unfortunately, the Web statistics program in use during the period studied did not differentiate operating systems or devices used to access sites, so no data could be analyzed to find evidence of the trend.
Are external Web sites (such as Google or Wikipedia) driving people to the Libraries' Web site, resources on the Libraries' servers?	Server logs for pages within each of the Libraries' subdomains.	Users rarely viewed pages beyond the domain and subdomain home pages. However, a few secondary pages accounted for a disproportionate number of visits. Users were directed to those pages largely through Google searches, and occasionally from Wikipedia pages or other academic institutions' sites.

outside systems. This suggests that perhaps the library would do best to focus its resources less on search (at least on general search) and more on delivery and fulfillment.

Data to Support Trends

Identifying the most prominent trends in the literature is helpful in framing a discussion, but

eventually that discussion will always come back to the individual institution: Can one really assume that trends identified by one author or institution are valid for users at the University of Minnesota? In the spirit of evidence-based decision making, the first Discoverability group sought evidence of the trends in the Libraries' own operations and systems by formulating specific questions. Table 1 outlines a few examples.

Setting a Direction for Discovery

As Table 1 illustrates, some of the questions could be addressed by data from multiple systems, but others were unlikely to be answered by any of local data. Although the group found strong anecdotal evidence of the five trends, it was difficult to support the trends' validity armed only with the local statistics and data.

Statistics offered some description of user behavior, but provided little insight into user intent. The complications encountered with the data, or lack thereof, indicated that the Libraries would benefit from closer attention to tools and processes related to efficient and ongoing data collection. Therefore, when the group synthesized the identified trends into a set of guiding principles for future investigations, the following principle was written to address the data problem. The principle reads:

In order to remain responsive, relevant, and useful to our users, we must aggressively measure and analyze user behavior through local system statistics. These efforts will complement our ongoing assessments utilizing focus groups, usability studies, and reviews of applicable literature.

Most of the other guiding principles identified by this first Discoverability group relate directly to the end user:

- Users draw little distinction between discovery and delivery; systems, data, and information objects should be optimized for fulfillment.
- In order to remain agile and responsive in a rapidly changing information environment, our systems and data structures should provide us with the greatest possible flexibility for frequent iteration and reuse by ourselves and others.
- In order to best facilitate our users' discovery of relevant information, we should strive to be end-user device/platform agnostic.

- Discovery should be organized around users rather than collections or systems. This organization should be based on realistic, evidence-based models of our users and their research tasks.
- Users are successfully discovering relevant resources through non-library systems. We need to ensure that items in our collections and licensed resources are discoverable in non-library environments.
- Users rely on system- and peer-generated recommendations to discover relevant resources. We should capture the data necessary to provide targeted suggestions to users and defer to network-level systems where critical mass already exists.

PHASE 2: DEVELOPING A VISION AND BEST PRACTICES

Grounded with the set of guiding principles, the Libraries were poised to launch a new phase of work. Two objectives in the Discoverability Phase 2 charge came directly from the guiding principles:

- Because users are successfully discovering relevant resources through non-library systems, the group was asked to inventory and analyze the internal and external data sources to which the Libraries provide access.
- Because discovery should be organized around users rather than collections or systems, the group was asked to identify and understand the user groups the discovery interface would be serving.

The sponsors assumed the next phase of the Discoverability process would result in a recommendation for the Libraries' next discovery environment. Therefore, the charge included that the group:

- Develop a vision for a future discovery environment based on the findings of Phase 1.
- Investigate and summarize how selected peer institutions have redefined and redesigned their discovery environments.

Exposing Local Data through the Discovery System

Web statistics collected during Phase 1 alerted the Libraries to the number of people finding locally produced data sources through search engines and external aggregations. However, in many cases a user would need to know that a data source existed in order to find it from within the Libraries' Web sites. In an ideal discovery environment, these local data sources should be ingested and searchable.

In addition to the library catalog, local data sources include discipline-specific repositories, digitized collections, and applications that extract and adapt catalog data for a particular purpose or audience. In many cases, these local data sources have been created over the course of years in different systems, and intended for a variety of uses. Some of the sources adhere to established metadata standards, while others rely on various homegrown metadata schemas. Some sources include digital objects and metadata describing those objects, while others include only metadata surrogates for physical and digital items. Some data sources are updated and used regularly, while others have outlived the need that initially spurred their creation.

During Phase 2, the group worked to document these many local data sources and their defining characteristics. Once defined, the group considered which characteristics made a particular data source a candidate for inclusion in the future discovery environment.

One important characteristic is the uniqueness of the data in each source. When a source does not contain objects or metadata unique to the local institution, the institution must evaluate

the cost/benefit of incorporating that source into a discovery environment. A local database that largely duplicates catalog data in a non-standard metadata schema is probably not worth the time and effort required to translate it into a form ingestible by a discovery tool. On the other hand, a local database that includes both unique digitized objects and metadata for those objects may be well worth the effort to prepare it for inclusion, even if the relative cost of doing so is greater.

Uniqueness is not the sole consideration: Larger databases with non-unique content that adhere to nationally or internationally accepted metadata standards and that exist in systems with extensive capabilities for data export (such as the local catalog) typically require less effort—or at least less new effort—to expose their content. Using a “grid of exclusivity,” an analyst can plot the uniqueness of a resource against the degree to which sources may be shared with external sources. The grid, shown in Figure 1, reveals the data sources with the highest exclusivity and the lowest degree of current sharing. For these sources, the effort to expose the data will have a greater benefit.

A discussion of how to expose worthwhile locally-created data leads inevitably to the question of where that data should be exposed. To begin to answer that question, the group created a list of external data aggregations with existing known value to users and their salient characteristics. Table 2 lists some types of external aggregators identified by the group.

Each type of aggregation requires a slightly different analysis to determine whether an institution should include its metadata and/or digital objects. Several factors to consider for each aggregator include:

- **Ease of access.** Is the aggregator's content freely accessible, or is authentication required?
- **Coverage.** Does the aggregator contain most of the content that users might re-

Setting a Direction for Discovery

Figure 1. The grid of exclusivity was useful in determining which data sources were most worth ingesting for the discovery layer

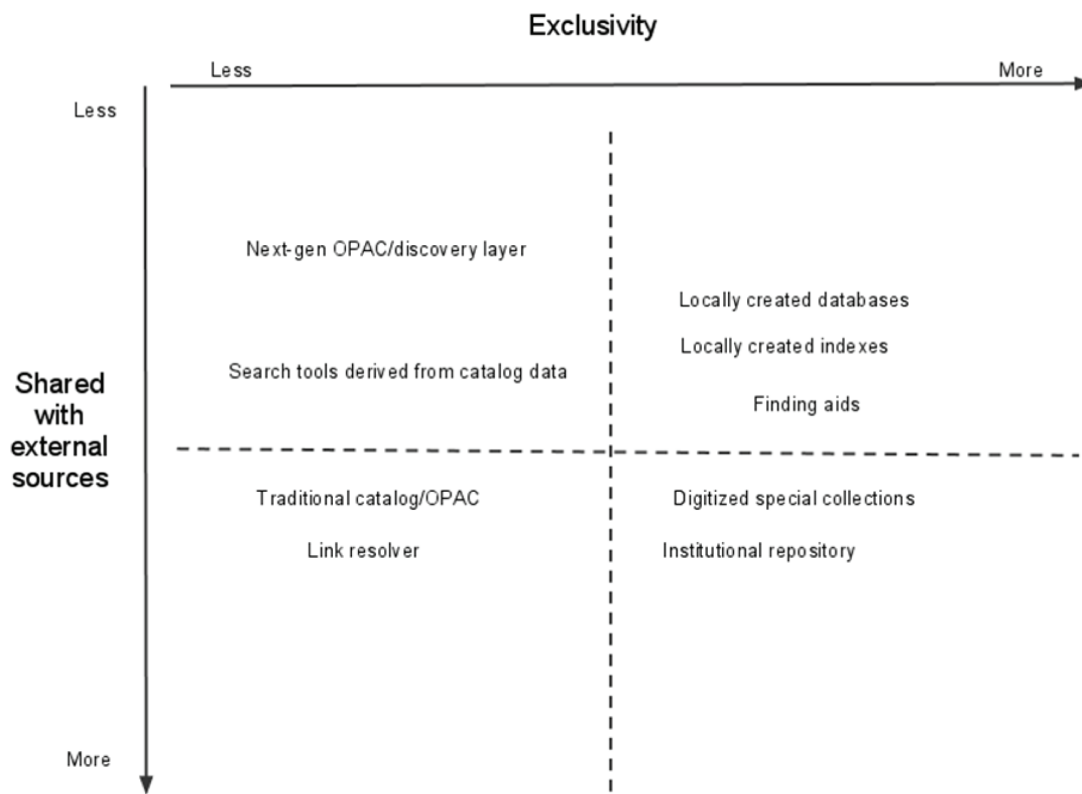


Table 2. Types of external aggregators investigated by the Phase 2 group

Aggregation Type	Examples
General metadata aggregations	WorldCat [®] , Google Scholar [™] , OAIster ^{®7} , Primo Central [™] , Google
General data object aggregations	HathiTrust, Wikipedia, Internet Archive, Flickr ^{®8} Commons
Disciplinary aggregations	AgEcon Search, ArXiv, Earth Prints
Form aggregations	Digital Dissertations, MERLOT (learning objects), ArtSTOR, ArchiveGrid ^{SM9}
Topical aggregations	Minnesota Reflections (state-wide repository for cultural heritage organizations), Ethic-Share (collaboration tool for the study of ethics)

- sonably expect, or are there substantial gaps in coverage?
- **Ease of contribution.** Can metadata and/or resources be contributed automatically via a protocol like OAI/PMH, or is more active effort required of contributors?
- **Standardization requirements for metadata and file formats.** Does the aggregator accept a variety of formats and standards, or must contributions adhere to a particular set of standards?

- **Cost.** Are there costs to contribute metadata or resources, and/or to search and retrieve them?
- **Rights management.** Are there use limitations imposed by the aggregator on resources it contains? Are legal rights to content managed responsibly?
- **Commitment to preservation.** Is the aggregator committed to preserving resources contributed to it?
- **Maintenance.** Does the aggregator devote effort to maintaining currency and accuracy for its data? Are contributors able to maintain and correct their contributions?
- **Local impact and interest.** Does the aggregator serve the needs of any particular local interest groups? How will increased exposure of local data impact demand on institutional resources?

Among these, standardization is perhaps the most important practical concern in analyzing the readiness of local data sources for broader exposure. This factor determines how easy it will be for any future discovery layer to ingest and expose the data. For example, metadata created according to a locally developed scheme, or metadata created according an inconsistently applied scheme, may pose difficulties for ingestion by external aggregations. While organizational capacity to create robust metadata may be limited, libraries should strive to create readily interoperable metadata whenever possible.

By making an inventory of internal data sources and external aggregations, the Libraries gained a more complete and nuanced picture of expectations for interactions among the discovery layer and other systems.

User Communities

The discovery layer connects users to the information they need. With a clear sense of both traditional library collections and more ad hoc

data sources in mind, members of the Phase 2 group turned their focus to the user communities served by that data environment.

The group determined that people at the University of Minnesota rarely have information needs or information seeking behaviors that differ from their counterparts at other large research universities. In late 2009, the Libraries surveyed users of the catalog (MNCAT Classic) and the discovery layer (MNCAT Plus) and found striking differences in the degree of satisfaction and primary reasons for using each tool based on the user's college affiliation or status (Chew, 2010). Both the survey results and anecdotal evidence led the group to conclude that Libraries users' needs and behaviors are distinct from each other based on factors like discipline (arts, engineering, health sciences, etc.) and status (undergraduate, graduate student, faculty, etc.). To ensure that a cross-section of users is considered, as decisions regarding discovery are made, the group developed a set of user personas. The personas are very specific, but fictional, users.

Personas are most effective when the set is small; many software applications are written with a single user persona in mind. The group sought to maximize the number of users represented while minimizing the number of personas future decision-makers must synthesize and accommodate. After casting a broad net based on affiliation, role, discipline, and major, the group trimmed the set of personas to those for whom significant literature or first-hand knowledge existed, and who were distinct from other personas in a meaningful way. Table 3 shows the final set of sixteen personas.

The sixteen personas are based on a combination of recent literature on information seeking behavior for specific groups and input from subject specialists at the University of Minnesota. The complete personas, available on a University of Minnesota Web site, include photos and paragraphs written from the individual's perspective, as shown in Figure 2.

Setting a Direction for Discovery

Table 3. Sixteen personas in four categories act as a cross-section of libraries users

Affiliates	Non-Affiliates	Health Sciences	Other Disciplines
Undergraduates Graduate Students Faculty	High School Students Engineering Professionals Commercial Users Archives Users	Clinicians Professional Students Researchers/Fellows	Arts and Humanities Visual Arts Engineering Law Physical/Life Sciences Social Sciences

Figure 2. Each persona includes a photo and paragraph written from the individual's perspective (University of Minnesota, 2010)

Affiliate Users

Undergraduates
Graduate Students
Faculty

Non-affiliate Users

High School Students
Engineering Professionals
Commercial Users
Archives Users

Health Science Disciplines


Clinicians
Professional Students - Health Sciences
Researchers/Fellows - Health Sciences

Other Disciplines

Arts and Humanities
Visual Arts
Engineering
Law
Physical/Life Sciences
Social Sciences

[Edit sidebar](#)

Graduate Students



Now that I'm a graduate student, I spend a lot of time doing the same things that faculty do - practice, I suppose. Certainly my research is much more in depth than it used to be, or that my students do in my classes. I'm so busy that I have to fit my research into any gaps I can find - it's hard to sit down and really get organized. I've heard other people mention databases that they use from the library, but I don't know how they ever found them. There are so many choices from the library and, like I said, I'm working around lots of other responsibilities so I end up getting frustrated. It would be great if I could search lots of databases at once. That way I could really make the most of the citations I'm working from. It would also help me when my research spreads into other disciplinary areas or focuses on data I'm not familiar with. What I really need to do get my research published and published in the right place and I'd love it if there were some easy tool from the library to help me with that.

From the literature

- More interested in numeric data than other groups (Desktracker stats)
- Often work from/need help with citations (Desktracker stats)
- Identify themselves more closely with academic researchers than undergraduates (Hampton-Reeves et al., 2009).
- Often lack basic research skills and work by trial and error. Largely unaware of many library services offered (University of Minnesota Libraries, 2006).

From our colleagues

[Kate Peterson, Information Literacy librarian](#)

- Graduate students are often overwhelmed by the availability of information/discovery and don't know when to stop (or how to be systematic) and thus want tools that can, for example, search multiple databases together or the 'one library search' idea, at least as a way to improve efficiency. This is mostly a social sciences and interdisciplinary challenge.
- This point from the Life Sciences profile is also relevant to many other graduate students: "Work is increasingly collaborative. Obstacles include different vocabularies, not knowing core journals, and not knowing where to publish (Marcus, Ball, Delserson, Hribar, & Loftus, 2007)

References

Hampton-Reeves, S., Mashiter, C., Westaway, J., Lumsden, P., Day, H., & Hewertson, H. (2009). Students' use of research content in teaching and learning. *Report for the Joint Information Systems Council.*

University of Minnesota Libraries. (2006). A multi-dimensional framework for academic support. Minneapolis: University of Minnesota. Retrieved from <http://purl.umn.edu/5540>

The Libraries' Vision for Discovery

The inventory of internal data sources, analysis of types of external aggregators, and development of user personas were all tools intended to be used by the next Discoverability group, which would be charged with recommending a specific product as the Libraries' next default discovery layer. In addition to these three practical tools, the Phase 2 group developed a high-level vision for discovery built upon the work and research from Phase 1, and the problems with having a decentralized discovery environment. The vision further synthesized the research assembled on user needs (including work from Phase 1) and local collection imperatives and opportunities. In developing and articulating the vision, the group was directed to ignore the limitations of products currently available.

The Libraries' Vision for discovery identifies seven key concepts. The seven concepts are:

1. Scopes, customized according to the user's community, past usage, and preferences.
2. Integrated metadata layer, which makes a broad range of resources available to the discovery layer.
3. Attention to processes and best practices needed to expose internal data to external systems.
4. Expand on discovery and delivery options through external data aggregations.
5. Object manipulation and personal curation, to facilitate delivery and to build on the concept of literature searching as an ongoing process in research.
6. Make use of systems that facilitate data-driven decision making.
7. Ongoing review and evolution of the discovery environment.

Below is a brief description of each of the concepts and the imposing factors, which informed the vision.

Local data sources are available in vast numbers of disparate sources, and present users with a complex, jumbled environment. If users want to be comprehensive in their searching, they need to pull information from different data sources, but may not be aware of all the existing sources.

Scopes provide a means for the discovery layer to present relevant resources and information based on a user's community affiliation. The scopes could be predetermined based on a set of characteristics associated with a user affiliation and a set of resources selected for that affiliation by a librarian, or they could be dynamic in nature and created or enhanced by the users themselves. Libraries are already using scopes as a means to organize sets of resources for specific users or user communities. For example, a library may present a set of resources for undergraduates or may list resource sets by specific discipline. The idea of scopes presented in the Phase 2 report expands this concept since it is dynamic and allows a mechanism for the user to create or enhance the set.

An integrated metadata layer would provide the backbone for the scopes. The integrated layer would allow disparate data sources to be presented in a single interface which could then be scoped. Users would see a consistent search interface and work in an environment that offers preselected data sources and provides the opportunity for users to select their own data sources.

The exploration of user communities and their needs suggests there is a strong desire for personalization in the search environment. The concept of personalization and object interaction is rooted in Web 2.0 and not new to library users. Users have come to expect that search algorithms will adapt to their behavior, and that users have the ability to manipulate and repurpose discovered items for themselves. Tools for personalization are interactive and offer opportunities for Libraries' staff to create new access points and enhance the discovery process for users.

As in Phase 1, the Phase 2 group emphasized the Libraries' need to make data-driven deci-

Setting a Direction for Discovery

sions. Systems and their functionality need to be reviewed and revised to meet the changing information needs and information seeking methods of library users. The process of reviewing functionality and seeking out improved tools should be fluid and continuous.

These key concepts provide the foundation for future exploration into discovery, and a framework for identifying and possibly choosing a discovery tool for the Libraries.

PHASE 3: DEFINING REQUIREMENTS

During the initial phases of the Discoverability process, the groups had the luxury of combining the observable with the intangible or aspirational. The Discoverability Phase 1 group gathered institutional data, but also explored trends in user behavior. Data only rarely confirmed the trends, but the group did not consider the trends any less valid because they weren't measurable by existing methods. The Discoverability Phase 2 group inventoried existing internal data sources and characterized external data sources, but also proposed a vision for a discovery layer that would meet the needs of diverse users. This vision, with its concepts of scope and an integrated metadata layer, describes an ideal beyond the capabilities of the tools currently on the market.

The group that convened for Phase 3 of the Discoverability process, in process as of this writing, has no such luxury. While the previous two groups were exploratory, the Discoverability Phase 3 group is a task force: The group is charged with conducting a request for proposal (RFP) process that will lead to the recommendation of a unified discovery environment for the University of Minnesota. The task force must reconcile the vision for an ideal discovery system with the reality of the current marketplace.

Outside Factors to Consider

If a discovery tool existed in a vacuum, the RFP process would be straight-forward. The task force could simply transform the previous group's vision into a list of requirements, issue the RFP, and wait for the responses to roll in. Unfortunately, the tool and the process are subject to many outside factors:

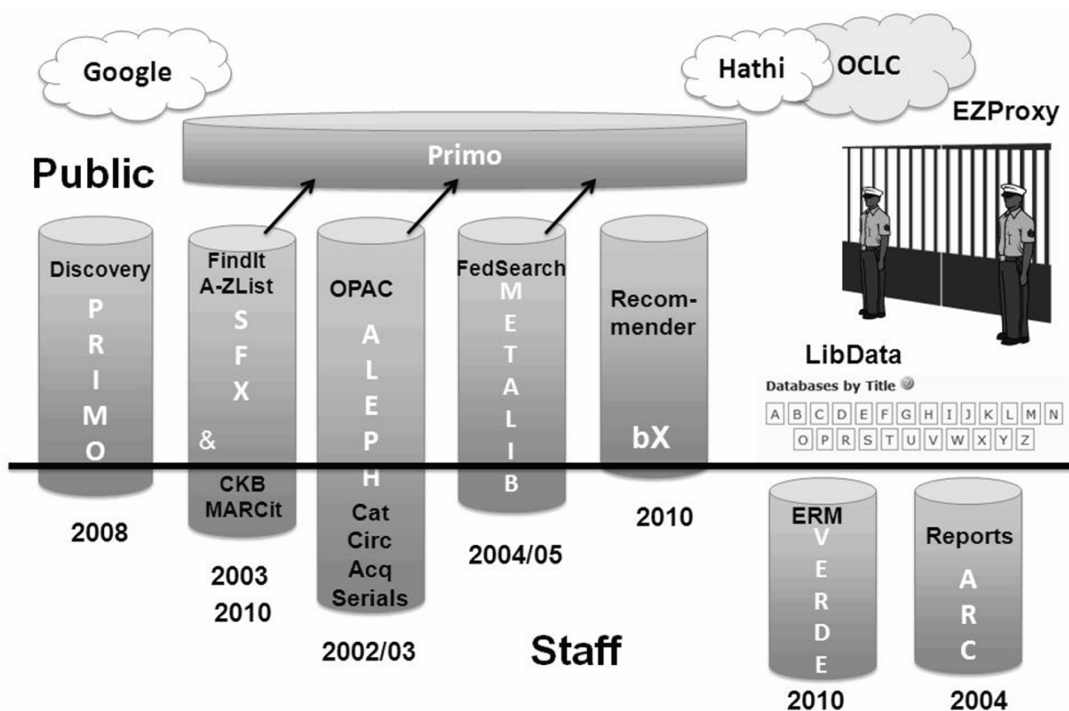
- The information technology environment at the institution
- The library information ecosystem, both present and anticipated future
- The current state of the art

The Information Technology Environment

After spending more than two years exploring the information seeking/discovery landscape, selecting a new discovery tool is the next logical step. At the same time, the current economic climate is quite different from that of mid-2008. Although the Libraries strive to continuously improve the user experience, it seems as though the timing could be better. Like academic institutions everywhere, the University of Minnesota and its Libraries are under increasing pressure to cut costs while maintaining service levels. However, the information technology of the Libraries relies upon the University's Office of Information Technology (OIT). OIT owns the server on which the current discovery system resides, and that server has reached the end of its life. OIT prefers not to host the next generation discovery system at all; the new discovery layer must run under a Software as a Service (SaaS) delivery model.

The SaaS requirement means that the known options are automatically narrowed. Although it is possible that a new or unknown candidate will appear once the RFP has been issued, at this writing the group assumes that the possible choices for a Web-scale hosted discovery tool are:

Figure 3. University libraries ecosystem as of Summer 2011, with year of implementation



- EBSCO Discovery Service™
- Hosted Primo® with Primo Central™¹⁰
- Summon™¹¹
- WorldCat®¹² Local

The University’s purchasing department has experience with SaaS RFPs, which simplifies the process of developing the RFP. A number of requirements are boilerplate so they can be directly copied to the RFP for a discovery system.

The Library Ecosystem

A discovery tool is the lens through which the user sees the library’s physical and virtual collections. It is also the user’s gateway to actually getting the items they need. Library staff members see the whole picture: the discovery tool is just one part of and must fit within the elaborate library ecosystem. In the course of responding to an increasingly digital information universe, systems

and pieces of software were grafted, overlaid, or linked to the core ILS. Figure 3 sketches the various systems and tools that make up the University of Minnesota Libraries discovery environment as of summer 2011.

The current Libraries ecosystem includes numerous public-facing services: the OPAC, federated search services, recommender services, FindIt (open linking from citations to full-text), and the A to Z list of available databases. Each service is provided by a separate tool, each was acquired at a different time and integrated with the Aleph® ILS, and each has its own back-end functions for staff. Further tools are needed by staff for electronic resource management (ERM) and generating reports. Additional discovery and delivery services are provided by Primo® and through LibData course pages. The increasing number of cloud-based resources, such as Google and HathiTrust, adds to the complexity, as does

Setting a Direction for Discovery

the need for authentication, which is handled by EZProxy^{®13}.

As of summer 2013, the Libraries will be adopting Ex Libris's[™] new unified resource management tool, Alma. Alma will replace the tools filling all of the functions on the staff side of the diagram, but it does not perform all of the functions those tools fill on the public side. Regardless of what the Libraries choose as a discovery layer, the functions of these systems still need to be fulfilled and the discovery system will need to integrate with both the current suite of tools and with Alma.

The State of the Art

At the most fundamental level, the biggest challenge to selecting a discovery solution is the current state of the art. Web-scale discovery remains a relatively immature market, and the available systems that work within information technology environment and library ecosystem fall short of the Discoverability Phase 2 group's vision. Current systems are capable of meeting certain needs, such as integrating data from disparate sources and filtering search results on multiple criteria. But based on information gathered from preliminary testing, marketing material, and resources such as the ALA Library Technology Report "Web Scale Discovery Services" (Vaughan, 2011a), other aspects of the vision do not appear to be achievable at this time.

Scopes, as described in the Vision, are still some time away. Most systems allow the user to save a search, but are not set up to create a profile, either by the Libraries based on pre-established user groups or by the user based on past usage or stated preferences. There is little evidence that systems can handle metadata at the level of granularity that would allow the group's vision of scopes to be realized. And while the ways in which users can interact with resources they discover continue to expand, many of the functions envisioned, such

as the ability to annotate or to create and share ad-hoc collections remain theoretical.

The vision is farther-reaching than the current market supports. The RFP development process, then, requires that the task force choose which aspects of the vision are most important and should be included in the RFP, whether any product is known to meet the requirement or not. The task force must also consider what questions to ask the vendors about how their products behave.

Defining and Prioritizing RFP Requirements

The group drafted initial RFP requirements based on the Libraries' vision for discovery, the known constraints of the University and Libraries environment, additional feedback from Libraries staff, and market research. The considerations outlined above made the already laborious process of compiling and refining these requirements for publication in the RFP even more lengthy and complex. In particular, the need for the system to work with the Libraries' current ILS as well the new back-end system that will be up and running in 2013 (which will be very different in terms of its functions and architecture) considerably complicated the list of requirements. The group was also concerned that the requirements anticipate the issues raised by many of the earliest adopters of Web-scale discovery services.

Key areas of the RFP include the following:

- The basic concept of "Web-scale" discovery suggested many questions about how desired outcomes would be achieved. The system will have its own massive index of article and repository content that will be integrated with local resources, providing a single search across the catalog, print and e-journals and articles, special collections and the collections of peer institutions. Related RFP requirements include questions about the types of metadata

that would be ingested and indexed, how the coverage of journals and databases indexed in the discovery tool would compare to Libraries subscriptions, and how variations in the types and formats of metadata in various sources would be accommodated in the tool.

- The group is aware that the market increasingly includes tools that have been created by companies that have traditionally been in the aggregation and publishing business. Concerns have been raised about potential bias in the content and presentation of the search results. The RFP thus includes requirements to describe how indexing and relevance rankings ensure high quality and unbiased results regardless of the source of the metadata and any relationships with publishers.
- While the move to a hosted tool frees the University from many system maintenance tasks, it also potentially limits the Libraries' ability to customize the tool. A number of RFP questions concern the availability of robust APIs that will allow customers to access and extend the system's capabilities. For example, the Libraries may wish to provide specialized searches that access user account information, or supply permanent URLs for resources and sets of resources.
- The group remains committed to finding a tool that is appropriate not only for novice users, but for experienced users of library tools. Related RFP requirements include questions about how users can identify peer-reviewed material and what types of shortcuts are provided for power users. Also relevant to this point are a number of questions about how a searcher knows which databases are included in a particular search and whether and how the user is referred to specific, relevant databases for additional searching. In addition, the RFP requests descriptions of any usability testing that the vendor has conducted.
- A concern that was raised repeatedly by Libraries staff and is echoed in the reports of early adopters of Web-scale discovery systems focuses on the ability to conduct reliable known-item searches. Early adopters note that the discovery service has not replaced the catalog, which continues to be an important tool for finding specific titles. Once the Libraries transition to Alma in 2013, users will no longer have the luxury of falling back on an OPAC should the discovery layer fail. Therefore, any discovery layer the Libraries adopt must address this challenge.
- As a large research institution with 14 separate library buildings and many specialized collections, University of Minnesota Libraries staff strongly desire the ability to search within specific collections, in addition to the ability to apply facets after a search. Related requirements ask about the ability to limit or pre-construct searches in various ways.
- The Libraries collections encompass material in over 300 languages, including over 155,000 in Chinese, Japanese, and Korean. The new discovery system must handle searching and indexing in multiple languages. Requirements outlined in the RFP ask about UTF-8 compliance, Unicode, and CJK (Chinese/Japanese/Korean) and other vernaculars, as well as normalization of searching to accommodate both use and non-use of accents and diacritics.
- The success of initiatives to improve access and discovery in the Libraries depends on the ability to analyze data from Libraries systems. The first Discoverability group uncovered the need for data and statistics that are not currently available. The RFP therefore includes detailed questions about

Setting a Direction for Discovery

the types and quality of statistics and reporting available from the discovery layer.

CONCLUSION

As the group readies the RFP for release and begins the process of finally selecting a discovery system from among the available options, its biggest challenge will be managing expectations about what is actually possible. For example, based on the architecture of existing Web-scale discovery systems, alphabetical browsing--a feature identified as highly desirable for locating known items--may not be available in any system. However, if a system can in other ways effectively overcome the obstacles to known item searches, it is perhaps a feature we could sacrifice.

More problematic is the Discoverability 2 group's vision of "scopes." Scopes attempt to address the needs of specific audiences, including faculty and graduate students who need to search comprehensively within a particular discipline. Yet early adopters report that while the new Web-scale systems they have adopted are effective for discovery across disciplines, they are less effective once research is focused in-depth on a specific discipline (Vaughan, 2011b). Scopes, as envisioned in Discoverability 2, may not be a feature in any currently available system. The group may instead look for configuration options and APIs that would facilitate developing something like scopes for targeted audiences in the near future.

Given the probable trade-offs and gaps between what was envisioned and what is actually available, one might well question the value of a process as lengthy and detailed as this. The phased approach has enabled Libraries staff to anticipate and more fully consider future challenges and changes for library search and discovery as well as information technology in general. These include the increasing de-coupling of "back-end" staff-facing systems from the discovery layer and the move to "the cloud." The life cycle of systems

is shorter than it once was, and all libraries will need to be more nimble; data need to be more portable and transformable. While the Libraries may or may not identify a system that embodies the full vision, the multi-phase process undertaken over the past three years will continue to serve the institution well. Having articulated the ideal, the University of Minnesota is in a better position to shape the future development of systems, and is better prepared to address their short-comings in the meantime.

ACKNOWLEDGMENT

The authors gratefully acknowledge the work of all three Discoverability groups, and particularly Cody Hanson, Heather Hessel, and Chiat Naun Chew, who each filled chair or co-chair roles during the Discoverability process.

REFERENCES

- Becher, M., & Schmidt, K. (2011). Taking discovery systems for a test drive. *Journal of Web Librarianship*, 5(3), 199–219. doi:10.1080/19322909.2011.583134
- Breeding, M. (2010). State of the art in library discovery 2010. *Computers in Libraries*, January-February. Retrieved from <http://www.librarytechnology.org/lgt-displaytext.pl?RC=14574>
- Brubaker, N., Leach-Murray, S., & Parker, S. (2011). Shapes in the cloud: Finding the right discovery layer. *Online*, 35(2), 20–26.
- Chew, C. (2010). Next generation OPACs: A cataloging viewpoint. *Cataloging & Classification Quarterly*, 48(4), 330–342. doi:10.1080/01639370903437709

- Chew, C., Fransen, J., Gangl, S., Hendrickson, L., Hessel, H., Mastel, K., Nelsen, A. . . . Peterson, J. (2010). *MNCAT Plus and MNCAT Classic survey: Results and analysis*. Retrieved from <http://purl.umn.edu/92473>
- Connaway, L. S., Dickey, T. J., & Radford, M. L. (2011). "If it is too inconvenient I'm not going after it." Convenience as a critical factor in information-seeking behaviors. *Library & Information Science Research*, 33(3), 179–190. doi:10.1016/j.lisr.2010.12.002
- Emanuel, J. (2011). Usability of the VuFind next-generation online catalog. *Information Technology and Libraries*, 30(1), 44–52.
- Fisher, K. E., & Julien, H. (2009). Information behavior. *Annual Review of Information Science & Technology*, 43(1), 1–73. doi:10.1002/aris.2009.1440430114
- Hanson, C., Hessel, H., Barneson, J., Boudewyns, D., Fransen, J., Friedman-Shedlov, L. . . . Traill, S. (2009). *Discoverability phase 1 final report*. University of Minnesota Libraries. Retrieved from <http://purl.umn.edu/48258>
- Hanson, C., Hessel, H., Barneson, J., Boudewyns, D., Fransen, J., Friedman-Shedlov, L. . . . West, A. (2011). *Discoverability phase 2 final report*. University of Minnesota Libraries. Retrieved from <http://purl.umn.edu/99734>
- Hanson, C., Nackerud, S., & Jensen, K. (2008). Affinity strings: Enterprise data for resource recommendations. *Code4Lib*, (5). Retrieved from <http://journal.code4lib.org/articles/501>
- Library Journal Reviews. (2011, December 7). *Discovering what works: Librarians compare discovery interface experiences*. Retrieved from <http://reviews.libraryjournal.com/2011/12/reference/discovering-what-works-librarians-compare-discovery-interface-experiences/>
- Pradhan, D. R., Trivedi, K., & Arora, J. (2011). *Searching online resources in new discovery environment: A state-of-the-art review*. Retrieved from <http://ir.inflibnet.ac.in/dxml/bitstream/handle/1944/1623/14.pdf?sequence=1>
- Prescott, M., & Veldof, J. (2010). A process approach to defining services for undergraduates. *portal: Libraries and the Academy*, 10(1), 29–56. doi:10.1353/pla.0.0085
- University of Minnesota Libraries. (2006). *A multi-dimensional framework for academic support: Final report*. University of Minnesota Libraries. Retrieved from <http://purl.umn.edu/5540>
- University of Minnesota Libraries. (2010). *Library user communities*. University of Minnesota Libraries. Retrieved from <http://purl.umn.edu/99734>
- Vaughan, J. (2011a). Investigations into library Web-scale discovery services. *Articles (Libraries), Paper 44*. Retrieved from http://digitalcommons.library.unlv.edu/lib_articles/44
- Vaughan, J. (2011b). Web scale discovery services. *Library Technology Reports*, 47(1).
- Yang, S. Q., & Wagner, K. (2010). Evaluating and comparing discovery tools: How close are we towards next generation catalog? *Library Hi Tech*, 28(4), 690–709. doi:10.1108/07378831011096312

ADDITIONAL READING

- Anderson, J. Q., & Rainie, L. (2008). *The future of the Internet III*. Retrieved from <http://pewinternet.org/Reports/2008/The-Future-of-the-Internet-III.aspx>
- Bhatnagar, G., Dennis, S., Duque, G., Henry, S., MacEachern, M., Teasley, S., & Varnum, K. (2010). *University of Michigan article discovery working group final report*. Retrieved June 16, 2011, from <http://www.lib.umich.edu/files/adwg/final-report.pdf>

Setting a Direction for Discovery

- Bowen, J. (2008). Metadata to support next-generation library resource discovery: Lessons from the eXtensible catalog, phase 1. *Information Technology & Libraries*, 27(2), 5–19.
- Brantley, P. (2008). Architectures for collaboration: Roles and expectation for digital libraries. *EDUCAUSE Review*, 43(2), 30–38.
- Cai, H., Dou, T., & Jiang, A. (2011). Effective approaches to the evaluation and selection of a discovery tool. In Xing, C., Crestani, F., & Rauber, A. (Eds.), *Digital libraries: For cultural heritage, knowledge dissemination, and future creation* (pp. 347–356). Berlin, Germany: Springer. doi:10.1007/978-3-642-24826-9_43
- DeFelice, B., Kortfelt, J., Mead, T., Mounts, M., Pawlek, C., Seaman, D. ... Wheelock, C. (2009). *An evaluation of Serials Solutions Summon as a discovery service for the Dartmouth College Library*. Retrieved May 19, 2011, from http://www.dartmouth.edu/~library/admin/docs/Summon_Report.pdf
- Dempsey, L. (2006). The library catalogue in the new discovery environment: Some thoughts. *Ariadne*, 48. Retrieved from <http://www.ariadne.ac.uk/issue48/dempsey/>
- Dentinger, S., Keclik, K., Barclay, A., Bruns, T., Larson, E., Quattrucci, A. ... Walker, N. (2008). *Resource discovery exploratory task force final report*. Retrieved from <http://staff.library.wisc.edu/rdetf/RDETF-final-report.pdf>
- DeRosa, C., Cantrell, J., Hawk, J., & Wilson, A. (2006). *College students' perceptions of libraries and information resources: A report to the OCLC membership*. Retrieved from <http://www.oclc.org/reports/pdfs/studentperceptions.pdf>
- Inger, S., & Gardner, T. (2008). *How readers navigate to scholarly content: Comparing the changing user behaviour between 2005 and 2008 and its impact on publisher Web site design and function*. Retrieved from <http://www.sic.ox14.com/howreadersnavigatetoscholarlycontent.pdf>
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). *2011 Horizon report*. Austin, TX: The New Media Consortium. Retrieved from <http://wp.nmc.org/horizon2011/>
- Kroski, E. (2008). On the move with the mobile Web: Libraries and mobile technologies. *Library Technology Reports*, 44(5).
- Law, D. (2007). Beyond the hybrid library: Libraries in a Web 2.0 world. In Earnshaw, R., & Vince, J. (Eds.), *Digital convergence: Libraries of the future* (pp. 107–118). London, UK: Springer. doi:10.1007/978-1-84628-903-3_9
- Marcus, C., Ball, S., Delserone, L., Hribar, A., & Loftus, W. (2007). *Understanding research behaviors, information resources, and service needs of scientists and graduate students: A study by the University of Minnesota Libraries*. Retrieved from http://conservancy.umn.edu/bitstream/5546/1/Sciences_Assessment_Report_Final.pdf
- Maron, N. L., & Smith, K. K. (2008). *Current models of digital scholarly communication: Results of an investigation conducted by Ithaka for the Association of Research Libraries*. Retrieved from <http://www.arl.org/bm~doc/current-models-report.pdf>
- Nackerud, S., & Scaletta, K. (2008). Blogging in the academy. *New Directions for Student Services*, (124): 71–87. doi:10.1002/ss.296
- Prabha, C., Connaway, L. S., Olszewski, L., & Jenkins, L. R. (2007). What is enough? Satisficing information needs. *The Journal of Documentation*, 63(1), 74–89. doi:10.1108/00220410710723894
- Research Information Network. (2008). *Discovering physical objects: Meeting researchers' needs*. Retrieved from <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/discovering-physical-objects-meeting-researchers->

Research Information Network. (2009). *E-journals: Their use, value, and impact: Final report*. Retrieved from <http://www.rin.ac.uk/our-work/communicating-and-disseminating-research/e-journals-their-use-value-and-impact>

Rowlands, I., Nicholas, D., Williams, P., Huntington, P., Fieldhouse, M., & Gunter, B. (2008). The Google generation: The information behaviour of the researcher of the future. *Aslib Proceedings*, 60(4), 290–310. doi:10.1108/00012530810887953

Schonfeld, R., & Housewright, R. (2010). *Ithaka S&R faculty survey 2009: Key strategic insights for libraries, publishers, and societies*. Retrieved from <http://www.ithaka.org/ithaka-s-r/research/faculty-surveys-2000-2009/Faculty%20Study%202009.pdf>

Shadle, S. (2008). The local catalog is dead! Long live the local catalog! *Serials Review*, 34(2), 85–87. doi:10.1016/j.serrev.2008.03.004

Smith, S. D., Salaway, G., & Caruso, J. B. (2009). *The ECAR study of undergraduate students and information technology, 2009*. Retrieved from <http://www.educause.edu/Resources/TheECARStudyofUndergraduateStu/187215>

Smithsonian Institution. (2010). *Information technology plan*. Retrieved from <http://www.si.edu/content/ocio/pdfs/SITP.pdf>

University of California Libraries Bibliographic Services Task Force. (2005). *Rethinking how we provide bibliographic services for the University of California*. Retrieved from <http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf>

Walker, J. D., & Jorn, L. (2007). *Net generation students at the University of Minnesota, Twin Cities*. Retrieved from http://www.oit.umn.edu/prod/groups/oit/@pub/@oit/@web/@evaluation-research/documents/asset/oit_asset_149639.pdf

Ward, J. L., Shadle, S., & Mofjeld, P. (2008). WorldCat Local at the University of Washington libraries. *Library Technology Reports*, 44(6), 4–41.

KEY TERMS AND DEFINITIONS

Data Object/Digital Object: A resource described as a data or digital object may be at varying levels of granularity, including: citation information, database entries/information, e-journal articles, digitized images, online finding aids, etc. Digital objects may be “born digital” but can also be digitized versions of physical objects.

Discovery and Delivery: The process of finding information about a resource or the existence of an item and obtaining it. Discovery refers to accessing the item digitally or finding a physical location for physical items. Users are now expecting the discovery and delivery to occur together which includes the act of access online content or requesting an item.

Discovery Layer: Interface used for discovery that allows metadata from disparate sources to be presented in a single discovery interface. The discovery layer allows end-users to work within a consistent interface regardless of the source the originating metadata.

Environment: The suite of tools and services presented to users for discovery and access to relevant information resources. The environment may include a library catalog, link resolver, Web sites, licensed resources and/or homegrown databases.

Local Data Sources: Local data sources are data contained or systems managed at a local level. Examples include: the ILS (Integrated Library Systems) for a specific library or consortium or local databases created for specific collections or sub-collections.

Setting a Direction for Discovery

Metadata: Structured data that describes content such as a digital object or a data object, as well as, a description for physical items/content. It may describe the characteristics of a particular resource or object as well.

Metadata Layer: A central index that includes metadata and full-text content, aggregated and made available as a unified searchable index. This central index includes data from multiple sources which may include the ILS, institutional repositories, aggregated collections from publishers or other sources.

Metadata Schema: A system used for the cataloging or structuring descriptive records. The schema defines specific elements and organizes them according to specific rules. Examples of metadata schema include: Dublin Core, Encoded Archival Description (EAD), Anglo-American Cataloging Rules (AACR2), etc.

Next Generation OPAC: An online public access catalog that provides a single search box that produces search results in a unified interface. Next Generations OPACs often incorporate Web 2.0 features such as spell check, facets, and relevancy rankings; along with enrichment features such as: tags, reviews, cover images, etc.

Personas: Set of very specific fictional users, describing knowledge, attitudes, habits, and similar information, used to determine how a product or software will serve a user type, population, or other demographic. Personas are often created to determine the needs, preferences, etc. of potential users.

Personal Curation: Refers to the ability to access, manipulate, and/or organize a digital object and interact it with it by adding personalized data such as tags, annotations, combining it with other objects to create a collection, or manipulating it for the user's own needs.

Scopes: The action of filtering, grouping, or organizing a set of resources in the discovery environment for searching or browsing, based on the tasks associated with various user communities. In effect, scopes are user-centered information silos.

User Communities: The diverse user population that libraries serve including: affiliates, non-affiliates, users of locally-managed interfaces (e.g., website, catalog), users of remote systems (e.g., WorldCat, Google Scholar), communities defined by discipline, geography, task, institutional role (e.g. faculty, staff, undergraduate, graduate student).

ENDNOTES

- ¹ Ex Libris is a trademark of Ex Libris Ltd.
- ² Primo is a registered trademark of Ex Libris Ltd.
- ³ Aleph is a registered trademark of Ex Libris Ltd.
- ⁴ EBSCO Discovery Service is a trademark of EBSCO Publishing.
- ⁵ SFX is a registered trademark of Ex Libris Ltd.
- ⁶ Google Scholar is a trademark of Google, Inc.
- ⁷ OAIster is a registered trademark of OCLC.
- ⁸ Flickr is a registered trademark of Yahoo! Inc.
- ⁹ ArchiveGrid is a service mark of OCLC.
- ¹⁰ Primo Central is owned by Ex Libris Ltd. or its affiliates.
- ¹¹ Summons is owned by ProQuest LLC
- ¹² Worldcat is a registered trademark of OCLC.
- ¹³ EZproxy is a registered trademark of OCLC.