Next Generation Library Systems

Convenient, Connected, User-Centric, Ubiquitous

Next Generation Library Taskforce
Dartmouth College Library

April 2013

Taskforce Members:

1. **Barb Bushor**, Bibliographic Records Supervisor, Cataloging and Metadata Services
2. **Richard Felver**, Business & Engineering Librarian, Feldberg Business & Engineering Library
3. **William Garrity**, Director, Academic and Campus Technical Services, and Associate CIO [until July 2012]
4. **Mark Mounts**, Business & Engineering Librarian, Feldberg Business & Engineering Library; Digital Library Technologies Group
5. **Barb Sagraves**, Head of Preservation Services
6. **David Seaman**, Associate Librarian for Information Management (Chair)
7. **Becky Torrey**, Digital Resources Coordinator, Acquisitions Services

To understand our collective needs, this Taskforce worked with a wide range of Library committees and departments. Examples include the Information Discovery and Access Group (IDAG); the Access Services Round Table (ASRT); Research & Instruction Services (RIS); the Digital Projects and Infrastructure Group (DPIG); and the Preservation, Acquisitions and Cataloging & Metadata Committee (PACC).

Image Credits:

Cover image from UVM: http://alumni.uvm.edu/vq/spring2009/complex.asp
Data Base: http://www.mandiberg.com/2011/12/07/burned-books/
Colorful people: http://webanalysis.blogspot.com/2010/12/online-personalization-issue-is-with.html#axzz2CPep1jrm
Open Access symbol: http://en.wikipedia.org/wiki/Open_access
Europeana PolyMath Graph: http://pro.europeana.eu/polymath-edm
Infinity bookshelf: http://blog.betterworldbooks.com/2008/12/15/infinity-bookshelf/
Business intelligence: http://www.computerweekly.com/blogs/open-source-insider/2010/10/open-source-business-intelligence-on-the-mobile-map.html
Word Cloud: http://www.flickr.com/photos/bonus/3772366436/sizes/m/in/photostream/

TABLE OF CONTENTS

Executive Summary

The Dartmouth College Library is ambitious to select, acquire, describe, curate, and deliver collections and services in a manner that is deeply integrated into the working lives of our users, and which results in excellent research and teaching. We place a premium on convenient access and relevant discovery. We recognize that research and scholarship is interdisciplinary and globally connected, and that Internet-enabled pedagogy is increasingly commonplace. Our library systems must be nimble enough to adjust to these evolving scholarly behaviors, and we expect that the pace of change is only likely to quicken.

Library systems are complex, encompassing a range of business workflows, discovery processes, physical inventories, and electronic resource collections. They must serve disparate user communities, from the first year student to the world-class subject expert. Cutting edge scholarship requires a research library not only to provide access to online journals but also to medieval manuscripts, as well as to the expert library staff who make these systems work.

In past decades these disparate needs led research libraries to the Integrated Library System, or ILS: a modular infrastructure, usually from a commercial vendor, which combines purchasing, description, and inventory functions with the ability to search the library holdings and discover materials. Since the development of these ILS systems, however, the Internet has massively widened our field of view. No longer do our collections exist predominantly in the same physical location that our users do; no longer is it enough to know from a library catalog that we own a scholarly journal – now we expect to read those journal articles online. The problems of managing information scarcity have given way to the problems of managing information super-abundance.

The network of campus and vendor computer systems in which the ILS must function has become more complex too. In order to drive the services we wish to provide, our ILS systems must interoperate with the Registrar's database of class enrollments, the course management system, identity management knowledge bases, institutional repositories, faculty profile systems, alumni portals, and other data sources. The ability to share data between systems is now a core competency, rather than an additional feature.

All of this change has led to new opportunities for library system design. Where the choice a decade ago was between several large commercial systems, now there is a thriving open-source community, as well as a new breed of cloud-based approaches to library business processes and scholarly discovery. We are ambitious to place the Dartmouth College Library at the center of the research, teaching, and publication processes of our users, and to do that we are undertaking this evaluation of the functional requirements of the systems on which both staff and users rely.

Introduction

Current Situation

Most academic libraries are operating an ILS that has developed over many years. Many of these systems are informed by core design assumptions rooted in an era when library collections were predominantly comprised of physical objects, and when the library catalog was the central inventory and discovery tool. At Dartmouth we have successfully used the same commercial ILS for 20+ years, and it has grown in a modular fashion in response to our changing needs. It provides our public catalog interface alongside a suite of acquisitions, cataloging, circulation, and inventory control functions.

Our existing Library infrastructure can be articulated as two main clusters of activity:

- *User Experience*: the interfaces and relationships through which users learn about and interact with the Library's collections, services, and human expertise, to create knowledge.

- *Business Processes*: the workflows and systems that allow Library staff to select, purchase, catalog, inventory, preserve, and control access to our materials and services, in pursuit of an excellent user experience.

Discovery systems and business processes are different functions within an ILS; indeed, that very integration of these tasks into a common infrastructure was seen as a move towards greater efficiency in past decades. Increasingly, we realize that the systems that control business processes could (again) be separable from those that enable discovery, as long as they communicate effectively with one another.

In recent years, strategic efforts have focused on improving discovery systems, including Dartmouth College Library's *Next Generation Library Report* (2008)[1], and the subsequent adoption of a unified discovery service.[2] This 2013 *Report* includes discovery processes within the larger context of the integrated systems we need to build and manage collections, staff, and services over time to serve excellent research and teaching.[3]

---

[1] http://www.dartmouth.edu/~library/admin/docs/NextGenLibraryReport.pdf

[2] Dartmouth College Library was one of two initial beta partners for this entirely new class of online discovery service in 2008; in 2013, *Summon* is used by over 500 libraries in 40 countries, with a unified index of over 1 billion items. http://www.serialssolutions.com/en/press_room/detail/2169

[3] Both the recent *Preservation, Acquisitions and Cataloging & Metadata Committee (PACC) Elimination Goals* and *The Save the Time of the Reader* project have also discovered and acted on workflow efficiencies within our existing library system.

The Wider Picture

After several decades of stability, with most U.S. research libraries using an ILS from one of several large commercial companies, there has been noticeable recent volatility in library systems conceptualization and design, with several diverging development paths:

> **Commercial Systems –** ILS vendors provide locally installed systems that one purchases and on whose modules one pays annual maintenance. These systems are relatively slow to change, although vendors are reacting to expressed customer needs with additional service offerings, including both new catalog interfaces and library services platforms.
>
> **Open Source Platforms** – This category of freely available software built collaboratively is diversifying and maturing. It now includes delivery systems, business processing infrastructures, and complete ILS suites.
>
> **Commercial "Web-Scale" Offerings** – A new model entirely, based on both business processing and discovery services that reside at the network level, rather than in systems that are hosted and maintained locally.

Both open source and the web-scale library services are expected to be ready for wide adoption by Association of Research Libraries (ARL) member libraries as early as 2014, so this is an excellent time to undertake an overview of contemporary library service needs.

Planning at Dartmouth

Increasingly, the ILS must interact with other campus and vendor services – repositories, publishing platforms, course management software, and publishers' systems – in order to be properly embedded in the teaching and research environments our users inhabit. Currently, our systems are far too isolated one from another. Any one of the development paths above could be the best fit for Dartmouth in the years ahead, but without some concerted thought about what we really want to do, we cannot ensure we have the best systems architecture for our staff and users. This Report then focuses not on a comparison of contemporary product offerings, but rather asks the question: "What do we want our systems to do for us?"

We recognize that large-scale systemic changes have major staffing implications, as well as technology and service ones. This report does not make direct recommendations about new and revised staffing roles, but does recognize that our ambitions to work in new ways and drive new services will require new skills and work practices.

Organization of Findings

In the conversations the Taskforce convened across Library departments and committees in 2012-13, what we heard predominantly was a focus on what users wanted, and what processes we needed to achieve these goals. The categories under which we have organized the feedback and observations are terms that we hope to be readily understandable to an academic audience generally. The orders of the sections, and of their desired functions, are not ranked according to importance or priority.

The User Experience

- *Discovery*
- *Personalization*
- *Re-Use*
- *Assessment and Analytics*

Business Processes

- *Collection Development*
- *Metadata Creation and Management*
- *Collection Management*
- *Systems Integration*

Overlaps and linkages inevitably arise between sections – for example, what we learn from *personalization* services becomes *business intelligence* that drives better *discovery* and *systems integration*.

Scenarios

To help understand the impact of various aspects of this report, we have provided some stories, or scenarios, in **Appendix I** that illustrate how the future might look when the systems and services described in this report are in place. These are not comprehensive – they do not embody all aspects of this report – but as vignettes they may be helpful.

Discovery

Current Situation

Discovery systems support an iterative set of tasks that comprise the information-seeking process. Although discovery includes "searching," it may also start with a referral or a recommendation. Post "search" tasks are also becoming critical to successful information seeking, including retrieving content for local uses (see ***Re-Use*** section).

Dartmouth currently uses a discovery platform that allows a single search over the library catalog's content, as well as most of our online journal articles, eBooks, dissertations, online newspapers, and other materials (468 million items are currently listed as part of the Dartmouth collection, of which 436 million have full text online).[4] *Finding* resources, therefore, is not a problem facing our students and faculty. The big problem is finding the *right* information when faced with overwhelming result sets.

Users have different discovery needs and habits. Undergraduates with short time frames for their projects want a few top resources, and hope to find the complete item or object immediately. Graduate students need to perform comprehensive literature searches in specialized domains, and often have to document this undertaking. Faculty members value different aspects of the discovery process, sometimes favoring serendipity and sometimes favoring precision and currency. Library staff need powerful search mechanisms to assist users or to enhance their work processes. However, our current discovery systems present the same results to all users, and are not sensitive to different user communities.

Contemporary discovery systems also need to be deeply and fully integrated into the online places and devices in which we research, teach, and work. Users want to discover sources related to a concept, find a known item, find materials related to that item, find materials by chance through serendipity, and remember the discovery processes used so they can be repeated. Moreover, our 21st century discovery systems should help one find more than books, journals, and other library holdings – they should also uncover unseen relationships between resources, and help users find subject specialists and experts in their field.

[4] Figures from Summon in March 2013: http://dartmouth.summon.serialssolutions.com/

Desired Functions

Convenient

- Search results should be fast and accurate – a user should get an immediate response and be able to trust that it is reliable.
- Content should be linked directly from metadata records.
- The discovery environment should help users navigate huge result sets, perhaps through graphical representations of what is discovered. It should also provide users with all instances of the same work and the ability to get them easily (e.g. all the available editions of Shakespeare's *Hamlet*).
- A user should never reach a dead end: they should always be provided other discovery options if no results are found.

Locally Relevant

- While the discovery layer must help us navigate the universe of scholarly materials, it should also be Dartmouth aware. For example, it could highlight Dartmouth authors; allow us to apply our own terminology (perhaps by highlighting content relevant to local courses); or link local blog entries, reviews and exhibits into discovery.
- We should find ways to emphasize the materials created locally (scholarship, digitized collections, etc.).
- Library experts need to be easily discovered. "Find a librarian" services should be deeply integrated into our content discovery systems.
- We need broader accommodation of local forms of classification.
- We need support for commentary, reviews, and tagging by local users.[5]

Responsive To User Needs

- Our discovery services need to be relevant to senior researchers and to incoming freshmen, to domain experts and to scholars venturing into new fields of study. One size does not fit all.
- Browsing and serendipitous discovery is as important online as it is on the physical bookshelf, and should be supported robustly.[6]
- We need the ability for a user to create virtual collections of items that one has read or would like to read.
- Library staff members need a "power search" that is easy to use and is able to search all data including "staff access only" data.

---

[5] An example can be seen here: http://gilfind.uga.edu/vufind/Record/3634193

[6] A discussion of Stanford's new system can be found here: http://www.ibiblio.org/bess/?p=248

- The library should be able to define the content universe, which includes all licensed, owned, shared, and free content. Library content should be discoverable by the use of discipline-specific vocabularies.

Helpful and Alert

- Helpful features include persistent URLs; access to the best versions of full text; display of resource relationships (linked data); location mapping for materials; and complementary data such as tables of contents, cover art, etc.
- Amazon.com-like "see inside the book/video/audio" previews should be commonplace.
- The systems should highlight what is new, and advise on "best bets" based on librarian and user recommendations.
- A Library Chat service, embedded within the discovery system's user interface, should be available to provide referrals to library experts.
- The system should help disambiguate authors' names.
- The system should present a view that is graphic and less list-like, and display relationships between results in a graphical interface.

Open and Configurable

- Discovery systems need to be open and configurable, allowing for customization where appropriate.
- Authentication should be transparent – a single sign-on should carry over through all the content sites that the discovery system allows you to discover.

Analog Friendly

- When no e-version is available, users need clear references to available print versions, or to information on how to borrow the item.
- One should have the ability to ask for a print item and not have to know where it is located.
- A user requesting an item not available as full text online should be able to apply time constraints as part of that request ("I only want to see information I can get within 2 hours; 48 hours; etc.")
- Way finding – we should be able to show a user exactly where the book, DVD, or musical score is shelved, using interactive navigation.[7]

[7] See as an example: http://josiah.brown.edu/record=b5866515~S7

Mobile

- Mobile integration is vital throughout our discovery layer. Interfaces should work well on cell phones, tablets, and other mobile devices.
- Mobile services should include the ability to text a book's call number to a cell phone from the discovery interface.

Summary

The largest change to our library infrastructure in recent years has been the new demands on discovery (and new service opportunities there too). The working habits and expectations of our users are changing rapidly and we need discovery systems flexible enough to keep up. Access to massive amounts of content (both electronic and print), and the need for discovery to be ever more convenient, are the new realities, as is the rise of mobile platforms to which we need to deliver a superior user experience.

Personalization

Current Situation

Our library systems currently offer only modest abilities for individuals to personalize services. In addition, our systems do not enable library staff to construct services that are customized according to what users allow us to know about their usage habits. Current infrastructures support little in the way of alerting services, or customized views of our holdings. Staff cannot personalize their daily work environment in the current systems either, and reporting functions are limited, both of which can hinder productivity.

Desired Functions

Customization

- We know a lot about students and faculty (what level of academic experience they have; what classes they are teaching and taking; what they have published) and the systems should talk to each other to make use of this information *when the user allows it*.
- Users should be able to opt in to Library services that make use of their personal data and user habits.
- Social media activities like recommending, tagging, reviewing, rating, and sharing should be possible in the ILS.
- Library users should be able to authenticate into library services with social site logins, as they can in many commercial systems.
- There is a need for all users to be able to customize their work environment by simplifying how they log in, change the look and feel of programs (both the data layer and the presentation layer), and access their most frequently used tools easily.

Smart Searching

- Users should have a discovery environment that can remember their preferred databases and make use of their saved queries.
- We should be able to integrate faculty profiles (including research, publication and teaching output) into finding tools to drive new services.

Helpful and Alert

- Users should be able to receive notifications based on saved preferences, supplied keywords, or even entire bibliographies.

- The library should provide an Amazon.com-like service where a user can see what others with similar interests have searched, how many times a given item has been used, and provide the top ten items on a given topic.

Summary

Systems need to be more aware of the individuals' needs, habits, and preferences. They need to be personalized recommender systems, as well as places to ask questions. Users often expect the software environment to intuit what they need, and customized feeds of information are increasingly necessary as the material made available through the library grows in bulk and complexity.

The ability to opt in and out of personalized services is critical, as such services often requires staff to track online information-seeking behaviors or to tap into course enrollment systems. However, such personalized recommender systems are so prevalent in the e-commerce and social network spaces that to ignore these behaviors in the ILS almost entirely (as we currently do) is to miss service opportunities.

Re-Use

Current Situation

Both library users and staff members want to re-use materials that they discover. This re-use may include re-publication, course reserves, mash-ups, inclusion in presentations, and integration into other data sets. While we have made advances recently in the discovery of materials in our collections, there has been less progress in the ease with which we can re-use library materials in other settings. Users too often lack convenient ways to export holdings from our online systems, and library materials often lack a clear articulation of the rights information that can inform and encourage re-use.

Within the Library, we are seeing frequently exporting metadata from one system to re-use in another. We download catalog records to re-use in our discovery tool, for example, or to include as metadata in an online collection. These wholesale and targeted exports are currently labor intensive and cumbersome. Library staff members need much more flexible processes as they transfer content from system to system.

Desired Functions

Export

- Robust data export and manipulation is important for metadata, citations, full text, and other digital objects.
- Exported result sets need to be manageable by any platform or service (mobile, RSS, web services, batch loading).
- We need to convey to our user what technical or legal limits, if any, there are on re-use of library content.
- Library-generated data should be open and available for anyone to use, when legally permissible and when it does not contain personally identifiable information.

Tools

- Tools to manipulate objects such as video, audio, images, and text should be integrated with the discovery service.
- If the user chooses to re-use content in another tool or platform, the content should be exportable in standard formats and with metadata (provenance, ownership, date and time of export, etc.).
- Any online infrastructure with which the user interacts, including the course management system, should be able to integrate materials from the library.

Experts

- We need to provide online and in-person assistance with copyright and licensing to promote richer re-use.
- We need to provide online and in-person assistance with creating and publishing new content from existing materials.

Summary

Creative users need malleable content with which to innovate; the current systems promote discovery, but hamper re-use. Part of the barrier includes uncertainty about copyright and licenses.

Assessment and Analytics

Current Situation

The Dartmouth College Library is embedding a culture of assessment across its services and operations. Increasingly, we measure the outcomes of our actions, and we make data-informed, evidence-based business decisions. The Library has an Assessment Committee that "encourage[s] the use and implementation of assessment methods to assure effective library programs, collections and services,"[8] and outcomes-based assessment is woven into our strategic objectives. We undertake surveys, focus groups, and other data-collecting activities, and download usage statistics from multiple vendors and organizations.

Many of our systems track usage data, or store information about what we buy, their cost, and what subject coverage we have. However, it is difficult to get data from our systems and vendors in a malleable form that can be integrated into meaningful reports. Too often, we lack the data integration we need to inform the best decisions about services and collections, to drive efficiencies, and to spot emerging trends. Our desire is to use data in a connected and malleable way, as an aid to our decision-making and service development.

Desired Functions

User Feedback

- We need processes to capture feedback from users, including recommendations and user reviews, which we can use as source data for service decisions.
- We need easier access to information about faculty and student scholarly output and interests. This data should be sourced from what they publish, what they teach, and from new "alt-metrics" services that gauge their network presence in other ways.
- Collections trends and activity should be exposed to the public in real time on our web site and on the flat screens in the various libraries – showing, for example, use of material by language or subject; highlighting various trending search topics; or showing usage by media type.

[8] http://www.dartmouth.edu/~library/home/committees/assessment/

Broad Access and Use of Data

- We need easy, aggregated access to usage information from multiple sources simultaneously (web, library catalog, interlibrary loan, etc.)
- We need to connect disparate sources of information for more comprehensive views of what is happening in our collections. For example, usage data should be interlinked with price data for an item or database; collections development data should interact with public service data, and so on.
- All staff should be able to build reports to match their needs. A "dashboard" display would be desirable to give a quick overview and a sense of trending topics or needs.
- All significant business intelligence data should be located centrally so that multiple users can access it for their own purposes.

New Business Efficiencies

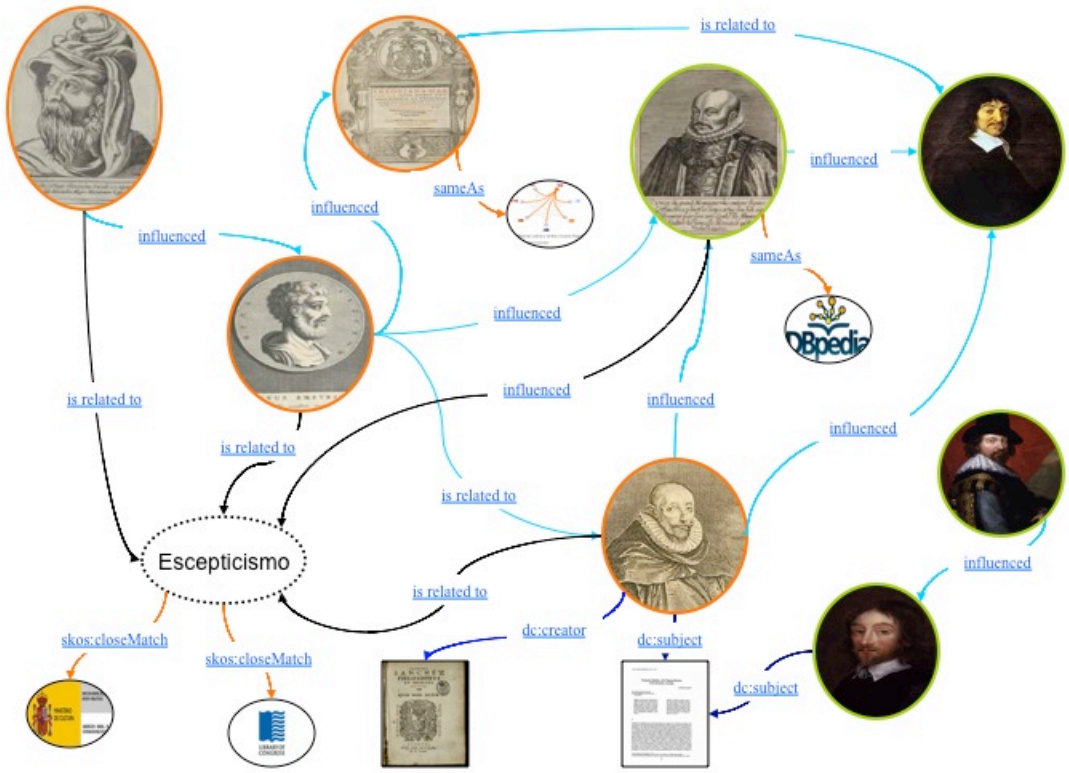
- We should be able to define new variables to measure and analyze results.
- Staff need reports on the expenditure of funds that are flexible and can be tracked over time, and automated report generation for topics we track on a regular basis.
- Our services increasingly require real-time data exchange between systems. For example, the Registrar's database should tell our systems what students are signed up for courses in order to control access to course reserves; patron information should load automatically and immediately from the campus identity management infrastructure; and real-time fund reporting should be commonplace, including scope of fund, purchases made, and usage of materials.
- We need better ways to extract, reformat, and analyze data from our systems in order to use it for external reporting or collaborations, and for internal cost-benefit analyses.

Asset Tracking

- We need statistics on use of databases by subject.
- We need statistics on the number of times an article has been viewed when, where, and by whom.
- We need statistics on high-use items held locally or consortially. And we need to see circulation data for an item in total but also across time periods (to see if a topic is coming back into, or going out of, favor).
- We must be able to track inventory and generate reports. For example, the system should allow us to track physical items in the collection as they move from an existing location, to temporary "swing space," to a new location.
- And we need the ability to track digital items in the collection, to know where and how they are stored, and to track changes in them over time.

Summary

We need much easier access to information about library usage and collections patterns, in order to drive better services and work environments. We need better pathways, embedded throughout our discovery and use systems, for users to tell us what they want so we do not have to intuit it; we need more seamless linking up of disparate data sets into an understandable and actionable view of our shifting business and service landscape. We are – as Columbia University's Jim Neal has observed[9] – an "information poor" information industry when it comes to the patterns of purchase and use that our infrastructure could provide us if this data were integrated.



[9] Neal, J. G. (2006). The Research and Development Imperative in the Academic Library: Path to the Future. *Portal: Libraries and the Academy* 6(1), 1-3. The Johns Hopkins University Press. Retrieved October 30, 2012, from Project MUSE database.

Collection Development

Current Situation

Selection

We acquire a wide array of library materials every year. Physical collections, including books, DVDs, and some microfilm, grow by over 30,000 titles per year. The majority of new titles are digital, largely in the form of e-journals and e-books, often acquired in large sets. Current holdings show 3.6 million e-books, and 44 million articles from scholarly publications.

Library bibliographers select physical and digital library materials, which are usually ordered electronically via the ILS; however, some orders continue to be made via email, the web, within vendor systems, by telephone, and through paper mail.

Ordering

All orders are processed through the ILS. However, processing workflows for physical materials are often distinct from those for digital materials, and too often the workflows for the acquisition of digital materials are forced to fit into processes designed for physical material acquisition.

Acquisitions Services receives lists of titles to order from library bibliographers and determines the best commercial source with which to place the order, based on material format, language, country of publication, vendor discounts and performance, and other criteria. Orders for English-language, U.S., Canadian, and UK books are purchased and administered through one vendor, and data is electronically transferred into the ILS to form the basis of the order and invoicing.

Acquisitions Services staff receives all physical materials for the circulating collections, sets up access for electronic materials, manages continuing resource subscriptions, creates, maintains, and monitors multiple vendor accounts, and approves all material invoices for payment.

The Library also develops collections of locally digitized materials in the Dartmouth Digital Library Program. These collections still need to be integrated into the data set used by library bibliographers during collection development and analysis.

Fund Accounting

The Information Resources (IR) budget for library materials is a subset of the overall library budget, and is comprised of a combination of funding sources, including restricted and unrestricted endowments, restricted gifts, and institutional

subvention. These resources are distributed to over 400 individual subject funds, and fund accounting is controlled within the ILS, tracking allocations, expenditures, and fund balances. The terms of the individual funds must be enforceable within the ILS, because many funds are drawn from restricted endowments.

Today, the ILS acquisitions modules only interface with the campus financial systems through Electronic Data Interchange (EDI) processing for accounts payable functions. While the majority of invoice payments are processed through EDI, many invoices must still be handled manually, due to various signatory or supporting document requirements based on invoice amount. Acquisitions Services also manages invoice payment for services that support library materials, such as metadata fees, Open Access author fees, resource sharing fees, and purchases from Special Collections endowments. These invoices also require special manual processing. Acquisitions Services currently manages invoices of record through paper files.

Library staff members regularly monitor and reconcile funds in both the Library and the campus financial systems, to ensure that funds and endowments are not overspent. The campus financial system tracks expenditures at the account level, while the ILS manages budgets at the individual fund and subject level. The library's 400+ funds "roll up" to a much smaller set of larger accounts in the institution's financial system. Matching these balances is challenging and time-consuming.

Desired Functions

Fund management

- Library fund accounting should be integrated with campus financial systems.
- Library fund management should be integrated with the Library's budget preparation process. All budget preparation is currently done outside the ILS.

Invoice management

- We need an electronic method for managing required signatures for invoices prior to payment.
- Invoices-of-record should be archived electronically and retrievable by various search criteria.
- Bills must be paid quickly and accurately. Fiscal year close activities must be supported.

Selection

- Bibliographers need to be able to generate reports of their program's activities.

- The ILS should support selecting, ordering, implementing, and managing commercial digital library resources.
- The Library needs to be able to choose which vendors will best supply the materials we need and create and maintain business relationships with any vendor.
- Library systems should integrate the management of patron-driven acquisitions. Examples include e-books that are "selected" from sets by user "clicks", or support for on-demand acquisitions processed by the Library's Resource Sharing office.

Data exchange

- Metadata needs to be easily transferrable between vendor databases and the ILS to facilitate the creation and management of orders, the presentation of metadata about on-order items to users, the ability for users to place holds on on-order titles, the avoidance of duplication in the ordering process, and the creation of business intelligence so that the Library can easily analyze and evaluate its vendors.

Summary

The scope of the library collection development program requires that a large number of bibliographers manage more than 400 separate funds, and requires that Acquisitions Services be able to make purchases from any vendor who provides quality service at the best cost. Better integration of fund accounting functions and campus financial systems would introduce efficiencies in workflows. There is little room for error in fund accounting workflows; systems designed to support these workflows should enforce campus policies regarding fund accounting and management.

Expenditure of staff time is necessary in almost all ILS processes, from discarding of material, to expedited services, to the comparison of our electronic holdings to the holdings in LOCKSS[10] and Portico[11]. We manage our inventory of holdings through the ILS, with little inventory management (except for circulation processes) occurring once material is in the stacks.

Any new systems should greatly increase the efficiency of our staff as they manage complex and massive collections of analog and digital items with the accompanying legal and financial data.

[10] LOCKSS: Lots of Copies Keep Stuff Safe: http://www.lockss.org/
[11] Portico: http://www.portico.org/digital-preservation/

Metadata Creation and Management

Current Situation

Metadata, or structured data about data, is central to library services. Librarians create and manage metadata to describe what an item is, where it can be accessed, how it can be used, whether it has been used, how it was acquired, what other items it is related to, whether it is received once or continuously (like a journal is), whether it has been lost, whether another user has recalled it, whether we have decided to discard it, and more. Metadata is also created and managed for locally digitized collections to record technical details about the object, although processes to manage these metadata are still under construction.

Currently, most descriptive metadata is managed in the ILS, which enforces policies protecting the integrity of these data. The ILS is the "database of record"; it defines the validity and credibility of the data. Other metadata – for Special Collections holdings, College administrative records, or locally digitized materials – are stored and managed in other local repositories, which are the databases of record for the data they hold. These various descriptive metadata repositories are not integrated; when we move metadata between these systems it is labor intensive and there is potential for data to be lost or corrupted. Under these conditions, it is increasingly difficult to guarantee that the "database of record" concept still holds, or to have the ability to record actions taken on the data over time.

Current systems require significant staff intervention. Material needs to be checked in manually and library records need to be updated in multiple places to communicate information to our local users and consortial partners.

Different formats currently are managed in different systems, some in the ILS and some in other commercial databases. For the most part, print journals are managed in the ILS, while electronic journals are managed in a separate commercial system. We manually enter most of our electronic holdings into this system, receive metadata for e-books and electronic serials here, and in turn benefit from services that utilize these data.

Vendors provide use counts to the library through external systems; these data are currently collected and analyzed outside the ILS. Google Analytics collects use data on local digital collections, also stored and processed externally. The Electronic Resource Management (ERM) module of the ILS also has a feature for gathering electronic resource usage statistics and creating spreadsheets with cost per use figures; however, it has a complex set-up and requires time-consuming maintenance.

Other important processes occur outside of the ILS entirely. Vendor services such as commercial binding, mass de-acidification, or film preservation often have tracking systems that do not integrate with the ILS; large journal packages are managed on

standalone spreadsheets outside of the ILS; and user requests for new acquisitions are made via a web form, which then goes to the relevant selector to make a decision.

Desired Functions

Records Creation and Maintenance

- Metadata repositories should be integrated to communicate with each other.
- Policies that ensure "database of record" validity and integrity must be definable and enforceable in any system where metadata are stored and managed.
- Metadata are never static; any system that stores metadata needs to support management of the data and the relevant policies.
- We need greater granularity in setting authorizations for adding, modifying and deleting data within any library metadata repository. In particular, we need meticulous control over the editing of bibliographic records, to ensure that library policies are enforced and that metadata are not corrupted.
- Previous versions of edited records should be stored and accessible, which record the identity of the editor and the nature of the changes made. We need to be able to revert to a previous version of the metadata if errors have occurred or data has been corrupted.
- We need to edit metadata while viewing the objects being described, while consulting the database into which the object's metadata is being added, and while consulting web-based reference materials.
- The metadata editing software should be aware of standard schema and capable of being "taught" locally defined schema. It should be able to work with multiple schemas and allow staff to move seamlessly from one to another.
- We need to be able to choose to use metadata from any source available in the marketplace, and not be limited to specific vendors' metadata stores.
- We need robust tools to create and edit metadata in batches.
- We need the ability to export and import large sets of library metadata – many library collaborative processes, and subscriptions to commercial services, require this routinely.
- We need to be able to use query and transformation tools to define complex conditions and tailor metadata for various uses.
- We need the ability to create an unlimited number of indexes, defined locally, and the flexibility to re-define existing indexes.
- We need the ability to "publish" local vocabularies or other data sets in a manner that conforms to linked data protocols.
- We need to create and store local administrative metadata to support data management and collection management processes.

- Much of the metadata stored in library systems is not intended for display to the public. We need deep granularity and control over what metadata are published to the public and when.
- Local policies about the use and re-use of our metadata must be enforceable by any library metadata repository. Library policies determine when, where, and how to publish these data or to make them available for harvesting, crawling, or other capture and use.

Summary

Library metadata represent an enormous investment, and is generated and validated through time-intensive processes performed by library staff. The metadata are as old as the collection, and must be protected and preserved in the same manner as the collection itself. All library services are dependent on the availability of accurate and valid metadata. Metadata creation and management functions are the fundamental core of any library system of the present and the future. We would expect any future library system to be much more efficient and integrated in its creation and use of metadata (see **Appendix II**).

Collection Management

Current Situation

Library collections are in a period of disruptive change. Users rely heavily on digital information sources for much of their information needs. Increasingly, we assume that library users are more tolerant of wait times for books to be delivered from remote storage facilities for their use. Shelf browsing is assumed to be of less value because it excludes reference to the digital materials in our collection. Other purposes for on-campus library space, such as learning commons, group study areas, media labs, and social spaces are causing libraries to re-examine the role of the on-campus physical collection. As preservation methodologies improve for e-journal and e-book materials, the need for libraries to hold print equivalents diminishes. Libraries are actively pursuing the creation of an infrastructure to support multi-institutional print repositories with collaborative collection retention, access, storage, and preservation activities and policies.

Collection management is a highly distributed function in the Library. It involves staff members in all of the libraries working in collaboration with Acquisitions, Preservation, and Cataloging and Metadata Services staff members. Collection Management functions are not segregated from other library services, but are an integral part of them.

Collection management activities are intensely data driven, and so the ability to manipulate and mine local library metadata is essential. Current practices often require that metadata be reported in MS Excel spreadsheets that are used during collections projects. A deep challenge currently facing the Library is how to get the metadata recorded in those project spreadsheets back into the ILS for further use.

The ILS contains a wealth of local business process data that needs to be mined to support decisions about the disposition and storage of physical collections in this rapidly changing environment. “Business process data” includes payment histories for continuing resources (such as serials), archived usage information, use policies (or restrictions on use), identification of multiple copies or formats or editions, license terms and restrictions, and the current physical storage location. Other information that may be relevant to shared print repositories include preservation actions on an item, the relationship between print holdings and digital holdings, how many other partner libraries also hold a title, and whether we have made a commitment to retain the title through collaborative collection development or other agreements.

Process management is also an essential part of collection management, but it is not well accommodated in the existing ILS. For example, most annual subscription renewal, cancellation, and change requests are currently handled in a separate

database that was locally created and designed specifically to map, track, and document the various stages. Only the final decision gets entered into the ILS.

Electronic resource acquisition offers an example of the growing need to track a complex decision-making process prior to purchase, which is not the case with print acquisition processes. Electronic resources often require comparing product offers from multiple sources, with various pricing options, discounts, license restrictions, archival availability, and many other considerations – an extremely labor intensive process. With the majority of our budget spent on electronic resources, we have a critical need to manage efficiently the array of information through a complex decision path. ERM processes will need to be integrated wholly into future library systems.[12]

While library users search metadata in web-based searchable indexes that provide instant results, library staff members work in an environment where metadata is discovered through old command-line interfaces and complex search strategies, requiring deep knowledge of the data structure in each record type. The staff-side reporting mechanisms feature non-instant results, limited export formats, and virtually no ability to manipulate result sets.

Desired Functions

Business Processes

- Data editing policies, developed to ensure the continuing integrity of the ILS, must be enforceable by the ILS.
- Library systems should be able to store business process information about Library materials on the item level to support decision-making.
- We need to archive our decisions to retain items or discard them, to store them locally or remotely, to add an item to a shared collection, or to keep an item for a defined number of years. These data should be discoverable through searchable indexes, protected from public view, and shareable with consortium and other partners.
- Library use data needs to be integrated into collections management processes, which include the ability to track processes and processing decisions from the initial offer through decision and purchase.
- Managing the physical collections requires immediate staff access, through robust staff-user interfaces, to all forms of library metadata, and the ability to manipulate these data directly.

---

[12] For an up- to-date report on critical needs and components of a successful ERM, see "Making Good on the Promise of ERM: a Standards and Best Practices discussion paper", a NISO White Paper. http://www.niso.org/apps/group_public/download.php/7946/Making_Good_on_the_Promise_of_ERM.pdf

- Information about locally produced digital collections needs to be integrated into these collection analysis functions, and not kept in separate management silos.

Inventory Control

- Staff members should be able to survey the contents of the shelves electronically as they inventory the collection in the stacks, and to be able to see in real time if an item is checked out, missing, or misplaced.
- Staff members should be able to key in a call number and create a bar code using mobile equipment.
- Staff members should be able to track an item as it moves through various workflows.

Electronic Resources Management

- Staff should be able to use an integrated infrastructure to manage all aspects of the electronic resources life cycle, including selection, decision making, and ongoing maintenance.
- Data from vendors and publishers should be importable and actionable.
- Staff should only have to enter data once and be able to link to it from any type of record.
- License terms should be easily shared via the discovery interface, allowing the fullest possible re-use of items and resource-sharing opportunities.
- A workflow utility should map electronic resource management workflows and be customizable to accommodate local practices.
- Electronic resource management should support current standards for digital information exchange.[13]
- An ERM system should differentiate between current subscriptions and previously subscribed materials when making cost-per-use calculations.
- Staff should be able to search for all types of records in the ILS; for example, to search for a license record to see which resources are associated with it, and to search by vendor or publisher to see all associated resources and packages.
- Usage statistics data should be easily incorporated into the ILS and simple to manage. This would allow us to calculate cost per use to help in collection decision-making.
- Staff should be able to identify information that needs to be available to users such as a service outage notice for a commercial database.
- Staff should be able to manage locally created collections as well as vendor and publisher provided collections.

[13] Examples include Project TRANSFER, ONIX-PL and Institutional Identifiers (NISO I^2).

- Staff members need to be able to track titles that move from one provider to another.

Summary

Collection management processes, at the full collection level, are a high priority for the library. An increasing proportion of the Information Resources budget is devoted to materials in digital form, and the ILS should support selection and acquisition processes for these materials, including the management of contracts and license metadata.

Collection management requires data for decision-making, and the ability to manipulate and share these data. Decision makers need access to a "dashboard" of library metadata selected and arranged according to customizable criteria to support large-scale collection management processes.

Systems Integration

Current Situation

The systems landscape in a contemporary research university is complex, and it has grown up incrementally. Too often, critical systems are silos of information, however excellent they may be at their core function. Our library infrastructure is no exception.

Current integration of information across systems is often time-consuming and requires programmers to write and maintain scripts and processes to move data from one system to another. While there is some integration between modules in the ILS – the financial module can interact with the ERM to produce cost per use information, for example – this level of machine-to-machine interaction is too often missing.

For our business and service requirements, the ILS needs to interact across the campus systems architecture and beyond, including student enrollment, HR systems, local and remote financial databases, vendor sites for ordering and invoicing, course management, electronic reserves, inter-library loan, discovery, and metadata providers, to name just the obvious candidates. A core service such patron loads – the loading of user information into (and out of) the ILS – draws on student and human resources systems, is largely a manual process, and therefore occurs only once a term. This is one example of many of a data exchange that should be automated and run frequently.

Desired Functions

Integrated User Services

- The library infrastructure should have easy, automated access to data from other campus sources to enrich the search, retrieval, management, re-use, recombination, and sharing of our collections.
- Real-time interactions between systems should be commonplace, to insure currency of information to all users.
- Faculty profiles should be able to be integrated into finding tools, including their research, publication and teaching output. Such integration requires the sharing of institutional information such as basic personnel data, courses taught, publishing records, and grants.
- Local collections, including archives, special collections, digitized collections, and digital publishing, should be seamlessly integrated into discovery tools.
- Library information and services should be integrated into whatever online tool is being used, regardless of where the tool comes from.

- The discovery environment needs to support new ways to link to related external materials, references, and resources (example: connections to HathiTrust Digital Library material in the library's discovery tools).
- Users need an integrated map of campus that shows spaces within buildings, both to find people's offices and to the physical location of a book.

Grants Systems Integration

- Researchers would like to be able to find funding opportunities along with locating research materials, and to submit the results of their research as required by funding agencies more easily than they do now.
- Grant proposals require a well-documented literature search, a current CV, and a record of previous related work. An ILS needs to support this kind of literature search, saving of citations, and uploading the results into the grant proposal form.
- Alerts about new funding opportunities and about relevant new research are more accurate and effective if the systems providing them can use the data in the faculty profile, such as metadata about the previous publications or grants.
- Integration into the new research administration system and compliance software that Dartmouth's Office of Sponsored Projects is implementing.

Financial Systems Integration

- Our library systems need real time financial data transfer to other trusted systems: for example, the current cumbersome process for transferring fines and fees to Student Accounts is unsatisfactory and time-consuming. Users should be able to pay fines through an on-line system that accepts credit cards.

Course Management Integration

- We want to be able to navigate from one infrastructure to another: from the ILS to Blackboard, etc. We want a more consistent way to move from platform to platform.
- We need to integrate online reserves with the course management system, to allow for the auto-population of course pages with reserve readings. And as items are added to course pages the same items are populated in the online reserves system.

Reporting Systems Integration

- We need better interoperability between library and institutional systems (HR, Financial, Registrar, Office of Institutional Research).

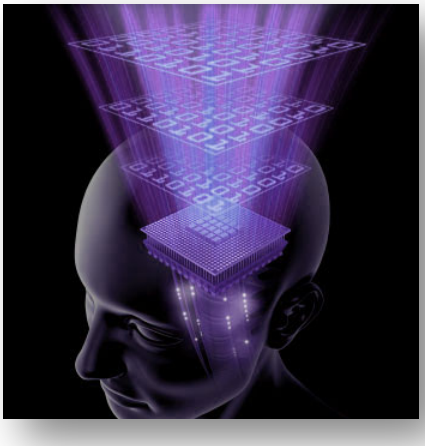
- Library systems need to interoperate with the software that produces institution-wide reports about faculty output, to help keep those reports up-to-date.
- We need to be able to import data from any external non-Dartmouth source, and to export to external systems.

Multi-Institutional Integration

- Permit multi-institutional integration of operations and staff. For example, sharing staff expertise in areas where multiple institutions have needs but cannot each justify a full time position; or shared permanent print collections of record distributed across multiple libraries; or consortial buying of e-books.

Summary

The need for different systems to connect and share data with each other in real time and without human intervention is felt across both the business processing and library services spheres of activity. Machine-to-machine integration and data sharing should be the norm, not the exception; any new systems should be designed to be part of an integrated campus network, and not silos of information based around a single work task.



Appendix I: Scenarios

Faculty Scenario #1: Deep and Rapid Discovery

Lucy is teaching a new class, "Emerson's America," and she is looking for a good hook to engage her students. She goes to the Library website and enters "Emerson" in the Library's discovery tool. She gets a lot of results, but she sees a facet that draws her attention to Emerson materials in the College Archives – that is something she might be able to use. A few clicks later into digitized materials, she reads that Emerson was twice invited to Dartmouth to speak: the first time, by the Social Friends just two weeks after his controversial "American Scholar" essay is published (a text her students will be reading).

From there, Lucy is able to scan the horizon of related material quickly: auto-didactic literature from the period; an early book on adult education published at Columbia Teachers' College; articles from a Dartmouth student publication about the visit; the finding aid to the Records of the Social Friends; a recent article on Emerson's lecture circuit; coverage from a local newspaper housed in the Storage Library; a diary from a minister at UCC at Dartmouth; and a Harvard faculty member's contemporary response to "American Scholar" in the *Boston Herald*. She is quickly able to develop an assignment based on the collections; she drag-and-drops links to all of the relevant material on to her course website. She also creates an alert that will email her early next year a reminder to return to the topic while she is on sabbatical to develop this into an essay.

Faculty Scenario #2: Highlighting and Managing Dartmouth Research

Sam is thrilled to hear that there is a database of Dartmouth faculty profiles that highlights their expertise and publications. She and her library liaison sit down to set up her profile; they import links to her publications via a new research management tool, add in her talks and potential collaborators, both local and remote, and import her photo from Flickr. Sam's profile captures her preferences for what she keeps private or makes public, her interests in sharing, tagging and reviewing content, and indicates when she will receive targeted research updates from the Library. Automated content feeds are created based on research queries she wishes to follow – the feeds will populate her profile with recent published research in her areas of interest. All format types are represented: journal articles, conference papers, books, preprints, white papers, etc. A video chat channel to her library liaison is available directly from her profile showing when the liaison is online and ready to help. Best of all, now that Sam has a faculty profile established she notices that after logging in to any library application – discovery and borrowing services, Web spaces, etc. – the environment is customized to her preferences and is aware of her research areas and department affiliations.

Staff Scenario #1: Business Intelligence for Space Planning and Relocation

Shelving on Level 5 is tight and George needs to de-accession some items and transfer others to the offsite depository library. Standing in the stacks, he calls up all the items on a set of shelves, and sees instantly what is checked out. Knowing what is checked out tells him what he needs to allocate sufficient shelf space for on-site content, while also marking materials for de-accession and transfer. He knows he will have to transfer some titles to the depository and he is able to run a report of titles that have never circulated and then mark the records for processing at a later date.

George checks the available shelving in the depository in real time as he makes these determinations. He can also analyze the usage patterns for this shelving range, by examining circulation records over time, and can see the annual growth rate of the shelving section.

Staff Scenario #2: Metadata for Manuscripts

Many people are part of the chain of events that allow users to discover manuscripts in Special Collections. The Manuscript Supervisor and the Processing Specialist analyze the collection and input data into the library metadata management system. The system auto-checks the metadata to ensure conformance with international standards and local practices. Transformation programs convert the format of the metadata to XML for publication, and automatically upload it to the ILS for use. Metadata specialists collaborate with Special Collections staff to maintain these metadata over time. All library staff members with responsibilities for the management of manuscript metadata can view, edit, and save metadata in the library metadata management system. Changes to metadata for manuscripts occur in a central repository, and propagate to all discovery services.

Staff Scenario #3: Working with Print and Digital Analytics

Mario, the Japanese Studies bibliographer, is reviewing his budget, to serve a new faculty member with a specialty in kabuki dancing. Mario quickly runs a report that combines user analytics of the print and electronic items in his area of interest. From this he discovers that our only book on Japanese theatrical makeup has been heavily used, so he adjusts his selection profile to direct more titles in this area. Reviewing the user analytics he also determines that a costly electronic journal has not been used since the retirement of Professor Huzar five years ago. He marks the record and will later run a report of journal titles for potential cancellation that he will discuss with his faculty.

Student Scenario #1: Research Coaching

Rebecca checks her phone – it is 8:30pm, so just enough time to get some work done on her assignment before heading off to a sorority planning meeting. She takes a look at the

assignment instructions on the course web site, and clicks the "Hey, I need a Research Coach!" button that appears on every course site. A scheduling calendar and a chat window open up that redirects her to an RWiT tutor, who steers Rebecca to the course research guide and suggests a couple of starting points listed there. The calendar also tells her when the librarian liaison for the course is available tomorrow. The next day, even before the scheduled meeting, Rebecca is contacted by the librarian who asks whether she had any questions from the previous evening, and offers her some further research sources.

Rebecca can't wait to tell her roommate, who was the one who gave her this great tip. Wow! She was right about that research coach!

Student Scenario #2: Connected Field Research

On an autumn day in 2015, Joe is conducting field research for his Environmental Studies senior thesis. He is revisiting a site near the Second College Grant where he took a water sample that contained high levels of arsenic, which puzzled him: the usual culprits – abandoned mines and factories – are not in the area.

He pulls out his phone and connects to the Library's discovery tool and gets results from a vast pool of information in many formats, and is prompted to try related searches on arsenic mortality and geographic information. First, Joe skims through field notes from a Cornell and Dartmouth faculty team that worked in the area 20 years ago, but finds only that they too identified arsenic in the water. He then taps into an NIH-funded research study on arsenic-related illness in New England. Cases have been reported nearby. He also sees that the Library digitized a local physician's nineteenth-century logbook, and he scans it and finds that there are symptoms listed that could be from arsenic, but it is almost impossible to tell from the evidence. Still, he now suspects that this problem could be one that has been here for a long time. He calls up a Sanborn map from 1910 and checks his location: nothing listed but the owner's name. He also finds that there is a 1788 surveyor's map in the Library, but it is not digitized. He makes a note to himself to check it when he is back on campus. He does a new search on the family name listed on the 1910 Sanborn map. Census information indicates they have been in the area for a long time. He pauses in his work to look at the spectacular foliage and notices a series of slight depressions nearby. He moves closer: it is they look kind of like small cellar holes. He notes the location.

A week later Joe is back in Hanover in Special Collections looking over the surveyor's map. It indicates the original owner's name and finds that the family trade is listed as "Tanner." Back to the map, the depression is right there, right on their land. He has found it: a late eighteenth-century tannery. The depressions he found were not cellar holes at all, but a tanner's pits that leached arsenic into the ground water for nearly a century.

Appendix II: Metadata Management

DISTRIBUTED LIBRARY METADATA (CURRENT)					*METADATA MANAGEMENT (FUTURE)*
ILS	**REPOSITORIES**	**ARCHIVES/RECORDS**	**RESOURCE SHARING**	**DISCOVERY ENVIRONMENTS**	***NEXT GEN SYSTEM***
Catalog of library holdings	digital library collections: images, text, metadata	Archives accessions	Illiad interface	Summon	*Unified metadata management enviornment*
Financial management/procurement	digital library publications: images, text, metadata	Archives metadata	Borrow Direct/Relais	Research Guides	*Integration with campus HR/personnel databases*
Commercial product management	digital library exhibits: images, text, metadata	Archives data management		Library Web	*Integration with campus financial systems*
License repository	manuscript finding aids	College Records Management		Get It (OpenURL resolver)	*Integraton of locally-produced and commercial library information resources*
Physical collection management				Borrow Direct	*Collection analysis tools*
Patron identification and accounts				DartDoc	*Ability to share collection data with partner libraries*
Circulation				Library Catalog "my account"	*Support for license storage and communication about license restrictions to users*
Reserves					*Integrated crosswalking and reformatting tools*