

Metadata to Support Next-Generation Library Resource Discovery: Lessons from the eXtensible Catalog, Phase 1

Jennifer Bowen

The eXtensible Catalog (XC) Project at the University of Rochester will design and develop a set of open-source applications to provide libraries with an alternative way to reveal their collections to library users. The goals and functional requirements developed for XC reveal generalizable needs for metadata to support a next-generation discovery system. The strategies that the XC Project Team and XC Partner Institutions will use to address these issues can contribute to an agenda for attention and action within the library community to ensure that library metadata will continue to support online resource discovery in the future.

Library metadata, whether in the form of MARC 21 catalog records or in a variety of newer metadata schemas, has served its purpose for library users by facilitating their discovery of library resources within online library catalogs (OPACS), digital libraries, and institutional repositories. However, libraries now face the challenge of making this wealth of legacy catalog data function adequately within next-generation Web discovery environments. Approaching this challenge will require:

- an understanding of the metadata itself and a commitment to deriving as much value from it as possible;
- a vision for the capabilities of future technology;
- an understanding of the needs of current (and, where possible, future) library users; and
- a commitment to ensuring that lessons learned in this area inform the development of both future library systems and future metadata standards.

The University of Rochester's eXtensible Catalog (XC) Project will bring these various perspectives together to design and develop a set of open-source, collaboratively built next-generation discovery tools for libraries. The XC Project Team seeks to make the best possible use of legacy library metadata, while also informing the future development of discovery metadata for libraries. During Phase 1 of the XC Project (2006–2007), the XC Project Team created a plan for developing XC and defined the goals and initial functional requirements for the system. This paper outlines the major metadata-

related issues that the XC Project Team and XC Partner Institutions will need to address to build the XC system during Phase 2. It also describes how the XC Team and XC Partners will address these issues, and concludes by presenting a number of issues for the broader library community to consider.

While this paper focuses on the work of a single library project, the goals and functional requirements developed for the XC Project reveal many generalizable needs for metadata to support a next-generation discovery system.¹ The metadata-related goals of the XC Project—to facilitate the use of MARC metadata outside an Integrated Library System (ILS), to combine MARC metadata with metadata from other sources in a single discovery environment, and to facilitate new functionality (e.g., faceted browsing, user tagging)—are very similar to the goals of other library projects and commercial vendor discovery software. The issues described in this paper thus transcend their connection to the XC Project and can be considered general needs for library discovery metadata in the near future.

In addition to informing the library community about the XC Project and encouraging comment on that work, the author hopes that identifying and describing metadata issues that are important for XC—and that are likely to be important for other projects as well—will encourage the library community to set these issues as high priorities for attention and action within the next few years.

The eXtensible Catalog Project

The University of Rochester's vision for the eXtensible Catalog (XC) is to design and develop a set of open-source applications that provide libraries with an alternative way to reveal their collections to library users. XC will provide easy access to all resources (both digital and physical collections) and will enable library content to be revealed through other Web applications that libraries may already be using. XC will be released as open-source software, so it will be available for free download, and libraries will be able to adopt, customize, and extend the software to meet their local needs. The XC Project is a collaborative effort between partner institutions that will serve a variety of roles in its development.

Phase 1 of the XC Project, funded by the Andrew W. Mellon Foundation and carried out by the University of Rochester River Campus Libraries between April 2006 and June 2007, resulted in the creation of a project plan for the development of XC. During XC Phase 1, the XC Project Team recruited a number of other institutions that will serve as XC Partners and who have agreed to contribute resources toward building and implementing XC during Phase 2. XC Phase 2 (October 2007 through

Jennifer Bowen (jbowen@library.rochester.edu) is Director of Metadata Management at the University of Rochester River Campus Libraries, New York, and is Co-Principal Investigator for the eXtensible Catalog Project.

June 2009) is supported through additional funding from the Andrew W. Mellon Foundation, the University of Rochester, and XC Partners. During Phase 2, the XC Project Team, assisted by XC Partners, will deploy the XC software and make it available as open-source software.²

Through its various components, the XC system will provide a platform for local development and experimentation that will ultimately allow libraries to manage and reveal their metadata through a variety of Web applications such as Web sites, institutional repositories, and content management systems. A library may choose to create its own customized local interface to XC, or use XC's native user interface "as is." The native XC interface will include Web 2.0 functionality, such as tagging and faceted browsing of search results that will be informed by FRBR (Functional Requirements for Bibliographic Records)³ and FRAD (Functional Requirements for Authority Data)⁴ conceptual models. The XC software will handle multiple metadata schemas, such as MARC 21⁵ and Dublin Core,⁶ and will be able to serve as a repository for both existing and future library metadata. In addition, XC will facilitate the creation and incorporation of user-created metadata, enabling such metadata to be enhanced, augmented, and redistributed in a variety of ways.

The XC Project Team has designed a modular architecture for XC, as shown in the simplified schematic in figure 1. XC will bring together metadata from a variety of sources (integrated library systems, digital repositories, etc.), apply services to that metadata, and display it in a usable way in the Web environments where users expect to find it.⁷ XC's architecture will allow institutions that implement the software to take advantage of innovative models for shared metadata services, which will be described in this paper.

XC Phase 1 activities

During the now-completed XC Phase 1, the XC Project Team focused on six areas of activity:

1. Survey and understand existing research on user practices.
2. Gauge library demand for the XC system.
3. Anticipate and prepare for the metadata requirements of the new system.
4. Learn about and build on related projects.
5. Experiment with and incorporate useful, freely available code.
6. Build a community of interest.

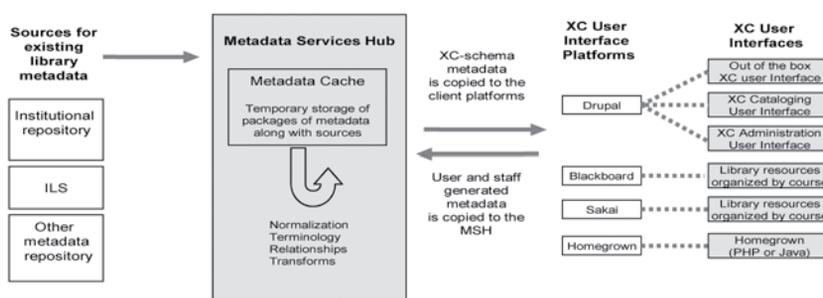


Figure 1. XC System Diagram

The XC Project Team carried out a variety of research activities to inform the overall goals and high-level functional requirements for XC. This research included a literature search and ongoing monitoring of discussion lists and blogs, to allow the team to keep up with the most current discussions taking place about next-generation library discovery systems and related technologies and projects.⁸ The XC team also consulted regularly with prospective partners and other knowledgeable colleagues who are engaged in defining the concept of a next-generation library discovery system. In order to gauge library demand for the XC system, the team also conducted a survey of interested institutions.⁹

This paper reports the results of the third area of activity during XC Phase 1—anticipating and preparing for the metadata requirements of the new system—and looks ahead to plans to develop the XC software during Phase 2.

XC goals and metadata functional requirements

The goals of the XC Project have significant implications for the metadata functionality of the system, with each goal suggesting specific high-level functional requirements for how the system can achieve that particular goal. The five goals are:

- Goal 1: Provide access to all library resources, digital and non-digital.
- Goal 2: Bring metadata about library resources into a more open Web environment.
- Goal 3: Provide an interface with new Web functionality such as Web 2.0 features and faceted browsing.
- Goal 4: Conduct user research to inform system development.
- Goal 5: Publish the XC code as open-source software.

An overview of each XC goal and its related high-level metadata requirements appears below. Each requirement is then discussed in more detail, with a plan for how the XC Project Team will address that requirement when developing the XC software.

Goal 1: Provide access to all library resources, digital and non-digital

Working alongside a library's current Integrated Library System (ILS) and its other Web applications, XC will strive to bring together access to all library resources, thus eliminating the data silos that are now likely to exist between a library's OPAC and its various digital repositories and commercial databases. This goal suggests two fairly obvious metadata requirements (Requirements 1 and 2).

Requirement 1—The system must be capable of acquiring and managing metadata from multiple sources: ILSs, digital repositories, licensed databases, etc.

A typical library currently has metadata pertaining to its collections residing in a variety of separate online systems: MARC data in an ILS, metadata in various schemas in digital collections and repositories, citation data in commercial databases, and other content on library Web sites. A library that implements XC may want to populate the system with metadata from several online environments to simplify access to all types of resources. To achieve Goal 1, XC must be capable of acquiring and managing metadata from all of these sources. Each online environment and type of metadata present their own challenges.

Repurposing MARC data

Repurposing MARC metadata from an existing ILS will be one of the biggest metadata tasks for a next-generation discovery system such as XC. In planning XC, we have assumed that most libraries will keep their current ILS for the next few years or perhaps migrate to a newer commercial or open-source ILS. In either case, most libraries will likely continue to rely on an ILS's staff functionality to handle materials acquisition, cataloging, circulation, etc. for the short term. Relying upon an ILS as a processing environment does not, however, mean that a library must use the OPAC portion of that ILS as its means of resource discovery for users. XC will provide other options for resource retrieval by using Web services to interact with the ILS in the background.¹⁰ To repurpose ILS metadata and enable it to be used in various Web discovery environments, XC will harvest a copy of MARC metadata records from an institution's ILS using the Open Archives

Initiative Protocol for Metadata Harvesting (OAI-PMH).¹¹ Using Web services and standard protocols such as OAI-PMH offers not only a short-term solution for reusing metadata from an ILS, but can also be used in both the short- and long-term to harvest metadata from any system that is OAI-PMH harvestable, as will be discussed further below.

While harvesting metadata from existing systems into XC creates duplication of metadata between an ILS and XC, this actually has significant benefits. XC will handle metadata updates through automated harvesting services that minimize additional work for library staff, other than for setting up and managing the automated services themselves. The internal XC metadata cache can be easily regenerated from the original repositories and services when necessary, such as to enable future changes to the internal XC metadata schema. The XC system architecture also makes use of internal metadata duplication among XC's components, which allows these components to communicate with each other using OAI-PMH. This built-in metadata redundancy will also enable XC to communicate with external services using this standard protocol.

It is important to distinguish the deliberate metadata redundancies built into the XC architecture from the type of metadata redundancies that have been singled out for elimination in the Library of Congress Working Group on the Future of Bibliographic Control draft report (Recommendation 1.1)¹² and previously in the University of California (UC) Libraries Bibliographic Services Task Force's final report.¹³ These other "negative" redundancies result from difficulties in sharing metadata among different environments and cause significant additional staff expense for libraries to enrich or recreate metadata locally. XC's architecture actually solves many of these problems by facilitating the sharing of enriched metadata among XC users. XC can also adapt as the library community begins to address the types of costly metadata redundancies mentioned in the above reports, such as between the OCLC WorldCat database¹⁴ and copies of that MARC data contained within a library's ILS, because XC will be capable of harvesting metadata from any source that uses a standard API.¹⁵

Metadata from digital repositories and other free sources

XC will harvest metadata from various digital collections and repositories, using OAI-PMH, and will maintain a copy of the harvested metadata within the XC metadata cache, as shown in figure 1. The metadata services hub architecture provides flexibility and possible economy for XC users by offering the option for multiple XC institutions to share a single metadata hub, thus allowing participating institutions to take full advantage of the hub's capabilities to aggregate and augment metadata from multiple sources. While the procedure for harvest-

ing metadata from an external repository is not technologically difficult in itself, managing the flow of metadata coming from multiple sources and aggregating that metadata for use in XC will require the development of sophisticated software. To address this, the XC Project Team is partnering with established experts in bibliographic metadata aggregation to develop the metadata services portion of the XC architecture. The team from Cornell University that has developed the software behind the National Science Digital Library's Metadata Management System (NSDL/MMS)¹⁶ is advising the XC team in the development of the XC metadata services hub, which will be built on top of the basic NSDL/MMS software.

The XC metadata services hub will coordinate metadata services into a reusable task grouping that can be started on demand or scheduled to run regularly. This XC component will harvest XML metadata and combine metadata records that refer to equivalent resources (based on Uniform Resource Identifier [URI], if available, or other unique identifier) into what the Cornell team describes as a "mudball." Each mudball will contain the original metadata, the sources for the metadata, and the references to any services used to combine metadata into the mudball. The mudball may also contain metadata that is the result of further automated processing or services to improve quality or to explicitly identify relationships between resources. Hub services could potentially record the source of each individual metadata statement within each mudball, which would then allow a metadata record to be redelivered in its original or in an enriched form when requested.¹⁷ By allowing for the capture of provenance data for each data element, the hub could potentially provide much more granular information about the origin of metadata—and much more flexibility for recombining metadata—than is possible in most MARC-based environments.

After using the redeployed NSDL/MMS software as the foundation for the XC metadata hub, the XC Project Team will develop additional hub services to support XC's functional requirements. XC-specific hub services will accommodate incoming MARC data (including MARC holdings data for non-digital resources); basic authority control; mappings from MARC 21, MARCXML,¹⁸ and Dublin Core to an internal XC schema defined within the XC Application Profile (described below); and other services to facilitate the functionality of the XC user environments (see discussion of Requirement 5, below). Finally, the XC hub services will make the metadata available for harvesting from the hub by the XC client integration applications.

Metadata for licensed content

For a next-generation discovery system such as XC to provide access to all library resources, it will need to provide access to licensed content, such as citation data and

full-text databases. Metasearch technology provides one option for incorporating access to licensed content into XC. Unfortunately, various difficulties with metasearch technology¹⁹ and usability issues with some metasearch products²⁰ make metasearch technology a less-than-ideal solution. An alternative approach would bring metadata from licensed content directly into a system such as XC. The metadata services hub architecture for XC is capable of handling the ingest and processing of metadata supplied by commercial content providers by adding additional services to handle the necessary schema transformations and to control access to the licensed content. The more difficult issue with licensed content may be to obtain the cooperation of commercial vendors to ingest their metadata into XC. Pursuing individual agreements with vendors to negotiate rights to ingest their metadata is beyond the original scope of XC's Phase 2 Project. However, the XC team will continue to monitor ongoing developments in this area, especially the work of the EthicShare Project, which uses a system architecture very similar to that of XC.²¹ It remains our goal to build a system that will facilitate the inclusion of licensed content within XC in situations where commercial providers have made it available to XC users.

Requirement 1 summary

When considering needed functionality for a next-generation discovery system, the ability to ingest and manage metadata from a variety of sources is of paramount importance. Unlike a current ILS, where we often think of metadata as mostly static unless it is supplemented by new, updated, and deleted records, we should instead envision the metadata in a next-generation system as being in constant motion, moving from one environment to another and being harvested and transformed on a scheduled basis. The metadata services hub architecture of the XC system will accommodate and facilitate such constant movement of metadata.

Requirement 2—The system must handle multiple metadata schemas.

An extension of Requirement 1 will be the necessity for a next-generation system such as XC to handle metadata from multiple schemas, as the system harvests those schemas from various sources.

Library metadata priorities

As a part of the XC survey of libraries described earlier in this paper, the XT Team queried respondents about what metadata schemas they currently use or plan to use in the near future. Many responding libraries indicated that they expect to increase their use of non-MARC 21 metadata within the next three years, although no library indicated the intention to completely move away from

MARC 21 within that time period. Nevertheless, the idea of a “MARC exit strategy” has been discussed in various circles.²² The architecture of XC will enable libraries to move beyond the constraints of a MARC-based system without abandoning their ILS, and will provide an opportunity for libraries to stage their “MARC exit strategy” in a way that suits their purposes.

Libraries also indicated that they plan to move away from homegrown schemas toward accepted standards such as METS,²³ MODS,²⁴ MADS,²⁵ PREMIS,²⁶ EAD,²⁷ VRA Core,²⁸ and Dublin Core.²⁹ Several responding libraries plan to move toward a wider variety of metadata schemas in the near future, and will focus on using XML-based schemas to facilitate interoperability and metadata harvesting. To address the needs of these libraries in the future, XC’s metadata services will contain a variety of transformation services to handle a variety of schemas. Taking into account the metadata schemas mentioned the most often among survey respondents, the software developed during Phase 2 of the XC Project will support harvested metadata in MARC 21, MARCXML, and Dublin Core (including Qualified Dublin Core).³⁰

Metadata crosswalks and mapping

One respondent to the XC Survey offered the prediction that “reuse of existing metadata and transformation of metadata from one format to another will become commonplace and routine.”³¹ XC’s internal metadata transformations must be designed with this in mind, to facilitate making these activities “commonplace and routine.” Fortunately, many maps and crosswalks already exist that potentially can be incorporated into a next-generation system such as XC.³² The metadata services hub architecture for XC can function as a standard framework for applying a variety of existing crosswalks within a single, shared environment. Following “best practices” for crosswalking metadata, such as those developed by the Digital Library Federation (DLF),³³ will be extremely important in this environment. As the DLF guidelines describe, metadata schema transformation is not as straightforward as it might first appear to be. While the DLF guidelines advise always crosswalking from a more robust schema to a simpler one, sometimes in a series of steps, such mapping will often result in “dumbing down” of metadata, or loss of granularity. This is a particularly important concern for the XC Project because a large percentage of the metadata handled by XC will be rich legacy MARC 21 metadata, and we hope to maintain as much of that richness as possible within the XC system.

In addition to simply mapping one data element in a schema to its closest equivalent in another, it is essential to ensure that the underlying metadata models of the two schemas being crosswalked are compatible. The authors of the Framework for a Bibliographic Future

draft document define multiple layers of such models that need to be considered,³⁴ and offer a general high-level comparison between the FRBR data model³⁵ and the DCMI (Dublin Core Metadata Initiative) Abstract Model (DCAM).³⁶ More detailed comparisons of models are also taking place as a part of the development of the new metadata content standard, Resource Description and Access (RDA).³⁷ The developers of RDA have issued documents offering a detailed mapping of RDA elements to RDA’s underlying model (FRBR)³⁸ and analyzing the relationship between RDA elements, the DCMI Abstract Model, and the Metadata Framework.³⁹

As a result of a meeting held April 30–May 1, 2007, a joint DCMI/RDA Task Group is now undertaking the collaborative work necessary to carry out the following tasks:

- Develop an RDA Element Vocabulary.
- Develop an RDA/Dublin Core Application Profile based on FRBR and FRAD.
- Disclose RDA Value Vocabularies using RDF/RDFS/SKOS.⁴⁰

These efforts hold much potential to provide a more rigorous way to communicate about metadata across multiple communities and to increase the compatibility of different metadata schemas and their underlying models. Such compatibility will be essential to enabling the functionality of future discovery systems such as XC.

An XC metadata application profile

The XC Project Team will define a metadata application profile for XC as a way to document decisions made about data elements, content standards, and crosswalking used within the system. The use of an application profile can facilitate metadata migration, harvesting, and other automated processes, and presents an approach to metadata that is more flexible and responsive to local needs than simply adopting someone else’s metadata guidelines.⁴¹ Application profiles facilitate the use of multiple schemas because elements can be selected for inclusion from more than one existing schema, or additional elements can be created and defined locally.⁴² Because the XC system will incorporate harvested metadata from a variety of sources, the use of an application profile will be essential to support XC’s complex system requirements.

The DCMI Community has published guidelines for creating a Dublin Core Application Profile (DCAP), which is defined more specifically as:

[a] form for documenting which terms a given application uses in its metadata, with what extensions or adaptations, and specifying how those terms relate both to formal standards such as Dublin Core as well as to less formally defined element sets and vocabularies.⁴³

The announcement of plans to develop an RDA/Dublin Core Application Profile illustrates the important role that application profiles are beginning to take to facilitate the interoperability of metadata schemas. The planned RDA/DC Application Profile will “translate” RDA into a standard structure that will allow it to be related more easily to other metadata element sets. Unfortunately, the RDA/DC Application Profile will likely not be completed in time for it to be incorporated into the first release of the XC software in mid-2009. Nevertheless, we intend to use the existing definitions of RDA elements to inform the development of the XC Application Profile.⁴⁴ This will allow us to anticipate any future incompatibilities between the RDA/DC and the XC application profiles, and ensure that XC will be well-positioned to take advantage of RDA-based metadata when RDA is implemented. This process may have the reciprocal benefit of also informing the developers of RDA of any RDA elements that may be difficult to implement within a next-generation system such as XC.

The potential value of RDA to the XC project—in terms of providing a consistent approach to bibliographic and authority metadata and facilitating FRBR-related user functionality—is very significant. It is hoped that at some point XC can become an early adopter of RDA and provide a mechanism through which libraries can move their legacy MARC 21 metadata into a system that is compatible with an emerging international metadata standard.

Goal 2: Bring metadata about library resources into a more open Web environment

XC will reveal library metadata not only through its own separate interface (either the out-of-the-box XC interface or an interface designed by the local library), but will also allow library metadata to be revealed through other Web applications. The latter approach will bring library resources directly to Web locations that library users are already visiting, rather than attempting to entice users to visit an additional library-specific Web location. Making library metadata work effectively in the broader Web environment (outside the well-defined boundaries of an ILS or repository) will require the following Requirements 3 and 4:

Requirement 3—Metadata must conform to the standards of the new Web environments as well as to that of the system from which it originated.

Achieving Requirement 3 will require library metadata in future systems to perform a dual function: to conform to both existing library standards as well as to Web

standards and conventions. One way to achieve this is to ensure that the two types of standards themselves are compatible. Coyle and Hillmann have argued persuasively for changes in the direction of RDA development to allow metadata created using RDA to function in the broader Web environment. These changes include the need to follow a clearly refined, high-level metadata model, to create data elements that can be manipulated by machines, and to move toward the use of URIs instead of textual identifiers.⁴⁵ After the announcement of the outcomes of the RDA/DC Data Modeling meeting, the two authors are considerably more optimistic about RDA functioning as a standard within the broader Web environment.⁴⁶ This discourse concerning RDA shows but a piece of the process through which long-established library metadata standards need to be reexamined to make library metadata understandable to both humans and machines on the Web. Moving away from AACR2 toward RDA, and ultimately toward incorporating standard Web conventions into library metadata, can be a difficult process for those involved in creating and maintaining library standards. Nevertheless, transforming library metadata standards in this way is essential to fulfill the requirements necessary for next-generation library discovery systems.

Requirement 4—Metadata must function effectively within the new Web environments as well as within the system from which it originated.

Not only must metadata for a next-generation system follow the conventions and standards used in the broader Web, but the data also needs to be able to function effectively in a broader Web environment. This is a slightly different proposition from Requirement 3, and will necessitate testing the metadata standards themselves to ensure that they enable library metadata to function effectively.

The XC Project will provide direct experience with using library metadata in two types of Web environments: content management systems and learning management systems.

Library metadata in a content management system

As shown in the XC architecture diagram in figure 1, the XC Project Team will build one of the primary user environments for XC on top of the open-source content management system, Drupal.⁴⁷ The XC Drupal module will allow us to respond to many of the needs expressed by libraries in their responses to the XC survey⁴⁸ by supplying:

- a Web application server with a back-end database;

- a user interface with Web 2.0 features;
- library-controlled Web pages that will treat library metadata as a native data type;
- a metadata interface for enhancing or correcting metadata in the system; and
- an administrative interface.

The XC Team will bring library metadata into the Drupal content management system (CMS) as a native content type within that environment, creating a Drupal “node” for each metadata record. This will allow XC to take advantage of many native features of the Drupal CMS, such as a taxonomy system.⁴⁹ Building XC interfaces on top of the Drupal CMS will also give us an opportunity to collaborate with partner libraries that are already active participants in the Drupal user community.

XC’s architecture will allow the possibility of developing additional user environments on top of other content management systems. Bringing library metadata into these new environments will provide many new opportunities for libraries to manipulate their metadata and present it to users without being constrained by the limitations of the current generation of library systems. Such opportunities will then inform the future requirements for library metadata in such environments.

Library metadata in a learning management system

Figure 1 illustrates two examples of XC user environments through learning management systems: XC interfaces to both the Blackboard Learning System⁵⁰ and Sakai.⁵¹ Much exciting work is being done at other institutions to bring library content into these Web applications.⁵² XC will build on projects such as these to reveal library metadata for non-licensed library resources from an ILS through learning management systems. Specifically, we plan to develop the capability for libraries to make the display of library metadata context-sensitive within the learning management system. For example, searching or browsing on a page for a particular academic course could be configured to reflect the subject area of the course (e.g., chemistry) and automatically present library resources related to that subject.⁵³ This capability will build upon the experiences gained by the University of Rochester through its work to develop its “CoURse Resources” system.⁵⁴ Such XC functionality will be integrated directly into the learning management system, rather than simply providing a link out to a separate library system.

Again, we hope that our efforts to bring library metadata into these new environments will encourage libraries to engage in further work to integrate library resources into broader Web environments and inform future requirements for library metadata in these environments.

Goal 3: Provide an interface with new Web functionality such as Web 2.0 features and faceted browsing

New functionality for users will require that metadata fulfill more sophisticated functions in a next-generation system than it may have done in an ILS or repository, in order to provide more intuitive searching and navigation. The system will also need to capture and incorporate metadata generated through tagging, user-contributed reviews, etc. Such new functionality creates the need for Requirements 5 and 6.

Requirement 5—Metadata must support functionality to facilitate intuitive searching and navigation, such as faceted browsing and FRBR-informed results groupings.

Enabling faceting and clustering

Much research has already been done regarding the design of faceted search interfaces in general.⁵⁵ When considered along with user research conducted at other institutions⁵⁶ and to be conducted during the development of XC, this data provides a strong foundation for the design of a faceted browse environment. The XC Project Team has already gained firsthand experience with developing faceted browsing through the development of the “C4” prototype interface during Phase 1 of the XC Project.⁵⁷ To enable faceting within XC, we will also pay particular attention to what others have discovered through designing faceted interfaces on top of legacy MARC 21 metadata. Specific lessons learned from those involved with North Carolina State University’s Endeca-based catalog,⁵⁸ Vanderbilt University’s Primo implementation,⁵⁹ and Plymouth State University’s Scriblio system⁶⁰ provide valuable guidance for the XC Project Team as we design facets for the XC system. Ideally, a mechanism should be developed to enable these discoveries to feed back into the development of metadata and encoding standards, so that changes to existing standards can be considered to facilitate faceting in the future.

Several new system implementations have used Library of Congress Subject Headings (LCSH) and LC subdivisions from MARC 21 records as the basis for deriving facets. The XC “C4” prototype interface provides facets for topic, genre, and region that are based simply upon one or more MARC 21 6XX tags.⁶¹ North Carolina State University’s Endeca-based system has enabled facets for topic, genre, region, and era using LCSH subdivisions as well, but this has necessitated a “massive cleanup” of subdivisions, as described by Charley Pennell.⁶² OCLC’s FAST (Faceted Application of Subject Terminology) project may provide another option for enabling such facets.⁶³ A library could populate its MARC 21 data with FAST headings, based

upon the existing LCSH in the records, and then use the FAST headings as the basis for generating facets. It remains to be seen whether FAST will offer significant benefit over LCSH itself when it comes to faceting, however, since FAST headings are generated directly from LCSH.

While MARC 21 metadata has some known difficulties where faceting and clustering are concerned (such as those involving LCSH), the XC system will encounter additional difficulties when implementing these technologies with less robust metadata schemas such as simple Dublin Core, and especially across metadata from a variety of schemas. The development of Web services to augment batches of metadata records in an automated manner holds some promise for improving the creation of facets from other metadata schemas. Within the XC system, such services could be added to the metadata services hub and run against ingested metadata. While designing extensive services of this type is beyond the scope of the next phase of XC software development, we will encourage others to develop such services for XC.

Another (but much less desirable) approach to augmenting metadata is for a metadata specialist to manually edit one record or group of records. The XC cataloging interface, built within the Drupal CMS, will allow record-by-record editing of metadata when necessary. While we see this editing interface as essential functionality for XC, we anticipate that libraries will want to use this feature sparingly. In many cases it will be preferable to correct or augment metadata within its original repository (e.g., the institution's ILS) and then re-harvest the corrected metadata, rather than correcting it manually within XC itself. Because of the expense of manual metadata augmentation and correction, libraries will be well-advised to rely upon insights gained through user research to assess the value of this type of work. For example, a library might decide to edit individual metadata records only when the correction or augmentation will support specific system functionality that is of high priority for the institution's users.

Implementing FRBR results groupings

To incorporate logical groupings of search results based upon the FRBR⁶⁴ and FRAD⁶⁵ data models over sets of diverse metadata within XC, we will encounter similar difficulties that we face with faceting and clustering. Various analyses of the MARC 21 formats have dealt extensively with the relationship between FRBR and MARC 21,⁶⁶ and others have written specifically about methodology for FRBRizing a MARC-based catalog.⁶⁷ In addition, various tools and Web services are available that can potentially facilitate this process.⁶⁸ Even with this extensive body of work to draw upon, however, the success of our implementation of FRBR-based functionality will depend upon both the quality and completeness of the system's metadata. Metadata in XC that originated as Dublin Core records may need significant augmenta-

tion to be incorporated effectively into FRBRized results displays. To maximize the ability of the system to support FRBR/FRAD results groupings, we may need to supplement automated grouping of resources with a combination of additional services for the metadata services hub, and with cataloger-generated metadata correction and augmentation, as described above.⁶⁹ The XC team will use the results of user research carried out during the next phase of the XC Project to inform our decision-making regarding what FRBR-informed results grouping users find helpful, and then assess what specific metadata augmentation services are needed for XC.

Providing FRBR-informed groupings of related records in search results will be easier when the underlying metadata incorporates principles of authority control. Of course, the vast majority of the non-MARC metadata that will be ingested into XC will not be under authority control. Again, this situation suggests the need for additional services or functionality to improve existing metadata within the XC metadata hub, the XC cataloging interface, or both. As an experiment in developing services to facilitate authority control, the XC Project Team carried out a pilot project in partnership with a group of software engineering students from the Rochester Institute of Technology (RIT) during Phase 1 of XC. The RIT students designed a basic name access control tool that can be used across disparate metadata schemas in an environment such as XC. The tool can ingest MARC 21 authority and bibliographic records as well as Dublin Core records, provide automated matching, and facilitate a cataloger's handling of problem reports.⁷⁰ The XC Project Team will implement the automated portion of the tool as a Web service within the XC hub, and the "cataloger facilitation" portion of the tool within the XC cataloging user interface. Institutions that use XC can then incorporate additional tools to facilitate authority control into XC as they are needed and developed.

In addition to providing a test case for developing XC metadata services, the RIT pilot project proved valuable by providing an opportunity for student software developers and catalogers to discuss the functional requirements of a cataloging tool. Not only did the experience enable the developers to understand the needs of the system's intended users, but it also presented an opportunity for the engineering students to demonstrate technological possibilities that the catalogers—who work almost exclusively with legacy ILS technology—may not have envisioned before participating in the project.

Requirement 6—The system must manage user-generated metadata resulting from user tagging, submission of reviews, etc.

Because users now expect Web-based tools to offer Web 2.0 functionalities, the XC Project has as one of its basic

goals to incorporate these functionalities into XC's user environments. The results of the XC Survey rank tools to support the finding, gathering, use, and reuse of scholarly content (e.g., RSS feeds, blogs, tagging, user reviews) eighth out of a list of twenty new desirable OPAC features.⁷¹ We expect to learn much more about the usefulness of Web 2.0 technology within a next-generation system through the user research that we will carry out during Phase 2 of the XC Project.

The XC system will capture metadata generated by users from any one of the system's user environments (e.g., Drupal-based interface, learning management system integration) and harvest it back into the system's metadata services hub for processing.⁷² The XC Application Profile will incorporate user-generated metadata, mapped into its own carefully defined metadata elements. This will allow us to capture and manage this metadata as discrete content, without inadvertently mixing it with other metadata created by library staff or ingested from other sources.

Goal 4: Conduct user research to inform system development

User research will be essential to informing the design and functionality of the XC software. To align XC's functional requirements as closely as possible with user needs, the XC Project Team will practice a user-centered design methodology that takes an iterative approach to defining the system's functional requirements. Since we will engage concurrently in the processes of user research and software design, we will not fully determine the system requirements for XC until a significant amount of user research has been done. A complete picture of the demands upon metadata within XC will thus emerge as we gain information from our user research.

Goal 5: Publish the XC code as open-source software

Central to the vision of the XC Project is sharing the XC software freely throughout the library community and beyond. Our hope is that others will use all or part of the XC software, modify it, and improve it to meet their own needs. New requirements for the metadata within XC are likely to arise as this process takes place. Other future changes to the XC software will also be needed to ensure the software's continued compatibility with various metadata standards and schemas. These changes will all affect the system requirements for XC over time.

Addressing Goals 4 and 5

While Goals 1 through 3 for the XC Project result in

specific high-level functional requirements for the system's discovery metadata that can be addressed and discussed as XC is being developed, Goals 4 and 5 present general challenges that must be addressed in the future. Goal 4 is likely to fuel the need to update the XC software over time as the needs of users change. Goal 5 provides a challenge to managing that updating process in a collaborative environment. These two goals suggest an additional general requirement for the system's metadata Requirement 7:

Requirement 7—The system's metadata must be extensible to facilitate future enhancements and updates.

Enabling future user needs

Developing XC using a user-centered design process in which user research and software design occur simultaneously will enable us to design and build a system that is as responsive as possible to the needs of users that are seeking library resources. However, user needs will change during the life of the XC software. These needs must be assessed and addressed, and then weighed against the desires of individual institutions that use XC and who request specific system enhancements.

To carry forward the XC Project's commitment to serving users, we will develop a governance model for the XC community that brings the needs of future users into the decision-making process by providing a method for continuing to determine and capture user needs. In addition, we will consciously cultivate a commitment to user research among members of the XC community. Because the XC software will be released as open source, we can also encourage XC partners to develop whatever additional functionality they need for their own institutions and make these enhancements available to the entire community of XC users. This approach is very different from the enhancement process in place for most commercial systems, and XC partner institutions may need to adjust to this approach.

Enabling future metadata standards

As current metadata standards are revised and new standards and schemas are created, XC must be able to accommodate these changes. New crosswalks will allow new metadata schemas to be mapped to the XC internal schema in the future. The XC Application Profile can be updated with the addition of new data elements as needed. The Drupal-based XC user environment will also allow institutions that use XC to create new internal data types to incorporate additional types of metadata. As the development of the Semantic Web moves forward⁷³ and enables smart linking between existing authority files and vocabularies,⁷⁴ XC's architecture can make use of the resulting Web services, either by incorporating them

through the XC metadata services hub or through the native XC user interface as part of a user search query.

Further considerations

The above discussion of the goals and requirements for XC has revealed a number of issues related to the development of next-generation discovery systems that are unfortunately beyond the scope of the next phase of the XC Project. We therefore offer them as a possible agenda for future work by the broader library community:

1. Explore the wider usefulness of Web-based metadata services and the need for an automated metadata services coordinator to control these functions. Libraries are already comfortable with basic “services” that are performed on metadata by an outside agency: For example, a library may send copies of its MARC records to a vendor for authority processing or enrichment with tables of contents or other data elements. The library community should encourage vendors and others to develop these and other metadata enrichment options as automated Web services.
2. Study the advantages of using statement-level metadata provenance, as used in the NSDL Metadata Management System and considered for use within the XC metadata services hub, and explore whether there are ways that MARC 21 could move toward allowing more granularity in recording and sharing metadata provenance.
3. To facilitate access to licensed library resources, encourage the development of more robust metasearch technology and standards so that technological limitations do not hinder system performance and search result usability. If this is not successful, libraries and content providers must work together to enable metadata for licensed resources to be revealed within open discovery environments such as XC and EthicShare.⁷⁵ This second scenario will enable libraries to directly address usability issues with the display of licensed content, which may make it a more desirable longer-term solution than attempting to improve metasearch technology.
4. The administrative bodies of the two groups represented on the DCMI/RDA Task Group (i.e., the Dublin Core Metadata Initiative and the RDA Committee of Principals) have a responsibility to take the lead in funding this group’s work to develop and maintain the RDA/DC Application Profile and its related registries and vocabularies. Beyond this, however, the broader library community must recognize that this work is essential to ensure that future library metadata standards will function in the broader Web environment, and offer additional administrative and financial support for it in the coming years.
5. To ensure that library standards work effectively outside of traditional library systems, catalogers and metadata experts must develop ongoing, collaborative working relationships with system developers. Such collaboration will necessitate educating each group of experts about the domain of the other.
6. Libraries should experiment with using metadata in new environments and use the lessons learned from this activity to inform the metadata standards development process. While current library automation environments by and large do not provide opportunities for this, the eXtensible Catalog will provide a flexible platform where experimentation can take place.⁷⁶ XC will make experimentation as risk-free as possible by ensuring that the original metadata brought into the system can be reharvested in its original form, thus minimizing concerns about possible data corruption. XC will also minimize the investment needed for a library to engage in this experimentation because it will be released as open-source software.
7. To facilitate new functionality for next-generation library discovery environments, libraries must share their new expertise in this area with each other. For example, library professional organizations (such as ALA and its associations) should form discussion groups and committees devoted to sharing lessons learned from the implementation of faceted interfaces and Web 2.0 technologies, such as tagging and folksonomies. Such groups should develop a “best practices” document outlining a preferred way to define facets from MARC 21 data that can be used by any library implementing faceting on top of its legacy metadata.
8. The library community should discuss and encourage mechanisms for pooling and sharing user-generated metadata among libraries and other interested institutions.

Conclusions

To present library resources via the Web in a manner that users now expect, library metadata must function in ways that have never been required of it before. Making library metadata function effectively within the broader Web environment will require that libraries take advantage of the combined knowledge of experts in the areas of cataloging/metadata and system development who share a

common vision for serving library users. The challenges to making legacy library metadata and newer metadata for digital resources interact effectively in the broader Web environment are significant, and work must begin now to ensure that we can preserve the investment that libraries have made in their legacy metadata.

While the recommendations within this report are the result of planning to develop one particular library discovery system—the eXtensible Catalog (XC)—these lessons can inform the development of other systems as well. The actual development of XC will continue to add to our knowledge in this area. While it may be tempting to wait and see what commercial vendors offer as their next generation of commercial discovery products, such a passive approach may jeopardize the future viability of library metadata. Projects such as the eXtensible Catalog can serve as a vehicle for moving forward by providing an opportunity for libraries to experiment and to then take informed action to move the library community toward a next generation of resource discovery systems.

Acknowledgments

Phase 1 of the eXtensible Catalog Project was funded through a grant from the Andrew W. Mellon Foundation. This paper is in partial fulfillment of that grant, originally funded on April 1, 2006, and concluding on June 30, 2007. The author acknowledges the contributions of the entire University of Rochester eXtensible Catalog Project Team to the content of this paper, and especially thanks David Lindahl, Barbara Tillett, and Konstantin Gurevich for reading and offering suggestions on drafts of this paper.

References and notes

1. Despite the use of the word “catalog” within the name of the eXtensible Catalog Project, this paper will avoid using the word “catalog” in the phrase “next-generation catalog” because this may misleadingly convey the idea of a catalog as solely a single, separate Web destination for library users. Instead, terms such as “discovery environment” and “discovery system” will be preferred.

2. The XC blog provides a list of XC Partners, describes their roles in XC Phase 2, and provides links to reports that represent the outcomes of XC Phase 1. “XC (eXtensible Catalog): An Open-Source Online System That Will Unify Access to Traditional and Digital Library Resources,” www.extensiblecatalog.info (accessed October 4, 2007).

3. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records* (Munich: K. G. Saur, 1998), www.ifla.org/VII/s13/frbr/frbr.pdf (accessed July 23, 2007).

4. IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), “Functional

Requirements for Authority Data: A Conceptual Model,” April 1, 2007, www.ifla.org/VII/d4/FRANAR-ConceptualModel-2ndReview.pdf (accessed July 23, 2007).

5. Library of Congress, Network Development and MARC Standards Office, “MARC 21 Formats,” April 18, 2005, www.loc.gov/marc/marcdocz.html (accessed September 3, 2007).

6. “Dublin Core Metadata Element Set, Version 1.1,” December 20, 2004, <http://dublincore.org/documents/dces> (accessed September 3, 2007).

7. University of Rochester River Campus Libraries, “eXtensible Catalog Phase 2,” (grant proposal submitted to the Andrew W. Mellon Foundation, July 11, 2007).

8. “Literature List,” eXtensible Catalog blog, www.extensiblecatalog.info/?page_id=17 (accessed August 27, 2007).

9. A summary of the results of this survey is available on the XC blog. Nancy Fried Foster et al., “eXtensible Catalog Survey Report,” July 20, 2007, www.extensiblecatalog.info/wp-content/uploads/2007/07/XC%20survey%20report.pdf (accessed July 23, 2007).

10. Lorcan Dempsey has written of the need for a service layer for libraries that would facilitate the “de-coupling” of resource retrieval from back-end processing. Lorcan Dempsey, “A Palindromic ILS Service Layer,” *Lorcan Dempsey’s Weblog*, January 20, 2006, <http://orweblog.oclc.org/archives/000927.html> (accessed August 24, 2007).

11. “Open Archives Initiative Protocol for Metadata Harvesting v. 2.0,” www.openarchives.org/OAI/openarchivesprotocol.html (accessed August 27, 2007).

12. Library of Congress, Working Group on the Future of Bibliographic Control, “Report on the Future of Bibliographic Control: Draft for Public Comment,” November 30, 2007, www.loc.gov/bibliographic-future/news/lcwg-report-draft-11-30-07-final.pdf (accessed December 30, 2007).

13. University of California Libraries Bibliographic Services Task Force, “Rethinking How We Provide Bibliographic Services for the University of California,” *Final report*, 34, <http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf> (accessed August 24, 2007).

14. “[WorldCat.org] Search for an Item in Libraries Near You,” www.worldcat.org (accessed August 24, 2007).

15. OCLC’s plan to create additional APIs to WorldCat as part of its WorldCat Grid project is a welcome development that may enable OCLC members to harvest metadata directly from WorldCat into a system such as XC in the future. See the following blog posting for an early description of OCLC’s plans, which have not been formally unveiled by OCLC as of this writing: Bess Sadler, “The Librarians and the Chocolate Factory: OCLC Developer Network Day,” *Solvitur ambulando*, October 3, 2007, www.ibiblio.org/bess/?p=88 (accessed December 30, 2007).

16. “Metadata Management System,” *NSDL Registry*, September 20, 2006, http://metadatarregistry.org/wiki/index.php/Metadata_Management_System (accessed July 23, 2007).

17. Diane Hillmann, Stuart Sutton, and Jon Phipps, “NSDL Metadata Improvement and Augmentation Services,” (grant proposal submitted to the National Science Foundation, 2007).

18. Library of Congress, Network Development and MARC Standards Office, “MARCXML: MARC 21 XML Schema,” July 26, 2006, www.loc.gov/standards/marcxml (accessed September 3, 2007).

19. Andrew K. Pace, "Category: Metasearch," *Hectic Pace*, <http://blogs.ala.org/pace.php?cat=150> (accessed August 27, 2007). See in particular the following blog entries: "MetaMeta," July 25, 2006; "More Meta," September 29, 2006; "Preaching to the Publishers," Oct 31, 2006; "Even More Meta," July 11, 2007; and "Still Here," August 21, 2007.
20. David Lindahl, "Metasearch in the Users' Context," *The Serials Librarian* 51, no. 3/4 (2007): 220–222.
21. EthicShare, a collaborative project of the University of Minnesota, Georgetown University, Indiana University–Bloomington, Indiana University–Purdue University Indianapolis, and the University of Virginia, is addressing this challenge as part of its plan to develop a sustainable online environment for the practical ethics community. The architecture of the proposed EthicShare system has many similarities to that of XC, but the project focuses specifically upon ingesting citation metadata from a variety of sources, including commercial providers. See Cecily Marcus, "EthicShare Planning Phase Final Report," July 2007, www.lib.umn.edu/about/ethicshare/University%20of%20Minnesota_EthicShare_Final_Report.pdf (accessed August 27, 2007).
22. Roy Tennant used this phrase in "MARC Exit Strategies," *Library Journal* 127, no. 19 (November 15, 2002), www.libraryjournal.com/article/CA256611.html?q=tennant+exit (accessed July 23, 2007); Karen Coyle presented her vision for moving beyond MARC to a more flexible, identifier-based record structure that will facilitate a range of library functions in "Future Considerations: The Functional Library Systems Record," *Library Hi Tech* 22, no. 2 (2004).
23. Library of Congress, Network Development and MARC Standards Office, "METS: Metadata Encoding and Transmission Standard Official Web Site," August 23, 2007, www.loc.gov/standards/mets (accessed September 3, 2007).
24. Library of Congress, Network Development and MARC Standards Office, "MODS: Metadata Object Description Schema," August 22, 2007, www.loc.gov/standards/mods (accessed September 3, 2007).
25. Library of Congress, Network Development and MARC Standards Office, "MADS: Metadata Authority Description Schema," February 2, 2007, www.loc.gov/standards/mads (accessed September 3, 2007).
26. "PREMIS: Preservation Metadata Maintenance Activity," July 31, 2007, www.loc.gov/standards/premis (accessed September 3, 2007).
27. Library of Congress, Network Development and MARC Standards Office, "EAD: Encoded Archival Description Version 2002 Official Site," August 17, 2007, www.loc.gov/ead (accessed September 3, 2007).
28. Visual Resources Association, "VRA Core: Welcome to the VRA Core 4.0," www.vraweb.org/projects/vracore4 (accessed September 3, 2007).
29. "Dublin Core Metadata Element Set, Version 1.1."
30. Other XML-compatible schemas, such as MODS and MADS, will also be supported initially in XC if they are first converted into MARC XML or Qualified Dublin Core. In the future, we plan to allow these other schemas to be harvested directly into XC.
31. Foster et al., "eXtensible Catalog Survey Report," July 20, 2007, 15. The original comment was submitted by Meg Bellinger in Yale University's response to the XC Survey.
32. Patricia Harpring et al., "Metadata Standards Crosswalks," in *Introduction to Metadata: Pathways to Digital Information* (Getty Research Institute, n.d.), www.getty.edu/research/conducting_research/standards/intrometadata/crosswalks.html (accessed August 29, 2007); see also Carol Jean Godby, Jeffrey A. Young, and Eric Childress, "A Repository of Metadata Crosswalks," *D-Lib Magazine* 10, no. 12 (December 2004), www.dlib.org/dlib/december04/godby/12godby.html (accessed July 23, 2007).
33. Digital Library Federation, "CrosswalkingLogic," June 22, 2007, <http://webservices.itscs.umich.edu/mediawiki/oaibp/index.php/CrosswalkingLogic> (accessed August 28, 2007).
34. Karen Coyle et al., "Framework for a Bibliographic Future," May 2007, <http://futurelib.pbwiki.com/Framework> (accessed July 23, 2007).
35. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records*.
36. Andy Powell et al., "DCMI Abstract Model," *Dublin Core Metadata Initiative*, June 4, 2007, <http://dublincore.org/documents/abstract-model> (accessed August 29, 2007).
37. Joint Steering Committee for Development of RDA, "RDA: Resource Description and Access: Background," July 16, 2007, www.collectionscanada.ca/jsc/rda.html (accessed August 29, 2007).
38. Joint Steering Committee for Development of RDA, "RDA-FRBR Mapping," June 14, 2007, www.collectionscanada.ca/jsc/docs/5rda-frbrmapping.pdf (accessed August 29, 2007).
39. Joint Steering Committee for Development of RDA, "RDA Element Analysis," June 14, 2007, www.collectionscanada.ca/jsc/docs/5rda-elementanalysis.pdf (accessed August 28, 2007). A revised version of the document was issued on December 16, 2007, at www.collectionscanada.gc.ca/jsc/docs/5rda-elementanalysisrev.pdf (accessed December 30, 2007).
40. "Data Model Meeting: British Library, London 30 April–1 May 2007," www.bl.uk/services/bibliographic/meeting.html (accessed July 23, 2007). The Task Group has outlined its work plan, including deliverables, on its wiki at <http://dublincore.org/dcmirdataskgroup> (accessed October 4, 2007).
41. Emily A Hicks, Jody Perkins, and Margaret Beecher Maurer, "Application Profile Development for Consortial Digital Libraries," *Library Resources and Technical Services* 51, no. 2 (April 2007).
42. Makx Dekkers, "Application Profiles, or How to Mix and Match Metadata Schemas," *Cultivate Interactive*, January 2001, www.cultivate-int.org/issue3/schemas (accessed August 29, 2007).
43. Thomas Baker et al., "Dublin Core Application Profile Guidelines," September 3, 2005, <http://dublincore.org/usage/documents/profile-guidelines> (accessed October 8, 2007).
44. Joint Steering Committee for Development of RDA, "RDA Element Analysis."
45. Karen Coyle and Diane Hillmann, "Resource Description and Access (RDA): Cataloging Rules for the 20th Century," *D-Lib Magazine* 13, no. 1/2 (Jan./Feb. 2007), www.dlib.org/dlib/january07/coyle/01coyle.html (accessed August 24, 2007).
46. Karen Coyle, "Astonishing Announcement: RDA Goes 2.0," *Coyle's InFormation*, May 3, 2007, <http://kcoyle.blogspot.com/2007/05/astonishing-announcement-rda-goes-20.html> (accessed August 29, 2007).

47. "Drupal.org," <http://drupal.org> (accessed August 30, 2007).
48. Foster et al., "eXtensible Catalog Survey Report," 14.
49. "Taxonomy: A Way to Organize Your Content," *Drupal.org*, <http://drupal.org/handbook/modules/taxonomy> (accessed September 12, 2007).
50. "Blackboard Learning System," www.blackboard.com/products/academic_suite/learning_system/index.Bb (accessed August 31, 2007).
51. "Sakai: Collaboration and Learning Environment for Education," <http://sakaiproject.org> (accessed August 31, 2007).
52. For example, the Library into Blackboard Project at California State Fullerton has developed a toolkit for faculty that brings OpenURL resolver functionality into Blackboard to create linked citations to resources. See "Putting the Library into Blackboard: A Toolkit for Cal State Fullerton Faculty," 2005, www.library.fullerton.edu/librarytoolkit/default.shtml (accessed August 31, 2007); and Susan Tschabrun, "Putting the Library into Blackboard: Using the SFX OpenURL Generator to Create a Toolkit for Faculty." The Sakaibrary Project at Indiana University and the University of Michigan are working to integrate licensed library content into Sakai using metasearch technology. See "Sakaibrary: Integrating Licensed Library Resources with Sakai," June 28, 2007, www.dlib.indiana.edu/projects/sakai (accessed August 31, 2007).
53. University of Rochester River Campus Libraries, "eXtensible Catalog Phase 2."
54. Susan Gibbons, "Library Course Management Systems: An Overview," *Library Technology Reports* 41, no. 3 (May/June 2005): 34–37.
55. Marti A. Hearst, "Design Recommendations for Hierarchical Faceted Search Interfaces," August 2006, <http://flamenco.berkeley.edu/papers/faceted-workshop06.pdf> (accessed August 31, 2007).
56. Kristin Antelman, Emily Lynema, and Andrew K. Pace, "Toward a Twenty-First Century Library Catalog," *Information Technology and Libraries* 25, no. 3 (September 2006): 128–138.
57. "C4," <https://www.library.rochester.edu/c4> (accessed September 28, 2007). As of the time of this writing, the C4 prototype is available to the public. However, the prototype is no longer being developed, and this prototype may cease to be available at some point in the future.
58. Charley Pennell, "Forward to the Past: Resurrecting Faceted Search @ NCSU Libraries," (PowerPoint presentation at the American Library Association Annual Conference, Washington, D.C., June 24, 2007), www.lib.ncsu.edu/endeca/presentations/200706-facetedcatalogs-pennell.ppt (accessed August 31, 2007).
59. Mary Charles Lasater, "Authority Control Meets Faceted Browse: Vanderbilt and Primo," (PowerPoint presentation at the American Library Association Annual Conference, Washington, D.C., June 24, 2007), www.ala.org/ala/lita/litamembership/litaigs/authorityalcts/2007annualfiles/MaryCharlesLasater.ppt (accessed August 31, 2007).
60. Casey Bisson, "Faceting and Clustering: An Implementation Report Based on Scriblio," (PowerPoint presentation at the American Library Association Annual Conference, Washington, D.C., June 24, 2007), http://oz.plymouth.edu/~cbisson/presentations/ALAAnnual_2-2007June24.pdf (accessed August 31, 2007).
61. "Subject Access Fields (6XX)," in *MARC 21 Concise Format for Bibliographic Data* (2006), www.loc.gov/marc/bibliographic/ecbdsbj.html (accessed September 28, 2007).
62. Pennell, "Forward to the Past: Resurrecting Faceted Search@ NCSU Libraries."
63. "FAST: Faceted Application of Subject Terminology," www.oclc.org/research/projects/fast (accessed August 31, 2007).
64. IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records*.
65. IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), "Functional Requirements for Authority Data."
66. Library of Congress, Network Development and MARC Standards Office, "Functional Analysis of the MARC 21 Bibliographic and Holding Formats," April 6, 2006, www.loc.gov/marc/functional-analysis/functional-analysis.html (accessed August 31, 2007); Martha M. Yee, "FRBRization: A Method for Turning Online Public Finding Lists into Online Public Catalogs," *Information Technology and Libraries* 24, no. 2 (June 2005): 77–95; Pat Riva, "Mapping MARC 21 Linking Entry Fields to FRBR and Tillet's Taxonomy of Bibliographic Relationships," *Library Resources and Technical Services* 48, no. 2 (April 2004): 130–143.
67. Trond Aalberg, "A Process and Tool for the Conversion of MARC Records to a Normalized FRBR Implementation," in *Digital Libraries: Achievements, Challenges and Opportunities* (Berlin/Heidelberg: Springer, 2006), 283–292; Christian Monch and Trond Aalberg, "Automatic Conversion from MARC to FRBR," in *Research and Advanced Technology for Digital Libraries* (Berlin/Heidelberg: Springer, 2003): 405–411; David Mimno and Gregory Crane, "Hierarchical Catalog Records: Implementing a FRBR Catalog," *D-Lib Magazine* 11, no. 10 (October 2005), www.dlib.org/dlib/october05/crane/10crane.html (accessed August 24, 2007).
68. Trond Aalberg, Frank Berg Haugen, and Ole Husby, "A Tool for Converting from MARC to FRBR," in *Research and Advanced Technology for Digital Libraries* (Berlin/Heidelberg: Springer, 2006), 453–456; "FRBR Work-Set Algorithm," www.oclc.org/research/software/frbr/default.htm (accessed August 31, 2007); "xISBN (Web Service)," www.worldcat.org/affiliate/webservices/xisbn/app.jsp (accessed August 31, 2007).
69. For example, MARC 21 data may need to be augmented to extract data attributes related to FRBR works and expressions that are not explicitly coded within a MARC 21 bibliographic record (such as a date associated with a work coded within a general note field); or to "sort out" the fields in a MARC 21 bibliographic record for a single resource that contains various works and/or expressions (e.g., a sound recording with multiple tracks), to associate the various fields (performer access points, analytical entries, subject headings, etc.) with the appropriate work or expression.
70. While the RIT-developed tool is not publicly available at the time of this writing, it is our intent to post it to Sourceforge (www.sourceforge.net) in the near future. The final report of the RIT project is available at <http://docushare.lib.rochester.edu/docushare/dsweb/Get/Document-27362> (accessed January 2, 2008).

71. Foster et al., "eXtensible Catalog Survey Report."

72. Note the arrow pointing to the left in Figure 1 between the user environments and the metadata services hub.

73. Jane Greenberg and Eva Mendez, *Knitting the Semantic Web* (Binghamton, NY: Haworth Information Press, 2007). This volume, co-published simultaneously as *Cataloging and Classification Quarterly* 43, no. 3/4, contains a wealth of articles that explore the role that libraries can, and should, play in the development of the Semantic Web.

74. Corey A. Harper and Barbara B. Tillett explore various methods for making these controlled vocabularies available in "Library of Congress Controlled Vocabularies and Their Appli-

cation to the Semantic Web," *Cataloging and Classification Quarterly* 43, no. 3/4 (2007): 63. The development of SKOS (Simple Knowledge Organization System), a Semantic Web language for representing controlled structured vocabularies, will also be valuable for XC. See Alistair Miles and Jose R. Perez-Aguiera, "SKOS: Simple Knowledge Organisation for the Web," *Cataloging and Classification Quarterly* 43, no. 3/4 (2007).

75. Marcus, "EthicShare Planning Phase Final Report."

76. The Talis Platform provides another promising environment for experimentation and development. See "Talis Platform: Semantic Web Application Platform," *Talis*, www.talis.com/platform (accessed September 2, 2007).