Moving Library Metadata toward Linked Data: Opportunities Provided by the eXtensible Catalog

Jennifer Bowen
University of Rochester
eXtensible Catalog
Organization
New York, USA
jbowen@library.rochester.edu

Abstract

To ensure that they can participate in the Semantic Web, libraries need to prepare their legacy metadata for use as linked data. eXtensible Catalog (XC) software facilitates converting legacy library data into linked data using a platform that enables risk-free experimentation and that can be used to address problems with legacy metadata using batch services. The eXtensible Catalog also provides "lessons learned" regarding the conversion of legacy data to linked data by demonstrating what MARC metadata elements can be transformed to linked data, and helping to suggest priorities for the cleanup and enrichment of legacy data. Converting legacy metadata to linked data will require a team of experts, including MARC-based catalogers, specialists in other metadata schemas, software developers, and Semantic Web experts to design and test normalization/conversion algorithms, develop new schemas, and prepare individual records for automated conversion. Library software applications that do not depend upon linked data may currently have little incentive to enable its use. However, given recent advances in registering legacy library vocabularies, converting national library catalogs to linked data, and the availability of open source software such as XC to convert legacy data to linked data, libraries may soon find it difficult to justify continuing to create metadata that is not linked data compliant. The library community can now begin to propose smart practices for using linked data, and can encourage library system developers to implement linked data. XC is demonstrating that implementing linked data, and converting legacy library data to linked data, are indeed achievable.

Keywords: eXtensible Catalog; linked data; RDF; Semantic Web; metadata; MARC; FRBR

Acknowledgements

The author appreciates the support of her colleagues, David Lindahl, Nancy Fried Foster, and Randall Cook on the eXtensible Catalog leadership team for their support, and for their willingness to incorporate linked data into the design of the XC software. Many thanks also to XC's partner institutions, especially Kyushu University, for their patience as we carried out this development; to Diane Hillmann for her feedback on a draft of this paper, and to Jon Phipps and Nannette Naught for contributing to some of the ideas in the paper.

1. Introduction

Within the past two years, calls for enabling library metadata to function as a part of the Semantic Web have been heard more and more frequently. Why is this important? With information seekers more often than not bypassing library systems to locate information directly on the web, it is essential that library metadata be understandable within the broader web environment, outside of closed systems that have been designed around the unique features of library metadata. As Karen Coyle describes it, we must enable library data to "…be integrated into the virtual working spaces of the users served by the library" (Coyle, 2010a, p. 5). To ensure that they are able to participate in this new environment, libraries need to prepare their existing metadata for use as linked data within the Semantic Web.

Linked data is a concept first articulated by Tim Berners-Lee in 2006. In his initial description of it, Berners-Lee outlined four rules, or expectations for behavior, for linked data:

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs so that they can discover more things (Berners-Lee, 2006).

Since 2006, linked data has sparked a community of activity on the Web ("Linked Data: Connect Distributed Data across the Web," n.d.).

While linked data has gained some momentum on the Web in general, the preparation of library metadata to function as linked data has not moved forward in quite as timely a manner. This may be because of the lack of a widespread understanding within the library community of how linked data works and why it is important. Library technology projects that were planned or underway before the possibilities of linked data were articulated will likely need to alter their development plans in midstream to support linked data. Making this kind of change "on the fly" may be difficult, especially if a project needs additional resources to accomplish this. Incorporating linked data into an existing project that does not otherwise require it may seem at best an unwelcome distraction, or at worst an unnecessary risk. Since best practices for linked data are still being formulated, any project that implements linked data in the near future faces uncertainty about whether it has done so correctly. This may set a project up for criticism later on, once best practices for linked data are established. Understandably, some projects may decide not to take such a risk.

Despite these challenges, the team working to develop the eXtensible Catalog (XC) decided to move forward with adding the implementation of linked data to an already well-defined project. In the case of XC, the nature of the software architecture makes it an ideal environment for facilitating linked data. XC provides a platform for bringing legacy library data into a linked data environment and for addressing potential difficulties with metadata. It can therefore facilitate the development and evolution of best practices for linked data. The experiences of the XC development team with repurposing MARC metadata provide valuable lessons for transforming legacy library metadata to a linked data environment. XC also provides a safe environment where the library community can experiment with implementing linked data without jeopardizing the integrity of existing metadata, and without changing established metadata practices before new practices can be evaluated.

2. Related Work

While the literature of the information science and metadata communities has included discussions of how libraries might participate in the Semantic Web for the past several years, it is only more recently that such discussions have been seen with any regularity within more general library and library cataloging literature in the U.S. Several topics and projects related to libraries and the Semantic Web are described in a special issue of the journal, *Cataloging & Classification Quarterly*, entitled "Knitting the Semantic Web," which includes a dozen papers by a variety of prominent scholars (Greenberg & Mendez, 2007). The authors represented in this volume present a strong case for the important contribution that libraries can make to the development of the Semantic Web, as well as for the benefit that libraries can derive from such efforts. Coyle has written extensively in mainstream library journals recently about the Semantic Web (Coyle, 2008, 2009, 2010c), and others have also made significant contributions to bringing libraries' potential role in the Semantic Web to the attention of the broader library community (Chudnov 2009a, 2009b, 2009c; Bradley, 2009). Yee has taken the discussion to a more detailed level in her attempt to model bibliographic data within RDF, thus exploring how traditional library cataloging might move into the Semantic Web domain (Yee, 2009).

Developments related to moving library data to the Semantic Web have now begun to move forward more rapidly as a variety of agencies within the library community have recently registered metadata element sets and vocabularies for FRBR (Dunsire, 2010), RDA (Hillmann, Coyle, Phipps, & Dunsire, 2010; Coyle, 2010b) and Library of Congress Subject Headings (Summers, Isaac, Redding, & Krech, 2008) in RDF-compatible registries. Beyond efforts such as these to convert legacy vocabularies and vocabularies to linked data, an increasing number of library-related agencies are taking steps toward making their entire collections of legacy library metadata available as linked data. The national libraries of Sweden, Hungary, and Germany have all recently announced the availability of their catalogs as linked data ("LIBRIS available as Linked Data at semweb@libris," 2008; Pohl, 2010; Horvath, 2010). Proponents of linked data within the library community are now conveying a sense of the urgency with which the library community must take advantage of these opportunities (Coyle, 2010a; Harper, 2009). It is into this newly-energized conversation about linked data and libraries that the eXtensible Catalog is finding an opportunity to contribute.

3. About the eXtensible Catalog

The eXtensible Catalog is open source, user-centered, next generation software for libraries. It comprises four software components that can be used independently to address a particular need, or combined to provide an end-to-end discovery system to connect library users with resources. Developed under the leadership of the University of Rochester in Rochester, New York, and supported through grants from The Andrew W. Mellon Foundation and by XC partner institutions, the XC software is now supported and managed by the eXtensible Catalog Organization (XCO), a sole-member LLC formed by the University of Rochester. The XCO employs a team of software developers to build new features into the XC software code, help community members incorporate their contributions, and ensure that the XC software serves the needs of libraries. The software is available for download via the XCO's website (eXtensible Catalog Organization, 2010).

To enable libraries to make the best possible use of their legacy metadata, XC software provides tools for normalizing MARC metadata, transforming it into other schemas, aggregating it, and preparing it for use in the XC user interface or in other applications (Lindahl, 2010). Previous papers have discussed in detail the goals and requirements for metadata within the XC (Bowen, 2008) and the specific functionality of XC software related to metadata and metadata services (Bowen, 2009). The XC Metadata Services Toolkit (MST), one of the four XC software toolkits, empowers libraries to easily manage their legacy metadata and prepare it for use in a variety of web applications.

The primary metadata-related objective of the XC Project has been to use the MST to transform metadata so that it can be used within XC's user interface. Throughout XC's development, however, the XC team has gained an understanding of the role that XC software can play toward encouraging the library community to implement linked data. To take advantage of this opportunity, we have worked to ensure that XC's software platform can also act as a framework for implementing linked data.

4. Issues Related to Legacy MARC Metadata

The problems and frustrations inherent in working with MARC data have been described in detail by others (Tennant, 2002a; Tennant, 2002b; Coyle, 2006; Coyle, 2010d; Spero, 2010). Many of these problems show up as a result of attempting to repurpose MARC metadata in other non-MARC environments. Such repurposing means that we are expecting metadata that conforms to standards originally intended for communicating card catalog data to now serve an entirely different purpose, a purpose that was never envisioned when much of the metadata was created.

To function as linked data, catalog data that originally was intended to be understood only by humans must now be understandable to machines, without human intervention. It is not surprising that repurposing legacy data in this environment can be problematic.

The XC team began with a goal to clean up and transform legacy MARC metadata so that it can be used, in combination with metadata from other sources, within the XC user interface. While this original goal did not include creating linked data, many of the issues that we encountered with our initial data transformation work are the same issues that will need to be addressed to convert MARC data to linked data. Our experiences working toward our initial goals can therefore also inform the more overarching goal of transforming MARC data to function as linked data. While the actual process of creating linked data requires providing URIs for data elements, registering vocabularies, and enabling the eventual output of the metadata as RDF, these tasks must take place within the context of mapping MARC data to other data elements, identifying and describing relationships between data elements and resources, and capturing the complexities of the original metadata. We expected to (and did!) encounter problems in all of these areas when mapping MARC to our own XC Schema. Some of these difficulties are related to the structure of MARC itself, while others are a result of cataloging practices used in conjunction with MARC. The issues that we faced in these three areas are described below.

4.1. Data Mapping

The XC MST software and related metadata services can process any type of XML data, normalize it, transform it from one schema to another, and merge records that represent the same entity by matching on identifiers. XC's MARC Normalization and Transformation services can handle the following situations very successfully:

- Converting MARC codes to corresponding vocabulary terms for display in a user interface
- Removing extraneous data that may interfere with other uses of a metadata element (i.e. as an identifier), such as text that is not part of an ISBN at the end of a MARC 020 field
- Mapping the majority of discrete MARC data fields and parsing them to FRBR Group 1 entities (IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998).

Specific difficulties with reusing legacy MARC metadata include, among other things, the lack of useful granularity to map free-text data to specific data elements. This is a problem even when MARC records are tagged correctly and completely. It is especially difficult to map transcribed data fields such as the MARC 245 field, which can contain multiple instances of information within the same subfield (e.g. multiple statements of responsibility, parallel titles). For example, in Figure 1, the English subtitle, the French parallel title and subtitle, and additional subtitle information containing the place and date of the event all appear in the same subfield (245 ‡b).

245 10 ‡a Proceedings : ‡b International Conference of the Association for the Advancement of Rehabilitation Technology, choice for all = Compte rendu : Conférence internationale pour le développement de la technologie en réadaptation, choix pour tous : ICAART 88, Montréal, Palais des Congrès, 25-30, June-juin 1988.

FIG. 1: Insufficient subfielding in MARC Field 245 for data parsing

The MARC 300 field, which contains multiple different metadata elements within the same subfield (e.g. illustrations, playing speed, etc.), also requires complicated parsing. Table 1 shows the mappings used by XC services for 300 subfield ‡b. Depending upon the type of MARC bibliographic record (as distinguished by the MARC Bibliographic Leader Value 06), the data in this subfield may correspond to data associated with one or another FRBR Group 1 entity (either FRBR expression or manifestation). While the data for some record types, including sound recordings, language material and musical scores, can be mapped successfully to specific data elements, other record types can be mapped much less reliably, which led us to define an XC-specific metadata element called "Other physical details" that is defined specifically to hold this indeterminate MARC-derived data.

TABLE. 1. XC mappings for MARC 300 Subfield ‡b

Type of MARC bibliographic record	FRBR entity mapping	XC Schema element
Sound recordings (Leader 06=i or j)	manifestation	rdvocab:soundCharacteristics
Language material, musical scores (Leader 06=a, c, d, or t)	expression	rdvocab:illustrativeContent
Other (any other Leader 06 value)	manifestation	xc:otherPhysicalDetails

At times, a single MARC bibliographic record may contain metadata that pertains to more than one FRBR work, or to more than one FRBR manifestation. When this occurs, it is desirable to parse the MARC data into multiple separate records to enable the metadata to be transformed accurately to linked data. Unfortunately, the original MARC records frequently lack the necessary information to parse this data accurately. For example, when a MARC record contains multiple analytics in 7XX fields (for additional works represented within the resource, such as a sound recording with multiple compositions in multiple tracks), it is possible to parse the 7XX fields themselves into separate records, but not to parse the other data elements appropriately between the records unless MARC linking subfields (which are defined in MARC, but seldom used for this purpose) are available to show which subject headings, performers, etc., go with each work. An example of this situation is shown in the abbreviated MARC record in Figure 2.

245 10 ‡a Composers' Forum ‡h [sound recording].
[…]
505 0 ‡a Mechanical etude / Yie-Eun Chun (3:11) (Yie-Eun Chun, piano) -- Piece for solo violin / Kyle Quarles (1:58) (Markiyan Melnychenko, violin) -- Nocturne I / Stephen Danyew (5:35) (Y. Chun, piano) -- Sanjo : for solo harp / Y. Chun (6:04) (Hillary Schefter, harp) -- Little orchard / Matt Barber (6:20) (Zuzanna Szewczyk, piano).
650 0 ‡a Piano music.
650 0 ‡a Violin music.
650 0 ‡a Harp music.
700 1 ‡a Chun, Yie Eun. ‡4 prf
700 1 ‡a Melnychenko, Markiyan. ‡4 prf
700 1 ‡a Schefter, Hillary. ‡4 prf
700 1 ‡a Szewczyk, Zuzanna. ‡4 prf
700 12 ‡a Chun, Yie Eun. ‡t Mechanical etude.
700 12 ‡a Quarles, Kyle. ‡t Piece, ‡m violin.
700 12 ‡a Danyew, Steve, ‡d 1983- ‡t Nocturnes, ‡m piano, ‡n no. 1.
700 12 ‡a Chun, Yie Eun. ‡t Sanjo.
700 12 ‡a Barber, Matthew Stephen. ‡t Little orchard.

FIG 2: Portion of a MARC record for a sound recording showing fields that are difficult to parse to FRBR entities

In this case, five different works are represented in the analytics (700/12 fields), but are covered by only three subject headings (650 fields) and four performers (700/1 fields). Even with the aid of the 505 field that explains who is performing on which work, it would be impossible for an automated service to parse this information accurately without manual intervention (i.e. record-by-record editing). Similarly, the use of a single MARC record to describe more than one format or version of a resource (e.g. both print and electronic versions described on a single record) is problematic in this regard because the metadata cannot be parsed automatically into separate manifestation records for the two formats.

Local MARC coding practices that were intended to get around the limitations of a particular user interface also cause considerable problems when mapping MARC metadata. One example is the practice of adding leading zeros to identifiers and accession numbers to facilitate a numeric data sort. Libraries also have been known to facilitate display of MARC metadata by breaking a single field into multiple MARC tags to make it more readable on a screen, or even to include formatting punctuation (such as a series of dashes to create a line ------) in a MARC tag to create a visual separation of data in a screen display. Problems such as this can be fixed using XC's Normalization Service, assuming that the problems occur fairly consistently across a group of metadata records. However, the need to deal with such problems complicates the mapping of MARC metadata for repurposing in other environments, whether for use within XC and/or for use as linked data. The library community has come a long way toward understanding that we must create metadata that can be deciphered outside of a specific software application, but such legacy practices will likely continue to come to light as we repurpose older metadata.

4.2. Describing Relationships

The XC metadata service that we have defined for XC to transform MARC data to the XC Schema successfully maps over 100 MARC tags to XC Schema records that represent the FRBR Group 1 entities work, expression, and manifestation ("MARCXML to XC Transformation Service Documentation, eXtensible Catalog," 2009). Our goal when creating this service was to retain as much of the richness inherent in MARC data as possible, and we feel that we succeeded by mapping the vast majority of the MARC data that we wanted to carry forward into XC Schema records. However, there is more to transforming MARC data to a FRBRized schema (and eventually to linked data) than simply mapping the data elements. It is also necessary to identify the relationship between each data element and a specific FRBR entity, and there are times when such relationships are difficult to identify in legacy MARC data. A MARC element may relate to an entity that is not defined in FRBR, or to a FRBR entity that is not explicitly identified within the original MARC record.

The MARC formats pre-date the FRBR data model, so there is not necessarily a close correlation between the metadata contained in the various MARC formats and the groups of entities defined in FRBR. The relationship between MARC and FRBR is described in detail in a study commissioned by the Library of Congress (Library of Congress. Network Development and MARC Standards Office, 2006). While the lack of a conceptual model behind the MARC formats complicates this process, data elements defined in RDA (RDA Toolkit: Resource Description & Access, n.d.) should present less of a problem in the future because they are tied closely to the FRBR entity-relation model and have been analyzed in terms of RDF as part of the RDA vocabularies project (Hillmann, Coyle, Phipps and Dunsire, 2010).

One example of MARC bibliographic/holdings elements that proved difficult to relate to a FRBR entity is summary holdings (866/868). Summary holdings is a useful and widespread display convention for describing a group of individual items/volumes that have common characteristics, such as being owned by the same library or shelved together in one location. While summary holdings cannot be related in a practical manner to any one of the FRBR Group 1 entities (Library of Congress. Network Development and MARC Standards Office, 2006, Table 3, 82-83), there may be other ways that the relationships represented in this data could be conveyed within RDF, perhaps as an ownership relationship between a resource and a specific library or institution. Such examples show that relying upon FRBR mappings for all MARC metadata elements will still not solve all of our relationship/description issues, since FRBR does not cover every metadata element that is important to libraries. Additional data modeling will likely be necessary to accommodate all MARC metadata elements in RDF.

In other cases, it is possible to identify the relationship represented by a MARC tag, but the surrounding metadata to specifically identify the related entity is not recorded in the original MARC record. For example, MARC tag 041/‡h encodes the original language of a resource that is a translation of another resource. In FRBR terms, the data in this MARC subfield represents one attribute of an expression that is related to the expression represented by the resource described in the MARC record (Library of Congress. Network Development and MARC Standards Office, 2006, Table 3, p. 15). The MARC record may or may not explicitly identify the related expression, making it impossible to provide an adequate description of the related resource, let alone to relate this data to a description of the related resource. It would likewise be difficult to translate this data element into RDF if the relationship to a specific resource cannot be identified. The new RDA standard defines a Derivative Expression relationship for this data, "Translation of (expression)" ("RDA: Resource Description and Access. Constituency Review," 2008, Appendix J, p. 11). Now that RDA data elements are registered in a way that enables RDF output ("RDA Relationships for Works, Expressions, Manifestations, Items: Show Detail," n.d.), it should be possible in the future to record RDA metadata so that this data element can be converted to linked data. However, it seems unlikely that language codes in the MARC 041/‡h could be repurposed in a non-MARC environment without manual intervention.

Our difficulties with identifying relationships in MARC data might seem to be an issue only for applications that are implementing MARC metadata in a FRBR-based environment. However, the relationship between a metadata element and some other entity (FRBR or not) must be articulated in order to use the metadata within any Semantic Web application. The structure of an RDF triple identifies the relationship between the triple's subject and object. Therefore, the inability to identify or describe the relationship between data in a MARC tag and another entity will likely complicate the automatic creation of RDF from MARC metadata even when FRBR is not involved.

4.3. Capturing Additional MARC Data Complexities

The MARC 880 field, "Alternate Graphic Representation" and the MARC 9XX Equivalence and Cross Reference Fields (used by Canadian libraries) may be difficult to convey as linked data because of their complexity. The 880 field replicates metadata in another field in the same record in a different script. Each 880 field identifies the other field in the record that it relates to, the script of the data included, and the text of the data element in the alternate script. 880 fields can theoretically parallel any other variable field in a MARC record, but are commonly used for transcribed fields, such as title and publication information, as well as for headings (access points).

If the metadata in an 880 field corresponds to an attribute of a Group 1 FRBR entity (as is the case with titles, publication information, etc.), the 880 field can be mapped appropriately to the same FRBR entity as its paired field. However, if 880 fields are used for more than a single instance of the same MARC tag, the paired relationships between the multiple fields will need to be maintained within the new environment. Since more than one relationship would need to be captured in the linked data (i.e. the relationship between the metadata field and the FRBR entity that it relates to, plus the relationship of the paired metadata fields to each other), it is unclear how these MARC fields would be converted to RDF triples.

The situation is more complex when the 880 field contains an alternative script for a heading (access point) that represents a FRBR Group 2 or Group 3 entity instead of containing an attribute for a FRBR Group 1 entity. In this case, simply mapping paired fields that both represent the same FRBR Group 2 entity to a record for a Group 1 entity would perpetuate (in library terms) the questionable practice of adding added entries to a bibliographic record for variant forms of a name, rather than using cross references in an authority record. To avoid repeating the alternate script entries in every relevant Group 1 entity record, these alternate scripts should be mapped from MARC bibliographic records to records that describe the Group 2 or 3 entities to which they apply. This would add another layer of complexity to the mapping process. To make this situation even more complicated, bibliographic records often contain more than one heading/access point of the same type (e.g. two subject headings), thus creating many instances of the situation described in the preceding paragraph. The library community will need to address these situations once it gains more experience modeling library data in RDF.

5. Establishing Best Practices for Linked Data

While working with legacy library metadata, the XC team encountered a variety of past metadata practices that were evidently intended to enable a specific application or user interface to function effectively but which now cause significant problems when reusing legacy metadata in other environments. Having seen these problems first-hand, we have attempted to avoid favoring the functional needs of our own software at the expense of adhering to sound metadata practices. It is especially important that decisions made regarding metadata within XC do not create additional problems for other applications since we intend for XC to process metadata for use within any number of other applications.

Because XC's software is able to process metadata for a variety of applications, the XC team is in a position to promote the establishment and use of appropriate practices to facilitate linked data. As we learn more about what best practices for linked data would actually entail, we hope that our findings will inform MARC coding practices, just as the recent OCLC Research/RLG Partnership study of MARC tag usage has recommended best practices for library MARC metadata practices (Smith-Yoshimura, Argus, Dickey, Naun, Ortiz, & Taylor, 2010, p. 13). Using XC to transform MARC data could also inform future MARC development. Best practices for linked data will take time to evolve, but XC is taking initial steps to adhere to what we understand to date to be best practices, as described below.

Two of these best practices, using registered schemas and registered vocabularies, relate to Berners-Lee's third rule for linked data:

3. When someone looks up a URI, provide useful information (Berners-Lee, 2006).

A registry provides useful information about a URI not only when a person looks it up, but even more importantly when a machine looks it up, i.e. when it is resolved by a web-based application. XC's efforts to conform to this rule within the XC Schema are described in the following section. In addition, XC can facilitate the ability for other library projects to conform to Berners-Lee's rules, as described in Section 5.3, Enabling Experimentation.

5.1. Data Elements from Registered Schemas

The eXtensible Catalog uses its own XML metadata schema, called the XC Schema, to enable the functionality of the XC User Interface. In designing this schema, the XC team identified and included as many data elements/properties as possible from other existing registered namespaces rather than defining all elements for the XC Schema from scratch. We plan to create an application profile for XC using the Singapore Framework (Baker and Johnston, 2008) to ensure that our metadata usage is understandable to others who want to use it, whether those users are humans or online applications. As described by Coyle, an RDF-compliant application profile allows elements and vocabularies to be taken from any suitable defined set (Coyle, 2010a, pp. 34-35). Using elements from other namespaces has simplified the process of defining a schema for XC because the majority of the data elements are already defined. The XC Schema contains all properties from the Dublin Core "dcmi-terms" namespace (DCMI, 2008), plus a subset of the data elements from RDA's properties and relationships. We can also add elements from other namespaces (or more data elements from the RDA element set) to our application profile as needed in the future.

To facilitate the functionality of the XC software, we have also defined a number of data elements for the XC Schema that are unique to XC (Bowen, 2009), and which we plan to register in the NSDL Metadata Registry ("The Registry!" n.d.). The NSDL Metadata Registry provides a "sandbox" for registering a data element or vocabulary provisionally, which makes it easy to begin working toward defining URIs for data elements while the element definitions are still being refined. Figure 3 shows the provisional registration of one XC data element and how it is registered in the Registry Sandbox. This particular XC element, "type007", is defined as a placeholder for the vocabulary used in the MARC Bibliographic format 007/00 field. ("The Registry! :: XC Elements :: Type from MARC 21 007 :: Show Detail," n.d.)

NSDLREGISTRY
Supporting Metadata Interoperability

Element Sets: XC Elements
Elements: Type from MARC 21 007

Detail | Statements | History

Metadata +

Detail

Label:	Type from MARC 21 007
Name:	type007
URI:	http://www.extensiblecatalog.info/Elements/type007 (RDF)
Description:	A term that defines the category of material, derived from MARC 21 Bibliographic 007/00
Comment:	Associated with the FRBR entity manifestation. Uses as its vocabulary encoding scheme the terms associated with codes used in MARC 21 Bibliographic 007/00.
Type:	property
Domain:	
Range:	
Status:	New-Proposed
Language:	English

FIG. 3. Proposed XC data element based upon the MARC 21 007/00 data element and vocabulary

An example of three XC Schema records for a linked work, expression and manifestation appear in Figure 4A, 4B, and 4C, followed in Figure 5 by the MARC record that was used as the source of the metadata in these records. The three records in Figure 4 show the use of data elements from a variety of namespaces (Dublin Core terms, RDA elements and roles, and newly-defined XC elements). By using metadata elements from a variety of identified schemas within the XC Schema, we hope to encourage libraries to experiment with using a variety of registered metadata elements in the same application, rather than inventing a totally new schema whenever existing schemas do not exactly fit local needs. The flexibility offered by application profiles may be one of the early benefits that libraries gain from applying Semantic Web techniques to their existing metadata.

```
<?xml version="1.0" encoding="UTF-8"?>
<xc:frbr xmlns:xc="http://www.extensiblecatalog.info/Elements"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:rdvocab="http://rdvocab.info/Elements"
xmlns:dcterms="http://purl.org/dc/terms/" xmlns:rdarole="http://rdvocab.info/roles">
<xc:entity type="work" id="oai:mst.rochester.edu:MST/MARCToXCTransformation/10081">
<dcterms:subject xsi:type="dcterms:LCC">PS3505.U334</dcterms:subject>
<dcterms:subject xsi:type="dcterms:DDC">811/.52</dcterms:subject>
<dcterms:subject xsi:type="dcterms:DDC">B</dcterms:subject>
<rdarole:author>Sawyer-Lauc<U+0327>anno, Christopher, 1951-</rdarole:author>
<rdvocab:titleOfTheWork>E.E. Cummings :</rdvocab:titleOfTheWork>
<xc:subject xsi:type="dcterms:LCSH">Cummings, E. E. (Edward Estlin), 1894-1962.</xc:subject>
<xc:subject xsi:type="dcterms:LCSH">Poets, American-20th century-Biography.</xc:subject>
</xc:entity>
</xc:frbr>
```

FIG. 4A XC Schema Work Record

```
<?xml version="1.0" encoding="UTF-8"?>
<xc:frbr xmlns:xc="http://www.extensiblecatalog.info/Elements"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:rdvocab="http://rdvocab.info/Elements"
xmlns:dcterms="http://purl.org/dc/terms/" xmlns:rdarole="http://rdvocab.info/roles">
<xc:entity type="expression" id="oai:mst.rochester.edu:MST/MARCToXCTransformation/10082">
<xc:titleOfTheExpression>E.E. Cummings :</xc:titleOfTheExpression>
<rdvocab:illustrativeContent>ill. ;</rdvocab:illustrativeContent>
<xc:workExpressed>oai:mst.rochester.edu:MST/MARCToXCTransformation/10081</xc:workExpressed>
</xc:entity>
</xc:frbr>
```


FIG. 4B XC Schema Expression Record


```
<?xml version="1.0" encoding="UTF-8"?>
<xc:frbr xmlns:xc="http://www.extensiblecatalog.info/Elements"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:rdvocab="http://rdvocab.info/Elements"
xmlns:dcterms="http://purl.org/dc/terms/" xmlns:rdarole="http://rdvocab.info/roles">
<xc:entity type="manifestation" id="oai:mst.rochester.edu:MST/MARCToXCTransformation/10083">
<xc:recordID type="LCCN">2004012234</xc:recordID>
<xc:identifier type="bnb">GBA466907</xc:identifier>
<xc:identifier type="Uk">012991839</xc:identifier>
<xc:identifier type="YBP">99809225526</xc:identifier>
<xc:recordID type="OCoLC">55518224</xc:recordID>
<xc:recordID type="NRU">2726457</xc:recordID>
<rdvocab:statementOfResponsibilityRelatingToTitleProper>Christopher Sawyer-
Lauc<U+0327>anno.</rdvocab:statementOfResponsibilityRelatingToTitleProper>
<dcterms:title>E.E. Cummings : a biography /</dcterms:title>
<rdvocab:placeOfProduction>Naperville, Ill. :</rdvocab:placeOfProduction>
<dcterms:publisher>Sourcebooks,</dcterms:publisher>
<dcterms:issued>c2004.</dcterms:issued>
<dcterms:extent>xvi, 606 p., [32] p. of plates :</dcterms:extent>
<rdvocab:dimensions>24 cm.</rdvocab:dimensions>
<dcterms:description>Includes bibliographical references (p. [545]-588) and indexes.</dcterms:description>
<dcterms:tableOfContents>Their son -- In just- Spring -- Poet of simplicity -- A smattering of languages, a first
taste of independence, and the truest friends -- The graduate student as poet -- Poetic excursions -- La
Guerre -- Singing of Olaf and Elaine -- Poet, painter, and papa in New York -- The enormous room -- Abroad --
An American in Paris -- Tulips and chimneys -- Marriage and unmarriage -- The life and times of Edward Seul -
- Anne -- Him -- Viva Cummings -- Worlds of was (and is) -- Carefully into growing and into being and into
loving -- Nourish my failure with thy freedom -- Never the soft adventure of undoom -- Freedom as
breakfastfood -- Allegiance only to the imagination -- We know who we are -- Xaipe and controversy --
Ordeals by audience -- When such marvels vanish -- Steering for dream.</dcterms:tableOfContents>
<dcterms:description>MAG Stacks copy signed by the author, Christopher Sawyer-
Lauc<U+0327>anno.</dcterms:description>
<xc:expressionManifested>oai:mst.rochester.edu:MST/MARCToXCTransformation/10082</xc:expressionMa
nifested>
</xc:entity>
</xc:frbr>
```


FIG. 4C XC Schema Manifestation Record

5.2. Registered Vocabularies

The XC Schema enables the use of URIs for vocabulary terms from registered vocabularies by recording a URI as an attribute for a vocabulary term contained within appropriate schema elements for controlled vocabularies. The first version of XC software assumes that the actual term will also be recorded as text in an XC Schema record, with or without the URI. This assumption could change, however, so that URIs could be recorded simply as "pointers", without the inclusion of the actual vocabulary term in each metadata record (Coyle, 2010a, p. 30). Using URIs for terms without the terms themselves would facilitate the maintenance of controlled vocabularies because metadata would not need to be updated when the form of a term or name heading changes. Another possible scenario is that a metadata registry would actually facilitate the update of systems that use the elements and vocabularies by enabling version control and by retaining older versions of vocabulary terms, as described by Coyle (Coyle, 2010a, pp. 31-32). In either model, the use of registered vocabularies could reduce the need for much time-consuming metadata maintenance.

Tag	Ind 1	Ind 2	Bibliographic Data
000			02011cam·a2200373·a·4500
001			2726457
005			20080708160757.0
008			041229s2004····iluaf····b····001·0beng··
010			‡a 2004012234
035			‡a (OCoLC)ocm55518224
035			‡a 2726457
040			‡a DLC ‡c DLC ‡d YDX ‡d UKM ‡d WIQ ‡d YBM
015			‡a GBA466907 ‡2 bnb
016	7		‡a 012991839 ‡2 Uk
020			‡a 1570717753 (alk. paper)
024	7		‡a 99809225526 ‡2 YBP
043			‡a n-us---
049			‡a RRRR
050	0	0	‡a PS3505.U334 ‡b Z84 2004
082	0	0	‡a 811/.52 ‡a B ‡2 22
100	1		‡a Sawyer-Lauçanno, Christopher, ‡d 1951-
245	1	0	‡a E.E. Cummings : ‡b a biography / ‡c Christopher Sawyer-Lauçanno.
260			‡a Naperville, Ill. : ‡b Sourcebooks, ‡c c2004.
300			‡a xvi, 606 p., [32] p. of plates : ‡b ill. ; ‡c 24 cm.
504			‡a Includes bibliographical references (p. [545]-588) and indexes.
505	0		‡a Their son -- In just- Spring -- Poet of simplicity -- A smattering of languages, a first taste of independence, and the truest friends -- The graduate student as poet -- Poetic excursions -- La Guerre -- Singing of Olaf and Elaine -- Poet, painter, and papa in New York -- The enormous room -- Abroad -- An American in Paris -- Tulips and chimneys -- Marriage and unmarriage -- The life and times of Edward Seul -- Anne -- Him -- Viva Cummings -- Worlds of was (and is) -- Carefully into growing and into being and into loving -- Nourish my failure with thy freedom -- Never the soft adventure of undoom -- Freedom as breakfastfood -- Allegiance only to the imagination -- We know who we are -- Xaipe and controversy -- Ordeals by audience -- When such marvels vanish -- Steering for dream.
590			‡a MAG Stacks copy signed by the author, Christopher Sawyer-Lauçanno.
600	1	0	‡a Cummings, E. E. ‡q (Edward Estlin), ‡d 1894-1962.
650		0	‡a Poets, American ‡y 20th century ‡v Biography.
945			‡a PromptCat
949			‡b 39087027239823
980			‡a 2005-01-12 ‡b 29.95 ‡e 25.16 ‡f 519971 ‡g 1
994			‡a 92 ‡b RRR
945			‡a MARS

FIG. 5 MARC Source Record for Figures 3A-3C

By using URIs in their metadata for those vocabularies terms that already have them, and registering other vocabularies themselves, libraries can increase the percentage of library-related vocabularies whose terms can be expressed as URIs. The NSDL Metadata Registry ("The Registry!" n.d.) enables libraries to self-register their vocabularies easily and at no cost, as described above. As more vocabularies are registered, libraries can use the XC Metadata Services Toolkit to populate metadata records with the URIs for their registered vocabularies.

5.3. Enabling Experimentation

The library community began a formal discussion of the use of URIs for controlled vocabularies within MARC records in early 2009. As of January 2010, no decision had been made to enable the encoding of URIs for controlled values within the MARC formats. As of this writing, the current status of the issue as follows:

> 1/17/10 - Results of MARC Advisory Committee discussion: Some participants were reluctant to experiment with encoding URIs in MARC records because of the large amount of effort for systems to support experimentation. This includes questions about how to explain, what to get back, how to define the relationship between a value and a URI. Some were interested in experimenting with a set of test records. Nothing will be finalized on this until issues are sorted out, but a document will be prepared with some guidelines and examples of how URIs might be used in MARC records so that those that wish to may experiment (RDA/MARC Working Group, 2009, Status/Comments).

While the MARC Advisory Committee has approved the definition of $u for URIs in some MARC fields, these URIs have been included for a purpose other than enabling the use of URIs as identifiers. Instead, the goal of their inclusion is "… facilitating user access to online information that is more current and easily maintained outside of the MARC record" (Bibliographic Standards Committee, ALA/ACRL/Rare Books and Manuscripts Section, 2008). It is unclear whether the current technique for recording these URIs will enable them to be used as identifiers.

Until the MARC formats allow for the inclusion of more URIs in MARC records, and for a URI to be recorded in a manner that will enable it to act as an identifier for a specific data element, libraries will be unable to move forward with experimenting with linked data in a MARC environment. On the other hand, XC provides a metadata platform that can already accommodate the automated application of URIs for registered vocabularies to metadata that originated as MARC. Because there is a considerable learning curve associated with linked data, libraries need to try out new techniques in a low-risk environment. They need to be able to start over, or make changes to their metadata, as the collective knowledge of applying Semantic Web practices to library data evolves. XC's software architecture enables libraries to engage in iterative processing of their metadata, without risking corruption of the original metadata. Libraries can use XC to experiment with linked data now, without waiting for such experimentation to be approved for a MARC environment, and without investing anything except for the time that they spend experimenting. XC can therefore lower the barriers to participating in linked data and enable libraries to move forward more quickly in this area.

6. Making Linked Data a Priority in Software Development

By allowing library metadata to be used independently of a particular online environment and alongside metadata from any number of non-library sources, linked data provides potential functionality for software applications that could hardly be imagined a few years ago. As described in this paper, the XC team did not begin with a goal of enabling Semantic Web-based metadata applications. Instead, we added this functionality to XC as we began to understand both the opportunities offered by linked data and the methodology needed to implement it. By describing some of the project management issues that this change created, we hope to encourage other existing projects to incorporate linked data as well.

6.1. Prioritizing Development Work

Because linked data has been an "add on" to XC, it has been necessary to balance its implementation in XC with other development goals. This has been difficult when enabling the application of linked data functionality within XC would take longer than NOT implementing linked data functionality. In such cases, adhering to the principles of linked data can sometimes be seen as an expendable feature when time is short. The extra time that it has taken us to implement linked data relates not to the complexity of linked data itself (the principles behind it are actually quite simple) but to other factors: the learning curve needed to learn about this new concept, the lack of existing models for its application in the library world, and the complexity of mapping legacy metadata to this new environment. As linked data is used more and more in library applications, these issues should be less of an obstacle for future projects.

The XC team has balanced our seemingly-conflicting project goals by focusing upon developing a robust software platform for metadata processing, and by ensuring that we do nothing that would prevent the application of linked data in the future. Because the XC metadata platform enables low-risk experimentation with metadata, others will likely develop more well-informed linked data applications using XC in the future than what we can achieve now. Our focus upon developing a robust platform may delay some of the formal output products related to implementing linked data in XC, such as a formal XC application profile and fully-registered XC elements and vocabularies. However, once the XC software itself has been fully tested and implemented in various library environments, we will also complete these related tasks.

6.2. Internal Use vs. Output of Linked Data

Many online systems do not maintain metadata standards within the systems themselves, but store metadata internally in relational databases rather than in record-based structures. As long as these systems enable metadata output in a standard schema or format, the system's internal metadata storage does not matter. The various XC software toolkits store metadata using a variety of methods, including MySQL and Solr. However, each XC software toolkit component contains its own OAI-harvestable repository. Each software component provides metadata that can be harvested by other XC components or by other applications, which provides considerable flexibility for using various XC software components with other applications. Because of this architecture, XC does not have the luxury of foregoing the use of a standard schema within the system itself. Instead, XC enables output in one or another XML schema at various steps within the system itself.

XC's architecture has encouraged us to adhere closely to standards for metadata harvesting, but not necessarily to implement Semantic Web technologies within an end-to end XC system. While it is a requirement for XML data to be expressed as RDF triples in order for it to participate within the linked data community (Coyle, 2010b, p. 32), XC's system components do not require RDF, nor does the XC user interface need RDF for its functionality at this time. As noted by the Variations/FRBR team, libraries currently operate in an XML rather than an RDF environment, which has led that project team to move their exploration of other metadata data presentations (including RDF) to the future ("FRBR XML Schemas released," 2010). We have made a similar decision, to make RDF output of XC Schema metadata a slightly longer-term deliverable. However, we are looking forward to finding out how well we have done to make XC metadata "RDF-ready".

7. Conclusions

An increasing number of proponents are now urging the library community to move forward with implementing linked data, and an increasing number of library-related applications are beginning to take advantage of linked data. However, a paradox exists here that is preventing this process from happening quickly. Libraries are tied to MARC-based systems that do not yet facilitate the creation of linked data. Without a body of library data converted to linked data, software developers have little incentive to create new applications that require it. And without a significant number of applications that take advantage of linked data, vendors of current systems have little incentive to implement linked data in a legacy environment. To move beyond this chicken-and-egg stalemate, the eXtensible Catalog facilitates converting legacy data to linked data outside of the current MARC environment, so that library data can function in non-MARC-based discovery applications as well as in the broader web environment.

In addition to creating software to assist libraries in converting legacy data to linked data, the eXtensible Catalog Project can provide the following "lessons learned" to the library community regarding the conversion of legacy data to linked data:

- Some MARC metadata can be converted to linked data fairly easily using batch processing. Although a significant amount of data in a MARC record cannot be converted to linked data without record-by-record editing, enough basic MARC data can be converted to identify a resource and its basic relationships, thus making the conversion process worthwhile.
- Libraries must make well-informed decisions about what information within a MARC record is worthy of cleanup and enrichment, whether automated or manual. The more experience that libraries gain with linked data, the more they will be able to make such well-informed decisions, enabling efforts to be targeted toward those aspects of legacy data that are actually reusable as linked data.
- Converting legacy metadata to linked data will require a team of experts, including MARC-based catalogers, specialists in other metadata schemas, software developers, and Semantic Web experts to design and test normalization/conversion algorithms, develop new schemas, and prepare individual records for automated conversion.

Library software applications that do not depend upon linked data to accomplish their goals may currently have little incentive to enable the use of linked data. However, this situation can change in a heartbeat: as soon as an exciting new application becomes available that needs linked data in order to function, the availability of such data will quickly become a high priority. Given recent advances in registering legacy library vocabularies, the availability of national library catalogs as linked data, and the recent availability of open source software such as XC to convert legacy data to linked data, libraries are likely to soon find it difficult to justify continuing to create metadata that is not linked-data compliant.

Making the case for linked data in the library world should become easier over time, as libraries begin to understand its value toward enabling Semantic Web-based applications. In the meantime, the library community can begin to propose smart practices for using linked data, and should encourage the developers of systems for libraries – both commercial and open source - to incorporate linked data. XC is helping to demonstrate that implementing linked data, and converting legacy library data to linked data, are indeed achievable goals.

References

Baker, Thomas and Pete Johnston. (2008, January 14). The Singapore Framework for Dublin Core Application Profiles. Retrieved March 25, 2010, from http://dublincore.org/architecturewiki/SingaporeFramework/.

Berners-Lee, Tim. (2006, July 27). Linked data: Design issues. Retrieved March 20, 2010, from http://www.w3.org/DesignIssues/LinkedData.html.

Bibliographic Standards Committee, ALA/ACRL/Rare Books and Manuscripts Section. (2008). Adding subfield $u for Uniform Resource Identifier to field 510 (Citation/References note) of the MARC 21 Bibliographic Format. MARC Proposal No. 2009-05. Retrieved April 3, 2010, from http://www.loc.gov/marc/marbi/2009/2009-05.html.

Bowen, Jennifer. (2008). Metadata to support next-generation library resource discovery: Lessons from the eXtensible Catalog, Phase 1. Information Technology and Libraries, 27(2), 6-19. Retrieved March 27, 2010, from https://urresearch.rochester.edu/handle/1802/5757.

Bowen, Jennifer. (2009). Supporting the eXtensible Catalog through metadata design and services. Retrieved March 27, 2010, from http://hdl.handle.net/1802/6377.

Bradley, Fiona. (2009). Discovering linked data. Library Journal, 134(7), 48-50.

Chudnov, Dan. (2009a April). The Geography of linked data and ready reference. Computers in Libraries, 29(4), 26-28.

Chudnov, Dan. (2009b May). Connecting linked data, OPACs, and online exhibits. Computers in Libraries, 29(5), 25-27.

Chudnov, Dan. (2009c September). In Which "linked data" means "a better web". Computers in Libraries, 29(8), 31-33.

Coyle, Karen. (2006, September 1). Murdering MARC. Coyle's InFormation. Retrieved March 23, 2010, from http://kcoyle.blogspot.com/2006/09/murdering-marc.html.

Coyle, Karen. (2008, May). Meaning, technology, and the Semantic Web. Journal of Academic Librarianship, 34(3), 263-264.

Coyle, Karen. (2009 April 15). Making connections. Library Journal, 134(7), 44-47.

Coyle, Karen. (2010a January). Understanding the Semantic Web: Bibliographic data and metadata. Library Technology Reports, 46(1). Chicago: American Library Association.

Coyle, Karen. (2010b Feb./Mar.). RDA vocabularies for a twenty-first-century data environment. Library Technology Reports, 46(2). Chicago: American Library Association.

Coyle, Karen. (2010c, March). Cataloging horizons. American Libraries, 41(3), 26.

Coyle, Karen. (2010d, March 5). MARC: From mark-pp to data. Coyle's InFormation. Retrieved April 2, 2010, from http://kcoyle.blogspot.com/.

DCMI. (2008, January 14). DCMI metadata terms. Retrieved April 7, 2010, from http://dublincore.org/documents/dcmi-terms/.

Dunsire, Gordon. (2010, February 24). Re: FRBRer & FRAD in Registry. JISCmail - DC-RDA Archives. Retrieved March 21, 2010, from https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind1002&L=DC-RDA&T=0&F=&S=&P=3612.

eXtensible Catalog Organization. (2010). Retrieved March 20, 2010, from http://www.extensiblecatalog.org/?page_id=28.

FRBR XML Schemas released. (2010, March 15). Variations/FRBR Project News. Retrieved April 2, 2010, from http://www.dlib.indiana.edu/projects/vfrbr/projectDoc/admin/news.shtml.

Greenberg, Jane, and Eva Mendez. (2007). Knitting the Semantic Web [Special issue]. Cataloging & Classification Quarterly, 43(3-4).

Harper, Corey A. (2009, November 13). Linking library data: Exposing our bibliographic heritage to the Semantic Web. Presented at the LYRASIS Bibliographic Services Conference, Worchester, MA. Retrieved April 7, 2010, from http://www.lyrasis.org/Classes-and-Events/~/media/Files/Lyrasis/Classes%20and%20Events/charper%20lyrasis%2020091113.ashx.

Hillmann, Diane, Karen Coyle, Jon Phipps, and Gordon Dunsire. (2010). RDA Vocabularies. D-Lib Magazine, 16(1/2). doi:10.1045/january2010-hillmann. Retrieved March 20, 2010, from http://www.dlib.org/dlib/january10/hillmann/01hillmann.html.

Horvath, Adam. (2010, April 12). Hungarian National Library published its entire OPAC and Digital Library as linked data. Retrieved June 30, 2010 from http://www.mail-archive.com/public-lod@w3.org/msg05008.html.

IFLA Study Group on the Functional Requirements for Bibliographic Records (1998). Functional requirements for bibliographic records. Munich: K.G. Saur. Accessed July 10, 2010 from http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records.

Library of Congress. Network Development and MARC Standards Office. (2006, April 6). Functional analysis of the MARC 21 bibliographic and holding formats. Retrieved March 25, 2010, from http://www.loc.gov/marc/marc-functional-analysis/functional-analysis.html.

Library of Congress. Network Development and MARC Standards Office. (2010). MARC 21 Format for bibliographic data. 1999 ed.. Washington, D.C.: Library of Congress. Retrieved March 27, 2010, from http://www.loc.gov/marc/bibliographic/ecbdhome.html.

LIBRIS Available as Linked Data at semweb@libris. (2008 December 3). Retrieve June 30, 2010 from http://blog.libris.kb.se/semweb/?p=7.

Lindahl, David. (2010). eXtensibleCatalog's Library. Retrieved March 20, 2010, from http://www.screencast.com/users/eXtensibleCatalog.

Linked Data: Connect Distributed Data across the Web. (n.d.). Retrieved March 20, 2010, from http://linkeddata.org/.

MARCXML to XC Transformation Service Documentation, eXtensible Catalog (2009 January 7). Retrieved June 29, 2010 from http://dev2.xcproject.org/learnmore/publications/marcxml-xc-transformation-service-documentation.

Pohl, Adrian. (2010 April 14). Linked open data prototype at the German National Library. Retrieved June 30, 2010 from http://permalink.gmane.org/gmane.culture.libraries.ngc4lib/7544.

RDA Relationships for Works, Expressions, Manifestations, Items: Show Detail. (n.d.). Retrieved March 25, 2010, from http://metadataregistry.org/schema/show/id/13.html.

RDA/MARC Working Group. (2009). Encoding URIs for controlled values in MARC records. MARC Discussion Paper No. 2010-DP-02. Retrieved April 2, 2010, from http://www.loc.gov/marc/marbi/2010/2010-dp02.html.

RDA: Resource Description and Access. Constituency Review. (2008, October 31). Retrieved March 25, 2010, from http://www.rdaonline.org/constituencyreview/.

RDA Toolkit: Resource Description & Access (2010. Retrieved July 10, 2010 from http://access.rdatoolkit.org/index.php.

Smith-Yoshimura, Karen, Catherine Argus, Timothy J. Dickey, Chew Chiat Naun, Lisa Rowlison de Ortiz, and Hugh Taylor. (2010). Implications of MARC tag usage on library metadata practices. Dublin, Ohio: OCLC Research and the RLG Partnership. Retrieved April 1, 2010, from http://www.oclc.org/research/publications/library/2010/2010-06.pdf.

Spero, Simon. (2010 February 23). MARCTHULU. Lightning Talk presentation at code4lib 2010, Asheville, North Carolina. Retrieved April 3, 2010, from http://www.archive.org/details/MARCTHULU.

Summers, Ed, Antoine Isaac, Clay Redding, and Dan Krech. (2008). LCSH, SKOS and linked data. In Metadata for Semantic and Social Applications: Proceedings of the 8. International Conference on Dublin Core and Metadata Applications, 22-26 September 2008: DC-2008 Berlin, 25-33. Retrieved April 7, 2010, from http://dcpapers.dublincore.org/ojs/pubs/article/view/916.

Tennant, Roy. (2002a). MARC must die. Library Journal, 127(17). Retrieved April 3, 2010, from http://www.libraryjournal.com/article/CA250046.html.

Tennant, Roy. (2002b). MARC exit strategies. Library Journal, 127(19), 27-28. Retrieved April 3, 2010, from http://www.libraryjournal.com/article/CA256611.html?q=tennant+exit.

The Registry! :: About. (n.d.). . Retrieved March 25, 2010, from http://metadataregistry.org/about.html.

The Registry! :: XC Elements :: Type from MARC 21 007 :: Show Detail (n.d.). Retrieved June 29, 2010 from http://sandbox.metadataregistry.org/schemaprop/show/id/22.html.

Yee, Martha. (2009). Can bibliographic data be put directly onto the Semantic Web? Information Technology and Libraries, 28(2), 55-80. Retrieved June 30, 2010 from http://escholarship.org/uc/item/91b1830k?query=yee%20semantic.