Proposal to Explore a Schema for Rare Materials Description

Submitted 15 June 2015 to the Bibliographic Standards Committee,
Rare Books & Manuscripts Section, Association of College and Research Libraries,
American Library Association

Summary

This paper explains the need for a schema to encode description of rare materials. The schema solution is envisioned as a companion to the content standards produced by the Rare Books and Manuscripts Section (RBMS) of the Association of College and Research Libraries (ACRL), in order to make this content machine-actionable and meet the discovery needs of rare materials users. A Bibliographic Standards Committee (BSC) task force is proposed to explore needed data elements and recommend a schema solution.

Background

The BSC is charged with enabling intellectual access to rare materials. The BSC has produced two content standards for rare materials description that reflect the specialized research needs of rare book and manuscript users:

1. *Descriptive Cataloging of Rare Materials* (DCRM): a suite of guidelines for resource description
2. *Controlled Vocabularies*: a taxonomy of rare materials features and genres, and a vocabulary for indexing these features and genres

The DCRM suite and the *Controlled Vocabularies* are sophisticated content standards, the use of which produces granular metadata regarding works, editions, and individual objects in many formats. Recent discussions have elicited strong community interest in aggregated, cross-collection access to this metadata.

Several BSC initiatives are underway to ensure that rare materials description is compatible with contemporary data models and technologies. For examples: The Descriptive Cataloging of Rare Materials Task Force is aligning DCRM with *Resource Description and Access* (RDA). The Controlled Vocabularies Editorial Group is working to produce the *Controlled Vocabularies* in Linked Data formats, in support of the Bibliographic Framework Initiative (BIBFRAME). A proposal to the American

Library Association (ALA) Committee on Cataloging: Description and Access (CC:DA) requests that references to descriptions are appropriately modeled as relationships between resources.

In sum, the BSC supports rare materials cataloging and discovery, and is committed to enabling this activity using Web and Linked Data technologies.

Problem

There is a gap between the granular content within rare materials catalogs and our ability to encode this metadata for machine-action. Among the encoding schemas which RBMS members employ, none fully reflect our content standards, nor the research needs of our users. The MARC Standards and BIBFRAME are designed for general-purpose libraries. Encoded Archival Description (EAD) is designed for archival collections. The Text Encoding Initiative (TEI) is designed for manuscripts. Schema.org's "Creative Work," is designed for the commercial sector. Each of these schema options include elements relevant to rare materials description, but none are designed to encode the full content resulting from the DCRM suite and the *Controlled Vocabularies*.

Without a schema designed specifically for rare materials description, two major problems arise:

1. inconsistent encoding, which diminishes data interoperability
2. high-level encoding, which diminishes discovery of granular content

Examples of inconsistent encoding in the MARC formats include:

- Location & local call numbers
    - 852 ## |a Institutional location |b Collection |c Shelving location VS
    - 099 ## |a Call number
- Subdivision of genre terms
    - Leather bindings (Binding) |z France |y 19th century VS
    - Leather bindings (Binding) |x France |x 19th century VS
    - Leather bindings (Binding) -- France -- 19th century
- Acquisition and provenance notes
    - 541 Immediate Source of Acquisition Note VS
    - 561 Ownership and Custodial History VS

    - 796/797 Donor notes
- Binding notes
    - 563 ## |a Note |5 Institution VS
    - 59x ## |a Note (for local copy)

The last example above illustrates the problem of high-level encoding:

- Notes
    - 500 (General note) OR 59x (Local notes) VS
    - 5xx More specific note fields

Inconsistent and high-level encoding make it difficult to index rare materials for discovery platforms. The end result is that most discovery platforms fall short of providing access to the granular content of rare materials description, and the desired cross-collection access is severely hindered.

Solution

To encode (and by extension, to build access to) the granular content of rare materials catalogs, the RBMS community needs a specialized encoding schema. A schema designed for rare materials will diminish the problems of inconsistent and high-level encoding. A schema designed for rare materials will enable development of discovery platforms and cross-collection access.

Given the work of the Bibliographic Framework Initiative to transition libraries away from the MARC formats, this is an opportune time to develop a schema solution.

This solution could be executed in several ways:

1. As a standalone schema, produced by RBMS
2. As an extension of an existing schema (e.g., BIBFRAME and/or the RDA Registry), recommended by RBMS
3. As an amalgamation of elements from existing schema, with mappings maintained by RBMS

Communities beyond RBMS will also have an interest in this solution. Likely partners include the Chartered Institute of Library and Information Professionals

(CILIP) Rare Books and Special Collections Group, the Library of Congress Bibliographic Framework Initiative, and the RDA Registry.

Proposal

A Bibliographic Standards Committee task force is proposed to explore needed data elements and recommend a schema solution. The task force will be charged with the following:

1. To determine a list of data elements complementary to the DCRM suite, the *Controlled Vocabularies*, and rare materials user needs
2. To compare this list with existing schema, and to determine areas of overlap and/or need
3. To recommend a schema solution (see the categories above)
4. To seek community partners for this solution

Conclusion

A consistent and granular schema for rare materials description will enable improved discovery of our collections. It will enable the full potential of content built with the DCRM suite and the *Controlled Vocabularies.* It is the logical outgrowth of these existing BSC standards.

The proposed task force will complement the work to align the DCRM suite with RDA, and to produce the *Controlled Vocabularies* as Linked Data. Together, these three initiatives will enable machine-actionable data in catalogs for rare materials, and will enable Web-based discovery applications that are appropriate for rare materials users.

BSC Contact

Allison Jai O'Dell, AJODELL@ufl.edu