RDA and the Semantic Web, Linked Data Environment

Barbara Tillett

Connecting to friends, colleagues, customers, and others on the Internet is an everyday experience for most people these days. We use email, Twitter, Facebook, and other social networking systems quickly and easily when there is wifi or an Internet service provider that reaches our geographic location – even as we move around. This change in our communication systems even extends sometimes to replacing phone calls with communications like Skype or Facetime. A former phone communication can now be a multi-media experience where you not only talk but also see each other (or groups of people), share pictures or videos or documents quickly and easily all at the same time. Where are libraries in this world?

Our collections are being digitized and information about our traditional and digital collections is being made accessible through the Web. Our bibliographic descriptions and the information that we provide about the people, families, and corporate bodies used in our descriptions is available for re-use. This has been demonstrated through the popular Virtual International Authority File (VIAF) and various linked data projects, such as the Library of Congress posting of the Library of Congress Subject Headings (LCSH) and other controlled vocabularies on the Web at id.loc.gov. These projects also have shown that the data libraries provide can be very useful in

a linked data environment for mashups and pathways to related information that may be of interest to the Web searcher – either through showing them added facets they may wish to consider to refine their search or suggesting new directions or related resources they may also like to see. Library data about our resources is no longer just meant to be stored in catalog drawers as the inventory tool to access an individual library's collections. It can now be made available to anyone, anywhere in the world, at any time.

Libraries share data because they find they collect similar things and can re-use the descriptions of the manifestations collected by other libraries or re-use the authority data about a person, family, or corporate body they also have represented in their collections. But unlike the start of sharing bibliographic and authority data among libraries (mostly by exchanging MARC formatted records or before that through buying catalog cards for the new acquisitions), the data can be shared globally rather than "exchanged" for redundant storage locally. That is, it can be made available in a linked data environment, so libraries do not need to replicate the same data over and over, but instead share it mutually with each other and with others using the Web. This environment helps reduce the library costs and expands the accessibility of information. Library data need not be just in the form of a citation in a bibliography or other linear listing, but the descriptive and authority data can be re-used and packaged in creative new ways that can be context-appropriate to a user's needs.

Library data is dynamic. It can actually start with the descriptive data and identifying information that the creator of a work provides, augmented by information from a rights management organization, a publisher or manufacturer or distributor, further enhanced by a cataloger to provide a classification and/or subject terms to help find that resource, while others may add more content information

or expand the relationships, and maybe even later the data can be enhanced by a scholar with some special expertise or knowledge about the resource or the topic it covers. Many other related resources can be connected to the description and pathways emerge in the bibliographic universe. The role of libraries in this chain of description and access is an important one to help organize information and to connect users to the information they are looking for. We are part of the information network and have a tremendous legacy to contribute. Since the late 1990's, the library world has worked towards increased sharing based on agreed international principles (IFLA's International Cataloguing Principles (*ICP*)), internationally accepted conceptual models with their user tasks and recommended mandatory data elements (IFLA's FRBR (*FRBR*) and FRAD (*FRAD*)) as the foundation for how we are re-visualizing our descriptions of our resources. These changes caused us to re-examine our cataloging standards, like ISBD (*ISBD*), and our cataloging codes, like the Anglo-American Cataloguing Rules (AACR) and led to changes to position libraries for the future. Certainly the reconceptualization of the AACR was a huge milestone to move libraries towards seeing their descriptions as more than citations intended only for a library catalog. The move to Resource Description & Access (RDA) is a move towards an internationally shared vision and internationally maintained content code intended for the digital environment that we now find ourselves in.

Resource Description & Access (RDA)

Building on the foundations for the IFLA conceptual models, especially of the Functional Requirements for Bibliographic Records (FRBR) user tasks, national level bibliographic record mandatory elements, and the entity-relationship conceptual model, and the

Functional Requirements for Autority Data (FRAD) expansion of that model, RDA instructs us how to build well-formed metadata that can be re-used for multiple purposes. RDA is also built on the foundations of the Statement of International Cataloguing Principles (ICP), and work continues to align the instructions into closer harmony with those principles through discussions with various communities, like the music library world, to adjust the Anglo-American Cataloguing Rules, 2nd edition (AACR2) rules that were carried forward into RDA into truly FRBR/ICP principle-based instructions, rather than case law of the past. This new cataloging code is a major shift in perspective to make our bibliographic and authority data more useful in today's linked data environment and beyond. Unlike codes of the past, RDA ties the description and access to the user tasks and the specific entities that we are describing. RDA enables us to identify the resource no matter what its format – book, sound recording, score, motion picture, game, map, photograph, data file, whatever. All resources share some basic identifying information, and we can add more identifying characteristics and relationships as needed. Yet RDA has ties with the codes of the past in the objectives to collocate all the works of a creator, all the expressions of a work, all the manifestations of an expression, all the items held. RDA enables us to indicate other related resources and information, to identify the persons, families, corporate bodies, and places, so we can collocate things by or about those entities. It enables us to link to other information about those entities on the web. Such information can then be used for multiple purposes, like creating mashups, as shown with VIAF, Library Thing, and several other services; providing users with useful groupings of information and links to related things of interest. RDA moves us beyond the vocabulary of catalog cards (main entry, headings, see references etc.) to more clearly label the identifying characteristics of

and relationships among entities to provide pathways for users and machines. It uses controlled vocabularies and makes them available in a linked data environment (through the Open Metadata Registry), so the Uniform Resource Identifier (URI) for each term can be used and identified with the various language versions of the term for the value being labeled.

Having URIs also can apply to standard phrases, like [publisher not identified], or [place of publication not identified]. One would hope future input systems for cataloging will suggest such standard terms through drop down menus or other devices, so catalogers would not need to key in that standard information. Such features of RDA and the Web-based vocabularies are intended to facilitate the development of future systems to make cataloging easier and to enable multilingual displays for users around the world. RDA also encourages the re-use of descriptive data from publishers, offering us the option to transcribe, that is, take data as we see it on the resource being described or from the data that accompanies the resource coming from the publisher. RDA avoids abbreviating so the user can understand what we're trying to tell them, in line with the ICP.

The new Bibliographic Framework Initiative from the Library of Congress recognizes that our past methods of communicating bibliographic and authority data through exchanging records are in need of an overhaul. With the help of interested parties and experts from around the world, we have started the process of rethinking our approach, exploring the possibilities in this Semantic Web, linked data environment. As we explore the new possibilities and build an internationally shared vision, we must always keep our users foremost in mind, as they are our customers, the reason we collect resources, the reason we organize information. We will continue to work towards collaborative approaches to reduce cataloging costs

and to enrich user experiences throughout the world as they seek information in our bibliographic universe.

References

International Federation of Library Associations and Institutions. ISBD Review Group and International Federation of Library Associations and Institutions. Cataloguing Section. Standing Committee. *ISBD International Standard Bibliographic Description*. Berlin: De Gruyter Saur, 2011. (Cit. on p. 141).

International Federation of Library Associations and Institutions. Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records*. München: Saur, 1998. http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf. (Cit. on p. 141).

Patton, Glenn E., ed. *Functional requirements for authority data: a conceptual model: final report december 2008: approved by the Standing committees of the IFLA cataloguing section and IFLA classification and indexing section*. München: Saur, 2009. (Cit. on p. 141).

Tillett, Barbara and Ana Lupe Cristán, eds. *IFLA Cataloguing Principles: the Statement of International Cataloguing Principles (ICP) and its Glossary in 20 languages*. München: Saur, 2009. (Cit. on p. 141).

BARBARA TILLETT, Library of Congress.
btil@loc.gov

ABSTRACT: Connecting to friends, colleagues, customers, and others on the internet is an everyday experience for most people these days. We use email, Twitter, Facebook, and other social networking systems quickly and easily when there is wifi or an internet service provider that reaches our geographic location – even as we move around. This change in our communication systems even extends sometimes to replacing phone calls with communications like Skype or Facetime. A former phone communication can now be a multi-media experience where you not only talk but also see each other (or groups of people), share pictures or videos or documents quickly and easily all at the same time. Where are libraries in this world?