RDA and Linked Data: Moving Beyond the Rules
Sarah Hess Cohen, Florida State University

In the twenty-first century information landscape, everything is interconnected. Jenn Riley (UNC-Chapel Hill) and Kimmy Szeto (SUNY Maritime College) presented a session which explored these connections and their implications for libraries, which will need to maintain and share data beyond their own walls, and indeed beyond the scope of the traditional library community. Riley began the session with an overview of the Semantic Web (which she described as an "authority file on steroids") and Linked Data, which uses URIs for identification. These URIs can then be used to discover more information about a thing, and also link to other URIs, which can lead to still further discoveries. RDA data elements could be the basis of machine interoperability of data in a Linked Data environment.

Szeto delved into a technical discussion of Linked Data, introducing the Resource Description Framework. This provides the structure to draw relationships between elements. An RDF statement "triple" is three ordered parts -- for example, two entities connected by a relationship, such as "resource – has property – value." The URI references can apply to things, classes of things, or properties. These URIs have human-readable labels.

RDA is a descriptive standard independent of metadata coding. RDA elements can be expressed in RDF as properties. RDA properties are designed to work with FRBR entities. Szeto gave an example in xml for manifestation data. By using this we can go beyond library data to participate in the wider information community. This framework can take a property from another source, such as GeoNames, DBpedia or Last.fm and use it in the description.

In the Semantic Web/Linked Data environment, which looks at the information world as a graph, the concept of a "record" isn't really meaningful. The "open world" assumption is that there can always be more information. In this environment we will be using many vocabularies, with connections between them, so we should expect that implementations will deal with data from multiple sources.

In order for this to become reality, library systems will first need improved infrastructure. We will need ways to identify trusted data sources, ways to find properties and classes defined by others, best practices for data caching, actual shared cataloging, and of course, better systems for data creation, management, sharing, and exposure. The library community should determine where best data creation happens, so that we can concentrate our attentions there.

The library community is using the Open Metadata Registry (http://www.metadataregistry.org) to share RDA values and element sets, although this initiative has not yet been endorsed by the JSC. Authority and bibliographic data and vocabularies are being exposed as Linked Data. There is also the W3C Linked Library Data Incubator Group, the Stanford Linked Data Technology Plan, and the LC Bibliographic Framework Transition Initiative. Meanwhile, there are music Linked Data initiatives outside the library community, such as Last.fm, BBC Music, MusicBrainz, Discogs, DBtune, and Magnatune. With these sources already in place, libraries can be Linked Data consumers as well as publishers, connecting us with the greater information community.