April 7, 2015 PRESENTATION TO ACCM

By Rhonda Super

Slide 1: We are all aware libraries are undergoing major changes. We are leaving MARC and transitioning into RDA. The purpose for this transition is to make our bibliographic information accessible on the Semantic Web.

“…the role of the library is increasingly moving from a focus on ‘stored knowledge’ to one where ‘smart knowledge’ is paramount.” Libraries will increasingly be linking and connecting to content, which means that local collections will need to be placed within a larger national, and even international, context1.

BIBFRAME is a tool under development by the Library of Congress to help us bring this about.

1Jisc. (2015, *March* 17). *Report: A national monograph strategy roadmap*. Downloaded March 30, 2015 from: <http://www.jisc.ac.uk/reports/a-national-monograph-strategy-roadmap>.

Slide 2 The Web, as it is familiar to us, uses the HTTP protocol to retrieve information resources. Documents, images, audio and video files are examples of information resources.

Slide 3 Information pulled from our MARC records using our OPAC interface is based on this model.

Slide 4 The Semantic Web uses the HTTP protocol to identify real world, non-information resources and the relationships between resources and non-information resources. People, places, abstract concepts, and relationships are examples of non-information resources.

Slide 5 The goal of the Semantic Web is to make data available free of proprietary software, single social networks, or web application so that data can be pulled from different sources and relationships between data resources and non-resources can be shown.

Slide 6 One way for us to do this is by connecting our records to the global Web of Data.

This is the Linking Open Data (LOD) Cloud Diagram. At the center is DBpedia, the Linked Data version of Wikipedia. The colors on the diagram represent broad topic areas. Library and cultural institutions are in the green area.

Slide 7 Linked Data refers to a set of best practices and techniques for publishing and connecting structured data on the Semantic Web using international standards of the W3C.

Data that conforms to these practices and techniques are also called Linked Data.

Linked Data relies on HIGH-QUALITY METADATA MANAGEMENT in order to enhance data usage.

Slide 8 RDF (Resource Description Framework) is the model used to implement Linked Data on the Semantic Web. As you can see, it consists of layers of technology including:

* + - Explicit metadata: They allow web pages to carry their meaning on their sleeves (openly, publicly),
    - Ontologies: They describe the main concepts of a domain and their relationships,
    - Logical reasoning: it makes it possible to draw conclusions from combining data with ontologies.

TRUST is also a major component of the Semantic Web: it relies on folks providing good, accurate information.

Slide 9 RDF is a graph technology.

The basic building block is a simple statement called a triple made up of a subject, predicate, and object. HTTP unique identifiers are assigned to subjects, predicates, and objects that have unique identifiers. Objects can also take the form of text which do not have unique identifiers, and when this happens the object is known as a typed literal.

Subjects and objects with HTTP identifiers are displayed as circles, literals are shown in rectangles, and predicates are the arrows, or arcs between them.

The predicates denote the relationship between a subject and an object. Arrows denoting predicates point away from the subject and towards the object.

A set of triples is called an RDF graph.

Slide 10 Libraries around the world are contributing datasets to the LOD Cloud. You can access these datasets on datahub.

Slide 11 The Library of Congress’s Authorities and Vocabularies are now linked in the Cloud and LC provides a search interface so you can access their unique identifiers.

Slide 12 OCLC administers the Virtual International Authority File (VIAF). They are providing a unique identifier which links the authority records of National libraries throughout the world who have submitted their records to VIAF.

Slide 13 OCLC is building tools and services based on these datasets.

Slides 14 Scroll down a Wikipedia page and you will see these VIAF and National Library identifiers are now available on Wikipedia pages …

And 15 thus increasing the exposure of library data to a broader linked global community.

This adheres to one of the major principles of Linked Data, to provide useful links to other datasets.

Slide 16 This Wikipedia page describes the editing community’s consensus of norms and practices for adding authority control to Wikipedia articles. It also provides a template for authority control. Click on the slide to access the link to the page.

Slide 18 The British Library is one of the forerunners in experimenting with how to represent library-related information on the web.

Slide 19 How does BIBFRAME fit into this? LC is developing BIBFRAME as a tool to transform our MARC records…

Slide 20 into RDF graphs thereby providing the means to connect bibliographic information to the LOD Cloud.

What you see on this slide is a small portion of the RDF graph for the MARC record on the previous slide transformed into a MARC RDF/XML record using BIBFRAME and then parsed using an RDF validator which produces the graph.

This set of slides continues with additional slides that list the references used for my presentation followed by a list of resources that those of you who would like to learn more may find interesting.

Many of these slides have hyperlinks to their sources which can be accessed by right-clicking on the slide and opening the hyperlink.

Valerie will introduce BIBFRAME, and Kevin will show you what our Continuing Resource Work Group has been doing.