

ALA Annual Conference 2009

Monday, 7/13/2009

8:00 to 10:00 a.m.: From Legacy Data to Linked Data: Preparing Libraries for Web 3.0 (Hilton, Lake Ontario)

Corey Harper, Metadata Services Librarian, NYU – Intro only.

>> Eric Miller, Sophira (sp?) semantic web/knowledge mgmt: Bridging between repositories via OS tools and standardization efforts. Bind/map different hashtags rather than arguing over which to choose.

Captcha takes 10 sec./person/day X the number of people using the web = huge time waste. Re-captcha uses indecipherable OCR content as captchas to decode them, while managing the task of getting access.

Linked data: "Sharing metadata, structuring URIs to link to specific chunks of data, so it can be reused on/incorporated in different sites in different contexts" Spreads cost of development among all users. Apply metadata to the chunks to make them discoverable. Apply identifier (URI) to a person, then link to data chunks. Data.gov.

Tim Burners-Lee: "Raw data now!"

Specific applications and config makes rigid, brittle systems for single purposes. Structure applications in a standard way so they can be used in any way -- no need to grasp all possible uses beforehand; flexible structure lets the func requirement emerge later. UI becomes a simple afterthought -- for ANY application of the data.

LC project to archive digital versions of legacy content. Old way: Structure a metadata database to provide access a specific way. New way: store content elements with exposed metadata. Managed via spreadsheet, columns annotated for human use, exposed to the web allows the data to be combined and reused.

Annotated directory structure allows access for selection UI of resources stored in the directories. "Identifiers." Trust in who applies the annotations is the key element in choosing which datasets to use. Librarians help by validating and properly identifying/annotating/tagging elements.

Freemix OS tool package. Recollection platform.

>> Diane Hillman, Cornell/Metadata Mgmt Associates: Must be machine-readable; old model of one-by-one human processing is too expensive for volume of data. RDA will be available to linked data -- wealth of relationships surfaced and named for further processing.

Traditional cataloging: attribute-value pairs in a record with one identifier. Statements must be used packaged as a record. Instead, RELATIONSHIPS get identifiers, making data usable for different purposes and formerly unanticipated applications.

NSDL Registry: URI with vocabulary domain linked to FRBR Entity: 'Author' links to 'Work.'

RDA appendices define relationships among data elements. Author subproperties: libretist, screen writer, etc. XML schema defines the relationships, incl. resolvable URIs. FRBR entities maintained by IFLA.

>> Jennifer Someone (Bowen?) from URochester: Metadata on the Bleeding Edge: New normal is not knowing what you're doing. extensiblecatalog.org.

URochester Extensible Catalog project. OS tools for resource discovery, metadata mgmt, exposing ILS data. Reqs: Reuse MARC & other data; define schema to support the application; based on FRBR; reuse Dublic Core. Interim solution pending RDA release.

Developed application profile matching DC standard: combine fields from multiple sources, with namespace to trace back to sources, but compiled according to need or desire. Express terms as RDF triples. Also defined new elements to enhance the compilation.

USing all DC properties as intended; added 11 RDA elements; 37 elements. Add attributes that link to authority records, e.g., so some parallel properties.

Map MARC data into new app profile, but DC didn't include stuff like scale for maps or frequency for for serials. Assigning persistent URI to each metadata property.

Could do their XC without linked data, but they want to be progressive. Where's the ROI???

OAI/PMH protocol (?). Can dump data in one format, modify format, & reload from structured DB, giving a solid experimental platform.

>> Rebecca Someone, LC: Locally controled lists for specific purposes, among other controlled vocab elements, e.g. code lists, xml schemas with att. values specified, etc.

Simple Knowledge Organization System (SKOS). RDF app for expressing knowledge organization systems through linked data. Concept Schemes -- individual entities. Controls that link concepts. Can relate attributes in different metadata sets to one another.

Labeling properties: properties of attributes. Relationship levels & types.

<http://id.loc.gov>. Computer-to-computer interaction mostly, but human-readable. Communicates policy for access & interaction to specify wide-scale approved uses. Visualization tab shows element relationships.

RDF definitions of element attributes