

Bibliographic Framework Initiative Approach for MARC Data as Linked Data

Sally McCallum
smcc@loc.gov
Library of Congress

IGeLU, Zurich, September 13, 2012

Outline

- The environment
 - MARC environment pros and cons
 - How about cataloging norms?
 - And future resources?
 - Take aways?
 - Bold venture
 - Goals
 - Approach
 - Modeling progress
-

The environment

The environment

- MARC

- Positives

- Fully permeates the environment, globally
 - Adjusts to different cataloging norms
 - Adapted to technical changes – XML, UNICODE

- Issues

- Accumulation of data elements
 - Structural limitations
 - Changes in communications
-

Cataloging norms

- New library cataloging norms
 - More identification and parsing of data
 - More controlled vocabularies
 - Use of terms over codes
 - Emphasis on relationships
 - More flexibility with authoritative headings
 - Is it possible to include the broader cultural community in library cataloging norms?
-

Cataloging norms

- Transcription pros and cons
 - Subject headings or terms

 - Non-traditional material
 - Cover images? Excerpts? Table of contents?
 - Reviews? Author biographies? Abstracts?
 - User supplied information (crowd sourcing)
-

Resources

- Printed resource production – not yet diminishing?
 - E-resources
 - Increasing from the publisher
 - Increasing in the collections
 - Casual resources
-

Systems

- More need for eresource access management
 - Licensing management
 - Rights management
 - And for eresource object management
 - Preservation
 - Basic retrieval/scholar retrieval needs
 - How to leverage the web and linked data
-

Take away?

- Flexibility
 - Changing communications options
 - Changing cataloging behaviors
 - Variety of resource types
-

Bold venture

Bold venture

- Bibliographic Framework Initiative
 - Replace MARC as the core vehicle for exchange of bibliographic description
-

“Requirements”

- Broad accommodation of content norms and data models
- New views of different types of metadata
 - descriptive, authority, holdings
 - coded data, classification data, subject data
 - preservation, rights, technical, archival
- Reconsideration of the activity relationships
 - exchange, internal storage, input interfaces and techniques

“Requirements”

- Enhanced linking
 - traditional = textual, identifiers
 - semantic technology = URIs
- Accommodate different types of libraries
 - large, small, research, public, specialized, ...
- MARC compatibility
 - maintenance of MARC 21 continued
 - enable reuse of data from MARC
 - provision of transformations to new models

Approach

- Orient to a Web and Linked Data exchange environment
 - Investigate use of semantic web standards
 - RDF (Resource Description Framework) data model
 - Various syntaxes: XML, JSON, N-triples, etc.
 - Organize experimentation with initial high level models
 - Collaborate
-

Why linked data orientation?

- Easier integration of catalog data into general purpose Web and social media
 - Improved positioning of library resource metadata within Web search engines
 - Facilitate reuse of data and incorporation into applications
 - Increased flexibility for describing resources
-

Modeling progress

Initial model development

- Contracted with Zepheira in May 2012
 - Two major tasks:
 - Review several related initiatives
 - Translate bibliographic data to a linked data form
 - Evolution not revolution
 - Basis for community discussion / dialog
-

Balancing a number of factors

- MARC 21 historical data and roles
 - Previous efforts for modeling bibliographic information
 - FRBR (RDA), Indecs (ONIX)
 - Previous efforts to express bibliographic information as linked data
 - British Library, Deutsche National Bibliothek, Library of Congress's ID, OCLC Worldcat, schema.org
-

Balancing a number of factors

- Using Web as model for expressing and connecting information
 - URIs, decentralization of data, annotation
 - Library community social and technical deployment probabilities
 - Adoption outside the library community
-

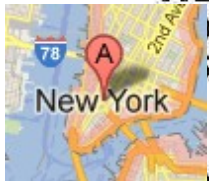
Balancing a number of factors

- Flexibility for future cataloging and use scenarios
 - Leverage machine technology for the mechanical while keeping the librarian expertise in control
-

Deconstructing MARC - Identifying MARC Resources (MARCR)



```
001 26132811
008 920608s1992 nyu 0 eng
010 $a92021087
020 $a1555701205
040 $aDLC$cDLC$dAGL
049 $aHNKA
050 00 $aZ678.9$b.C58 1992
082 00 $a025/.00285$220
1 $aCohn, John M
10 $aPlanning for automation :$ba how-to-do it manual for
librarians /$cJohn M. Cohn, Ann L. Kelsey, Keith Michael
Fiels
260 $aNew York $bNeal Schuman Publishers,$cc1992
300 $avi, 116 p. ;$c20 cm
440 0 $aHow-to-do it manuals for libraries ;$uno. 25
504 $aIncludes bibliographical references and index
650 0 $aLibraries$xAutomation$XManagement$XHandbooks, manuals, etc
650 0 $aLibrary planning$XHandbooks manuals, etc
700 10 $aKelsey, Ann L
700 10 $aFiels, Keith Michael
```



Search








Edward Maynard

Author

- 2 (missing this field)
- 1 Bryan, W. B.
- 1 Georgetown (Washington, D Corporation.
- 1 Kahn, Tom.
- 1 Maynard, Edward,

Cooper, W. W. Corcoran, John A. B. Dahlgren, Stephen A. George Washington Riggs, Edwin McMasters Stanton, a

Subject

-  Carter, Timothy W.
-  Cooper, Samuel, (★)(✓)
-  Corcoran, W. W. (✓)
-  Dahlgren, John Adolphus Bernard, (★)(✓)
-  Douglas, Stephen A. (✓)
-  Ewing, Thomas, (★)(✓)
-  Kearny, Philip, (★)(✓)
-  Mackay, George
-  Mann, Horace, (★)(✓)
-  McClellan, George Brinton, (★)(✓)

Details

Visualization

 **Cooper, Samuel, 1798-1876**

URI(s)

> <http://id.loc.gov/authorities/names/n8611443>

Instance Of

- > [MADS/RDF PersonalName](#)
- > [MADS/RDF Authority](#)
- > [SKOS Concept](#) 

Washington, (★)(✓)

..., (★)(✓)

Maynard tape primer), his patent of it partners, Thomas L. Smith and J. Washington Tyson, to get firearms manufactured utilizing the his improvements in the breech-loading rifle and the sale of his patent rights. Correspondents include (usetts Arms Company), George Mackay, Edward Riddle, Thomas L. Smith, and Joseph Washington documentina Mavnard's Washington, D.C.. dental practice. Patients mentioned include Samuel

Cooper, Samuel, 1798-1876 

VIAF ID: 38401361 (Personal)

Permalink: <http://viaf.org/viaf/38401361>

 Preferred Forms



100 1 _ [⌘a Cooper, Samuel, ⌘d 1798-1876](#)

High level model- Phase 1

- 4 Core classes

- Work

- Resource reflecting the conceptual essence of the cataloging item
- Roughly equivalent to FRBR work or expression

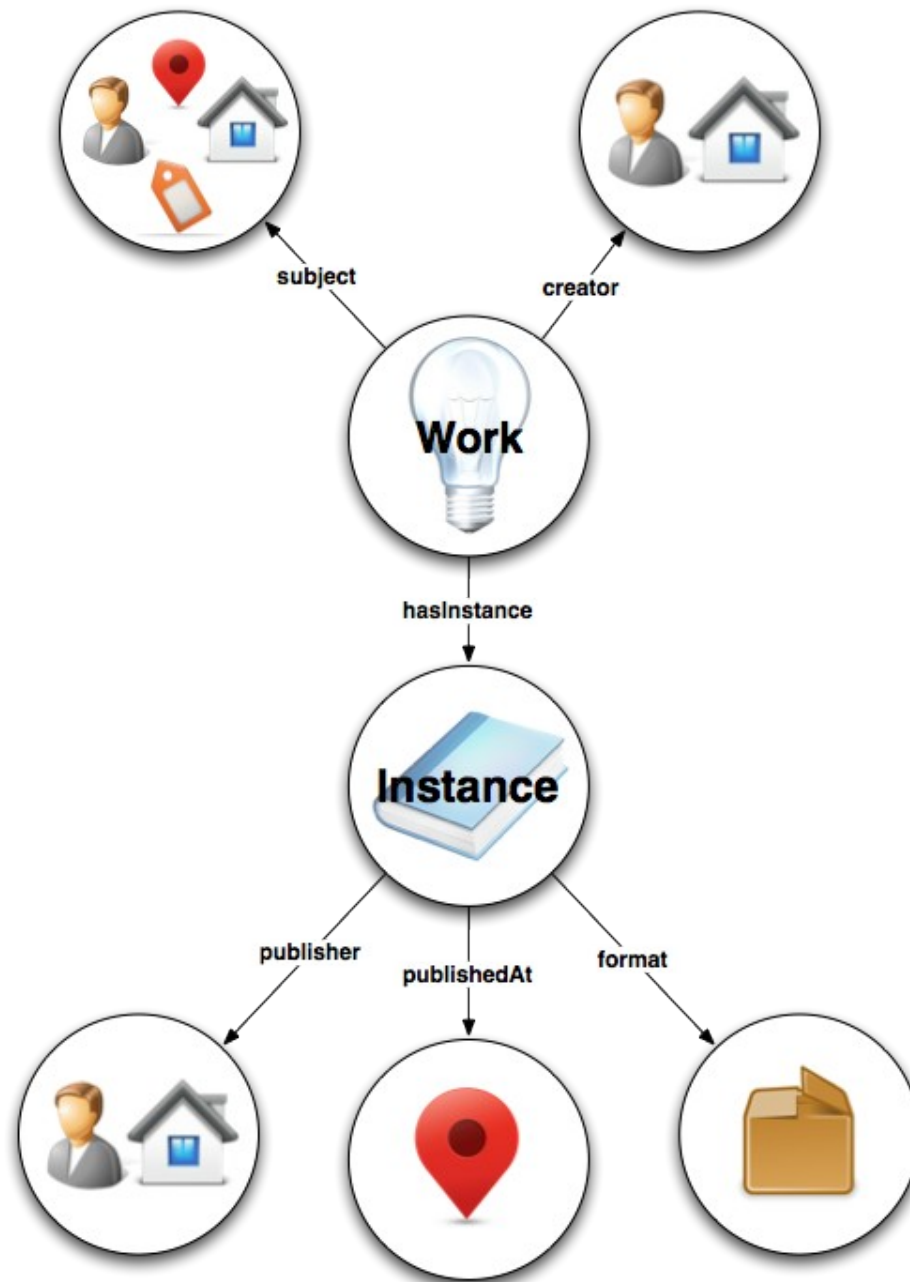
- Instance

- Resource reflecting an individual, material embodiment of the Work
-

High level model

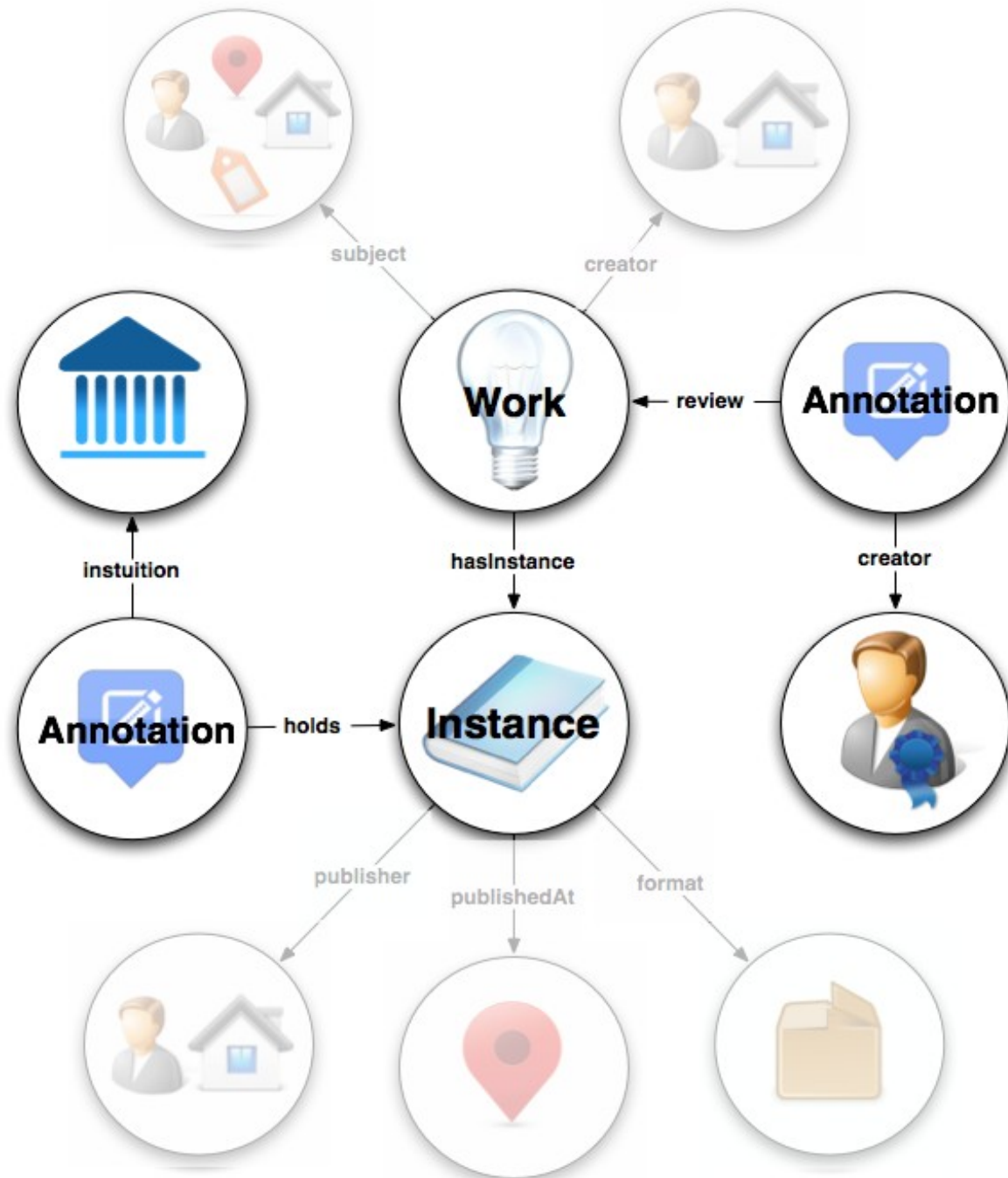
- Authority
 - Resource reflecting key authority concepts that have defined relationships reflected in the Work and Instance
 - E.g., People, Places, Topics, Organizations

 - Annotation
 - Resource that “decorates” other MARCR resources
 - E.g., holdings, cover images, reviews
-



Annotations

- Assertions about the other core class elements
 - Works – reviews, table of contents, abstract, excerpts
 - Instances – book cover images, web site opening page, holdings
 - Name authority – author biographical information
 - Administrative metadata
-



Initial model development

- Other Zepheira tasks:
 - Develop prototype services and tools to enable experimentation
 - Create a roadmap for moving forward toward refinement, redevelopment, or alternative approaches
-

“Phase 1.5” – early experimentation

- LC – preliminary work
 - Very small group of early experimenters
 - October-November
 - Working with high level model, vocabularies, conversion tools
 - Creative development of syntaxes and configurations
 - Adjust model
-

Model development

- Make model, mappings, and tools available and encourage broader experimentation?
 - Parallel Phase 2 to refine model and keep folding in experience based changes

 - Follow the progress
 - www.loc.gov/marc/transition
 - Join the discussion
 - bibframe@listserv.loc.gov
-