When URIs become authority:
Benefits and challenges of library Linked Data
**BIBFRAME: Proposed Roadmap**

**Short-Term Goal:** Refine BIBFRAME Model and Prepare for Transition

- Landscape Assessment
- Early Experiments
- Model Development
- Pipeline Development

**Medium-Term Goal:** Community Education and Transition

- Extended Group Testing
- BIBFRAME Model Refinement & Application Profiles
- Annotation Demonstration and Stress Testing

**Long-Term Goal:** Adoption

- Completed Tasks

**Community/Market Adoption**

- Education & Outreach Materials
- Stakeholder Transition Strategy, Education and Transition
- Library Vendor and Search Strategy
- Library Vendor and Search Adoption
- Use Cases and Validation
- End User Facing Tool Development
- Value Proposition for Transition
- Authority Services
- BIBFRAME Vocabulary Governance Model

---

Kevin Ford  
NDMSO, Library of Congress  
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013  
ALCTS CaMMS Subject Analysis Committee  
ALA Midwinter, Seattle, WA
Authoritative URIs

OUTLINE

Introduction

Authorities and Linked Data
   The ideal and the URI

BIBFRAME Authority
   The reality and the abstraction-layer

LCSH as Linked Data

LCSH as Linked Data meets actual bibliographic description

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
Identifying Things

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA

28 January 2013
Identifying Things

http://id.loc.gov/authorities/names/n92062246

http://id.loc.gov/authorities/names/n79007751

http://id.loc.gov/authorities/n80016154

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
<Report id = "http://bibframe/work/frbr-report">
  <title>Functional requirements for bibliographic records :</title>
  <titleRemainder>final report / IFLA Study Group on the Functional Requirements for Bibliographic Records ; approved by the Standing Committee of the IFLA Section on Cataloguing.
  </titleRemainder>
  <creator resource = "http://id.loc.gov/authorities/names/nr98013265" />
</Report>
What is BIBFRAME?

The Bibliographic Framework Initiative will re-imagine and implement a bibliographic environment for a post-MARC world.

Key dates:

May 2011 – Initiative announced.
October 2011 – LC publishes general plan
May 2012 – Zepheira LLC contracted to evaluate related initiatives and data modeling
November 2012 – LC publishes high-level model document

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
Core Classes

Work

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills
Core Classes

Work

Resource reflecting the conceptual essence of the cataloged item
BIBFRAME

Core Classes

Work

Instance

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills
Core Classes

Work

Instance

Resource reflecting a material embodiment of a BIBFRAME Work
Core Classes

BIBFRAME Work

Instance

Resource reflecting a material embodiment of a BIBFRAME Work

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
BIBFRAME

Core Classes

Work

Instance

Authority

Resource reflecting key authority concepts that have defined relationships to Works and Instances
Authorities are not designed to compete or replace existing authority efforts but rather provide a common, light-weight abstraction layer over various different Web based authority efforts to make them even more effective.
The Ideal

<Report id = "http://bibframe/work/frbr-report">
  <title>Functional requirements for bibliographic records : </title>
  <titleRemainder>final report / IFLA Study Group on the Functional Requirements for Bibliographic Records ; approved by the Standing Committee of the IFLA Section on Cataloguing.
  </titleRemainder>

  <creator resource = "http://id.loc.gov/authorities/names/nr98013265" />
</Report>

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
Functional requirements for bibliographic records: final report / IFLA Study Group on the Functional Requirements for Bibliographic Records; approved by the Standing Committee of the IFLA Section on Cataloguing.

IFLA Study Group on the Functional Requirements for Bibliographic Records
http://www.ifla.org/
http://id.loc.gov/authorities/names/nr98013265

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
Present definition of BIBFRAME Authority not wrong but...

   the objective, as described in the primer, is achievable without the “abstract layer” construct.

Nevertheless, for other reasons, the abstraction layer becoming unavoidable

100, 110, 111 ... $e$
700, 710, 711 $i$
RDA emphasis on true, faithful transcription

In short:

   The “abstraction layer” construct may have always been inevitable.
The ideal has always been envisioned separately from practice/reality.
Reality turns up things you weren't expecting.
<Play id = "http://bibframe/work/burial-thebes">
  <title>Burial at Thebes</title>
  <creator resource = "http://bibframe/auth/person/heaney" />
</Play>

<Person id="http://bibframe/auth/person/heaney">
  <label>Heaney, Seamus, 1939-</label>

  <relationToWork>trans. buy</translationToWork>

  <hasIDLink resource="http://id.loc.gov/authorities/names/n79099140" />
</Person>
<Play id = "http://bibframe/instance/burial-thebes">
    <title>Burial at Thebes</title>
    <placeOfPublication resource = "http://bibframe/auth/place/nyc" />
</Play>

<Place id="http://bibframe/auth/place/nyc">
    <transcribedLabel>New York</transcribedLabel>
    <hasIDLink resource="http://id.loc.gov/authorities/names/n79007751" />
</Place>

<Geographic rdf:about="http://id.loc.gov/authorities/names/n79007751">
    <authoritativeLabel>New York (N.Y.)</authoritativeLabel>
</Geographic>
Quick facts about LCSH at ID.LOC.GOV:

Available since 2009
Approximately 415,000 resources

Quick facts about LCSH usage in LC's Voyager database:

24,893,391 total LCSH headings
5,692,450 unique LCSH headings

5,692,450 (million) > 415,000 (thousand)

LCSH – because of pre-coordination (which is a good thing) is basically an infinite system

ID.LOC.GOV has, at best, a subset of the possible.
ID.LOC.GOV does have the data from which one creates headings (well, mostly).

ID.LOC.GOV is missing, however, at least, about 5.3 MILLION known concepts
LCSH & Linked Data

Questions

How to best do authority control / LCSH in a real-life Linked Data ecosystem?
How best to do authority control / LCSH in BIBFRAME?

To pre-coordinate or post-coordinate
  No contest: arguments presently overwhelming favor pre-coordination (and rightly so)

Still, distinct flavor for “faceting” with BIB data presently
  Audience designations
  Group characterization of authors
  Medium
  Temporal notions

URIs for the all possibilities?
  Possible
  Changes to LCSH practice?
More about actual subject usage in the catalog*:

<table>
<thead>
<tr>
<th>Components used ($a + subdivs)</th>
<th>Number of Subject Headings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6,781,752</td>
</tr>
<tr>
<td>2</td>
<td>10,468,040</td>
</tr>
<tr>
<td>3</td>
<td>5,760,110</td>
</tr>
<tr>
<td>4</td>
<td>1,611,110</td>
</tr>
<tr>
<td>5</td>
<td>244,322</td>
</tr>
<tr>
<td>6</td>
<td>25,715</td>
</tr>
<tr>
<td>7</td>
<td>2,140</td>
</tr>
<tr>
<td>8</td>
<td>171</td>
</tr>
<tr>
<td>9</td>
<td>21</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

* I'd be interested to know how these stats compare to LCSH usage in the LC database before the Airlie House Conference recommendations

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALA Midwinter, Seattle, WA
## LCSH Usage Stats

More about actual subject usage in the catalog:

<table>
<thead>
<tr>
<th>Subfield usage</th>
<th>No. of subject headings using the same pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>6,781,717</td>
</tr>
<tr>
<td>ax</td>
<td>4,883,546</td>
</tr>
<tr>
<td>az</td>
<td>3,096,913</td>
</tr>
<tr>
<td>av</td>
<td>2,042,296</td>
</tr>
<tr>
<td>azx</td>
<td>1,046,032</td>
</tr>
<tr>
<td>axz</td>
<td>887,856</td>
</tr>
<tr>
<td>azv</td>
<td>773,085</td>
</tr>
<tr>
<td>azy</td>
<td>762,346</td>
</tr>
<tr>
<td>axx</td>
<td>709,172</td>
</tr>
<tr>
<td>azz</td>
<td>679,781</td>
</tr>
<tr>
<td>axv</td>
<td>458,786</td>
</tr>
<tr>
<td>ay</td>
<td>301,921</td>
</tr>
<tr>
<td>azzx</td>
<td>285,397</td>
</tr>
<tr>
<td>azzv</td>
<td>218,017</td>
</tr>
<tr>
<td>ayy</td>
<td>183,559</td>
</tr>
</tbody>
</table>
<Book id = "http://bibframe/work/acw">
  <title>The American Civil War: A Compilation of primary source materials</title>
  <subject resource = "http://id.loc.gov/authorities/subjects/sh85140253" />
</Book>

<Topic rdf:about="http://id.loc.gov/authorities/subjects/sh85140253">
  <authoritativeLabel>
    United States--History--Civil War, 1861-1865—Pamphlets
  </authoritativeLabel>
</Topic>
<Book id = "http://bibframe/work/acw">
   <title>The American Civil War: A Compilation of primary source materials</title>
   <subject resource = "http://bibframe/subject/acw-diaries" />
</Book>

<Topic id = "http://bibframe/subject/acw-diaries">
   <label>United States--History--Civil War, 1861-1865—Diaries</label>
   <geographic id = “http://id.loc.gov/authorities/names/1234567” />
   <topic id = “http://id.loc.gov/authorities/subjects/1234567” />
   <temporal id = “http://id.loc.gov/authorities/subjects/987654” />
   <genreForm id = “http://id.loc.gov/authorities/subjects/456783” />
</Topic>

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
<Book id = "http://bibframe/work/acw"/>
  <title>The American Civil War: A Compilation of primary source materials</title>
  <subject resource = "http://bibframe/subject/acw" />
  <genreForm resource = "http://bibframe/genreForm/diaries" />
</Book>

<GenreForm id = "http://bibframe/genreForm/diaries">
  <label>Diaries</label>
  <hasIDLink id = "http://id.loc.gov/authorities/subjects/sh99001602" />
</GenreForm>

<Topic id = "http://bibframe/subject/acw">
  <label>United States--History--Civil War, 1861-1865</label>
  <hasIDLink id = "http://id.loc.gov/authorities/subjects/sh85140205" />
</Topic>

Kevin Ford
NDMSO, Library of Congress
Email: kefo@loc.gov / Twitter: @3windmills

28 January 2013
ALCTS CaMMS Subject Analysis Committee
ALA Midwinter, Seattle, WA
<Book id = "http://bibframe/work/acw">
    <title>The American Civil War: A Compilation of primary source materials</title>
    <subject resource = "http://id.loc.gov/authorities/subjects/sh85140205" /> 
    <genreForm resource = "http://id.loc.gov/authorities/subjects/sh99001602" /> 
</Book>
Questions

What might the bibliographic universe be like if every known, legitimate LC subject heading had its own record/resource?

What are the benefits (or not) of removing form subdivisions from subject headings in favor of faceted identification of genre/form in the bibliographic record/resource?

General thoughts on the notion of a BIBFRAME Authority?

Thank you.

Kevin Ford

Email: kefo@loc.gov
Twitter: @3windmills