Authority Control: Where Is It Heading?

Dr. Sherry Vellucci
Rutgers University
vellucci@scils.rutgers.edu
Overview of Today’s Talk

- Authority Control for Names
  - Conceptual models
  - International context
  - Research initiatives

- Authority Control for Subjects
  - Conceptual models
  - International context
  - Subjects on the Web
  - Research initiatives

© 2006 Dr. Sherry Vellucci
Authority Control (AC)

- Traditional Functions
  - Ensures that access points are unique and consistent in content and form
  - Provides a network of linkages for variant and related headings in the catalog
  - Improves precision & recall for database searches
Reasons for Authority Control Success

- AC operates within a well-defined and bounded universe—the library catalog
- Creation of access points based on principles & standardized practices that guide the process
- Authority work is aided by reference to authoritative lists
- Performed by highly trained individuals
  - Part of library culture
  - Understand cause and effect in the information retrieval process
Name Authority Conceptual Models

- FRANAR (Name of Working Group)
  - Functional Requirements & Numbering of Authority Records
- FRAR (Name of Model)
  - Functional Requirements for Authority Records
FRANAR (Functional Requirements & Numbering of Authority Records)

- International Working Group of IFLA
- Extend work of FRBR & define the Functional Requirements of Authority Records
- Study the feasibility of an International Standard Authority Data Number (ISADN)
FRANAR (cont’d)

- Define possible use and users, to determine for what types of authority records such an ISADN is necessary.
- Identify what data elements are necessary for internationally shared authority records.
- Identify possible structure for the number and the type of management that would be necessary.
Feasibility of an ISADN

- International Standard Authority Data Number first proposed in the 1980s
- Assigned to a heading or to a record?
- Focus shifting away from a single form of heading for an entity
- Focus shifting away from sharing physical records
The Goal of the FRANAR Model

- To provide an understanding of how authority files function currently
- To clarify the underlying concepts to provide a basis for refining and improving on current practice in the future
Functions of the Authority File

- Document decisions
- Serve as reference tool
- Control forms of access points
- Support access to bibliographic file
- Link bibliographic and authority files
Users

- Authority record creators and reference librarians
- Library patrons
Users

- Authority record creators and reference librarians
- Library patrons

User Tasks

- Find
  - Find an entity or set of entities corresponding to stated criteria
- Identify
  - Identify an entity
- Contextualize
  - Place a person, corporate body, work, etc. in context
- Justify
  - Document the authority record creator’s reason for choosing the name or form of name on which an access point is based
Bibliographic Entities

Names and/or Identifiers

known by

basis for

Controlled Access Points
FRBR Bibliographic Entities

Group 1
- WORK
- EXPRESSION
- MANIFESTATION
- ITEM

Group 2
- PERSON
- FAMILY
- CORPORATE BODY

Group 3
- CONCEPT
- OBJECT
- EVENT
- PLACE

FRAR Model
- ACCESS POINT
- RULES
- AGENCY

 known by

 assigned

 basis for

 governed by

 created / modified by

 (Diagram from FRANAR Report)
FRAR Model

(Diagram from FRANAR Report)
FRANAR Terminology

Access Point
What’s Next for FRANAR?

- Draft for worldwide review (October 2005)
- Resolution of comments
- Publication of the FRAR model
- Completion of recommendations on numbering
- Recommendations for changes to existing IFLA publications
FRAR resources


Access Control in an International Context

Focus on Names
Authority Control in the International Community

- Connecting International Authority Files
  - Problems:
    - Language
    - Scripts
    - Cataloging rules & standards
    - Record formats
    - Local systems
  - Solutions:
    - Access vs. Authority Control
    - Display switches
    - Pool authority record resources
    - ISADN & FRANAR

Virtual International Authority File (VIAF)
Old IFLA UBC Authority Principles

- Each country responsible for authority headings for its own personal and corporate authors
- National authority records available for everyone to use
- *Same form and structure used worldwide*
New View of UBC

- User perspective - to display script/language(s) of one’s own country
- Bibliographic agencies still responsible for control in their own countries (or region/cataloguing rules/etc.)
- Link forms established in “national” or “regional” authority files to create a virtual international authority file
Same Entity/Variant Scripts

Library

發音記号表

Arхив

 발음기호표

Bibliothèque

لا توجد نتائج بحث عن
Switching for Displays

- Library Default
- User selected preference
- Client set-up, “cookies”
  - Language
  - Script
  - Culture (country)/spelling
    - Catalog vs. Catalogue
Local System - User View

User’s cookie specifies Cyrillic/Russian preferred

Display 700 form of name, building on local system’s authority structure (700 = Cyrillic/Russian)

Чайковский, Петр Ильич
Authority Control

And More Research …
Retrievability: The Measure of Success for Authority Control

- Research studies in this area include:
  - Data quality
  - Automation
  - Foreign and non-roman languages
  - Internationalization

- Management studies include:
  - Workflow, time, and cost
Authority Data Quality Research

- What errors exist in authority data and how does this impact retrieval?
- What impact does lack of cross-references have on retrieval?
- What impact do variant author names have on retrieval?
- How do forms of corporate names appearing on corporate Web sites compare to corporate names found in the Library of Congress Name Authority File?
Patton, Reynolds and Choudhury present a detailed report of the process used to create an automated authority record for the Lester S. Levy Collection of Sheet Music

To test the dependability of a new authority control module in the library’s automated system, Greever compared the module’s effectiveness to the currently used manual procedures for authority control.

Foreign & Non-Roman Names

- Ruiz-Perez, et al., investigated how Spanish names were handled in three different databases and identified mistakes that would have a negative impact on locating and retrieving works by Spanish authors.


- Hu used a transactional analysis method to examine problems in cataloging Chinese names, including transliterating Chinese characters and establishing entries for personal names.

Bolick also studies Chinese names, examining the reasons why vendor software produced negative results for non-unique Chinese names.


Wang examined the problems of Chinese names in different languages, i.e., English transliteration, transliterated name from spoken Chinese dialect, westernized Chinese, and pure Chinese.

A-C Management Issues

- Kulczak used the OCLC database to evaluate the quality of front-end authority work for monographs in order to determine whether authority work was necessary at the copy cataloging stage.


- Ellero investigated the use of the World-Wide-Web as a source of information for authority records.


- Jin adopted a narrower research focus and examined the process of constructing corporate name headings using data from the corporate body’s home page.

  Jin, Qiang. “Creating up-to-Date Corporate Name Authority Records by Using Official Corporate Home Web Pages.” *Cataloging & Classification Quarterly* 38, no. 3/4 (2004):281-90;
Bangalore and Prabha investigated the copy cataloging process by measuring the time and effort expended.


Santizo surveyed academic libraries to find information about the type of decisions and level of responsibility of authority control required of copy catalogers.


Tsui and Hinders conducted a cost-benefit analysis for outsourcing the authority control process.

Controlled Vocabularies in an International Context

Focus on Subjects
FRSAR

- A further extension of FRBR into the area of subject access
- Both subject thesauri and classification
- Working group chaired by Marcia Zeng (Kent State University), Maja Žumer (University of Ljubljana) and Athena Salaba (Kent State University)

http://www.ifla.org/VII/s29/wgfrsar.htm
FRSAR Terms of Reference

- Build a conceptual model of Group 3 entities within the FRBR framework
  - Object
  - Concept
  - Event
  - Place
- Provide a frame of reference for relating data recorded in subject authority records to the needs of users of those records
- Assist in assessment of potential for international sharing and use of subject authority data in the library sector and beyond
FRSAR User Tasks Pilot Studies

by Marcia Zeng, Athena Salaba
Maja Zumer
FRSAR Activities

- Determine Use of Subject Authority Data
- Identify User Tasks
  - User surveys on the Use of Subject Authority Data
    - Pilot with Non Library & Information Professionals
    - Library & Information Professionals
  - FRSAR User Tasks
- Identify Subject Entities
Use of Authority Data

Direct use of authority data
- Search/Find authority data
- Browse/Explore authority data
- Identify/Verify term
- Obtain entity information
- Select term

Indirect use of authority data
- Access to bibliographic information through authority data

Users
- Information professionals
- Searcher/end-user
Q1. Which of the following best describes your work? (ranked by frequency)

- 1 Creating & maintaining semantic tools (thesauri, taxonomies, ontologies and other controlled vocabularies)
- 2 Offering consultations in information architecture, navigation, design of retrieval tools & interfaces
- 3 Supervising/directing metadata project(s)
- 4 Using semantic tools in metadata creation
- 5 Offering consultations on vocabulary control issues
- 6 Quality control of metadata
- 6 Offering consultations on metadata issues
- 6 Assisting others in finding information

(Presentation by Zeng, Saluba, & Zumer, March 2006 @ The Semantic Technology Conference)
Q2. In what ways do you use the semantic tools? (ranked by frequency)

- 1. to find/identify relevant information resources using specified terms
- 2. to access metadata information
- 3. to find/identify appropriate terms when searching for information
- 3. to navigate or modify search queries through broader and narrower terms
- 4. to assign subject terms and/or class labels in metadata records
- 5. to explore a topic through browsing synonymous and related terms
### 1. Background

1.1. Which of the following best describes your work? (Please, check all that apply)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Response Percent</th>
<th>Response Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assisting others in finding information (e.g., reference work)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creating metadata and cataloging records</td>
<td>79.1%</td>
<td>617</td>
</tr>
<tr>
<td>Creating and maintaining subject authority records (in relation to metadata and cataloging record creation)</td>
<td>42.2%</td>
<td>326</td>
</tr>
<tr>
<td>Supervising metadata and/or cataloging projects</td>
<td>46.5%</td>
<td>359</td>
</tr>
<tr>
<td>Quality control of metadata and/or cataloging records</td>
<td>60.5%</td>
<td>472</td>
</tr>
<tr>
<td>Consulting on metadata issues</td>
<td>22.3%</td>
<td>174</td>
</tr>
<tr>
<td>Indexing and/or abstracting</td>
<td>11.3%</td>
<td>90</td>
</tr>
<tr>
<td>Creating indexes</td>
<td>6.5%</td>
<td>51</td>
</tr>
<tr>
<td>Creating and maintaining controlled vocabularies (lists of subject headings, thesauri, classifications, lexicons, etc.)</td>
<td>18.8%</td>
<td>147</td>
</tr>
<tr>
<td>Consulting on vocabulary control issues</td>
<td>10.4%</td>
<td>82</td>
</tr>
<tr>
<td>Selecting and acquiring resources in relation to collection development</td>
<td>22.7%</td>
<td>177</td>
</tr>
<tr>
<td>Consulting on information architecture and design of retrieval tools and interfaces</td>
<td>10.3%</td>
<td>82</td>
</tr>
<tr>
<td>Teaching and training information professionals</td>
<td>10.6%</td>
<td>82</td>
</tr>
<tr>
<td><strong>Total Respondents</strong></td>
<td>780</td>
<td></td>
</tr>
<tr>
<td><strong>(skipped this question)</strong></td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Q1. Which of the following best describes your work?

1. Which of the following best describes your work? (Please, check all that apply)

- assisting others in finding information (e.g., reference work)
  - Response Percent: 30.1%
  - Response Total: 235

- creating metadata and cataloging records
  - Response Percent: 79.1%
  - Response Total: 617

- creating and maintaining subject authority records (in relation to metadata and cataloging record creation)
  - Response Percent: 42.9%
  - Response Total: 335

- supervising metadata and/or cataloging project(s)
  - Response Percent: 46%
  - Response Total: 359

- quality control of metadata and/or cataloging records
  - Response Percent: 60.6%
  - Response Total: 473

© 2006 Dr. Sherry Vellucci
(cont.) Q1. Which of the following best describes your work? (2)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Percentage</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>consulting on metadata issues</td>
<td>22.3%</td>
<td>174</td>
</tr>
<tr>
<td>indexing and/or abstracting</td>
<td>11.5%</td>
<td>90</td>
</tr>
<tr>
<td>creating indexes</td>
<td>6.5%</td>
<td>51</td>
</tr>
<tr>
<td>creating and maintaining controlled vocabularies (lists of subject headings, thesauri, classifications, taxonomies, etc.)</td>
<td>18.8%</td>
<td>147</td>
</tr>
<tr>
<td>consulting on vocabulary control issues</td>
<td>16.4%</td>
<td>128</td>
</tr>
<tr>
<td>selecting and acquiring resources in relation to collection development</td>
<td>22.7%</td>
<td>177</td>
</tr>
<tr>
<td>consulting on information architecture and design of retrieval tools and interfaces</td>
<td>10.5%</td>
<td>82</td>
</tr>
<tr>
<td>teaching and training information professionals</td>
<td>18.6%</td>
<td>145</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>6.8%</td>
<td>53</td>
</tr>
</tbody>
</table>
Q2. In what ways do you use controlled vocabulary information?

- 2a. In cataloging and metadata creation (87%)
- 2b. In subject authority work (77%)
- 2c. In searching or helping others search bibliographic information (81%)
Q2. In what ways do you use controlled vocabulary information?

2a. In cataloging and metadata creation

![](image)

© 2006 Dr. Sherry Vellucci

July 27, 2006 data
Q2. In what ways do you use controlled vocabulary information?

2b. In subject authority work

<table>
<thead>
<tr>
<th>Activity</th>
<th>Response Percent</th>
<th>Response Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>to normalize and standardize controlled vocabulary terms</td>
<td>87.2%</td>
<td>525</td>
</tr>
<tr>
<td>to modify subject authority data when needed</td>
<td>62.3%</td>
<td>375</td>
</tr>
<tr>
<td>to maintain subject authority data for future use</td>
<td>57.6%</td>
<td>347</td>
</tr>
<tr>
<td>to establish and update term relationships</td>
<td>52.7%</td>
<td>317</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>3.5%</td>
<td>21</td>
</tr>
</tbody>
</table>

Total Respondents: 602

(skipped this question) 178
(cont.) Q2. In what ways do you use controlled vocabulary information?

2C. In searching or helping others search bibliographic information

<table>
<thead>
<tr>
<th>Activity</th>
<th>Response Percent</th>
<th>Response Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>to select and/or verify appropriate search terms for retrieval</td>
<td>86.4%</td>
<td>545</td>
</tr>
<tr>
<td>to explore relationships (equivalent, hierarchical, and associative) among terms</td>
<td>61.8%</td>
<td>390</td>
</tr>
<tr>
<td>to understand terms and their relationships</td>
<td>59.3%</td>
<td>374</td>
</tr>
<tr>
<td>to find and identify relevant documents/resources on a specified topic</td>
<td>76.9%</td>
<td>485</td>
</tr>
<tr>
<td>to navigate and/or browse bibliographic and metadata records using equivalent, hierarchical, and associative terms</td>
<td>63.9%</td>
<td>403</td>
</tr>
<tr>
<td>to modify searching through synonymous (equivalent) terms</td>
<td>60.9%</td>
<td>384</td>
</tr>
<tr>
<td>to modify searching through broader and narrower (hierarchical) terms and related (associative) terms</td>
<td>65.8%</td>
<td>415</td>
</tr>
</tbody>
</table>

Other (please specify)                                                   | 2.4%             | 15             |

Total Respondents                                                        | 631              |

(Skipped this question)                                                  | 149              |
**User Tasks**

**FRBR (1998):**
- **Find** entities of Group 1 that have entities from Group 1, 2, 3 as their subject
- **Identify**
- **Select**
- **Obtain**

**FRAR (2005):**
- **Find** one entity or entities
- **Identify** an entity
- **Contextualize**, place in context, explore relationships
- **Justify** the form of an access point

**FRSAR (2006 draft):**
- **Find** subject entity or entities
- **Identify**
- **Select**
- **Obtain** Additional information about the subject entity
- **Explore** Bibliographic records or resources about this subject entity
Interoperability Between Vocabularies

- Huge demand for accessing information that has been indexed with another language and/or vocabulary. The buzzword is ‘Mapping’. The Semantic Web is just one application.
Interoperability (CONT.)

Factors Affecting Interoperability
- Multilingual Controlled Vocabularies
- Searching
- Indexing
- Merging Databases
- Merging Controlled Vocabularies
- Storage and Maintenance of Relationships among Terms in Multiple Controlled Vocabularies
Three Approaches in the Development of Multilingual Thesauri:

- Building a new thesaurus from the bottom up
  - starting with one language and adding another language or languages
  - starting with more than one language simultaneously
- Combining existing thesauri
  - merging two or more existing thesauri into one new (multilingual) information retrieval language to be used in indexing and retrieval
  - linking existing thesauri and subject heading languages to each other; using the existing thesauri and/or subject heading languages both in indexing and retrieval
- Translating a thesaurus into one or more other languages

1. IFLA Guidelines for Multilingual Thesauri
Examples are in Multiple Languages

<table>
<thead>
<tr>
<th>English (British)</th>
<th>English (USA)</th>
<th>Dutch</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>cranes (birds)</td>
<td>cranes (birds)</td>
<td>kraanvogels</td>
<td>grue (oiseau)</td>
</tr>
<tr>
<td>cranes (lifting equipment)</td>
<td>cranes (lifting equipment)</td>
<td>hijskranen&lt;br&gt;SN voor andere typen kranen, zie aldaar</td>
<td>grue (appareil de levage)</td>
</tr>
<tr>
<td>water taps</td>
<td>water faucets</td>
<td>waterkranen</td>
<td>robinet à eau</td>
</tr>
<tr>
<td>gas taps</td>
<td>gas faucets</td>
<td>gaskranen</td>
<td>robinet à gaz</td>
</tr>
<tr>
<td>taps</td>
<td>faucets</td>
<td>kranen&lt;br&gt;SN voor kranen als hijswerktuig gebruik hijskranen</td>
<td>robinet&lt;br&gt;NT robinet à eau&lt;br&gt;NT robinet à gaz</td>
</tr>
<tr>
<td>NT water taps</td>
<td>NT water faucets</td>
<td>NT waterkranen</td>
<td></td>
</tr>
<tr>
<td>NT gas taps</td>
<td>NT gas faucets</td>
<td>NT gaskranen</td>
<td></td>
</tr>
</tbody>
</table>

Cranes is a homograph in English does not necessarily mean that equivalent terms in other languages are also homographs. The Dutch term kranen is a homograph too, but with the meanings cranes (lifting equipment) and taps.
And on That Note …

- Authority Control / Access Control
  - Still an important part of information search and retrieval, especially for the semantic Web
  - The process may become more automated, but humans are still needed
  - Will see new types of and terms for authority control used in metadata (URI, DOI, subject as a work attribute, biographical/contextual information …)
  - Hang in there – Your knowledge and skills are vital in the digital environment

Thank You!

Sherry Vellucci, PhD
Rutgers University
School of Communication, Information and Library Science
4 Huntington Street
New Brunswick, NJ 08901
732-932-7500 ext. 8232
vellucci@scils.rutgers.edu

And Thanks to Barbara Tillett, Glenn Patton and Marcia Zeng who provided many of the slides