USE CASE 2 DEMO

Rebecca Younes, Cornell University
Linked Data for Libraries Workshop
Stanford University, February 23, 2015
Use Case 2

• See and search on works by people to discover works of interest based on connection to people, and to understand people based on their relation to works
• Links between catalog data and profile systems for the enhancement of each
• Catalog data: biblio-centric
• Profile system: people-centered
VIVO

- Faculty profile system used at Cornell
- Originally developed and implemented at Cornell
- Now adopted by a growing number of institutions of various types world-wide
- Information on people, research, publications, teaching activities, events, and departments within the institution
- vivo.cornell.edu
VIVO technology

- Open source semantic web application
- Vitro: general-purpose web-based ontology and instance editor with customizable public browsing
- VIVO = Vitro + ontologies
- www.vivoweb.org
Cornell Use Case 2 Demo

- Demonstrate links between CUL catalog and VIVO
- Round-trip from catalog to VIVO and back to catalog
- Sample data: Cornell thesis records
- Convert records to RDF
- Ingest RDF into Vitro as a browse and search interface
Stages of data transformation

1. **MARC record**
2. **MARC to MARCXML converter**
3. **MARC XML**
4. **Pre-processing**
5. **Normalized MARC XML**
6. **LC BIBFRAME converter**
7. **BIBFRAME RDF**
8. **Post-processor**
9. **Linked Data**
10. **LOD**
MARC thesis record

The Suggestive Influence Of Lineups On Memory

The study was designed to test whether making a selection from a target-absent lineup could alter memory for the actual perpetrator. Three different lineup presentations were used: simultaneous, sequential, and elimination. In addition, three different memory tests were used in which participants saw two of the following three faces: the actual thief, a suspect from the original lineup, and a novel suspect. The sequential lineup resulted in the fewest initial false identifications, while the simultaneous lineup appeared to protect the most against memory alteration. Logistic regression models were utilized to estimate the likelihood of making a particular identification during the memory test. Participants were more likely to identify the innocent original suspect when they were older, non-white, and were shown the sequential lineup. Participants were more likely to make an accurate identification of the thief when they were younger and had made an identification from the initial lineup.

Lineup
Misinformation Effect
Memory

http://hdl.handle.net/1813/38774 Connect to full text. Access to electronic
This study was designed to test whether making a selection from a target-absent lineup could alter memory for the actual perpetrator. Three different lineup presentations were used:

1. **Control Lineup**: Participants were shown lineups with one target and three non-targets.
2. **Alter Lineup**: Participants were shown lineups with two targets and two non-targets. One of the targets was the correct perpetrator, and the other was a non-perpetrator.
3. **Misinformation Lineup**: Participants were shown lineups with two targets and two non-targets. One of the targets was the correct perpetrator, and the other was a non-perpetrator. Following the lineup, participants were given misleading information about the perpetrator.

Participants were then asked to identify the perpetrator from a new lineup. The results showed that participants were more likely to identify the non-target as the perpetrator in the Misinformation Lineup compared to the Control and Alter Lineups. This suggests that misinforming participants about the perpetrator can alter their memory, leading to incorrect identifications.
Stages of data transformation

1. **Pre-processing**
   - MARC record
   - MARC to MARCXML converter
   - Normalized MARC XML

2. **Post-processing**
   - Linked Data
   - LOD

3. **LC BIBFRAME converter**
   - BIBFRAME RDF

4. **MARC XML**
Thesis advisor in Bibframe
Stages of data transformation

1. MARC record
2. MARC to MARCXML converter
3. MARC XML
4. Normalized MARC XML
5. LC BIBFRAME converter
6. BIBFRAME RDF
7. Linked Data
8. LOD

Pre-processing
Post-processor
LD4L Linked Data fragment

VIVO linked open data

advisor (foaf:Person)

author (foaf:Person)

owl:sameAs

madsrdf:identifiesRWO

VIVO URL (bf:Authority)

bf:hasAuthority

institution (bf:Organization)

advisor (bf:Person)

author (bf:Person)

pav:authoredBy

bf:creator

CUL catalog record

work (bf:Work)

bf:hasInstance

bf:hasInstance

bf:workTitle

bf:instanceTitle

bf:systemNumber

manuscript (bf:Instance)

electronic (bf:Instance)

bf:Title

bf:Title

bf:Title

bf:systemNumber

CUL bib id

bf:subject

bf:Topic

bf:Topic

bf:Topic
Sample post-processor functions

- Asserts that the bf:Work is an ld4l:Thesis
- Creates foaf:Person from thesis author
  - Links bf:Person to foaf:Person via madsrdf:identifiesRWO / madsrdf:isIdentifiedByAuthority
  - Links work to foaf:Person via pav:authoredBy
- Creates foaf:Person from thesis advisor
  - Re-uses VIVO URI rather than minting a new URI
  - Links bf:Person to foaf:Person via madsrdf:identifiesRWO / madsrdf:isIdentifiedByAuthority
  - Requires entity resolution where same advisor has multiple advisees
Post-processor functions, cont.

- Creates foaf:Organization from degree-granting institution
  - Requires entity resolution where same institution has granted degrees to multiple individuals
- Adds rdfs:labels
  - Application-specific – Vitro depends on them for display
  - Multiple input sources: bf:label, bf:title, bf:Title
LD4L data in Vitro

- Live at ld4l.library.cornell.edu/ld4l-vitro
- Simulates catalog records augmented with LOD
- Links to external LOD – in this case, VIVO
- Then VIVO brings us:
  - Back to additional catalog records
  - Out to the LOD cloud
Future post-processing challenges

• Scale up to real world magnitude of library catalog data
• Extend to full variety of different types of catalog records
• Entity resolution in the real world
• Integrate catalog updates into existing triplestore
• Link to other external data sources for people, organizations, and works
Additional connections to global identifiers for works, organizations, and people

VIVO linked open data

institutions (foaf:Organization)

advisors (foaf:Person)

authors (foaf:Person)

VIVO URL (bf:Authority)

ORCID

VIAF

ISNI

WorldCat Person

schema:creatorOf

WorldCat Work

schema:exampleOfWork

WorldCat URL

manuscript (bf:Instance)

electronic (bf:Instance)

catalyst id

catalyst bib id

catalyst title

catalyst subject

catalyst institution