Overview

● Background and overview
● New: DASH! Production integration
● Newer: BANG! Data analysis and prototype
● Newest: BAMWOW investigation
● Implementation, data, and usability overview
● Lessons learned and future work
Background and Motivation
Background

- Linked Data for Production: Closing the Loop (LD4P3).
  - “LD4P3 aims to "close the loop" to create a working model of a complete cycle for library metadata creation, sharing, and reuse.”
  - Funded by Andrew W. Mellon Foundation
- Discovery
  - Using linked data to support and enhance discovery
  - Continue experiments
  - Work on integration into production
Motivation

- Expand to more open-ended discovery
- Meet users closer to where they are with respect to search terms and categories
- Enable use of library and external data to provide context and relevant relationships
- Support discovery pathways through the catalog using external data
Approach

- Multiple phases of work: production, experimental
  - Each phase has (or will) incorporate user research
- Phases
  - Discovery Across Subject Headings! (DASH!)
    - Context for authors and subjects
  - Bibliographic Access Newly Gui’d! (BANG!)
    - Works and what they may mean for discovery
  - Browsing Across Music With Obtainable Wikidata! (BAMWOW!)
    - Musical works and Wikidata
Implementation Approach

- Extend existing design options and navigational pathways
- Client-side display (for these phases)
- Employ shared identifiers and URIs as the bridge between library catalog data and other sources
Usability and User Feedback Approach

- Decide on objective of user test
- Recruit target audience
- One-on-one interviews: in person or over zoom
- Task list, think aloud
- Debrief and identify key takeaways or patterns
- Iterate and test again
(Partial) overview of discovery system

Discovery layer

Bibliographic index

Author index

Subject index

Authorities

Additional indices not displayed
Discovery Across Subject Headings! (DASH!)
New!
Now in production
DASH! Motivation

- How can contextual information be used to decorate author and subject entities in the catalog?
- How can we use this information to provide more discovery pathways into the catalog?
- How can we extend navigational structures in the catalog to add this context?
Middlemarch

Eliot, George, 1819-1880.
Works by: 170 Title(s)
Works about: 305 Title(s)
Full record

Availability
Qlin Library
PR4662.A2 C37 2019
0 Text
Checked out, due 05/11/21

Request item
Request scan of article/chapter
Middlemarch

"George Eliot; edited with notes by David Carroll; with an introduction by David Russell."

Author, etc.: Eliot, George, 1819-1880, author
Format: Book
Language: English
Copyright date: ©2019
Subject: City and town life > Fiction.
Married people > Fiction.
Young women > Fiction.
England > Fiction.
Bildungsromans.
Didactic fiction.
Domestic fiction.
Romance fiction.
Description: xlii, 810 pages; 20 cm.
ISBN: 0198815514
Where we started

Eliot, George, 1819-1880.

Works by: 176 Title(s)
Works about: 385 Title(s)

Reference info

Alternate form(s):
Cross, Marian Evans, 1819-1880
Cross, Mary Ann, 1819-1880
Eliot, Dzhordzh, 1819-1880
Eliot, Dżordż, 1819-1880
Elįyt, Jąj, 1819-1880
Elįyt, Jój, 1819-1880
Evans, Marian, 1819-1880
Evans, Mary Anne, 1819-1880
Leaves, M. E. (Marian Evans), 1819-1880
Leaves, Marian Evans, 1819-1880
Leíveis, M. E. (Marian Evans), 1819-1880
Leíveis, Marian Evans, 1819-1880

Headsings type: Personal Name

Description
Place of Death: Chelsea (London, England)
Field: Fiction
Occupation: Novelists
Birth Place: Warwickshire (England)
Country: England
Eliot, George, 1819-1880.

Mary Ann Evans (22 November 1819 – 22 December 1880; alternatively Mary Anne or Marian), known by her pen name George Eliot, was an English novelist, poet, journalist, translator and one of the leading writers of the Victorian era. She wrote seven novels: Adam Bede (1859), The Mill on the Floss (1860), Silas Marner (1861), Romola (1862–63), Felix Holt, the Radical (1866), Middlemarch (1871–72) and Daniel Deronda (1876). Like Charles Dickens and Thomas Hardy, she emerged from provincial England; most of her works are set there. She is known for their realism, psychological insight, sense of place and detailed depiction of the countryside. (From DBPedia.)

Occupation: Novelist
Place of Birth: Warwickshire (England)
Place of Death: Chelsea (London, England)
Citizenship: United Kingdom of Great Britain and Ireland*
Educated at: Bedford College*

* From Wikidata

Library Holdings

Total Works By: 171 Titles
Total Works About: 432 Titles
Books (565)
Journals/Periodicals (5)
Manuscripts/Archives (1)
Microforms (5)
Musical Scores (3)
Non-musical Recordings (2)
Theses (25)
Videos (1)

Browse related items by call number:
PR4650 - PR4698
Russo-Japanese War, 1904-1905

Works about: 498 Title(s)

Reference Info

Broader Term: Eastern question (Far East) Works about: 1014
Alternate form(s): Japan > History > War with Russia, 1904-1905
Japanese-Russian War, 1904-1905
Russia > History > War with Japan, 1904-1905

Headings type: Topical Term
The Russo-Japanese War (Russian: Русско-японская война, romanized: Русско-японская война; Japanese: 日露戦争, romanized: Nichiro sensō, lit. 'Japanese-Russian War') was fought between the Empire of Japan and the Russian Empire during 1904 and 1905 over rival imperial ambitions in Manchuria and Korea. The major theatres of military operations were the Liaodong Peninsula and Mukden in Southern Manchuria, and the seas around Korea, Japan, and the Yellow Sea. (From DBpedia.)

Reference Information

Broader Term: Eastern question (Far East) Works about: 1021

Heading Type: Topical Term

Alternate Form(s): Japan > History > War with Russia, 1904-1905
Japanese-Russian War, 1904-1905
Russia > History > War with Japan, 1904-1905

Library of Congress Subject Headings (LCSH)
Now in the Cornell Library Catalog

- Author knowledge panel
- Author details
- Subject browse
- Subject details
Bibliographic Access Newly Gui’d! (BANG!)
Newer!
Experimental Phase
BANG Motivation

● How can/do work relationships enable end users
  ○ To get more context for a library resource?
  ○ Find related library resources?
    ■ (Cornell Catalog already supports “Other forms of this work using OCLC WorkID clusters from 2015)
● Which work relationships can we find in existing BIBFRAME datasets? Are they…
  ■ Consistent?
  ■ Meaningful?
Work Relationships and Aggregations

● Sources
  ○ Library of Congress (LOC) Hubs
  ○ ShareVDE Program for Cooperative Cataloging (PCC) data

● Focused on top level aggregations as a start
● Retrieve clusters of related Cornell catalog resources
Clusters of Related Catalog Records

Hub/Opus

Work → Instance

Work → Instance

ISBN

catalog/id1

catalog/id2
BIBFRAME data analysis

- The Hub & Opus have potential to provide useful clustering of related catalog items
- Can provide specific relationships (e.g. translations, or all forms of same higher level opus or hub)
- Need further analysis to determine overlap with existing OCLC work id set.
  - Individual examples where other sources provide more clusters than OCLC work ids
  - Individual examples where other sources provide fewer clusters than OCLC work ids
Dune

Dune is a 1965 epic science fiction novel by American author Frank Herbert, originally published as two separate serials in Analog magazine. It is tied with Roger Zelazny's This Immortal for the Hugo Award in 1966 and it won the inaugural Nebula Award for Best Novel. It is the first installment of the Dune saga. In 2003, it was described as the world's best-selling science fiction novel. Since 2020, the names of planets from the Dune novels have been adopted for the real-life nomenclature of planets and other features on Saturn's moon Titan.

Author(s): Herbert, Frank.
Language: English.
Subject: Dune (Imaginary place) • Fiction.

**Copies of "Dune"**

<table>
<thead>
<tr>
<th>Title</th>
<th>Location</th>
<th>Format</th>
<th>Condition</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dune</td>
<td>Berkeley, Calif.</td>
<td>Paperback</td>
<td>GD</td>
<td>2nd ed.</td>
</tr>
<tr>
<td>Dune</td>
<td>New York</td>
<td>Paperback</td>
<td>GD</td>
<td>40th anniversary ed.</td>
</tr>
<tr>
<td>Dune</td>
<td>New York</td>
<td>Paperback</td>
<td>GD</td>
<td>50th anniversary ed.</td>
</tr>
<tr>
<td>Dune</td>
<td>New York</td>
<td>Hardcover</td>
<td>GD</td>
<td>Special/hardcover edition</td>
</tr>
</tbody>
</table>

**Translations**

<table>
<thead>
<tr>
<th>Title</th>
<th>Language</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dune</td>
<td>Thai</td>
<td>The Dune, the Desert, Frank Herbert, Frank Herbert</td>
</tr>
<tr>
<td>Dune</td>
<td>Korean</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
<tr>
<td>Dune</td>
<td>French</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
<tr>
<td>Dune</td>
<td>Spanish</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
</tbody>
</table>

**Adaptations**

<table>
<thead>
<tr>
<th>Title</th>
<th>Platform</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dune</td>
<td>Movie</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
<tr>
<td>Dune</td>
<td>TV Series</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
</tbody>
</table>

**Items about "Dune"**

<table>
<thead>
<tr>
<th>Title</th>
<th>Platform</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dune</td>
<td>Book</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
<tr>
<td>Dune</td>
<td>Insight Editions</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
</tbody>
</table>

**Items by Herbert, Frank**

<table>
<thead>
<tr>
<th>Title</th>
<th>Format</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dune</td>
<td>Paperback</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
<tr>
<td>Dune Messiah</td>
<td>Paperback</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
</tbody>
</table>

**Sequels**

<table>
<thead>
<tr>
<th>Title</th>
<th>Format</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dune Messiah</td>
<td>Paperback</td>
<td>Dune (Imaginary place) • Fiction.</td>
</tr>
</tbody>
</table>
Browsing Across Music With Obtainable Wikidata!
(BAMWOW!)
Newest!
Experimental Phase
BAMWOW! Motivation

- Can we incorporate contextual information around musical works into the library catalog?
  - Partner with MLA to leverage already defined data creation projects to see if that data can be used to enhance catalog
  - Gain experience with both music-specific statements in Wikidata and more broadly properties that might apply to all/other types of works
BAM WOW!!
Implementation, Data, and Usability Overview
User testing and user representative feedback

- Possibilities for improving the catalog discovery experience
- DASH! testing (on earlier prototype) helped re-design of info button for more clarity
- DASH! user representative feedback identified which data sources and features to keep
- BANG! testing compared production to prototype
  - Found prototype fared better with several work to work relationships
  - Further work to see which relationships of most benefit to users
Interface implementation details

- Client-side requests
  - *Lookup URIs for string* headings (DASH!, BAMWOW!)
  - *Use URI in MARC* to retrieve related entities (BANG!)
- "Disallowing" or hiding certain information (DASH!)
- Pre-existing separate author and subject indices
  - Note: No dedicated Work index our Blacklight implementation; operates at the publication/manifestation level
- Extend or link from existing catalog pages (e.g. existing browse and item views)
DASH! client side requests

- Eliot, George, 1819-1880
- Lookup URI in LOC
- n79045512

- George_Eliot
- Query with QID, label
- Q131333
- Image
- Description
- Citizenship
- Educated at
- Pseudonyms

- Classification
- Retrieve JSON-LD
DASH! Disallowing Information

? "Saunders, George, 1958-"
: ~

? "Eliot, George, 1819-1880"
: ~

- "image"
- "description"
- "citizenship"
- "education"
- "pseudonyms"
BANG! Sinopia to Catalog

Dune

Work

Dune 1965

OCLC

catalog/57525

Dune 1982

ISBN, OCLC

catalog/58460

Dune Online

LCCN, OCLC

catalog/15139747
BANG!: Connecting Sinopia with Our Catalog

- Sinopia Work URIs in 758 of BIB
- ISBNs, LCCNs, and/or OCLC numbers in Sinopia Instances
Lessons Learned and Future Work
Data reliance and integration

- Going beyond “nice-to-have decorations”
- Recognize/account for strengths/weaknesses in datasets
- Do we support searching on what (up until now) has been considered “external” data?
- Provenance and controlling data usage
Bibliographic data models and user needs

● What’s the most ideal way to present to the user Work level information consistently… how far do we separate/flatten the data describing works vs. publications vs. items?

● The catalog has to balance user needs, rather than faithfully reflect any/all (sometimes complicated) data model/s.
How might this change cataloging practices

- Identifiers in our data make it easier to connect across datasets.
  - Strings to things AND how they are related
- Data quality, comprehensiveness, consistency for new use cases
- Engagement/strategic partnerships with external data producers.
- We can begin to benefit while still heavily relying on MARC
Q&A: This is more of a comment than a question.

Everyone talks about the hoverboard. This is what I want from Back to the Future.
References

- Documentation/wiki
  - LD4P3 Discovery work
- Email
  - astridu@stanford.edu
  - hjk54@cornell.edu
  - sf433@cornell.edu
Additional Slides
DASH! Author Info

Author

- Description
- Educated at
- Citizenship
- Pseudonyms
- Image

LOC Author

LOC ID

P18

P244

P69

P27

P742
## Clusters with >=2 Catalog Records

<table>
<thead>
<tr>
<th>System</th>
<th>Description</th>
<th>Count</th>
<th>Clusters by</th>
</tr>
</thead>
<tbody>
<tr>
<td>ShareVDE Opus</td>
<td>10,000,486 total opera</td>
<td></td>
<td>3,258 ISBN clusters by opus</td>
</tr>
<tr>
<td>LOC hub</td>
<td>2,481,410 total hubs</td>
<td></td>
<td>9,430 LCCN clusters by hub</td>
</tr>
<tr>
<td>OCLC Work ID</td>
<td>513,078 total work ids (in catalog)</td>
<td></td>
<td>195,108 clusters by work id</td>
</tr>
<tr>
<td>Data source</td>
<td>Description for groups</td>
<td>Number of groups matching at least two Cornell catalog records</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>ShareVDE PCC</td>
<td>ISBNs aggregated by <strong>Opus</strong></td>
<td>3,258</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LCCNs aggregated by <strong>Opus</strong></td>
<td>1,226</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISBNs related by work to work relationships</td>
<td>184</td>
<td></td>
</tr>
<tr>
<td>LOC Hubs</td>
<td>ISBNs aggregated by <strong>hub</strong></td>
<td>6,452 (extrapolated)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LCCNs aggregated by <strong>hub</strong></td>
<td>9,430 (extrapolated)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ISBNs related by <strong>hub to hub</strong> relationships</td>
<td>5,459 (extrapolated)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LCCNs related by <strong>hub to hub</strong> relationships</td>
<td>8,189 (extrapolated)</td>
<td></td>
</tr>
</tbody>
</table>