Production: Knowledge panels and author/subject/work pages

This page provides additional details around the integration of information from Wikidata, DBPedia, and the Library of Congress (LOC) linked data service into the Cornell production catalog. We began this work as part of LD4P2 and LD4P3 discovery explorations around knowledge panels and author and subject pages. After multiple discussions with our library user representatives and modifications based on their feedback, we coordinated with the Discovery and Access team, the group responsible for our catalog implementation, to move forward with this integration. These features are subject to change as we continue to support and build on them.

Examples and screenshots

On the page for this Middlemarch item, the user can click on "author info" to open up the knowledge panel for George Eliot.

Clicking on "view full info" will lead to the author page for George Eliot, as you can see below.

Users can also reach this author page through our author browse list here:

Similarly, users can reach the subject page for a heading from the subject browse list (as shown below). We are not currently implementing separate subject info buttons from the item view page. Clicking on the "subject info" button next to the heading "Russo-Japanese War" will take the user to the subject page.

When there is a one to one correspondence between a Name/Title authorized access point and a Wikidata Work, the user can see information from Wikidata inline with the rest of the catalog record, e.g page for The Magic Flute.

Catalog numbers:	Köchel catalogue (1st edition) : 620
First performance date:	30 Sep, 1791
First performance location:	Theater auf der Wieden
Librettist:	Emanuel Schikaneder

Some of this information comes from the Discogs database ↗.
Highlight the Discogs data.

Some of this information comes from Wikidata ↗
Remove the Wikidata highlighting.

For items in the catalog that contain more than one work and have Name/Title authorized access points and matching Wikidata Works, e.g. this Vivaldi con moto item, the user can click on "work info" to open up the knowledge panel for any of the works.

Included work:

Vivaldi, Antonio, 16 minor. Work info »
Vivaldi, Antonio, 16 major. Work info »
Vivaldi, Antonio, 16 major. Work info »
Vivaldi, Antonio, 16 major. Work info »
Vivaldi, Antonio, 16 major. Work info »
Vivaldi, Antonio, 16 minor. Work info »

Vivaldi, Antonio, 1678-1741. | Concertos, violin, string orchestra, RV 232, D major

Works: 1 Title
Works about: 0 Titles

Catalog numbers: Rinaldi catalog : R. op. 36, n. 4
Fanna catalogue : F. I n. 18
Ryom-Verzeichnis : RV 232
Pincherle catalog : P. 200

Tonality: D major

Instrumentation: violin, string orchestra, continuo group

View full info »

Clicking on "view full info" will lead to the work page for any of the works, as you can see below.

Vivaldi, Antonio, 1678-1741. | Concertos, violin, string orchestra, RV 232, D major

Catalog numbers:	Rinaldi catalog : R. op. 36, n. 4 * Fanna catalogue : F. I n. 18 * Ryom-Verzeichnis : RV 232 * Pincherle catalog : P. 200 *
Tonality:	D major *
Instrumentation:	violin, string orchestra, continuo group *

* From Wikidata ↗

Library Holdings

Total Works: 1 Title
Total Works About: 0 Titles

Reference Information

Alternate Form(s):	Vivaldi, Antonio, 1678-1741.	Concertos, violin, string orchestra, F. I, 18, D major Vivaldi, Antonio, 1678-1741.	Concertos, violin, string orchestra, P. 200, D major Vivaldi, Antonio, 1678-1741.	Concertos, violin, string orchestra, R. op. 36, no. 4, D major

Data integration

- Knowledge panels
    - We now include an image and short description from Wikidata.
- Author, works, and subject browse pages:
    - We now include
        - An image from Wikidata
        - A description from DBPedia (or a description from Wikidata if the former is not available)
        - Additional Wikidata information for people such as citizenship, education, and pseudonyms
        - Additional Wikidata information for works such as thematic catalog numbers for musical works, first performance date/location, and dedications.
        - The LOC classification number which is used to generate a link to our call number browse.

Implementation

All of the data on the page is retrieved using dynamic lookups against external data sources to retrieve information. We also check against our configuration to see which data or authorities should show information from external sources. (This process is described in greater detail below). Currently, the author knowledge panel and author and subject pages rely on a string lookup of the authorized heading string against the id.loc.gov suggest service. This lookup returns a URI which we then use to retrieve LOC information such as the classification number and to query Wikidata. DBPedia queries use a combination of Wikidata QID and label searches to retrieve descriptions. Additionally, we have requested the inclusion of LOC identifiers or URIs within the catalog Solr indices and, in the future, we hope to be able to use that information directly instead of relying on string lookups to get LOC URIs .

The image below shows how information is being retrieved for the heading for George Eliot. Blue arrows represent queries or JSON retrieval. The gray boxes include the string heading or the identifiers. Th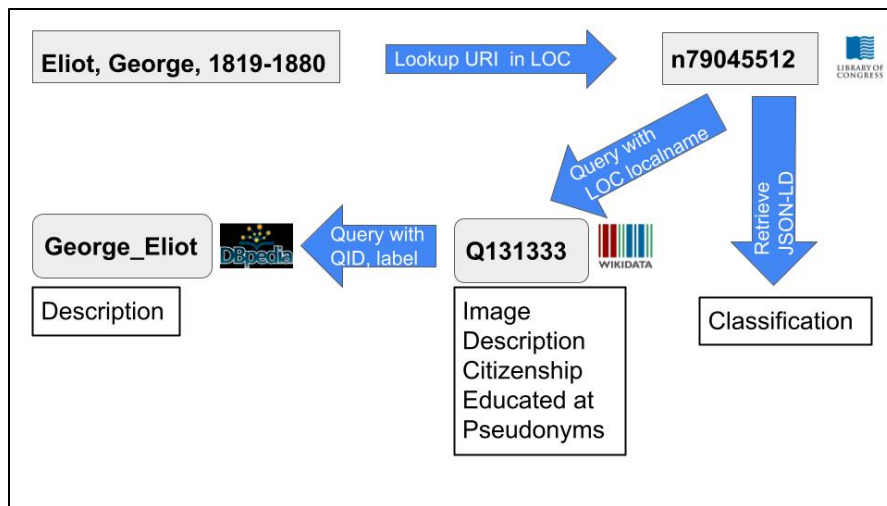e image displays the local names for the URIs to make the image easier to read. The white boxes with black borders underneath the URIs show the properties we are retrieving for that URI.

When we don't want to show specific data

We implemented a system whereby we can exclude information from being displayed for specific authors or subjects. Using a YAML file, we can designate whether we wish to hide all external data or whether we wish to hide certain properties for a specific author or subject heading.

For example, if we want to not display the image and description for the heading "Eliot, George, 1819-1880" but wish to display the remaining properties we retrieve from Wikidata, such as citizenship and place of education, we would include the following in the YAML file:

```
? "Eliot, George, 1819-1880"
:
  - "image"
  - "description"
```

If we did not want to show any of the information we retrieve from Wikidata or DBPedia for this specific heading, we would instead write the following in the YAML file:

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

What happens when external data services are not available

If the external lookups or APIs are not functioning, we try to ensure the knowledge panel and author and subject pages still load with the information from our own browse indices.

Additional changes

We also updated aspects of the design of both the author knowledge panel and the subject and author pages. We incorporated library holdings search information into the author and subject pages to allow users to more easily find related information and resources. Our author and subject browse indices already included the "total works by" and "total works about" information. In order to ensure that the library holdings information search results reflected works by and about the authors, we employed an advanced search using a boolean OR to query for results where the heading was the author or the subject (using the same author and subject fields search fields that are used for the links for "total works by" and "total works about" links).

Code

Related sections of code include:

- Knowledge panel
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/assets/javascripts/pops.js
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/controllers/kpanel_controller.rb
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/views/kpanel/panel.html.erb
- Author and subject pages
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/assets/javascripts/author_browse.js
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/assets/javascripts/subject_browse.js
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/assets/javascripts/retrieve_lcsh.js
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/controllers/browse_controller.rb
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/views/browse/_author_info.html.erb
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/views/browse/_subject_info.html.erb
- Configuration for what not to display for specific headings
    - https://github.com/cul-it/blacklight-cornell/blob/dev/app/controllers/handle_external_data_controller.rb
    - https://github.com/cul-it/blacklight-cornell/blob/dev/public/excludeEntities.yml
- Related additions to routes.rb

    ```
    get 'panel'=> 'kpanel#panel'
    get 'kpanel/index'
    # Permissions for showing certain external info
    get '/check_permissions' => 'handle_external_data#check_permissions', as: 'check_permissions'
    ```

Additional information about specific code changes can be found in the Discovery and Access documentation for this work here.

Thanks and acknowledgements

Special thanks to Tim Worrall, our lead developer on this work. Many thanks as well to our library catalog user representatives and to the entire Discovery and Access team. Thanks to Frances Webb who provided insight into our Solr indices and to Melissa Wallace who provided feedback around design.

On the LD4P3 front, thanks also to our discovery team (Discovery On the Ground), including but not limited to: Astrid Usong for UX work and contributions, Greg Delisle for server infrastructure, Steven Folsom and Jason Kovari for metadata feedback, Tim Worrall for development and design work, Michelle Futornick for contributing to discussions, Hilary Thorsen for her query pointers, and Dave Eichmann for his data indexing work and providing data query support. Thanks also to Lynette Rayle for her support as a colleague and friend. Thanks to Kevin Ford for answering our LOC data questions. As always, thanks to our LD4P3 project directors: Jason Kovari, Simeon Warner, Tom Cramer, Dave Eichmann, Phil Schreur.