Discovery (WP4)

Summary and Background

The use of linked data within a traditional library catalog is the key to broader implementation of linked data within the academic library community. Without a concrete demonstration of the benefits of linked data, administrators are reluctant to commit the time, effort, and funds for the transition from MARC. LD4P2 will enhance the Blacklight open source search engine with linked data features, such as (1) knowledge panel in search results to present contextual information powered by linked data; (2) browsing based on authority files and links to related entities in external data; (3) semantic search, which could include laying out alternative terms in a “no results” page, suggesting semantically related terms in a type-ahead, or providing richer, geo-based browse by leveraging URIs for places; and (4) microdata on item pages to enable machine crawling.

References and links

Partner institutions involved in this area of the grant have individually and collaboratively worked on the usability and user interface design questions around the use of linked data in discovery. We have also been experimenting with prototypes and development to better understand how the integration of linked data can be implemented and what are the resulting possible user interactions. Discussions and participation in the LD4 Discovery Affinity Group have also led to fruitful conversations and refinement of grant-related work. The Blacklight LD Working Meeting is another collaborative community effort to which we hope to contribute and from which we hope to garner additional ideas for further work.

- Google drive for discovery work
    - LD4P2 Discovery Stanford work 8/28/18 meeting notes
    - "On the ground" discussions and work: Collaboration/meetings discussing grant work
        - "On the ground" meeting notes
- Blacklight-LD Working Meeting - September 2019
- LD4 Discovery Affinity Group

- Grant (overall) questions/discussions

    - Initial brainstorming/analysis
    - Knowledge panel examples
    - Browsing examples and related examples
    - Articles/related work review (library catalog design and linked data interfaces)
    - User interviews (mix of researchers/faculty) (Astrid Usong)
    - Knowledge Panel Guerilla testing and results (focusing on students) (Astrid and Ally)

        - Also see Discovery Affinity Group page (referenced below)
- Cornell

    - Overall draft plan for discovery work
    - Cornell code on GitHub
    - KPAOW! (Knowledge Panels Area Of Work)

        - Knowledge panels or KPAOW
        - Motivating search examples
        - Mockups for evaluation - thank you Astrid Usong !!
        - KPAOW Mockup Usability Results . A big thank you to Evelyn Hudson, Library Research Assistant
        - KPAOW Demo (With Discovery Theatre Intro + Closing)
            - Version Without Theatre
        - Lessons learned
    - BAM! (Browsing Assorted Materials)
        - Mockup evaluations (Astrid Usong, Huda Khan and Filip)
        - BAM! Demo
        - Lessons Learned
    - SMASH! (Semantic Maneuvering Across Subject Headings)
        - Usability Results
        - Lessons Learned
        - Annif Use and Explanation
        - SMASH! demo
    - WHAM! (Why Having AutoSuggest Matters)
        - Organizing document with links to meetings and documentation
        - Indexing Process Documentation Google slides and PDF
        - Motivating Use Cases
        - Script documentation
        - Usability Results
        - Lessons Learned
        - Packaging as Blacklight gem
        - WHAM! demo
        - A note on Discogs work
            - We examined whether it made sense to move the AJAX calls in the JavaScript into a controller. After doing some initial work on this, there would be both performance and UX improvements. So if we do anything with Discogs going forward, we should complete this update to the code.
    - Cornell Focus Group meeting
        - Google drive with notes
        - Report outs and notes
    - Code note
        - Code for these features is integrated in this branch: https://github.com/ld4p/blacklight-cornell/tree/ld4p2

        - One of our team members, John Skiles Skinner, had worked on some subject facet exploration. This is not included in the main LD4P2 code but is available in this branch: https://github.com/LD4P/blacklight-cornell/tree/era-subject-browse
- Stanford
    - SearchWorks LD (Stanford) Guerilla testing and results (Astrid Usong)
    - SearchWorks LD code (thanks Jessie Keck for links below)
    - SearchWorks LD prototype
    - SearchWorks LD Work Cycle demo

Description from the grant

Blacklight has hundreds of installations worldwide and particularly deep adoption in North American research libraries, and in conversations at ALA Midwinter 2018 about the next phase of LD4P, Blacklight discovery captured the imaginations of a number of libraries.

Potential discovery enhancements to Blacklight include (1) knowledge panel in search; (2) browsing based on authority files and links to related entities in external data; (3) semantic search; and (4) microdata on item pages to enable machine crawling.

Items 1 (knowledge panel in search results to present contextual information powered by linked data) and 4 (microdata) have proven implementations in other discovery environments, and will be the first targets for development.

Item 2 (browse) offers the richest potential benefits to users. This effort will require balancing the capabilities of the data, user needs, design of an intuitive and useful interface, and technical constraints of current technologies. Challenges and questions around designing and implementing an effective browse experience include:

- Is it viable to integrate a network graph powered by semantically enriched library metadata? Would such a tool prove useful and intuitive to library users in a general purpose discovery environment?
- How can concatenated, LCSH subject headings best be represented in a browse interface? Is it possible to break them into discrete elements (such as with FAST) and enable toggling on and off discrete elements? If so, will it be intuitive?
- For geographic browse, what kinds of visualizations will allow for users to navigate for works about or by creators in certain areas? Will points-on-maps provide enough / useful visualization, or will regions be needed (and more useful)? How will we cluster or show density of terms—via a cluster of pins or perhaps a heatmap?
- Can we implement a NetFlix or Amazon-style browse that will meet patron needs and expectations for scholarly browse, and also be based on the catalog data we have?
- Does the growth in use of mobile device users of the catalog change anything we might do for browse user interfaces?

Item 3 (semantic search) also has promise, though the best examples of semantic search happen in more tightly scoped information domains (such as medical search using MeSH). Analysis and experimentation will be required to see if subject searches and headings will lend themselves to a semantic approach in a general catalog environment. We intend to explore geographic search as a first and promising entry point to this approach. Analysis will include how patrons interact with the library catalog, especially for subject searches and searches resulting in zero results, to create a better, user-centered design; log analysis of current searches, as well as data analysis to determine what search arguments could be transformed to alternative searches with richer results using linked data. It will also include exploring technical approaches to semantic search, including adding semantically-related terms at index time, leveraging dynamic look-ups of linked terms via external services, and performant approaches to including semantic type-ahead / autosuggest of search terms.

This analysis and experimentation will be done by a team combining user experience designers, metadata experts, and software engineers. Design and development will happen in an iterative manner as the group explores objectives and possible approaches; trials implementation; assesses effectiveness; and refines for better results.

Development of the Blacklight enhancements will be supported by the developers, UX designers and technical services at Cornell and Stanford. The development of semantic search and browse mechanisms, including visualizations where useful, will be supported by the Iowa team.