Mapping Linked Data Subject Headings in the Library Catalog

Patrick Burns
patrick@diyclassics.org
Institute for the Study of the Ancient World Library
United States of America

Gabriel McKee
gm95@nyu.edu
Institute for the Study of the Ancient World Library
United States of America

David M. Ratzan
dr128@nyu.edu
Institute for the Study of the Ancient World Library
United States of America

Tom Elliott
te20@nyu.edu
Institute for the Study of the Ancient World
United States of America

The MARC-based library catalog is text-based: in order to find relevant information, a researcher must use text to identify an author, title, or subject. Our team has been expanding the paradigm of text-based discovery by exploring visual discovery, building a browsable map of our library catalog based on record authority and linked open data. This geospatial representation of our holdings not only offers researchers a new mode of discovery; it also opens the door for new avenues of research by highlighting unexpected connections and virtually collocating materials that are classified and shelved separately. The process is as follows. 1. Newly catalogued items are assigned a stable URI from Pleiades, the open-access ancient world gazetteer (Bagnall et al., 2016), reflecting our Institute's scholarly focus and library holdings. 2. This identifier is added as a subject heading in the item's MARC record. 3. This data is exported, cross-referenced with geoJSON records containing latitude and longitude data (Baumann, 2014), and mapped using Leaflet.js (Agafonkin, 2016). The mapped representation of the catalog serves as an alternative mode of discovery for researchers, who can now browse for library materials by focusing on, for example, a general region of the Near East, a city in ancient Egypt, or a specific archaeological site in China. The experimental nature of this map-based discovery has become all the more viable as the Library of Congress has recently added Pleiades to its list of subject heading authorities. (Library of Congress, 2016) So, while the inclusion of linked data in MARC records has seen some adoption in recent years (Papadakis et al., 2015), this imprimatur from the LOC for Pleiades opens up massive potential for geographical linked data specifically. Other projects have mapped LOC subject headings (e.g. Freeland et al., 2008; Bennett et al., 2011), but this is the first project to do so directly from linked data embedded in MARC records. Accordingly, this project heeds the recent call of the MIT Future of Libraries report (MIT Libraries, 2016) to provide "comprehensive, accessible, digital content" in library discovery that supports the ability to "combine, manipulate, [and] visualize" library data for the global community. Our poster includes: an explanation of the linked-data principles underlying the project, our visualization workflow, and an example of mapped catalog data from books acquired in the fourth quarter of 2016, presented both as a static map on the poster as well as a live, browsable demo on a tablet.

Bibliography

Agafonkin, V. (2016). Leaflet : An Open-Source JavaScript Library for Interactive Maps. http://leafletjs.com/ (accessed 28 March 2017).

Bagnall, R.S., et al. (2016). Pleiades: A Gazetteer of Past Places. http://pleiades.stoa.org/ (accessed 28 March 2017).

Baumann, R. (2016). Pleiades-Geojson. https://github.com/ryanfb/pleiades-geojson (accessed 28 March 2017).

Bennett, R., O'Neill, E.T., Kammerer, K., and Shipengrover, J. D. (2011). mapFAST: A FAST Geographic Authorities Mashup with Google Maps, The Code4Lib Journal, 14. http://journal.code4lib.org/articles/5645 (accessed 28 March 2017).

Freeland, C., Kalfatovic, M., Paige, J., and Crozier, M. (2008). Geocoding LCSH in the Biodiversity Heritage Library, The Code4Lib Journal, 2. http://journal.code4lib.org/articles/52 (accessed 28 March 2017).

Library of Congress, Network Development and MARC Standards Office. (2016). Technical Notice September 15, 2016: Additions to the MARC Code Lists for Relators, Sources, Description Conventions. https://www.loc.gov/marc/relators/tn160914src.html (accessed 28 March 2017).

MIT Libraries, Ad Hoc Task Force on the Future of Libraries. (2016). Institute-Wide Task Force on the Future of Libraries—Preliminary Report https://future-of-libraries.mit.edu/sites/default/files/FutureLibraries-PrelimReport-Final.pdf (accessed 28 March 2017).

Papadakis, I., Kyprianos, K., and Stefanidakis, M. (2015). Linked Data URIs and Libraries: The Story So Far, D-Lib Magazine, 21(5/6). http://www.dlib.org/dlib/may15/papadakis/05papadakis.html (accessed 28 March 2017).

5 The Library Linked Data Cloud. 5.1 A Look Back and Forward. 5.2 Next Steps. Â From a linked data perspective, this result implies that no resolvable URI can be assigned to nearly 94% of the LCSH headings accessible from the WorldCat cataloging data because they do not have unique identifiers or authoritative descriptions. This analysis also reveals that most of the unauthorized headings are topical subjects. They are an expected consequence of the LCSH design, which has been labeled â€˜syntheticâ€™ by classification experts because it features a small set of editorially maintained terms and a rich set of production rules applied by catalogers to create unique headings that The Linked Data Service provides access to commonly found standards and vocabularies promulgated by the Library of Congress. This includes data values and the controlled vocabularies that house them. Datasets available include LCSH, BIBFRAME, LC Name Authorities, LC Classification, MARC codes, PREMIS vocabularies, ISO language codes, and more. Â Everything Audio Recordings Books/Printed Material Films, Videos Legislation Manuscripts/Mixed Material Maps Notated Music Newspapers Periodicals Personal Narratives Photos, Prints, Drawings Software, E-Resources Archived Web Sites Web Pages 3D Objects. Suggestions enabled. Search. The project maps existing Library of Congress Subject Headings (LCSH) terms for University of Pennsylvania manuscripts. The LCSH terms, plus other MARC metadata identified as significant, were rendered as free text and keywords stored in more. The project maps existing Library of Congress Subject Headings (LCSH) terms for University of Pennsylvania manuscripts. The LCSH terms, plus other MARC metadata identified as significant, were rendered as free text and keywords stored in TEI XML to a target set of keywords (a controlled vocabulary known as the Bibliophilly keywords) based on those use