2019 LD4 Conference on Linked Data in Libraries

shortcut to this page: http://bit.ly/ld4conference

Photos by Merrilee Proffitt unless otherwise indicated.
<table>
<thead>
<tr>
<th>Date</th>
<th>2019 May 10-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Joseph B. Martin Conference Center at Harvard Medical School, Boston, Massachusetts, USA</td>
</tr>
<tr>
<td>Map and directions</td>
<td>Harvard Longwood Campus Map</td>
</tr>
<tr>
<td></td>
<td>Directions to Martin Center</td>
</tr>
<tr>
<td></td>
<td>Parking options (Please note: Parking will not be available at the Martin Center itself, and driving is discouraged due to Boston traffic.)</td>
</tr>
<tr>
<td>Visiting Boston</td>
<td>There are no organized tours as part of the conference, but these links can help you plan your own tours: Greater Boston Convention &amp; Visitors Bureau <a href="https://www.bostonusa.com">https://www.bostonusa.com</a>, Freedom Trail: <a href="https://www.thefreedomtrail.org/">https://www.thefreedomtrail.org/</a>, Black Heritage Trail: <a href="https://www.nps.gov/boaf/learn/historyculture/upload/BOAFmap1.pdf">https://www.nps.gov/boaf/learn/historyculture/upload/BOAFmap1.pdf</a></td>
</tr>
<tr>
<td>Participants</td>
<td>Participant List</td>
</tr>
<tr>
<td>Friendly space policy</td>
<td>Friendly space policy for LD4 Conference</td>
</tr>
<tr>
<td>Twitter</td>
<td>#ld4</td>
</tr>
</tbody>
</table>

[ [ Friday, May 10 ] [ Saturday, May 11 ] [ Accommodations ] [ Participation ] [ Travel Stipends ] [ Program Committee ] ]

**Friday, May 10**

<table>
<thead>
<tr>
<th>Room</th>
<th>Rotunda</th>
<th>Room 215</th>
<th>Room 217</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Video playlist</td>
<td>Video playlist</td>
<td>Video playlist</td>
</tr>
<tr>
<td>08:30</td>
<td>Check-in &amp; Coffee</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:00</td>
<td>Welcome</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
LD4P began as a tool for the transformation of library metadata production from workflows based in the MARC formats to linked data. As the project evolves, however, it becomes an opportunity to reevaluate the library’s role in a developing, worldwide information ecosystem.

While much of the world knows about Wikipedia, the emergence of Wikidata as a key global structured data project has only recently emerged as a key way to engage libraries, archives and museums. Andrew will discuss the ways in which the Wikimedia movement has adopted Wikidata and how it interfaces to institutions and collections through linked open data initiatives and innovative reuse. He will talk about notable projects that showcase these and a vision for the future which includes projects such as Structured Data on (Wikimedia) Commons, a global citation database in WikiCite and unified collections contributions workflows now being engineered tougher by the Wikimedia community and GLAM institutions.

Through considering a case study written for the Design for Diversity project, we will consider how metadata and aggregation across collections has simultaneous potentials: to perhaps surface a more diverse range of histories and cultures; to perhaps surface those histories but through metadata that still lacks cultural relevance or respect; or to perhaps only re-inscribe the largely white, largely male histories represented in U.S. library, archive, and museum collections.

Amanda Rust will first briefly introduce the Design for Diversity project, and Dorothy Berry will then discuss her work making African American materials more discoverable through digitization and metadata aggregation in Umbra Search.

Break
The larger questions of how and where to augment discovery processes with linked data are potentially relevant to end-users and institutions both within and beyond GLAM areas. We propose a session where we can present the design questions, processes, and findings from our LD4P2 work and begin a discussion around related discovery interfaces that use external data and/or linked data to support end-users.

This talk will present on work at UW-Madison to enhance the library catalog with Linked Data-derived info cards for the authors, contributors and topical names found in bibliographic works. This feature was added to the production search interface for the UW-Madison Libraries in January and Steve will discuss how the feature works as well as an assessment of it by library staff and patrons.

Tom Cramer, Jessie Keck, Huda Khan, Dan Scott, and Simeon Warner on knowledge panels.

I will present analysis and recommendations compiled by the SHARE-VDE Transformation Council’s working groups regarding SHARE-VDE data modelling, with special focus on SuperWorks/BF:Works and Master Instances. Further activities and analysis planned thanks to the SVDE Working groups directions will also be briefly outlined.

This presentation explores the intersection of philosophical theories of time and their experimental serializations in RDF and related languages. Implications for adoption of particular models is addressed as well as future areas for exploration.

We consider rich semantics beyond typical approaches to linked data. For instance, we have explored supporting access to the NEH/LC NDNP digitized historical newspaper collection. Retrieval with dynamic, structured “community models” seems more promising than traditional indexing. We also applied rich semantic modeling to scientific research reports so that they would be implemented as knowledgebases rather than as text. Upper ontologies, frame semantics, and object-oriented programming-language semantics are among the approaches that may be incorporated into such “direct representation”.

This presentations describes a six-month project to build an intermediate knowledge layer between a knowledge base of classical texts built using FRBRoo and the CITE architecture and the Perseus catalog of published works. The process entails automatic extraction of works and expressions from multiple MODS records and their recombination into FRBRoo statements, which were then made available through a SPARQL endpoint.

UC San Diego’s participation in the LD4P2 Cohort will revolve around describing Open Access (OA) resources. This presentation will cover some of the driving forces behind this focus on Open Access materials, the opportunities and challenges we see for describing them with the BIBFRAME model, and our current training and preparation for the release of the Sinopia editor.

The Wikidata Tutorial will provide a hands-on introduction to the basics of Wikidata and the Wikidata Query Tool. Participants will learn the basic functions of both and how to make simple manual edits to Wikidata items and create simple database queries. The workshop will also discuss how institutions can use Wikidata effectively and employ automated tools like the Distributed Game and OpenRefine. A laptop or mobile device will be necessary to fully participate in the workshop.
LITMUS is an open platform, and its success is due to the expertise of its contributors, including Facilitator Nancy Fallgren of the Rare Materials Affinity Group. This session will explore the process of creating linked data for a unique collection, with a focus on Dublin Core metadata and RDF. The presentation covers current work related to the EU-funded LITMUS (Linked Irish Traditional Music) project at the Irish Traditional Music Archive (ITMA) in Dublin, Ireland.

This presentation covers current work related to the EU-funded LITMUS project at the Irish Traditional Music Archive (ITMA) in Dublin, Ireland. LITMUS is a two-year effort to create linked data tools specific to the needs of Irish traditional music and dance. These linked data tools included the first ontology specific to Irish traditional song, instrumental music and dance as well as two bi-lingual thesauri for traditional instruments and tune types. In this presentation, I will focus on linked data considerations particular to Irish and other traditional musics, and how these challenges were met within the context of the LITMUS project. Examples drawn from LITMUS will illustrate over-arching concepts related to linked data relationships in music and consider future directions for such work.

This presentation will describe steps taken to integrate Discogs as a copy cataloging source in an RDF cataloging tool for describing Sinatra 45 RPM vinyl records. Because Discogs does not natively offer its data as RDF, it might not seem as an obvious choice for use. However, this data was chosen for conversion to RDF in this particular project will be provided, along with lessons learned about generalizing this workflow.

This session will not be recorded. Representatives of the Rare Materials Affinity Group will participate in a facilitated discussion session on the topic of how linked data might not play well with special collections cataloging. We will examine ways that linked data promises to make our materials more accessible to users, and identify potential areas of tension with long-standing descriptive bibliographical tradition.

For those with a basic understanding of the design and organization of Wikidata items, this session introduces the best practices for working with collections of data and their associated ontologies. Ideally, attendees will bring a laptop device to better experiment and try out the tools introduced, but a mobile device can also suffice. The session will cover key tools for discovering, uploading, fixing and reusing Wikidata content for collections, and include notable case studies with GLAM institutions such as the Smithsonian Institution and The Metropolitan Museum of Art.

Tools to be covered include:

- Wikidata user-selectable gadget such as Recoin, EasyQuery and Hotcat.
- WikiProjects that
- SQID for browsing Wikidata entity relationships
- Wikidata Query advanced methods
- Wikidata Graph Builder for investigating ontologies
- TABernacle for editing and fixing collections data
- Quickstatements for bulk uploading and item modification
- Wikidata Game for mobile-friend crowdsourced Wikidata contributions and fixes
- Petscan for advanced querying across Wikimedia projects
- PAWS and Python tools for advanced scripting and needs

**Library of Congress Special Topics**
Facilitator: Greg Reeve

This is a pilot project the USF Libraries Linked Data Team has done to experiment transforming the digital collections into linked data. The team chose a small oral history collection to work on and the team was able to reconcile the data, transform the data into triples and design SPARQL queries to support basic search. Throughout the process, the team was inspired to streamline the workflow and further enhance the final product.

A case study of introducing linked data concepts to a cataloging request and the successful partnership between cataloging and a campus data archive.

**Omeka Tutorial**
Facilitator: Michelle Futornick

In this hands-on workshop (using the sandbox at https://omeka.org/s/download/#sandbox) we will begin with a brief introduction to the popular digital humanities publication platform Omeka S and how it implements Linked Open Data. We will focus on content creation and, of course, metadata creation options using a variety of LOD vocabularies, and how content and metadata are represented in the JSON-LD-based API. To get the most out of this tutorial, participants should bring a laptop and have access to a few files (images or other content) and some metadata for those files.
The National Archives API: A Five-Year Journey from Idea to Imperative (Video)
Facilitator: Will Kent
This session is about the story of the National Archives’ first catalog API—our design choices, use cases, philosophies—and how it has evolved over 5 years of development and use, and become engrained in our daily work. It is a story not just about the API itself, but how the act of designing an API from scratch has provoked us to change old ways of thinking about discovery, reference, and, ultimately, archival work itself.

Digital Collections & Institutional Repositories 2 (Video)
Facilitator: Michelle Durocher
This lightning talk will discuss the Perseus Catalog, a research project of the Perseus Digital Library, and current work to convert the legacy metadata collection and related bibliographic data to linked data standards.
Starting in 2018 as part of the Cultivating a Latin American Post-Custodial Archival Praxis grant funded by the Andrew W. Mellon Foundation, the LLILAS Benson Post Custodial project team began working on developing and migrating the Latin American Digital Initiatives (LADI) digital repository to the Drupal 8 /Islandora 8 (formerly CLAW)/Fedora 4 repository framework. One core component of this work includes investigations and implementation of linked data capabilities for better discoverability, access, and analysis.

Managing Local Data (Video)
Facilitator: Michelle Futornick
Implementation of BIBFRAME or other linked data cataloging workflows at scale will require institutions to address both how the data will be managed over time and how administrative activities that have been driven by MARC data will need to adapt. Bring your ideas, concerns, and questions to this conversation about what is necessary in the short-term as we experiment with new approaches to description and what is required in the long-term to make adoption of best practices feasible. Potential discussion questions include:

- If you are cataloging materials in a linked data editor, does your institution plan to represent those materials in the current integrated library system, or other local systems?
- Is it best practice to continue to manage a MARC dataset alongside RDF entities or other linked data datasets? Or, will it be best to transform historic data so that bibliographic data is uniformly represented in a local system? What are the risks and benefits associated with each of the options?
- What are the functional requirements for managing linked data for description in a system such as FOLIO that is intended to be format agnostic for bibliographic data?

Saturday, May 11

<table>
<thead>
<tr>
<th>Room</th>
<th>Rotunda</th>
<th>Room 216</th>
<th>Room 217</th>
</tr>
</thead>
<tbody>
<tr>
<td>Video playlist</td>
<td>Video playlist</td>
<td>Coffee</td>
<td>Coffee</td>
</tr>
</tbody>
</table>
### Authorities

**Videos:** Hanson; Panigabutra-Roberts; Chou; Ho plus Q&A beginning; Q&A continued

**Facilitator:** Christine Fernsebner Eslao

While at Johns Hopkins University, I oversaw the creation of a specialized authority file that was used to normalize user-submitted data and generate a standardized list of faculty names associated with the electronic theses and dissertations collection. Using Python and RDF, simple workflows were used to normalize variant forms that appeared in DSpace as well as update the RDF file when new variations were entered by submitters.

“Enhancing Opaquenamespace.org: Refinement of Local Name Authority Files and Workflows” will discuss the creation and development of Opaquenamespace.org, a controlled vocabulary manager for LOD local name authorities and vocabularies. She will discuss workflows and processes for URI creation, management, and dissemination, as well as address mistakes that were made as early adopters of LOD for other institutions to consider before creating and managing their own local name authority files. This session will not be recorded.

My talk will focus on my experience in learning how to apply Wikidata to build an author’s profile in Scholia. I will also discuss the learning curve to create new Wikidata and to update the existing ones to connect with the new items. I had also explored different tools, queries and visualization in Wikidata. In conclusion, I will discuss my take away and possible scenarios we can use Wikidata for our collections at UT Libraries.

This presentation examines and explores how to improve the linking relationships of persons with more than one identity in VIAF through definition comparisons and case studies in terms of the semantic web. In light of identity management in cyberspace, library authority data should include the perspectives of privacy and security.

This presentation will describe plans that are being made for the creation of an “authority control” app for names within repositories at Texas A&M University. The purpose of this app is to serve as a tool for library staff, external metadata providers, and repository users to identify and disambiguate individual persons, as well as link to contextual data in external sources (e.g., Library of Congress Name Authority File, ISNI, ORCID, etc.).

### Special Formats 2 (Video)

**Facilitator:** Mary Seem

How do we get from here to there? Our existing infrastructure is heavily based on MARC, but the MARC format itself has evolved over time that can help smooth the path to adoption of linked data. This presentation will describe recent work by the PCC on incorporating linked data into MARC and suggest directions that cataloguing practice can take in response to these developments.

The talk will cover the process of creating video game controlled vocabulary in Wikidata and the translation of the vocabulary to 6 languages (English, German, French, Japanese, Chinese, and Korean).

The Harvard Map Collection is home to over 400,000 sheet maps dating back to the late 15th Century. Over 5,000 of these maps have been scanned and are freely available for viewing through the Harvard Library image delivery service but remain largely undiscoverable outside of the Harvard Library online catalog. This talk will focus on describing an LD4P2 subproject that is exploring the use of the LD4P Sinopia Linked Data Editor with the BIBFRAME 2.0 ontology and Wikidata items to produce linked data descriptions for scanned cartographic materials to enable broader discovery and re-use of these images on the open web.

Serials cataloging is NLM’s bread and butter because of the importance of serials to the medical community and internal dependencies related to article indexing. This is an update on the status of NLM’s LD4P2 Cohort task to improve serials cataloging in BIBFRAME.

The work of the American Beat Generation author William S. Burroughs is notorious among collectors and librarians for its scale and complexity. In this presentation, we describe our objectives and experiences in representing, enhancing, and publishing the information in Brian E. Schottlaender’s “Anything But Routine: A Selectively Annotated Bibliography of William S. Burroughs, version 4.0” as linked data using BIBFRAME 2.0 and ARM ontologies.

**Women Writers in Review** is a collection of 18th- and 19th-century reviews, publication notices, literary histories, and other texts responding to works by early English-language women writers. It is published by the Northeastern University Women Writers Project with support from the Digital Scholarship Group at Northeastern University Libraries, and grant funding from the National Endowment for the Humanities. This session provides an introduction to the data, an explanation of how a digitized evaluation of reception of these works enriches the understanding of the data, and reviews potential ways of modeling it on Wikidata. It also addresses how development of a Wikidata model for this purpose could be adapted to other scholarly digital collections.

### Application Profiles (Video)

**Facilitator:** Jennifer Baxmeyer

The concept of the (metadata) application profile has for two decades been a central focus of attention in the Dublin Core community and has underpinned many of DCMI’s development efforts. There continues to be significant community interest in developing tools to help people create and document application profiles and, more recently, in technologies for validating data produced according to profiles. This talk will describe a new initiative to respond to this interest.

This talk presents the work of the LD4P Profiles Working Group to develop initial protocols for collaborative development of profiles in the LD4P Sinopia shared environment. Attendees will (1) get a better sense of what a BIBFRAME profile is and how it can be managed; (2) understand the challenges of working with profiles on a shared environment; (3) learn about strategies and best practices for profile re-use; and (4) see a live-demo of the tools and the workflow.

### Break
This presentation highlights some key lessons from our experiences in the OCLC Research’s Linked Data Wikibase Prototype (“Project Passage”) regarding Wikidata’s multilingualism support.

At WikiCite 2018 Christine Fernsebner Eslao and Honor Moody worked to develop tools for preliminary reconciliation and ingest of local MARC authority records into Wikidata using MarcEdit and OpenRefine. In the 6 months since they’ve learned that it’s more complicated than they first thought, and “easy to set up” doesn’t always mean what Magnus Manske thinks it does.

This session will take the form of an open discussion of the opportunities and challenges presented by Wikibase, the open source software for managing structured data. A brief overview and survey of current activity related to libraries will be presented and the presenters will engage in an open discussion of possibilities, critical issues and questions on the use of Wikibase within the library and broader GLAM context. Particular emphasis will be placed on opportunities for its use with marginalized community data, specifically their experience piloting Wikibase for Canadian Indigenous community data.

The Canadian Federation of Library Associations launched the BIBFRAME Readiness Workgroup in late 2018 with the mission to assess the readiness of Canadian libraries to adopt BIBFRAME. In 2019, the workgroup will survey Canadian libraries to measure their readiness for a transition to, and their understanding of, BIBFRAME. These findings, analyzed by demographics including library type, library size, participant role, and region, will inform the Canadian library community's efforts to support BIBFRAME adoption.

Brief report from the EBW initiative, and further steps to facilitate international discussion and exchange of experiences on BIBFRAME implementations.

A Linked Open Data Working Group formed in 2011 within the international Ex Libris customer organization IGeLU, and in 2013 became a joint working group with the separate North American customer organization ELUNA. Its purpose has been to work with the vendor to advocate, provide use cases, and consult on development for changes to its products to enable customers to do linked data development with our data, and also to incorporate linked data based features into the products themselves. This talk explores how the relationship with Ex Libris has evolved over time and with changes in the larger library and linked data environment.

Lunch and Birds of a Feather
Topics to be chosen by participants
Birds of a Feather sessions will not be recorded.

Pick up lunch

Birds of a Feather 1 (Rotunda) | Birds of a Feather 2 (Room 214)

Birds of a Feather 3
Data: organizing analysis/assessment of data sources (Wikidata, DBPedia, etc.) and how to connect MARC records to URIs from relevant data sources.

Knowledge panels / displays of external data in discovery interfaces: assessing success / ideas / questions

Affinity Group "in-person"

Break
Accommodations

We have reserved a block of rooms at the Inn at Longwood Medical, which is in walking distance to the conference center.

Inn at Longwood Medical
342 Longwood Avenue, Fenway Kenmore
Boston, Massachusetts 02115

The Inn at Longwood Medical is now sold out; to check for cancellations call them directly (617) 731-4700, our booking code is STAN0519
Linked data promises to expose the richness of library collections to the world, and to open up new pathways to knowledge based on previously unlinked data. After a decade of experimentation and pilot projects, what are the next steps to move to large-scale production of linked data? How can the library community learn from and contribute to other communities who are working toward similar goals? The 2019 LD4 Conference, to be held May 10-11 in Boston, Massachusetts, USA, aims to bring practitioners together to collaborate on creating pathways to implementation of linked data in libraries.

The language of the conference is English.

Participation

Participation is by open application; the application period ended February 28, 2019. The Program Committee invited participation that will bring diverse and broad perspectives, from both aspiring and experienced practitioners including:

- Librarians working (or aspiring to work with) linked data
- Representatives from adjacent cultural heritage spaces including archives and museums
- Ontologists and data modelers
- Software engineers and user experience professionals interested in library metadata
- Open knowledge advocates
- Others who share an interest in implementing linked data in libraries

We had a large response to our call for applications, and the Program Committee chose the applications that best fit within the conference scope.

For questions about the LD4 Conference, please contact the conference co-chairs at ld4conf_chairs@googlegroups.com

Travel Stipends

If you have received a travel stipend, follow Travel Stipend Instructions to claim your stipend.

Program Committee

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justin Sègbédji Ahinon</td>
<td>AfricArxiv</td>
</tr>
<tr>
<td>Asaf Bartov</td>
<td>Wikimedia Foundation</td>
</tr>
<tr>
<td>Amber Billey</td>
<td>Bard College</td>
</tr>
<tr>
<td>Nurunnaby Chowdhury</td>
<td>Centre for Open Knowledge</td>
</tr>
<tr>
<td>Christine Fernsebner Eslao</td>
<td>Harvard University</td>
</tr>
<tr>
<td>Michelle Futornick, Co-Chair</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Myung-Ja (MJ) K Han</td>
<td>University of Illinois</td>
</tr>
<tr>
<td>Will Kent</td>
<td>Wiki Education</td>
</tr>
<tr>
<td>Jason Kovari, Co-Chair</td>
<td>Cornell University</td>
</tr>
<tr>
<td>Andrew Lih</td>
<td>Wikimedia DC</td>
</tr>
<tr>
<td>Merrilee Profitt</td>
<td>OCLC Research</td>
</tr>
<tr>
<td>Greg Reeve</td>
<td>Brigham Young University</td>
</tr>
<tr>
<td>Mary Seem</td>
<td>The Frick Art Reference Library</td>
</tr>
<tr>
<td>Lars G. Svensson</td>
<td>Deutsche Nationalbibliothek</td>
</tr>
</tbody>
</table>