Notes on Operations

Collaborative Learning on Linked Data through a Virtual Study Group

Xiping Liu, Sharon Reidt, Jodene Pappas, Jill J. Crane, and Ada Laura Ramirez

This paper examines the challenges and successes of a virtual Linked Data Study Group that began at a multi-campus academic institution in 2018, and later grew to include a total of seven librarians from multiple institutions across the country. It describes the group's planning for their monthly meetings and the discussions at the meetings which covered such topics as Linked Data basic concepts, BIBFRAME Editor, Sinopia Editor, and Wikidata. It also presents a collaborative project the group undertook after two years' learning. The paper concludes with a summary of what the Linked Data Study Group has achieved thus far, the challenges they faced, and their future plans.

Linked Data emerged in the library's digital and cataloging landscape more than ten years ago. Since then, librarians have been eager to learn and understand how Linked Data works. This desire drove four cataloging librarians from the University of Houston Libraries to form a study group. The group created a study plan, with monthly learning themes and reading lists, to prepare for the anticipated switch from MARC to BIBFRAME. Topics ranged from exploring basic Linked Data concepts to using tools for MARC to BIBFRAME conversion.

After the first three months of discovery and learning, the group presented their initial work at the state library association annual conference in 2019. Since then, three additional librarians from other institutions who were also interested in learning about Linked Data joined the group. The expanded group decided on monthly topics of discussion and started their first meeting in late fall that year. In the following months, the group reviewed the basics of Linked Data, explored Sinopia Editor and Wikidata as accessible means of gaining hands-on Linked Data practice, and joined the library metadata community's discussion on the newly published book *Linked Data for the Perplexed Librarian*.[1] This work led to a presentation at the 2020 LD4 Conference on Linked Data in Libraries. In the following year, the study group collaborated on a Wikidata project as part of the PCC (Program for Cooperative Cataloging) Wikidata Pilot. The study group is currently taking a hiatus before deciding what their future plan is.

The paper begins with a brief overview of the literature on Linked Data's development and librarians' collaborative efforts in Linked Data exploration, training, and education. It then describes how the study group was initially started, how it expanded, and what the group worked on in the two years since its inception. It describes a collaborative project the group undertook as a result of the two years' learning. The paper ends by presenting the benefits of a virtual study group, the challenges faced by the team, and their future plans.

Literature Review

Linked Data is a concept that was first introduced by Tim Berners-Lee in 2001 as the key component of Semantic Web, a web of structured data which will replace the traditional web of documents and allow for meaningful searching, data sharing

Xiping Liu (xipingl@gmail.com), former Resource Description Librarian, University of Houston Main Campus. **Sharon Reidt** (sharon.reidt@gmail.com), former Metadata Management Librarian, University of Connecticut. **Jodene Pappas** (pappas_j@utpb.edu), Head of Technical Services, University of Texas Permian Basin. **Jill J. Crane** (jcrane1@udayton.edu), Coordinator of Cataloging and Metadata, University of Dayton. **Ada Laura Ramirez** (alramirez@uh.edu), Senior Library Specialist, University of Houston Main Campus.

and data linking on the web.[2] GLAM (galleries, libraries, archives, and museums) institutions have been exploring the application of Linked Data for their carefully curated metadata ever since. Notably, the Library of Congress (LC) initiated the BIBFRAME (Bibliographic Framework Initiative) project in 2012. It is designed to replace MARC (Machine Readable Cataloging) formats for bibliographic description and make library resources more visible on the web.[3] The Library of Congress also initiated a pilot project in 2015 in which LC catalogers began cataloging in BIBFRAME Editor, a Linked Data based metadata editor developed as part of the BIBFRAME Initiative. Since 2019, over one hundred LC catalogers have joined the pilot project. At the January 2022 BIBFRAME update forum, the Library of Congress reported that 90 percent of the LC catalogers will solely work in BIBFRAME Editor by the end of 2022 without double entering the data in MARC format.[4]

The LD4P (Linked Data for Production) project was another attempt by several well-known universities in the United States to develop standards, guidelines, and infrastructure to communally produce metadata as Linked Open Data.[5] LD4P is currently in Phase 3: Linked Data for Production: Closing the Loop (LD4P3), aiming to "close the loop" to create a working model of a complete cycle for library metadata creation, sharing, and reuse.[6] One important tool developed through LD4P is Sinopia Editor, another Linked Data editor for metadata creation. Currently, LD4P's Sinopia Cataloging Affinity Group is actively engaging in the cataloging community to practice cataloging through Sinopia Editor.

The Canadian Linked Data Initiative (CLDI) was formed to leverage the existing collaboration between the technical services departments of Canada's top five research libraries.[7] CLDI investigated a variety of areas such as grant application, education and training, identifiers, digital projects, BIBFRAME Editor, etc. through different working groups.[8] The cross institutional collaboration allowed the staff members to cover more ground in a coordinated manner and share skill sets and documentation, thus strengthening the existing relationships while forging new ones.[9] One of CLDI's future goals is to build a culture of learning and experimentation within the organizations, allowing the group to move forward both individually and collectively.[10]

In an International Linked Data Survey for Implementers conducted by OCLC Research in 2014, 2015, and 2018 respectively, 143 institutions from twenty-three countries responded to the surveys and reported on their Linked Data projects.[11] Survey respondents indicated that their chief motivations for publishing Linked Data are: to expose data to a larger audience on the web, to demonstrate what could be done with datasets as Linked Data, and to simply try it out by exposing some local data as Linked Data.[12] Survey responders also listed "steep learning curve for staff," "lack of resources," and "lack of tools" as the barriers and challenges during their Linked Data implementation.[13] Based on the survey results, most of the linked data projects or services are done entirely in-house, but still through collaborations with external groups. Twenty-two percent of the respondents carried out the project as part of a multi-institutional implementation.[14]

In 2017, the Association for Library Collections and Technical Services (ALCTS, now part of ALA Core) released the *Core Competencies for Cataloging and Metadata Professional Librarians*. Under "Knowledge of trends in the cataloging and metadata profession," it lists Linked Data as one major trend that cataloging and metadata professionals need to be aware of.[15] In a report on the cataloging and metadata professional development survey published in 2017, Tosaka and Park found that respondents listed Linked Data and BIBFRAME within the top four continuing education topics they were interested in exploring.[16] When examining how the cataloging and metadata community perceive the implications of the Semantic Web, the authors found a strong consensus that Linked Data implementation would represent a new opportunity for the profession, citing "improved user services" and "improved data/resource discovery" as the top two potential benefits of the Semantic Web.[17] When asked about professional competencies important for the future of cataloging and metadata librarians, respondents listed "ability to learn and use software" and "ability to collaborate with people within the organization and beyond" as the top two competencies.[18]

In the book *Linked Data for the Perplexed Librarian* published in 2020, the authors made recommendations on building collaborative partnerships in experimenting on Linked Data projects, because it has shown that the most successful of the largest projects in the Linked Open Data community for libraries, archives, and museums are not one-person operations. Project teams can be formed both within one organization, among local GLAM professionals, or even virtually.[19]

The University of California at Los Angeles' Continuing Resources Study Group's activities around BIBFRAME sets another good example for collaborative learning of Linked Data. Balster reported that their study group began actively investigating Linked Data as early as 2014 in order to better understand the BIBFRAME model.[20] One member of the UCLA group designed training sessions on understanding the basic principles of Linked Data and the Semantic Web and two other members took a Library Juice Academy course series to gain further technical skills. The UCLA group identified tools to convert MARC to Linked Data and held weekly discussions analyzing the conversion results.[21] The UCLA study group inspired the initial incarnation of the Linked Data study group.

Initial Efforts

In Summer 2018, ALCTS hosted an E-forum on Linked Data.[22] During this E-forum, a librarian from the University

of California at Los Angeles shared her successful experience of learning about BIBFRAME through a study group. After the E-forum, the resource description librarian from the University of Houston reached out to the cataloging librarians from other university systems and proposed the idea of a similar study group. All four cataloging librarians had previously established working relationships through emails and various projects though they had never met personally. The desire to better understand Linked Data and BIBFRAME led to a quick agreement to form their own study group. Two librarians volunteered to lead the team and create a study plan for the whole study group.

As a first step, the newly formed study group set up practical goals. During the establishment of these goals, considerations were made for schedules and individual workloads. The goals were focused on understanding the basic concepts of Linked Data, experimenting with BIBFRAME Editor, and if possible, identifying a collection from their shared catalog and exploring ways to convert MARC records in those collections to BIBFRAME, as a prototype for applying Linked Data.

The study group scheduled virtual monthly meetings and listed each member's responsibilities which included preparing the reading materials and leading the virtual discussion. Each learning topic would consist of both a required reading list and an optional one, so everyone could choose how far they would like to delve into the topic based on their availability. The librarians also planned to devote the first two meetings to the basics of Linked Data and BIBFRAME.

The study group relied heavily on the learning materials at the Library of Congress's Cataloger's Learning Desktop to create the reading list. The learning materials were originally designed for the catalogers at the Library of Congress for their BIBFRAME Pilot Project so they provided a wide range of resources, from basic to advanced readings. At the time, these resources met the needs of the study group because they offered flexibility with respect to granularity.

The first meeting was focused on the definition of Linked Data. The group discussed introductory concepts including the Semantic Web, Resource Description Framework (RDF), and triples. Semantic triples, or triples, are the three parts of an RDF statement: these are the *subject* or resource being described, the *predicate* or property, and the *object* entered as text or a Uniform Resource Identifier (URI).[23] Group members each shared their understanding of Linked Data and discussed why the Library of Congress had chosen to replace MARC format with the BIBFRAME model. They talked about the possibility of undertaking a Linked Data project locally but also recognized the challenges each library faces in implementing Linked Data. Finally, the group translated a catalog record into triple statements.

In the next meeting, the study group moved on to discuss the BIBFRAME Editor. Everyone was asked to experiment with the BIBFRAME Editor before the meeting so they could share feedback. Discussion topics included what participants liked and disliked about the Editor, what improvements could be made to the Editor, and what challenges catalogers have faced in moving from editing MARC to editing BIBFRAME. Group members talked about the concepts of BIBFRAME classes and properties and gave a few examples before cataloging a book in the BIBFRAME Editor together. They then used the Library of Congress's MARC to BIBFRAME Comparison Viewer to compare the MARC format with RDF, as shown in figure 1 and figure 2.

In a meeting later that Fall, the group was able to identify a small set of records from the shared catalog and applied the Linked Data tool in MarcEdit to add URIs to the corresponding MARC fields, see figure 3.

In April 2019, the group presented a poster at the state library association annual conference, where they made connections with other librarians who were also interested in learning about Linked Data.[24] After the presentation, the group took a hiatus because of a library-wide system migration to Ex-Libris Alma.

Joint Study Group

A cataloging services librarian and another two librarians from smaller academic libraries and a regional public library system had also been researching and studying Linked Data on their own by viewing online webinars, examining websites such as the LD4 and Library of Congress sites, and reading professional literature. Each of them, as solo cataloging librarians, read articles and attended webinars and conferences on Linked Data, but struggled to understand technical and practical details. Attending the conferences and discussions allowed these librarians to meet and inspire each other to learn more about Linked Data implementation through self-directed projects.

Over the previous few years, the cataloging services librarian took an introductory metadata class where she learned key metadata schemes including Metadata Object and Description Schema (MODS) and Dublin Core, and how to convert these records and MARC records into MARC XML syntax.[25] These records could then be converted into a Linked Data-ready metadata record in RDF/XML syntax. She chose a small collection of monographs, videos and archival collection from her library that focused on a specific theme as a pilot to create Linked Data records. During the state library association annual conference (the same conference in which the above poster was presented), she presented on her progress as a solo librarian. Aware that the lack of funding and supporting staff presented a barrier to Linked Data implementation inspired the Cataloging Services Librarian to start a small group for "solo" catalogers

Figure 1. MARC format in the Library of Congress's MARC to BIBFRAME Comparison Viewer

Figure 2. Turtle serialization

to learn together. She issued an invitation to the other two librarians to overcome the lack of funding and staff by working and studying together through virtual monthly meetings. Later at the conference, they saw the poster presented by the librarians from the study group from the University of Houston Libraries and suggested that they join together virtually.

Two of the librarians from different institutions initiated a plan in the Fall of 2019, brainstormed goals, and presented them to the rest of the group for agreement. Their goals were to: (1) stay up to date with the current development of Linked Data and BIBFRAME; (2) seek opportunities to convert and publish respective collections into Linked Data; (3) explore Alma's Linked Data capabilities; and (4) publish an article or present at a conference. As part of the study group's programming, a monthly schedule of topics and topic facilitators were drafted for the academic year.

This fulfilled the group's only rule, that everyone should share in leading the monthly discussions. As a facilitator, each member would identify the learning materials (including webinars, reading material, etc.) and identify discussion topics. The group decided to meet through Zoom and scheduled a recurring meeting at the beginning of each month. Study materials and documents—including the study plan for each month, links to resources, and questions to review before and during monthly meetings—were shared through Google Drive. The newly expanded group of six librarians from the various institutions had their first meeting in November 2019. They called themselves the Linked Data Study Group (LDSG).

```
=001  991022558209705701
=005  20201027190723.0
=008  140124s2014\\\\ilua\\\\\b\\\\001\0\eng\\
=010  \\$a2013050031
=015  \\$aGBB482058$2bnb
=016  7\$a016815813$2Uk
=020  \\$a9781555709679 (alk. paper)
=020  \\$a1555709672 (alk. paper)
=020  \\$z9780838919880 (PDF)
=020  \\$z083891988X (PDF)
=020  \\$z9780838919897 (ePub)
=020  \\$z0838919898 (ePub)
=020  \\$z9780838919903 (Kindle)
=020  \\$z0838919901 (Kindle)
=035  \\$a(TxHU)b68024630-01uho_inst
=035  \\$a(OCoLC)859186517
=040  \\$aDLC$beng$erda$cDLC$dYDXCP$dBTCTA$dBDX$dERASA$dYNK$dKSU$dCUS$dIQU$dILC$dUKMGB$dTHD
=042  \\$apcc
=049  \\$aTHDM
=050  00$aZ675.U5$bM5755 2014
=100  1\$aMoniz, Richard J.$0http://id.loc.gov/authorities/names/no2011171317$1http://viaf.org/viaf/161584526
=245  10$aFundamentals for the academic liaison /$cRichard Moniz, Jo Henry, and Joe Eshleman.
=264  \1$aChicago :$bNeal-Schuman, an imprint of the American Library Association,$c2014.
=300  \\$aviii, 200 pages :$billustrations ;$c23 cm.
=336  \\$atext$2rdacontent.
=337  \\$aunmediated$2rdamedia.
=338  \\$avolume$2rdacarrier.
=490  1\$aALA fundamentals series.
=504  \\$aIncludes bibliographical references and index.
=505  0\$aFaculty/staff orientation meetings -- Subject expertise -- Communication with faculty -- Online tutorials -- Faculty assistance -- Collection development -- Accreditation and new courses -- Evaluation.
=520  \\$aThe role of the library and librarians on campus has changed in the past two decades. Though still relevant to academic enterprise, the transformation o with regard to the future of library services in this technological age.
=650  \0$aAcademic librariesxRelations with faculty and curriculum.$0http://id.loc.gov/authorities/subjects/sh85076594
=650  \0$aAcademic librariansxProfessional relationships.
=650  \0$aAcademic librariansxEffect of technological innovations on.
=700  1\$aHenry, Jo.$0http://id.loc.gov/authorities/names/n2014004484$1http://viaf.org/viaf/307425304
=700  1\$aEshleman, Joe.$0http://id.loc.gov/authorities/names/n2014004487$1http://viaf.org/viaf/307425307
=830  \0$aALA fundamentals series.$0http://id.loc.gov/authorities/names/n2004141779
=907  \\$a.b68024630
```

Figure 3. A sample MARC record with URI added in the 100, 6xx and 7xx fields

Review over the Previous Year and Sinopia Editor

During the first meeting, the LDSG decided to review the lessons of the previous year that had been undertaken by the University of Houston librarians, focusing on the basic concepts of Linked Data and BIBFRAME. This served to develop a shared understanding of the fundamental work of the group before moving forward with additional learning objectives. For example, some members had a small collection of records that they wanted to convert to Linked Data. Others did not have a particular project in mind but were eager to find opportunities to practice Linked Data with real world examples.

During the ALA Annual Conference that year, the LD4P (Linked Data for Production) project team presented Sinopia Editor as a Linked Data cataloging tool. The group decided to evaluate the tool together and try cataloging a few books using the Sinopia Editor. In preparing for this meeting, all members read articles about the development of the Sinopia Editor and its Profile Editor and reviewed the BIBFRAME Editor so they could compare the two tools (see appendix B for a detailed reading list). As part of their collaborative evaluation, the group walked through the steps together to catalog a title and discussed the positives and negatives of using the Sinopia Editor, as seen in figure 4. They agreed that the type head and automatic linking function were an improvement compared to the current MARC format, but the switch between templates for different entities was clumsy.

Wikidata

Their third study session focused on Wikidata, a knowledge base readable by both humans and machines that serves as a data source for Wikipedia and other Wikimedia projects. Wikidata has gained increasing attention from the library community for its Linked Data and identity management capabilities and was the focus of a recent PCC initiative. The meeting facilitators provided an overview of Wikidata and walked study group members through a guided tour of the Wikidata website. They reviewed items and statements and demonstrated the steps needed to add and edit items; several of them had the opportunity to create Wikidata items for their libraries.

After discussing the "how" of Wikidata, they discussed the "why." They reviewed the Association of Research Libraries White Paper on Wikidata, which focuses on the use of Wikidata as a global discovery tool for institutions'

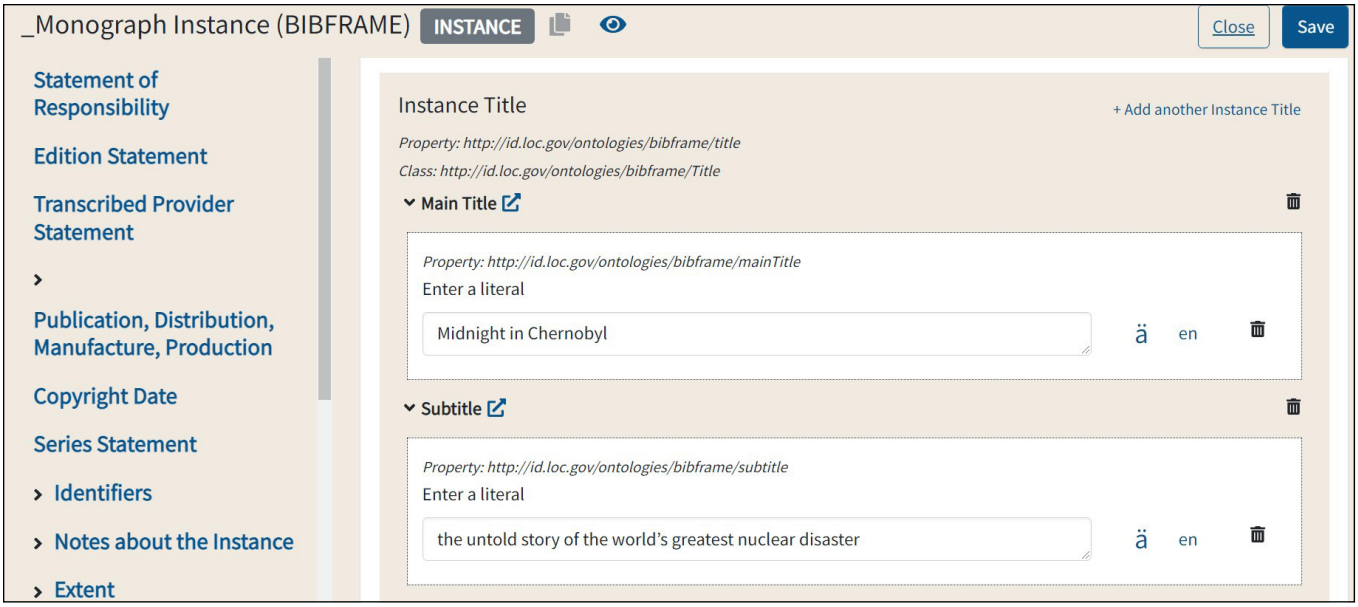

Figure 4. Catalog in Sinopia Editor's Monograph Instance Template

collections.[26] They discussed the University of Virginia's use of Listeria within Wikidata to highlight people, places, and things affiliated with UVA, as shown in figure 5. Wikidata's potential in providing an alternate form of authority control with fewer barriers than the Name Authority Cooperative Program, as a means of promoting research and scholarship at the group members' universities, and as a way to collect statistics on faculty publications and scholarship was another topic of conversation during this session.

Study group members were enthusiastic about exploring Linked Data concepts via Wikidata to test its value for their individual institutions. After their Wikidata meeting, they individually explored Wikidata and WikiProjects pursued at other academic libraries and shared their discoveries via email. One WikiProject that caught their eye was the Stanford Libraries' WikiProject, shown in figure 6. Many aspects of the project seemed translatable to projects group members could start at their own respective universities.

A member of the study group invited the coordinator of the WikiProject at the Stanford Libraries to meet with the group, and she graciously agreed. Since Wikidata is also a topic of interest to members' colleagues outside of the study group, these colleagues were invited to participate in the presentation although they do not regularly attend the LDSG meetings.

During their meeting, the WikiProject coordinator shared an overview of her training and Wikidata expertise development. She shared the history, progression, and organizational structure of Stanford Libraries' WikiProject. She gave practical advice on how to get started using Wikidata and how members of the group could start WikiProjects of their own.

Participation in Outside Webinars and Presentations

When the ALCTS New Members Interest Group (ANMIG) announced the virtual read-along program for the book *Linked Data for the Perplexed Librarian*, the study group members chose to participate in the program.[27] This was a great opportunity for the group to interact with the authors, ask follow-up questions, and ultimately enhance their knowledge about Linked Data. This follow-up discussion enabled them to ask questions of each other about the book itself and about the ANMIG discussion.

Members of the study group also attended a webinar hosted by the Georgia Library Association's Technical Services Interest Group titled "Linked Data for the Real World: Leveraging Metadata for Cataloging," it was presented by Robin Fay. In the presentation, Fay provided an overview of Linked Data and described Linked Data's potential in freeing up catalogers' time and increasing access and discoverability of library materials.[28]

Earlier in the year, a member had suggested that the group consider presenting at the LD4 conference and could share how they pursued a collaborative approach to learning Linked Data. Since the conference was slated to be held in College Station, Texas, the majority of study group members would be able to attend. Members of the group reacted favorably to this idea and decided to submit a proposal. Two librarians volunteered to submit the proposal, and everyone agreed to work on the presentation if the proposal was accepted.

Later, the group learned that their LD4 proposal had been accepted by the conference organizers. When the

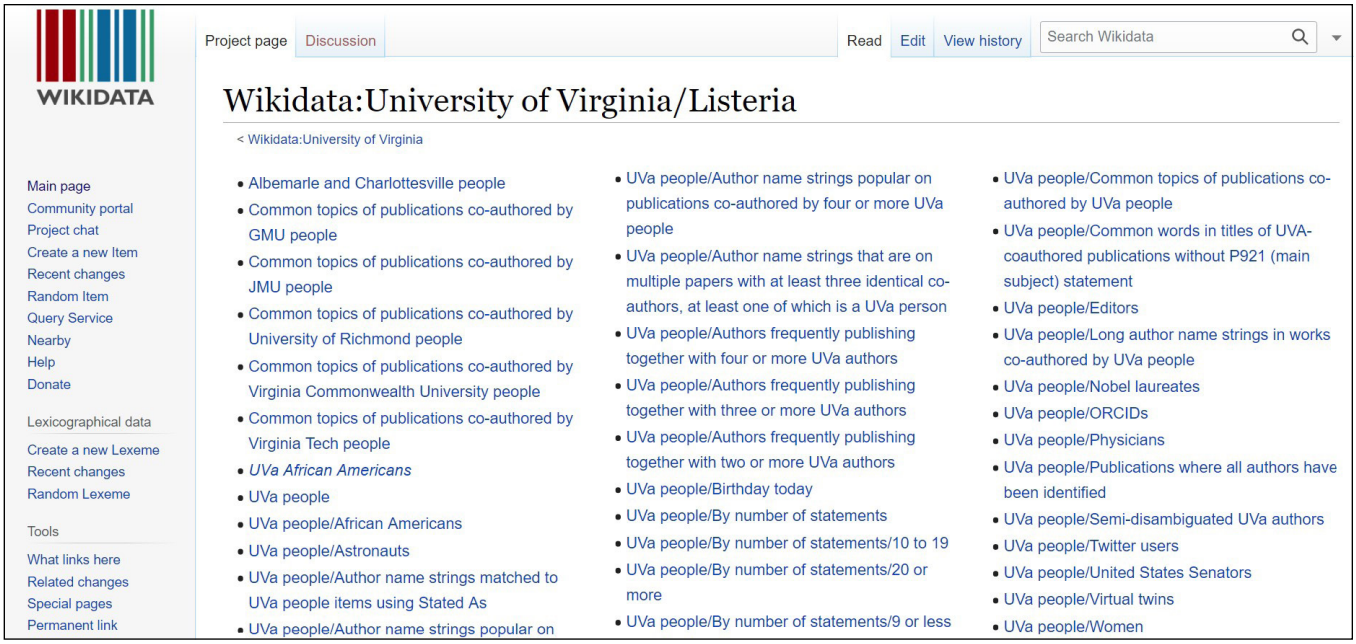

Figure 5. Wikidata: University of Virginia/Listeria

Figure 6. Wikidata: WikiProject Stanford Libraries

decision was made to switch to a virtual conference due to the COVID-19 pandemic, LDSG members agreed to continue to present as planned and welcomed the opportunity for everyone in the group to participate in the presentation. During the meetings leading up to the conference, members discussed and reviewed the materials; in-between meetings they added content to the slides. The group met outside their regular meeting hours to review the presentation slides, approve their final draft, and perform several run-throughs. The study group's presentation, "Being Solo No More: Collaborative Learning through a Virtual Study Group on Linked Data," was delivered on July 21, 2020.[29]

PCC Wikidata Pilot

In September 2020, the LDSG began their second year, continuing monthly meetings and adding a new librarian into the group. Since then, the Linked Data Study Group had spent part of the previous year investigating Wikidata's Linked Data functionalities, having read two articles and a book, as well as learned about a Wikidata project at Stanford University. When the call went out from the PCC for participants in their Wikidata Pilot, it seemed to be a good fit for the LDSG to participate as a group. By participating in the pilot, LDSG could achieve several objectives. It would allow the group to support the learning objectives of the pilot in assessing Wikidata's viability as a means of facilitating identity management.[30] It would give LDSG members a chance to develop hands-on experience using Linked Data. Participation would also provide an opportunity to judge the value of Wikidata and, by extension, Linked Data.

Roughly seventy-five institutions from all over the world participated in the pilot, bringing in a diverse range of projects, with most institutions focusing on creating personal and corporate entities for their faculty, departments, schools, and colleges. Some are highlighting local collections by creating entities for local agents and works appearing in these collections.

After discussing possible collaborative projects, LDSG decided to pursue a project that would benefit a third-party

organization, the Conflict Archive on the INternet (cain.ulster.ac.uk) or CAIN. The CAIN website and project gather resources regarding the Northern Ireland conflict, also known as The Troubles. This archive was deemed as important to amplify in Wikidata due to the historical and political importance of the data, which span decades, from 1968 to the present.[31] At the time, CAIN faced funding challenges; working with its data offered the possibility of assessing Wikidata's added value. One of the librarians reached out to the archive's aggregator and director for CAIN/Ulster University for permission to re-use and publish the information to Wikidata. Each librarian attended the Wikidata Pilot Project kickoff, signed up for the Pilot listserv, and attended training programs and the LD4 Wikidata Affinity Group meetings.

After receiving permission from the CAIN archive's director, the group chose to focus solely on the organizations listed on the website. Narrowing the scope made the task feasible for a one-year project. One of the study group librarians created a shared Google spreadsheet with a list of organization names contained on the CAIN site. The names were reconciled with Wikidata so that each organization listed on the spreadsheet had a link to its Wikidata entry (where one existed). A Wikidata Project Page was created under the auspices of the PCC Wikidata Project. This included a metadata application profile that outlined specific mandatory and optional statements to create basic or more complete items for each of the entries. Unlike the traditional name authority records which follow the policy and standards of the Library of Congress, the LDSG created the metadata application profile based on their judgment on what worked best for CAIN and what would drive traffic to their website.

The study group's next meeting involved walking through the process of adding Wikidata records and setting up an optional weekly working hour in addition to the monthly meetings. The weekly working hour served to specifically work on Wikidata together and discuss questions around the project. From the Google spreadsheet of organizations mentioned on the CAIN archive/website, LDSG members created or augmented Wikidata items.

Monthly study group meetings for the rest of the academic year focused on Wikidata and the pilot program. Questions about syntax, common usage, and errors regarding specific Wikidata statements and qualifiers were discussed and decided upon. Additionally, LDSG members learned about deprecation of statements, disambiguation, and Wikidata scripts and gadgets meant to simplify item creation.

By July of 2021, the LDSG created or updated 283 items for entities in Wikidata. In addition, 102 items were reviewed and found to have already met the basic or complete standard for entry into Wikidata. Only thirty-five of the 420 items on the spreadsheet were not created, mostly due to a lack of information about those organizations on CAIN's website. The LDSG hopes the increased internet visibility that adding these items will allow more people to learn about this important archive/website and motivate the creators of CAIN to continue to add information.

LDSG's communication with CAIN has been sporadic, and it is difficult to determine what, if any, impact Wikidata has played on the archive. They have not been able to ascertain whether increased presence on Wikidata has increased traffic to the CAIN website. It's also unclear if LDSG's Wikidata work played a part in easing CAIN's funding challenges; as of February 2021 the archive's future was no longer in jeopardy.[32] The inability to gauge Wikidata's effectiveness made the project less satisfying.

Working together on a group project through the PCC Wikidata Pilot Project allowed the LDSG to apply Linked Data concepts through Wikidata and learn how Wikidata works in a practical way. Creating Wikidata entities is very similar to creating the Library of Congress Name Authority Records. However, the ease of adding attributes to Wikidata items gives more flexibility to metadata practitioners. The weekly working hours kept the project moving forward and helped to create a learning environment through the questions asked during entry creation. Participation in the pilot gave LDSG members the opportunity to share information with other project members and with the colleagues at their institutions. The project provided a concrete sense of accomplishment, both individually and as a team. It gave several LDSG members the confidence and the skillset to create Wikidata projects for the benefit of their institutions. The resources provided by the PCC and the projects created by peer institutions also served to make Wikidata less daunting and enabled colleagues outside of technical services to get up to speed. The group trusts the work benefited the global community, as well as enabling the team to learn about Northern Ireland's conflict. Working on the PCC Wikidata Pilot also gave the LDSG some direction for the next year of study. With renewed focus, the team aims to learn more about efficiently creating and manipulating Linked Data through SPARQL queries, batch loading, OpenRefine, and other tools.

Year Three of the Joint Linked Data Study Group

As the academic year (and the PCC Wikidata Pilot Project) drew to a close, the LDSG members discussed whether to continue meeting for a third consecutive year. Each member valued the group's work and felt there were Linked Data concepts they would benefit from studying with each other. Rather than pursue a single topic, as they had in year two, they opted to take a more generalist approach, similar to what they had done in their first year.

Group members brainstormed ideas on a shared Google doc, then reviewed and selected the topics they wished to pursue. In year three, they have met on a regular, if not quite

monthly, basis. Topics covered over the course of the year include Wikidata, SPARQL, and OpenRefine.

In the future, the LDSG will continue to meet remotely and use their meetings to continue discussing new developments relevant to Linked Data and BIBFRAME, along with any conferences, webinars, and reading from which they have learned. The LDSG plans to explore practical ways to convert records into Linked Data ready records, both individually and as a group. They also plan to continue working together on future conference presentations.

Discussion

Not only have the study group members exceeded the goals and expectations they set for themselves, but they have also built a successful template for how other librarians can partner to learn more about Linked Data (or other professional topics of interest). The basic structure for forming a successful study group consists of four steps: finding others to study with, deciding on desired goals and outcomes, establishing a study plan and meeting schedule, and maintaining contact between meetings.

The first and most important step to replicate is for one or more individuals to reach out to colleagues to gauge interest in forming a group. The study group can be centered at one's own institution (as was the first iteration of the Linked Data Study Group) or with like-minded colleagues from across institutions.

Part of what made the Linked Data Study Group so successful from the perspective of its participants was the ability to study the theory of Linked Data with colleagues who approached the topic from a beginner's perspective. This enabled them to ask questions freely without fear of judgment. When assembling a study group, give some thought as to whether your group would benefit from having participants from similar or disparate levels of experience.

Once the group has been formed, it's essential to decide upon the aim and goals of the group and establish a study plan and schedule. Now is the time to choose how and how often you will meet, how meetings will be led, how you will communicate and share information as a group, and what your study plan will be. For the Linked Data Study Group, meetings took place virtually once a month during the academic year, and group members took turns leading individual meetings. The key is to find what works for your group.

During this step, you may wish to consider whether or not to pursue a group project. The Linked Data Study Group found that the group projects they pursued helped them in their efforts to turn theory into practice. A group project can help solidify understanding of the concepts being studied. Maintaining contact between meetings helps members gauge progress and iron out scheduling difficulties and conflicts that arise. It also allows members to share timely information about upcoming webinars and conference presentations centered on the study topic.

The study group went above and beyond their goals for the group in several respects. They have met for two academic years, and they typically meet every month of the academic year. In addition to their regularly scheduled meetings, they inform each other of outside learning opportunities, and each of them attended Linked Data programs hosted by ANMIG and the Technical Services Interest Group of the Georgia Library Association.

Their preparation before meetings, along with the meetings themselves, enabled them to build a solid foundation of Linked Data concepts despite the limited time and resources at their disposal. They increased their knowledge of the theoretical concepts behind Linked Data through curated readings provided by the topic facilitators and by reading through *Linked Data for the Perplexed Librarian*. Because they were interested in practice as well as theory, they found ways to develop hands-on practice using Linked Data concepts and gained experience using Sinopia Editor and Wikidata.

Another key accomplishment achieved during their first year was their group presentation at the July 2020 LD4 Conference. They spent time outside of their normal meetings to work on the presentation, and because the conference was virtual, each of them was able to present. Through their presentation, they shared their study group's experiences, and gave participants information on how to create informal Linked Data study groups of their own. In their second year, they expanded their hands-on experience using Linked Data through collaboration on a WikiProject. While they are no longer actively adding items to Wikidata as part of this project, they hope to revisit it in the future to assess the value of Wikidata.

Both Linked Data and group participation presented challenges to the LDSG. Studying Linked Data can be frustrating because moving from theory to practice is difficult. The members of the study group, like many librarians, have limited opportunities to experiment with Linked Data in the cataloging work that they do. One reason Wikidata resonated so strongly with the group is that it is an accessible tool for gaining practical experience with Linked Data. At times it was difficult to strike a balance between work commitments and participation in the group. Scheduling between two time zones across five different institutions also complicated matters. Finding a Linked Data project that they could work on together was challenging due to the limited resources at their disposal, system dependencies and constraints, and the varying amounts of time they are able to devote to activities that fall outside of their normal work obligations.

Participation in the study group has proven immensely beneficial to each of them, regardless of Linked Data expertise level or experience in metadata. The LDSG members have helped each other gain a solid grasp of Linked Data

concepts during the course of their first year studying together as a group. Having a theoretical framework has given them the confidence to explore practical Linked Data applications to assess the potential benefits of Linked Data, and has enabled them to speak confidently about Linked Data concepts with colleagues at their institutions who do not work in metadata. During their second year, they expanded their Linked Data knowledge from the conceptual to the practical as they undertook a WikiProject. By showcasing their knowledge of and experience with Linked Data, they will be in a better position to facilitate and advocate for the adoption of Linked Data projects at their respective institutions and during professional interactions. They are now equipped to act as Linked Data ambassadors at their organizations, and are poised to justify the adoption of BIBFRAME and other initiatives to library leadership who may lack an understanding of Linked Data. The Wikidata projects they've undertaken at their individual institutions have increased cross-departmental collaboration, yet these projects would not have been possible without their participation in the PCC Pilot.

Participation in the PCC Pilot and their work with Wikidata gave them much needed hands-on experience working with Linked Data. Other practical applications have been limited, and for now most of the members' Linked Data experience remains theoretical. This theoretical knowledge mirrors the current status of Linked Data in libraries. BIBFRAME is on the horizon, but its implementation date remains unknown. It is encouraging to know that they are not alone. Since they are approaching the topic from the same level of expertise, they feel safe asking questions of one another and acknowledging when they do not know the answers to questions.

They have developed an informal network among themselves. In addition to their monthly meetings, they routinely share information via email outside of these meetings. They discuss WikiProjects at other institutions and share information on webinars and presentations related to Linked Data. Participation in the group allows them to network with librarians outside their respective institutions and at different kinds of academic libraries (both large public universities and small liberal arts colleges). Working with each other on a monthly basis has strengthened these relationships. Now in their third year, they continue to grow as a group, sharing information and learning from one another as they go. The collaborative learning process has given them confidence and provided motivation and inspiration to continue to stretch their knowledge and move forward with studying Linked Data.

Conclusion

By the close of summer 2021, each member of the LDSG had a solid foundation in Linked Data theory and developed strong peer-to-peer information sharing relationships with colleagues across institutions. Their success is directly attributable to their participation in the study group, which has come at no cost, excluding staff time, and has given them opportunities to use their knowledge to the benefit of their institutions. Although the study group plans to remain relatively static in their membership, they believe that the work they've done is easily reproducible and would encourage librarians interested in learning more about Linked Data or other professional topics to connect with like-minded colleagues to form their own independent study groups. This approach allows members to pursue a topic of interest to them and gives them the freedom to develop a curriculum for their shared benefit. Members have ownership of their learning progress and provide each other with a supportive learning environment. Participating in an independent study group creates networking opportunities, and there is the possibility of cross-institutional collaboration if working with colleagues from other organizations. The Linked Data Study Group set the perfect example for the academic library community.

References

1. Scott Carlson et al., *Linked Data for the Perplexed Librarian* (Chicago: ALA Editions, 2020).
2. "What is Linked Data," World Wide Web Consortium, accessed May 17, 2022, https://www.w3.org/standards/semanticweb/data#summary.
3. "Bibliographic Framework Initiative," Library of Congress, accessed May 17, 2022, https://www.loc.gov/bibframe/.
4. Beacher Wiggins et al., "BIBFRAME January 2022 Update Forum," Library of Congress, https://www.loc.gov/bibframe/news/bibframe-update-mw2022.html.
5. "Linked Data for Production (LD4P)," LD4P Public Website, accessed May 30, 2022, https://wiki.lyrasis.org/pages/viewpage.action?pageId=74515029.
6. "Linked Data for Production: Closing the Loop (LD4P3)," LD4P3 Public Website, accessed May 30, 2022, https://wiki.lyrasis.org/pages/viewpage.action?pageId=187176106.
7. Marlene van Ballegooie, Juliya Borie and Andrew Senior, "The Canadian Linked Data Initiative: Charting a Path to a Linked Data Future," *The Serials Librarian* 72, no. 1–4 (2017): 207, https://doi.org/10.1080/0361526X.2017.1292751.

8. van Ballegooie, Borie, and Senior, "The Canadian Linked Data Initiative," 207.
9. van Ballegooie, Borie, and Senior, "The Canadian Linked Data Initiative," 212.
10. van Ballegooie, Borie, and Senior, "The Canadian Linked Data Initiative," 212.
11. Karen Smith-Yoshimura, "Analysis of 2018 International Linked Data Survey for Implementers," *Code4Lib Journal* 42 (November 2018): sec. overview, https://journal.code4lib.org/articles/13867.
12. Smith-Yoshimura, sec. what and why Linked Data is published.
13. Smith-Yoshimura, sec. barriers.
14. Smith-Yoshimura, sec. overview.
15. Cataloging Competencies Task Force, "Core Competencies for Cataloging and Metadata Professional Librarians," January 2017, https://alair.ala.org/handle/11213/7853.
16. Yuji Tosaka and Jung-ran Park, "Continuing Education in New Standards and Technologies for the Organization of Data and Information: A Report on the Cataloging and Metadata Professional Development Survey," *Library Resources & Technical Services* 62, no. 1 (January 2018): 4, https://doi.org/10.5860/lrts.62n1.4.
17. Tosaka and Park, "Continuing Education," 10–11.
18. Tosaka and Park, "Continuing Education," 11–12.
19. Carlson et al., *Linked Data for the Perplexed Librarian*, 122.
20. Kevin Balster, "BIBFRAME for Serials: Activities of the UCLA Continuing Resources Study Group," *Serials Review* 41, no. 4 (October 2015): 257–59, https://doi.org/10.1080/00987913.2015.1095033.
21. Balster, "BIBFRAME for Serials," 258.
22. Anne Washington, Andrew Weidner and Xiping Liu, "Linked Data for the Rest of Us" (E-forum, Association for Library Collections and Technical Services, July 24, 2018).
23. "Basic RDF," World Wide Web Consortium, accessed December 2, 2022, https://www.w3.org/TR/PR-rdf-syntax/#basic.
24. Xiping Liu et al., "Learn It Together and Learn It Well: Forming a Linked Data Study Group at University of Houston Libraries" (poster presented at the Texas Library Association Annual Conference, Austin, Texas, April 2019).
25. Oksana L. Zavalin, "Introduction to Metadata and Linked Data," ALA, February 2018, online course.
26. Association of Research Libraries, "ARL White Paper on Wikidata Opportunities and Recommendations: Opportunities and Recommendations" (April 18, 2019), http://www.arl.org/wp-content/uploads/2019/04/2019.04.18-ARL-white-paper-on-Wikidata.pdf.
27. Association for Library Collections and Technical Services New Members Interest Group, "Virtual Read-Along Program for the Book *Linked Data for the Perplexed Librarian*" (online webinar held on April 20, 2020).
28. Robin Fay, "Linked Data for the Real World: Leveraging Metadata for Cataloging" (presentation, Georgia Library Association's Technical Services Interest Group, May 19, 2020).
29. Jill Crane et al., "Being Solo No More: Collaborative Learning through a Virtual Study Group on Linked Data" (PowerPoint presentation, LD4 Conference on Linked Data in Libraries, College Station, Texas, July 21, 2020, online).
30. "Wikidata: WikiProject PCC Wikidata Pilot," Wikidata, August 15, 2022, https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/Participants.
31. "CAIN Archive, Conflict and Politics in Northern Ireland," Ulster University, accessed June 7, 2022, https://cain.ulster.ac.uk/.
32. "The Future of Cain," Ulster University, accessed August 20, 2022, https://cain.ulster.ac.uk/cainfuture.htm.

Appendix A: Timeline

University of Houston Study Group

Summer 2018—First contact proposing a study group amongst cataloging librarians at the University of Houston, creating goals, schedule, and reading list.

September 2018—First meeting of University of Houston study group. Introductory concepts: Semantic Web, Resource Description Framework (RDF), and triples.

October 2018—Discussion of BIBFRAME Editor, likes and dislikes, suggested improvements, challenges. Exercise using Library of Congress' MARC to BIBFRAME Comparison Viewer.

November 2018—Exercise in using Linked Data tool in MarcEdit to add URIs to MARC fields.

April 2019—Poster presentation at state library association annual conference.

Linked Data Study Group (LDSG)

Spring 2019—Connections made between posters, presentations and previous interactions of those interested in studying Linked Data together through a virtual study group.

September 2019—The new LDSG agreed to goals, study plan and programming over the next academic year.

November 2019—Reviewed concepts and lessons from the previous year (Linked Data and Semantic Web, RDF, BIBFRAME).

December 2019—Evaluate the Sinopia Editor and its Profile Editor, discuss positives and negatives.

February 2020—Overview of Wikidata, guided tour of Wikidata website, demonstration of how to add and edit items. Examination of how Wikidata is being used at the University of Virginia. Review of ARL White Paper on Wikidata. Call for proposals for the LD4 conference, should we present?

April 2, 2020—Meeting with coordinator of the WikiProject at the Stanford Libraries.

April 16, 2020—Discussion about our proposal for Linked Data for Production (LD4) conference presentation about LDSG.

April 30, 2020—Discussion on the book: *Linked Data for the Perplexed Librarian*.

May 28, 2020—Discussion on LD4 presentation and potential individual project for the summer. Discussion on taking part in PCC Wikidata Pilot Project as a group.

July 2020—Writing and developing LD4 presentation.

July 21, 2020—LD4 Conference Presentation (online).

September 2020—Share thoughts after the Pilot kickoff meeting. Discuss our project ideas. Encourage participation in the LD4-Wikidata Affinity Group meetings and work hours.

October 2020—Review the project page together, add everyone's Wikipedia ID, review the basic and extended properties, walk through the process of creating a Wikidata entry for an organization together. Set up goals and working hours for the next few months.

November 2020—Discuss Wikidata Dashboard and any questions around creating/editing Wikidata items. Update from CAIN webmaster. Learn about scripts and gadgets on Wikidata.

March 2, 2021—Updates from members, director of the CAIN site, and the progress on our Wikidata project. Discussed how to handle incorrect statements in Wikidata, and how to deprecate a statement for an incorrect statement and assign to a deprecated rank. Suggestions from another PCC project participant were shared.

April 13, 2021—After individual updates and comments, a link to news on CNN about recent riots in Northern Ireland was shared (since it was relevant to our project). One member also suggested that members review British English vs. American English. Disambiguation page. Ex. https://www.wikidata.org/wiki/Q207829. This page may help with some of the questions that have come up while creating or updating CAIN items.

May 11, 2021—Updates, questions and discussion of the use of Wikidata by the broader library community, a topic that was included in a recent PCC membership meeting.

June 8, 2021—Update on the Wikidata project. One question in past meetings included how to handle items with very brief descriptions on the CAIN site. At the suggestion of one member, it was decided to skip creating Wikidata items for organizations with very brief descriptions. Add "CAIN's entry is very brief" in the "Level of Completeness" field and "n/a (see Notes)" in the "Wikidata Item" field. The LDSG has created/updated 283 items out of 420 organizations as of 6/3/21. Encouragement to continue and finish strong was expressed. Ideas for next year's LDSG meetings were asked for and could be added to a newly created Google document.

July 27, 2021—Final report on our PCC Wikidata project, including the numbers on how many items were created/updated from the list of organizations on the CAIN website. A wrap-up email will be sent to the director of the CAIN website. The members shared thoughts about this project, the good, the challenges, what we can do better next time (if we want to work on a similar project again), etc.

* In addition to these meetings, from October 2020 through July 2021, optional weekly working hours were scheduled for members of the LDSG to virtually meet and work on the PCC Wikidata Project together.

September 28, 2021—Discussion of the announcement that the PCC Wikidata Pilot project has been extended until December 2021. Since we have completed our group project, do we want to start on individual projects? We will continue bi-weekly Wikidata working hours for the LDSG group through the end of the year and bring our own projects. One member of the group has batch added Wikidata items for the remaining CAIN orgs via OpenRefine. This can be covered during one of the regular monthly meetings.

December 6, 2021—Discussed future meeting topics suggested on the shared Google document. These topics included SPARQL queries, Open Refine, batch creating Wikidata items with Open Refine, Python, BIBFRAME and Sinopia Editor, the RDA toolkit, and JSON and JSON-LD. Are the following topics still the ones we would like to learn? Are there any new topics? Order of learning? What approaches should we take? Should we assign one person to lead the discussion for each topic? This person will also be responsible for researching and providing learning resources, discussion questions? We can take our time and extend our learning time if needed. The members also discussed writing a paper about our experience with the Wikidata Project and decided to continue the bi-weekly LDSG Wikidata working hours to work on our own Wikidata projects.

Appendix B: Sample Workform for Starting a Linked Data Study Group

Timeline	Action	Notes/Suggestions
Pre-meeting organization	Contact possible members to propose study group, suggest topics, and create a meeting schedule.	Suggest meeting dates over a specific time period (i.e., one academic or calendar year). Discuss working and communication styles. Ex: Allow each member to lead a topic or meeting.
First meeting (in-person or virtual)	Discuss goals, study plan and program for individual meetings. Create shared spaces.	Create shared drives; folder(s) for members to include suggested readings, activities, or learning opportunities. Appendix C in this document includes a suggested reading list.
Second meeting	Discuss introductory concepts: Semantic Web, Resource Description Framework (RDF), and triples.	
Subsequent meetings	Discuss topical readings, tour Linked Data websites, learn programs, complete exercises, etc. according to agreed upon study plan and program.	
	Refine curriculum.	Consider creating individual or shared Linked Data projects to practice creating metadata. Create a "working hour" every week or month to work on projects.
Second year/time period	Review what has been learned and accomplished, create a new study plan and program for the next time period.	Set new learning goals, consider new tools and strategies that will move the group forward.
Remain aware and open to newly published literature, to learning opportunities, and open projects your study group may be able to participate in as a group or individually.		

Appendix C: Suggested Reading List

Understanding the Basics

"BIBFRAME Training at the Library of Congress." *Library of Congress*. https://www.loc.gov/catworkshop/bibframe/.

Hawkins, Les. "Setting the Stage Linked Data Overview." *Library of Congress*, http://www.loc.gov/catworkshop/bibframe/Setting-the-stage1.mp4.

Baxmeyer, Jennifer, et al. "Linked Data Infrastructure Models: Areas of Focus for PCC Strategies." *Library of Congress*. http://www.loc.gov/aba/pcc/documents/LinkedDataInfrastructureModels.pdf.

Naun, Chew Chiat, et al. "LTS and Linked Data: A Position Paper." Cornell University Library, 16 December 2015. https://ecommons.cornell.edu/handle/1813/41481.

Van Hooland, Seth and Ruben Verborgh. "Linked Data for Librarians Part 1 Module 1: Introduction." *Linked Data for Librarians,* http://course.freeyourmetadata.org/introduction/.

Understanding BIBFRAME

"BIBFRAME Update Forum at the ALA Annual Conference 2022." *Library of Congress*, https://www.loc.gov/bibframe/news/bibframe-update-an2022.html.

Shieh, Jackie. "Embedded URI in MARC: An Essential for Linked Data." Microsoft PowerPoint presentation. http://downloads.alcts.ala.org/ce/10262016_MARCtoBIBFRAMEseriesPart3_Embedded_URI_MARC_Slides.pdf.

MarcEdit Libguide. University Library, University of Illinois at Urbana-Champaign. http://guides.library.illinois.edu/c.php?g=463460&p=3168070.

Smith, MacKenzie, et al. "BIBFLOW: A Roadmap for Library Linked Data Transition." University Library, University of California, Davis. https://www.library.ucdavis.edu/bibflow/roadmap/.

Reese, Terry. "MarcEdit MARCNext: Linked Records Tool." YouTube, September 2, 2014. https://youtu.be/ifhxNT1TxVU.

Hanson, Eric M. "A Beginner's Guide to Creating Library Linked Data: Lessons from NCSU's Organization Name Linked Data Project." *Serials Review* 40, no. 4 (2014), 251–58. https://doi.org/10.1080/00987913.2014.975887.

Reese, Terry. "MarcEdit MARCNext: JSON Object Viewer." YouTube, September 2, 2014. https://youtu.be/wyijGEn8sr0.

Reese, Terry. "MarcEdit MARCNext: BIBFRAME testbed." YouTube, September 2, 2014. https://youtu.be/2BTkjjowF1s.

Sinopia Editor

"Training Videos Based on Library of Congress Tools." LD4P. https://github.com/LD4P/sinopia/wiki/Training-Videos-based-on-Library-of-Congress-tools.

"Profile Editor Compared in LC BF and Sinopia." https://docs.google.com/document/d/1oNvUWFuMHzgGsb8ad175FeqlhKtADdb0Jzx6f933GCA/edit.

Nelson, Jeremy. "Developing Sinopia's Linked-Data Editor with React and Redux." *Code4Lib Journal* 45 (August 2019). https://journal.code4lib.org/articles/14598. The introduction and conclusion provide useful information about the background of the Sinopia editor. The rest is too technical for our purposes and can be skipped.

Linked Data for Production: Pathway to Implementation. Yale University. https://web.library.yale.edu/ld4p/charter.

Wikidata

Wikidata: University of Virginia/Listeria. https://www.wikidata.org/wiki/Wikidata:University_of_Virginia/Listeria.

"Wikidata Getting Started: Helpful Resources, Labels, Aliases, Descriptions." Microsoft PowerPoint presentation. https://drive.google.com/file/d/1UHOtsOquIGOuVHIPKTsA0dq4jrsh5P0o/view.

Wikidata: Introduction. https://www.wikidata.org/wiki/Wikidata:Introduction.

"Introduction to Wikidata for Librarians Structuring Wikipedia and Beyond." OCLC. Microsoft PowerPoint presentation. https://www.oclc.org/content/dam/research/events/2018/oclcresearch-lih-fernandez-introduction-to-wikidata-wip-2018.pdf.

"ARL White Paper on Wikidata Opportunities and Recommendations." ARL Task Force on Wikimedia and Linked Open Data. https://www.arl.org/wp-content/uploads/2019/04/2019.04.18-ARL-white-paper-on-Wikidata.pdf.

Allison-Cassin, Stacy and Scott, Dan. "Wikidata: A Platform for Your Library's Linked Open Data." *Code4Lib Journal* 40 (May 2018). https://journal.code4lib.org/articles/13424.

Scott, Dan. "Creating and editing libraries in Wikidata." https://coffeecode.net/creating-and-editing-libraries-in-wikidata.html.

"Wikidata: Wikiproject Stanford Libraries." https://www.wikidata.org/wiki/Wikidata:WikiProject_Stanford_Libraries.

Carlson, Scott, Cory Lampert, Darnell Melvin, and Anne Washington. *Linked Data For The Perplexed Librarian*. Chicago: ALA Editions. 2020.

Appendix D: Members of the Linked Data Study Group:

Jill Crane, Coordinator of Cataloging & Metadata, University of Dayton

Martha Hood, Associate Director, Assessment & Planning, University of Houston Clear Lake

Xiping Liu, former Resource Description Librarian, University of Houston Main Campus

Marla McDaniel, Catalog Librarian, University of Houston Downtown (2018–2019)

Jodene Pappas, Head of Technical Services, University of Texas Permian Basin

Ada Laura Ramirez, Senior Library Specialist, University of Houston Main Campus

Sharon Reidt, former Metadata Management Librarian, University of Connecticut (2019–2022)

Susan Vandale, Technical Services Librarian, Dickinson College