Linked Data Report and Recommendations for Oregon State University Libraries and Press

November 17, 2016
Michael Boock

Executive Summary

Oregon State University Libraries and Press (OSULP) demonstrates the value of openness in a variety of ways. We worked with faculty to pass an open access policy to provide the broadest possible access to university scholarship. We teach faculty and students to use data tools that promote the sharing and distribution of research data. We help students make their theses, dissertations, final research projects, and data openly available. We work with scholars to publish digital humanities projects, conference proceedings, textbooks, journals and other research outputs on the open web. We have a commitment to the use of open source software and make our code openly available; and, with the development of Oregon Digital as a linked data platform, we openly share our metadata for the widest possible machine distribution and reuse on the open web.

As a production-ready metadata solution with the potential for improving user access to related resources, the use of linked data within Oregon Digital and the ScholarsArchive@OSU institutional repository remains a promising and worthwhile endeavor. Its use ensures normalization and consistency of metadata and enables quality faceted browsing of digital objects within discovery interfaces. Longer term, our use of linked data has the potential to link our digital content with other data on the open web, providing users with added context for items within our collections and connecting them with related resources that reside outside our own collections. Also, the fact that linked data is more easily understood by computers gives that data an opportunity to interact with other systems and data sources more intelligently. As one librarian I spoke with said, "Once you have semantic concepts embedded into your data, things like authority control, data sharing, alternative language support – these things become easier."

Given the significant investment OSULP has already made in linked data by building Oregon Digital on a linked data platform, its future promise of reducing silos of library information, and the research value of this work, I propose that OSULP should continue to create, maintain and publish linked data for, at least, our digital collections and, in the future, the institutional repository. The ability for OSULP to fully benefit from the work we've already done depends on our ability to make it interoperable with other linked datasets on the web. OSULP should also begin tracking the emerging use of linked data in library catalogs and discovery services.

To do so, OSULP would benefit by hiring a "next generation" cataloger/metadata librarian and increasing the pool of staff with an understanding of linked data and its uses. The next cataloger/metadata person hired by OSULP should be someone with linked data expertise, familiar with BIBFRAME and the RDF data model, who also has knowledge of MARC, dcterms, and other current cataloging standards and metadata schemes. We also need more staff with at least a basic level of linked data expertise who can take incoming metadata and prepare it for bulk ingest into our digital asset management systems. Training on basic linked data concepts and the tools we've put in place should be provided to library faculty and staff who are in positions to take on more of this work. SCARC could take a leadership position for the library in terms of metadata in general, and linked data specifically.

Introduction

The Linkeddata.org website describes the purpose of linked data as "using the Web to connect related data," and Wikipedia describes linked data practices as those "exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF." With the development of Oregon Digital as a linked data metadata infrastructure, OSU and UO have built expertise and established practices upon which future digital collections, repository functionality, and interoperability with other related resources on the web can be built. Steiner (2010) explains that "the Internet architecture needs a new layer that takes care of data interoperability for interconnecting pieces of machine-processable data." OSU has taken a lead in libraries by expressing its digital collections metadata as linked data. Making this data interoperable with other data on the web depends on others expressing their data in this way and building the links between them. Very little of this work has been done.

In putting together this report, I spoke with several staff from the Special Collections and Archives Research Center, Emerging Technologies and Services, and Teaching and Engagement departments who work with linked data, either developing Oregon Digital and ScholarsArchive@OSU (ETS), creating metadata schema (SCARC and TED), or assigning metadata (SCARC). I also spoke with leading experts in the field of linked data for libraries. The purpose of these conversations was to gain an understanding of current and potential future uses of linked data within our library and within the broader library community. This report describes the state of OSU Libraries and Press (OSULP) linked data projects, discusses current and future benefits of using linked data at the library, recommends how the library should use linked data now and in the future, and describes linked data staffing and training needs at OSULP.

Linked Data Basics

Linked data uses RDF graphs to makes statements about things using triples. Triples break statements into three elements: the subject, predicate, and object of the statement. This is an example of a triple that represents the statement that this report was written by me:

```
    ( Report )

       │ Authored by
       ▼

  ( Michael Boock )
```
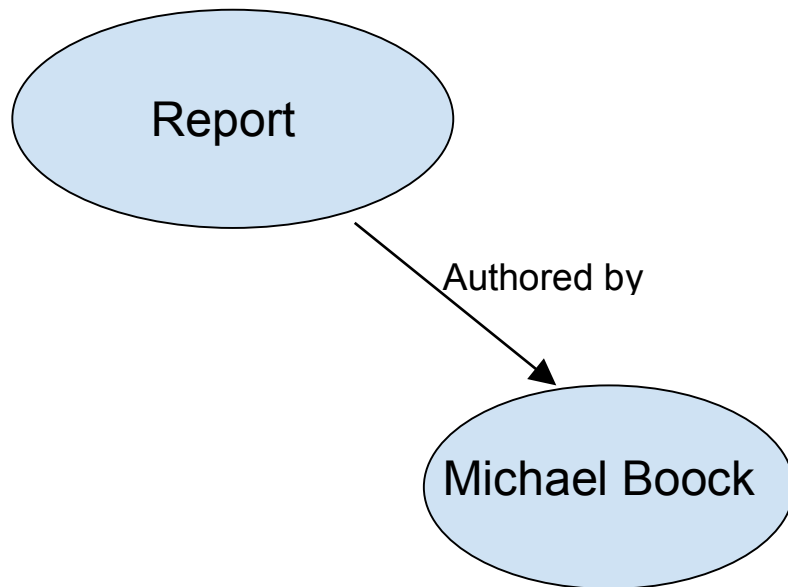
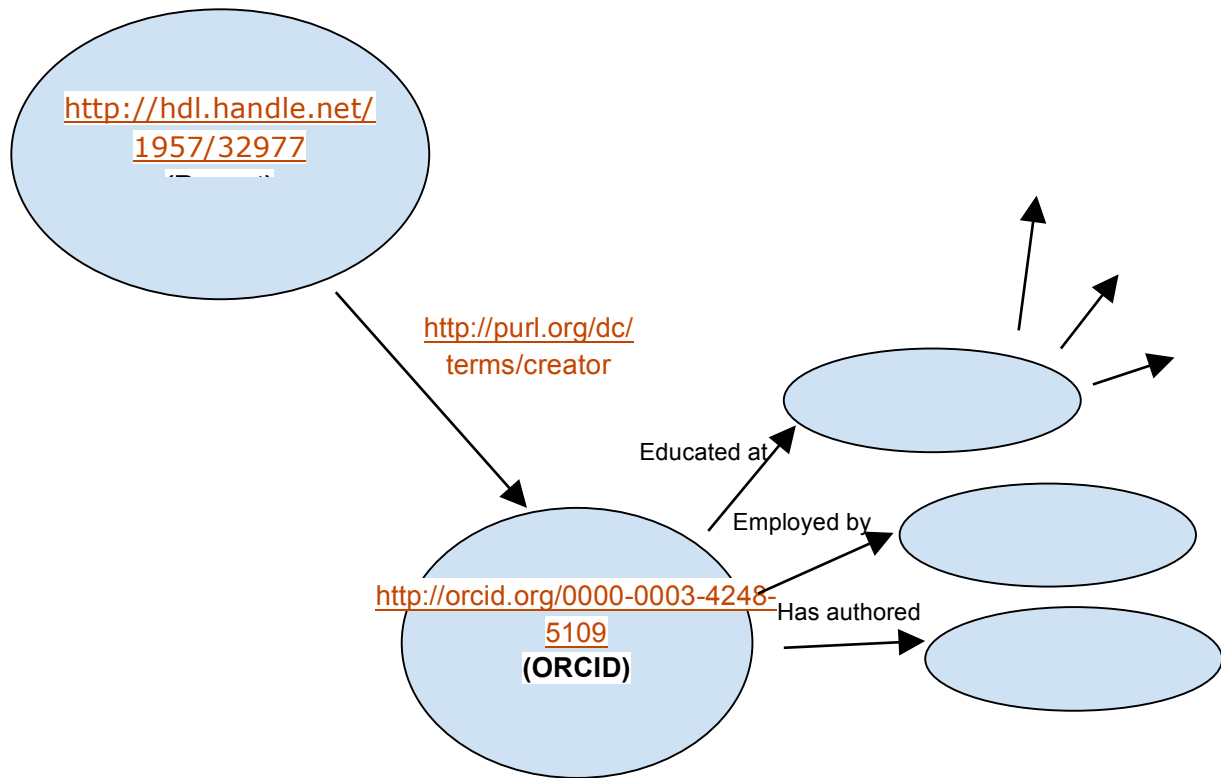Linked data uses identifiers in the form of HTTP URIs to reference concepts and relationships that exist in RDF graphs and library metadata. URIs also provide "a standard mechanism for finding information about concepts and relationships (Johnson, 2011)." Resolvable URIs make linked data machine-readable and usable, so the information contained in the graph above must have URIs in place of each string of text:
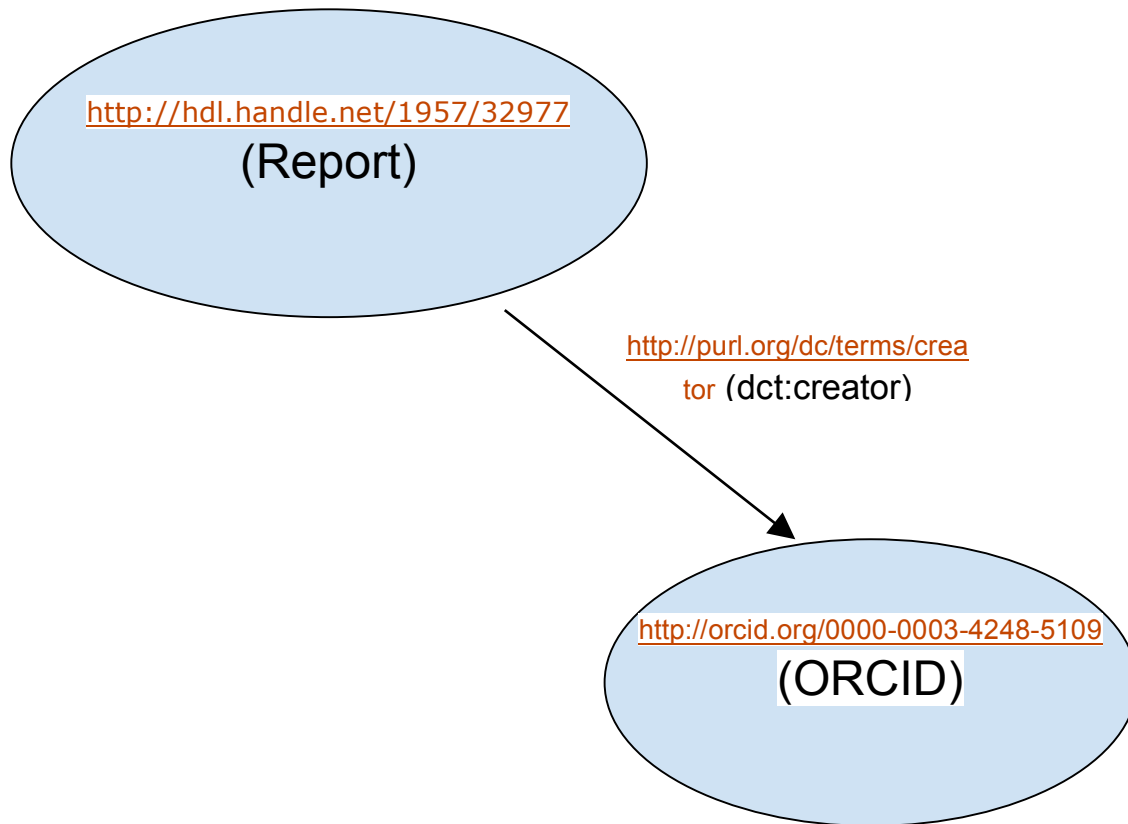
In this graph, Report is the subject, dct:creator is the predicate, and my ORCID is the object value.

The value of containing information/metadata in RDF graphs is that each element within the graph can be connected (automatically, by machines, once they are set up to do so) to other information on the web that is similarly structured. RDF graphs can be endlessly expanded. So, for example, using RDF graphs, computers can see all of the other works I've authored that ORCID has listed, where I went to school, where I was employed, etc.

http://hdl.handle.net/1957/32977

http://purl.org/dc/terms/creator

Educated at

Employed by

http://orcid.org/0000-0003-4248-5109
(ORCID)

Has authored

Using RDF graphs, relationships and added context can be provided to machines and users. Otherwise hidden links can be explored. A key future challenge is choosing how to visualize links to other data within our user interfaces, displaying it in a way that makes sense to the user, and providing links to related information that resides elsewhere on the web.

How Do We Use Linked Data at OSU and What Are the Current Benefits?

At this time, OSULPs linked data efforts only have internal functions. We do not yet take advantage of the potential linked data provides for leading users to related data on the open web or for leading others to our content. Within our digital collections we use linked data published elsewhere and we create and publish our own linked data on the open web, but we do not yet extend the graph to lead our users to other URIs, resources, and information that are part of other graphs. Nor, as far as we know, do other repositories or information systems extend their graphs to our content.

This is because, at least in the cultural heritage space, OSU/UO are among the first libraries to build a digital collections infrastructure on a linked data platform. Karen Smith-Yoshimura with OCLC Research conducted a survey of libraries in 2014 and 2015 in order to learn about linked data projects and services in place at libraries. The surveys found that "of the 112 linked data projects or services described by the 2015 respondents most are primarily experimental in nature (Yoshimura, 2015)." There is no question that we are at the bleeding edge in this area,

particularly given that we've not only experimented with and tested the use of linked data but have implemented it in a production environment.

Oregon Digital

In 2013, OSU and the UO committed to the use of linked data as the metadata infrastructure for Oregon Digital. Staff from both institutions identified dozens of fields in use in CONTENTdm that could be normalized into predicates (Simic, 2016). The pre-migration metadata review also revealed many "local fields where no predicate existed but that were determined to contain sufficiently important information for preservation (ibid)." If predicates were not available for fields, the local fields and their controlled vocabularies were self-published in OpaqueNamespace, a locally developed linked data service. The Controlled Vocabulary Manager, also locally developed as a Ruby on Rails application, is used to manage these controlled vocabularies. The benefits of publishing field names (predicates) and their controlled vocabularies in opaquenamespace.org as URIs is that those local fields (e.g. OSU Department and College) and their locally managed controlled vocabularies are then available to be used by other systems such as ScholarsArchive@OSU, archival finding aids, or even the library's discovery system.

In August 2015 the migration of content and metadata to the Oregon Digital hydra platform was completed. Where predicates and controlled vocabularies were available from respected and well established linked data sources such as Library of Congress and Geonames, identifiers (URIs) were assigned in place of the text strings that resided in CONTENTdm. For example, all instances of the text string of Pauling, Linus, 1901-1994 and its variants were replaced with the official Library of Congress Name Authority URI: http://id.loc.gov/authorities/names/n79144786. In the process of migration, all instances of Pauling, Linus; Pauling, Linus, 1901-; Pauling, Linus, 1901-1994; as well as misspellings and typing errors were replaced with this URI, resulting in normalization and consistency of the metadata and improved searching and browsing for items in Oregon Digital that include or are about Linus Pauling. It is important to note that this kind of metadata cleanup work could be done without turning to linked data, but the use of linked data *enforces* metadata validation and consistency. The transition to identifiers has resulted in an enforced structure for library staff who apply metadata.

Linked data also reduces or eliminates ambiguity about the metadata that we use. Linked data often contains contextual information in machine readable form as well as information about who owns and maintains the data. Changes to controlled vocabularies are effectuated across systems and handled more efficiently. When the former Library of Congress Subject Heading "Afro-American" changed to "African American", libraries around the world had to take steps to replace instances of these subject headings in their individual library catalogs. In a more fully implemented linked data world, when a subject heading like this changes, the identifier changes too. Then any library that uses the identifier rather than the "Afro-American" text string in their catalog or repository will either redirect to the new, correct term "African American", or, the identifier for the term is sunsetted by the owner with information about why or what identifier for the term should be used in its stead. Another example of a potential benefit that hasn't yet come

to fruition relates to our use of Geoname URIs in place of text strings for locations. The use of geonames can provide users with context about geographic locations such as what county a point is in, what state, what mountain it is near, in what water basin, etc.

ScholarsArchive@OSU

The use of linked data is more of a challenge in ScholarsArchive@OSU than it is in Oregon Digital because much of its metadata is created externally; that is, authors who are not expected to take the extra time required to select URIs rather than type text create it. While linked data in the new ScholarsArchive@OSU is not currently a development focus, it is expected to be used where possible for things like geographic names, material types, formats, and possibly authors.

Triples and controlled vocabularies with URIs are already being created for Oregon State University academic units for use in the new ScholarsArchive@OSU. The primary purpose of this controlled vocabulary will be to enable improved reporting on the publications (e.g. theses, dissertations, articles) produced by departments, schools, and/or colleges over specific periods of time. For example, we will be able to identify what percentage of faculty from specific colleges or academic units have deposited articles to the institutional repository to meet the terms of the Open Access Policy. Making this available as linked data will enable other academic units and machines to automatically retrieve, display, and reuse this data. This namespace can also be used for assigning academic units to archival collections.

How Are Other Research Libraries Using or Planning to Use Linked Data?

The University of California San Diego and Northwestern University are the only other libraries I could find that are building digital content management systems (using hydra/fedora, as we are) on linked data concepts and technologies. Metadata in UCSD digital collections "follows the RDF data model, and is serialized as RDF/XML (University of California San Diego Library, n.d.)." A peer library is "just getting started using Linked Data in repository applications, via the Hydra stack and Fedora 4. [They] are using it to increase discovery and interoperability (through the use of as many standard schemas, vocabularies, and ontologies as possible), as well as to try to improve things like name and subject authority by standardizing on known URI-based authorities." They are also exploring the use of linked data for their digital collections and data repository.

UNLV exports metadata out of their CONTENTdm digital asset management system and creates and publishes triples for that metadata for potential use by other systems. As of 2014, they had experimented with applications that use the data but as of this date they don't yet have any services in production. Another peer library I spoke with has done little more than what they describe as "proof of concept work generating URIs for subject data from the catalog for use in other applications." They would like to structure their controlled vocabularies as linked data and publish their own data so that it can be pointed to by other systems and computers on the web. Many librarians I spoke with at other libraries say that they would like to use linked data for their digital collections and institutional repository systems but the vendor (Digital Commons) or open source (DSpace) systems they use don't work with linked data.

Erin Faulder, currently at Cornell and formerly at Tufts University, met with a group of us in the spring to talk about Tufts' use of linked data standards in assigning academic department and college metadata for those responsible for creating their archival holdings. Tufts identifies one major benefit of this implementation as the ability to make this data interoperable with other repositories and sources "to collectively enrich the description of related collections and records creators (Faulder, 2014)." We are building on this work in the development of the OSU academic units controlled vocabulary described above.

At least in the U.S., the Library of Congress (LC) and OCLC are at the forefront in terms of publishing library bibliographic metadata as linked data. The LC Linked Data Service publishes name and subject authority data to the web as linked data which we use in Oregon Digital. OCLC is working to make the wealth of member library bibliographic metadata available from WorldCat available on the web so that "the data can be read and embedded in more websites and online tools than traditionally formatted bibliographic data (OCLC, 2016)."

LC and the Program for Cooperative Cataloging are working to transition the use of MARC to BIBFRAME. BIBFRAME is conceived as "a general model for expressing and connecting bibliographic data (Library of Congress, 2016)." One person I spoke with speculated about what it might look like to completely replace MARC in our discovery systems and "have a catalog that is entirely linked data." Would users be able to engage in a more serendipitous discovery environment that "shows people related content on the web in new and dynamic ways?" What would we gain and what would we lose in this environment? Tom Johnson, OSU's former digital applications librarian began experimenting with this idea several years ago by creating a linked dataset for all of the bibliographic data within our local catalog. He also created a subset of linked data for the university's theses and dissertations using metadata from the local catalog and the institutional repository. He got so far as testing a name authority service built on top of that dataset.

Barriers

Many of the people I spoke with about the future of linked data in libraries agree that "the fundamental infrastructure to support the linked data ecosystem in libraries doesn't yet exist." Several agree that it all still remains very "idealistic and hopeful" in spite of the concept of the semantic web being around since the early days of the web. The greatest barrier to libraries proactively using and publishing linked data may be the lack of linked data expertise within the profession. No doubt relatedly, there is also an unwillingness among libraries, at least to this point, to undertake creation and publication of metadata as linked data even though doing so makes it potentially more discoverable, accessible, and easier to consume by other libraries, as well as by commercial services like google that already use linked data. A peer remarked that "establishing usage and harvesting norms (e.g., making sure things are taken in the correct context)" is their library's biggest challenge in creating and publishing linked data.

Because linked data relies on URIs as identifiers, if other systems upon which we rely for metadata resolution are down, our system can't display it or use it unless the data is cached locally. This means that staff and users who are entering metadata can't select identifiers that live elsewhere when those systems go down. One person I spoke with said that the Library of

Congress, which offers a wide variety of linked data services, is "drowning under requests" and their services aren't even used very much, especially in comparison with how they might one day be used if the semantic web comes more fully to fruition. LC is probably not going anywhere, but if something like Geonames goes away, then the services we build upon them go away too. Do we cache entire datasets so we have them if the other sites go down or if the other sites are slow? If so, we commit to storing and maintaining duplicate data ourselves, which goes against the whole idea of not duplicating resources and effort. Nevertheless, it may be worthwhile to do so if only so our users who assign metadata are able to get their work done more efficiently.

Linked data requires trusting external organizations and even individuals to continue to maintain their services. We have no written agreements with linked data providers. If a linked data service goes down, there is no recourse. While we haven't encountered major problems so far, we have experienced the occasional hiccup, as when geonames goes down occasionally. We pull in URIs from other systems at the point of need (the point at which we are assigning metadata). Doing so results in a slower process. LC and Geonames in particular are slow parsing the information we request at the point of assigning metadata and giving it back to us as URIs and labels. Also, terms disappear in the public interface when another system that manages the URIs goes down.

There aren't many linked data tools or software, especially for libraries, that are available for us to use. As noted above, we built the Controlled Vocabulary Manager, which is only now in its initial release phase and took more than a year to build, because there were no ontology management tools available.

Next Steps for OSULP

Assuming that "the value and usefulness of [linked] data increases the more it is interlinked with other data (Bizer, 2009)," the ability for OSU to fully benefit from the linked data work we've already done depends on our ability to build on this work by making the data interoperable with other linked datasets on the web, enabling our users to explore related information on the web more easily. Craemer (2014) says, "[Linked data] can provide direct access to data in ways that are not currently possible, and provides unanticipated benefits that will emerge later as the stores of linked data expand exponentially." There is an increasing amount of data on the open web that is linked. Already, in addition to OCLC WorldCat and LC metadata, Wikipedia has made all of its information available as linked data through dbpedia. BBC Things provides "data on the places, people and organisations that appear in BBC programmes and online content" and "powers large parts of the BBC website, including BBC News and Sport (Murphy, 2016)."

As Smith-Yoshimura (2015) says, "Linked data efforts demonstrate that the institution is taking the initiative in laying the groundwork for a future, different environment." Because of the linked data work with which we're already engaged, OSULP is recognized as a leader in linked data among research libraries. With the development of Oregon Digital as a linked data platform, OSU and UO have established a strong base upon which future functionalities and

interoperability with other related resources on the web can be built. This may take time. Utilizing linked data for its intended purpose--connecting interrelated data on the web--might, indeed, fail in the end if there is no work in the broader library community to make existing and future linked datasets interoperable. The potential benefits to our users, the benefit of preparing our metadata for the semantic web, and the research benefits we realize by being seen as "thought leaders" in this space make it worth continued effort.

Training and Positions Needed

Through my discussions with OSULP staff, it became immediately clear that they would like to have someone in the library with a high level of linked data knowledge coupled with the technical know how to put linked data services into effect. This requires a thorough understanding of RDF and the use of identifiers, someone who can see things at a high level about the choices we make in this space, someone who can explore more ways to use and interact with linked data. Others mentioned that we need someone with the ability to work with programmers to justify design schema selections and ontology choices, someone to explain why these are selected and why they are worthwhile. This requires an ability to work with and communicate with programmers to make systems work. Too, if we are to continue to take a leadership role in the use of linked data in libraries, we need someone who can spend time thinking about the big ideas of linked data and what it is capable of. Could this be the focus of the next Grey Family Chair?

In addition to this higher level expertise, others mentioned that we also need more staff with at least a basic level of linked data expertise who can take incoming metadata and prepare it for bulk ingest into digital asset management systems. Even the people who already work regularly with linked data don't necessarily understand it. Training should be offered to SCARC and possibly Cataloging staff on everything from what linked data is to how it is effectuated in Oregon Digital. ETS staff too should be given opportunities to learn more about linked data. Staff involved in its creation and development should take opportunities to communicate the value of linked data for the projects we're doing to the rest of the library. There is the sense that the rest of the library doesn't understand what the people who are working on digital collections and repositories do and why they do it.

The BIBFRAME replacement of MARC is forthcoming, and Cataloging staff should become acquainted with it and the linked data principles upon which it is based. Tom Johnson, Maura Valentino, and Karen Estlund, former Head of the University of Oregon's Digital Scholarship Center, provided some training to interested library staff two years ago. That should be built upon. There is a wealth of library linked data training resources available on the web. The Program for Cooperative Cataloging "recommends the recently released BIBFRAME Training at the Library of Congress module 1, parts 1 and 2, which consists of an overview of Semantic Web and linked data concepts and data model (PCC, 2015)."

One person who works with linked data suggested that he'd like to understand more about the technical side of linked data: e.g. What is an N-Triple or FoaF and what does that code look

like? Conversely, programmers don't have background on descriptive metadata. There is room for knowledge to be gained on each side of the programming divide in relation to metadata and linked data.

Bibliography

Berners-Lee, T. (2006). linked data - Design Issues. Retrieved July 23, http://www.w3.org/DesignIssues/LinkedData.html

Bilder, Geoffrey; Lin, Jennifer; Neylon, Cameron (2015): What exactly is infrastructure? Seeing the leopard's spots. https://dx.doi.org/10.6084/m9.figshare.1520432.v1Retrieved: Oct 06, 2016

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. Semantic Services, Interoperability and Web Applications: Emerging Concepts, 205-227.

Coombs, K. (2014). WorldCat Discovery API and Linked Data. OCLC Developer Network. Retrieved October 06, 2016, from http://www.oclc.org/developer/news/2014/worldcat-discovery-api-and-linked-data.en.html

Estlund, K. & Johnson, T. (2013). Linked data in Oregon Digital. Presentation at UO Knight Library. Retrieved on October 18, 2016, from https://docs.google.com/presentation/d/1IZo8znenf7lOTqVg7_Tcr-keXTPuSNzr9eQ5zDVJiKo/edit#slide=id.g296f17e7e_114

Faulder, E., Ferrante, F. & Wilczek, E. (2014) Using EACCPF at an Institutional Archives, Journal of Archival Organization, 12:1-2, 21-34, DOI: 10.1080/15332748.2015.997348

Cramer, T. (2014). *The Linked Data For Libraries (LD4L) Project: A Progress Report*. Lecture presented at Coalition for Networked Information (CNI) Fall 2014 Membership Meeting, Washington, DC.

Johnson, T. (2013). Indexing linked bibliographic data with JSON-LD, BibJSON and Elasticsearch. *Code4lib Journal*, *19*, 1-11.

Johnson, T. and Boock, M. (2012). Linked Data Services for Theses and Dissertations, Proceedings of the 15th International Symposium on Electronic Theses and Dissertations, Lima, Peru. Available http://hdl.handle.net/1957/32977

Library of Congress. Bibframe. Retrieved November 1, 2016, from https://www.loc.gov/bibframe/

Linked Data - Connect Distributed Data across the Web. (n.d.). Retrieved October 06, 2016, from http://linkeddata.org/

Linked data. (n.d.). Retrieved November 02, 2016, from
https://en.wikipedia.org/wiki/Linked_data

Miller, E. (2015, April 14). *BIBFLOW: A Roadmap for Library Linked Data Implementation*.
Lecture presented at Coalition for Networked Information (CNI) Spring 2015 Membership
Meeting, Seattle, WA. Retrieved September 30, 2016, from
https://www.youtube.com/watch?v=Z-g_yJ1FL0U

Murphy, R. Opening up the BBC's Linked Data with /things. Retrieved October 31, 2016 from
http://www.bbc.co.uk/blogs/internet/entries/afdf2190-4e60-3dfc-b15f-fc17f88c85a1.

OCLC Data Strategy. https://www.oclc.org/worldcat/data-strategy.en.html Retrieved November
1, 2016.

Program for Cooperative Cataloging Standing Committee on Training. Report on Available
Linked Data Training Resources. September 9, 2015.
https://www.loc.gov/aba/pcc/sct/documents/PCCSCTFinalReportonAvailableLinkedDataTraining
Resources.docx. Retrieved November 2, 2016

Sanderson, R. (2013, April 3). *RDF: Resource Description Failures & Linked Data Letdowns*.
Lecture presented at Coalition for Networked Information (CNI) Spring 2013 Membership
Meeting, San Antonio, Texas.

Simic, J., Sato, L., Wick, R., & Mellinger, M. (2016, October 7). Case Study: Oregon Digital
Migration from CONTENTdm to Hydra for digital collections. Lecture presented at Hydra
Connect 2016, Boston. Retrieved from
https://docs.google.com/presentation/d/1QvVlsdnph_kcajF7WPhyHPDNyCYaAdqcLe9G-
5ze94E/edit?usp=sharing

Singer, R., 2009. Linked Library Data Now!, Journal of Electronic Resources Librarianship, 21,
114-126. doi:10.1080/19411260903035809

Steiner, T., Troncy, R. and Hausenblas, M (2010). How Google is using Linked Data Today and
Vision For Tomorrow. Linked Data in the Future Internet at the Future Internet Assembly. Ghent.
Retrieved from:
http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/37430.pdf

Tennant, R. Putting linked data to work. OCLC Next. http://www.oclc.org/blog/main/putting-
linked-data-to-work/

University of California San Diego Library. Metadata in the UC San Diego Digital Collections.
(n.d.). Retrieved October 28, 2016, from http://libraries.ucsd.edu/services/data-curation/sharing-
discovery/describe-your-data/repository-metadata.html

Van de Sompel, H., & Nelson, M. L. (2015). Reminiscing about 15 years of interoperability efforts. *D-Lib Magazine*, *21*(11), 1.

Vila Suero, D., 2011. Library Linked Data Incubator Group: Use Cases. Available
http://www.w3.org/2005/Incubator/lld/XGR-lld-usecase-20111025/

Westrum, A., Rekkavik, A., and Tallerås, K., 2012. Improving the presentation of library data using FRBR and Linked data, Code4Lib Journal, 16. Available
http://journal.code4lib.org/articles/6424

Yoshimura, K. S. (2016). Analysis of International Linked Data Survey for Implementers. D-Lib Magazine, 22(7), 6.

Appendix A

Internal Interviews:
Brian Davis, Digital Production Unit Supervisor
Mike Eaton, title
Korey Jackson, Grey Family Chair for Innovative Technologies
Maura Valentino, Metadata Librarian
Steve Van Tuyl, Digital Repository Librarian
Ryan Wick, title
Hui Zhang, Digital Applications Librarian

Appendix B

Interview questions:
1. How do you use linked data and in what applications?
2. How do you plan to use linked data and in what applications?
    a. Digital collections
    b. Institutional repository
    c. Data repository
    d. Library catalog/discovery system
    e. Other
3. What do you see as the benefits of using linked data?
    a. How does it benefit users currently, if at all?
    b. How do you see it benefitting users in the future?
    c. How does it benefit library internal management of information?
    d. How might it benefit library internal management of information in the future?
4. Rank on a scale of 1-5 (5 being the top score) the reasons you believe publishing linked data is important:
    a. Data re-use: One expectation of linked data is that it will enable broader use of local data by more communities.
    b. Increased discoverability: One goal for publishing linked data is to increase the discoverability of the institution's resources.
    c. New knowledge creation: Some see repositioning library knowledge work and providing access to its resources through the semantic web as a network activity enriching researchers' understanding.
    d. Thought leadership: Linked data efforts demonstrate that the institution is taking the initiative in laying the groundwork for a future, different environment. This may include good publicity and feedback among library linked data communities, demonstrating linked data possibilities and the influence on others' linked data projects.
    e. Preparation for the semantic web: Preparing existing metadata and facilitating metadata remediation for the potential future semantic web.
    f. Operational success: May include working well with other services that use the data and inspiring others to contribute content.
    g. Organizational development: Even absent metrics demonstrating linked data's value to others, linked data projects may provide professional development for staff.
    h. Organizational transformation: Changing catalog data management from MARC-based records to RDF modeling and linked data principles will require new workflows across the library community, now hampered by data duplication.
5. Rank on a scale of 1-5 (5 being the top score) what you see as barriers to creating linked data?
    a. Steep learning curve for staff
    b. Inconsistency of legacy data
    c. Selecting appropriate ontologies to represent data

       d. Establishing the links

       e. Little documentation or advice on how to build the systems

       f. Lack of tools

       g. Immature software

       h. Ascertaining who owns the data

6. What do you see as detriments (if any) of creating and publishing linked data?
7. What linked data skills/positions do we need in the library?
8. Are there other questions I should be asking?
9. Who else should I talk to?