**ALLISON-CASSIN** — **Sounding**

<0 figures>

**Sounding It Out: The Mariposa Folk Festival and a Linked Open Data Digital Library Prototype**

Allison-Cassin, S., Suhonos, M. J., Ruest, N. and St. Onge, A.

Big data, ubiquitous computing, and the quotidian nature of the digital point to the ways metadata touches all aspects of human life, making it increasingly clear that the quality of metadata and the number and strength of linkages surrounding an object in a library collection is vital. Content discovery is dependent upon the data, search interface, and algorithms that allow us to locate material, causing an awareness of the power of the knowledge structures that underlie these systems more crucial than ever, especially given that these structures are masked by the seemingly silent processing of machines (Kittler, 1999; Lessig, 2006; Olson, 2002). Despite this reality, library bibliographic control, as the professional practice of creating metadata for library collections, continues to be walled and trained along root-tree hierarchies. Nontextual materials are forced within systems that are conceptualized for text-based print resources. Linked Open Data (LOD)—as a form of structured metadata that allows for dynamic interlinking of data across the Web, creating an active web of relationships, rather than a static set of documents and as a technology created for the Web—is a way to rewire these entrenched systems.

**Objectives**

The primary goal of this project is to create a series of experimental, rhizomatic digital networks through the development of LOD prototypes of archival material related to the Mariposa Folk Festival. These prototypes will surface the multitude of relationships and contextual information normally silent in library catalogues and conceptualize this data in a fundamentally different way by activating the rich possibilities of Semantic Web technologies. Models currently being explored for the prototypes include the sonification of the network, and a network conceptualized around ideas of memory, narrative, and time.

**Harmony and Dissonance**

Much of the literature on LOD has focused on the harmonization of data as a means to create semantic relationships on the Web. Open standards provide ‘a common framework that allows data to be shared and reused across application, enterprise, and community boundaries,’ and shifts the Web from being a web of documents to a web of data (W3C, 2013) by connecting discrete pieces of data rather than links between pages (Lange and Pattuelli, 2014). Information encoded as linked data interoperates with other data coded the same way, making for a rich integration of information from various data sources, allowing the Web to function like a giant database. This results in more comprehensive and accurate search results than current federated search systems (Byrne and Goddard, 2010).  The use of standardized vocabularies and schemas creates harmony within an information system, but this comes at a cost. Voices that do not fit within the system must be left out. We are grappling with the implication of these choices and the difficult issue of creating specialized schemas in an environment that prizes standards above all else.

An increasing number of organizations are making LOD available, but there remains a dearth of projects putting this data into action, particularly coming from the academic library sector in Canada. With no standard, out-of-the-box mechanisms for creating and working with LOD, many in the community, while aware of the supposed benefits, are unsure of how to get there (Lampert and Southwick, 2013), leading to increasing calls for the library community to begin to prototype and experiment (Jordan et al., 2014; Salo, 2013; Singer, 2009), and leading Askey et al. (2013) to challenge libraries to be open to failure. Experimentation and divergent thinking is vital to innovation in a sector that is slow to change.

**The Case for Music**

Music, as a sounding art that takes place in time, is not easily captured in the manner of textual or image-based documents, making it challenging to align with practices based on the creation of textual metadata. Musical works give rise to numerous possible instantiations in various media, such as recordings, scores, and moving images, making connecting all these resources challenging (Smiraglia, 2001). Additionally there are complex interpersonal relationships of mixed authorship. A single performance may have a composer, lyricist, performers, and others, such as audio engineers. Because musical objects have a dense network that is often invisible in library systems, it is an ideal case for LOD experimentation.

The Mariposa Folk Festival is an internationally recognized Canadian music festival that began in the summer of 1961 and continues to be a key venue for the promotion of traditional music. York University holds the archives of the Mariposa Foundation. In 2009 York University Libraries received a grant from the Canadian federal government to digitize and provide public access to live sound recordings and other archival documents to celebrate the fiftieth anniversary of the Mariposa Folk Festival. Numerous questions related to the challenges of copyright, metadata, and cultural memory from this project are being explored.

**Methodology**

We are making use of practice-based research approaches as well as theoretical methods to extend the theory and practice of library systems. This can be thought of as a ‘feedback loop’ where theorizing and prototyping perform a continuous loop, and the generation of questions is the focus over the mobilization of a final, polished product.

The initial period involves the development of the conceptual models and the dataset necessary for the deployment of the prototypes. Once the existing metadata have been mapped, the team will begin to build in enhancements using standard schemas and ontologies. Much of the time in the first phase is focused on developing metadata, mapping, and cleanup, with the possible extensions to the York Digital Library’s Islandora platform.

The prototypes are the focus of the second phase. They are intended to extend technical capacity and theoretical ideas around information and mediation. If LOD is to fundamentally change the way information is navigated on the Web, then we must work to change not only the structures of the data but the conceptual models in which they are deployed.

**Challenges and Implications**

An early challenge that has developed is that much of the existing data requires checking and correction. We have reduced the size of the dataset to ensure detailed and clean metadata. This smaller grouping of documents is being grouped around a theme and time period. As much as LOD is promising, a great deal of manual work is still required.

The real work of this project is around the development of the conceptual models, in particular trying to approach the network from the idea of the sonic and the implications of ‘sound’ on the network idea and attempting to move away from the hegemonic ‘nodes and edges’ approach to LOD.

Copyright has previously proved challenging and remains a challenge for this project. Rather than choosing to not include sound documents, we are attempting to leverage community relationships, link to alternate sources, and find solutions.

**Conclusion**

This talk outlines our initial steps in creating Linked Open Data network prototypes. While we are at the initial stages of the LOD implementation, our work builds upon years of praxis and expert knowledge of library and archival descriptive practices, archival preservation, repository software, and library systems. We will discuss our plans for our LOD networks and other means of pushing the boundaries on what is considered library discovery.

**References**

**Askey, D., Jordan, M., Steeves, C. and Suhonos, M. J.** (2013). IT Experimentation, Innovation, and Failure. Panel discussion at *Access 2013,* St. John’s, Canada, 23–26 September 2013.

**Byrne, G., and Goddard, L.** (2010). The Strongest Link: Libraries and Linked Data. *D-Lib Magazine*, **16**(11/12), doi:10.1045/november2010-byrne.

**Jordan, M., Sahadath, C. and Suhonos, M. J.** (2014). Linked Open Data and Libraries: When? or NOW! BCLA Conference, Vancouver, BC.

**Kittler, F.** (1999). *Gramophone, Film, Typewriter*. Stanford University Press, Stanford, CA.

**Lampert, C. K., and Southwick, S. B.** (2013). Leading to Linking: Introducing Linked Data to Academic Library Digital Collections. *Journal of Library Metadata*, **13**(2–3): 230–53, doi:10.1080/19386389.2013.826095.

**Lange, L. and Pattuelli, M. C.** (2014). Linked Jazz: Building with Linked Open Data. EDUCAUSE Review Online, 30 June 2014, http://www.educause.edu/ero/ article/linked-jazz-building-linked-open-data.

**Lessig, L.** (2006). *Code: Version 2.0*. Basic Books, New York.

**Olson, H. A.** (2002). *The Power to Name : Locating the Limits of Subject Representation in Libraries*. Kluwer Academic Publishers, Boston.

**Salo, D.** (2013). Soylent Semantic Web Is People! (with notes). Slideshare slides presented at the *SWIB (Semantic Web Im Bibliotheken; Semantic Web in Libraries)*, Hamburg, Germany, 26 November 2013, http://www.slideshare.net/cavlec/soylent-semantic-web-is-people-with-notes.

**Singer, R.** (2009). Linked Library Data Now! *Journal of Electronic Resources Librarianship*, **21**(2): 114–26, doi:10.1080/19411260903035809.

**Smiraglia, R. P.** (2001). *The Nature of ‘a Work’: Implications for the Organization of Knowledge*. Scarecrow Press, Lanham, MD.

**W3C Semantic Web Activity Homepage**. (n.d.). http://www.w3.org/2001/sw/ (accessed 13 October 2014).