

2017

## **Stronger Together: Embracing Google and Linked Data in Law Libraries**

Elizabeth Manriquez

Follow this and additional works at: [https://digitalcommons.law.uw.edu/law-lib\\_borgeson](https://digitalcommons.law.uw.edu/law-lib_borgeson)

---

### **Recommended Citation**

Elizabeth Manriquez, *Stronger Together: Embracing Google and Linked Data in Law Libraries* (2017), [https://digitalcommons.law.uw.edu/law-lib\\_borgeson/71](https://digitalcommons.law.uw.edu/law-lib_borgeson/71)

This Article is brought to you for free and open access by the Law Librarianship Program at UW Law Digital Commons. It has been accepted for inclusion in Borgeson Paper Archive by an authorized administrator of UW Law Digital Commons. For more information, please contact [lawref@uw.edu](mailto:lawref@uw.edu).

# ***STRONGER TOGETHER: Embracing Google and Linked Data in Law Libraries***

---

Elizabeth Manriquez

Submitted to Professor Richard Jost to fulfill course requirements for Current Issues in Law Librarianship,  
LIS 595, and to fulfill the graduation requirement of the Culminating Experience Project for MLIS  
University of Washington Information School Seattle, Washington May 23, 2017

## *Embracing Google and Linked Data in Law Libraries*

### *Abstract*

The Marc format ushered in the age of automation for libraries. However, the standard is dated. The Library of Congress began developing BIBFRAME, a linked data ontology, five years ago to supplant the aging standard. It isn't merely a replacement for MARC but the beginning of a new age of automation. Coupled with improved search engines and the open access movement, linked data can revolutionize the way all libraries do business. This paper provides an overview of linked data technologies, surveys the early results of a linked data initiative, and argues for the adoption of this technology in law libraries.

### *Introduction*

“Google can bring you back 100,000 answers. A librarian can bring you back the right one.”<sup>1</sup>

The information revolution is upon us and there is no greater innovator than the Internet. For most researchers, gone are the days of searching through dusty tomes for some tidbit of information required to bolster a proposition. Now students and faculty merely turn to their laptops and boot up. Sophisticated researchers navigate databases and library catalogs with aplomb, but most patrons turn to Google before consulting any other resource for information. Many librarians view Google as the enemy, providing limited or misleading information. Simple answers are rarely the best answers. However, Google can be our ally in the fight against misinformation and alternative facts. By publishing librarian curated information to the semantic web, libraries can lead patrons to the resources they need to examine queries in depth. Imagine a student tasked with a research paper on a local hero, such as Norman Lenfest<sup>2</sup> or Molly Brown.<sup>3</sup> When this student enters the name into Google, she will typically be greeted with the immediately recognizable Wikipedia entry page, isolated news clips, and scattered non-scholarly blog postings. However, if his/her local library has published their data to the semantic web, s/he can instead see catalog holdings for a biography, a photograph from the library's historical image collection, or a cookbook named for the local hero. Not only will this information transform the student's paper, it will change the student. Positive results incentivize repeat behavior. The student is more likely to turn to the library catalog the next time s/he is conducting research. This foothold into a student's mind is all libraries need to foster lifelong relationships with patrons. Akin to walking through the looking glass<sup>4</sup>, opening the wardrobe door<sup>5</sup>, or running into platform 9 ¾<sup>6</sup>, a world of change begins with a single chance. Libraries cannot afford to miss their chance.

---

<sup>1</sup> Quote widely attributed to Neil Gaiman. YouTube video, *Neil Gaiman on Libraries*, Excerpt starts at 1 minute 20 seconds of 1 minutes 56 seconds, Uploaded by: indyPL (Indianapolis Public Library), Uploaded April 19, 2010. <https://www.youtube.com/watch?v=uH-sRluCQ6g>.

<sup>2</sup> A Snohomish, WA resident born in 1894. The Snohomish Historical Society possess several photographs of residents from this period. These photographs are available through the online catalog and accessible by search engines such as Google. <http://link.sno-isle.org/portal/Norman-Lenfest/nKQOowhLzig/>

<sup>3</sup> The “Unsinkable” Molly Brown was a Colorado resident best known for surviving the sinking of the Titanic. <http://www.mollybrown.org/learn/about-molly-brown/>

<sup>4</sup> See generally Lewis Carroll, *Alice Through the Looking-Glass and What She Found There*, Palazzo Editions Ltd., 2015.

<sup>5</sup> See generally C.S. Lewis, *The Lion, the Witch and the Wardrobe*, 1st HarperCollins ed., 1994.

<sup>6</sup> See generally J.K. Rowling, *Harry Potter and the Sorcerer's Stone*, 1st American ed., 1998.

## *Embracing Google and Linked Data in Law Libraries*

Negative economic, societal, and government factors threaten the very existence of libraries.<sup>7</sup> Now is the time to partner with the “enemy” and use that relationship to our advantage.

For centuries, librarians have been organizing information. Meticulously classifying titles, authors, and places of publication; compiling lists, catalogs, and eventually card catalogs. The development of the Machine-Readable Cataloging (“MARC”) standard changed the method of organization, but not the principles. The organization of centuries of information was not simply a persnickety exercise in futility. It facilitated access to that information. Automation was a response to the technological revolution birthed by computer. Now we are in the midst of another revolution, the development of the semantic web. Libraries can lead this revolution with the tools they have spent centuries developing: organization; innovation; and determination.

Librarians fifty years ago developed a standard to increase the access and visibility of their collections, the MARC standard. The past fifty years have seen innumerable changes in academia, technology, and the world at large. The time has come for a new standard to increase library access and visibility: linked data. Through advancements such as the Google Knowledge Graph<sup>8</sup> and Bing’s Satori<sup>9</sup>, search engines are becoming more powerful and intuitive. They are learning to search using knowledge rather than form. When libraries publish their data to the semantic web, using the vocabularies these improved search engines understand, they can increase their relevancy both on the web and in society generally.

This paper will begin by providing a very brief background on MARC, Functional Requirements for Bibliographic Records (“FRBR”), and Resource Description and Access (“RDA”). This is done to provide reference for the structure and evolution of cataloging standards.<sup>10</sup> Part II will introduce the basic mechanics of metadata, ontologies, and linked data.<sup>11</sup> Then I will discuss how Google and Bing are already using linked data through the Google Knowledge Graph and Satori. Part III will present my research and findings. Part IV will discuss the implications for libraries generally, and law libraries specifically. Part V will introduce some criticisms and conclude.

---

<sup>7</sup> See Philip Kennicott & Peggy McGlone, *Trump wants to cut the NEA and NEH. This is the worst-case scenario for arts groups*, Wash. Post, March 16, 2017, accessed electronically:

[https://www.washingtonpost.com/lifestyle/style/with-elimination-of-nea-and-neh-trumps-budget-is-worst-case-scenario-for-arts-groups/2017/03/15/5291645a-09bb-11e7-a15f-58d4a988474\\_story.html?utm\\_term=.9ee7c411870c](https://www.washingtonpost.com/lifestyle/style/with-elimination-of-nea-and-neh-trumps-budget-is-worst-case-scenario-for-arts-groups/2017/03/15/5291645a-09bb-11e7-a15f-58d4a988474_story.html?utm_term=.9ee7c411870c)

, May 23, 2017; Oleg Kagen, *Is the Internet Killing Libraries?*, Medium (April 3, 2017),

<https://medium.com/@lifeinoleg/is-the-internet-killing-librarians-88076b86fdda>; and Charles I. Guarria & Zhonghong Wang, *The economic crisis and its effect on libraries*, 112 *New Library World*, Issue: 5/6, pp.199-214 (2011).

<sup>8</sup> See <https://www.google.com/intl/es419/insidesearch/features/search/knowledge.html>.

<sup>9</sup> See <https://blogs.bing.com/search/2013/03/21/understand-your-world-with-bing/>.

<sup>10</sup> This is merely a brief overview of the standards necessary to understand the mechanics of linked data and how it relates to the bigger picture of cataloging standards. The author acknowledges this introduction omits many concepts, rules, and standards to cataloging. It is beyond the scope of this paper to provide a more in depth history or explanation of cataloging principles. See Karen M. Spicher, *The Development of the MARC Format*, 21 *Cataloging & Classification Quarterly* No. 34, pp. 75-90 (1996), for a more detailed history of the MARC format.

<sup>11</sup> Again, this is the briefest introduction to a complicated technological concept. It would require reams of paper to fully explain the computer science of linked data. This explanation merely introduces the concept to non-computer scientists to provide a basis for my greater argument regarding its importance to modern libraries. See Tim Berners-Lee, James Hendler & Ora Lassila, *The Semantic Web*, *Scientific American*, May 17, 2001, <https://pdfs.semanticscholar.org/566c/1c6bd366b4c9e07fc37eb372771690d5ba31.pdf>, for a more detailed explanation of linked data and the semantic web.

*Part I: When Did It All Begin?*

**A. MARC**

While the MARC standard has endured for half a century, it was not the first attempt at automation for librarians at the time.<sup>12</sup> These earlier efforts emphasized the need for a database of machine-readable records. The Library of Congress (“LC”) responded to this need with a task force headed by Henriette Avram.<sup>13</sup> Avram was not a librarian, but a systems analyst and computer programmer.<sup>14</sup> Thus, began the marriage of computer science and library science. The task force delivered a report in June, 1965.<sup>15</sup> It proposed a format that “would conform to current LC cataloging practices so that existing work patterns could be appropriated for the preparation of conversion copy.”<sup>16</sup> This format was comprised of both variable fields and fixed fields.<sup>17</sup> Fixed fields were given to elements such as publication date, author, and title to allow for fast searching,<sup>18</sup> while variable fields provided organization of “catalog card text for machine access.”<sup>19</sup>

Following the release of the report, the Council on Library Resources awarded a grant to LC “for a pilot project in the creation and distribution of LC cataloging data in machine-readable form. The MARC Pilot Project, under the direction of Henriette Avram, was formed to test a system for the distribution of records to a sample group of libraries. These libraries would define uses for the data and report to LC on the adaptation of records for local use and the efficiency of the distribution system.”<sup>20</sup> This pilot project was relatively successful and spanned about a year, commencing in November, 1966, and concluding in October, 1967.<sup>21</sup> The results of this early effort led to the formation of MARC II, the standard we are most familiar with today.<sup>22</sup> There were many advances from MARC I and MARC II, namely the accommodation of all types of material, “such as monographs, serials, maps, or music.”<sup>23</sup> Other changes included the use of three digits in variable fields, instead of the two digit fields MARC I contained.<sup>24</sup> These three digit fields allowed for “tagging schemes...; tags were coded to reflect both the function and the type of the information contained in the field. The first digit continued to group fields by function, such as title, physical description, or series, while the second digit identified, where appropriate, the type of information, such as personal, corporate, or uniform title.”<sup>25</sup>

The standardization of data introduced by LC and MARC facilitated “the widespread use of machine-readable bibliographic data.”<sup>26</sup> It allowed for the spread of automation which had

---

<sup>12</sup> Spicher, *supra* note 10, at 76-78.

<sup>13</sup> *Id.* at 81.

<sup>14</sup> *Id.*

<sup>15</sup> *Id.* at 82.

<sup>16</sup> *Id.*

<sup>17</sup> *Id.*

<sup>18</sup> *Id.*

<sup>19</sup> *Id.*

<sup>20</sup> *Id.* at 83.

<sup>21</sup> *Id.* at 84.

<sup>22</sup> *Id.*

<sup>23</sup> *Id.*

<sup>24</sup> *Id.* at 85.

<sup>25</sup> *Id.*

<sup>26</sup> *Id.* at 86.

been cost prohibitive for many libraries, and most importantly it redefined interlibrary communication.<sup>27</sup> The intention of LC in the formation of MARC was “to produce catalogue records capable of processing by computer, not only for searching (its primary function) but also for editing, updating, transferring between systems and sharing.”<sup>28</sup> Even at this early juncture, libraries were sharing bibliographic data for the greater good. When she was interviewed decades after the rollout of MARC II, Henriette Avram stated, “I believe overall librarians have done wonders. The very fact that so early on they, as a profession, understood the need for and adopted standards has given them a leadership role in resource sharing.”<sup>29</sup> This role as resource sharers was the first steps towards building two standards needed for linked data to succeed: cooperation between libraries and relationships between resources.

## **B. FRBR**

In the 1990s, LC and other major institutions involved in bibliographic standards, both national and international, recognized the need to review cataloging traditions with a perspective towards the emerging digital environment.<sup>30</sup> In response, the International Federation of Library Associations (“IFLA”), developed “terminology, relationships and user tasks” that they published in a report known as FRBR.<sup>31</sup> FRBR further explored the relationships between resources, which were changing concurrently with technologies and user demands.<sup>32</sup> The FRBR model involves separating materials into a cascading system of Group 1 entities known as “Work, Expression, Manifestation, and Item (WEMI).”<sup>33</sup> Beginning with the most abstract (work) and continuing to the most concrete (item), the model places importance on the connections between different materials, the entities involved in their creation, and the relative attributes of each.<sup>34</sup> This building on relationships forms the foundations for linked data ontologies and standards.

---

<sup>27</sup> *Id.*

<sup>28</sup> Richard Gartner, *Metadata: Shaping Knowledge from Antiquity to the Semantic Web*. Switzerland: Springer, 2016 at 29.

<sup>29</sup> Lucia J. Rather & Beacher Wiggins, “Mother Avram’s Remarkable Contribution” *Henriette D. Avram: Close-up on the career of a towering figure in library automation and bibliographic control*, 20 *American Libraries*, No. 9, pp. 855-861, (1989), at 860.

<sup>30</sup> IFLA Study Group on the Functional Requirements for Bibliographic Records, and International Federation of Library Associations Institutions. Section on Cataloguing. Standing Committee. *Functional Requirements for Bibliographic Records: Final Report*. München: K.G. Saur, at 6 (1998).

<sup>31</sup> *Id.*

<sup>32</sup> Rachel Ivy Clarke, *Breaking Records: The History of Bibliographic Records and Their Influence in Conceptualizing Bibliographic Data*, 53 *Cataloging & Classification Quarterly*, No. 3-4, pp. 286-302, at 294-295 (2015).

<sup>33</sup> IFLA Study Group on the Functional Requirements for Bibliographic Records, and International Federation of Library Associations Institutions. Section on Cataloguing. Standing Committee, *supra* note 30, at 18.

<sup>34</sup> *See Id.*

### **C. RDA**

Where MARC is a communication format<sup>35</sup>, and FRBR is a conceptual model for the data, RDA emerges as a content standard for data, the what and how of recording information.<sup>36</sup> While I will not explore the mechanics of RDA, I include it in this examination because it serves as an important “half step” between the old and the new standards of cataloging. For example, RDA “relies on linking of records for each work, each manifestation of that work, each expression of each manifestation, and possible each item. These records are then related to one another using FRBR relationships. Relationships therefore exist at the record level, relating a collection of data elements to another collection of data elements, rather than at the level of the data elements themselves.”<sup>37</sup> This practice is also known as “entity relation (ER) modeling.”<sup>38</sup> As we will discuss later in this paper, linked data functions at a much more granular level, linking the data elements themselves. However, the creation of a content standard that acknowledges relationships even at the record level is still an important milestone in the journey to complete linkage. RDA is also an important nexus for linked data ontologies what emerged in response to the linked open data initiative, such as FRBRer.

Armed with an awareness of these basic standards in cataloging, we can begin to explore the mechanics of linked data and its departure from the standards of the past.

#### *Part II: How Does It Work?*

### **A. A Meta Moment**

When speaking in terms of cataloging, we often use the word “data” to describe the information entered into the MARC record. However, in the digital age what we are really discussing is metadata. Metadata is simply put “data about data.”<sup>39</sup> It describes the informative and relevant description about the original data.<sup>40</sup> Where the book is the data, the date acquired by the library is the metadata. There are three categories of metadata: descriptive, administrative, and structural.<sup>41</sup> While administrative and structural data are essential to the function of an information system, this paper will focus on descriptive metadata.<sup>42</sup> It “is the metadata designed to help us discover and locate the data it refers to.”<sup>43</sup> For instance, in libraries and archives, metadata is the information available in the item record, such as “the author and the title of a

---

<sup>35</sup> American Library Association's ALCTS/LITA/RUSA Machine-Readable Bibliographic Information Committee in conjunction with Network Development and MARC Standards Office Library of Congress, *The MARC 21 Formats: Background and Principles*, Revised November 1996, accessed online May 23, 2017, <https://www.loc.gov/marc/96principl.html#one>.

<sup>36</sup> Thomas Baker, Karen Coyle & Sean Petiya, *Multi-entity models of resource description in the Semantic Web: A comparison of FRBR, RDA and BIBFRAME*, 32 *Library Hi Tech*, Issue: 4, pp.562-582, at 570 (2014).

<sup>37</sup> Clarke, *supra* note 32, at 295.

<sup>38</sup> Baker, *supra* note 36 at 568.

<sup>39</sup> Jenn Riley, *Understanding Metadata: What is Metadata? And What is it for?*, National Information Standards Organization, pp. 1-45, at 1 (2017).

<sup>40</sup> *See Id.*

<sup>41</sup> Gartner, *supra* note 28 at 6.

<sup>42</sup> *Id.*

<sup>43</sup> *Id.*

book, who published it, when it was published and what its subject is. It may also cover information that can help us to identify it unambiguously, such as its ISBN (International Standard Book Number).”<sup>44</sup> Linked data is built on the foundation of good descriptive metadata. By linking ontologies and descriptive metadata, computers can move from understanding content to deriving knowledge about content.<sup>45</sup> This movement from information to knowledge is made possible through semantic links. For instance, two components of information with descriptive metadata assigned can immediately form a connection if they share an attribute.<sup>46</sup> These connections “express some meaning about the relationships between the information they join together. A network of these rapidly allows us to begin to answer the ‘how’ questions that are the domain of knowledge.”<sup>47</sup> Metadata alone cannot accomplish this, just as RDA and FRBR alone cannot form the links necessary to weave the semantic web, all data must be expressed in a linked data ontology based on Resource Description Framework (“RDF”).<sup>48</sup>

## **B. More than Semantics, how Linked Data works**

In 2001, Sir Tim Berners-Lee, James Hendler, and Ora Lassila first published their vision for the semantic web.<sup>49</sup> Through the use of eXtensible Markup Language (“XML”) and the RDF, they reasoned the semantic web could improve cooperation between man and machine.<sup>50</sup> Rather than linking documents, the semantic web depends on the linking of data through the use of RDF and ontologies.<sup>51</sup> Ontologies, or vocabularies, provide the language and “RDF provides the grammar for a language of data. The language is used to assert ‘facts’ (or ‘claims’) about the world in the form of three-part, sentence-like statements called ‘triples.’”<sup>52</sup> These triples are comprised of a subject, a predicate, and an object.<sup>53</sup> Figure 1 demonstrates an example of a triple. The relationship between the subject, Margaret Atwood, and the predicate, *The Handmaid’s Tale*, is described by the predicate, “Is the Author Of.” Triples are simple expressions and nearly all metadata can be deconstructed into these easier to process “molecules of information.”<sup>54</sup> These structures are also fluid, as shown in Figure 2 where a new triple is formed and the subject becomes the object. Figure 3 then shows the linking of all three pieces of data. This linking of seemingly disparate pieces of data can happen an infinite number of times, creating exponentially more triples as more data is introduced to the semantic web.

---

<sup>44</sup> *Id.*

<sup>45</sup> Gillian Byrne & Lisa Goddard, *The Strongest Link: Libraries and Linked Data*, 16 D-Lib Magazine, No. 11/12, (2010), accessed electronically May 23, 2017, <http://www.dlib.org/dlib/november10/byrne/11byrne.html>.

<sup>46</sup> Gartner, *supra* note 28 at 11.

<sup>47</sup> *Id.*

<sup>48</sup> See <https://www.w3.org/standards/semanticweb/data>.

<sup>49</sup> See Berners-Lee, *supra* note 11.

<sup>50</sup> *Id.*

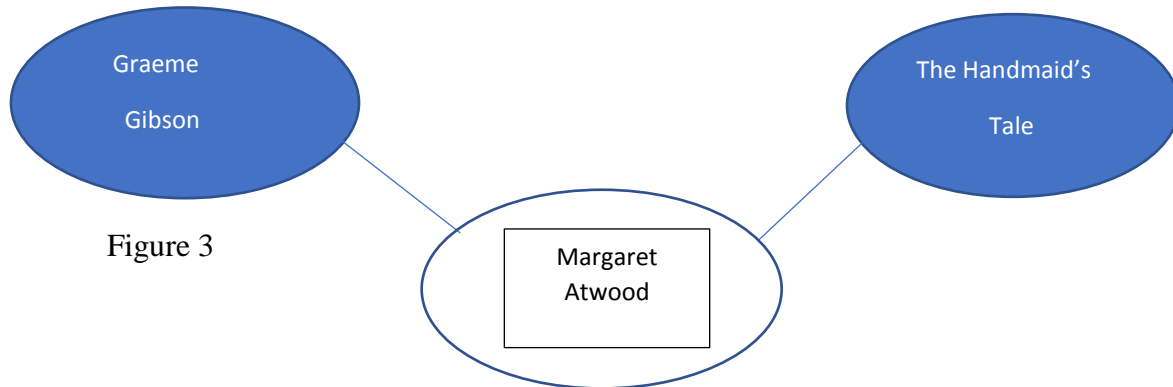
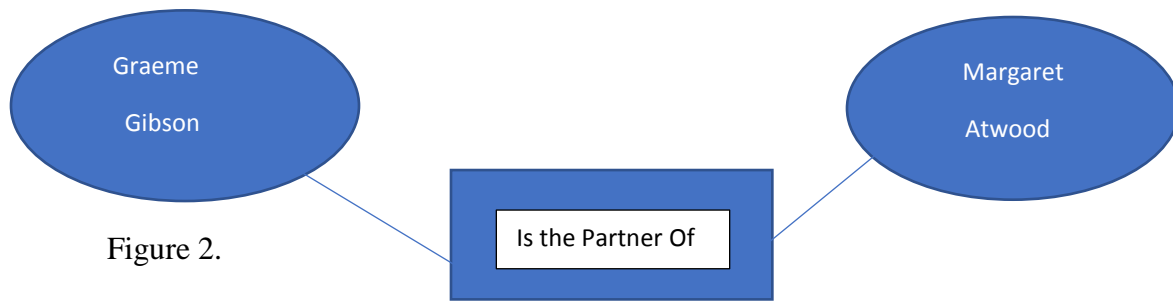
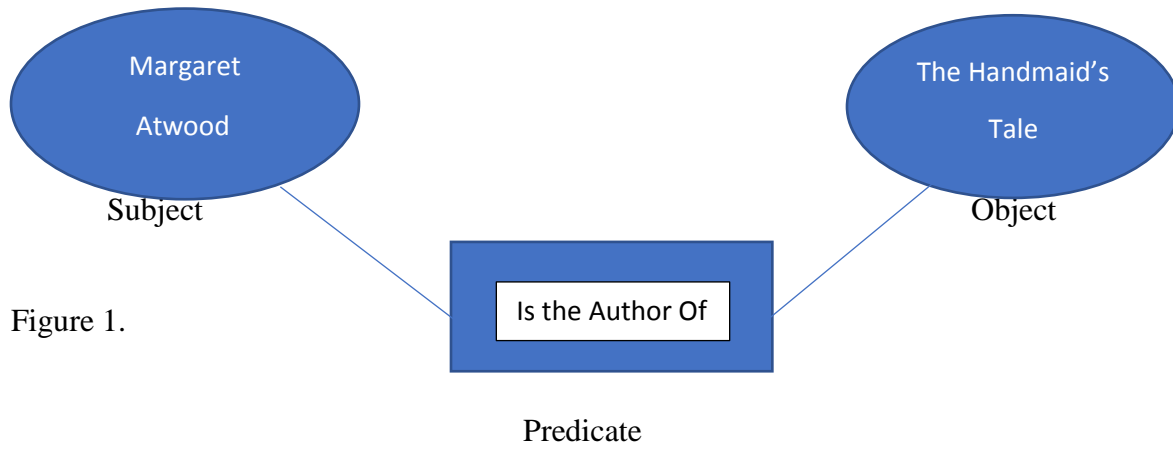
<sup>51</sup> *Id.*

<sup>52</sup> Baker, *supra* note 36 at 564.

<sup>53</sup> Clark, *supra* note 32 at 295-296.

<sup>54</sup> Gartner, *supra* note 28 at 89.





However, one more step is necessary to conform to linked data standards, the introduction of Uniform Resource Identifiers (“URIs”).<sup>55</sup> “For the model to work successfully, however, subjects and predicates must be represented by URIs - unique character strings used to ‘distinguish one resource from all other resources.’ Every resource, physical or digital, concrete or abstract, is assigned a standardized character string that uniquely identifies that resource among other resources. A URI may also specify location of a resource, especially in the case of digital resources. A URI is a *single* data element that uniquely describes (and more often than not, locates) a resource.”<sup>56</sup> These URIs are derived from a specific ontology, or vocabulary, or language of linked open data. These ontologies are the final piece necessary for the creation of knowledge.

### **C. Ontologies: Now You’re Speaking My Language**

Since the introduction of linked data, several ontologies have emerged. One already introduced in this paper is FRBRer. Others include, but are not limited to: schema.org, DBpedia, and BIBFRAME.<sup>57</sup> This paper will only discuss BIBFRAME for two reasons: my research focused on an initiative using this vocabulary; also, it is the vocabulary developed by LC to transition from MARC format. It has also been argued that, “BIBFRAME’s vocabulary is more compatible with the open web and with potential variations in bibliographic concepts than is FRBRer.”<sup>58</sup> Similar to other ontologies, BIBFRAME builds on the basic framework of FRBR while incorporating the standard triples of RDF. Figure 4 illustrates the BIBFRAME model. BIBFRAME combines the FRBR work and expression for their entity work; they define Work as “the highest level of abstraction, a Work, in the BIBFRAME context, reflects the conceptual essence of the cataloged resource: authors, languages, and what it is about (subjects).”<sup>59</sup>

The FRBR concept of manifestation is similar to the BIBFRAME concept of instance.<sup>60</sup>

---

<sup>55</sup> Clark, *supra* note 32 at 296.

<sup>56</sup> *Id.*

<sup>57</sup> BIBFRAME is an abbreviation for Bibliographic Framework; Baker, *supra* note 36 at 563.

<sup>58</sup> Karen Coyle, *FRBR: Before and After*, ALA Edition, Chicago, 2016, at 147.

<sup>59</sup> Baker, *supra* note 36 at 572-573; and See <https://www.loc.gov/bibframe/docs/bibframe2-model.html>.

<sup>60</sup> *Id.*

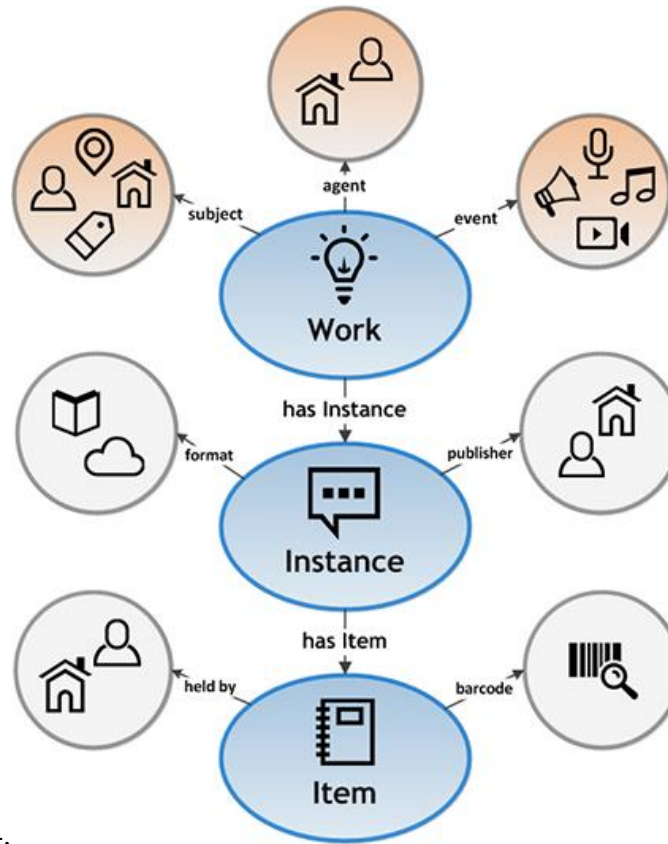


Figure 4.

LC explains Instance as it relates to Work. “A Work may have one or more individual, material embodiments, for example, a particular published form. These are Instances of the Work. An Instance reflects information such as its publisher, place and date of publication, and format.”<sup>61</sup> Similar to the FRBR concept of item, a BIBFRAME Item is “an actual copy (physical or electronic) of an Instance. It reflects information such as its location (physical or virtual), shelf mark, and barcode.”<sup>62</sup> BIBFRAME provides a structure and LC features many tools on their website to assist libraries interested in converting their legacy records to URIs and then publish their catalog online.<sup>63</sup>

#### **D. Google Knowledge Graph**

Once someone has chosen an ontology, used URIs to denote their things, and published this data online, there is only more factor needed to increase data linkage and visibility: search engines. Traditional search engines operate by using a combination of algorithms and crawlers.<sup>64</sup> These crawlers browse the documents, also known as web pages, published to the

<sup>61</sup> See <https://www.loc.gov/bibframe/docs/bibframe2-model.html>.

<sup>62</sup> *Id.*

<sup>63</sup> *Id.*

<sup>64</sup> See <http://www.bbc.co.uk/guides/ztbjq6f>.

internet looking for content that matches a query.<sup>65</sup> The search engine then builds an index using the Uniform Resource Locators (“URL”s).<sup>66</sup> Similar to URLs, which are the address of the various web documents, URIs, which are the address of various web data, are searchable. However, searching URIs requires an advanced search engine. Google, Microsoft, and Yandex each launched their own semantic search engines within the last five years. They are respectively: Knowledge Graph; Satori; and Object Answer.<sup>67</sup> These search engines, “want to understand the user queries semantically and serve their information needs precisely from their knowledge repositories. They want to answer many of the information needs directly. They are working to build large knowledge repositories about real world entities and concepts.”<sup>68</sup> These entities include “people, places, books, movies, events, arts, science, etc.”<sup>69</sup>

Results of semantic search engines have already permeated our experiences with the web. For example, traditional search engines take a query and match the terms against its internal index and deliver a result page consisting of a list of web documents that it “thinks” best meets the criteria in the query. This result page is known as a Search Engine Result Page (“SERP”).<sup>70</sup> In 2012, SERPs began to change appearance as semantic search engines began to auto-populate the page with query answers.<sup>71</sup> One example of these improved SERPs was discussed by Google at their Keynote speech during the Google I/O conference in 2013. According to Amit Singhal, Google Knowledge Graph will be able to answer, converse and anticipate user queries.<sup>72</sup> This is accomplished through three types of searches: single entity search; entity list search; and attribute search. Single entity searches involve a query “about a particular entity.”<sup>73</sup> Searching virtually any specific entity will return a SERP with the list of webpages on the left-hand side of the SERP while the right-hand side of the SERP is populated with pictures and attributes of the entity.<sup>74</sup>

See Figure 5 for an illustration of single entity search. Here the entity is “bananas.” The right-hand of the SERP is populated with images and attributes, such as nutritional information.

---

<sup>65</sup> *Id.*

<sup>66</sup> *Id.*

<sup>67</sup> *Supra* note 8; *Supra* note 9; and See [https://en.wikipedia.org/wiki/Knowledge\\_Graph](https://en.wikipedia.org/wiki/Knowledge_Graph).

<sup>68</sup> Ahmet Uyar & Farouk Musa Aliyu, *Evaluating search features of Google Knowledge Graph and Bing Satori: Entity types, list searches and query interfaces*, 39 *Online Information Review*, Issue: 2, pp.197-213, at 197 (2015).

<sup>69</sup> *Id.* at 198.

<sup>70</sup> *Id.* at 200.

<sup>71</sup> Amit Singhal, *Keynote Speech at Google I/O* (streamed live by Google Developers May 15, 2013), [https://www.youtube.com/watch?v=9pmPa\\_KxsAM#t=1h51m10s](https://www.youtube.com/watch?v=9pmPa_KxsAM#t=1h51m10s).

<sup>72</sup> *Id.*

<sup>73</sup> *Supra* note 68 at 200-201.

<sup>74</sup> *Id.*

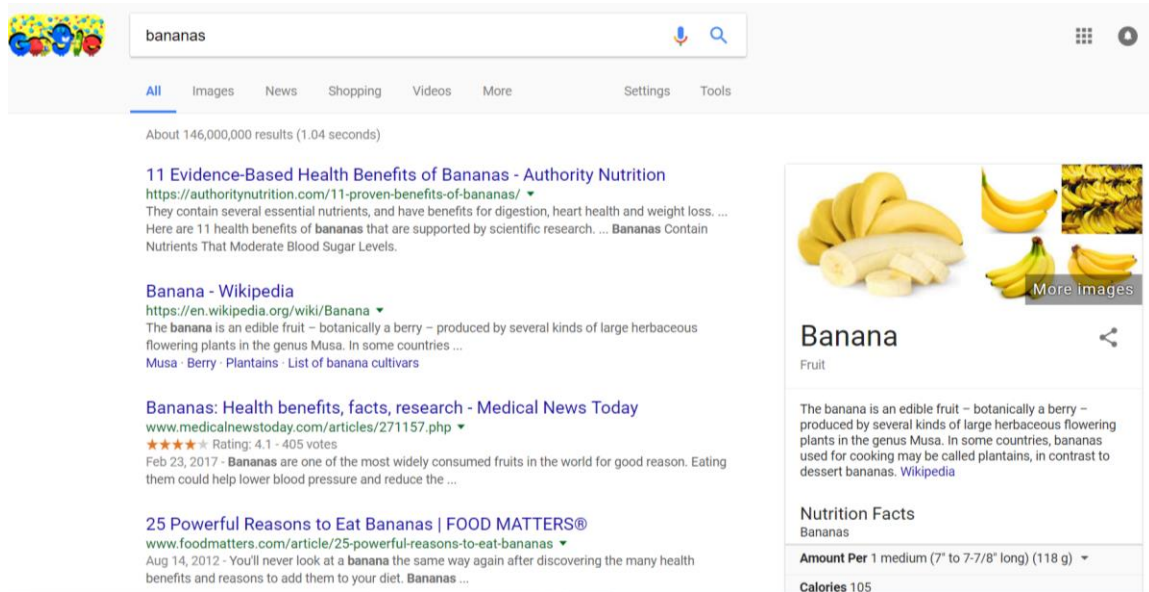


Figure 5

The second type of query is entity list search. These queries involve a user constructing a search involving terms such as “how many” where the answer will involve multiple entities. When these queries are entered into Google Knowledge Graph, “a list of entities is presented in a carousel at the top of the SERP.”<sup>75</sup> Each entity is represented by a picture in the carousel.<sup>76</sup> Attributes are not listed; however, selecting an entity from the carousel will result in the right-hand SERP populating with information about the chosen entity.<sup>77</sup>

The final type of query is attribute search. An attribute search seeks information about an entity, but asks a specific question about said entity rather than simply searching using a term.<sup>78</sup> Google Knowledge Graph and other “semantic search engines try to provide direct answers to factual queries.”<sup>79</sup> The search engine identifies the appropriate entity “and then retrieve[s] the requested result by identifying the relevant attribute.”<sup>80</sup> The answer, the attribute that answers the factual inquiry, is then displayed prominently in a box at the top of the SERP with the right-hand portion of the SERP populated with the entity information.

A recent study of the functionality of semantic search engines Knowledge Graph and Satori using a data set of 877 geographic entities divided into four categories of complexity.<sup>81</sup> These categories were labeled: simple; moderate; complex; and more complex.<sup>82</sup> Neither search

<sup>75</sup> *Id.*

<sup>76</sup> *Id.*

<sup>77</sup> *Id.*

<sup>78</sup> *Id.*

<sup>79</sup> *Id.*

<sup>80</sup> *Id.*

<sup>81</sup> *Id.* at 207.

<sup>82</sup> *Id.*; The authors of the study described the categories thusly, “(1) Simple queries. The query has a simple grammatical structure with one verb and unambiguous intent. The query should target a single entity, a single attribute or a set of entities belonging to a class and satisfying only one condition. A few examples would be: ‘What is the capital of Utah?’, ‘What state is Austin in?’, ‘How big is Alaska?’, ‘How long is the Colorado river?’, ‘What are all the rivers in Texas?’, ‘What are the cities in California?’; (2) Moderate queries. The query has one or two verbs. In addition, it should have one keyword that requires conditional or selective processing. The keyword may

## *Embracing Google and Linked Data in Law Libraries*

engine was able to answer complex or more complex queries, however they fared better when asked moderate and simple queries.<sup>83</sup> Satori was able to answer 7 percent of moderate queries.<sup>84</sup> Both search engines were able to answer more than half of the simple queries posed to them.<sup>85</sup> While these results may seem unremarkable, certain factors encourage proponents of semantic search engines. This study was conducted in 2014, only two years after semantic search engines were introduced. As the authors noted in their conclusion, the technology surrounding these search engines is ‘rapidly evolving.’<sup>86</sup> As time progresses, these semantic search engines will expand their indexes and “learn” to return answers to more complex queries. Evidence of this can be found in this very study; the authors noted in their conclusion that through the course of their original research, some entities were not originally searchable, but became so later within the time period of the study.<sup>87</sup> Second, the authors note that user queries are “rarely... grammatically complex” and as user awareness of semantic searching increases, so will their ability to craft complex queries.<sup>88</sup> Essentially, the technology and the user base must both evolve. As both parties learn to speak the others’ language, the conversations they share will naturally become more complex and lead to deeper answers.

### *Part III: Where Are We Now?*

To evaluate whether law libraries should implement linked data standards into their catalogs, I studied the effectiveness of an existing linked data initiative. The initiative I chose is the Library.Link Network.

Library.Link Network is an initiative spearheaded by Zepheira Technologies in collaboration with several major stakeholders in library publishing and distribution, such as SirsiDynix, Innovative Interfaces, NoveList, and EBSCO. Their mission is to increase library visibility through web relevance. According to their website, since 2014 they have collaborated with “1000+ library systems.”<sup>89</sup> Zepheira has partnered with a diverse set of libraries. Many were members of consortia, or represented multi-branch libraries; there were also several international library systems. These redundancies decreased the possible pool of participants considerably. Based on their participating partner list, I sent 88 introductory emails asking for survey participation. I received 50 affirmative responses from these emails and surveys were then sent to these libraries. Of these 50 surveys sent, 35 were completed for a 70% participation

---

specify an ambiguous intent (major cities), require determination of maximum/minimum values (largest city, longest river) or involve calculation of a conditional statement (states bordering Texas). A few example queries would be: ‘What are major rivers in Texas?’, ‘What is the biggest city in Arizona?’, ‘What is the biggest state?’, ‘What is the capital of the smallest state?’, ‘What state has the most cities?’, ‘States bordering Iowa?’; (3) Complex queries. The query requires two conditional calculations or selective processing. It can have one or two verbs. It may have compound or nested grammatical structures. For example, ‘How high is the highest point in the largest state?’, ‘What are the highest points of states surrounding Mississippi?’, ‘Which states does the longest river cross?’; (4) More complex queries. The query requires more than two conditional calculations or selective processing, for example: ‘What states border states that border states that border Florida?’ and ‘What are the largest cities in the states that border the largest state?’”

<sup>83</sup> *Id.* at 208

<sup>84</sup> *Id.*

<sup>85</sup> *Id.*

<sup>86</sup> *Id.* at 211

<sup>87</sup> *Id.*

<sup>88</sup> *Id.*

<sup>89</sup> See <http://library.link/>.

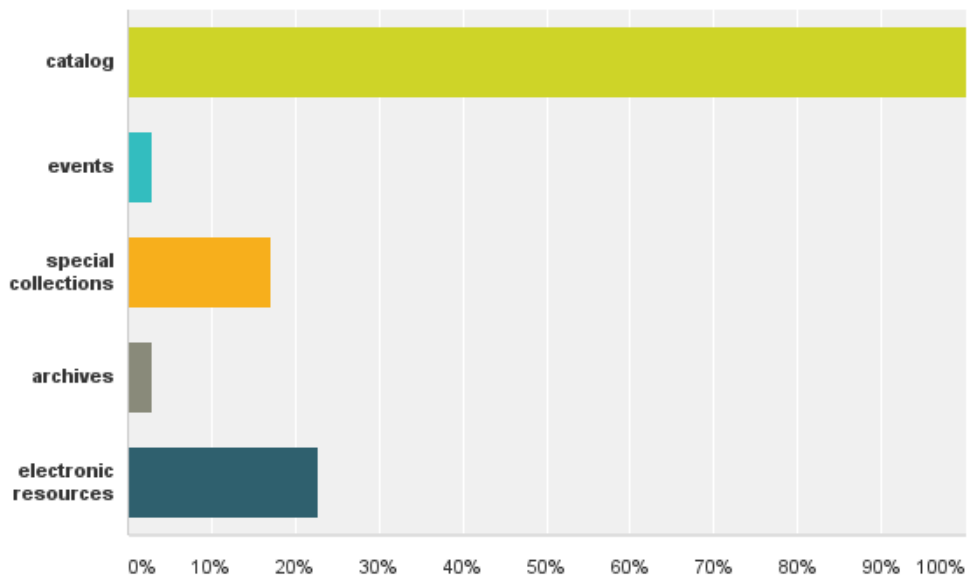
*Embracing Google and Linked Data in Law Libraries*

rate. Surveys were sent via a web link to ensure the anonymity of participants. Email addresses, names and library characteristics were not associated with responses to protect the privacy of participants. All but one survey was sent to a public library system. They varied in size, number of branches, and community demographics. District, county, community, and one academic libraries were represented. The majority of libraries reside in the continental United States, however there were respondents from Canada as well.

The survey consisted of ten questions.<sup>90</sup> These questions were selected to gauge the effectiveness of the initiative from the perspective of the participants.

*Research Results*

A. What content/services did your library include in your linked data migration?



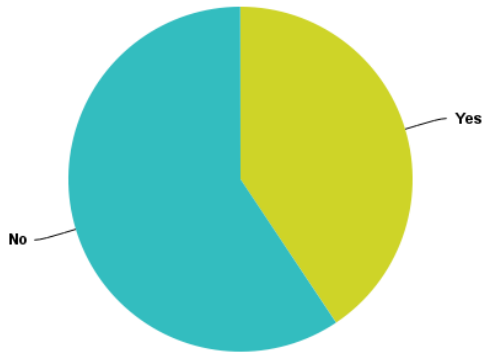
All participating libraries converted their catalog records to linked data, however only one library included events and archives in their conversion. Nearly 23% of respondents included electronic resources in their data conversion and 17% included special collections.

---

<sup>90</sup> See APPENDIX 1

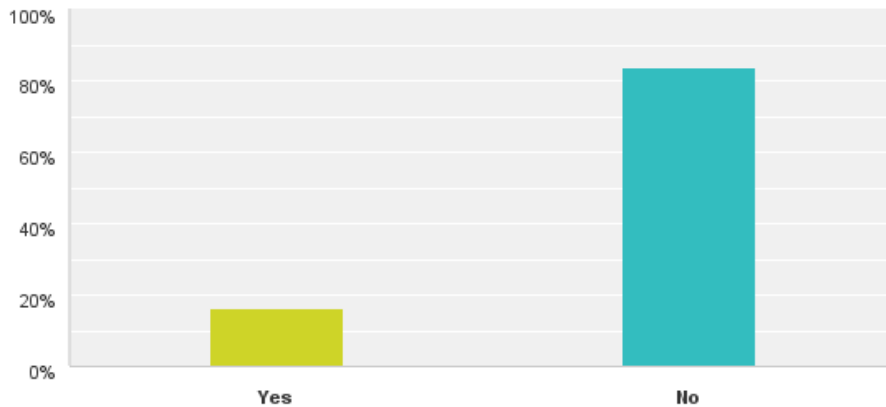
*Embracing Google and Linked Data in Law Libraries*

B. Did you see increased traffic to your website after publishing your data to the Web?



Approximately 60% of respondents replied negatively to this question. Many respondents stated they could not confirm an increase in traffic because it was too soon after implementation, their analytics did not isolate linked data redirects, or they have not experienced an increase in traffic.

C. Was there an increase in the number of patrons visiting the library after your data was published, compared to numbers regularly seen during that time period?

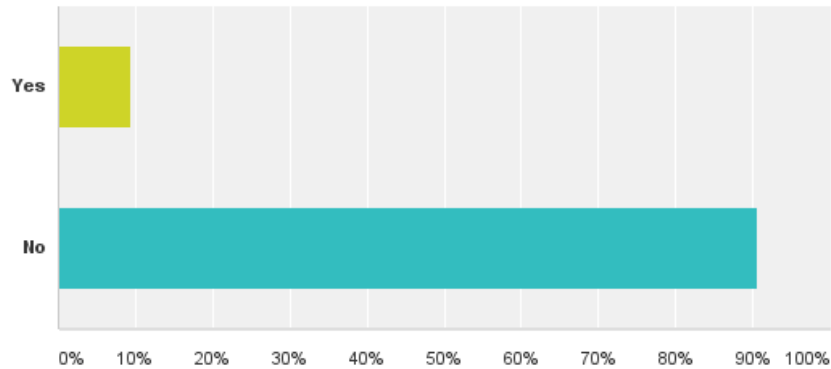


Approximately 16% of respondents replied in the affirmative to this question. The majority of respondents stated they could not confirm a connection between linked data and an increase in patron traffic. Reasons cited included: too soon after implementation to measure; lack of an analytic to measure foot traffic; and intervening factors. Many respondents described other changes to their library, such as moving to a new building and introducing a summer reading program, that may have contributed to a rise in foot traffic. Without holding all other factors constant, it's difficult to measure whether an increase or decrease in traffic can be attributed to the addition of linked data.



## *Embracing Google and Linked Data in Law Libraries*

- D. Was there an increase in new patron registrations after your data was published, compared to numbers regularly seen during that time period?



While over 9% of respondents stated they did see an increase in new patron registrations, the overwhelming majority stated there was no increase or they lacked a tool to measure a causal relationship to linked data.

- E. If you chose to include events in your linked data conversion, did you see increased attendance at library events following publishing your data to the web?

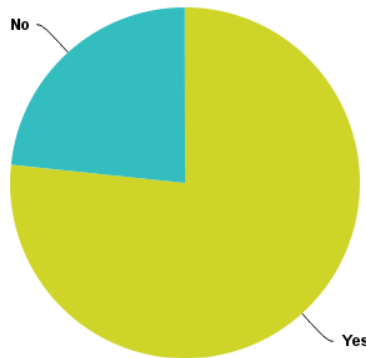
Only one participant who responded to the survey chose to include events in their linked data conversion, however this respondent did report an increase in attendance following publication.

- F. If you chose to include special collections in your linked data conversion, did you see increased requests to access these collections, compared to numbers regularly seen during this time period?

Of the libraries surveyed, six chose to include their special collections in the migration to linked data. Two of these libraries, approximately 33%, confirmed an increased interest in special collections following the conversion. One respondent theorized the uniqueness of their special collection holdings, as opposed to more commonly held titles, caused search engines to identify and return these holdings to a greater degree.

## *Embracing Google and Linked Data in Law Libraries*

- G. A main mission of The Library.Link Network is to improve web visibility of library data, do you feel your library is more visible now that your data is published online?



Despite the lack of citable evidence to the success of linked data in their libraries, respondents clearly felt their library was more visible online as a result of publishing their data to the semantic web. The libraries who responded in the negative stated they believed it was too early to tell whether the experiment was working. Many libraries cited a lack of analytics to accurately state whether linked data contributed to an increase in visibility. Others stated that their resources were more visible, but their holdings would only appear when the search query including the name of the library as well, for example, “The Lord of the Rings ‘X’ public library.” Others offered anecdotal evidence to support their belief in increased visibility: “We have received comments from people who have found the links through search engines;” “We do see results in Google now;” and “We have received calls and email from non-[library] users that discovered items in our collection through Google, Bing, etc.”

- H. Question 8: What was the biggest challenge you faced during this process?

Most respondents to this question stated the lack of reliable analytics made this a difficult project to continue. They said they lacked a method to “evaluate the effectiveness of linked data. Even with Google analytics, we haven’t been able to determine how often people are getting redirected to our catalog.” Some felt the lack of immediate results made the project difficult to justify to administration, “Our biggest challenge was managing expectations during the process. Some people expected immediate results and linked data is really a long-term project. We’re going to need to see thousands of libraries publishing their data in a variety of different linked data vocabularies in order to see a big impact.”

Multiple libraries cited the need to include the library name in the search query as an obstacle, “Implementation requires a lot of patience. If you bake a cake, you mix the ingredients and short while later it comes out of the oven and you can eat it. Discovery online takes several months to the better part of a year before institution / library data begins to rise in relevance and becomes discoverable in any meaningful way, based on our experience. This is such a new endeavor that there are bound to be hiccups and learning curve. Geo-spatial location of library assets is a key goal of the discovery process. Despite geo-tagging of our data assets, it is often

## *Embracing Google and Linked Data in Law Libraries*

necessary to include the name of the library in the user's search to produce a positive hit. This will eventually be addressed but currently is a hurdle to reaching the most desirable demographic to us: the user that wasn't thinking of us when they began his/her search.” Many thought the process of publishing data was simple as vendors handled most of the implementation work, but struggled to educate their staff and community on what linked data is, how it works, and why it is important to the library. Another library stressed the need for better search engines. Most libraries were forward looking in their evaluations, but stressed the need for patience.

### I. Question 9: What advice would you give to a library considering publishing their data to the web?

Overwhelmingly, respondent advice included: have patience; have an outside company do the conversion for you; and participate. Several libraries stressed the importance of patience, “Do not expect immediate results. If you have the funds to take on a project like this, please do. We need to make ILS companies see that we are looking to the future and want linked data to become a standard going forward.” Several respondents noted that participating in a project of this nature is a risk and it may be some time before there is a visible return on investment. Others stressed the importance of managing expectations at all levels and recommended gathering major stakeholders before deciding to participate, “Do your research and get a clear picture of how it all works, including talking to libraries that are participating. Involve key staff members so that they also understand.”

Many libraries recommended cleaning up data before having an outside company perform the conversion. As one library stated, “1. Be patient 2. Make sure you have a clean house before you invite guests over (clean up your catalog bib/item data to ensure you're presenting an accurate collection to the entire world) 3. You are investing in the future and the sooner you do it, the sooner relevancy will occur (see #1).”

Most librarians' recommendations stressed the importance of libraries working together, not just by participating, but by talking to one another to improve results. One library stated, “I'd recommend talking to other, like libraries – those in a similar geographic area and with a similar collection size – who've published their data and are regularly assessing the impact.” Even the most enthusiastic recommenders also reiterated the importance of patience, “Give it a try! But don't expect immediate results. I think this effort is very promising and will become important to libraries in the coming years.” While not all respondents were as enthusiastic, “There are a lot of grand promises and great sounding ideas, but there haven't been a lot of results” overall, most respondents expressed a belief in the promise of linked data, “Do it. This is transformative in ways that we haven't yet begun to imagine.”

### J. Question 10: Can you provide an example of how The Library.Link Network initiative has improved your library's visibility?

Answers to this question were primarily in the negative; many responders stated it was too soon to tell, that they lacked the analytics necessary to provide an example, or that visibility was not their main purpose for experimenting with linked data. These respondents cited a broader impetus for their involvement in the Library.Link Network, one stated, “I don't think that ‘visibility’ is the point. While it's a helpful side-effect, the real potential is for library data to enter an information universe in new ways.”

## *Embracing Google and Linked Data in Law Libraries*

Notwithstanding the negative replies, many libraries provided examples of linked data working within their libraries. Most cited examples such as: “If you know to do it, you can search for an author or a title combined with our name and get the link to our catalog.” Some were able to provide concrete examples of linked data leading a patron to their catalog: “Sure! This past weekend, a patron was able to find a nonfiction title we own through a Google search. He reserved that title through our catalog and is waiting for it to become available” and “One example we had early on was the image of Norman Lenfest. This is part of our historical image collection. After we first published, I saw the link network data appear for this title near the top of the search engine results. After a couple of weeks checking this title, the search engine figured out the right place to go was our catalog and that became the top result and the linked data entry fell down in ranking.” Finally, one library stated that in January, 2016, they saw a dramatic increase of visits to their catalog from out of state searchers. Where they would normally expect about 15% of visitors to come from out of state, this number jumped to 75% that month. Holding all other factors constant, this jump in out of state researchers can legitimately be attributed to the library’s use of linked data.

### *Part IV: What Do Libraries and Linked Data Have in Common?*

Libraries and linked data have a common goal: increased visibility of information. Where they diverge is their methods of evaluation of the information provided. Linked data is based on algorithmic functions between indexed data and quantity of linkages. In comparison, libraries spend vast amounts of time evaluating the quality of information: provenance, reliability, and accountability. In the digital age, neither approach is adequate to fulfill their shared mission: information dissemination. Libraries that have not chosen to convert their catalog holdings to linked data are only accessible by patrons who are aware of them, know where they are located, and have the sophistication to navigate a library catalog. Although many Integrated Library Systems now feature a Google-like search box, they do not function like Google and patrons expect the same experience they receive from Google. Publishing catalog holdings to the web allows less sophisticated users to use Google for their research and still have access to the materials available in their local library, be it public or academic.

The libraries surveyed for this article were mostly public in nature, but academic libraries are launching linked data initiatives as well. The most visible of these initiatives is Linked Data for Libraries (“LD4L”).<sup>91</sup> This project is a multiphase alliance between Cornell, Harvard, and Stanford spanning four years of development.<sup>92</sup> They acquired other participating libraries since the project launched in 2014, including Columbia, Princeton, the University of Iowa, and the Library of Congress.<sup>93</sup> Their mission “is to create a Scholarly Resource Semantic Information Store (SRSIS) model that works both within individual institutions and through a coordinated, extensible network of Linked Open Data to capture the intellectual value that librarians and other domain experts and scholars add to information resources when they describe, annotate, organize, select, and use those resources, together with the social value evident from patterns of usage.”<sup>94</sup> LD4L proposed to meet this mission by producing three deliverables: SRSIS ontology;

---

<sup>91</sup> See <https://wiki.duraspace.org/pages/viewpage.action?pageId=41354028>.

<sup>92</sup> *Id.*

<sup>93</sup> *Id.* and *supra* note 91.

<sup>94</sup> See <https://wiki.duraspace.org/display/ld4lGW/Linked+Data+for+Libraries+%28LD4L%29+Gateway>.

SRSIS Semantic editing, display, and discovery system; and Project Hydra compatible interface to SRSIS.<sup>95</sup> Now in the final phase of the project, LD4L has available on its website materials for any library exploring converting their legacy MARC records, LibGuides, and other resources into linked data.<sup>96</sup> These materials include access to a general-purpose extension to BIBFRAME model, domain-specific ontologies and VitroLib, a linked data instance editor.<sup>97</sup>

LD4L is the largest scale linked data project involving academic institutions, but several other schools have launched university specific initiatives. These include George Washington University, who experimented with adding URIs to the MARC subfield zero<sup>98</sup>; the University of Nevada, Las Vegas, whose linked data project studied “the feasibility of developing a common process that would allow the conversion of our collection records into linked data preserving their original expressivity and richness” and experimented with publishing data from their collection the web;<sup>99</sup> and the University of Illinois’ project, Linked Open Data for Special Collections, recently resulted in the team presenting a poster at the 2016 Annual Meeting of the Association for Information Science and Technology (“ASIS&T”) in Copenhagen.<sup>100</sup>

While these are examples of academic libraries experimenting with linked data, there is little evidence of law libraries following suit. Implementing linked data in law libraries requires two communities traditionally resistant to change to accept a technology they may not understand, and whose success is still considered theoretical even by some current participants.<sup>101</sup> Law schools have taught by the “case method” for nearly 150 years, despite changes in society, education, and technology.<sup>102</sup> Downturns in the economy, falling job prospects, and falling enrollments are finally spurring law schools to change their model of scholarship, but these changes are coming at a characteristically slow rate.<sup>103</sup> As law schools begin to reevaluate how to produce the next generation of attorneys, the profession itself is

---

<sup>95</sup> See <https://wiki.duraspace.org/display/ld4l/Expected+Outcomes>.

<sup>96</sup> See <https://wiki.duraspace.org/pages/viewpage.action?pageId=74515029>.

<sup>97</sup> *Id.*

<sup>98</sup> Jackie Shieh & Terry Reese, *The Importance of Identifiers in the New Web Environment and Using the Uniform Resource Identifier (URI) in Subfield Zero (\$0): A Small Step That Is Actually a Big Step*, 15 *Journal of Library Metadata*, Issue 3-4, pp. 208-226, (2015).

<sup>99</sup> See <https://www.library.unlv.edu/linked-data/2014/06/learn-more-about-unlv-linked-data-project-ala-las-vegas-2014.html>

<sup>100</sup> See <http://publish.illinois.edu/linkedspcollections/2016/11/08/project-team-presents-at-asist-2016-in-copenhagen/>.

<sup>101</sup> Survey respondents often expressed frustration regarding the lack of tangible results. Comments included: “Right now, I can't provide any evidence that this is a cost effective program. While I believe the potential is there, it's going to be difficult to convince my administrators to stay the course;” “We still aren't sure if it is worth the cost; our items aren't more visible;” and “Implementation requires a lot of patience. If you bake a cake, you mix the ingredients and short while later it comes out of the oven and you can eat it. Discovery online takes several months to the better part of a year before institution / library data begins to rise in relevance and becomes discoverable in any meaningful way, based on our experience. This is such a new endeavor that there are bound to be hiccups and learning curve. Geo-spacial location of library assets is a key goal of the discovery process. Despite geo-tagging of our data assets, it is often necessary to include the name of the library in the user's search to produce a positive hit. This will eventually be addressed but currently is a hurdle to reaching the most desirable demographic to us: the user that wasn't thinking of us when they began his/her search.”

<sup>102</sup> Steven C. Bennett, *When will Law School Change?*, 89 *Nebraska Law Review*, Issue 1, pp. 87-130, (2010).

<sup>103</sup> David Segal, *What They Don't Teach Law Students: Lawyering*, *The New York Times*, Nov. 19, 2011. Accessed online: May 23, 2017. <http://www.nytimes.com/2011/11/20/business/after-law-school-associates-learn-to-be-lawyers.html>; Hannah Hayes, *Recession Places Law School Reform in the Eye of the Storm*, 18 *Perspectives*, No. 4, pp. 8-9, 14, (2010).

witnessing its own revolution in the from the open access movement.<sup>104</sup> As primary and secondary sources become readily available from free databases, the need for law libraries lessens even more. No longer will law libraries hold the keys to the kingdom of legal knowledge. In addition to an aversion to change, there have been relatively few attempts at creating legal ontologies necessary for the unique nature of the law.<sup>105</sup> Without the vocabulary required for data to speak to each other, we are left with only grammar. Despite movements toward legal ontologies and open access to legal materials, the industry is slow to change. Legal librarians will no doubt maintain their hold to the kingdom for the foreseeable future.

In contrast, public and academic libraries face a more immediate threat. These libraries, long functioning as the only proprietor of knowledge in a community, “were inclined to focus more on preserving the past than on inventing the future.”<sup>106</sup> The prevalence of Google and the explosion of information accessible from the palm of the hand has forced libraries to re-evaluate this strategy.<sup>107</sup> Change has been slow coming, often taking the shape of a search box resembling Google without the accompanying immediacy of answers; however, many libraries are considering new technologies and practices to reach their patron base effectively.<sup>108</sup> The development of library specific technologies and the influx of a new generation of “digital natives” in the librarian profession has accelerated the acceptance of new technologies in public and academic libraries.<sup>109</sup> This younger generation of librarians are more likely to: consider new technologies; believe in their ability to master their functions; and incorporate them into their workflows and policies.<sup>110</sup>

The changes in these two arenas of knowledge are reflective of changing economic landscape, but also a refocusing on production of the most equipped law school graduates possible. Law firms are demanding graduates with a trove of practical skills already in their quiver.<sup>111</sup> Law schools are slowly beginning to respond to this demand. It is time for law libraries to follow their lead and explore the best methods for connecting law students with the library’s resources. Despite advances in the appeal of interfaces and the addition of a myriad of facets, students are not finding improved Online Public Access Catalogues (“OPAC”s) as intuitive and easy to navigate as librarians and vendors predicted.<sup>112</sup> If a technology is difficult to navigate or failing to produce meaningful results, users will revert to their former method of research.<sup>113</sup> Linked data will bridge the gap between the old habits and the new. It makes

---

<sup>104</sup> Much can and has been written about the Open Access Movement and its application to the legal realm. It is too broad a topic to cover in this paper, but See Michael W. Carroll, *The Movement for Open Access Law - Symposium*, 10 Lewis & Clark Law Review, no.4, pp. 741-760, (2006), for more on the topic.

<sup>105</sup> One example of a legal ontology in development can be found at: <http://www.estrellaproject.org/>. This project occurred in Europe, resulting in 6 workpackages of deliverables. However, they have not published an activity report since 2008.

<sup>106</sup> Sharon Gray Weiner, *Resistance to Change in Libraries: Application of Communication Theories*, 3 Libraries and the Academy, Issue 1, pp. 69-78, (2003). Accessed online May 23, 2017: <https://muse.jhu.edu/article/38574>.

<sup>107</sup> *Id.*

<sup>108</sup> See Hollie M. Osborne & Andrew Cox, *An investigation into the perceptions of academic librarians and students towards next-generation OPACs and their features*, 49 Program, Issue 1, pp.23-45, (2015).

<sup>109</sup> See Jenny Emanuel, *Digital Native Librarians, Technology Skills, and Their Relationship with Technology*, Information Technology and Libraries, pp. 20-33, (Sep.2013).

<sup>110</sup> *Id.*

<sup>111</sup> Ed Finkel, *Training a New Breed of Lawyers*, Before the Bar, ABA For Law Students (Nov. 1, 2014), <http://abaforlawstudents.com/2014/11/01/training-new-breed-lawyers/>.

<sup>112</sup> *Supra* note 108 at 41.

<sup>113</sup> See generally <http://learningcenter.unc.edu/handouts/changing-habits/>.

## *Embracing Google and Linked Data in Law Libraries*

catalog information available through the method of research most Americans begin with: Google. Finding the resource through the search engine is only the beginning. Students must then venture into the library to retrieve the resource. Once a law student is in the physical space of the library, the library staff has the opportunity to provide an essential service available to law students: formal and informal instruction.<sup>114</sup> Law librarians maintain and organize resources, but they also provide assistance with formulating search queries, introducing new databases, and fostering the development of critical thinking.<sup>115</sup> These vital services remain undiscovered for those students who rarely enter the library, such as those who instead confine themselves to their casebooks and electronic databases they have perfunctory knowledge of manipulating. The unfortunate result of these services remaining undiscovered is a waste of resources and opportunity to produce a law graduate fully armed with the tools needed to succeed in his/her career.

### *Part V: Why Adopt this Technology?*

The Semantic Web is not without its drawbacks. Ontologies, specifically legal ontologies are slowly developing. BIBFRAME has been in development for over five years and the Semantic Web “has no concept of quality control.”<sup>116</sup> Open data is precisely that, open on the web waiting for any triple to form. Conversion can also be expensive, sometimes prohibitively so.<sup>117</sup> Ironically, slow adoption of a technology prevents economies of scale allowing for a price drop, resulting in few libraries adopting the technology. This phenomenon is especially cruel when the technology is dependent on mass participation to succeed.

And succeed it must. Libraries are not merely warehouses of information, a brick and mortar internet. Libraries are much more; they are the depositories of past events, current issues, and tomorrow’s ideas, a temple of curation. “Curation is often confused with preservation, but there is much more to it than this alone. Curation involves identifying those elements of a culture that particularly define it and choosing which ones are important; it then describes and adds context to these, making connections between them, so that they can be understood by all those who have an interest in them. Finally, it involves disseminating a culture, making it accessible. All of these are in addition to ensuring that these elements will continue to exist for a long time in the future. Going through these steps ensures above all that a culture can be understood when it is transmitted between generations. It is thanks to the curatorial efforts of our forebears that any culture beyond the most ephemeral has any existence at all.”<sup>118</sup> We must continue the mission of our forebears and continue to curate information. Partnering with Google Knowledge Graph and other semantic search engines to produce “quality information resources appropriate for the local

---

<sup>114</sup> Ruth Levor, *The Unique Role of Academic Law Libraries*, Toolkit for Academic Law Libraries, Academic Law Libraries Special Interest Section, American Association of Law Libraries. Accessed online: May 23, 2017, <http://www.aallnet.org/sections/all/storage/marketingtoolkit2004/unique-role.pdf>.

<sup>115</sup> *Id.*

<sup>116</sup> Coyle, *supra* note 58 at 140.

<sup>117</sup> “The technology of data management evolves less quickly than some other computer technologies because it often requires a rather costly conversion of data and interfaces to take advantage of new capabilities. For this reason, few can afford to jump on a new technology bandwagon when it first appears. Instead, most wait until such a technology has matured before adoption.” *Id.* at 137.

<sup>118</sup> Gartner, *supra* note 28 at 12.

## *Embracing Google and Linked Data in Law Libraries*

researcher”<sup>119</sup> benefits our patrons and our industry. Embracing the newest technology to disseminate information is imperative. Shunning the printing press because it had no proven results would have been disastrous for the modern world.<sup>120</sup> Linked data can make curated information available to greater segment of the population, resulting in a more educated populace.

Linked data and the Semantic Web can usher in a renaissance in librarianship. The greater visibility of library resources can begin a process by which patrons formerly unaccustomed to seeking information from the library begin to incorporate the it into their normal routine of information gathering. This can only be accomplished by mass participation of libraries of all types. Fifty years have passed since the MARC format changed librarianship, but the words of Henriette Avram still ring true: “...We must work together with the academic/research network under development to build one logical academic/research/library network for the nation.”<sup>121</sup> The time has come for computer and information science to renew their vows, a marriage born of necessity can bloom anew in the modern age.

---

<sup>119</sup> Jina Choi Wakimoto, *Scope of the Library Catalog in Times of Transition*, 47 *Cataloging & Classification Quarterly*, Issue 5, pp. 409-426, at 424 (2009).

<sup>120</sup> See <http://www.historyguide.org/intellect/press.html>.

<sup>121</sup> Rather, *supra* note 29 at 860.



Appendix: Survey

- 1) What content/services did you decide to include in your transform into linked data?
  - a) catalog
  - b) events
  - c) special collections
  - d) archives
  - e) electronic resources

Please share why you chose these resources:

- 2) Did you see increased traffic to your website after publishing your data to the Web?
  - a) Yes
  - b) No

Approximately, how much was the increase:

- 3) Was there an increase in the number of patrons visiting the library after your data was published, compared to numbers regularly seen during that time period?
  - a) Yes
  - b) No

Approximately, how much was the increase:

- 4) Was there an increase in new patron registrations after your data was published, compared to numbers regularly seen during that time period?
  - a) Yes
  - b) No

Approximately, how much was the increase:

- 5) If you chose to include events in your linked data conversion, did you see increased attendance at library events following publishing your data to the web?
  - a) Yes
  - b) No

Approximately, how much was the increase:

- 6) If you chose to include special collections in your linked data conversion, did you see increased requests to access these collections, compared to numbers regularly seen during this time period?
  - a) Yes
  - b) No

Approximately, how much was the increase:

*Embracing Google and Linked Data in Law Libraries*

7) A main mission of The Library.Link Network is to improve web visibility of library data, do you feel your library is more visible now that your data is published online?

Why or Why Not?

8) What was the biggest challenge you faced during this process?

9) What advice would you give to a library considering publishing their data to the web?

10) Can you provide an example of how The Library.Link Network initiative has improved your library's visibility?