

## **Exploring methods for linked data model evaluation in practice<sup>1</sup>**

by:

Robin Elizabeth Desmeules

Cataloguing Librarian, McGill University (corresponding author)

<https://orcid.org/0000-0002-5801-518X>

Clara Turp

Discovery Systems Librarian, McGill University

<https://orcid.org/0000-0002-4472-7549>

Andrew Senior

Coordinator, eResources and Serials, McGill University

<http://orcid.org/0000-0003-3917-6591>

---

<sup>1</sup> The authors would like to thank Eka Grguric and Jessica Lange for their feedback on the survey tool, and thank the anonymous reviewers and editors at *The Journal of Library Metadata* for their thoughtful comments and suggestions on earlier drafts of this article.

## **Abstract**

Ontology development and data modeling are core components of any linked data project. Through our own experiments building a linked data ontology for our collections, we wondered: how are our peers in the linked data community evaluating their ontologies? Are participants engaging in ontology evaluation? What methodologies and evaluation criteria are they using? Are they documenting and sharing their processes? In this paper, we present findings from a survey conducted in the fall of 2018, aimed at professionals from libraries, archives, and museums (LAM) who were part of the data modeling team on linked data projects. The purpose of this survey was to better understand the reality of ontology evaluation in the context of a linked data project. We found that our colleagues were engaging in data modeling as part of linked data projects in a variety of different tasks and roles. There was some ambiguity with respect to evaluation, possibly in part due to the iterative nature of the modeling process. Evaluation is engaged iteratively and informally through use cases, competency questions, and testing of the data in the application. On the whole, not being shared widely outside of a project. The identified barriers to evaluating their models included: lack of knowledge, resources, and documentation.

**Keywords:** linked data, data modeling, ontology, evaluation, assessment

## Introduction

Ontology development and data modeling are core components of any linked data project. Through our own experiments building a linked data ontology for our collections, we wondered: how are our peers in the linked data community evaluating their ontologies? Are participants engaging in ontology evaluation? What methodologies and evaluation criteria are they using? Are they documenting and sharing their processes?

In this paper, we present findings from a survey conducted in the fall of 2018, aimed at professionals from libraries, archives, and museums (LAM) who were part of the data modeling team on linked data projects. The purpose of this survey was to better understand the reality of ontology evaluation in the context of a linked data project. While there are theoretical frameworks and possible methodologies concerning ontology evaluations, little detail exists in the literature about which methodologies are preferred by practitioners or even if they actively engage in ontology evaluation processes. Moreover, as institutions continue to implement linked data, we hope that sharing this scan with our peers will bolster the existing literature. Our study centres on three main research questions: 1) Are participants engaging in ontology evaluation in the context of linked data projects? ; 2) Which evaluation methodologies did they choose and what criteria formed the basis of their selection? ; 3) Are participants documenting and sharing their evaluation processes?

There are two primary benefits of this research: first, the study will inform the community about how evaluation forms a part of practical implementation of linked data models; second, the results of this research will highlight existing evaluative procedures which will potentially help professionals evaluate the ontologies used in their linked data projects.

## Literature Review

The only consensus about the definition of “ontology” is that there is no singular definition, (Noy & McGuinness, 2001) and this lack of definitional clarity holds true for linked data, particularly since the term “ontology” may also often be used interchangeably with “vocabulary” or “data model.” In the following section, we offer some common definitions of a linked data ontology, and provide a context for how the foundational principles of the semantic web affect the structure and shape of linked data ontologies.

According to the World Wide Web Consortium’s (W3C) Linked Data Glossary, an ontology is “A formal model that allows knowledge to be represented for a specific domain.” (<https://www.w3.org/TR/ld-glossary/#ontology>) Noy and McGuinness, in their “Ontology Building 101” guide offers a similar definition, stating an ontology is “a formal explicit description of concepts in a domain of discourse.” (Noy & McGuinness 2001, 3). As previously mentioned, the fundamental principles of the Semantic Web influence the structure -- and, therefore, definition -- of a linked data ontology. (Pattueli, Provo & Thorsen 2015, 266). In the *Semantic Web for the Working Ontologist*, the authors explain how semantic models and standards are built with the express purpose to facilitate reuse, common understanding, and collaborative gathering of knowledge: “The Semantic Web standards have been created not only as a medium in which people can collaborate on models. Models that they can use to organize the information that they share. Models that they can use to advance the common collection of knowledge.” (Allemang & Hendler, 2011, 14) These standards, developed by the W3C, call for ontologies to be reusable, and built on open standards, like RDF. These principles have a direct influence on a linked data ontology’s structure. Pattueli, Provo and Thorsen, in their article “Ontology Building for Linked Open Data: A Pragmatic Perspective,” note that “...the semi-

formal nature of LOD ontologies has been key to making them suitable for widespread adoption, scale, and information integration in an open world, proving the well-known mantra by Jim Hendler, one of the Semantic Web's founding fathers: "a little semantics goes a long way." (Pattueli, Provo, & Thorsen, 2015, 270; Heath & Bizer 2011) They continue to explain how this lower degree of formality and use of open standards (particularly via RDF) not only makes it easier to reuse vocabularies, but also makes it much easier to extend ontologies "by mapping and mixing terms from different vocabulary sources...LOD ontologies are agile and adaptive, more than prescriptive, allowing for multiple views and multiple interpretations of the data to coexist." (268) This ties into the AAA principle of linked data modeling on the semantic web identified in the introduction to *The Semantic Web for the Working Ontologist*: "Anyone can say Anything about Any topic." (Allemang & Hendler, 2011) Albeit, this semi-formal structure of linked data ontologies often lends itself to being used synonymously with vocabulary (Pattueli, Provo & Thorsen 2015, 266). The authors of the *Semantic Web for the Working Ontologist* equate ontology with the term "semantic models," (1) offering a definition similar to the "ontology" defined by those of W3C and Noy & McGuinness, where a model as an abstraction used to help people communicate, explain and make predictions about the world, and that models mediate among multiple viewpoints (13). In "The Linked Data Cookbook," a W3C Note with a blow-by-blow methodology for publishing government data as linked data (including how to model the data), they never mention "ontology," offering instead a section on how to "Model the data," and using "vocabulary" throughout, including offering key vocabularies to investigate for reuse in the data modeling process (Linked Data Cookbook).

For the purposes of this paper, regardless of whether it is called an "ontology," "vocabulary," or "data model," we understand it as an abstraction that describes and facilitates

representation of a particular domain and/or worldview and/or environment. In the specific context of linked data, this abstraction must be realized using open standards and also, whenever possible, reuse existing vocabularies/ontologies. This structure in turn shapes the methods through which it may be assessed, and these are discussed in the following section.

### **What is data modeling for the semantic web? And how is it assessed?**

Literature on linked data projects in libraries, archives, and museums (LAM) is growing as more institutions experiment and implement linked data, and share their results with the greater community. We can glean some trends about what constitutes the ontology modeling process from existing documentation and case studies from LAM institutions. First, that building an ontology is an iterative, experimental process. (Noy & McGuinness, 2001; Pattuelli, Provo & Thorsen 2015; Allemang & Hendler, 2011; W3C) With testing of some sort at the core of this process, an iterative methodology in many ways keeps assessment at its very heart, but not necessarily in an explicit, formal sense. In other words, ontology builders are encouraged to check in and refine their ontology as they develop it, but not necessarily in a specific way. To further muddy the waters, there is no singular method for building an ontology (Noy & McGuinness, 2001), which makes shared case studies invaluable and highly desired. (Pattuelli, Provo & Thorsen, 2015)

But what kinds of methods are suggested for testing throughout the building process? A key way is to involve subject experts in the modeling process, in order to rely on their domain expertise during construction (Noy & McGuinness 2001; Pattuelli, Provo & Thorsen 2015; Linked Data Cookbook; Allemang & Hendler, 2011). For example, the Linked Data Cookbook suggests to “test the assumptions in the schema with subject matter experts familiar with the data.” Noy & McGuinness, building on the methodology of Gruninger and Fox (1995), used

competency questions as a “litmus test” for their building process: “does the ontology contain enough information to answer these types of questions? Do the answers require a particular level of detail or representation of a particular area? These competency questions are just a sketch and do not need to be exhaustive” (5). Competency questions are also advocated by Allemang & Handler in their chapter “Good and Bad Modeling Practices.” (2011) Another identified method of both building and assessment involves developing use cases, which is the strategy adopted by the various groups involved with the Linked Data for Production (LD4P) projects creating extensions to BIBFRAME. ([Linked Data for Production \(LD4P\)](#) ; see their individual groups for publications and presentations detailing their methods). In all of these contexts, competency questions and use cases serve two functions: to help in the development of the ontology, as well as to serve as a kind of benchmark to test against.

A final identified way of assessing an ontology involves testing the data itself throughout the modeling process. This could take the form of checking against use cases and competency questions, and user testing of the data in the application. Again, this kind of testing seems to fall under the notion of iterative, experimental design, but without an explicit methodology or formalized assessment process. For example, assessing the quality of the data could also entail some larger scale linked data quality measures like those mentioned in Talleras (2017) and Zaveri et al, (2015). However, this is complicated by the fact that data quality may also be affected by the quality of the original data before transformation, or the effect of the interface itself on accessing the data. Many of the key benchmarks of linked data quality may have more to do with the technical infrastructure than the modeling process itself.

Several studies in software engineering point to these and other criteria and methodologies of ontology assessment (Brank, Grobelnik and Mladeni, 2005; Degbelo, 2017;

Gomez-Perez, 2001; Orbst, Ceusters, Mani, Ray and Smith, 2007; Raad & Cruz, 2015), but much of these discussions have not trickled over into the LAM literature. This could be for several reasons: the lower degree of formality of linked data ontologies might make some tools and methods non-applicable, the semantic web might make them non-applicable, the “craft-like”, experimental nature of so many linked data projects might make this kind of formalized assessment impossible (no time or energy), and possibly these processes seem unnecessary because of assumptions about success in production being the ultimate assessment.

The notion of success for a project overall, as well as challenges encountered is discussed in the literature for linked data, through the multi-year OCLC Research International Linked Data Survey for Implementers (Smith-Yoshimura, 2018; Smith-Yoshimura, 2017). Though a rich source of information on linked data projects from around the world, from over a hundred institutions, it does not explicitly address assessment methods in ontology development, focusing instead on holistic technical, production, and other concerns. Nonetheless, several findings are of note for the purposes of this project. What we do learn from the data from this survey, ran in 2014, 2015, and 2018, is a consensus that a key barrier to the success of a linked data project in production are: selecting appropriate ontologies, steep learning curve, and little documentation or advice on how to build systems. (Smith-Yoshimura, 2018). This paper seeks to check in about these challenges, but from the data modeling perspective. We suspect to have similar findings with respect to modeling challenges, particularly given the identified need for more case studies in linked data identified by Pattuelli, Provo & Thorsen: “Building ontologies in the time of linked data has a craft-like component to it, as it requires a more relaxed use of semantics than previously required, relying on still scarce methodological guidelines. LOD development, especially in the library, archive, and museum (LAM) community, involves a good deal of trial



and error, so sharing lessons learned is needed and valued.” (2015, 266). Moreover, it is hoped that the results presented in this article can help remedy the dearth of detailed documentation for future projects, and therefore answering the call above to share lessons learned.

The purpose of this study is to explore assessment methods in ontology development in LAM explicitly. Given the diversity of approaches and methods of building an ontology, and the highly experimental nature of linked data projects, we wanted to know what precisely our colleagues were doing, and the challenges they faced along the way. We also hope our findings will help bolster the work of our colleagues as they build and implement linked data projects.

## **Methodology**

We chose to build a survey for our study in order to get the widest possible reach with a minimal time commitment by the participant. The further advantage of a survey was the ease of dissemination and recruitment. Although we also considered conducting semi-structured interviews, we opted to reserve this more in-depth methodology for a future phase of the project. We developed a 12 question survey comprised of both quantitative and qualitative questions, using the institutional instance of LimeSurvey (see Appendix 1). The question types include: free text, multiple-choice, and yes/no, and only the first question about data modeling experience was mandatory. We built and tested the survey throughout the summer and fall of 2018.

We obtained approval from our Research Ethics Board for our project, and distributed the survey electronically through several key listservs for technical services workers in libraries, archives, and museums in November, 2018, including: ACRL Technical Services Interest Group, ALCTS Metadata Interest Group, ASIS&T-L, Code4lib, DLF, LITA, Metadata Librarians, and the Canadian Technical Services Network. We were unable to obtain numbers of subscribers from all of the listservs, but based on the available information we estimate that the survey

reached approximately 5000 people. The survey was originally sent out for two weeks, then extended by two weeks to December 7th to increase the amount of full responses. Our target sample size was relatively small (10), and we received 18 responses from people who had participated in data modeling for a linked data project. We performed some descriptive statistical analysis for the multiple choice questions using Microsoft Excel and the available online tools included in LimeSurvey. The free text questions were coded manually by reading through the responses for common themes and coding them throughout the answers in the entire survey.

## **Results**

We had a total of 54 respondents, with 40 completed responses. Of those 40 completed responses, 18 identified participating in the data modeling of a linked data project. The other 22 respondents were guided to the exit screen. Of the 18 people with modeling experience, most of them offered comments when requested. The 14 incomplete responses were not included in our analysis, as per our consent statement.

### **Experience**

The first section of the survey explored data modeling experience with linked data projects. People reported that they were a part of anywhere from one project (seven respondents) to 12 projects (one respondent), with an average 3.47 projects per person, and a median of two projects per person. Of the remainder of the participants, three respondents participated in two projects, one in three projects, two in four projects, one respondent in five, one respondent in seven, and one in 11 projects. (See Figure 1)

We then asked them to list their roles and tasks for each project, and 16 respondents expanded on their experiences. We identified the following overarching themes in their responses, grouped around major phases of a linked data project (note that respondents often

listed multiple tasks and roles): Data modeling/Ontology development, Linked Data Application and Infrastructure Development, Data Conversion and Transformation, and Education (See Appendix 2 for respondents' answers grouped accordingly). Not all participants performed the same tasks: some focused on the data model, some worked with the data, and others developed the application -- although those with multiple projects under their belts often did a combination of many of these types of activities.

The majority of tasks reported were part of the modeling process, including but not limited to data modeling, vocabulary selection, identifying use cases, and editing existing ontologies. People who mentioned modeling tasks and roles also noted that they were often tasked with overseeing the modeling process. The second largest category of reported tasks involved data conversion and transformation, where people reported converting data into a new schema (e.g., XML to RDF), conversion analysis, preparing datasets for this transformation, as well as reconciliation and name authority work. There were also a select number of respondents who reported building applications and infrastructure in order to run linked data in production. This ranged from building a proof of concept system in Fedora, to implementing linked data in a particular system, and also building or overseeing infrastructure (web) or tools to support these applications. As part of the experimental nature of many of these projects, a final theme involved learning through experimentation, which included projects that were listed as "proof of concept" or "experimental" by the respondents, with one also reporting leading a study group on linked data. Some of these experimental projects were not completed, or were on the verge of completion.

Respondents also reported working a wide range of capacities, including as consultants, programmers, and project leads, working in collaborative environments both within their

respective institutions, as part of larger organizations, as well as part of international partnerships.

## **Ontology Development**

The second general section of the survey concerned the process used by respondents to develop their data model or ontology. Although reusing existing ontologies whenever possible is a core tenet of linked data best practices (W3C; Linked Data Cookbook; Noy & McGuinness; Pattuelli, Provo & Thorson), it is not always the best fit for a particular dataset. For this reason, we asked for information on how they used (or didn't) existing ontologies in their models. Six participants (33.33%) reported that they built a new ontology, whereas 13 (72.22%) adapted and reused multiple existing ontologies, and four respondents (22.22%) adapted and reused a single ontology. A number of participants reported using a combination of strategies (see Table 1). There were a further five respondents who opted for "Other," with four of them noting that they extended an existing ontology by adding "local bespoke properties" to an existing ontology, or added local properties/extended the ontology (See Figure 2).

The sources of information used for ontology selection is also of interest, particularly given the importance of reuse, and the sheer number of existing ontologies. Many respondents chose multiple sources of information for ontologies. Ten respondents selected "Consulted Colleagues" (55.56%), 12 selected "Consulted literature" (66.67%), 10 selected "Other" (55.56%), and no respondent chose 'Social Media' as a place to learn about available ontologies (Figure 3). Five participants chose "Consulted Colleagues," "Consulted Literature," and "Other." Five participants selected "Consulted Colleagues" and "Consulted Literature". Four participants only chose "Other," and stated that they relied on personal expertise in the field, Library of Congress resources, Europeana expertise, or web searching. The "Other" option provided

additional sources of information for ontologies, such as: online resources from standards communities such as W3C, Europeana and Library of Congress; team investigation; ontology “Catalogues” (Linked Open Vocabularies (LOV) and BioPortal) and documentation.

## **Evaluation**

In the final section of the survey, participants were asked about evaluating their ontologies. Of interest to us were their evaluation methods and strategies, as well as any challenges or barriers encountered along the way. Of the 18 respondents, 11 (61.11%) evaluated their model, five (27.78%) did not evaluate it, and two (11.11%) didn’t answer.

Most respondents approached their evaluation process broadly, as part of the project as a whole using real-life workflows and principles of iterative design. The comments concerning their ontology and vocabulary choice indicated they used mostly existing ontologies and models that were widely adopted. Participants also judged their evaluation process as being informal. Of the 11 participants who evaluated their data model, ten (90.91% of respondents to this question) provided more detail about their evaluation methods. Responses varied: seven comments noted or implied their evaluation happened iteratively throughout the project; nearly all respondents who evaluated (nine people) mentioned that testing happened practically, by reviewing or comparing the data (four people), with some making note that it was in the application (six people). Subject experts (one person), use cases (two people), competency questions (one person), and user testing (two people) were also listed by various respondents, but not systematically or in concert with one another (See Table 2).

## **Challenges and Barriers**

We asked all survey participants if they encountered challenges and barriers to evaluating their work. They were also asked to describe these challenges, as well as list what they thought may have helped them evaluate their project.

A total of ten people identified challenges to evaluation, regardless of whether they had conducted an evaluation or not. Of the five people who did not evaluate their model, three (60%) said they had encountered challenges, whereas seven (63.64%) of the participants who evaluated their model said they had encountered challenges or barriers to evaluation. Both groups identified similar challenges to evaluation: lack of skills/knowledge gaps, lack of resources (staff, time, hosted space), lack of actual completed examples of projects in production to draw from, and the difficulties in bringing a project into production. One respondent explains: “mostly it feels like these projects have not achieved a critical mass of use or analysis to do any evaluation. One project is ongoing so we're not at the evaluation stage, another project feels like everyone has just moved on and is not interested in reflecting back on it.” Moreover, some participants felt as if evaluation was best done on the final overall project to reflect back and improve, which is implied in this comment: “One project is ongoing so we’re not at the evaluation stage.” Clearer documentation, including examples and application instructions, would have helped respondents evaluate, as one respondent explains: “a hosted sandbox with documentation for linked data projects, like an online demonstration exercise, would have provided an opportunity to learn about evaluation as part of the exercise.”

### Sharing evaluation process

Finally, respondents were asked if and how they shared their evaluation methodology (Figure 4). Of the 11 participants who evaluated their data model, four (36.36%) said they shared their

evaluation process and three provided details on their methods: in a paper (three); in a presentation (three); on Twitter (two); on GitHub (one); on blogs (two); and other (two). Seven (63.64%) respondents indicated they had not shared their evaluation. Two of the participants shared their evaluation process through multiple sources, such as paper, presentation, Twitter, GitHub, blog, and “meet-ups, symposia and workshops,” and one participant shared through their working group as part of a large-scale pilot project.

## **Discussion**

### **Project Experience**

Respondents reported a wide range of project experience, both in the number of projects in which they were involved (over half, 58.8%, were involved in only 1 or 2 projects) and the varying roles they undertook within those projects. Given that any single linked data project requires significant investment of resources, these results are not surprising, but it is important to ask whether this difference in experience level led to differences in the methods of ontology selection and evaluation of data models. While we could find no common theme between how respondents developed their model and their experience, surveyed practitioners with greater experience (either explicitly stated or inferred through a greater number of linked data projects) appear more likely to engage in some form of evaluation: Of the respondents who provided information, all who had taken part in more than 3 projects indicated they had evaluated their data model. The implication of this is that the greater a practitioner’s experience in linked data projects, the greater the likelihood that they will engage in some form of evaluation to strengthen the final product. It also suggests that as experience of linked data projects increases within our communities of practice, there will be a greater integration of evaluation into the design process.

### **Ontology Creation/Selection**

Respondents tended to follow best practices recommending the reuse of existing ontologies. Even with those who said they created a new ontology, all but one also adapted and reused a single ontology or multiple existing ontologies. In our results, respondents indicated their reuse of ontologies, but also that they adapted them to their own realities (combining with other ontologies, and adding local properties).

The results do highlight a certain lack of clarity around the difference between selecting and extending ontologies, versus creating a wholly new ontology. Arguably, this is a matter of perception. For example, at which point does reusing other ontologies or component properties result in creating a new ontology? Does adding local or “bespoke” properties equal adaption and reuse? One person considers several extensions as not being an inherently different ontology, whereas another person considers extensions to warrant it being a new ontology in and of itself. This aspect wasn’t completely clear for the authors, as we suspect was equally the case for the respondents. The way the survey question was presented may also have factored into this variation in responses.

### **Evaluation methodologies and criteria**

Responses to our survey suggest that ontology evaluation is taking place, although not uniformly, with variations in the degree, methods, and timing of the evaluation between respondents (Table 2). What is more, the reported evaluation methods align with recommendations outlined in the literature with some projects implementing iterative and experimental process of selecting and creating ontologies (Noy & McGuinness, 2001; Pattuelli, Provo & Thorsen 2015; Allemang & Hendler, 2011; W3C), following the guidelines and best practices of employing subject experts (Noy & McGuinness 2001; Pattuelli, Provo & Thorsen 2015; Linked Data Cookbook; Allemang & Hendler, 2011), using competency questions (Noy &



McGuinness 2001) and/or creating use cases (Allemang & Hendler, 2011). These results suggest that practitioners are informed and engaging in literature-outlined best practices to varying degrees.

Some respondents specified engaging in either formal or informal evaluations of their projects. Two indicated that their evaluation was informal. One respondent noted they thought our questions were more geared to formal evaluations but that theirs were informal and open “We are very involved in informal evaluation. We discuss issues regarding [...] the ontology with noted ontologists in the field. [...] Our informal processes are completely open [...] we use a wiki, Slack, social media Google Groups and other formats to receive and incorporate feedback.” Asking whether evaluation was formal versus informal in our question might have yielded more specific responses.

A major theme in our responses was the process of iterative assessment of their ontology. This kind of assessment may occur in a number of ways. Two examples of iterative processes used to evaluate involve comparing the data against the model or running the project into production. The verification of data following the initial modeling phase features highly in responses, and this makes sense for a semi-formal ontology as it is applied and practical. The applied approach involves direct interaction with the data, it doesn’t require a lot of time, and is consistent with the constraints and pressures identified by our respondents and elsewhere in the literature around linked data projects.

One respondent’s answers seemed to indicate they tested as they moved into production. Their response highlighted evaluation that was undertaken via an application-- in other words, testing how the data works in production resulting in revisions to the model: “Of course, instances arise during the course of a project or system implementation that were unanticipated

and caused the need for properties/values to be scrutinized, subtracted, revised, or added. These mostly came up as we moved forward practically as opposed to an official part of the data modelling process.” As mentioned earlier, hopefully, as more projects run into production and people gain more experience, an increasing number of projects will be evaluated. Since evaluation through usage is both a required step and an opportunity to test, more projects will naturally be evaluated through increased usage and iterative design practices. However, a potential challenge in evaluation in the later stages is ascertaining whether problems stem from the selected ontology(ies), the data model, or the application itself. The final product application may mask weaknesses in the data model or conversely, a poor application of the model might not mean the data model is at fault. Furthermore, an application might not meet user needs but still be a good application!

Some respondents highlighted other evaluation methods. One respondent listed the use of SPARQL queries and data transformations as a way of exposing gaps and understanding usability. They go on to say that “For ontology-driven software projects, end user testing if possible”. One respondent’s group also checked their model through translation, comparing how the data transformed using various tools and converting BIBFRAME metadata using SVDE. The double challenge faced in these approaches--and this was specifically mentioned in one answer--is ensuring there is both sufficient data to test against and that access to end users for testing is possible. This suggests that as part of evaluation processes, the availability of sufficient test data should be an early concern in a project.

A final theme in several responses is that of uncertainty around evaluation. While some respondents were quite clear they had engaged in evaluation of their model, some responses suggest respondents were still unsure if their work qualifies as evaluation, or sufficient

evaluation. This might have something to do with the iterative nature of linked data modeling and the extent to which it implies a kind of evaluation on-the-fly. For example, the legitimacy of evaluation through the application might be asserted as linked data projects running in production become more common; however, at present the dearth of case studies may hide the relevance of this testing method. Responses also revealed some ambiguity in our own question - what does evaluation mean in this context? We kept the question intentionally open in order to see if there was a specific way of understanding evaluation, and it did reveal different understandings of the term, as well as some uncertainty about whether they actually conducted an evaluation, as indicated in the following response: “I think I evaluated it ... These mostly came up as we moved forward practically as opposed to an official part of the data modeling process.”

### **Documenting and sharing evaluation processes**

Documenting and sharing linked data work, including steps taken and lessons learned, is a key theme in the literature on linked data projects (Pattueli, Provo & Thorsen 2015, 266). When asked what would help our respondents with their evaluations, they echoed this request to share, with one respondent proposing “a hosted sandbox with documentation,” where others also sought more documentation, with one specifying that it be “shared by the community with specific, easy to understand application instructions.”

In light of this identified need, our results indicate that evaluation processes are not being shared to the same extent that evaluation is taking place. There is a decreasing trend across responses from respondents’ involvement in projects (18 respondents), evaluation (11 respondents), and sharing the evaluation (4 respondents). This leads us to hypothesize that

evaluation is being engaged in informally, and on the whole, not being shared widely outside of a project.

There are several possible reasons for this. One possibility is that the barriers to evaluation—expressed in some responses as a lack of time, gaps in knowledge, or uncertainty in the process—may account for the lower number of shared evaluations. For example, practitioners with less modeling or project experience may be less confident in sharing their work and results. While it is plausible that the knowledge and confidence acquired through multiple projects would lead to a greater likelihood in sharing evaluation processes, this was not reflected in the experience of respondents who shared their evaluation, as the number of projects varied widely: from 1, 2, 4 to 12 projects. Additionally, if the project is still in progress, or was halted before running into production, sharing evaluation methods could feel premature, especially if exploratory.

Another possible reason that some respondents were not sharing could actually be the result of the lack of existing documentation outlining evaluation methods throughout a given project. Put simply, they may not realize it is something that *should* be shared. Moreover, given one of the primary challenges identified in the responses was time, respondents simply may have lacked the space and resources to document the evaluation stage. Finally, given the “craft-like” component of linked data ontology development, sharing evaluation methods might seem to be non-generalizable, and perhaps perceived as unhelpful.

Whether the results are generalizable, from a finished project, or even from an experienced team, there is still an identified need from the community to document and share the evaluation process of our ontologies as we progress through a linked data project. This call for greater documentation is identified in the literature, and echoed in our responses. Moreover, a

primary incentive for sharing evaluation methods is that it will help future linked data implementers and experimenters find their way. Lastly, sharing evaluation methods contributes to the scholarly discussion on developing best practices and guidelines for the community. (Pattueli, Miller, Lange, Fitzell, & Li-Madeo, 2013; Pattueli, Provo & Thorsen 2015)

## **Limitations**

The authors targeted the survey to professionals having worked with at least one linked data project. This resulted in the number of respondents progressing to the full survey being rather low compared to the number of overall initial respondents. Of the 40 completed responses, less than 50% (18) had been a part of modeling for a linked data project, and of the 20 respondents who indicated they had not been part of a linked data project, none left additional comments or feedback. This potentially missed respondents who are experimenting with ontology assessment outside of formal projects. It also meant that our sample size was ultimately small overall.

The authors acknowledge the survey design contained an inherent ambiguity around the phrasing of “Adapt and Reuse” for single or multiple existing ontologies, as respondents were left to decide where to indicate a response for an extension of an existing ontology with purpose-built properties, either under “Adapt and Reuse a Single Existing Ontology,” or under “Other.” While a high number of respondents indicated an adaptation and reuse of multiple ontologies, four respondents chose “Other” and indicated they had added new/local properties. It wasn’t possible to glean from these responses whether their authors had extended one or more ontologies for their project(s). There was a similar ambiguity around our use of the term “evaluation”, specifically around whether evaluation was intended to be “formal” or “informal”, and as such this might have influenced the way respondents answered. Nonetheless, our results

provide a valuable foundation for future research, as our findings still offer a glimpse and starting point for future work.

## **Conclusion**

In this paper, we presented the findings from a survey on evaluation methodologies for ontology/data modeling in the context of linked data projects. The purpose of this survey was to better understand the reality of ontology evaluation through the experience of our LAM colleagues. The results demonstrated that our colleagues focused on evaluating the linked data project as a whole, including the ontology. Evaluation through the use of the finalized application was recurrent throughout responses, and an important evaluation method in iterative design principles. Challenges such as the lack of documentation, lack of expertise as well as the experimental aspect of many of the linked data projects carried out by our colleagues may partly explain the fewer instances of evaluations being shared, and why 61.11% of the respondents said that they evaluated their ontology, with 36.36% of this group sharing their evaluation.

The aim of this survey was to get an overview of what LAM professionals were doing. While this overview could inform the community about a few different methodologies carried out during linked data projects, some possible next steps could be to identify gaps in the documentation and build tools that would help others better evaluate their linked data models.

## **Works Cited**

- Allemang, Dean, and Jim Hendler. 2011. "Good and Bad Modeling Practices." In *Semantic Web for the Working Ontologist*, 307–24. Elsevier. <https://doi.org/10.1016/B978-0-12-385965-5.10014-7>.
- Brank, Janez, Marko Grobelnik, and Dunja Mladeni. 2005. "A Survey of Ontology Evaluation Techniques," Proceedings of the conference on data mining and data warehouses (SiKDD 2005), 166-170  
<http://ailab.ijs.si/dunja/sikdd2005/papers/BrankEvaluationSiKDD2005.pdf>
- Degbelo, Auriol. 2017. A Snapshot of Ontology Evaluation Criteria and Strategies. In *Proceedings of the 13th International Conference on Semantic Systems (Semantics2017)*, Rinke Hoekstra, Catherine Faron-Zucker, Tassilo Pellegrini, and Victor de Boer (Eds.). ACM, New York, NY, USA, 1-8. DOI: <https://doi.org/10.1145/3132218.3132219>
- Gómez-Pérez, Asunción. 2001. "Evaluation of Ontologies." *International Journal of Intelligent Systems* 16 (3): 391–409. [https://doi.org/10.1002/1098-111X\(200103\)16:3<391::AID-INT1014>3.0.CO;2-2](https://doi.org/10.1002/1098-111X(200103)16:3<391::AID-INT1014>3.0.CO;2-2).
- Gruninger, Michael and Fox, Mark.S. 1995. Methodology for the Design and Evaluation of Ontologies. In Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, IJCAI-95, Montreal. Retrieved from: <http://www.eil.utoronto.ca/wp-content/uploads/enterprise-modelling/papers/gruninger-ijcai95.pdf>
- "Linked Data Cookbook - Government Linked Data (GLD) Working Group Wiki." n.d. Accessed September 25, 2019.

[https://www.w3.org/2011/gld/wiki/Linked\\_Data\\_Cookbook#Cookbook\\_for\\_Open\\_Government\\_Linked\\_Data](https://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook#Cookbook_for_Open_Government_Linked_Data).

Linked Data for Production (LD4P) Website

<https://wiki.duraspace.org/pages/viewpage.action?pageId=74515029> Accessed September 25, 2019.

Heath, Tom, and Christian Bizer. 2011. “Linked Data: Evolving the Web into a Global Data Space.” *Synthesis Lectures on the Semantic Web: Theory and Technology* 1 (1): 1–136. <https://doi.org/10.2200/S00334ED1V01Y201102WBE001>.

Noy, Natalya F., & McGuinness, Deborah L. 2001. Ontology development 101: A guide to creating your first ontology. Retrieved from [http://protege.stanford.edu/publications/ontology\\_development/ontology101-noy-mcguinness.html](http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html).

Obrst, Leo, Werner Ceusters, Inderjeet Mani, Steve Ray, and Barry Smith. 2007. “The Evaluation of Ontologies.” In *Semantic Web*, edited by Christopher J. O. Baker and Kei-Hoi Cheung, 139–58. Boston, MA: Springer US. [https://doi.org/10.1007/978-0-387-48438-9\\_8](https://doi.org/10.1007/978-0-387-48438-9_8).

Paris, Pierre-Henri, Nathalie Abadie, and Carmen Brando. 2017. “Linking Spatial Named Entities to the Web of Data for Geographical Analysis of Historical Texts.” *Journal of Map & Geography Libraries* 13 (1): 82–110. <https://doi.org/10.1080/15420353.2017.1307306>.



Pattueli, M. Cristina, Matt Miller, Leanora Lange, Sean Fitzell, and Carolyn Li-Madeo. 2013.

“Crafting Linked Open Data for Cultural Heritage: Mapping and Curation Tools for the Linked Jazz Project.” *The Code4Lib Journal*, no. 21 (July).

<https://journal.code4lib.org/articles/8670>.

Pattueli, M. Cristina, Alexandra Provo, and Hilary Thorsen. 2015. “Ontology Building for Linked Open Data: A Pragmatic Perspective.” *Journal of Library Metadata* 15 (3–4): 265–94. <https://doi.org/10.1080/19386389.2015.1099979>.

Raad, Joe, and Christophe Cruz. 2015. “A Survey on Ontology Evaluation Methods:” In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 179–86. Lisbon, Portugal: SCITEPRESS - Science and Technology Publications.  
<https://doi.org/10.5220/0005591001790186>.

Smith-Yoshimura, Karen. 2016. “Analysis of International Linked Data Survey for Implementers.” *D-Lib Magazine* 22 (7/8). <https://doi.org/10.1045/july2016-smith-yoshimura>.

Smith-Yoshimura, Karen. 2018. “Analysis of 2018 International Linked Data Survey for Implementers.” *The Code4Lib Journal*, no. 42 (November).  
[https://journal.code4lib.org/articles/13867#\\_edn4](https://journal.code4lib.org/articles/13867#_edn4).

Tallerås, Kim. 2017. “Quality of Linked Bibliographic Data: The Models, Vocabularies, and Links of Data Sets Published by Four National Libraries,” *Journal of Library Metadata*, 17:2, 126-155, DOI: [10.1080/19386389.2017.1355166](https://doi.org/10.1080/19386389.2017.1355166)

W3C. “Best Practices for Publishing Linked Data.” <https://www.w3.org/TR/ld-bp/> n.d. Accessed September 25, 2019.

Zaveri, Amrapali, Anisa Rula, Andrea Maurino, Ricardo Pietrobon, Julian Lehmann, and Sören Auer. 2015. “Quality Assessment for Linked Data: A Survey.” *Semantic Web* 7: 63–93. <https://doi.org/10.3233/SW-150175>.

## **Appendix 1: Survey**

### **Exploring methods for linked data model evaluation**

#### **Purpose of the research**

The purpose of the survey, "Exploring methods for linked data model evaluation," is to investigate the types of assessment practitioners in linked data undertake when building data models and better understand the reality of ontology evaluation in the context of a linked data project.

#### **Survey procedures**

The survey should take about 10 minutes to complete. You will be asked a series of open-ended questions relating to linked data projects you may have worked on.

#### **Benefits and risk**

Participants may benefit from thinking about their assessment processes. There are no perceived risks to answering this survey.

#### **Participation**

Your participation in this survey is entirely voluntary. At any point in time, you may opt to not answer any question. You may also choose to remove your consent or withdraw from the survey at any time by selecting "Exit and clear survey", in which case your responses will not be saved. As survey responses are anonymous, it will not be possible to have answers withdrawn once the survey is submitted.

#### **Confidentiality and anonymity**

The information being gathered is anonymous and your responses will not be traced back to you. Data collected via this survey may be used in research articles or professional presentations. Free text responses that are shared will have all possible identifying information removed.

If you have questions about this survey, please contact any of the researchers from this project: Clara Turp, McGill University Library & Archives ([clara.turp@mcgill.ca](mailto:clara.turp@mcgill.ca)), Robin Desmeules, McGill University Library & Archives ([robin.desmeules@mcgill.ca](mailto:robin.desmeules@mcgill.ca)), or Andrew Senior, McGill University Library & Archives ([andrew.senior@mcgill.ca](mailto:andrew.senior@mcgill.ca)).

If you have any ethical concerns or complaints about your participation in this survey, and wish to speak with someone not on the research team, please contact the McGill Ethics Manager at 514-398-6831 or [lynda.mcneil@mcgill.ca](mailto:lynda.mcneil@mcgill.ca).

#### **Consent**

By clicking Next, you consent that you are willing to answer the questions in this survey. Please save or print a copy of this document to keep for your own reference.

There are 18 questions in this survey

# Linked Data Experience

## 1. Have you ever been involved in data modeling for a linked data project?

Please choose **only one** of the following:

- Yes
- No

### 1. 1. How many projects have you been part of?

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? )

Only numbers may be entered in this field.

Please write your answer here: [answer]

### 1.2 Please tell us more about your involvement with these project(s), including but not limited to your role, and the tasks you undertook.

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? )

Please write your answer here: [answer]

### 1.1. Thank you for your interest. Is there anything you wanted to share with us about evaluating linked data models?

**Only answer this question if the following conditions are met:**

Answer was 'No' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? )

Please write your answer here: [answer]

## Data Modelling

Please answer the following questions by thinking back to a specific linked data project you were involved with.

[free text]

## 2.1 When building your data model for this particular project, did you:

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? )

Check all that apply

Please choose **all** that apply:

- Build a new ontology
- Adapt and reuse multiple existing ontologies
- Adapt and reuse a single existing ontology
- Other:

## 2.2 How did you choose your ontolog(ies)? Please select the option(s) that best describe your approach:

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? )

Check all that apply

Please choose **all** that apply:

- Consulted colleagues
- Consulted social media
- Consulted the literature
- Other:

## Data model evaluation

Please answer the following questions by thinking back to a specific linked data project you were involved with.

### 3. Did you evaluate your data model?

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? )

Please choose **only one** of the following:

- Yes

- No

### **3.1 Did you encounter any challenges or barriers that prevented you from evaluating your model?**

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'No' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? )

Please choose **only one** of the following:

- Yes
- No

#### **3.1.1 Please describe any challenges and/or barriers that you may have encountered.**

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '8 [Q2version2Evaluation]' (3.1 Did you encounter any challenges or barriers that prevented you from evaluating your model?)

Please write your answer here:

#### **3.1.2 What would have helped you evaluate your project?**

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'No' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? ) *and* Answer was 'Yes' at question '8 [Q2version2Evaluation]' (3.1 Did you encounter any challenges or barriers that prevented you from evaluating your model?)

Please write your answer here:

#### **3.1 Please describe how you evaluated your data model, including the following: your method of evaluation, the elements of the project you evaluated, and at which stages in the design and implementation you evaluated it.**

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'Yes' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? )

Please write your answer here:

### 3.2 Did you encounter any challenges or barriers during the evaluation process?

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'Yes' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? )

Choose one of the following answers

Please choose **only one** of the following:

- Yes
- No

#### 3.2.1 Please describe any challenges and/or barriers that you may have encountered.

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'Yes' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? ) *and* Answer was 'Yes' at question '12 [Q2version3Evaluation]' (3.2 Did you encounter any challenges or barriers during the evaluation process?)

Please write your answer here:

#### 3.2.2 What would have helped you evaluate your project?

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? ) *and* Answer was 'Yes' at question '12 [Q2version3Evaluation]' (3.2 Did you encounter any challenges or barriers during the evaluation process?)

Please write your answer here:

### 3.3 Did you share your evaluation process?

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'Yes' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? )

Please choose **only one** of the following:

- Yes
- No

### **3.3.1 How did you share your evaluation process? Check any that apply and use comments to share links and provide additional details.**

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'Yes' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? ) *and* Answer was 'Yes' at question '15 [Q5v1Evaluation]' (3.3 Did you share your evaluation process?)

Comment only when you choose an answer.

Please choose all that apply and provide a comment:

- In a paper
- In a presentation
- On Twitter
- On GitHub
- On blogs
- Other:

### **3.4 Is there anything else you would like to share with us?**

**Only answer this question if the following conditions are met:**

Answer was 'Yes' at question '1 [Q1Experience]' (1. Have you ever been involved in data modeling for a linked data project? ) *and* Answer was 'Yes' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? )

Please write your answer here: [answer]

### **3.2 Is there anything else you would like to share with us?**

**Only answer this question if the following conditions are met:**

Answer was 'No' at question '7 [Q1Evaluation]' (3. Did you evaluate your data model? )

Please write your answer here: [answer]

Thank you!

Submit your survey.

Thank you for completing this survey.



## **Appendix 2: Identified project tasks and roles grouped by theme**

Question: Please tell us more about your involvement with these project(s), including but not limited to your role, and the tasks you undertook.

### **A. Data modeling**

- parse out the goals of the system or the project
- suggest appropriate vocabularies, properties, and syntax standards
- reviewing others' edits/editions/extensions to existing ontologies
- giving input to consultants on the data model
- crosswalking from current standards
- identifying library use case
- our team developed fields and protocols for the data

### **B. Made ontology...**

- From scratch
  - we created our own ontology.
  - creating new ontologies from scratch with no reuse of concepts from other ontologies,
  - evaluating and selecting existing ontologies for import into a new ontology,
  - creating new ontologies with extensive reuse of concepts from other ontologies using the MIREOT technique common in the biomedical ontologies community,
- Used existing ontology
  - selecting an ontology to use for a small data analysis project,
  - attempted to find an ontology that would fit our project.
- Modified existing ontology
  - My tasks have included editing, augmenting, and extending existing ontologies,

### **C. Made a data model**

- design a conceptual metadata model
- designed a data model
- modeling content according to linked data best practices and principles.
- Data modeling
- data modeling for specific projects that are linked data related.
- done modeling for projects that preliminarily want to collect metadata but want the model / application profile to allow for eventual re-use/conversion of metadata to linked data
- modeled my institution's new linked data-based digital library
- modeling of data describing the library's subscription databases

### **D. Data munging**

- oversee production of metadata/linked data
- evaluating RDF serialization formats for usability
- transforming table-shaped data to RDF using tools such as Karma
- conversion of a legacy classification system from print to RDF/XML

- authority data project for linked data quality improvement,
- doing all the actual data creation
- promoting strings to things, reconciling URIs from multiple sources
- mapping MODS XML to RDF
- creating OWL implementations of existing taxonomies,
- mapping XML schema to existing ontologies for data transformation,
- conversion analysis both locally and advising on the transformation council
- converting the XML-encoded records into RDF using an XSLT transform.
- tasked with "linking and querying people by relationship"

#### E. Building applications

- creating application layers for ontology-driven software applications
- creation of a linked data visualization tool
- implementation of linked data in our DAMs system
- setting up a fuseki installation
- creating an application that could query the relationships.
- created a basic visualization to show "two-removed" social circles of individuals
- JSON-LD / RDFa website implementation.
- proof-of-concept application partially built with Fedora

#### F. Education

- led a Linked Data Study group
- led an experimental project to test the feasibility of transforming digital collections metadata into linked open data and publishing it
- "proof of concept"

## Figures

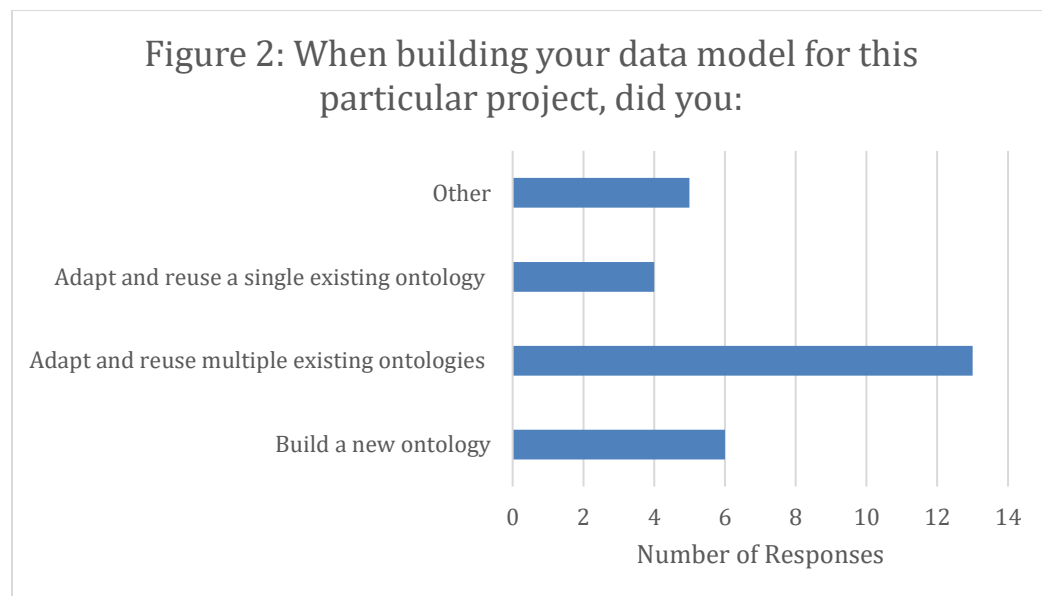
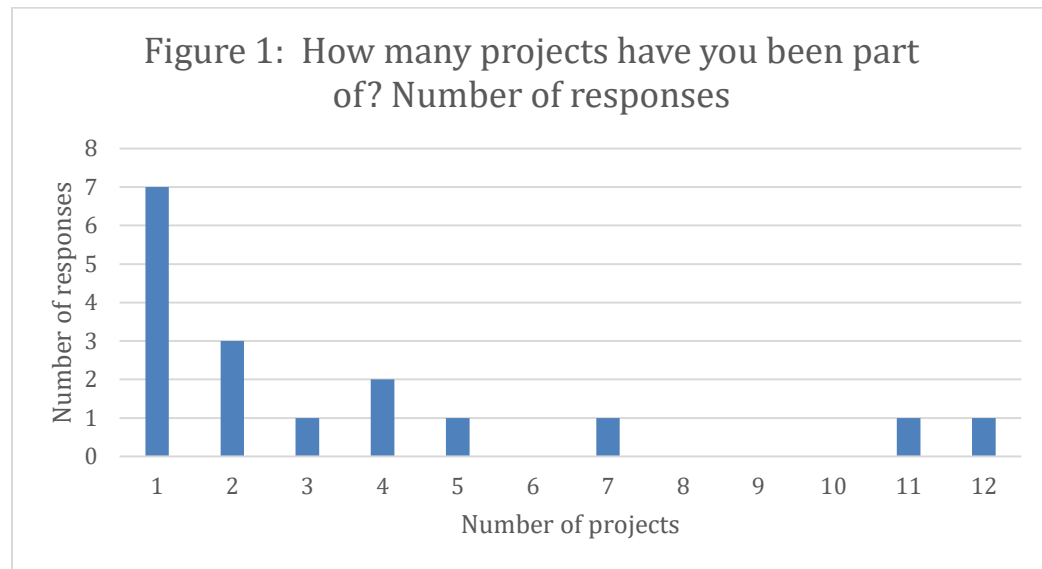


Figure 3: How did you choose your ontolog(ies)?

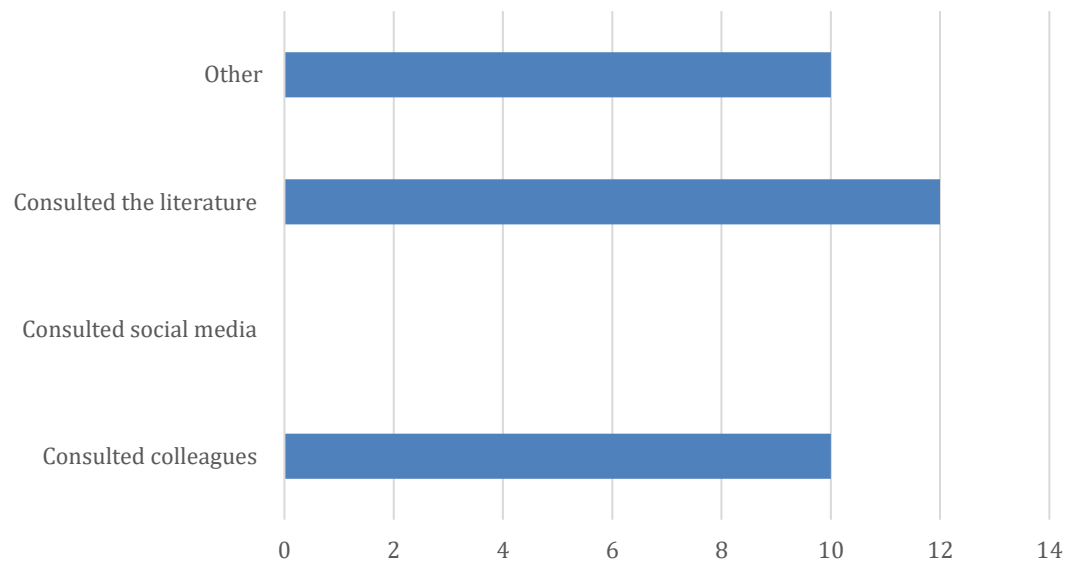
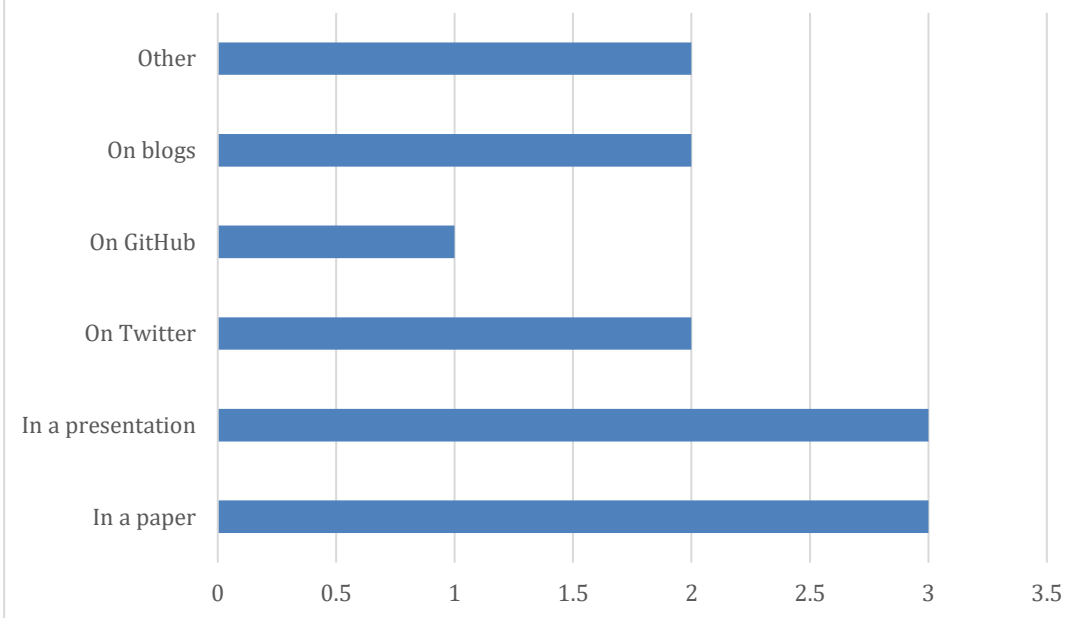


Figure 4: How did you share your evaluation process?



## **Figure Captions:**

Figure 1: How many projects have you been a part of? Distribution of number of projects undertaken by respondents.

Figure 2: When building your data model for this project did you...Number of respondents that either: adapted and reused a single ontology, adapted and reused multiple ontologies, or built a new ontology.

Figure 3: How did you choose your ontolog(ies)? Distribution of selection methods of ontologies. Respondents could choose more than 1 response.

Figure 4: How did you share your evaluation process? Distribution of different scholarly transmission methods. Respondents could choose more than 1 response.

**Table 1. Ontology Development Methods**

Respondent	Build a new ontology	Adapt and reuse multiple existing ontologies	Adapt and reuse a single existing ontology	Other
1		X		Also created minimal bespoke properties for local needs
2				
3		X		
4	X	X	X	extend an existing ontology
5		X		
6				
7	X		X	
8		X		
9		X		added local properties as needed
10			X	
11		X		
12	X	X		
13		X		
14	X	X	X	
15	X			
16		X		
17		X		Created some new properties but not an entirely new ontology
18	X	X		

**Table 2. Ontology Evaluation Methods**

Respondent	Iterative design	Test in the application	Comparing the data (application not mentioned)	Consult subject experts	Use cases	Competency questions	User testing
1	X	X					
2		X				X	X
3			X		X		
4	X		X				
5	X	X					
6	X		X				
7		X	X				
8	X			X			
9	X	X					X
10	X	X			X		