Ontologies in the Time of Linked Data

Hilary Thorsen, Stanford University, Stanford, CA, thorsenh@stanford.edu
M. Cristina Pattuelli, Pratt Institute, New York, NY, mpattuel@pratt.edu

ABSTRACT

This paper discusses some of the methodological issues one encounters when creating and using ontologies in the rapidly expanding Linked Open Data (LOD) landscape. Over the years the notion of applied ontologies has transitioned from that of a logically formalized knowledge system with varying degrees of inferencing power to that of a lightweight knowledge representation tool. This shift is reflected in the current lexicon where different actors in the LOD community use the term *ontology* interchangeably with more generic terms like *vocabulary*[1] or even *namespace* or *data schema*. Applied ontologies have been a key area of research in the context of Semantic Web initiative since the late 1990s. The Semantic Web has recently found a new stream of development in the Linked Data initiative, which is considered its natural evolution (Allemang and Hendler, 2011). While a good deal of literature has been devoted to investigating ontology engineering for the Semantic Web, not enough attention has yet been paid to understanding the nature and role that ontologies play in the linked data context, especially from the lens of knowledge organization research. Based on our ongoing work creating Linked Open Data applications and services for digital resources in the domain of the performing arts, we describe methodological steps and lessons learned in line with the spirit of the linked data initiative, where an agile and pragmatic approach to development is combined with the practice of learning from one another.

BACKGROUND

LOD is a W3C and community-based initiative working on extending the Web as we know it by meaningfully connecting data from heterogeneous sources. The central idea of linked data is achieved by making data processable by machines and accessible seamlessly using the Web itself as a unifying discovery space. In an effort to realize the "Web of Data", as Tim Berners-Lee defines the LOD environment (Bizer, Heath, and Berners-Lee, 2009), a great number of LOD datasets have been created and made freely available for sharing, reuse and interlinking. A visualization of the LOD ecosystems is provided by the LOD cloud[2], showing the dense interlinking between heterogeneous sets of data that continues to exponentially grow in different domains. Ontologies form the backbone of this linked data environment and are key to supporting its open and distributed infrastructure. To help understand the

---

[1] In the context of this paper, the terms *ontology* and *vocabulary* are also used interchangeably.

[2] http://lod-cloud.net/

nature and role of ontologies in this emerging information context, we begin by describing the technological framework that supports LOD development.

LOD T̲ECHNOLOGY S̲TACK

The LOD infrastructure relies on a rather small set of existing open standards that are deeply engrained in the fabric of the Web: a naming standard to uniquely identify resources using the URI (Uniform Resource Identifier) and the common Hypertext Transfer Protocol HTTP. The RDF (Resource Description Framework) model serves as the unifying platform to represent and exchange data (Berners-Lee, 2009). In other words, RDF is the common framework for representing resources (Schreiber and Raimond, 2014). In the context of LOD, a resource is anything that can be identified or named including any object, event or unit of information that can be referenced by a URI. The basic blocks of the LOD infrastructure are simple statements, called RDF triples, composed of three atomic elements: a subject, a predicate and an object. Each element of a triple is paired with a URI that performs a referential function, making it both human readable and machine processable. Because of this global naming convention, any object of an RDF triple can become the subject of another triple, creating chains of relationships and representing the information space as a graph or network. As a consequence, the multitude of ontologies that populate the LOD ecosystem all share the same underlying semantics made up of relatively simple modeling constructs. The implications of relying on open standards and a common data model to facilitate data interoperability and integration are far-reaching.

ONTOLOGY DEVELOPMENT METHODOLOGY

The process of building applied ontologies has been addressed rather extensively, especially within the framework of the Semantic Web (Noy and McGuinness, 2001; Gómez-Pérez, Corcho-García and Fernández-López, 2003). In the context of linked data development, methodological guidelines for creating and using RDF-based ontologies have yet to be established. However, best practices are emerging through shared documentation and lessons learned. Most recently, Villazón-Terrazas et al. (2011) propose a six-stage methodology that consists of: 1) specification, 2) modeling, 3) generation, 4) linking, 5) publication, and 6) exploitation. We will use these sequential steps as a general framework to address the process of building and using an ontology in the context of the Linked Jazz Project, which provides a real-world application scenario.

ONTOLOGY S̲COPE AND P̲URPOSE

The primary goal of Linked Jazz[3] is to leverage linked data principles and technologies to uncover the dense web of relationships between artists in the jazz community. The project relies on transcriptions of oral histories to identify relevant entities (jazz musicians) as well

---

[3] https://linkedjazz.org/

as the professional and personal relationships that occur among them (Pattuelli, Miller, Lange, Fitzell, and Li-Madeo, 2013). While still evolving and expanding to new areas of cultural heritage, Linked Jazz has delivered a rich LOD dataset representing over 9,000 artist entities and their connections. These connections are assigned a specific relationship type. The source of the relationship is the occurrence of a mention in the transcript text. In other words, whenever the subject of an oral history mentions someone, a triple is created that expresses a claim that this individual *knows of* the person they cite. For example, the claim that *Sam Rivers (*subject*) knows of* (predicate) *Dizzy Gillespie* (object) is represented by the triple below:

```
<http://linkedjazz.org/resource/Sam_Rivers>
<http://purl.org/vocab/relationship/knowsOf>
<http://linkedjazz.org/resource/Dizzy_Gillespie>
```

SPECIFICATION AND MODELING

The specification of the content domain was based on the analysis of our data sources, oral histories in the field of jazz history. Jazz artists were identified as the primary entities. A small set of classes and properties was required to model these entities and their properties. One type of relationship (*rel:knowsOf*) was sufficient to describe the connections among musicians as derived from the data sources (Figure 1). The modeling process was driven by one of the main principles and established practices in LOD development: the reuse of existing and publicly available LOD semantics. The value of adopting terms, whenever possible, from existing RDF ontologies is considered a powerful way to make it easier for applications to process and integrate linked data (Heath and Bizer, 2011).
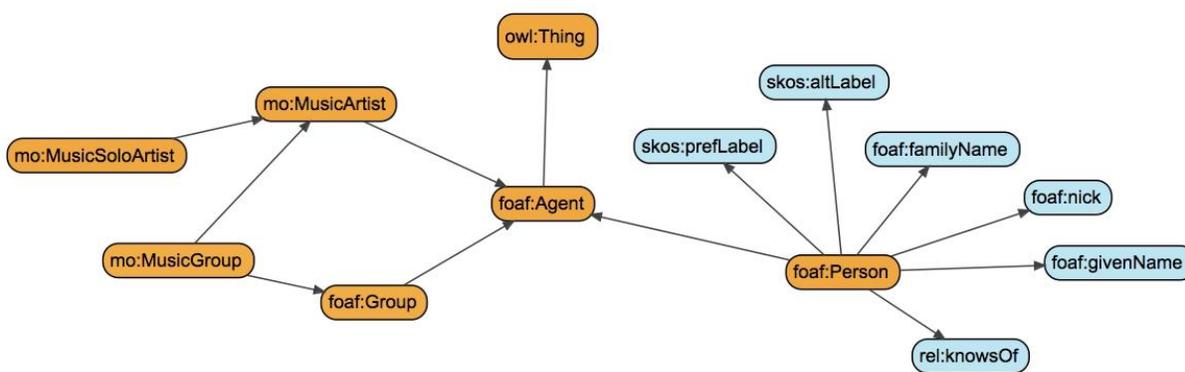

Figure 1 Diagram of core set of classes and properties.

As the project progressed and more semantic complexity was required to enrich the original set of data with new layers of meaning, the need arose to integrate the original

core with a broader range of personal and professional relationships. More specifically, the nature of the basic connections held by musicians was further specified and their semantics assigned using a crowdsourcing approach (Pattuelli, Miller, Lange, and Thorsen, 2013). Through a dedicated platform[4], different types of personal and professional relationships are manually contributed by volunteers with the goal of enriching the core Linked Jazz dataset with accurate and granular predicates. Crowd annotations are automatically mapped to a set of predicates derived from existing LOD vocabularies. These predicates, taken from the Relationship Vocabulary and Music Ontology, include *rel:knowsOf, rel:hasMet, rel:acquaintanceOf, rel:closeFriendOf, rel:influencedBy, mo:collaborated_with,* and *rel:mentorOf.*

The capability to easily extend our evolving data model is an important trait of the modeling practices in the LOD context. Elements from different vocabularies are easily integrated in an existing ontology through the process of mixing and matching. Multiple vocabularies at once can serve as sources of semantics and enrich an ontology in a layered fashion without the need for community agreement for adoption. Properties with overlapping scope can coexist without hampering the consistency of the schema. Again, the openness and decentralized nature of the RDF framework are key to enable inclusion of terms from external sources without the need for coordination or formal agreements on the adoption of a specific schema.

Both reuse and extensibility are distinctive traits of LOD ontologies that mark a clear departure from traditional computational ontologies. LOD ontologies are characterized by several features that facilitate the reuse of elements from existing RDF-vocabularies and their integration into a target knowledge system. First, they are lightweight in terms of their level of formality and typically small in size making it easy to manage and maintain them. Second, they rely on well-established and W3C-governed representation systems including SKOS (Simple Knowledge Organization System), RDFS (RDF Schema), and OWL (Web Ontology Language) as reference models and sources of essential concepts (e.g., *owl:Thing* usually serves as the ontology root). Having their basic representation framework grounded on sound and widely adopted standard vocabularies has the benefit of enforcing stability and facilitating interoperability and adoption by service providers and users.

This high degree of flexibility is especially suitable for supporting the dynamic nature of the evolving web of linked data. As discussed earlier, the openness of the management and discovery environment (the Web itself), the light-weight nature of LOD standards, and the unifying role played by the RDF platform are key to enabling terms from published

---

[4] https://linkedjazz.org/52ndStreet/

vocabularies to be easily included in an existing ontology. Rather than duplicating effort by defining concepts and properties that have been already defined elsewhere, we can point to the selected elements in an existing ontology via a URI. For example, the predicates that describe the personal and professional relationships in our extended dataset are derived from the Music Ontology[5] as well as from the Relationship Vocabulary[6]. To this end, the identification of suitable vocabularies as sources of reusable semantics represents an important methodological step in the modeling stage of the ontology building process.

VOCABULARY SELECTION

As part of the acquisition process, deciding which LOD ontologies to reuse can be challenging. As the LOD initiative has grown exponentially in the last few years, RDF-based ontologies have proliferated creating considerable overlap between them. Some of these are general-purpose ontologies published by trusted and stable bodies, such as the W3C RDF Concepts Vocabulary[7] and the Geospatial Vocabulary[8], as well as the Dublin Core Ontology[9]. These general ontologies typically describe entities, such as people, organizations, events, and geographic locations. They serve as reference vocabularies and "anchors" for designing a conceptual model. DBPedia's ontology[10] is particularly useful for its cross-domain and extensive thematic coverage. Created to map the massive amounts of data extracted from Wikipedia, the DBpedia ontology has become a de facto reference vocabulary due to the popularity of the DBpedia dataset in the LOD landscape.

Libraries (including the Library of Congress), museums (especially the International Council of Museums, which developed CIDOC-CRM[11], or the International Committee for Documentation's Conceptual Reference Model), and archives (which created the EAC-CPF, or the Encoded Archival Context-Corporate Bodies, Persons, and Families XML schema),[12] have also published RDF ontologies to reflect the representational needs of the communities that create them. Some ontologies originally developed by individuals as independent projects have gained popularity through grassroots adoption. This is the case of FOAF[13] (Friend of a Friend), which contains commonly used predicates for basic descriptions of people, and the Relationship Vocabulary, which provides predicates for describing relationships between people. While these vocabularies offer general terms, it is often necessary to include more granular terms from domain-specific ontologies, such as

---

[5] http://musicontology.com/specification/
[6] http://vocab.org/relationship/.html
[7] http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/
[8] http://www.w3.org/2005/Incubator/geo/XGR-geo/
[9] http://dublincore.org/documents/dcmi-terms/
[10] http://mappings.dbpedia.org/server/ontology/classes/
[11] http://cidoc-crm.org/docs/cidoc_crm_version_5.0.4.pdf
[12] http://labs.regesta.com/progettoReload/wp-content/uploads/2013/10/eac-cpf.html
[13] http://xmlns.com/foaf/spec/

the Music Ontology, to enrich their expressivity (Raimond, Abdallah, Sandler and Giasson, 2007).

To help navigate the proliferation of ontologies, a few services exist to search for vocabularies. We relied on Linked Open Vocabularies (LOV)[14] because it is a stable, carefully-curated repository maintained by the Open Knowledge Foundation. To be included in LOV, a vocabulary needs to fulfill a number of criteria: it must have stable URIs; be available on the web; use publication best practices; follow standard formats; contain high quality metadata and documentation; be published by an identifiable and trusted entity; and adhere to versioning policies. Users can suggest vocabularies to become a part of LOV. LOV curators review vocabularies, validate them, and put them in the LOV database. As part of the validation process, the vocabulary authors are contacted to ensure the vocabulary meets the requirements of LOV and has been published according to best practices. Once a part of the database, an automatic script checks for vocabulary updates on a daily basis and stores all the older versions of the vocabulary as well as the newest. LOV curators review the vocabularies on an annual basis. LOV code and data is published with a Creative Commons 4.0 license (CC BY 4.0). Users and machines can access the data by downloading a data dump, running SPARQL queries on the SPARQL Endpoint, or using the LOV APIs (Application Program Interfaces).

In addition, several features facilitate evaluation of the vocabularies. The LOV search engine is quite robust with a ranking algorithm based on term popularity in datasets and within LOV. All terms are indexed and full text can be searched. It is possible to search for vocabularies, terms (classes/properties), and agents and to filter by subject tag and language. The number of classes, properties, instances, and datatypes of the vocabulary as well as the language or languages it is expressed in are provided, which gives an overview of the vocabulary's characteristics. Popularity of the vocabulary is indicated by the number of datasets it is used in, as well as the number of incoming links to it. Incorporating popular terms means less work for machines as tools that understand those vocabularies can be reused. Popularity also indicates that the vocabulary is likely to be trustworthy. Links to the publisher and a timeline of the vocabulary's version history can be used to help determine the vocabulary's stability, maintenance, and government structure.

While vocabulary selection can be straightforward, a few challenges may also arise. Overlapping predicates in vocabularies are quite common. If the overlap is slight and the semantic meaning of the terms is very similar, these similar predicates can more easily be accommodated in lightweight ontologies. If, however, the semantics for the terms are different, it is not possible to include both, as the original semantics of terms from other ontologies must be preserved. In our case, for example, we needed a predicate to describe

---

[14] http://lov.okfn.org/dataset/lov/

the collaborative relationship between musicians. Two options were available—one in the Music Ontology (*mo:collaborated_with)* and one in the Relationship Vocabulary (*rel:collaboratesWith)*. The definition of *rel:collaboratesWith* was more general, "A property representing a person who works towards a common goal with this person,"[15] while the *mo:collaborated_with* definition, "Used to relate two collaborating people on a work,"[16] was more specific to the music context, our domain of interest. In addition, as a sub-property of *foaf:knows*, the predicate *rel:collaboratesWith* is assigned to the class *foaf:Person*, whereas the domain and range of *mo:collaborated_with* is the class *foaf:Agent*, which means collaborations can take place not only between people, but also between organizations, for instance, a band. While *mo:collaborated_with*, was deemed the most appropriate representational choice, we had to ensure that its original semantics was preserved, including having *foaf:Agent* as its domain constraint in order to prevent conceptual inconsistencies.

An overview of the range of source vocabularies selected for the Linked Jazz ontology is shown in Table 1.

Table 1. Source vocabularies

lj:	http://linkedjazz.org/ontology – namespace for LJ: Linked Jazz Ontology
dbpedia-owl:	http://dbpedia.org/ontology – namespace for DBPEDIA: DBpedia ontology
event:	http://purl.org/NET/c4dm/event.owl# – namespace for EVENT: Event Ontology
foaf:	http://xmlns.com/foaf/0.1/ – namespace for FOAF: 'Friend of a Friend' vocabulary
mo:	http://purl.org/ontology/mo/ – namespace for MO: The Music Ontology
owl:	http://www.w3.org/2002/07/owl# – namespace for OWL: Web Ontology Language
rel:	http://purl.org/vocab/relationship – namespace for REL: Relationship vocabulary
rdf:	http://www.w3.org/1999/02/22-rdf-syntax-ns# – namespace for RDF: Resource Description Framework
rdfs:	http://www.w3.org/2000/01/rdf-schema# – namespace for RDFS: RDF Schema
skos:	http://www.w3.org/2004/02/skos/core# – namespace for SKOS: Simple Knowledge Organization System

---

[15] http://vocab.org/relationship/.html#collaboratesWith
[16] http://musicontology.com/specification/

GENERATION

In the event that a suitable term for reuse cannot be found, it becomes necessary to create a new one. Best practices call for reusing terms in domain-specific non-ontological resources, such as a thesaurus or glossary, if possible and converting them into RDF ontologies (Vila-Suero, Gómez-Pérez, Montiel-Ponsoda, Gracia, & Aguado-de-Cea, 2014). There are instances, however, when expressive needs cannot be fulfilled by existing sources and local terms need to be coined. Our ontology development experience presents an interesting use case. The Music Ontology, one of our ontology sources, offers a suitable predicate, *mo:collaborated_with*, to express the relationship of collaboration between two musicians, essential to our area of interest. Based on the analysis of the content domain and the recommendation of domain experts, we intended to represent this relationship at a more granular level including various types of collaboration deemed relevant by our intended user community. We could find very few predicates in existing standard vocabularies to represent more nuanced degrees of collaboration suitable to our context. After evaluating a few modeling alternatives, we made the decision to mint a small set of sub-properties of the property *mo:collaborated_with*, following best practices for creating persistent human and machine readable URIs through a human and machine readable pattern: http://(domain)/(type)/(concept)/(reference). The newly minted sub-predicates include:

> lj:playedTogether (http://linkedjazz.org/ontology/playedTogether);
> lj:touredWith (http://linkedjazz.org/ontology/touredWith);
> lj:inBandWith (http://linkedjazz.org/ontology/inBandWith);
> lj:bandLeaderOf (http://linkedjazz.org/ontology/bandLeaderOf);
> lj:bandMemberOf (http://linkedjazz.org/ontology/bandMemberOf)

A comprehensive view of the current state of the Linked Jazz conceptual model, including the newly coined local sub-properties, is shown in Figure 2.
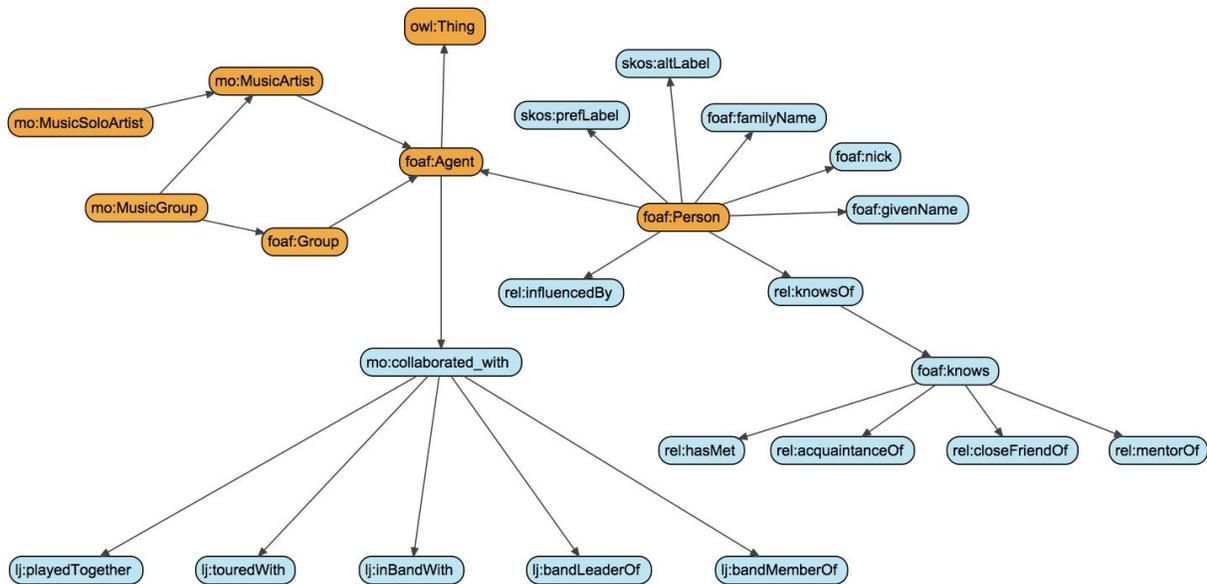
Figure 2. Diagram of the Linked Jazz conceptual model.

Efforts are now underway to expand the current conceptual model to include additional aspects of our domain of interest. More descriptive power will soon be needed as we are in the process of creating data mashups by integrating the existing dataset with data from a number of external sources representing discography and music performances. The data enrichment offered by mashing up datasets will open up unanticipated opportunities for information discovery and analysis, bringing together a wide range of information from temporal and spatial data (e.g., time periods, dates, events, geographic locations, etc.) to music-specific data (e.g., professional roles, instruments, recordings, music venues, etc.) . Data mashups are a common method employed to augment LOD datasets and offer integrated views of the data. An extended ontology will be needed to harmonize the added semantics and ease the process of integrating data. Based on the analysis of the data sources and of their underlying schemas, we have identified candidate entities and properties, created an inventory of terms, and performed crosswalks and mapping useful for supporting the selection of terms and informing modeling decisions. Entities suitable for inclusion are: *mo:MusicalWork; mo:Recording; event:Event; event:Place; mo:Performance; mo:Session; mo:Release;* and *foaf:Organization* (subclass: *mo:CorporateBody*).

LINKING

Interlinking data lies at the core of the LOD paradigm and is one of its distinctive elements. As discussed earlier, web resources are identified by a standard naming convention, the URI, and connected via meaningful links, which are also uniquely identified by a URI. These typed links are predicates drawn from RDF-based ontologies that provide the semantic

glue to connect the web of linked data. Linking an ontology to other vocabularies provides a way of extending its domain coverage and the breadth of its potential use. In our scenario, for example, we use the relationship *mo:collaborated_with,* derived from the Music Ontology, that relates two individuals collaborating on a work. This predicate has also been described in the Relationship Ontology as *rel:collaboratedWith* with a rather similar scope. A link has been specified in our ontology of the type *owl:sameAs* to connect the two predicates. This "identity link" makes it possible for both ontologies to publish information about the same entity, for example a jazz musician, and thus to maximize query results. Another type of connector, the "relationship link" *rdfs:seeAlso,* plays a critical role in enabling interlinking between terms and pointing to related resources that might be of interest to the user. Similarly, links can be created at the class level using r*dfs:subClassOf* or *owl:equivalentClass* relations. These relationships can then be exploited through a reasoner so that new semantics can be inferred by consuming applications.

Links are also created at the instance level. This process is performed by determining equivalent URIs and pairing them using the identifying link *owl:sameAs.* Reusing existing URIs for resources is a highly recommended LOD best practice to create interlinking between different LOD datasets. The reuse of URIs links a dataset with other related data and enters it into a semantically richer data ecosystem where equivalent and related entitles can be effortlessly discovered*.* In our context, for example, we co-referenced proper names of jazz musicians with equivalent entities from the DBpedia, VIAF and MusicBrainz:

```
<http://linkedjazz.org/resource/Sam_Rivers>
  <owl:sameAs>
<http://viaf.org/viaf/164803380/>
  <owl:sameAs>
<https://musicbrainz.org/artist/f86342be-eef7-445b-90c9-
250bdf3f0b3b>
```
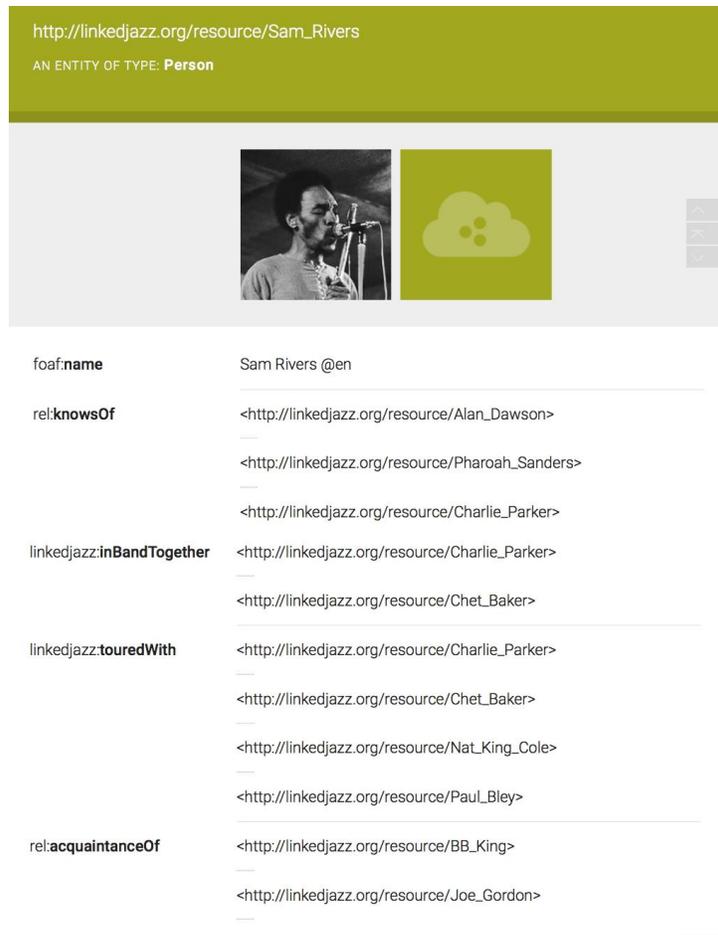
PUBLICATION AND EXPLOITATION

There are novel aspects of ontology development, still rather neglected in the literature, that have a direct impact on the way linked data is discovered and consumed, including ontology publication. Publishing ontologies on the web—making their classes and properties linkable—is an essential requirement of LOD practice. Ontologies should be published in RDFS or OWL, so that they can be interpreted by machine and in HTML, so that humans can understand them (Heath and Bizer, 2011). A visual form of the vocabulary is also helpful for providing an overview. Each concept and entity needs to be identified by a global identifier using the namespace URI for the ontology. These should be published online with their associated definitions and links to other related or identical concepts and entities, so that they can be discovered and verified (dereferenced, in LOD jargon) by a human user as a way to facilitate, share, and reuse trusted semantics. Documentation

should be updated regularly and each term should be published with examples and use cases to make it more easily understood and more likely to be adopted by the community. In this documentation, *rdfs:label* provides the entity's main title; *rdfs:comment* contains the entity's role; *rdfs:isDefinedBy* demonstrates explicitly the link between the entity and its namespace; and *vs:term_status* indicates the entity's status (Vandenbussche, Vatant, 2012).

Metadata about the vocabulary itself should be provided in RDF (Vandenbussche, Vatant, 2012). The vocabulary's URI serves as its identifier. A prefix identifying the namespace is often used as an abbreviation. Either Dublin Core or RDFS terms can be used to provide the title and description of the vocabulary. Since ontologies are regularly updated, dates of first publication, last modification, current version, and latest changes should be provided. Information about rights and licensing facilitate adoption.

While publishing the final version of the Linked Jazz Ontology is the next goal of the project, its current elements can already be perused online. This is made possible thanks to a web application, LodView, which we leverage to publish the Linked Jazz dataset according to defined standards for Linked Open Data. More specifically, LodView addresses the LOD requirement that each URI has to be "dereferenceable." This means that each identifier would resolve to a web page so that it could be looked up by user applications. LodView provides the platform to publish the URIs (entities and predicates) in our dataset as web pages (Figure 3).

Figure 3. Dereferenced URI for the person entity Sam Rivers.

Once data has been published, it can be accessed through different systems, including Semantic Web browsers and Semantic Web search engines. However, the most common access method is via SPARQL (SPARQL Protocol and RDF Query Language), a semantic query language for databases. SPARQL enables a query-answering type of retrieval where answers are automatically returned from posed queries. Query answering is particularly relevant to the LOD context because it provides a mechanism that enables users and applications to directly interact with ontologies and data. SPARQL queries are made against an endpoint, which is a data repository able to receive a query, interpret it, and return data. Figure 4 displays an example of SPARQL query from our SPARQL endpoint providing access to the Linked Jazz dataset, which has been made available on the Web. LOD applications can then be built on top of SPARQL query endpoints to transform the data returned into visualizations or other types of discovery services.

Figure 4. SPARQL Query returning Sam River's relationships.

The range and complexity of the queries that can be formulated in SPARQL is proportional to the level of complexity of the ontology underlying a set of data. SPARQL queries can generally access the classes, subclasses, properties, and sub-properties provided by the structure of RDFS, OWL, and SKOS vocabularies. As discussed earlier, LOD ontologies are considered lightweight knowledge organization systems. In general, their class structure is rather shallow and the representation constraints imposed to the properties are limited to domain and range. While formally defined ontologies carry a higher level of expressivity that allows for deeper exploitation of their semantics, LOD ontologies have limited reasoning capabilities. Their extensibility as well as the agile implementation and ease of use make up, to some extent, for less expressive power. The lack of complexity that characterizes LOD ontologies has been key to the rapid adoption of LOD technologies and to the continuous growth of LOD datasets. The entire LOD technical framework makes the consumption of ontologies a rather straightforward process. Systematic research has yet to be conducted on the implications of relying on informal and often vague semantics when it comes to issues of semantic consistency in large scale contexts of use.

CONCLUSION AND FUTURE WORK
In this paper, we have discussed our efforts towards the creation of an ontology for LOD applications in the domain of music with the goal to identify distinctive traits of ontology engineering in the context of linked data development. As linked data methods become increasingly pervasive as a popular means to represent web resources, more investigation

is needed to understand the challenges and the opportunities presented by engineering ontologies in this new context. Our next step will be to consolidate and publish the Linked Jazz ontology in line with LOD principles and best practices. As described in our scenario, the ontology building process can be relatively simple in terms of data acquisition and modeling, when compared to traditional practices. Our process has progressed in an incremental fashion, driven by pragmatic considerations concerned with the purpose and intended use of the ontology. We have adopted a bottom-up and ad hoc approach, common to LOD development. Evaluation activities will follow the implementation of the ontology to assess its performance and usability. Piloting and testing are intrinsically part of LOD practices as well as sharing and building upon one another's efforts. In this spirit, we hope others will benefit from and contribute to the work presented here.

REFERENCES

Allemang, D., & Hendler, J. A. (2011). *Semantic Web for the working ontologist: Effective modeling in RDFS and OWL*. Waltham, MA: Morgan Kaufmann/Elsevier.

Berners Lee, T. (2009, February). Tim Berners-Lee: the next Web. Retrieved from http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html.

Bizer, C., Heath, T. & Berners-Lee, Tim. (2009). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5(3), 1-22.

Gómez-Pérez, A., Corcho-García, O., and Fernández-López, M. (2003). *Ontological* Engineering. New York: Springer.

Heath, T. and Bizer, C. (2011). *Linked Data: Evolving the Web into a Global Data Space.* Morgan & Claypool.

Noy, N. and McGuinness, D. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. Retrieved from http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.

Pattuelli, M.C., Miller, M., Lange, L., Fitzell, S. and Li-Madeo, C. (2013) Crafting Linked Open Data for Cultural Heritage: Mapping and Curation Tools for the Linked Jazz Project. *Code4Lib Journal,* (21*)* 2013-07-15.

Pattuelli, M. C., Miller, M., Lange, L., and Thorsen, H. (2013). Linked Jazz 52nd Street: A LOD crowdsourcing tool to reveal connections among jazz artists. *Proceedings of Digital Humanities 2013*, 337-340.

Pattuelli, M. C. and Rubinow, S. (2013). The knowledge organization of DBpedia: A case study. *Journal of Documentation*, (*69)*6.

Raimond, Y., Abdallah, S., Sandler, M., Giasson, F. (2007) The Music Ontology. *ISMIR*, 417-422.

Schreiber G. and Raimond Y. (2014). *RDF 1.1 Primer. World Wide Web Consortium*. Retrieved from http://www.w3.org/TR/rdf-schema/

Vandenbussche, P. and Vatant, B. (2012). Metadata Recommendations for Linked Open Data Vocabularies. Retrieved from
http://lov.okfn.org/Recommendations_Vocabulary_Design.pdf

Vila-Suero, D., Gómez-Pérez, A., Montiel-Ponsoda, E., Gracia, J., and Aguado-de-Cea, G. (2014). Publishing Linked Data on the Web: The Multilingual Dimension. In *Towards the Multilingual Semantic Web* (101-117). Berlin: Springer.

Villazón-Terrazas, B., Vilches-Blázquez, L., Corcho, O., & Gómez-Pérez, A. (2011). Methodological guidelines for publishing government linked data. In D. Wood (Ed.), *Linking government data* (27–49). New York: Springer.