

Metadata Recommendations For Linked Open Data Vocabularies

Authors	Pierre-Yves Vandenbussche Bernard Vatan
Version	1.1
Date	2012-08-19
Modifications	VOAF namespace updated

Abstract

This document describes metadata recommendations for vocabulary description on the Web (in RDFS or OWL). We document the properties and classes used to describe metadata about the vocabulary itself and about vocabulary elements.

Introduction

As the Web of data is growing, emerging initiatives try to organize it [Heath and Bizer, 2011]. A particular project, the Linked Open Data¹ (LOD), helps to improve the quality of data shared by promoting their links and reuse [Bizer et al., 2009]. The data published on the Semantic Web is based on the use of shared vocabularies and ontologies that allow to structure and describe data. These vocabularies act as tools for building formalized reusable data. To facilitate the reuse, we propose some recommendations about metadata for such vocabularies.

Scope

We must first clarify the meaning of "Vocabulary" in the context of this document. We do not stick to the definition of vocabulary in linguistics meaning but approach it in terms of its use in the Semantic Web. In this context, a vocabulary is basically synonymous² of ontology. However, we differentiate vocabulary from an ontology by characteristics enabling reuse and integration by other vocabularies:

- Small size
- Low formal constraints (basically RDFS and a fistful of OWL)
- Few instances except for examples
- Rich user documentation (Labels, comments, definition, description, etc.)

By linking and reusing each other, vocabularies contribute to the growth of an awesome ecosystem: "*The Linked Open Vocabularies*".

Namespaces

Prefix	Namespace URI
owl	http://www.w3.org/2002/07/owl#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
dc	http://purl.org/dc/terms/
foaf	http://xmlns.com/foaf/0.1/
vs	http://www.w3.org/2003/06/sw-vocab-status/ns#
vann	http://purl.org/vocab/vann/

¹ See <http://linkeddata.org/>

² See <http://www.w3.org/standards/semanticweb/ontology>

cc	http://creativecommons.org/ns#
voaf	http://purl.org/vocommons/voaf#

Vocabulary Metadata

The recursive nature of RDF can be used to describe the vocabularies themselves. There is now a broad range of metadata vocabularies for a detailed description of vocabularies as well as their elements [LOV-Metadata]. In a good practice publication, description of the vocabulary metadata is contained in the RDF as properties of a resource representing the vocabulary.

Identification

A vocabulary is identified by its URI as any resource on the Web of data. To simplify the writing of its elements, it is often associated with a prefix that identifies the namespace.

Element	Usage
voaf:Vocabulary	Type of the resource which describes the vocabulary
vann:preferredNamespacePrefix	Indication of the prefix to be used when referring the vocabulary namespace. A prefix should be as short as possible (usually less than 5 characters) and avoid conflicts with existing ones (cf. the LOV dataset).
vann:preferredNamespaceUri	Indication of the namespace to be used for this vocabulary

```

<owl:Ontology rdf:about="http://purl.org/vocommons/voaf">
  <rdf:type rdf:resource="http://purl.org/vocommons/voaf#Vocabulary"/>
  <vann:preferredNamespacePrefix>voaf</vann:preferredNamespacePrefix>
  <vann:preferredNamespaceUri>http://purl.org/vocommons/voaf#</vann:preferredNamespaceUri>
</owl:Ontology>

```

VOAF vocabulary identification.

Title and description

A vocabulary is designated by its terms and a description that facilitates the understanding of its role and purpose. The most common way to name it is to use the attribute dc:title, but it should be noted that some applications prefer to find a more generic attribute rdfs:label, and that in the current DCMI specifications, dc:title is not a subproperty of rdfs:label.

Element	Usage
dc:title	Main vocabulary label
dc:description	Role and intention description of the vocabulary

```

<owl:Ontology rdf:about="http://purl.org/vocommons/voaf">
  <dc:title xml:lang="en">Vocabulary of a Friend</dc:title>
  <dc:description xml:lang="en">A vocabulary to describe linked data vocabularies and their relations.</dc:description>
</owl:Ontology>

```

VOAF vocabulary title and description.

Version and modification

Like any model, a vocabulary may change. In order to provide the change information to its users, it is important to specify the dates of first publication, last modified, the current version of the vocabulary and the latest changes. We will return in the "Documentation" section on the management of access to different versions of a vocabulary.

Note that we currently lack a vocabulary, more expressive than Dublin Core to declare the overall status of a vocabulary: is it a "work in progress"? Is it a draft of a version which will not be changed (e.g. a standard)? What are the coordinates of the curator, the address of a forum for discussion of the vocabulary? etc.

Element	Usage
dc:issued	Date of the first vocabulary publication
dc:modified	Last vocabulary modification date
owl:versionInfo	Current version of the vocabulary

rdfs:comment

Information about this vocabulary (e.g. change log between two version)

```
<owl:Ontology rdf:about="http://purl.org/vocommons/voaf">
  <dc:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-12-21</dc:issued>
  <dc:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2011-11-16</dc:modified>
  <owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">1.1</owl:versionInfo>
  <rdfs:comment xml:lang="en">-Version 1.1: depreciation of voaf:exampleDataset replaced by
voaf:dataset
-Version 1: creation</rdfs:comment>
</owl:Ontology>
```

VOAF vocabulary version and modification informations

Rights and property

Information and property rights provides information to users about vocabulary use and adaptation.

Element	Usage
dc:rights	Property rights of the vocabulary
cc:licence	Licence of the vocabulary
dc:creator	Creator(s) of the vocabulary
dc:contributor	Contributor(s) of the vocabulary
dc:publisher	Publisher of the vocabulary

```
<owl:Ontology rdf:about="http://purl.org/vocommons/voaf">
  <dc:rights>Copyright © 2011 Mondeca</dc:rights>
  <cc:license rdf:resource="http://creativecommons.org/licenses/by/3.0/" />
  <dc:contributor>
    <foaf:Person rdf:about="http://labs.mondeca.com/foaf/mondeca.rdf#lrozat" />
  </dc:contributor>
  <dc:contributor>
    <foaf:Person rdf:about="http://labs.mondeca.com/foaf/mondeca.rdf#pyv" />
  </dc:contributor>
  <dc:creator>
    <foaf:Person rdf:about="http://labs.mondeca.com/foaf/mondeca.rdf#bvatant" />
  </dc:creator>
  <dc:publisher>
    <foaf:Organization rdf:about="http://dbpedia.org/resource/Mondeca" />
  </dc:publisher>
</owl:Ontology>
```

VOAF vocabulary rights and properties information

Vocabulary Elements (Classes and Properties)

In addition to the formal definition of classes and properties constituting a vocabulary, some attributes must be present to ensure their proper use. A vocabulary element has a label and a comment to describe its role and purpose. Belonging of an element to its vocabulary eases its visualization and comprehension. Finally the status of an element is an important information that informs users of its stability.

Element	Usage
rdfs:label	Element main title
rdfs:comment	Role of the element
rdfs:isDefinedBy	Explicit link between an element and the namespace it belongs to
vs:term_status	Element status ("stable", "testing", "unstable", "deprecated") In case of "deprecated", the property "dc:isReplacedBy" may be used to identify the new element which replaces the deprecated one

```

<owl:Class rdf:about="http://purl.org/vocommons/voaf#VocabularySpace">
  <rdfs:label xml:lang="en">Vocabulary Space</rdfs:label>
  <rdfs:comment xml:lang="en">A vocabulary space defines any relevant grouping of vocabularies
  e.g., designed for similar purposes or domains, or designed by the same publisher or the same project, etc.
  A vocabulary can belong to zero, one or more vocabulary spaces.</rdfs:comment>
  <rdfs:comment>Dublin Core properties isPartOf and hasPart are used to link a vocabulary to a
  vocabulary space</rdfs:comment>
  <rdfs:isDefinedBy rdf:resource="http://purl.org/vocommons/voaf#" />
  <vs:term_status>testing</vs:term_status>
</owl:Class>

```

VOAF vocabulary elements metadata.

Documentation

To be interpretable by both machines and humans, a vocabulary must be made available in RDF format, but also HTML. The namespace of the vocabulary (the base URI) must provide access, ideally through a process of content negotiation, to the formal vocabulary and to its documentation.

There are now a number of applications that generate HTML documentation from the RDFS or OWL vocabulary description, which simplifies maintenance procedures. These tools include [neologism] [OntoSpec] [LODE] and [Parrot]. It's always good to find more of this HTML reproduction Automatic formal content, presentation documents indicating the general context of the vocabulary, as part of its development, using the data, the prospects for further development or maintenance, current Trustee, if any, items or contributions to conferences etc citing the vocabulary. A graphical form of UML diagram or equivalent is also a valuable service to the overall understanding of the vocabulary.

The document [W3C-BP-Publishing-RDF-Vocabs] clearly describes how to set up content negotiation for publishing vocabularies on the Web. The tool [Vapour] is designed to validate of the process that allows machines to access the formal document and humans to its documentation.

Content negotiation should refer to the latest version of the vocabulary while previous versions can be stored for archive, to allow backward compatibility with applications using earlier versions.

Conclusion

Adding those metadata does not represent a big effort (you might start from an existing vocabulary like VOAF): a small effort for your vocabulary but a big added value for the SW community.

Little semantics goes a long way [Hendler 1997].

References

- [Heath and Bizer, 2011] Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space: Theory and Technology. Volume 1. Morgan & Claypool Publishers (2011).
- [Bizer et al., 2009] Bizer, C., Heath, T., Berners-Lee, T.: Linked data-the story so far. Int. J. Semantic Web Inf. Syst. 5 (2009) 1-22.
- [LOV-Metadata] http://labs.mondeca.com/dataset/lov/details/vocabularySpace_Metadata.html
- [neologism] <http://neologism.deri.ie/>
- [OntoSpec] <http://moustaki.org/ontospec/>
- [LODE] <http://www.essepuntato.it/lode>
- [Parrot] <http://ontorule-project.eu/parrot/parrot>
- [Vapour] <http://validator.linkeddata.org/vapour>
- [W3C-BP-Publishing-RDF-Vocabs] <http://www.w3.org/TR/swbp-vocab-pub/>
- [Hendler 1997] <http://www.cs.rpi.edu/~hendler/LittleSemanticsWeb.html>