Report of the TDWG Vocabulary Management Task Group (VoMaG) v1.0

Title: Report of the TDWG Vocabulary Management Task Group (VoMaG) v1.0

Date issued: 2013-09-20

Date modified: 2013-09-20

Abstract: This report constitutes the outcome of the TDWG Vocabulary Management Task Group (VoMaG). It consists of an introduction and four sections. The introduction describes the remit of VoMaG and the three task areas that emerged based on its charter, each of which is dealt with in separate sections. The first area (Section 2) relates to the status of the TDWG Ontology and its relationship to existing TDWG standards. It also examines the Darwin Core Namespace Policy as a means to maintain technical specification standards that define vocabularies, and provides recommendations for actions that would resolve several long-standing issues. The second area (Section 3 and Section 4) deals with the use of Semantic MediaWiki as a community platform for developing and maintaining vocabularies and with the NCBO BioPortal as a tool for sharing ontologies. It looks at the core features provided by MediaWiki, the enhancements provided by the Semantic MediaWiki extension, the implementation of the ViBRANT/GBIF terms wiki on Biowikifarm, and the advantages of BioPortal for ontologies (as opposed to vocabularies). The third area (Section 5) addresses the requirements for a framework for managing vocabularies, identifying the main components and a suggested workflow. Sections include recommendations for consideration by TDWG.

Editors: Steve Baskauf, Éamonn Ó Tuama, Dag Endresen, Gregor Hagedorn

Part of TDWG Standard: not part of any standard

Creator: TDWG Vocabulary Management Task Group (VoMaG)

Identifier: http://www.gbif.org/orc/?doc_id=5422

Document Status: Type 3 informational document

Contents

1 Introduction to VoMaG

The Vocabulary Management Task Group (VoMaG), convened under the Technical Architecture Group (TAG) of Biodiversity Information Standards (TDWG) was formally established at the TDWG 2012 conference in Beijing. The purpose and scope of VoMaG is briefly introduced here. For further information, please consult the charter[2]. Much of the prior, substantive work that informed VoMaG grew out of the related TDWG RDF/OWL task group[3] and the EU funded ViBRANT project[4] which includes a work package dedicated to biodiversity vocabularies.

VoMaG seeks to address the best practices for the collaborative development and maintenance of vocabularies used to describe biodiversity resources. Due to the high core membership (around 40, as of writing) of the group, it was decided that the main means of collaboration would, in the absence of a fully functioning TDWG wiki platform, be a dedicated section of the GBIF Community Site[5]. The latter site includes pages dedicated to the main themes outlined in the charter as well as records of any discussions around them. This document, synthesising the content on the VoMaG community site, constitutes the main deliverable of the group and provides a set of recommendations to TDWG on a number of topics relating to the management of vocabularies/ontologies.

Based on the charter which provides a broad outline of the kinds of areas that might be in scope for discussion by the task group, a set of three topics emerged which formed the main sections of this report. These are: i) a review of the TDWG ontologies including status, maintenance issues, etc., ii) a proposal for a platform based on Semantic MediaWiki for developing and maintaining vocabularies, and iii) a general framework for managing vocabularies covering standards, best practices, architecture, etc. In particular, we do not address the creation of new vocabularies or ontologies but rather the infrastructure to support these.

The terms vocabulary and ontology are often used somewhat loosely in an overlapping fashion. In this document, "vocabulary" is used for any list of terms including lists of term names and definitions, while the word "ontology" is reserved for a semantically richer description of terms expressed in a standard way such as the web ontology language (OWL). Our main focus is on vocabularies as simple groupings of concepts (terms) and their definitions without implying any semantic relationships amongst them. The goal is to enable convergence on a shared terminology in support of data integration and interoperability across the biodiversity domain. We view the development of such vocabularies as an independent activity and possible precursor to the more

---

[2] http://community.gbif.org/pg/file/read/28812/vocabulary-management-group-vomag-charter-v1
[3] http://code.google.com/p/tdwg-rdf/
[4] http://vbrant.eu/
[5] http://community.gbif.org/pg/groups/21382/

complex task of developing true ontologies which seek to model a domain of knowledge by identifying classes of entities and properties, and the relations among them, thus supporting knowledge representation and reasoning as an ultimate goal. The GBIF task group report on Knowledge Organisation Systems (Catapano et al., 2011) supports this approach.

Likewise, the words *term* and *concept* are used interchangeably following widespread use (Clarke and Zeng, 2012). We recognise that, strictly speaking concept refers to the abstract unit within the vocabulary and a concept can have many terms associated with it (e.g., the name of the concept in different languages). In this document, we explore the suitability of using skos:Concept to declare terms. Such skos:Concept entities can describe either property-terms or value-terms.

We also refer extensively to two ways of modelling vocabularies, namely, Resource Description Framework (RDF)[6] and Simple Knowledge Organization System (SKOS)[7]. RDF is a language/model for representing information about resources on the World Wide Web. SKOS is an application of RDF typically used to represent content schemes such as thesauri, taxonomies and other classification schemes. In the context of this article, we also note that the word "vocabulary" has two common meanings: i) for the set of property-terms in a metadata element set (e.g., Dublin Core, Darwin Core) and ii) for "value-vocabularies" that may provide the actual content for metadata elements. For example, the following states that the value of the Dublin Core metadata element (or property) "dc:language" is "eng", itself a vocabulary term from the content vocabulary/controlled value-list specified by ISO 639-3.
```
<dc:language rdf:datatype="http://purl.org/dc/terms/ISO639-3">eng </dc:language>.
```

Likewise, in the following, the value of the Dublin Core metadata element "dc:subject" is a URI for the concept "Agrobiodiversity" provided by the AGROVOC[8] thesaurus. AGROVOC is expressed in SKOS and the URI dereferences to the SKOS representation of the Agrobiodiversity concept
```
<dc:subject>http://aims.fao.org/aos/agrovoc/c_37977<dc:subject>.
```

As far as we are aware, SKOS has not been used for describing property-terms/metadata element sets (as opposed to content or value vocabularies) but we think a subset of the SKOS elements, at least, are useful for that purpose, e.g., skos:Concept, skos:prefLabel; skos:definition; skos:example.

2 Review of TDWG Ontologies

2.1 Introduction
In order to review the TDWG Ontologies, it is necessary to understand the process that TDWG has put in place for ratifying its standards. We subsequently distinguish between the TDWG "vocabularies" and TDWG "ontologies", consider the status of the ontologies, covering the various extant versions, their structure, origins, use and maintenance.

---

[6] http://www.w3.org/TR/rdf-primer/
[7] http://www.w3.org/TR/skos-primer/
[8] http://aims.fao.org/website/AGROVOC-Thesaurus/

2.2 TDWG Standard

Biodiversity Information Standards TDWG[9] is an organization which promotes the adoption of standards for the exchange of biodiversity data. Existing current standards[10] fall into two categories: Technical Specifications and Applicability Statements (see Status and Categories page[11]). TDWG standards must go through a formal process of development, expert review, public review, and approval by the TDWG Executive before they are ratified.

2.3 TDWG Vocabularies

There is no formal definition of a "vocabulary" in the context of TDWG standards. A TDWG vocabulary could be considered to be a standard of type "Technical Specification" whose purpose is to define terms that describe biodiversity resources. As of September 2013, there is one TDWG vocabulary which is ratified as a "Current Standard": Darwin Core[12] (Wieczorek et al. 2012). Vocabularies that are "Draft Standards" are: Audubon Core[13] which has completed public review and is undergoing final revisions, and Natural Collections Description (NCD)[14] whose development has stalled. Access to Biological Collection Data (ABCD)[15] is a "Current (2005) Standard" which is recommended for use but has not gone through the current standards ratification process. It is a specific XML schema rather than a general purpose vocabulary.

2.4 The TDWG Ontology

When TDWG made the deployment of Life Science Identifiers (LSIDs) a priority[16], the development of vocabularies expressed in RDF (Resource Description Framework; see Beginner's Guide to RDF[17]) was intended to progress in tandem with LSID adoption (see LSID Vocabularies[18]). These RDF vocabularies (written using a form of RDF called Web Ontology Language or OWL) were, in aggregate, known as "The TDWG Ontology" and were envisioned as an integral part of an overall Standards Architecture for TDWG (see TAG Wiki[19]). Although the TDWG Ontology's name includes the word "ontology", terms in its component vocabularies do not have extensive semantic restrictions and the well-defined relationships found in formal ontologies.

The parts of the TDWG Ontology were intended to be either based on an existing standard or to eventually be incorporated as part of an evolving standard (see Ontology Governance[20]). Thus the TaxonName and TaxonConcept ontologies were based on (but not actually part of) the Taxon Concept Transfer Schema (TCS), a "Current (2005) Standard", and the Institution and Collection ontologies were created as part of the development of the NCD "Draft Standard". However, although the Technical Architecture Group of TDWG sanctioned (and coordinated) its development, no part of the TDWG Ontology itself has actually become part of any ratified TDWG standard.

---

[9] http://www.tdwg.org/
[10] http://www.tdwg.org/standards/
[11] http://www.tdwg.org/standards/status-and-categories/
[12] http://rs.tdwg.org/dwc/
[13] http://terms.gbif.org/wiki/Audubon_Core_Term_List
[14] http://www.tdwg.org/standards/312/
[15] http://www.tdwg.org/standards/115/
[16] http://www.tdwg.org/standards/150/
[17] http://code.google.com/p/tdwg-rdf/wiki/Beginners
[18] http://wiki.tdwg.org/twiki/bin/view/TAG/LsidVocs
[19] http://wiki.tdwg.org/TAG
[20] http://wiki.tdwg.org/twiki/bin/view/TAG/TDWGOntologyGovernance

In developing the ontology, TDWG opted for a multi-layered model[21] consisting of a "BaseOntology" to describe fundamental abstract concepts, a "CoreOntology" for the most common and general concepts within the biodiversity domain addressed by TDWG, and multiple domain ontologies that describe particular classes of biodiversity resources, e.g., TaxonName, TaxonConcept. The domain ontologies have a sub-class relationship with base and core ontologies.

2.4.1 Versions of the ontology

The "TDWG Ontology" consists of a number of individual ontology documents written in OWL/RDF and dated mostly between 2008-05-08 and 2008-07-07. The individual ontologies have base URIs in the form "http://rs.tdwg.org/ontology/[something]#" where [something] is descriptive of the particular ontology. For example, "http://rs.tdwg.org/ontology/voc/TaxonConcept#" is the base URI for the TaxonConcept ontology.

If an ontology URI is dereferenced by a semantic client (i.e., one which requests content-type: application/rdf+xml), the particular document located at http://rs.tdwg.org/ontology/ is returned. These documents declare an XSL stylesheet "human.xsl" which reformats the data into a web page when the consuming client dereferences the URI requesting content-type: text/html (e.g. the client is a web browser). It is possible (although difficult) to surf through the complex TDWG website from the TDWG homepage through the largely defunct TDWG wiki and arrive at the web pages generated from these ontology documents.

The human.xsl stylesheet, dated 2009-12-09, notes that ontology development has moved to the Google Code site and that the documents found at http://rs.tdwg.org/ontology/ might not be the most recent version. At the Google Code site, the individual ontologies can be viewed at http://code.google.com/p/tdwg-ontology/source/browse/trunk/ontology/ and lower levels of the document tree. The ontology documents found at the Google Code site are managed as a Subversion repository and have a revision number of 19. Following is a comparison of the documents at the Google Code site and the http://rs.tdwg.org/ontology/ directory:

1 Most, if not all, documents appear to be identical between the two sites. Thus, it does not appear that any significant development happened at the Google Code site after the content was moved there.
2 The Google Code site has several additional documents with ".owl" extensions that correspond to other documents having ".rdf" extensions, namely: TaxonConcept.owl, TaxonName.owl, and TaxonRank.owl. These documents are not present on the TDWG site. The two sets of documents have exactly the same $Rev: numbers (1211) and Last Modified dates (2007-06-26 10:36:21 +0100). However, they do NOT have the same content. The .rdf versions appear to be the later documents because they include additional terms not found in the .owl versions and also contain subclass declarations linking them to the Core ontology document which are not found in the .owl versions.
3 The Google Code site does not contain the CSS and XSL files used to render the content as XHTML. It also does not contain several subdirectories (/tapir and /images) found on the TDWG site. The /images subdirectory does not appear to contain significant content but rather content needed for the rendering as XHTML.

---

[21] http://wiki.tdwg.org/twiki/bin/view/TAG/TDWGOntology

There is one additional copy of the ontology found at http://rs.tdwg.org/ontology2/. This version appears to be identical to the one present at the Google Code site and also indicates the same Subversion revision number (19).

2.4.2 Structure and origins

Terms in the ontologies have subclass relationships to terms in the Base and Core ontologies. Two ontologies (Digital Image and Procedure) have apparent errors caused by uncorrected cut and paste problems and are therefore probably not functional. The Taxon Concept and Taxon Name ontologies are based on the Taxon Concept Transfer Schema TDWG standard[22]. The Collection ontology classes represent classes from the draft TDWG Natural Collections Description (NCD) standard[23], although the base URIs disagree with those used in subproperty declarations in the Darwin Core normative RDF.

2.4.3 Use of the ontology

Because the implementation of LSIDs was never widespread, many parts of the ontology were never used widely. However, certain parts have been adopted for use for particular purposes. The Collection ontology terms are in use in the Biodiversity Collections Index[24] metadata. The terms tc:hasName and tc:accordingTo as well as several other terms from the TaxonName and TaxonConcept ontologies have been used by several metadata providers to specify the basic components of Taxon Concepts as they are described in TCS. Because of problems with the Darwin Core Taxon class terms, they are not suitable for use as object properties. Some terms from the TaxonName and TaxonConcept ontologies may be used in place of the Darwin Core Taxon terms to express Darwin Core data as RDF[25].

2.4.4 Maintenance of the ontology

Because the TDWG Ontology was envisioned as a core component of TDWG's technical architecture, its maintenance was to be overseen by the Technical Architecture Group (TAG) itself. The TAG consists of representatives of subgroups of TDWG and active members of task groups created by the TAG[26,27]. The TAG bears responsibility for ensuring that standards created by task groups are compatible with standards created by other task groups and with standards outside of TDWG. Unlike Darwin Core which has a Task Group charged with its maintenance, there is no group specifically charged with maintaining the ontologies. Therefore, the ontologies are in a generally unmaintained state.

The original vision that the various ontologies within the TDWG ontology would either become ratified standards or that ratified standards would become part of the broader TDWG ontology has not happened. Since the original development of the ontology, Darwin Core has emerged as the *de facto* backbone of the TDWG technical architecture. The development of an RDF guide for Darwin Core would allow it to serve some of the needs that the TDWG ontology was designed to serve.

---

[22] http://www.tdwg.org/standards/117/
[23] http://www.tdwg.org/standards/312/
[24] http://www.biodiversitycollectionsindex.org/
[25] http://code.google.com/p/tdwg-rdf/wiki/TaxonInRDF
[26] http://www.tdwg.org/activities/tag/
[27] http://www.tdwg.org/about-tdwg/process/

However, there are other parts of the ontology that fall outside of the scope of Darwin Core, leaving the TDWG ontology as the only existing means to express those concepts. For that reason, a critical examination should be undertaken to determine which parts of the TDWG ontology should be retained and which parts should be deprecated, and to delineate what group should be responsible for the maintenance of those parts that are retained.

2.5 The Darwin Core Namespace Policy as a mechanism for maintaining standards

In the absence of a formal mechanism for maintaining ratified TDWG standards, the Darwin Core Namespace Policy which defines a policy for how changes to Darwin Core terms are to be made is considered as a general process for wider adoption for all TDWG vocabularies.

2.5.1 Mechanism for maintenance of TDWG standards

The draft TDWG Standards Documentation Specification[28] states that "TDWG provides no formal versioning mechanism for standards. Once a standard has been ratified it cannot be changed in any substantive way; it must be superseded by a standard with a different name." Since the Documentation Specification itself has never been ratified as a TDWG standard, it is effectively a recommendation, not a requirement. Thus there is no official mechanism for maintaining TDWG standards.

2.5.2 What is the Darwin Core Namespace Policy?

In practice, it is unrealistic that a technical standard as complex as Darwin Core should go through a re-ratification process with every minor change. Therefore, Darwin Core chooses to ignore the recommendation of the draft Documentation Specification and has instituted a Namespace Policy[29] which describes the process by which Darwin Core can be changed.

Section 3 of the Namespace Policy describes the Term Change Policy for Darwin Core. In summary, minor editorial errata can be corrected immediately, as long as the "TDWG Architecture Group" (i.e. the TAG) does not object to the changes. Subject to TAG approval, substantive editorial errata will be corrected after public notice and a possible 30-day public comment period to ensure that existing systems and applications are not adversely affected. Requests for semantic changes to existing terms or for addition of new terms require a more lengthy process involving a public comment period, a possible official decision by the TAG[30], and the possible issuance of a new term URI.

The Darwin Core Namespace Policy has several advantages. Minor errors in the standard can be corrected rapidly without a protracted process. It also provides a mechanism for suggesting, tracking, evaluating, and implementing changes to particular terms without requiring a re-examination of the entire standard.

---

[28] http://www.tdwg.org/standards/147/
[29] http://rs.tdwg.org/dwc/terms/namespace/index.htm
[30] http://rs.tdwg.org/dwc/terms/history/decisions/index.htm

The Term Change Policy[31] process has been invoked several times since Darwin Core was ratified and thus demonstrates a mechanism by which a standard can evolve incrementally. It forms a model for maintaining a Technical Standard which is likely to be followed by subsequent standards. The group responsible for the proposed Audubon Core standard has plans to implement a similar policy, if one is not imposed by TDWG itself.

2.5.3 Darwin Core Namespace Policy in practice

The Namespace Policy specifies that requests for changes should be made to the TAG via its mailing list[32]. However, in practice, change requests are submitted via the Issue Tracker at the Darwin Core Google Code site[33].

The mechanism for public comment is not specified in the Namespace Policy. In practice, public comment on a proposed change has been announced on the tdwg-content email list[34]. The submitter of a proposed change may or may not have posted notification of the proposal on the tdwg-content list.

The Term Change Policy specifies that correction of errors and addition of terms will not result in "decisions" recorded at the Darwin Core Decision History page[35]. However, in practice, all substantive changes to the standard have been recorded there.

2.5.4 Disadvantages of the Darwin Core Namespace Policy as currently implemented

The discussion of changes via the tdwg-content email list can be protracted and confusing. Emails to the list are archived[36] but there is no summary of the discussion and it is difficult to follow the thread of the discussion without considerable work.

The Namespace Policy does not specify any timetable for action on proposed changes. Thus proposed changes may languish for months or years if there is no clear consensus. This delay may actually be considered an advantage as it may contribute to the stability of the standard by requiring a clear imperative to make changes.

According to the Namespace Policy, the TAG is to play a major role in the decision-making process regarding proposed changes. However, the composition and function of the TAG itself is not clear. This has led to misunderstandings and claims of "conspiracy" in the functioning of TDWG.

2.6 Recommendations

The following recommendations are proposed.

---

[31] http://rs.tdwg.org/dwc/terms/namespace/#classesofchanges
[32] http://lists.tdwg.org/mailman/listinfo/tdwg-tag
[33] http://code.google.com/p/darwincore/issues/list
[34] http://lists.tdwg.org/mailman/listinfo/tdwg-content
[35] http://rs.tdwg.org/dwc/terms/history/decisions/index.htm
[36] http://lists.tdwg.org/pipermail/tdwg-content/

2.6.1 The fate of the TDWG Ontologies

Recommendation 2.1. Given that the purpose of the TDWG Ontology being placed on Google Code was to facilitate development of the ontologies but that such development has not happened, clearly label the site as defunct to make it clear that no further development at that site will take place. If the decision is made to continue development of parts of the ontology, a link should be given on the Google Code site to the site where such development is occurring.

Recommendation 2.2. Since the documents located at http://rs.tdwg.org/ontology2/ appear to have no use and are apparently not linked to any TDWG web pages, they should be deleted. If it is felt that the URIs of these documents might be in use, use HTTP 301 redirects to the authoritative versions at http://rs.tdwg.org/ontology/.

Recommendation 2.3. Since the documents located at http://rs.tdwg.org/ontology/ are the ones which are served when the term URIs are dereferenced, and since the terms of at least some of the sub-ontologies are in active use, the various ontology documents should be maintained there indefinitely.

Recommendation 2.4. The documents having extensions .owl and which are found on the Google Code site (TaxonConcept.owl, TaxonName.owl, and TaxonRank.owl) should be clearly marked as deprecated using XML comments. These comments should note that it is the particular file (e.g. TaxonConcept.owl[37]) that is being deprecated and not the ontology (e.g. http://rs.tdwg.org/ontology/voc/TaxonConcept) itself. The syntax shown in the following example is recommended:

```
<owl:Ontology rdf:about="http:// rs.tdwg.org/ontology/voc/TaxonConcept.owl">
   <owl:deprecated
rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean">true</owl:deprecated>
   <dcterms:isReplacedBy rdf:resource="http://rs.tdwg.org/ontology/voc/TaxonConcept.rdf" />
   <rdfs:comment>The TaxonConcept ontology described by the document TaxonConcept.owl is
deprecated and replaced by TaxonConcept.rdf</rdfs:comment>
    ...
</owl:Ontology>
```

Recommendation 2.5. The code
```
<?xml-stylesheet type="text/xsl" href="human.xsl"?>
```
should be removed from all of the documents served from http://rs.tdwg.org/ontology/ . The XSL stylesheet produces XHTML that refers human viewers to the non-functional Google Code site and the navigation on the resulting XHTML references parts of the TDWG website that are effectively no longer in use (i.e. the LSID Vocabularies pages). More generally, the method of generating human-readable representations of RDF/XML using XSLT is not generally advisable because it requires the content-type of the file to be incorrectly reported as text/xml. RDF/XML documents with declared XSL stylesheets that are correctly identified by a server as content-type: application/rdf+xml are not transformed by most browsers and in some cases generate an error that prevents the underlying page source from being viewed. The standard 303 redirect method of content negotiation (http://www.w3.org/TR/cooluris/) is preferable.

---

[37] https://code.google.com/p/tdwg-ontology/source/browse/trunk/ontology/voc/TaxonConcept.owl

Recommendation 2.6. At a minimum, the page http://wiki.tdwg.org/twiki/bin/view/TAG/LsidVocs should indicate clearly that the LSID vocabularies and TDWG Ontology are no longer under active development. Ideally, all of the pages regarding the TDWG Ontology should also be annotated so that a visitor would easily know that ontology development is no longer occurring.

Recommendation 2.7. Update the modified date for each document that is changed to reflect the date of modification. Increment the version number at the 0.1 unit level.

Recommendation 2.8. Because there are existing well-known vocabularies (e.g. FOAF) that duplicate many parts of the ontology, deprecate all of the ontologies except the following ones which don't have well-known alternatives and appear to be in use by at least one provider:

> http://rs.tdwg.org/ontology/voc/TaxonName
> http://rs.tdwg.org/ontology/voc/TaxonConcept
> http://rs.tdwg.org/ontology/voc/TaxonRank
> http://rs.tdwg.org/ontology/voc/Common
> http://rs.tdwg.org/ontology/voc/Collection
> http://rs.tdwg.org/ontology/voc/CollectionType
> http://rs.tdwg.org/ontology/voc/Institution

The deprecation should be part of the RDF itself and should include comments that elaborate on the reason for the deprecation. The syntax shown in the following example is recommended:

```
<owl:Ontology rdf:about="http://rs.tdwg.org/ontology/voc/Team">
   <owl:deprecated
rdf:datatype="http://www.w3.org/2001/XMLSchema#boolean">true</owl:deprecated>
   <rdfs:comment>This ontology is no longer under development and is no longer recommended
for use</rdfs:comment>
     ...
</owl:Ontology>
```

In cases where it is clear that a particular part of the ontology has been superseded by another well-known vocabulary, the recommended replacement vocabulary should be noted in the comment associated with the deprecation property. The TDWG RDF Task Group should seek to identify candidate vocabularies that may be recommended for this purpose.

Recommendation 2.9. Remove all declarations of subclass relationship between http://rs.tdwg.org/ontology/Core#Concept, http://rs.tdwg.org/ontology/Base#BaseThing, http://rs.tdwg.org/ontology/Base#Name, or http://rs.tdwg.org/ontology/Base#DefinedTerm and classes defined in the ontologies listed in Recommendation 2.8. This makes it possible for those ontologies to continue to be used by existing applications without tying those terms to other parts of the ontology that are subject to deprecation.

Recommendation 2.10. The TDWG Executive should appoint a new review manager for the stalled Natural Collections Description (NCD) draft TDWG standard. The first task of the new manager would be to evaluate the status of the standard. The review manager should consider who the likely users are of the standard, whether the existing form of the draft standard (including the Collection,

CollectionType, and Institution ontology documents) is consistent with current best practices, and then recommend one of the following actions to the TDWG Executive:

- withdraw the standard from the standards track (i.e. kill it);
- submit the existing draft to another round of anonymous review;
- have the authors revise the existing draft as necessary then request permission to go to public review.

The Executive should assess the review manager's recommendation and respond to the review manager accordingly. The ultimate goal of this recommendation is to take the Collection, CollectionType, and Institution ontologies out of limbo and either transform them into a *bona fide* TDWG standard or kill them.

Recommendation 2.11. The TAG should form a Task Group which is charged with determining the fate of the TaxonName, TaxonConcept, and TaxonRank ontologies. The Task Group could consider a range of options, including:

- Stripping down the ontologies to create a minimal vocabulary that would maintain the critical terms (notably including tc:hasName and tc:accordingTo) which are in use and eliminate the terms that don't work as designed. The existing URIs should be maintained to avoid "breaking" applications that use them.
- Fixing the ontologies so that they reflect the spirit of the existing TCS standard while facilitating metadata transfer in a Linked Data/Semantic Web context.
- Creating a new TCS standard and deprecating the "Current (2005) Standard".

Alternatively, this task could be placed upon the RDF/OWL Task Group, but since the task includes consideration of the TCS Standard, additional participants representing taxonomic interests outside of the RDF group would probably need to be included.

The ultimate goal of this recommendation is to take the TaxonName, TaxonConcept, and TaxonRank ontologies out of limbo and transform them into a *bona fide* TDWG standard or replace them with something better.

2.6.2 General adoption of the Darwin Core namespace policy

Recommendation 2.12. The TDWG Executive should kill the stalled TDWG Standards Documentation Specification[38] as a proposal on the standards track while leaving the existing document as an informative document that provides guidance for the formatting of human-readable documents (at least until it is replaced with something else). The Vocabulary Management Task Group (VoMaG) should commission a new author team tasked with writing a new Standards Documentation Specification standard which will include the following components:

- Guidelines for the formatting of human-viewable documents (i.e., a replacement for the existing draft standards specification) which are likely to exist in forms such as web pages and PDF documents.
- Guidelines for the formatting of computer-readable documents as RDF. These guidelines would not be guidelines for the construction of RDF which specifies the nature of the relationships and properties included in the term definitions, but rather specify the

---

[38] http://www.tdwg.org/standards/147/

properties that a RDF document would contain to identify the nature of the document, version and modification metadata, licensing and provenance information, links to explanatory web pages and standards repository URLs, etc. that would prevent the kind of confusion that currently exists with the TDWG Ontology documents.

- Guidelines for the creation of ancillary websites that allow questions, comments, examples, etc. that would help in clarifying how the standard could be used. The Darwin Core Google Code site is an example of such a site.

The new Standards Documentation Specification standard should not contain information about the maintenance of standards (i.e. the evolution or versioning of the standard). That should be considered as a separate issue (as is addressed in recommendation 2.13 below).

Recommendation 2.13. The VoMaG should commission an author team charged with writing a general Vocabulary Maintenance Specification modelled after the Darwin Core Namespace Policy which would be applicable to all vocabularies. This specification should include a timetable for taking action on proposed modifications and guidelines for determining the extent of modifications that would trigger the release of an update of the standard. It should also clarify who is responsible for carrying out the maintenance of a ratified standard and mechanisms for ensuring that the required maintenance occurs (see also recommendation 5.10).

Recommendation 2.14. TDWG should maintain a publicly viewable list of members of the Technical Architecture Group (TAG). Membership can include those who are automatically members (task/interest group conveners), co-opted/appointed members, and volunteers. Once a year members would have to declare their interest in continuing to serve or be dropped from the list for presumed lack of interest. The tdwg-tag email list should be adjusted accordingly.

Recommendation 2.15. According to the TDWG Standards Development Process guidelines[39], each standard that is under development is the responsibility of a chartered Task Group. The guidelines state that each Task Group convener must submit to the Executive Committee an annual status report against its charter. If the Task Group is not fulfilling the requirements of its charter, the Task Group may be reorganized or disbanded.

In accordance with these guidelines, at least once a year (probably at the annual meeting), the Executive Committee should consider the progress of each standard that falls within the Draft Standards category. The Task Group's annual status report should assist the Executive Committee in making this assessment. If it appears that no progress has been made in the previous year, the Executive Committee should place the draft standard on a one-year probation and determine an appropriate course of action based on the reason(s) for lack of progress. Some reasons and courses of action include:

- If the Task Group convener is no longer interested in leading the Task Group, he/she should be replaced by another core member of the group who is willing to take on the responsibility.

---

[39] http://www.tdwg.org/about-tdwg/process/

- If the Review Manager is no longer interested in managing the review, he/she should be replaced by the Executive Committee.
- If the lead author of the standard document(s) is no longer interested in working on moving the standard forward, the sponsoring Task Group convener should appoint a new lead author.

If there is still no advancement after another year has elapsed, the draft standard should be killed and be removed from the list of draft standards[40]. If the Task Group has no other charge in its charter than the development of the killed standard, the Task Group should be disbanded.

Recommendation 2.16. At least once a year (probably at the annual meeting), the TAG should meet with the Task Group convener to review the progress of each unresolved proposed modification to an existing vocabulary standard. The following actions are possible:
- Move the proposal to public comment.
- Shelve the proposal for another year.
- Kill the proposal if it clearly does not merit adoption.

At the present, the only standard to which this would apply would be Darwin Core, but if the Darwin Core Namespace Policy were adopted as a general policy, it would also apply to Audubon Core and any other vocabulary that should be adopted in the future.

Recommendation 2.17. Modify the Darwin Core Namespace Policy regarding Decisions to reflect the actual practice of recording all decisions including the addition of new terms.

Recommendation 2.18. Modify the Darwin Core Namespace Policy so that it reflects the actual practice of using an issue tracker rather than emailing the tdwg-tag list.

3 Semantic MediaWiki as a platform for developing and maintaining vocabularies

3.1 Introduction
A vocabulary management system must support many tasks, among them, the definition, annotation, discussion, translation and export of terms or concepts in standard formats. In particular, it should be capable of importing externally defined vocabularies to enable their re-use. At the implementation level, a web-based platform is highly desirable to support remote collaboration. Based on prior investigations[41] for the ViBRANT project, it was decided to develop a prototype using MediaWiki[42], the wiki engine that underpins Wikipedia, in conjunction with the Semantic MediaWiki[43] extension which allows data items to be tagged and queried within the wiki

---

[40] http://www.tdwg.org/standards/
[41] http://vbrant.eu/sites/vbrant.eu/files/ViBRANT_D4.2b — Ontology Tools- Status report 2012.pdf
[42] http://www.mediawiki.org/
[43] http://semantic-mediawiki.org/

pages. As MediaWiki is maintained by a large Open Source community and used for Wikipedia and other Wikimedia Foundation projects (e.g., Wikisource, Wikispecies, Wikidata) there is every possibility of long-term sustainability.

3.2 Features of MediaWiki

By its wiki nature, MediaWiki provides an excellent collaborative platform. Notable additional features include:

- Saving of page versions (based on edits) allowing them to be compared.
- A discussion page associated with each content page.
- A "recent changes" feature that lists all changes and a "contributions" feature that allows viewing of an individual's contributions.
- A "watch list" allowing one to keep track of changes to particular pages; an email alert is also available.

3.3 Enhancement provided by Semantic MediaWiki

The Semantic MediaWiki extension turns MediaWiki into a collaborative database by allowing the textual content in the wiki pages to be tagged, thus enabling richer organisation, evaluation and re-use (e.g. by re-grouping) of the marked-up elements. The tags provide a form of annotation and can carry rich semantics when drawn from well known vocabularies or ontologies. Functions provided by the Semantic MediaWiki extension include:

- Semantic data annotation.
- The import/re-use of external vocabularies.
- The exchange of data using RDF/XML and other formats.
- Facility to browse wiki content and display it visually in various ways (e.g., word clouds, charts, graphs, maps, tree lists, etc.).
- Simplified content entry / editing using web forms.

3.4 Advantages of Semantic MediaWiki as vocabulary platform

In summary, Semantic MediaWiki offers the following advantages as a platform for developing and managing vocabularies:

- Enabling (remote) collaboration by a community of practitioners for developing vocabulary concepts.
- Ease of contribution of concepts (labels, definitions, etc.) using web forms and their subsequent grouping into various concept schemes (vocabularies).
- Translation of terms into multiple languages, facilitated by provision of web forms.
- Support for import of external vocabularies thereby encouraging re-use of concepts.
- Facility to export vocabularies in a variety of formats including RDF.

3.5 The ViBRANT/GBIF Terms Wiki

An instance of Semantic MediaWiki known as the ViBRANT/GBIF Terms Wiki is hosted by Biowikifarm[44] and accessible at http://terms.gbif.org. This takes advantage of the features outlined above to provide a collaborative platform for managing vocabularies. The wiki has been populated with several test vocabulary examples including Darwin Core and Audubon Core. Each concept is provided with its own page which includes label, definition and examples as well as any translations.

---

[44] http://biowikifarm.net/

Concepts are organised into collections or more formally into concept schemes (vocabularies). Pages for collections and schemes are automatically generated from tagging concepts with membership attributes. Thus the page for the Darwin Core concept scheme (vocabulary) (http://terms.gbif.org/wiki/Darwin_Core) and the Darwin Core GeologicalContext collection (http://terms.gbif.org/wiki/Darwin_Core_GeologicalContext) are composed directly from individual concept pages such as: Earliest Age Or Lowest Stage (http://terms.gbif.org/wiki/dwc:earliestAgeOrLowestStage). The concept scheme page (Figure 1) allows browsing of the constituent concepts (e.g., by concept name, by label) for the scheme as a whole and by contributing collection and provides output of the scheme in both SKOS and RDF.

The biowikifarm hosting the terms wiki is presently run at the BGBM in Berlin, Germany and the support policy is backed by commitments from further Natural History Institutions (Staatliche Naturwissenschaftliche Sammlungen Bayerns, Munich; Museum für Naturkunde, Berlin). Its longevity is therefore backed by 3 institutions active in TDWG. The hosting service can therefore be reasonable expected to be available to TDWG for a significant time.

Figure 1. The Darwin Core concept scheme page of the ViBRANT/GBIF terms wiki allows browsing of the constituent concepts (e.g., by concept name, by label) and provides output of the scheme in both RDF and SKOS.

3.6 Recommendations

Recommendation 3.1. TDWG should adopt the prototype ViBRANT/GBIF Terms Wiki (the Semantic MediaWiki (Krötzsch et al. 2007) instance at http://terms.gbif.org which is hosted by Biowikifarm), rebranding it as http://terms.tdwg.org and maintaining it as the preferred (but not exclusive) community platform for developing and maintaining TDWG vocabularies.

4 BioPortal as a platform for sharing biodiversity ontologies

4.1 Introduction

The NCBO BioPortal[45] provides a software platform for publishing ontologies used in biology and biomedical research. This platform was developed for sharing ontologies expressed using the web ontology language (OWL) (Noy et al., 2009; Whetzel et al., 2011). Participants from the Biodiversity Information Standards (TDWG) Technical Architecture Group (TAG), Biological Science Collections (BiSciCol)[46] and the Genomics Standards Consortium (GSC) agreed at a joint workshop[47] (in May 2012) to establish a so-called "slice" in the NCBO BioPortal for Biodiversity KOS and ontology resources. The VoMaG convenors presented these plans at a webinar[48] organized by the National Center for Biomedical Ontologies (NCBO) in October 2012. The BiSciCol team and NCBO established the new biodiversity slice[49] in BioPortal in March 2013.

4.2 Advantages of using the NCBO BioPortal for biodiversity ontologies

The NCBO BioPortal provides a specialized platform (repository) for publishing and visualizing ontologies. At present (September 2013), BioPortal includes more than 350 ontologies published from a wide range of different user groups including, e.g., the biomedical communities, international agricultural research, Open Biomedical Ontologies (OBO) Foundry (Smith et al., 2007), and the World Health Organization (WHO)[50]. The so-called BioPortal "slices" will group ontologies from different user groups together, but you may create cross-linking and annotation for individual ontology entities between any ontology in BioPortal. These features of BioPortal can be used for integration of biodiversity terminology and the terminology from other biological sciences and user groups.

4.3 Recommendations

Recommendation 4.1. For developing and publishing ontologies based on the TDWG vocabularies, the newly established Biodiversity Information Standards "slice" (http://bis.bioportal.bioontology.org/ontologies) of the NCBO BioPortal should be used.

---

[45] http://bioportal.bioontology.org/
[46] http://biscicol.org/
[47] http://biocodecommons.org/workshops/sob.html
[48] http://www.bioontology.org/GBIF-vocabulary-management-for-biodiversity-informatics
[49] http://bis.bioportal.bioontology.org/ontologies
[50] http://www.bioontology.org/wiki/index.php/BioPortal_FAQ?pop=true#Ontology_Groups

5 A framework for managing vocabularies

5.1 Introduction

This section introduces the essential elements in a workflow that help ensure that a vocabulary, once developed by a community, is expressed in a well-known format and made discoverable in an appropriate repository for use by others.

The management routines for Darwin Core seem to work well given the successful uptake of this standard. Another proven successful guiding principle for the management of biodiversity vocabularies is the Dublin Core[51] vocabulary from the Dublin Core Metadata Initiative (DCMI). The DCMI vocabulary management system task group[52] recently came up with seven requirements for a DCMI vocabulary management system[53] that can be summarized as (1) portability, (2) version control, (3) open and established infrastructure, (4) support for multiple languages and scripts, (5) reuse of existing URIs and PURLs for all properties, classes and value vocabularies, (6) easy maintenance and revision, (7) interoperability and RESTful APIs.

DCMI also addressed the advent and continued expansion of the Linked Open Data world and its need for metadata expressed in RDF brings additional challenges for vocabulary management, particularly around the danger of proliferating vocabularies (Dunsire et al., 2012). There is a requirement for a framework for supporting discovery, evaluation and reuse of vocabularies. This has informed the current priorities of the DCMI Vocabulary Management Community which include best practice guidelines for vocabulary evaluation, selection, reuse; vocabulary preservation; best practice for extension of vocabularies; and developing and managing multi-lingual vocabularies.

The GBIF Knowledge Organization System (KOS) task group (Catapano et al. 2011, Lapp et al. 2011) provided recommendations for the uptake of KOS technologies for biodiversity information management. The KOS report recommended the issuance of persistent Globally Unique Identifiers for all terms and relations. It also recommended the use of SKOS as a minimal mechanism for expressing and sharing multi-lingual vocabularies. It highlighted the simplicity of SKOS as useful for including domain experts without experience in ontology development while also noting that this simplicity limited the expressiveness required to support semantic reasoning.

The CSIRO Solid Earth and Environment GRID is one example of a vocabulary service[54] using SKOS/RDF. It also provides guidelines on formalising a SKOS vocabulary[55].
The W3C is developing DCAT[56], a specification for an RDF vocabulary for describing data catalogues published on the web. There appears to be some impetus behind this as the European Commission, in turn, is developing a DCAT application profile[57] for data portals in Europe and will identify which DCAT properties to use, which controlled vocabularies, and the minimum metadata element set for describing public sector datasets. The DCAT specification itself could provide a model to follow for

---

[51] http://dublincore.org/documents/dcmi-terms/
[52] http://wiki.dublincore.org/index.php/Vocabulary_Management_System_Task_Group
[53] http://wiki.dublincore.org/index.php/Draft_Vocabulary_Management_System_Task_Group_Report
[54] https://www.seegrid.csiro.au/wiki/Siss/
[55] https://www.seegrid.csiro.au/wiki/Siss/VocabularyFormalizationInSKOS
[56] http://www.w3.org/TR/vocab-dcat/
[57] https://joinup.ec.europa.eu/asset/dcat_application_profile/description

describing the vocabulary catalogue (a type of data catalogue) of a repository and thus inform development of the TDWG vocabulary management framework. The elements required to describe a concept vocabulary would overlap with those required to describe a dataset but would most likely require some additional metadata elements from dedicated vocabularies such as the Linked Open Data vocabularies.

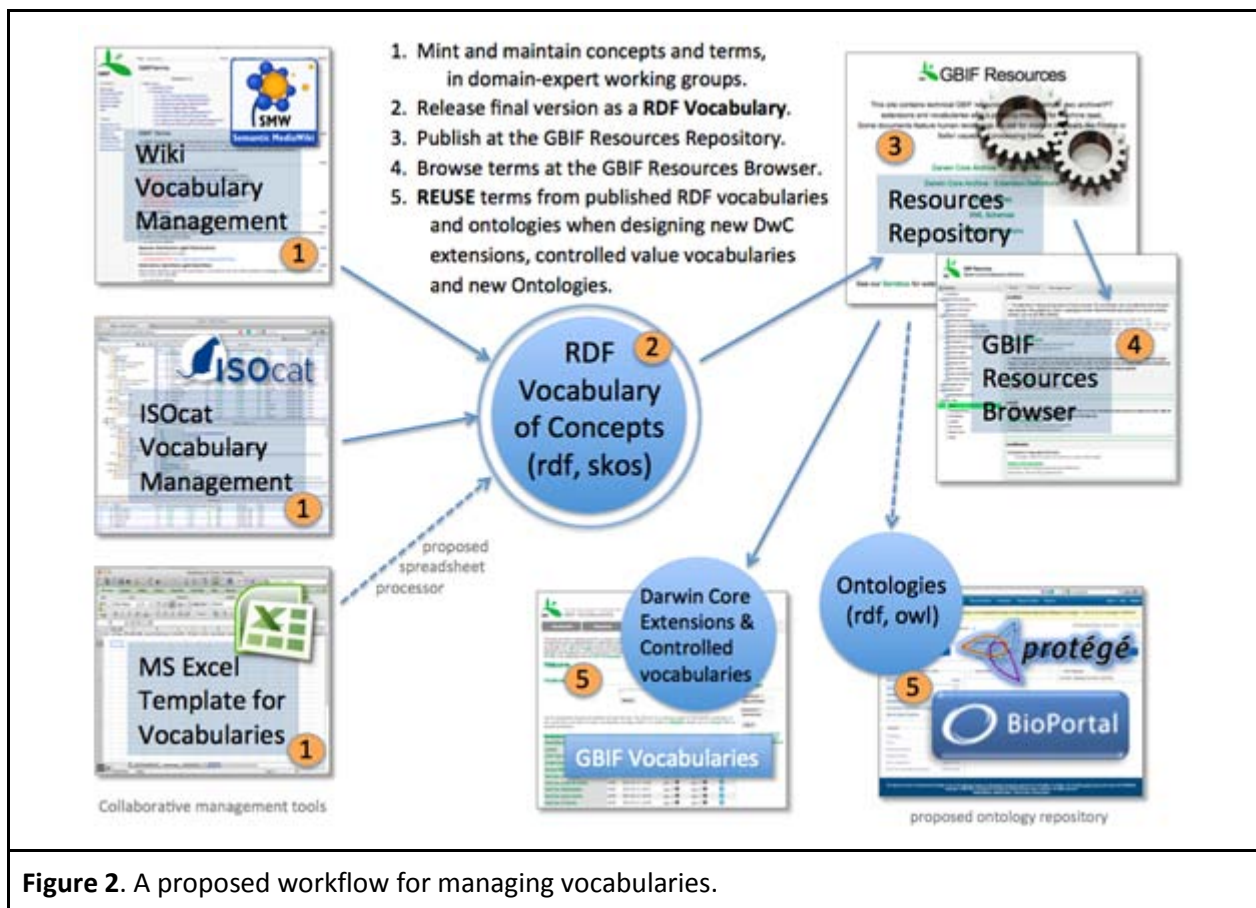## 5.2 A workflow for managing vocabularies

While the GBIF resources repository[58] (Figure 2) is used in an example of a workflow for managing vocabularies, the general principles apply to all repositories. The following steps are envisaged:

1. Domain expert groups develop vocabularies using whatever platform/tool they choose. These tools can range from spreadsheet templates to dedicated systems such as the ViBRANT/GBIF Terms Wiki[59].

2. Regardless of the tool used to develop the vocabulary, it must have a machine-readable expression, e.g., in SKOS/RDF.

3. The machine-readable expression (e.g., SKOS/RDF) is published (made available for discovery and download) in one or more repositories. To aid discovery, the vocabulary file should include metadata that describes the main characteristics of the vocabulary. Thus, the vocabulary itself and its constituent terms all require globally unique, persistent URIs.

4. A resources browser acts as an interface to a repository and allows a user to search/browse for vocabularies and their terms.

5. Client applications re-use the terms from the published machine-readable expression, e.g., in developing new Darwin Core extensions for the GBIF Integrated Publishing Toolkit, or in developing ontologies using systems such as the BioPortal.

---

[58] http://rs.gbif.org/
[59] http://terms.gbif.org/

Figure 2. A proposed workflow for managing vocabularies.

5.3 Recommendations

Recommendation 5.1. A TDWG vocabulary must have a globally unique namespace. The namespace can, for example, follow the Darwin Core practice of using the TDWG domain (http://rs.tdwg.org/dwc/terms/) or the Dublin Core practice of using PURL (http://purl.org/dc/terms/). It is not recommended to use the namespace bound to a particular application/technology such as the GBIF/TDWG terms wiki (http://terms.gbif.org/VOCABULARY/) as the identifier for the concept vocabulary.

Recommendation 5.2. Each term within a TDWG vocabulary must have its own identifier consisting of a dereferenceable URI.

Recommendation 5.3. As part of its documentation, a vocabulary must include a machine-readable expression such as SKOS or plain RDF.

Recommendation 5.4. Translations of a vocabulary to other (natural) languages should be made available in a separate machine-readable expression such as SKOS/RDF which is based on the main machine-readable expression in, e.g., RDF.

Recommendation 5.5. As part of its documentation, a vocabulary must include machine-readable metadata that describe the main characteristics of the vocabulary. Such metadata can be expressed e.g. in RDF. We strongly recommend using the minimal metadata element set proposed for describing Linked Open Data vocabularies[60]. These are listed in Annex 1, Table 1.

Recommendation 5.6. In addition, each term/concept within a vocabulary should, at a minimum, have properties for label, definition and parent vocabulary, expressed in either RDF or SKOS. These are listed in Annex 1, Tables 2 and 3.

Recommendation 5.7. Once ratified, the machine-readable expression of a TDWG vocabulary must be published in the TDWG resources repository at http://rs.tdwg.org/.

Recommendation 5.8. If a translation of a TDWG vocabulary is available, and endorsed by TDWG, it must be published in the TDWG resources repository.

Recommendation 5.9. TDWG should provide a "resources browser" interface to its repository at http://rs.tdwg.org/. The purpose of the browser is to support re-use by allowing a user to search/browse the vocabularies in the repository, to look up terms and their definitions, including translations. Several example applications/prototypes are available including the GBIF Resources Browser, the GBIF Term Browser and the Linked Open Data Vocabularies.

Recommendation 5.10. A ratified TDWG vocabulary must publish a list of experts who are responsible for maintaining and evolving the vocabulary in response to community feedback.

Recommendation 5.11. All the TDWG concept vocabularies must be published under an open license such as the Creative Common licenses. CC0[61] is recommended whenever possible.

Recommendation 5.12. TDWG should re-use terms (and their identifiers) from other published vocabularies rather than re-inventing them.

Acknowledgements

---

[60] http://lov.okfn.org/dataset/lov/Recommendations_Vocabulary_Design.pdf
[61] http://creativecommons.org/publicdomain/zero/1.0/
[62] http://community.gbif.org/pg/groups/21382/

References

Catapano T, Hobern D, Lapp H, Morris RA, Morrison N, Noy N, Schildhauer M, and Thau D (2011). Recommendations for the use of knowledge organization systems by GBIF. Released on 4 February 2011. Global Biodiversity Information Facility (GBIF), Copenhagen. Available at http://www.gbif.org/orc/?doc_id=2942&l=en, verified 23 April 2013.

Clarke SGD and Zeng ML (2012). Standard Spotlight: From ISO 2788 to ISO 25964: the evolution of thesaurus standards towards interoperability and data modeling. Information Standards Quarterly 24(1): 20-26. DOI: 10.3789/isqv24n1.2012.04

Dunsire G, Harper C, Hillmann D, and Phipps J (2012). Linked Data Vocabulary Management: Infrastructure Support, Data Integration, and Interoperability. Information Standards Quarterly, 24(2/3): 4-13. DOI: 10.3789/isqv24n2-3.2012.02

Krötzsch M, Vrandečić D, Völkel M, Haller H, Studer R (2007). Semantic Wikipedia, Web Semantics: Science, Services and Agents on the World Wide Web 5(4): 251-261. ISSN 1570-8268. DOI: 10.1016/j.websem.2007.09.001

Lapp H, Morris RA, Catapano T, Hobern D, and Morrison N (2011). Organizing our knowledge of biodiversity. Bulletin of the American Society for Information Science and Technology 37(4): 38-42. DOI: 10.1002/bult.2011.1720370411

Noy NF, Shah NH, Whetzel PL, Dai B, Dorf M, Griffith N, Jonquet C, D.L. Rubin, Storey M-A, Chute CG, and Musen MA (2009). BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Research 37: 170–173. DOI: 10.1093/nar/gkp440 PMCID: PMC2703982.

Smith, Barry, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J Goldberg, et al (2007). The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration. Nature Biotechnology 25: 1251–1255. doi:10.1038/nbt1346.

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012). Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE 7(1): e29715. DOI: 10.1371/journal.pone.0029715

Whetzel PL, Noy NF, Shah NH, Alexander PR, Nyulas C, Tudorache T, and Musen MA (2011). BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic Acids Research 39: 541–545. DOI: 10.1093/nar/gkr469.

Annex 1. Vocabulary metadata

The metadata elements in Table 1 are recommended as a minimal set for describing Linked Open Data Vocabularies[63]

Table 1. Metadata elements recommended for describing Linked Open Data vocabularies.

Identification	voaf:Vocabulary	Indicates that a vocabulary is the resource being described by the metadata
	vann:preferredNamespecePrefix	The prefix for the vocabulary namespace
	vann:preferredNamespaceUri	The namespace for the vocabulary
Title and Description	dct:title	The title (label) of the vocabulary
	dct:description	A description of the purpose of the vocabulary
Version and modification	dct:issued	The date when the vocabulary was first published
	dct:modified	The date when the vocabulary was last modified
	owl:versionInfo	The current version of the vocabulary
	rdfs:comment	Comments or further information about the vocabulary
Rights and property	dct:rights	Intellectual property rights
	cc:license	Copyright license
	dct:creator	Creator of the vocabulary
	dct:contributor	Contributor to the vocabulary
	dct:publisher	Publisher of the vocabulary

Each element within a vocabulary should, at a minimum, be described with the following properties:

Table 2. A minimal set of vocabulary element properties expressed in RDF.

Vocabulary element properties	rdfs:label	The label (title) of the element
	rdfs:comment	The definition of the element
	rdfs:isDefinedBy	The namespace (URI) of the vocabulary the element belongs to

Alternatively, more expressive vocabulary element properties are available in SKOS:

Table 3. A minimal set of vocabulary element properties expressed in SKOS.

Vocabulary element properties	skos:prefLabel	The label (title) of the element
	skos:definition	The definition of the element
	skos:inScheme	The namespace (URI) of the vocabulary the element belongs to
	skos:example	An example of the defined term

---

[63] http://lov.okfn.org/dataset/lov/Recommendations_Vocabulary_Design.pdf

The prefixes used in tables 1, 2 and 3 refer to the following namespaces.

Prefix	Namespace URI	25
cc	http://creativecommons.org/ns#	
dct	http://purl.org/dc/terms/	
foaf	http://xmlns.com/foaf/0.1/	
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#	
skos	http://www.w3.org/2004/02/skos/core#	
vann	http://purl.org/vocab/vann/	
voaf	http://purl.org/vocommons/voaf#	