



ViBRANT status report (M4.39)

Deliverable D4.2, Ontology Tools

M4.39 - Term browser

Develop a term browser for improved discovery of vocabulary terms used for documentation of biodiversity information resources.

Leading partner: Global Biodiversity Information Facility (GBIF)

Compiled by: Éamonn Ó Tuama and Dag Endresen

August 2012

GBIF Term Browser

The prototype application (<http://kos.gbif.org/termbrowser/>) allows a user to browse for terms defined in widely used vocabularies such as Darwin Core, Dublin Core, FOAF, etc., including, where available, translations. The idea is to allow someone to quickly see if a term already exists so they can re-use it rather than minting it afresh. This application is purely for searching and browsing and does not replace the need for tools such as ISOcat¹ and SemanticMediaWiki² as collaborative platforms for actually creating vocabularies (and translations).

The term browser application is based on Solr³, the technology also used for the GBIF occurrence/names index. It allows highly flexible, faceted searching⁴ (Figure 1). The facets provided are vocabulary, type (class or property) and language.

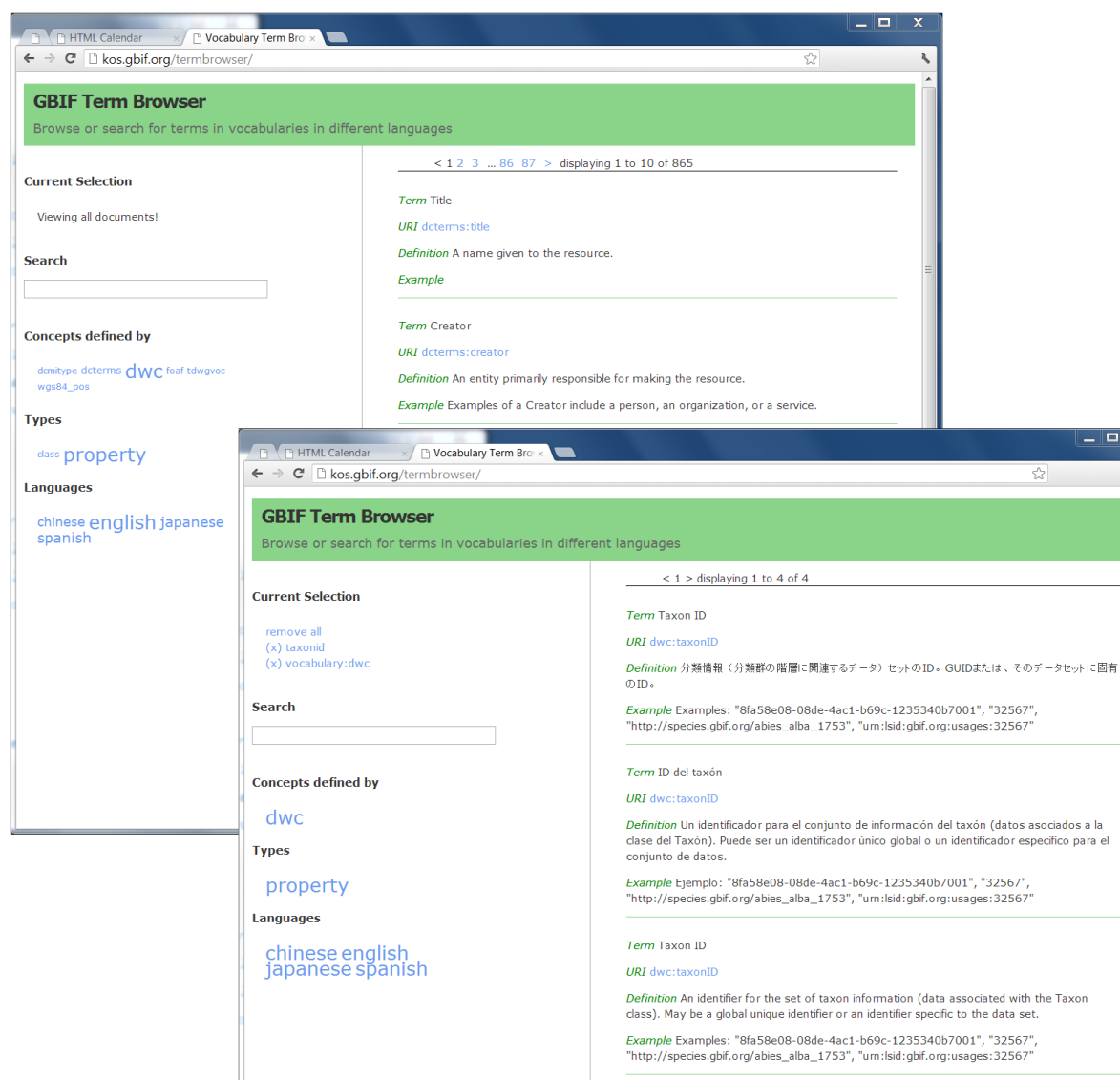


Figure 1. The GBIF term browser provides an interface for faceted searching. For example, in the upper screen, no filter has been applied and all 865 terms are available for browsing. In the lower right screen, all Darwin Core terms containing the string "taxonid" have been selected. If required, a further language filter could then be applied.

¹ <http://kos.gbif.org/isocat/interface/>

² <http://semantic-mediawiki.org/>

³ <http://lucene.apache.org/solr/>

⁴ http://en.wikipedia.org/wiki/Faceted_search

Populating the term browser

To populate the term browser with the vocabulary terms, the appropriate authoritative published definitions (in RDF) are transformed into a simpler XML format using XSLT for indexing by Solr (Figure 2). Currently, the term browser holds six vocabularies:

Darwin Core: <http://rs.tdwg.org/dwc/rdf/dwcterms.rdf>

DC Terms: <http://dublincore.org/2012/06/14/dcterms.rdf>

DCMI Type Vocabulary: <http://dublincore.org/2012/06/14/dctype.rdf>

FOAF: <http://xmlns.com/foaf/spec/index.rdf>

WGS84_pos: http://www.w3.org/2003/01/geo/wgs84_pos

Tcore (TDWG vocabularies): <http://rs.tdwg.org/ontology/voc/Collection.rdf>

```
<?xml version="1.0" encoding="UTF-8"?>
<vocabulary xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:dwcattributes="http://rs.tdwg.org/dwc/terms/attributes/">
  <concept>
    <language>English</language>
    <type>property</type>
    <label xml:lang="en">Accepted Name Usage</label>
    <name namespace="dwc:acceptedNameUsage">http://rs.tdwg.org/dwc/terms/acceptedNameUsage</name>
    <definition xml:lang="en">The full name, with authorship and
      date information if known, of the currently valid (zoological) or
      accepted (botanical) taxon.</definition>
    <example xml:lang="en">Example: "Tamias
      minimus" valid name for
      "Eutamias minimus"</example>
    <vocabulary namespace="http://rs.tdwg.org/dwc/terms/">DwC</vocabulary>
  </concept>
  <concept>
    <language>English</language>
    <type>property</type>
    <label xml:lang="en">Accepted Name Usage ID</label>
    <name namespace="dwc:acceptedNameUsageID">http://rs.tdwg.org/dwc/terms/acceptedNameUsageID</name>
    <definition xml:lang="en">An identifier for the name usage
      <?xml version="1.0" encoding="UTF-8"?>
      <vocabulary xmlns:skos="http://www.w3.org/2004/02/skos/core#"
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dcterms="http://purl.org/dc/terms/"
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:dwcattributes="http://rs.tdwg.org/dwc/terms/attributes/">
        <concept>
          <language>Chinese</language>
          <type>property</type>
          <label xml:lang="zh-Hans">公认使用名称</label>
          <name namespace="dwc:acceptedNameUsage">http://rs.tdwg.org/dwc/terms/acceptedNameUsage</name>
          <definition xml:lang="zh-Hans">目前有效 (动物学) 的或公认 (植物学) 的分类单元全称, 如已知来源和日期信息则需注明.</definition>
          <example xml:lang="en">Example: "Tamias
            minimus" valid name for
            "Eutamias minimus"</example>
          <vocabulary namespace="http://rs.tdwg.org/dwc/terms/">DwC</vocabulary>
        </concept>
        <concept>
          <language>Chinese</language>
          <type>property</type>
          <label xml:lang="zh-Hans">公认使用名称编号</label>
          <name namespace="dwc:acceptedNameUsageID">http://rs.tdwg.org/dwc/terms/acceptedNameUsageID</name>
          <definition xml:lang="zh-Hans">目前有效 (动物学) 的或公认 (植物学) 的分类单元的名称使用 (根据来源所取名称的存档记录方法) 标识符.</definition>
          <example xml:lang="en">Example:
            "8fa58e08-08de-4ac1-b69c-1235340b7001"</example>
          <vocabulary namespace="http://rs.tdwg.org/dwc/terms/">DwC</vocabulary>
        </concept>
      </concept>
    </concept>
  </concept>

```

Figure 2. The authoritative version of each vocabulary added to the term browser is transformed from the original RDF to a simpler XML format for indexing by Solr.

Translations

The current SKOS version of the translations of Darwin Core terms to Chinese, Japanese and Spanish is available at this URL:

http://rs.gbif.org/terms/dwc/dwc_translations.rdf.