



ANDS Prototype Controlled Vocabulary Service

Last Updated: 14 December 2011

Background.....	1
Overview.....	2
Technical Architecture.....	2
Services.....	3
Service Contact.....	5

Background

ANDS Prototype Controlled Vocabulary Service

The ANDS controlled vocabularies prototype service will allow research organisations to create, manage and query "controlled vocabularies" relating to their research datasets. The service has now been deployed as a rudimentary and experimental prototype that provides query access to the widely used ABS ANZSRC Field of Research Codes. Feedback is sought to inform a cycle of quick iterations over the next few months.

The service is designed for Australian organisations managing research datasets that require use of controlled vocabulary terms in their data or metadata values. Use of controlled vocabularies in the metadata fields to describe research data collections improves access, findability and reuse of research data.

ANDS and ABS Collaborative Project

The ANDS controlled vocabularies service is a joint undertaking with the ABS (Australian Bureau of Statistics). ABS are the custodians of the widely referenced ANZSRC FOR (The Australian and New Zealand Standard Research Classification, Fields of Research. Catalog no. 1297.0) codes. ABS and ANDS are collaboratively developing the service and working to make the service compatible with related current and future services at ABS and other agencies.

Service Timeline

ANDS and ABS expect the prototype controlled vocabularies service to be refined through several prototype iterations in Q1 and Q2 2012 before entering a formal pilot phase in Q2 2012 and transitioning to a full production

service in late 2012.

Overview

Service Development

Currently, the prototype controlled vocabulary service provides query access to the ANZSRC FOR codes. In future releases the service will provide access to other controlled vocabularies and also provide facilities for vocabulary owners to create and manage vocabularies through the service.

The prototype service is a M-2-M (machine to machine) service, but future versions will be matched with an accessible web interface.

Service Design

The prototype controlled vocabularies service development is informed by a Vocabulary Services Technical Working Group (TWG) convened by ANDS and the Office of Spatial Policy and staffed by experts and interested parties from ABS and other organisations managing research data collections.

Service Profile

The current prototype controlled vocabulary service is a machine-to-machine service: i.e. there is no web page for individual researchers to visit to query or create controlled vocabulary items. There is only a software interface that research software or data archive software systems can call as part of their automated data management workflows.

Quick updates and alternative prototypes are flagged over the next few months which will include new web service and graphical user interfaces.

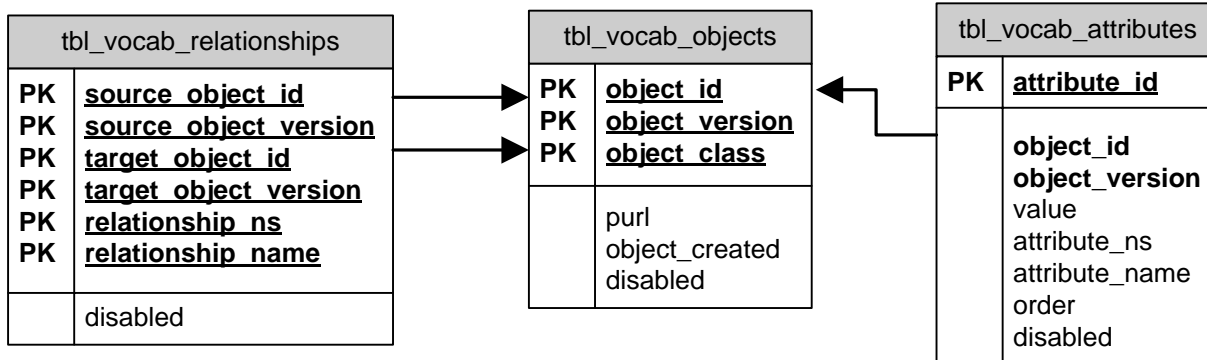
Technical Architecture

Architecturally, the technology underlying this prototype is similar to other services provided by ANDS. The Vocabulary Service prototype is essentially a RESTful interface providing traversal of a RDF/SKOS concept hierarchy. Each concept must be tagged with an identifier (in this case but not limited to, using persistent URLs provided by purl.org) and the service must be capable of resolving those persistent URLs to concepts/schemes which are hosted by the service.

The technologies supporting this prototype implementation are Apache's HTTPd using simple mod_rewrite rules to direct queries to a PHP web application (based on the CodeIgniter framework). Underlying this access layer, the storage engine is a relational database (PostgreSQL used at ANDS, but multiple RDBMS are supported). Content negotiation is supported by the mod_rewrite module.

These technologies are shared in most of the ANDS Online Services and so provide internal developers with a more rapid prototyping environment and ease of system maintenance.

Information Model



When initial discussion regarding this service were discussed, the future possibility of supporting multiple knowledge organisation systems (other than just SKOS) was introduced (hence the support of various namespaces to describe relationships/attributes). This might also be used for internal tool support (for example describing concept metadata as an attribute within an internally-defined namespace).

Forethought was also given to versioning, ownership and management of vocabularies and concepts. This is illustrated by the choice of keys for each concept / concept scheme. Whilst they are logically distinct, the output of Concepts and Concept Schemes (in SKOS) are syntactically similar and so they are treated as similar “vocab_objects” with similar relationships and attribute relationships.

This model also provides for concept schemes which might not use the same persistent identifier system (i.e. the purl is supplied/generated at ingest and the system does not mandate the structure of the identifier).

Services

Delivery Services

Service: getConceptSchemes

Description:

Returns the Concept Schemes (or “vocabularies”) which this service is able to deliver. ConceptSchemes are provided in SKOS format and in no particular order. Response only includes attributes (such as labels and scopeNotes) and does not return relations (such as related concepts).

Parameters: <none>

Example Usage:

Request: HTTP GET <http://services.andis.org.au/vocab/getConceptSchemes>

Response:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#">

  <rdf:Description rdf:about="http://purl.org/au-research/vocabulary/anzsrc-for/2008/">
    <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#ConceptScheme"/>
    <skos:prefLabel>1297.0 - Australian and New Zealand Standard Research Classification (ANZSRC),
2008</skos:prefLabel>
    <skos:altLabel>Field of Research Codes</skos:altLabel>
  </rdf:Description>

</rdf:RDF>
```

Service: getConceptByIdentifier

Description:

Returns zero or more Concepts (or a Concept Scheme) which match the identifier. Concepts are provided in SKOS format. Optionally, the depthParam allows you to request related records within a certain radius be included as well.

Parameters:

identifier (mandatory) – an identifier (such as a PURL) which uniquely identifies this concept (or ConceptScheme)

depthParam (optional) – tokenised string which defines a radius of concepts to include around the requested identifier (defaults to 0 in all directions, i.e. only returns the single identifier you requested). The format of this parameter is described in a subsequent section of this document.

Example Usage:

Request: HTTP GET <http://services.ands.org.au/vocab/getConceptByIdentifier?identifier=http://purl.org/au-research/vocabulary/anzsrc-for/2008/>

Response:

```
<?xml version="1.0"?>
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">
    <rdf:Description rdf:about="http://purl.org/au-research/vocabulary/anzsrc-for/2008/">
      <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#ConceptScheme"/>
      <skos:prefLabel>1297.0 - Australian and New Zealand Standard Research Classification (ANZSRC) ,
2008</skos:prefLabel>
      <skos:altLabel>Field of Research Codes</skos:altLabel>
      <skos:hasTopConcept rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/01"/>
      <skos:hasTopConcept rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/02"/>
      <skos:hasTopConcept rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/03"/>
      ...
    </rdf:Description>
  </rdf:RDF>
```

Service: *getConcepts*

Description:

Provides an alternate method of referencing Concepts (or ConceptSchemes) in the Vocabulary Service. This service uses a more “friendly URL” and also provides a resolution point for PURL redirects.

Returns zero or more Concepts (or a Concept Scheme) which match the request parameters. Concepts are provided in SKOS format. Optionally, the *depthParam* allows you to request related records within a certain radius to be included as well.

Request structure: `getConcepts/<vocabulary id>/<version>/<term identifier>/`

vocabulary id (mandatory) – name of the vocabulary (i.e. “anzsrc-for”)

version (mandatory) – specific version of this vocabulary (i.e. “2008”)

term identifier (optional) – identifier of the term which we are requesting (i.e. “720104”)

Parameters:

depthParam (optional) – tokenised string which defines a radius of concepts to include around the requested term (defaults to 0 in all directions, i.e. only returns the single term you requested). The format of this parameter is described in a subsequent section of this document.

Example Usage:

Request: HTTP GET <http://services.ands.org.au/vocab/getConcepts/anzsrc-for/2008/04>

Response:

```
<?xml version="1.0"?>
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">
    <rdf:Description rdf:about="http://purl.org/au-research/vocabulary/anzsrc-for/2008/04">
      <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
      <skos:prefLabel>EARTH SCIENCES</skos:prefLabel>
      <skos:narrower rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/0401"/>
      <skos:narrower rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/0402"/>
      ...
      <skos:narrower rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/0406"/>
      <skos:narrower rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/0499"/>
      <skos:inScheme rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/">
      <skos:topConceptOf rdf:resource="http://purl.org/au-research/vocabulary/anzsrc-for/2008/">
    </rdf:Description>
  </rdf:RDF>
```

Format of depthParam

The depthParam parameter allows you to define the graph radius of concepts to include, relative to the requested term. All identifiers which fall within the defined radius will be included in the response.

String format: n:0:b:0:r:0 (Default)

n:0..99 (narrower concepts)

b:0..99 (broader concepts)

r:0..99 (other related concepts)

So ?depthParam=n:99:b:0:r:0 would give you the requested concept and all concepts which are narrower than it.

Service Contact

For more information on the prototype controlled vocabularies service, please contact services@ands.org.au. For more information on the work of the Vocabulary Services Technical Working Group (TWG) or to discuss the operation of the service, please contact Stuart Hungerford at ANDS (stuart.hungerford@ands.org.au).