# Publishing controlled vocabularies for access and reuse

COAR webinar, December 15, 2016
Rowan Brownlee, Australian National Data Service (ANDS)
rowan.brownlee@ands.org.au



1

Hi everyone.  My name is Rowan Brownlee and I work for the Australian National Data Service, otherwise known by its acronym ANDS.  I'm based in Canberra, at the Australian National University.

In the previous webinar in October, we heard about the aims and activities of the COAR vocabularies program and development of the Resource Type vocabulary. There was a discussion of what is meant when we use terms like taxonomy, ontology, controlled vocabulary and thesaurus.  There was also discussion of the value of vocabularies in supporting metadata quality and repository interoperability.

# Overview

- Making vocabularies available for access and reuse
- What is SKOS.  Why was it developed?  Why use it?
- Expressing a classical vocabulary in SKOS
- Extending and combining SKOS
- Tools for creating, managing, publishing and accessing SKOS vocabularies
- Vocabulary registry interoperability
- Australian Vocabulary Special Interest Group (AVSIG)
    - Getting started with vocabularies

- Speakers in the previous webinar also talked about how the COAR vocabularies are being made available online for machine and human access. It is this last part, the topic of access for reuse which I'll be focusing upon in this webinar.

- Firstly I'll take an example of what might be described as a classical vocabulary and illustrate how it may be expressed in SKOS.  This will involve an overview of the origins of SKOS.

- Then I'll discuss the flexibility of the SKOS vocabulary, how it may be extended or combined with properties and classes from other RDF vocabularies such as DC and OWL,

- Then, through the lens of the ANDS vocabulary service, I'll talk about tools for creating, managing, publishing and accessing vocabularies.

- Then I'll talk a little about vocabulary registry interoperability.

- Along the way I'll mention the Australian Vocabulary Special Interest Group. This recently formed group, known by its acronym AVSIG has as one of its areas of focus the topic of getting started with vocabularies.  Its great to have tools and technologies.  We also benefit from plain-language guidance and support.

# Why publish vocabularies on the web?

- Vocabularies are relevant to the Internet
- Vocabularies help improve search precision within an indexed collection
- Hierarchical and associative relationships assist collection browsing
- Published vocabularies become independently accessible resources for reuse
- Published vocabularies are visible
- Vocabulary visibility supports discussion about duplication and consolidation
- Published vocabularies may be mapped to each other
- Vocabularies provide points of relation and connection across the Semantic Web

3

Before talking about how SKOS may be used to publish vocabularies, I'll first mention some of the reasons why vocabulary publication is a good idea.

Many organisations use published vocabularies, or develop and maintain their own, and these vocabularies are a key component of their information systems.  Galleries, Libraries, Archives, Museums.  Also research organisations.  Australian examples of research organisations which use vocabularies include CSIRO (Commonwealth Scientific and Industry Research Organisation), BoM (Bureau of Meteorology), IMOS (Integrated Marine Observing System), and GA (Geoscience Australia).

Many vocabularies are developed to support the needs of a particular organisation's systems, or a network of systems within a particular domain (such as a library network).  However, the purposes for which these vocabularies were developed remain relevant within the context of the Internet.  A controlled vocabulary helps to improve the precision of results when searching an indexed collection.  Hierarchical and associative relationships help when browsing collections.

When published online, controlled vocabularies become accessible resources. Rather than locked up in a specific institutional system, or accessible only to members of a particular network, these vocabularies become available for direct access and reuse.

Publishing a vocabulary makes it more visible.  Visibility may assist a domain community to discuss whether multiple similar vocabularies are really needed.  As an

example, Geoscience Australia are publishing a considerable number of vocabularies used by that organisation.  It is likely that other organisations in the same domain have their own versions of these types of vocabularies.  As more organisations publish their vocabularies, might there be options for consolidation?

There may be good reason for choosing one domain vocabulary over another, even if they are quite similar.  If that is the case, that is fine, and relationships can still be expressed between similar or identical concepts in different vocabularies.  This mapping of vocabularies assists users to traverse pathways across separate collections which may be indexed using separate but aligned vocabularies.

Vocabularies are also important to the idea of the Semantic Web.  The semantic web entails the expression of precise and meaningful statements about resources.  These precise and meaningful statements include expression of a resource's conceptual aspects.

Controlled vocabularies reflect agreement on terminology for labeling concepts. When there is agreement on common language for concepts, then the discovery, linking, understanding and reuse of resources is improved.  As an example, the Australian Integrated Marine Observing System publishes data and metadata which includes terminology from marine vocabularies.  These vocabularies are used by international initiatives such as ODIP (Ocean Data Interoperability Platform) and enable data collected in Australia to be shared and understood more widely within the marine community.

If you're going to publish a vocabulary, it helps to have an agreed and supported model.  Just as a vocabulary is an expression of agreement on meaning, a community standard is an expression of agreement on how a vocabulary structure will be expressed.  SKOS is an example of a model which provides a simple means to enable vocabulary access and reuse on the Internet.  A great strength of SKOS is its simplicity.

# Why use SKOS?  Why not use OWL?

- OWL provides a formal and precise way to represent domain knowledge.
- Vocabularies are not formally precise representations of domain knowledge.
- Vocabularies are not as complex as ontologies
- SKOS is simpler than OWL
- Translation to OWL can take considerable effort and expertise
- Translation to OWL may introduce misleading logical precision
- SKOS focuses on features typical of vocabularies
- SKOS provides a low-cost migration path to the semantic web

If we agree that controlled vocabularies are of value, and that there are good reasons for publishing vocabularies on web, how then should they be published?  What model should be used for expressing the typical features of a vocabulary?

The Web Ontology Language, otherwise known as OWL predates SKOS.  Why not use OWL?

OWL provides a formal and precise means of representing domain knowledge.  OWL enables a great deal of detail to be recorded about a knowledge domain, including rules and axioms, all of which in combination provide a basis for complex computer-based reasoning.  Creating an ontology using OWL can involve quite a bit of work if there's a lot that should be modelled.  It takes considerable domain expertise.

Controlled vocabularies are not formally precise representations of domain knowledge.  Rather they are more like *"informal structures reflecting the intuitive knowledge of human users in a form useful for resource discovery (e.g., through supporting query expansion)."* ( p2. Key choices in the design of Simple Knowledge Organization System (SKOS))  They don't have the complexity that requires the expressiveness of an ontology language.  Because they are less formal, less precise, it can take considerable work to translate a vocabulary to OWL.

If done incorrectly, conversion of a controlled vocabulary to an OWL ontology may introduce misleading logical precision.  As an example, a vocabulary may use a

broader term relationship, indicated by BT.  BT may indicate a generic, partitive or instantial relationship.  The way in which BT has been used by the thesaurus may not be documented.  Even if there are policies governing how BT should used in developing the thesaurus, these policies may not have been uniformly followed over time.

In one of the papers which I reference at the end of this presentation, the authors discuss AGROVOC.

*"The AGROVOC thesaurus of multilingual agricultural terminology, the product of many people working over many years from multiple perspectives, was straightforwardly converted into a hierarchy of OWL classes many years before the finalization of SKOS. While the maintainers of AGROVOC-in-OWL intended to increase its ontological precision over time, through editorial correction and refinement, it eventually proved to be more practical simply to convert AGROVOC back into the formally less ''committed'' form of a SKOS concept scheme, leaving it to designers of specific implementations to upgrade parts of the thesauri into class-based ontologies when required to support reasoning."* ( p3. Key choices in the design of Simple Knowledge Organization System (SKOS))

Rather than seeking to capture every possible characteristic of every vocabulary, SKOS focuses on typical features of many vocabularies. The developers of SKOS wanted to ensure compatibility between SKOS and existing thesaurus standards (ISO 2788 and ISO 5964,14).  SKOS therefore reflects standard thesaurus construction principles. It was intended to be able to be used and understood by people who already used and understood vocabularies.  In using SKOS,  they wouldn't need to model an ontology to publish a vocabulary for use by people and software.  SKOS was intended to provide a low-cost migration path to the semantic web.

*Simple Knowledge Organization System*

Scope: knowledge organization systems (KOS) such as thesauri, classification systems, subject heading lists…

SKOS is for representing KOSs in RDF in a *simple* way

In the previous webinar, Timo provided definitions for a controlled vocabulary, taxonomy and thesaurus.

Each of these can be represented using SKOS.  The emphasis here is on taking a simple approach.  A simple model that expresses the key features of controlled vocabularies should have a greater chance of being used and supported.
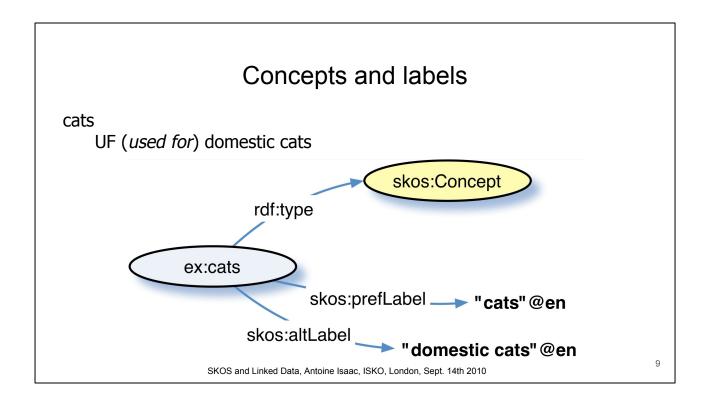
# Thesaurus example

Animals

cats

    UF (*used for*) domestic cats
    RT (*related term*) wildcats
    BT (*broader term*) animals
    SN (*scope note*) used only for domestic cats

domestic cats
    USE cats

wildcats

Here is an example which expresses typical properties of a thesaurus. Concepts, relationships between concepts and documentation guiding how a concept should be interpreted or applied.

In the case of this thesaurus, "animals", "cats" and "wildcats" are preferred, and "domestic cats" is non-preferred. The thesaurus creator has also decided that "cats" is equivalent to "domestic cats" and that "cats" should be used instead of "domestic cats". It may be that the thesaurus creator sees "cats" and "domestic cats" as synonyms, or that the further specificity of "domestic cats" is not required, and that it is instead sufficient to only use the broader term "cats".

There are other relationships between concepts. "Cats" has (as a broader term), "animals". There is also some sort of associative relationship between "cats" and "wildcats". Wildcats cannot be a type of cat, since "cats" covers domestic cats only. Since someone who is interested in "cats" may possibly also be interested in "wildcats", an associative relationship is expressed to assist exploration of the subject.
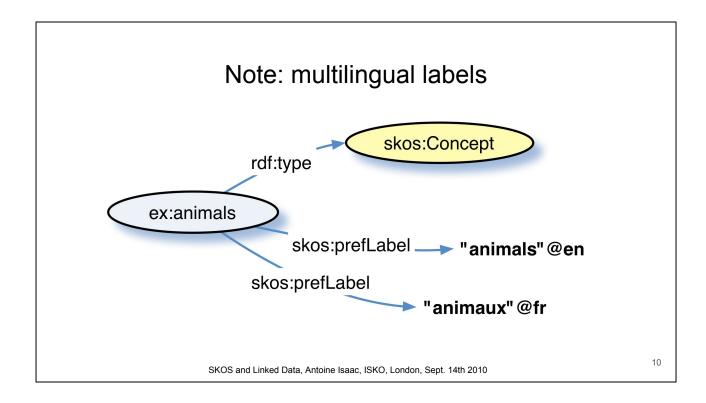
Next I'll illustrate how these typical properties may be expressed using SKOS.

# Concepts and labels

cats
    UF (*used for*) domestic cats



skos:Concept

rdf:type

ex:cats

skos:prefLabel ⟶ **"cats"@en**

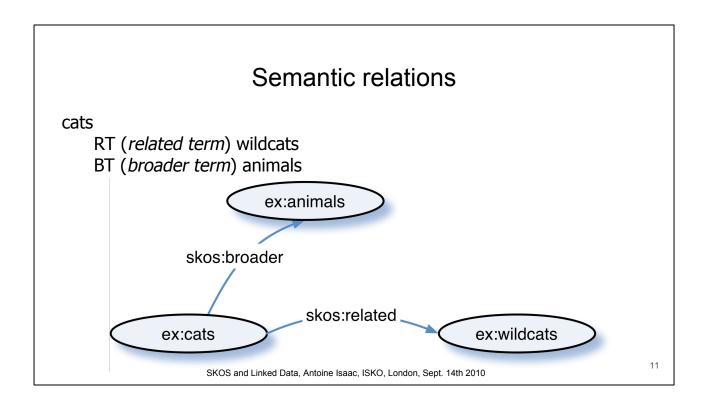skos:altLabel ⟶ **"domestic cats"@en**

"Cats" is a resource which is of type skos:Concept.

SKOS concepts have labels, including preferred labels and alternative labels. In this example, the preferred label is "cats" and the alternative label is "domestic cats". (This use of skos:prefLabel and skos:altLabel is equivalent to the earlier USE and USED FOR relationship of "cats" and "domestic cats").

SKOS can also indicate the language of a label, in this case English.

# Note: multilingual labels

A SKOS concept may have many preferred labels, but at most one in a particular language. In this example, we see the French and English language preferred labels for the concept "animals". This supports multilingual search.

# Semantic relations

cats
    RT (*related term*) wildcats
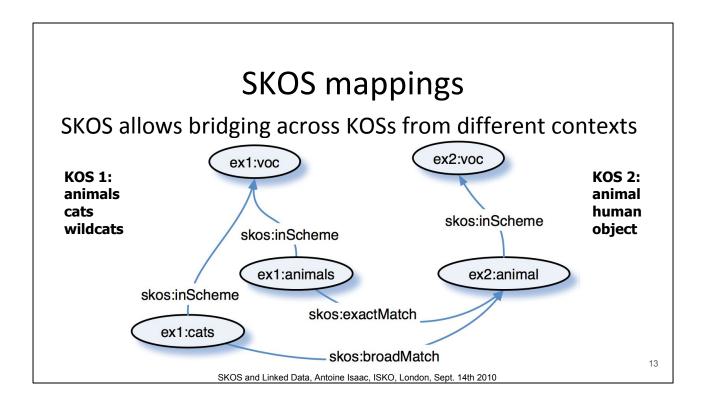    BT (*broader term*) animals

SKOS expresses the types of hierarchical and associative relationships found in a traditional thesaurus.  Animals is a broader concept to cats.  Wildcats is related to cats

A SKOS graph

animals
cats
      UF domestic cats
      RT wildcats
      BT animals
      SN used only for domestic cats
domestic cats
      USE cats
wildcats

SKOS and Linked Data, Antoine Isaac, ISKO, London, Sept. 14th 2010

12

Here we see the original thesaurus extract and SKOS version together. The SKOS concepts, their labels, relationships and documentation.

In SKOS terms, a vocabulary is a "concept scheme", an aggregation of interrelated concepts. Two properties from the SKOS Core specification which support the expression of a relationship between a concept and a concept scheme are skos:inScheme, and skos:hasTopConcept. These properties (not illustrated on this slide) provide a means of asserting that this concept is contained within this concept scheme.

# SKOS mappings

## SKOS allows bridging across KOSs from different contexts



KOS 1:
animals
cats
wildcats

ex1:voc

ex2:voc

KOS 2:
animal
human
object

skos:inScheme

skos:inScheme

ex1:animals

ex2:animal

skos:inScheme

ex1:cats

skos:exactMatch

skos:broadMatch

SKOS and Linked Data, Antoine Isaac, ISKO, London, Sept. 14th 2010

13

SKOS supports the assertion of relationships between concepts that are in separate concept schemes.

In this example, concept scheme one has animals, cats and wildcats. Concept scheme two has a concept "animal" which is judged to be sufficiently the same as the "animals" in concept scheme one. Because they are the same concept it is reasonable to assert an "exactMatch" relationship between "animal" and "animals". Because these concepts enjoy an exactMatch relationship, it is reasonable to assert that "cats" from concept scheme 1 has a "broadMatch" with "animal" from concept scheme two.

Apart from broadMatch and exactMatch, SKOS also provides narrowMatch, relatedMatch and closeMatch.

# SKOS mappings: Agrovoc

Table 1 below provides some figures about the vocabularies to which AGROVOC is aligned:

| | Resource | Topics | Total # of Linked concepts | Languages | Linked Resource available as LOD? | Type of link (and # of linked concepts) |
|---|---|---|---|---|---|---|
| 1 | ASFA | Fisheries | 1784 | | Yes | skos:closeMatch (38), skosLexactMatch (1741) |
| 2 | Biotechnology Glossary (FAO) | Biotechnologies | 793 | EN, ES, FR, +3 more | Yes | skos:closeMatch (793) |
| 3 | Chinese Agriculture Thesaurus (CAT) | Agriculture | | | Yes | skos:narrowMatch (137) skos:broadMatch (10153) skos:exactMatch (10325) |
| 4 | DBpedia | General | 11009 | EN, ES, FR + 8 more | Yes | skos:closeMatch (11009) |
| 5 | Dewey Decimal Classification (DDC) | General | 401 | EN, ES, FR + 8 more | Yes | skos:closeMatch (2) skos:exactMatch (399) |
| 6 | EUROVOC | General EU | 1 269 | EN, ES, FR + 21 more | Yes | skos:exactMatch (1269) |
| 7 | GEMET | Environment | 1 175 | EN, ES, FR + 30 more | Yes | skos:exactMatch (1175) |
| 8 | GeoNames | Geographical entities | 206 | EN, ES, FR + 63 more | Yes | skos:exactMatch (206) |
| 9 | Geopolical Ontology | Geopolitical entities | 253 | AR, CH, EN, ES, FR, RU | Yes | skos:exactMatch (253) |
| 10 | Library of Congress Subject Headings | General | 1 075 | EN | Yes | skos:exactMatch (1075) |

14

Here is an example of mapping using the agricultural vocabulary Agrovoc.

As of December 5 2016, Agrovoc had been mapped or aligned with 16 other vocabularies.  Some of these are general thesauri and others are environmental or agricultural.
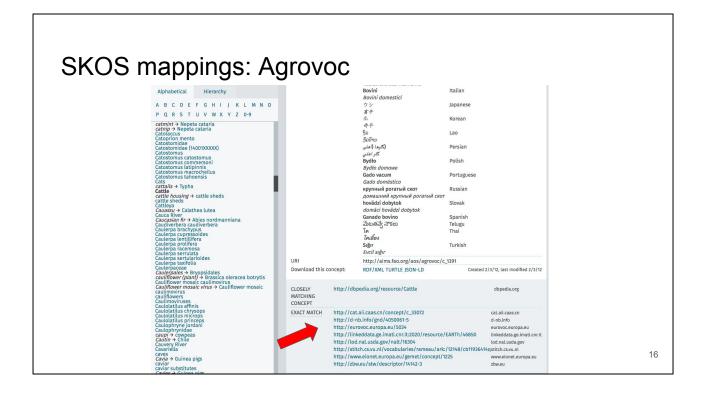Ten of these are displayed on this slide.  The column to the right displays the type of mapping, and the number of concepts mapped.  Many of these are major vocabularies commonly used in collection indexing.  This mapping of vocabularies assists users to traverse pathways across separate collections which may be indexed using separate but aligned vocabularies.

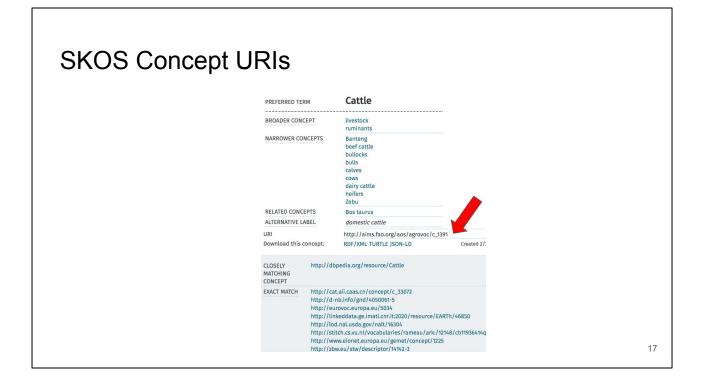http://aims.fao.org/standards/agrovoc/linked-open-data

# SKOS mappings: Agrovoc

Here is an example of a SKOS concept from Agrovoc. The concept is "Cattle." In this part of the display we can see broader and narrower relationships, preferred and alternative labels. There are also quite a few preferred labels in languages other than English. Agrovoc is a multilingual vocabulary.

http://oek1.fao.org/skosmos/agrovoc/en/page/c_1391

# SKOS mappings: Agrovoc

Further down the page we see mappings to other vocabularies. There is an exact match mapping between "Cattle" as a concept in Agrovoc, and equivalent concepts in these other resources.

# SKOS Concept URIs

| | |
|---|---|
| PREFERRED TERM | **Cattle** |
| BROADER CONCEPT | livestock<br>ruminants |
| NARROWER CONCEPTS | Banteng<br>beef cattle<br>bullocks<br>bulls<br>calves<br>cows<br>dairy cattle<br>heifers<br>Zebu |
| RELATED CONCEPTS | Bos taurus |
| ALTERNATIVE LABEL | *domestic cattle* |
| URI | http://aims.fao.org/aos/agrovoc/c_1391 |
| Download this concept: | RDF/XML TURTLE JSON-LD        Created 2/: |

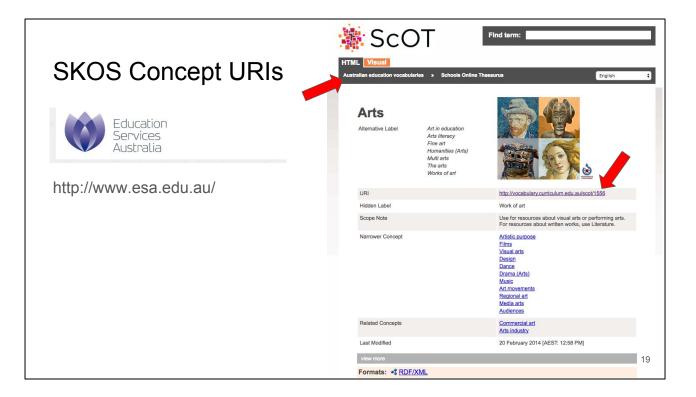| | |
|---|---|
| CLOSELY MATCHING CONCEPT | http://dbpedia.org/resource/Cattle |
| EXACT MATCH | http://cat.aii.caas.cn/concept/c_33072<br>http://d-nb.info/gnd/4050061-5<br>http://eurovoc.europa.eu/5034<br>http://linkeddata.ge.imati.cnr.it:2020/resource/EARTh/46850<br>http://lod.nal.usda.gov/nalt/16304<br>http://stitch.cs.vu.nl/vocabularies/rameau/ark:/12148/cb11936414q<br>http://www.eionet.europa.eu/gemet/concept/1225<br>http://zbw.eu/stw/descriptor/14142-3 |

We also see the concept's URI, or Uniform Resource Identifier.  A SKOS concept has a URI which identifies that resource on the Internet. It enables interaction with the resource through use of specific protocols, in this case HTTP. If I pasted that URI in a browser window, I would see this html page.  If a software application was accessing this URI, it could instead request the concept description in a structured format suitable for machine processing.  Formats such as RDF/XML, TURTLE or JSON-LD.

URIs are very important.

In this example, the concept ID is c_1391.  The concept ID is that very last piece of the URI.  A concept ID like c_1391 is opaque.  There's nothing about that ID which would prompt me to think - 'Ah yes, that must mean Cattle".  The vocabulary creator could have instead decided to use a semantically meaningful identifier.  Instead of c_1391, they might have used "Cattle".  But then what would happen if the preferred label changed?  At best the URI may become misleading or confusing, as it would appear out of synch with the preferred label.  Also, in a multilingual thesaurus, which language would be chosen for a semantically meaningful concept ID?  In designing a URI pattern, it is worth considering the use of an ID which has no semantic relationship to the concept that it is identifying.

In this example, it is also useful to examine other parts of the URI.  Agrovoc is produced by the Food and Agriculture organisation, and the URI contains that organisation domain.  What would happen if the organisation changed its name?

## SKOS Concept URIs

Here is another example of a URI pattern.  This is a concept from the Schools Online Thesaurus.  The example concept has the preferred label "Arts".  In this case the concept ID is opaque, and unlike the case of Agrovoc, the organisation name forms no part of the URI.

Schools Online Thesaurus is produced by Education Services Australia, or ESA. Rather than use ESA in the URI, they use "curriculum".  Education Services Australia could change its name or its URL, and this change would have no impact on the concept URIs in the ScOT thesaurus.  In contrast to a business name, the term "curriculum" is enduring and it also communicates something about what the company is doing in this domain.

Organisations change their name.  URIs should never change.  If a URI changes existing links break.

SKOS Concepts

- have labels (such as preferred and alternative)
- may be related to each other (broader, narrower, related)
- may be documented (such as scope notes, definitions, examples)
- are organised into concept schemes
- may be mapped across concept schemes
- are identified using URIs
- may be grouped into collections

These are some areas that I've talked about so far. SKOS concepts have labels. Concepts may have hierarchical and associative relationships within a vocabulary. Concepts within one vocabulary may be mapped to concepts in other vocabularies. Concepts are organised into concept schemes, using SKOS properties such as inScheme and topConceptOf. Concepts are also identified using URIs.

Now I'll talk a little about how SKOS Concepts may also be grouped into Collections.

# SKOS Collections

```
☐ ⚹  ...................... nails (fasteners)
☐ ➡  ......................... <nails by form>  ◄────────────
☐  ............................. barbed nails
☐  ............................. box nails
☐  ............................. cement-coated nails
☐  ............................. Chinese nails
☐  ............................. common nails
☐ ⚹  ............................. cut nails
☐  ............................... cut slating nails [N]
☐  ............................. dog nails
☐  ............................. flat nails
☐  ............................. gutter spikes [N]
☐  ............................. helical nails
☐  ............................. hook nails
☐  ............................. hurdle nails
☐  ............................. jobent nails
☐ ⚹  ............................. <nails by form: head type>  ◄──────────
☐ ⚹  ............................. brads
☐  ............................... casing nails
☐  ............................... chatônes (decorative bosses)
☐  ............................... clout nails
```

This is an extract from the Getty Art and Architecture Thesaurus.
http://www.getty.edu/vow/AATHierarchy?find=chisel+nails&logic=AND&note=&page=1&subjectid=300048926
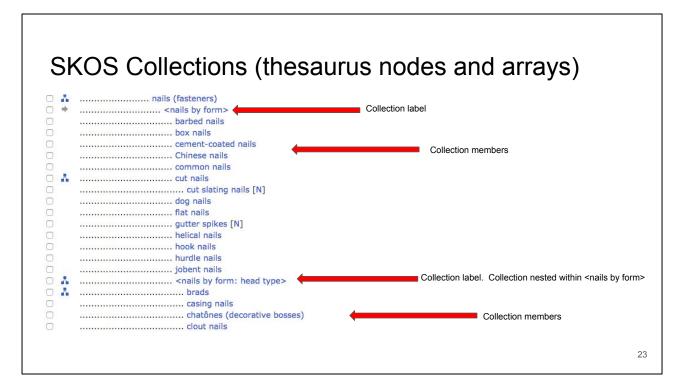
At the top level of this extract is the subject "nails (fasteners)". Immediately beneath that is "nails by form", and you'll see that "nails by form" is surrounded by greater-than/less-than brackets. Below "nails by form" is an alphabetic listing of nails. This alphabetic sequence continues until we see another heading within greater-than/less-than brackets. This time, "nails by form: head type". And then, as before, another alphabetic sequence of nails.

This example extract of thesaurus nodes and arrays taken from
www.willpowerinfo.co.uk/Will_ISKO2012_paper.doc
*The ISO 25964 Data Model for the Structure of an Information Retrieval Thesaurus*
by Leonard Will

# SKOS Collections (thesaurus nodes and arrays)



```
☐ ⬡  ...................... nails (fasteners)
☐ ➡  .......................... <nails by form>  ⬅——————————— Node label
☐     .............................. barbed nails
☐     .............................. box nails
☐     .............................. cement-coated nails  ⬅——————————— Members of array
☐     .............................. Chinese nails
☐     .............................. common nails
☐ ⬡  .............................. cut nails
☐     .............................. cut slating nails [N]
☐     .............................. dog nails
☐     .............................. flat nails
☐     .............................. gutter spikes [N]
☐     .............................. helical nails
☐     .............................. hook nails
☐     .............................. hurdle nails
☐     .............................. jobent nails
☐ ⬡  .............................. <nails by form: head type>  ⬅——— Node label.  Nested within <nails by form>
☐ ⬡  .................................. brads
☐     .................................. casing nails
☐     .................................. chatônes (decorative bosses)  ⬅——————— Members of array
☐     .................................. clout nails
```

22

Within a thesaurus, groups of concepts which have a common parent may be organised into arrays.  Node labels are thesaurus entries which indicate how the concepts have been grouped.  The concept "nails (fasteners)" has many narrower concepts.  One option would have been to display one single very long alphabetic list.  AAT instead uses nodes to help with navigation, browsing and display.  Nodes are not concepts.  Rather they are used to group arrays of concepts, providing further guidance to users of the vocabulary.  In the AAT example, members of an array are ordered alphabetically.  Other ordering methods could be used, such as some numeric sequence.  Alternatively an array could be unordered.  Arrays can also be nested.  In this example, the array labeled <nails by form> has a sub-array labeled <nails by form: head type>.

This use of nodes and arrays provides an additional means of organising concepts within a thesaurus.  It helps people navigate the thesaurus by breaking up a potentially very long list of narrow terms into further categorised groups of terms.
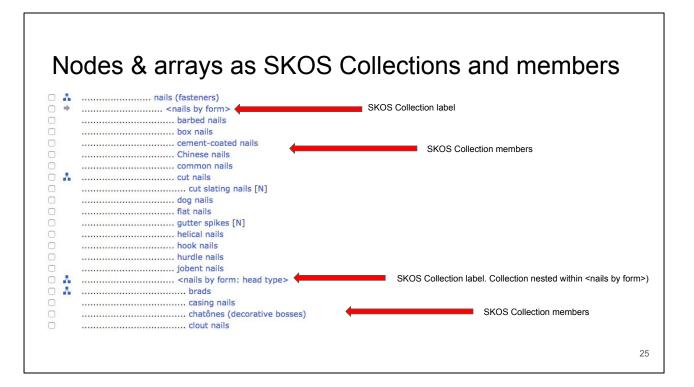
# SKOS Collections (thesaurus nodes and arrays)

```
☐  ⊹   ....................... nails (fasteners)
☐  ➡   .......................... <nails by form>          ⟵━━━━━━━━━━━  Collection label
☐             ............................... barbed nails
☐             ............................... box nails
☐             ............................... cement-coated nails      ⟵━━━━━━━━━━━  Collection members
☐             ............................... Chinese nails
☐             ............................... common nails
☐  ⊹          ............................... cut nails
☐             ................................. cut slating nails [N]
☐             ............................... dog nails
☐             ............................... flat nails
☐             ............................... gutter spikes [N]
☐             ............................... helical nails
☐             ............................... hook nails
☐             ............................... hurdle nails
☐             ............................... jobent nails
☐  ⊹          ............................... <nails by form: head type>   ⟵━━━━━  Collection label.  Collection nested within <nails by form>
☐  ⊹          ................................. brads
☐             ................................. casing nails
☐             ................................. chatônes (decorative bosses)   ⟵━━━━━  Collection members
☐             ................................. clout nails
```

23

So too the SKOS Collection can provide another means of expressing relationships between concepts within a thesaurus.  In this example, "nails by form" could be a Collection label, and the narrow terms below it, members of the Collection.  They are still narrower concepts of "nails (fasteners)", but they are also members of the Collection.

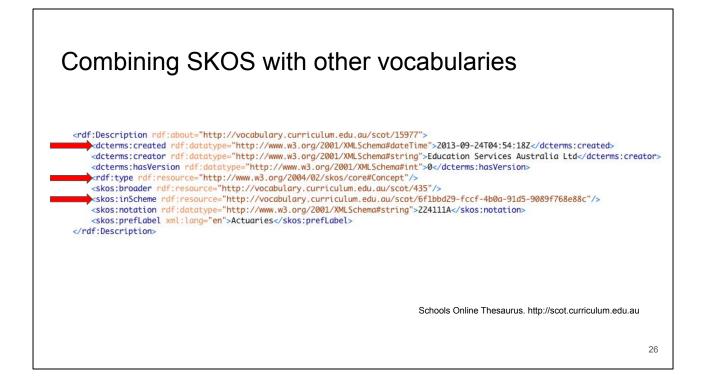# Nodes & arrays as SKOS Collection & members



Here is another view of how SKOS Collections may express the structure of nodes and arrays. In the AAT example, the node label <nails by form> is a type of Guide Term. Guide Term is a class defined within the Getty ontology. A Guide Term is a subclass of a Thesaurus Array, which is also a class within the Getty ontology. Thesaurus Array is however a subclass of a SKOS Collection. The node label <nails by form> is a label for a SKOS Collection.

http://vocab.getty.edu/aat/300048926
http://vocab.getty.edu/resource?uri=http%3A%2F%2Fvocab.getty.edu%2Fontology%23GuideTerm
http://vocab.getty.edu/resource?uri=http%3A%2F%2Fpurl.org%2Fiso25964%2Fskos-thes%23ThesaurusArray

# Nodes & arrays as SKOS Collections and members

☐ ⛓ ...................... nails (fasteners)
☐ ➡ ......................... <nails by form> ◄━━━━━━━━━━━━━━ SKOS Collection label
☐ ............................... barbed nails
☐ ............................... box nails
☐ ............................... cement-coated nails ◄━━━━━━━━━━ SKOS Collection members
☐ ............................... Chinese nails
☐ ............................... common nails
☐ ⛓ ............................... cut nails
☐ ................................... cut slating nails [N]
☐ ............................... dog nails
☐ ............................... flat nails
☐ ................................... gutter spikes [N]
☐ ............................... helical nails
☐ ............................... hook nails
☐ ............................... hurdle nails
☐ ............................... jobent nails
☐ ⛓ ............................... <nails by form: head type> ◄━━━━━━ SKOS Collection label. Collection nested within <nails by form>)
☐ ⛓ ................................... brads
☐ ................................... casing nails
☐ ................................... chatônes (decorative bosses) ◄━━━━━ SKOS Collection members
☐ ................................... clout nails

25

Back to the thesaurus view, <nails by form> is a label for a SKOS Collection which has member concepts including barbed nails, box nails and Chinese nails. <nails by form: head type> is a SKOS Collection nested within the <nails by form> Collection.

# Combining SKOS with other vocabularies

```
<rdf:Description rdf:about="http://vocabulary.curriculum.edu.au/scot/15977">
   <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2013-09-24T04:54:18Z</dcterms:created>
   <dcterms:creator rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Education Services Australia Ltd</dcterms:creator>
   <dcterms:hasVersion rdf:datatype="http://www.w3.org/2001/XMLSchema#int">0</dcterms:hasVersion>
   <rdf:type rdf:resource="http://www.w3.org/2004/02/skos/core#Concept"/>
   <skos:broader rdf:resource="http://vocabulary.curriculum.edu.au/scot/435"/>
   <skos:inScheme rdf:resource="http://vocabulary.curriculum.edu.au/scot/6f1bbd29-fccf-4b0a-91d5-9089f768e88c"/>
   <skos:notation rdf:datatype="http://www.w3.org/2001/XMLSchema#string">224111A</skos:notation>
   <skos:prefLabel xml:lang="en">Actuaries</skos:prefLabel>
</rdf:Description>
```

Schools Online Thesaurus. http://scot.curriculum.edu.au

26

In that previous example we could see an example of classes and properties from the Getty ontology being combined with SKOS classes and properties.

The Resource Description Framework or RDF supports combination of classes and properties from different vocabularies. In this example, SKOS is combined with Dublin Core Terms. DC expresses date of creation, name of creator and version. There is an rdf type property indicating that this type of resource is a SKOS concept. RDF makes it easy to combine RDF vocabularies, and SKOS was designed with this in mind. The designers of SKOS believed that there was no need to define properties and classes where it was possible to reuse those from existing and widely used RDF vocabularies. Why define a SKOS class for creator, when DC already provides one. Reuse of existing RDF vocabularies also assists with interoperability. If a community is already using DC, then it is known and understood. If I use DC, it is likely that systems implemented by that community will know how to make sense of the DC that I provide.

I'll mention at this point that the term "vocabulary" is somewhat overloaded. There are vocabularies like LCSH with are full of descriptive concepts. There are also RDF vocabularies like SKOS and OWL and DC. In both cases it makes perfect sense to combine. I might combine concepts from LCSH and Schools Online Thesaurus to adequately express the conceptual aspect of a resource. I might also combine classes and properties from various RDF vocabularies to adequately describe the resource.

# Extending SKOS (eg. XKOS)

http://www.ddialliance.org/Specification/RDF/XKOS

Apart from combining SKOS with other RDF vocabularies, SKOS was designed to be able to be extended to meet more specific needs.

I mentioned earlier that rather than seeking to capture every possible characteristic of every vocabulary, SKOS focuses on typical features of many vocabularies, to provide a simple and low-cost migration path to the semantic web.  For some domains, SKOS is insufficient.  For example, the statistics community have defined extensions to SKOS to meet the requirements for expressing statistical classifications.   XKOS, an extension of SKOS for the statistical community includes several dozen additional properties, including additional semantic relations.

# Extending SKOS (eg. XKOS)



Figure 5 – XKOS annotations

## 4.7 On Specializing the SKOS Model

SKOS is intended to serve as a common denominator between different modeling approaches. As such, the current vocabulary specification will allow many existing KOSs to be ported to the Semantic Web. However, the great variety of KOS models makes it impossible to capture every detail of these models while still retaining the first "S" ("simple") in "SKOS".

Applications that require finer granularity will greatly benefit from SKOS's being a Semantic Web vocabulary. SKOS can indeed be seamlessly *extended* to suit the specific needs of a particular KOS community while retaining compatibility with applications that are based on the core SKOS features.

This can mostly be done by *specializing* existing SKOS constructs into more-specific ones. Users can create their own properties and classes and attach them to the standard SKOS vocabulary elements by using the rdfs:subPropertyOf and rdfs:subClassOf properties from the RDF Schema vocabulary [*RDF-PRIMER*].

http://rdf-vocabulary.ddialliance.org/xkos.html
https://www.w3.org/TR/skos-primer/#secskosspecialization

Here is an example from XKOS which follows the SKOS recommended practice for defining extensions.  In this case the XKOS community wish to record more specific types of information than that which is catered for by a SKOS scope note. Two additional note properties are defined as sub-properties of the SKOS scope note property.  These new sub-properties are inclusion note and exclusion note.  In addition, Core content note and additional content note are sub-properties of the xkos inclusion note. Although all of these types of notes could be recorded as generic SKOS scope notes, the type of each note would be lost.  It would be difficult if not impossible to understand which scope note is the inclusion note or the exclusion note.

SKOS provides guidelines on how to define extensions in a way that provides support for the more granular and specific needs of particular domain communities, while retaining compatibility with applications which are based on the core SKOS features.

# SKOS-XL

- SKOS Core
  - A Concept is a Resource
  - A Label
    - is a property of a Concept
    - cannot be related to other labels
- SKOS-XL
  - A Label is a Resource
  - Supports relationships between labels
    - skosxl:labelRelation
  - Users may define relationships which specialise skosxl:labelRelation
    - ex:isAcronymOf (e.g. FAO isAcronymOf Food and Agriculture Organisation)

SKOS itself has been extended in the form of SKOS-XL, or SKOS extension for labels.

In SKOS Core, labels are properties of concepts. Labels cannot be described any further. Relationships cannot be expressed between labels. A label doesn't have a URI. A concept has a URI, and a label is a property of that concept.

In SKOS XL, a Label is a Resource. It has a URI and can be described with properties and relationships.

SKOS XL includes skosxl:LabelRelation. This is a super property for applications defining their own relations. An example provided in the documentation concerns full form and acronym. FAO is an acronym of Food and Agriculture Organisation. In SKOS Core, FAO may be an alternative label for the full form, but SKOS does not provide a means of documenting this relationship. SKOS-XL supports the definition of a relationship (such as ex:isAcronymOf), to explicitly assert that FAO-URI ex:isAcronymOf Food and Agriculture Organisation-URI.

At the time SKOS was developed, there was discussion of including a set of label relationship types. They decided to leave this to an extension, as they didn't believe that they'd come up with a comprehensive set of relationship types.

# Overview

- Making vocabularies available for access and reuse
- What is SKOS.  Why was it developed?  Why use it?
- Expressing a classical vocabulary in SKOS
- Extending and combining SKOS
- Tools for creating, managing, publishing and accessing SKOS vocabularies
- Vocabulary registry interoperability
- Australian Vocabulary Special Interest Group (AVSIG)
- Getting started with vocabularies

So far I've talked about SKOS, its development, how a vocabulary may be expressed in SKOS and how SKOS may be combined and extended.  Next, through the lens of the ANDS Vocabulary Service, I'll talk about some tools for creating, managing, publishing and accessing SKOS vocabularies.

# Research Vocabularies Australia (RVA)

- ● ANDS provides software and support for

  - ○ Creating and editing
    - ■ Edit existing vocabularies or create new

  - ○ Publishing
    - ■ Description, access, licensing, versioning

  - ○ Discovery
    - ■ Metadata harvested by google

  - ○ Integration
    - ■ Application programming interface and widget

The ANDS vocabulary service is called Research Vocabularies Australia, or RVA.

RVA helps to create and manage vocabularies.  I can publish a vocabulary in formats that are usable by people and machines.  This means that I can browse a published vocabulary through a web user interface, and I can use RVA to draw vocabularies into my local systems, for use within my organisation.  RVA also helps to make the vocabularies more findable, more visible, as information about the vocabularies is harvested by Google.

# Using a GUI to edit SKOS

RVA provides a means of editing existing vocabularies or creating new vocabularies. The PoolParty vocabulary editor is a user-friendly web-based tool. It outputs SKOS, but you don't have to know much about SKOS to use it. Here is an example of a vocabulary. The left-hand pane shows the vocabulary's hierarchical structure. The right-hand pane shows details about a highlighted concept. It is a straightforward process to create concepts and express relationships between them.

# Using a GUI to edit SKOS

## Cats
*https://lod.zoology.org.au/def/animals/2*

| Details | Notes | Documents | Linked Da |

**URI Generation**
URI: https://lod.zoology.org.au/def/animals/<Increment>

## Wildcats
*https://lod.zoology.org.au/def/animals/3*

| Details | Notes | Documents | Linked Dat |

34

I can easily define a URI pattern and the software will implement this pattern as each concept is defined. In this example I have a base URI, and then each concept is individually identified by addition of a running number. PoolParty generates these numbers.

# Using a GUI to edit SKOS



I don't have to be concerned about SKOS syntax. I can define concept labels set relationships between concepts, and document concepts. The editing software creates the underlying SKOS.

# Some examples of software for editing SKOS

Here are a few examples of SKOS editors.  Most of these are commercial.  Vocbench is an open-source product.  I'd be interested to hear of experiences using these and other editing software, whether commercial or open-source.

It is good to have user-friendly editing tools, whether commercial or open source.  I was recently talking with someone from an organisation which has repeatedly encountered problems maintaining vocabulary management systems which require highly technical knowledge on the part of the vocabulary creator.

# Research Vocabularies Australia (RVA)

- ● ANDS provides software and support for

    - ○ Creating and editing
        - ■ Edit existing vocabularies or create new

    - ○ Publishing
        - ■ Description, access, licensing, versioning

    - ○ Discovery
        - ■ Metadata harvested by google

    - ○ Integration
        - ■ Application programming interface and widget

Onto vocabulary publishing

# Publishing



RVA also provides a means of publishing vocabularies.

# Publishing

RVA has a workflow that enables authenticated users to pull in a vocabulary from the ANDS instance of PoolParty, or publish a vocabulary which has not been managed within PoolParty.

# Publishing

AODN Platform Vocabulary

**Vocabulary Title** ⓘ
AODN Platform Vocabulary

**Vocabulary Acronym** ⓘ
AODN Platforms

**Vocabulary Description** ⓘ
A controlled vocabulary for platforms that can be used in Marine Community Profile metadata. This vocabulary will be used by the faceted search in the AODN/IMOS portal 1-2-3.

**Vocabulary Licence** ⓘ
CC-BY

**Vocabulary Creation Date** ⓘ
2015-09-23

**Revision Cycle** ⓘ
Periodic

**Note** ⓘ
Notes

**PoolParty Project Info**

**PoolParty Project ID**
1DCE1494-A022-0001-FFBD-12DE19E01FEB

**Versions**

| TITLE | STATUS | |
|---|---|---|
| Version 1.0 | superseded | ✖ |
| Version 1.1 | superseded | ✖ |
| Version 1.2 | superseded | ✖ |
| Version 1.3 | superseded | ✖ |
| Version 2.0 | current | ✖ |

+ Add a version

**Related** ⓘ

| TITLE | TYPE | |
|---|---|---|
| eMarine Information Infrastructure (eMII) | party | ✖ |
| eMII_Finney.Kim_Admin | party | ✖ |
| Natalia_Atkins | party | ✖ |
| eMII_Atkins.Natalia | party | ✖ |

40

In this example I've chosen to draw in a vocabulary from PoolParty. This shows the metadata editing page for a vocabulary titled AODN Platform Vocabulary. AODN is the acronym for the Australian Ocean Data Network. Some of the metadata is pre-populated from the file drawn in from PoolParty. This includes title, description, creators and publisher.

# Publishing

This is a view of the RVA publishing portal showing the published vocabulary. We can see what the vocabulary is about, licence conditions for reuse, how it may be accessed, and previous versions.

In this example, access options include file download and access via Application Programming Interface. We can also see that there are four earlier versions of the vocabulary - versions 1.0 and 1.1, 1.2 and 1.3. In this example, the publisher has chosen to create and release versions and has also decided that earlier versions may remain available for use. RVA supports choice regarding versioning and access to current and superseded versions. At this stage, versioning is at the level of the vocabulary as a whole. ANDS is intending to explore options for managing finer grained versioning, such as versioning at the concept level.

# RVA Repository & endpoints

- OpenRDF Sesame
  - RDF repository & SPARQL 1.1 query endpoints
- SISSVoc
  - Linked data API

| Endpoint template | Type | Query |
|---|---|---|
| /collection | List | List of all Collections and OrderedCollections |
| /concept | List | List of all Concepts |
| /concept/broader?anylabel={text} | List | List of all Concepts broader than the one where a label matches the text |
| /concept/broader?uri={baseConcept} | List | List of all Concepts broader than the one identified by this IRI |
| /concept/broaderTransitive?anylabel={text} | List | List of all Concepts broaderTransitive than the one where a label matches the text |
| /concept/broaderTransitive?uri={baseConcept} | List | List of all Concepts broaderTransitive than the one identified by this IRI |
| /concept/narrower?anylabel={text} | List | List of all Concepts narrower than the one where a label matches the text |
| /concept/narrower?uri={baseConcept} | List | List of all Concepts narrower than the one identified by this IRI |
| /concept/narrowerTransitive?anylabel={text} | List | List of all Concepts narrowerTransitive than the one where a label matches the text |
| /concept/narrowerTransitive?uri={baseConcept} | List | List of all Concepts narrowerTransitive than the one identified by this IRI |
| /concept/topConcepts | List | List of all topConcepts |
| /concept/topConcepts?scheme={schemeIRI} | List | List of all topConcepts that belong to the given ConceptScheme |
| /concept?anylabel={text} | List | List of all Concepts where a label matches the text |
| /concept?labelcontains={text} | List | List of all Concepts where a label contains the text, any language, case-insensitive |
| /conceptscheme | List | List of all ConceptSchemes |
| /resource?uri={IRI} | Item | Describe the given resource |

42

Hosting and publishing vocabularies requires technology for storing and providing access.

Sesame provides a framework for hosting and providing query access to vocabularies published through the ANDS service. Queries can be made using SPARQL which is the generic RDF query language.
SISSVoc is used to provide a linked data API. SISSVoc was developed to provide an API that matches SKOS, that is aligned with the structure of SKOS. These endpoint templates allow a general user to explore a SKOS vocabulary without having to know the SPARQL query language.
SISSVoc is part of the Spatial Information Services Stack (SISS), developed by CSIRO (Commonwealth Scientific and Industrial Research Organisation). In the references at the end of this presentation I've included a link to a paper describing its origins and design.

# Skosmos

## Skosmos

Open source web-based SKOS browser and publishing tool

### Features

- search and browse vocabularies
- alphabetical index
- thematic index
- structured concept display
- visualized concept hierarchy
- multilingual user interface

Try the Skosmos demo!

### Use cases

Access to controlled vocabularies for indexing, information retrieval and vocabulary development.

### Current users

- Finto
- FAO / AGROVOC
- Rhineland-Palatinate spatial data initiative classifications
- UNESCO Thesaurus
- University of Oslo Library thesauri

finto

16 40
THE NATIONAL
LIBRARY
OF FINLAND

http://skosmos.org/

Although ANDS doesn't use Skosmos, it is an example of a vocabulary publishing platform that is well worth mentioning. Skosmos has its origins in Finland.

In Finland there was a law passed in 2011 concerning interoperability of information systems in the public sector. It required that everyone should work with agreed definitions. The law appeared, but the actual terminological work had not been undertaken. There followed considerable project activity.

Given the National Library of Finland's previous experience in providing access to information through descriptive metadata, they were identified as the agency best placed to work in the area of descriptive interoperability metadata. The Finto project focused on the publication, development and use of controlled vocabularies to serve the whole of the public sector including memory institutions, government as well as media and research. It has been quite an undertaking, and I'm sure much valuable experience has been learned and expertise developed. Experience and learning no doubt of interest to those involved in COAR vocabulary initiatives.

Skosmos is a vocabulary publishing tool used by finto, providing vocabulary access for humans and machines. In the links at the end of this presentation I'll include a paper describing work undertaken by the NLF to convert traditional thesauri to SKOS and another paper giving background on Skosmos.

# Research Vocabularies Australia (RVA)

- ● ANDS provides software and support for

  - ○ Creating and editing
    - ■ Edit existing vocabularies or create new

  - ○ Publishing
    - ■ Description, access, licensing, versioning

  - ○ Discovery
    - ■ Metadata harvested by google

  - ○ Integration
    - ■ Application programming interface and widget

Onto vocabulary discovery

# Discovery

RVA helps to make vocabularies more discoverable. The metadata from the RVA publishing portal is harvested by Google. I can search Google for marine platform vocabularies and find the AODN example. The AODN Platform vocabulary is hosted and published by RVA.

# Discovery

Google points me back to the landing page for the vocabulary.  I can see what it is, what it is about, access conditions and how to get access.  Because the vocabulary is hosted by RVA, I can browse the vocabulary through the web portal.
RVA provides information that helps me to make an informed choice about whether the vocabulary may be suitable for my purposes.

# Research Vocabularies Australia (RVA)

- ● ANDS provides software and support for

  - ○ Creating and editing
    - ■ Edit existing vocabularies or create new

  - ○ Publishing
    - ■ Description, access, licensing, versioning

  - ○ Discovery
    - ■ Metadata harvested by google

  - ○ Integration
    - ■ Application programming interface and widget
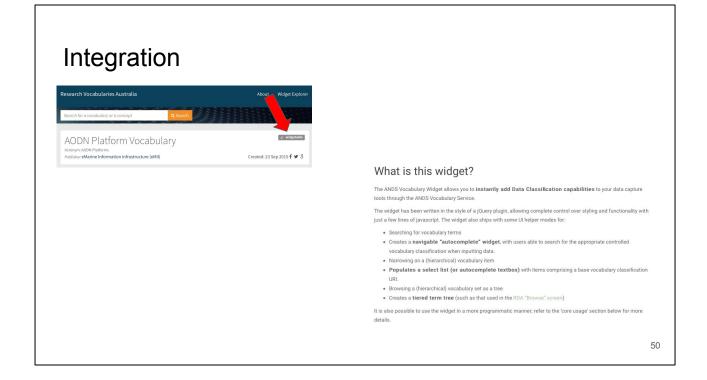
Onto integration

# Integration

I've talked about how RVA services can be used to support vocabulary management, publication and discovery. RVA services can also be integrated with local services.

As an example, the Australian Ocean Data Network (AODN) has a portal which provides access to marine and climate science data. Not surprisingly, the AODN portal uses AODN vocabularies. If the vocabularies are hosted by ANDS, how do they end-up in the system that drives the AODN portal?

# Integration

The answer lies in the fact that RVA provides technical means to draw vocabularies into local systems. Access can be provided by an application programming interface (API), a sparql endpoint or a widget.

As an example, AODN edit and publish their vocabularies using RVA services. When a new version of a vocabulary is published, the AODN portal uses the RVA application programming interface to get hold of it. This publishing workflow allows AODN to focus on their core services, rather than having to maintain additional services for vocabulary creation and management. They use the ANDS services to maintain their vocabularies and to get access as needed.

# Integration



## What is this widget?

The ANDS Vocabulary Widget allows you to **instantly add Data Classification capabilities** to your data capture tools through the ANDS Vocabulary Service.

The widget has been written in the style of a jQuery plugin, allowing complete control over styling and functionality with just a few lines of javascript. The widget also ships with some UI helper modes for:

- Searching for vocabulary terms
- Creates a **navigable "autocomplete" widget**, with users able to search for the appropriate controlled vocabulary classification when inputting data.
- Narrowing on a (hierarchical) vocabulary item
- **Populates a select list (or autocomplete textbox)** with items comprising a base vocabulary classification URI.
- Browsing a (hierarchical) vocabulary set as a tree
- Creates a **tiered term tree** (such as that used in the RDA "Browse" screen)

It is also possible to use the widget in a more programmatic manner; refer to the 'core usage' section below for more details.

50

Apart from an API and SPARQL endpoint, RVA provides a widget which may be incorporated into a web-form to provide data classification capability to local data capture and description systems. The widget has a number of configuration options including search, autocomplete and browse a hierarchical tree.

# Integration

There is also a widget explorer where you may try out the widget using vocabularies hosted by the ANDS service.

# RVA service components. Pick & mix

| | Edit using PoolParty | Publish to portal (available for reuse) | Describe in portal (without publishing) |
|---|---|---|---|
| 1 | | | ✓ |
| 2 | ✓ | ✓ | |
| 3 | | ✓ | |
| 4 | ✓ | | |
| | | | |

I've talked about the RVA components, and the AODN example provides an illustration of how those components can work together.  Editing in PoolParty, publishing to the RVA portal, integrating with the AODN portal via the RVA API.

That's not the only way to use RVA services

The RVA services can be used separately or in combination.  Here are a few possible examples, all of which are valid.

1.  I know of a vocabulary which I want more people to know about.   I can describe the vocabulary in the RVA portal.  The description can include a link to wherever the vocabulary is published.
2.  I want to edit a vocabulary using PoolParty, and I also want to publish that vocabulary to the RVA portal so that others may access and reuse it.
3.  I want to publish a vocabulary and make it accessible for reuse.  I don't want to use the PoolParty editor.  I already have tools for editing my vocabulary
4.  I want to edit a vocabulary.  I'm not in a position to publish it.  Maybe it will become publicly accessible, or maybe it has to remain in-house.  I don't know at the moment, but I want to get started

I don't have to use the editor to use the publishing portal.  I don't have to use the

publishing portal to use the editor.  I can edit a vocabulary and choose not to make it publicly accessible.  I can upload a vocabulary to the publishing portal without using the ANDS editing software.
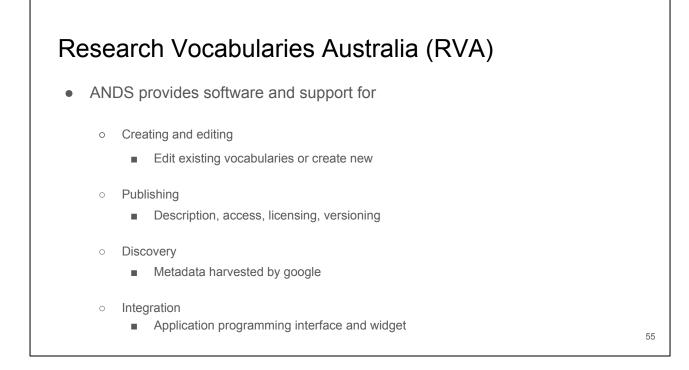
# How to access RVA services

- RVA portal
    - Anyone may search, browse and access the vocabularies described in the RVA portal
    - Anyone may self-register to describe or upload a vocabulary.
- RVA software interfaces
    - Anyone may use the linked data API, SPARQL endpoint and widget
- PoolParty editor
    - Commercial software licensed for use by Australian research organisations including
        - universities
        - research institutes
        - collecting organisations (such as libraries, archives, galleries, museums)
        - government agencies

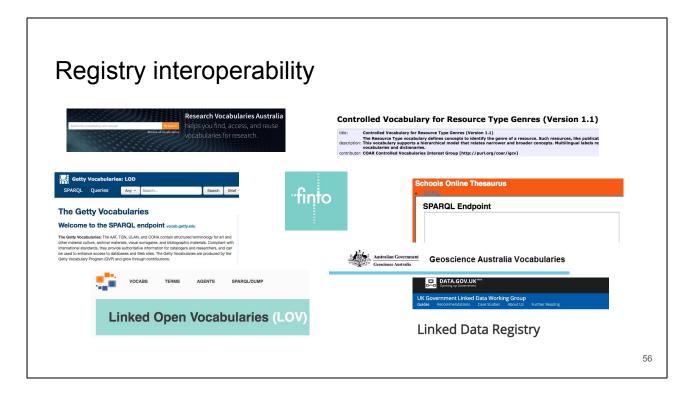Here is some information about how to access RVA services.

The RVA portal and machine interfaces are accessible to anyone.  Anyone may use vocabularies published in RVA according to the license conditions outlined in the description of each vocabulary.
The PoolParty editor is available to Australian research organisations including collecting organisations, government agencies and universities.

# Research Vocabularies Australia (RVA)

- ANDS provides software and support for

  - Creating and editing
    - Edit existing vocabularies or create new

  - Publishing
    - Description, access, licensing, versioning

  - Discovery
    - Metadata harvested by google

  - Integration
    - Application programming interface and widget

That was an overview of the ANDS vocabulary service and its tools and technologies.

Next I'll talk about vocabulary registry interoperability.
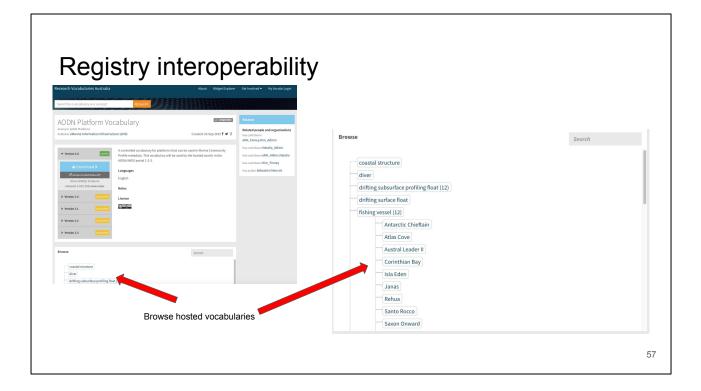
# Registry interoperability

Anecdotally It appears that activities are increasing regarding publishing of vocabularies in structured format for use by people and machines. Here are a few examples of organisations and initiatives that are either routinely providing access to vocabularies, or are in the process of developing services.

ANDS has an interest in interoperability and how various registries may relate to each other. As an example, RVA supports search for vocabularies and individual concepts. I can search RVA to find out if there are any vocabularies in RVA whose descriptive metadata contains the terms that I've searched on. This search will also retrieve concepts whose labels contain one or more of the terms that I've searched on. What if I could search a registry of vocabularies and discover not just vocabularies hosted by that service, but information about vocabularies hosted elsewhere?

Another example concerns the ANDS vocabulary widget. It has been developed to work with vocabularies hosted by ANDS. Would it be feasible for it to work with vocabularies hosted by other registries? Could I drop the ANDS vocabulary widget in a web form and have it retrieve terminology hosted by COAR?
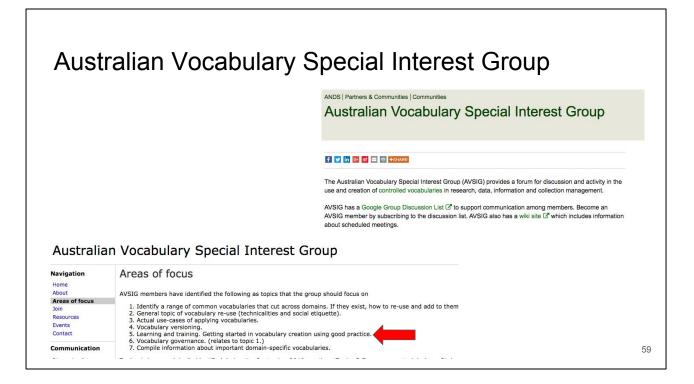
# Registry interoperability

A further example concerns the ANDS vocabulary browse interface. Each vocabulary hosted by RVA has a page of metadata, and a navigable browse tree. I can point and click to drill into a vocabulary and have a look to see if it may suit my needs.

Registry interoperability

This vocabulary is described but not hosted
Top concepts recorded as flat text

This is an example of a vocabulary which is described in RVA, but not hosted by RVA. The descriptive metadata was copied and pasted from the COAR site, and the top concepts were manually input as text. This vocabulary cannot be browsed further via the RVA user interface. What if I could browse the COAR vocabulary via the ANDS web interface, even though the COAR vocabulary is not published by RVA?

ANDS is planning to provide API access to registry metadata. This will allow our partners to push vocabularies and descriptive metadata to the ANDS service without needing to manually input via the RVA user interface. They'll also be able to use the API to edit existing RVA metadata.

Perhaps there are options around harvesting descriptive metadata from other vocabulary registries for inclusion in RVA. This would certainly lessen the maintenance task involved in keeping descriptive metadata and links up to date.

# Australian Vocabulary Special Interest Group

## Australian Vocabulary Special Interest Group

The Australian Vocabulary Special Interest Group (AVSIG) provides a forum for discussion and activity in the use and creation of controlled vocabularies in research, data, information and collection management.

AVSIG has a Google Group Discussion List ⬚ to support communication among members. Become an AVSIG member by subscribing to the discussion list. AVSIG also has a wiki site ⬚ which includes information about scheduled meetings.

### Australian Vocabulary Special Interest Group

**Navigation**
Home
About
**Areas of focus**
Join
Resources
Events
Contact

**Communication**

### Areas of focus

AVSIG members have identified the following as topics that the group should focus on

1. Identify a range of common vocabularies that cut across domains. If they exist, how to re-use and add to them
2. General topic of vocabulary re-use (technicalities and social etiquette).
3. Actual use-cases of applying vocabularies.
4. Vocabulary versioning.
5. Learning and training. Getting started in vocabulary creation using good practice.
6. Vocabulary governance. (relates to topic 1.)
7. Compile information about important domain-specific vocabularies.

59

Finally I'll mention the Australian Vocabulary Special Interest group. AVSIG was established a few months ago and is intended to provide a forum for discussion and activity concerning vocabularies. Although it is an Australian group, anyone from anywhere is welcome to join. Membership is open to anyone from Australian and international research organisations, including universities, research institutes, collecting organisations such as libraries, museums and galleries, and government agencies. Membership is also open to people from commercial organisations such as publishers and database vendors, and industry more broadly.

We have a quarterly videoconference and there have been some good presentations so far. We record presentations and make the videos, along with meeting discussion notes available online.

Although supported by ANDS, the group's direction is intended to be driven by its members. People identified a number of areas of interest, and this helps in organising discussions and events around those topics. One of these areas of interest is in learning and training - getting started in vocabulary creation, and ensuring good practice. I'll shortly be calling for expressions of interest from AVSIG members to work on this area of learning and training. If anyone attending this webinar is interested, you'd be welcome to join AVSIG and participate.

# Links

- Key Choices in the Design of Simple Knowledge Organization System (SKOS)
  - https://arxiv.org/abs/1302.1224
- SKOS Primer
  - https://www.w3.org/TR/skos-primer/
- ANDS Research Vocabularies Australia (RVA)
  - https://vocabs.ands.org.au/
  - https://documentation.ands.org.au/display/DOC/Research+Vocabularies
- SISSVoc: A Linked Data API for access to SKOS vocabularies
  - http://www.semantic-web-journal.net/system/files/swj880.pdf
- Publishing SKOS Vocabularies with Skosmos
  - skosmos.org/publishing-skos-vocabularies-with-skosmos.pdf
- Reuse of library thesaurus data as ontologies for the public sector
  - http://library.ifla.org/819/1/086-lappalainen-en.pdf
- Australian Vocabulary Special Interest Group (AVSIG)
  - http://www.ands.org.au/partners-and-communities/ands-communities/avsig

Thanks for your attention.  Here are some links relating to the presentation.  That concludes the presentation.