

Taxonomy standards and architecture

A brief introduction to SKOS

Taxonomy Boot Camp London

17 October 2018

Presented by

Heather Hedden

Senior Vocabulary Editor

Gale, A Cengage Company

Controlled Vocabulary Standards

Standards are of two basic types

1. *Standards for design* - supports an expected experience and results by varied users without requiring training
2. *Standards for specifications* (measurements, protocols, coding, etc.) – supports exchange and interoperability.

Standards for controlled vocabularies of each type

1. *Standards for design*: ISO 25964 Information and documentation - Thesauri and interoperability with other vocabularies;
ANSI/NISO Z39.19 Guidelines for Construction, Format, and Management of Monolingual Controlled Vocabularies
2. *Standards for specifications and interoperability*: Dublin Core, MARC, ZThes, DD 8723-5, RDF, OWL and **SKOS**)

<http://accidental-taxonomist.blogspot.com/2017/06/standards-for-taxonomies.html>

SKOS Background

SKOS (Simple Knowledge Organization System) model

- A World Wide Web (W3C) recommendation.
- Released in 2005 as a working draft and in 2009 as a recommendation.
- “A common data model for sharing and linking knowledge organization systems via the Web” and
“A data-sharing standard, bridging several different fields of knowledge, technology and practice” <https://www.w3.org/TR/skos-reference/>
- Encoded using RDF (Resource Description Framework) as a triple (subject-predicate-object).
- Can be expressed in various serialization formats: XML, N3, Turtle, JSON-LD.
- To enable easy publication and use of such vocabularies as linked data.
- Can also be queried using standard query language SPARQL.



W3C Home | SWDWG Home

SKOS Home
Introduction to SKOS
Frequently Asked Questions
Use Cases & Requirements
Specifications & Documentation
RDF Vocabularies
Find & Share Data
Software Tools & Applications
Development & Participation
Proposals & Issues
Tutorials, Presentations & Papers
Validation Services
Translations
Version History
Mailing List & Contact

SKOS Simple Knowledge Organization System - Home Page

SKOS is an area of work developing specifications and standards to support the use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading lists and taxonomies within the framework of the Semantic Web ... [\[read more\]](#)

Announcements

Alignment between SKOS and new ISO 25964 thesaurus standard (2012-12-13)

ISO 25964-1, published in 2011, replaced the previous thesaurus standards ISO 2788 and ISO 5964 (both now withdrawn). Members of the Working Group responsible for ISO 25964 have gone on to consider the implications for SKOS users. They have developed a set of linkages between the elements of the ISO 25964 data model and the ones from [SKOS](#), [SKOS-XL](#), and [MADS/RDF](#). This new "[Correspondence between ISO 25964 and SKOS/SKOS-XL models](#)" document can be accessed at <http://www.niso.org/schemas/iso25964/> and replaces the [previous table](#) presented in the SKOS Primer.

From Chaos, Order: SKOS Recommendation Helps Organize Knowledge (2009-08-18)

Today W3C announces a new standard that builds a bridge between the world of knowledge organization systems - including thesauri, classifications, subject headings, taxonomies, and folksonomies - and the linked data community, bringing benefits to both. Libraries, museums, newspapers, government portals, enterprises, social networking applications, and other communities that manage large collections of books, historical artifacts, news reports, business glossaries, blog entries, and other items can now use **Simple Knowledge Organization System (SKOS)** to leverage the power of linked data. The [Semantic Web Deployment Working Group](#) also published today two Group Notes with the Recommendation, updating the [SKOS Primer](#) and [SKOS Use Cases and Requirements](#). Read the [press release](#) and [testimonials](#) and learn more about the [Semantic Web Activity](#). ([Permalink](#))

Current specifications...

- [SKOS Reference](#) W3C Recommendation, 18 August 2009
- [SKOS Primer](#) W3C Working Group Note, 18 August 2009
- [SKOS Use Cases and Requirements](#) W3C Working Group Note, 18 August 2009
- SKOS RDF files: see the RDF vocabularies section

SKOS has been developed by the W3C Semantic Web Deployment Working Group (SWDWG).

SKOS Elements

Concepts	Labels & Notation	Documentation	Semantic Relations	Mapping Properties	Collections
Concept	prefLabel	note	broader	broadMatch	Collection
ConceptScheme	altLabel	changeNote	narrower	narrowMatch	orderedCollection
inScheme	hiddenLabel	definition	related	relatedMatch	member
hasTopConcept	notation	editorialNote	broaderTransitive	closeMatch	memberList
topConceptOf		example	narrowerTransitive	exactMatch	
		historyNote	semanticRelation	mappingRelation	
		scopeNote			

Example URI: `skos:prefLabel`

SKOS in Elements in XML

Example concept from the NASA Thesaurus in SKOS in XML

<https://www.sti.nasa.gov/nasa-thesaurus/#.WoG0uainGM8>

```
<skos:Concept rdf:about="#37804">
  <skm:termUpdate>add</skm:termUpdate>
  <skos:prefLabel>A-3 aircraft</skos:prefLabel>
  <skos:altLabel>A3D aircraft</skos:altLabel>
  <skos:altLabel>Skywarrior aircraft</skos:altLabel>
  <zthes:termID>37804</zthes:termID>
  <zthes:termVocabulary>NASA Thesaurus</zthes:termVocabulary>
  <skm:UF rdf:resource="#182148" rdf:ID="r37804-182148" />
  <skm:UF rdf:resource="#185586" rdf:ID="r37804-185586" />
  <skos:broader rdf:resource="#39556" rdf:ID="r37804-39556" />
  <skos:broader rdf:resource="#41969" rdf:ID="r37804-41969" />
  <skos:broader rdf:resource="#45801" rdf:ID="r37804-45801" />
  <skos:broader rdf:resource="#47765" rdf:ID="r37804-47765" />
  <skos:related rdf:resource="#38167" rdf:ID="r37804-38167" />
</skos:Concept>
```

A-3 aircraft	
UF	A3D aircraft Skywarrior aircraft
GS	attack aircraft . bomber aircraft . . A-3 aircraft jet aircraft . A-3 aircraft McDonnell Douglas aircraft . Douglas aircraft . . A-3 aircraft monoplanes . A-3 aircraft
RT	∞ aircraft

<https://www.sti.nasa.gov/thesvol1.pdf>

SKOS in Elements in Vocabulary Management Software

The screenshot displays the 'Concept Manager' interface. On the left is a navigation tree with 'Botany' selected under 'Training Thesaurus'. The main area shows the 'Botany' concept details, including its assigned class 'Concepts Class' and 'SKOS/OWL'. The 'Top Concept In Scheme' is 'Training Thesaurus'. The interface is divided into several sections:

- Narrower (SKOS):** Lists 'Applied botany' and 'Research Botany'.
- Related (SKOS):** Lists 'Botany'.
- sameAs (OWL):** Currently empty.
- AP Subjects Def. (Property Paths):** Contains a definition: 'The study of plant life. Includes research, commentary, announcements, publications, exhibits, and content aimed at enthusiasts and the general public. See "Forestry" for the science of the management and conservation of forests.'
- Preferred Labels:** Shows 'Botany' as the preferred label in English.
- Alternative Labels:** Shows 'Plant science' as an alternative label in English.
- Properties:** Shows the source 'http://dbpedia.org/resource/Botany'.
- Definition (SKOS):** Shows a partial definition: 'the scientific study of plants, including their physiology, structure, genetics, ecology'.

Red and blue boxes highlight the 'Narrower (SKOS)', 'Related (SKOS)', 'Preferred Labels', and 'Alternative Labels' sections respectively. A white box at the bottom left contains the text 'SKOS model in Synaptica'.

SKOS model in Synaptica

Budgeting

 Add to Collection

 Add to Blacklist

 Delete Concept

 <http://advanced.poolparty.biz/FoodandRecipes/94>

Details Notes Documents Linked Data Triples Visualization Quality Management History

SKOS  +

Broader Concepts

[Finance](#)



Narrower Concepts

[Budget analysis](#)



Related Concepts




[Accounting](#)



Top Concept of Concept Schemes




Preferred Label


 Budgeting


en

Alternative Labels

 Budget management

en

 Budget planning

 Budgets



Hidden Labels



Scope Notes



Definitions



SKOS model in PoolParty

Thesaurus Model vs. SKOS Model

A taxonomy or thesaurus comprises terms/concepts

Thesaurus model:

Terms

Of two *types*:

1. Preferred terms
2. Nonpreferred terms

Content is indexed/tagged to preferred terms only.

SKOS vocabulary model:

Concepts

Which each *have*:

- Preferred labels (one in each language)
- Alternative labels

Content is indexed/tagged to concepts as described by preferred labels.

The preferred term/label is displayed. Selecting a nonpreferred term or alternative label redirects and points to the associated concept/preferred term.

Thesaurus Model vs. SKOS Model

Hierarchical relationships

- Thesaurus designation of BT / NT (Broader term / Narrower term)
- SKOS designation: Broader concept / Narrower concept

Associative relationships

- Thesaurus designation of RT (Related term)
- SKOS designation: Related concept

SKOS for Multilingual Vocabularies

Terms in multiple languages

- Most vocabulary management software supports multilingual vocabularies.
- The SKOS model is especially suited for multilingual vocabularies.
 - A single concept has a preferred label in each language and alternative labels in each language.
 - Preferred labels must be exact translations for the vocabulary to function bi-directionally.
 - Alternative labels do not link to translations of each other, so may vary for each language.

Example: EuroVoc - Multilingual Thesaurus of the European Union

Alphabetical

Hierarchy

- └─ canned vegetables
- └─ deep frozen products
- └─ dried products
 - └─ dried fish
 - └─ dried fruits
 - └─ prunes
 - └─ raisins
 - └─ dried meat
 - └─ dried milk
 - └─ dried vegetables
 - └─ dried whey
 - └─ freeze dried products
 - └─ soup mixes
- └─ fermented products
- └─ fresh-cut produce
- └─ frozen products
- └─ geotextiles
- └─ prepared foods
 - └─ icecream
 - └─ **infant foods**
 - └─ instant foods
 - └─ soyfoods
- └─ processed animal products
- └─ processed foods
 - └─ desserts

products > foods > prepared foods > infant foods

products > processed products > prepared foods > infant foods

PREFERRED TERM

infant foods

BROADER CONCEPT

prepared foods

ALTERNATIVE LABEL

*baby foods**infant formulas**weaning foods*

IN OTHER LANGUAGES

أغذية الأطفال

Arabic

婴儿食品

Chinese

婴幼儿食品

离乳食品

婴儿配方食品

kojenecká výživa

Czech

*výživa pro děti**dětská strava**výživa pro kojence***Aliment pour nourrisson**

French

*Aliment pour bébé**Aliment infantile**Lait infantile**Aliment de sevrage pour enfant***Kindernahrungsmittel**

German

Kleinkindernahrung

SKOS model: Alternative labels and other languages

http://oek1.fao.org/skosmos/agrovoc/en/page/c_3850

SKOS Hidden Labels

Displayed vs. non-displayed variants

SKOS model also has **Hidden Label** (hidden Label) for these uses.

Non-displayed variants are useful for:

- Common misspellings, slang, or deprecated, or potentially offensive terms not displayed to users but can match searches
- Auto-categorization support but not intended for manual indexing
- Search support but not intended for type-ahead display
- Changed preferred name, making the former preferred name nonpreferred, but not meeting the criteria for a displayed variant. **Pates (Food) USE Pates.**

Questions/contact

Heather Hedden

Senior Vocabulary Editor

Gale, A Cengage Company

20 Channel Center Street

Boston, MA 02210 U.S.A.

heather.hedden@cengage.com

www.cengage.com, www.gale.com

www.cengage.co.uk, www.gale.cengage.co.uk

www.hedden-information.com

<https://accidental-taxonomist.blogspot.com>