

# Information—Integration—Intelligence Solutions



## SKOS and Linked Data

**Bob DuCharme**  
**Government Linked Data Working Group**  
**June 29-30, 2011**



# Introductions

- Representing myself, not TopQuadrant
- Deliverable 2.3 Best Practices for Publishing Linked Data
  - Issue 2: Vocabulary Selection
  - Issue 6: Legacy Data



# Two points

- There are a lot of thesauri, taxonomies, and controlled vocabularies in government data
- For people who work with them and are otherwise intimidated by the full stack of linked data/semweb technology, SKOS can be an easy sell



# Searching usa.gov for “taxonomy”

GOVERNMENT > **Web** Images Recalls Forms Change Text Size **Buscador.USA.gov**

**Search.USA.gov**  **SEARCH**

Advanced Search | Busque en español

*Related Topics*

- Education
- National Center for Biotechnology Information
- Nomenclature
- Scientific Classification
- Soil Classification

Results 1-10 of about 332,000 for 'taxonomy' Web results by **bing**

**Taxonomy home - National Center for Biotechnology Information**  
 Provides a **taxonomy** browser, **taxonomy** resources and other information.  
<http://www.ncbi.nlm.nih.gov/ta...> - Cached

Ascomycota	PubMed
Tree	All Databases
Common Taxonomy Tree	Nucleotide
Homo Sapiens	Disclaimer

**NCBI Taxonomy Homepage - National Center for Biotechnology Information**  
 The NCBI **Taxonomy** database is a curated set of names and classifications for all of the organisms that are represented in GenBank.  
<http://www.ncbi.nlm.nih.gov/.../taxonomyhome.html> - Cached

**NCBI: Taxonomy - National Center for Biotechnology Information**  
 Databases NCBI Website Search. A database of static NCBI web pages, documentation, and online tools. These pages include such content as specialized online sequence analysis ...  
<http://www.ncbi.nlm.nih.gov/.../taxonomy> - Cached

**Taxonomy Medicare Provider-Supplier Enrollment**  
 The Healthcare Provider **Taxonomy** Code Set is a hierarchical code set that consists of codes, descriptions, and definitions. Healthcare Provider **Taxonomy** Codes are ...  
[https://www.cms.gov/.../08\\_Taxonomy.asp](https://www.cms.gov/.../08_Taxonomy.asp) - Cached

**Taxonomy - Gonorrhoea - STD information from CDC**  
**Taxonomy** of the Family Neisseriaceae - Gonorrhoea Lab information from CDC.  
<http://www.cdc.gov/.../taxonomy.htm> - Cached

**Taxonomy - Arizona Department of Economic Security**  
**Taxonomy** Introduction. The Arizona Dictionary and **Taxonomy** of Human Services is a collaborative effort to develop a common language by major funders and planners throughout ...  
<http://www.azdes.gov/taxonomy/> - Cached

[PDF] **Using Taxonomy in ProviderOne**  
 P1PR009 - 04/10 Visit our website <http://hrsa.dshs.wa.gov/ProviderOne> Using **Taxonomy** in ProviderOne Phase 1 implementation The Department of Social and Health ...  
<http://hrsa.dshs.wa.gov/.../P1PR009%20taxonomy.pdf> - Cached

**MyFlorida.com - Taxonomy - Floridian**  
 Hot Topics  
<http://www.myflorida.com/.../floridian> - Cached

[PPT] **Taxonomy**  
 Carolus Linnaeus . Described organisms with two word names, instead of polynomials; Developed binomial nomenclature; First word = genus name; Second word = species name  
<http://step.nn.k12.va.us/.../taxonomy.ppt> - Cached



# Terminology

- Controlled vocabulary
- Taxonomy
- Thesaurus
- Ontology

- **(skipping nine slides)**



# Simple Knowledge Organization System

- Controlled vocabulary
- Taxonomy
- Thesaurus
- Ontology

SKOS: the W3C's OWL ontology for creating thesauri, taxonomies, and controlled vocabularies.

# Advantages of SKOS

- It's a standard
- Makes tracking of multi-lingual vocabularies much easier
- Extensible
- Easier to connect different vocabularies (URIs, ease of aggregating RDF, etc.)





# Who is using SKOS?

- AGROVOC
- New York Times: People, Organizations, Locations, Subject Descriptors
- Library of Congress subject headers
- AGFA drug admin. forms
- NASA: many categories

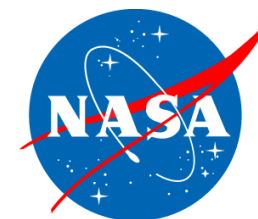


Food and Agriculture  
Organization of the  
United Nations

*for a world without hunger*

The  
New York  
Times

**AGFA** 



- TopQuadrant's Enterprise Vocabulary Net (EVN)
- PoolParty
- SKOSed Protégé plugin
- iQvoc
- TemaTres



# SKOS properties

prefLabel

broader

definition

altLabel

related

hasTopConcept

note

hiddenLabel

example

editorialNote

broadMatch

narrower

member

scopeNote

notation

narrowMatch

semanticRelation

historyNote

exactMatch

mappingRelation

changeNote

inScheme

closeMatch

topConceptOf

relatedMatch

narrowerTransitive

broaderTransitive

memberList



# SKOS properties (very minimal)

prefLabel  
broader  
definition



# SKOS properties (very minimal)

prefLabel

broader or definition

“If you have data where you’ve just defined a set of terms and either a tree relationship of them or just their definitions, that’s useful data to other people, especially if you publish it using a standard like SKOS.”

# SKOS properties

prefLabel  
broader  
definition  
altLabel  
related  
note

# SKOS properties

prefLabel

broader

definition

altLabel

related

note

foo:bar

foo:baz



# Message to evangelize

(For government agencies that already have taxonomies, thesaurii, dictionaries...)

“Data modeling (deciding what pieces of information to keep track of and their relationships) is the difficult part.”

*“You've already done that. SKOS gives you a standard way to share this work you've done so that others can take advantage of it.”*