

Introduction to SKOS

Bob DuCharme
October 6, 2011





Introductions

- Presentation and all its URLs:
<http://www.snee.com/skos/20111006/>
- Me: Solutions Architect at TopQuadrant;
formerly XML, SGML guy at RIA, Moody's,
LexisNexis, Innodata Isogen
- Weblog: <http://www.snee.com/bobdc.blog>

Also available...





Simple Knowledge Organization System

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



Terminology

- **Controlled vocabulary**
- Taxonomy
- Thesaurus
- Ontology



Controlled vocabularies

yes

no

Mr.

Ms.

Mrs.

Miss

Dr.

AL

AK

AR

CA

CO

CT

.

.

.

WI

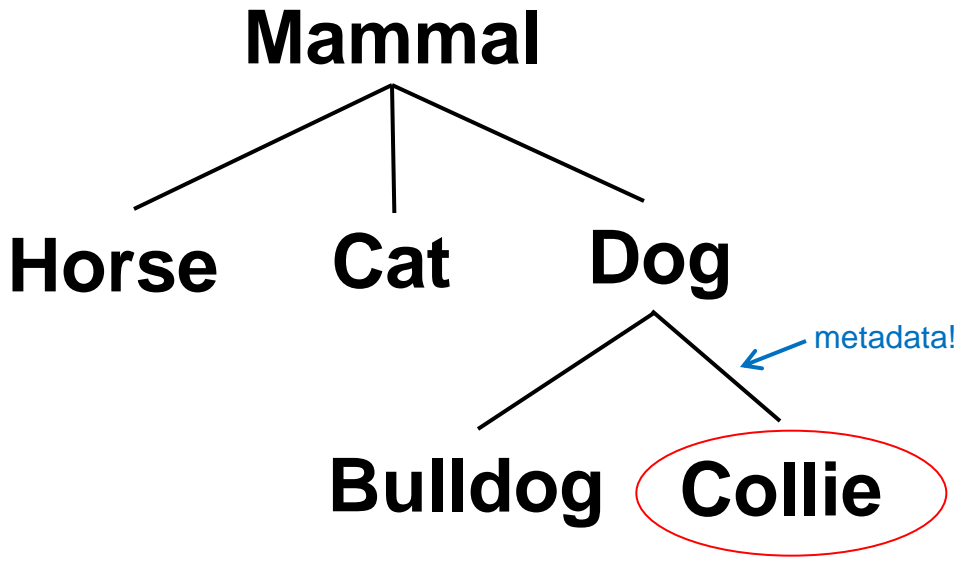
WY



Terminology

- Controlled vocabulary
- **Taxonomy**
- Thesaurus
- Ontology

Taxonomies



**Above: subset-of relationship.
Alternatives: part-of, instance-of.**



Taxonomies: business value

← → ↻ 🏠 🌐 www.lowes.com/pc_Lighting+Fans_4294935637_4294937087_ ☆ 🔍

Welcome to Lowe's | Log In or Register | Store Locator | My Account ▾ | 🛒 My Cart \$0.00 (0 items) | Checkout

LOWE'S
Let's Build Something Together™

Get Pricing & Availability Find My ZIP
Find a store near you for great selections at neighborhood prices.

Departments Savings Projects Home Ideas Help Center

SHOP. CLICK. PICK UP. 20 MINUTES GUARANTEED ▾

Home : Lighting & Fans Print

Lighting & Fans

- Bath Lighting (1083)
- Ceiling Fans & Accessories (1567)
- Ceiling Lights (2402)
- Fans (100)
- Fluorescent Lighting & Accessories (119)
- Lamps, Shades & Accessories (747)
- Light Bulbs (592)
- Lighting Accessories & Parts (2255)
- Night Lights (58)
- Outdoor & Landscape Lighting (1112)
- Security Lighting (130)
- Specialty & Accent Lighting (161)
- Timer & Lighting Controls (41)
- Wall Sconces (223)
- Work & Shop Lights (33)

Up to 50% Off Lighting
Brighten up your home décor this holiday season with great values on select lighting.

▶

Popular Collections

Tiffany Crystal Lyndsey Eastview Lola

New Designs in Ceiling Fans
The 2011 ceiling fans have

Harbor Breeze 52" Armitage White

Portfolio 7.9" Low-Voltage LED

New Lighting At Lowe's
New styles and designs have

Taxonomies: business value

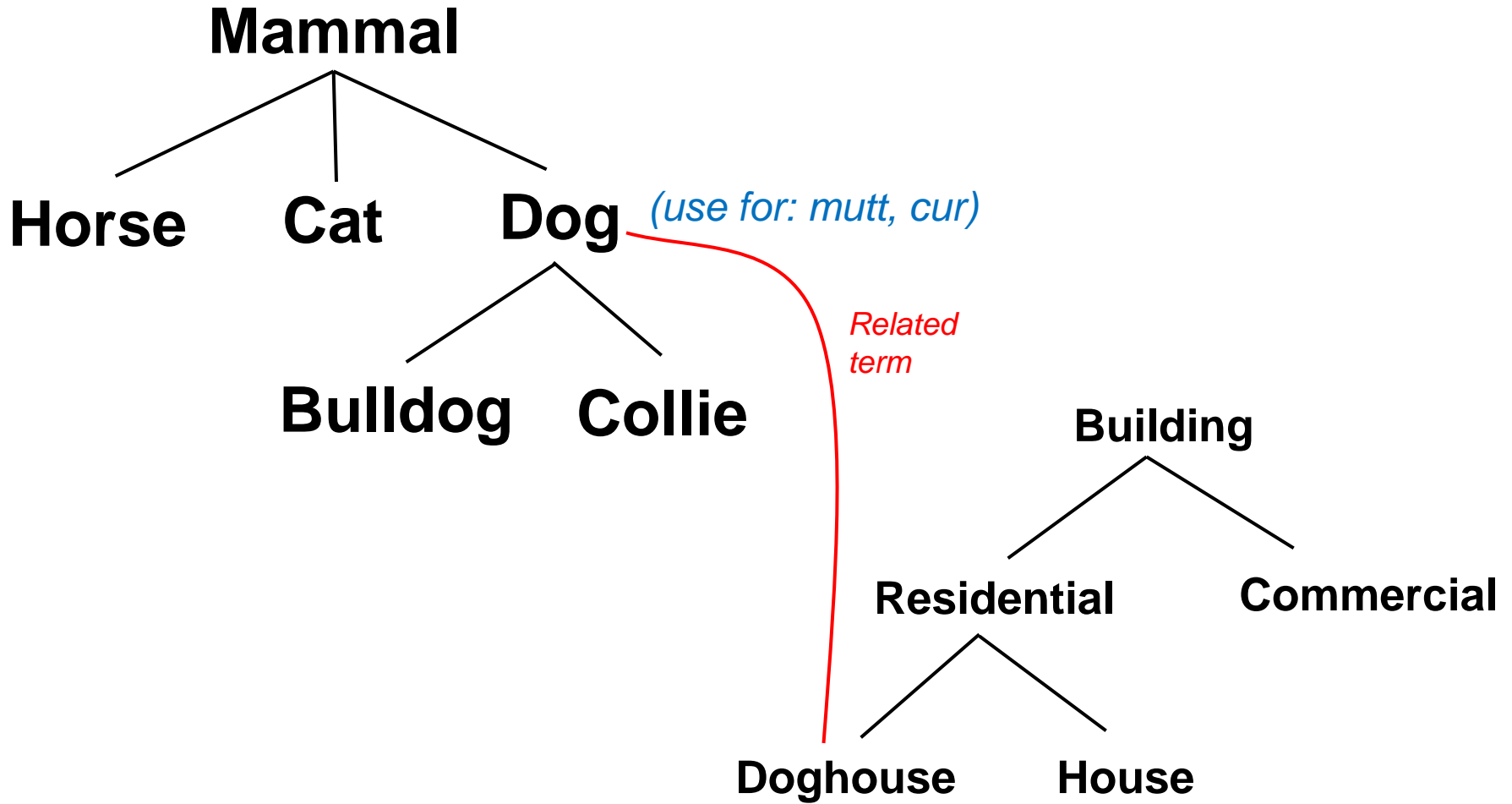
The screenshot shows the Home Depot website interface. At the top, there's a navigation bar with 'STORE FINDER', 'CALL 1 (800) HOME-DEPOT', 'SERVICES', and 'HELP'. Below this is a search bar and a 'MY CART' icon showing 0 items. A large promotional banner for 'SAVINGS ON MOST ORDERS OVER \$249' is visible. Below the banner, there's a section titled 'REDUCING THE COST TO REDO YOUR BATH' with a 'SPECIAL FINANCING FOR SELECT BATH PURCHASES' offer. A navigation menu is open, showing categories like 'INDOOR LIGHTING', 'OUTDOOR LIGHTING', and 'FANS'. The 'Lighting & Fans' category is highlighted. At the bottom, there's a promotional banner for '10% OFF ALL ENERGY STAR® QUALIFIED APPLIANCES \$398 OR MORE' with an image of two washing machines.



Terminology

- Controlled vocabulary
- Taxonomy
- **Thesaurus**
- Ontology

Thesaurus





Terminology

- Controlled vocabulary
- Taxonomy
- Thesaurus
- **Ontology**



Ontology

- Define terms and structures
- And relationships: Requires, contraindicates...
- Classes of things to keep track of, potential properties they might have
- OWL: W3C RDF-based standard for specifying ontologies

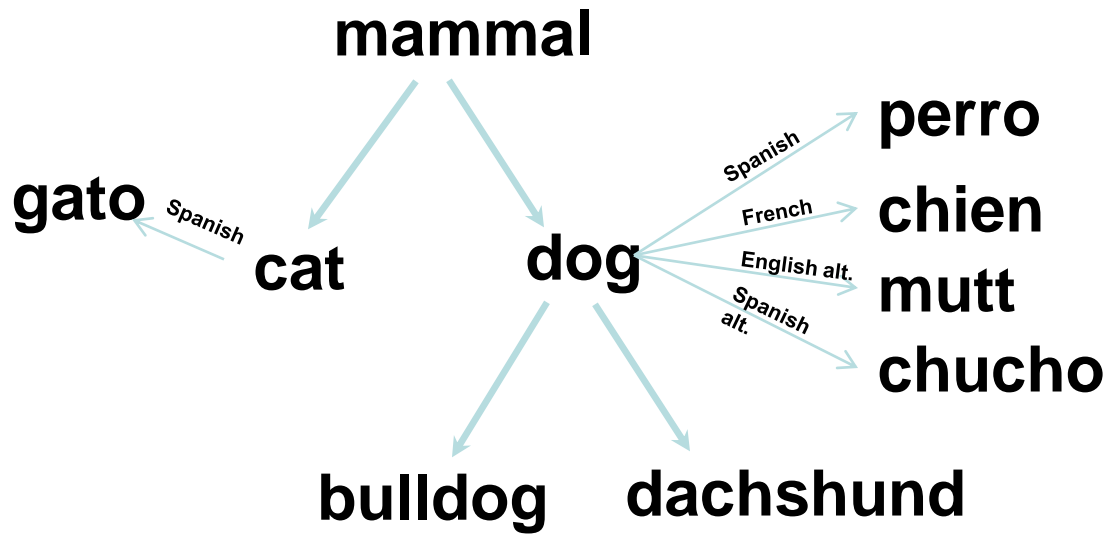


Simple Knowledge Organization System

- Controlled vocabulary
- Taxonomy
- Thesaurus
- Ontology

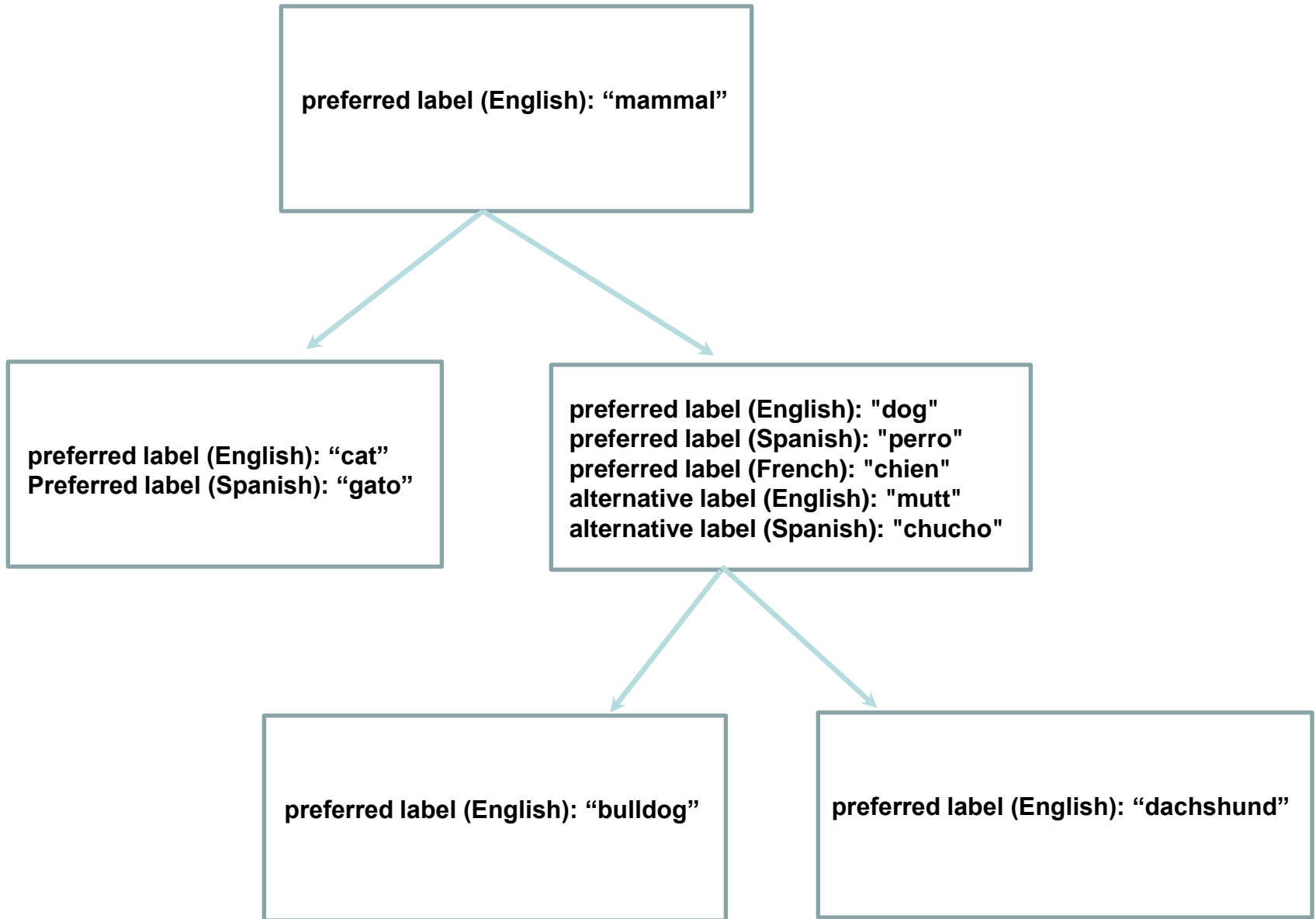
SKOS: the W3C's OWL ontology for creating thesauri, taxonomies, and controlled vocabularies.

Managing terms?





No, managing concepts





More metadata for concepts

- Standardized
 - Definition
 - Related concepts
 - History note
 - More...
- Unstandardized
 - Whatever you want!



More metadata: example

<http://myCompany.com/animals/c43209101>

preferred label (English): "dog"

preferred label (Spanish): "perro"

preferred label (French): "chien"

alternative label (English): "mutt"

alternative label (Spanish): "chucho "

history note: "Edited by Jack on 5/4/11 "

related term: <http://myCompany.com/shelters/c3048293>

product: <http://myCompany.com/vaccines/c2197503>



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



The
New York
Times

AGFA 





Other taxonomy standards

- **ANSI/NISO Z39.19-2005** Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies
- **ISO 2788:1986** Guidelines for the Establishment and Development of Monolingual Thesauri
- **ISO 5964:1985** Guidelines for the Establishment and Development of Multilingual Thesauri
- **ISO 25964** (combines 2788 and 5964)
- **Zthes**
- **MADS/RDF** (“Metadata Authority Description Schema in RDF”) Library of Congress

What is SKOS-XL?

<http://myCompany.com/animals/c43209101>

preferred label (English): "dog"

preferred label (Spanish): "perro"

preferred label (French): "chien"

alternative label (English): "mutt"

alternative label (Spanish): "chucho"

history note: "Edited by Jack on 5/4/11 "

related term: <http://myCompany.com/shelters/c3048293>

product: <http://myCompany.com/vaccines/c2197503>



Date of last edit?

Who edited it?

Concept

SKOS-XL: SKOS Extension for Labels

<http://myCompany.com/animals/c43209101>
preferred label: <http://myCompany.com/animals/c4329201>
preferred label: <http://myCompany.com/animals/c4329202>
alternative label: <http://myCompany.com/animals/c4329205>
history note: "Edited by Jack on 5/4/11 "
related term: <http://myCompany.com/shelters/c3048293>
product: <http://myCompany.com/vaccinations/c2197503>

Concept

<http://myCompany.com/animals/c4329201>
literal form: "dog" (English)

Label

<http://myCompany.com/animals/c4329205>
literal form: "chucho" (Spanish)
last edited: "2011-05-13"
editor: " Mary Jones"

Label



Your own metadata

<http://myCompany.com/animals/c43209101>

preferred label (English): "dog"

preferred label (Spanish): "perro"

preferred label (French): "chien"

alternative label (English): "mutt"

alternative label (Spanish): "chucho "

history note: "Edited by Jack on 5/4/11 "

related term: <http://myCompany.com/shelters/c3048293>

product: <http://myCompany.com/vaccinations/c2197503>

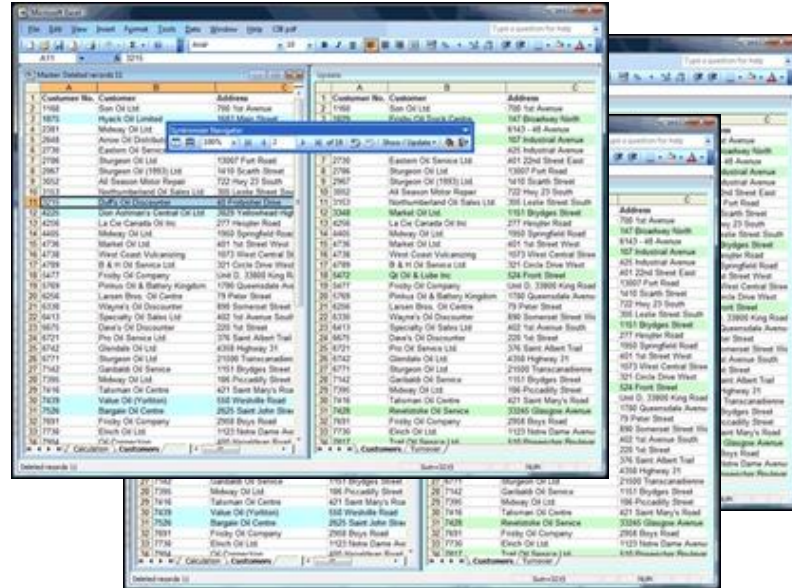
foo code: "5L-MN1-003"

Some SKOS using Turtle format

```
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
@prefix myco: <http://myCompany.com/vocab#> .
```

```
<http://myCompany.com/animals/c43290101>  
  skos:prefLabel "dog"@en ;  
  skos:prefLabel "perro"@es ;  
  skos:prefLabel "chien"@fr ;  
  skos:altLabel "mutt"@en ;  
  skos:altLabel "chucho"@es ;  
  skos:historyNote "Edited by Jack on 5/4/11" ;  
  skos:related <http://myCompany.com/shelters/c3048293> ;  
  myco:product <http://myCompany.com/vaccines/c2197503> ;  
  myco:fooCode "5L-MN1-003" .
```

Networked vocabularies: why

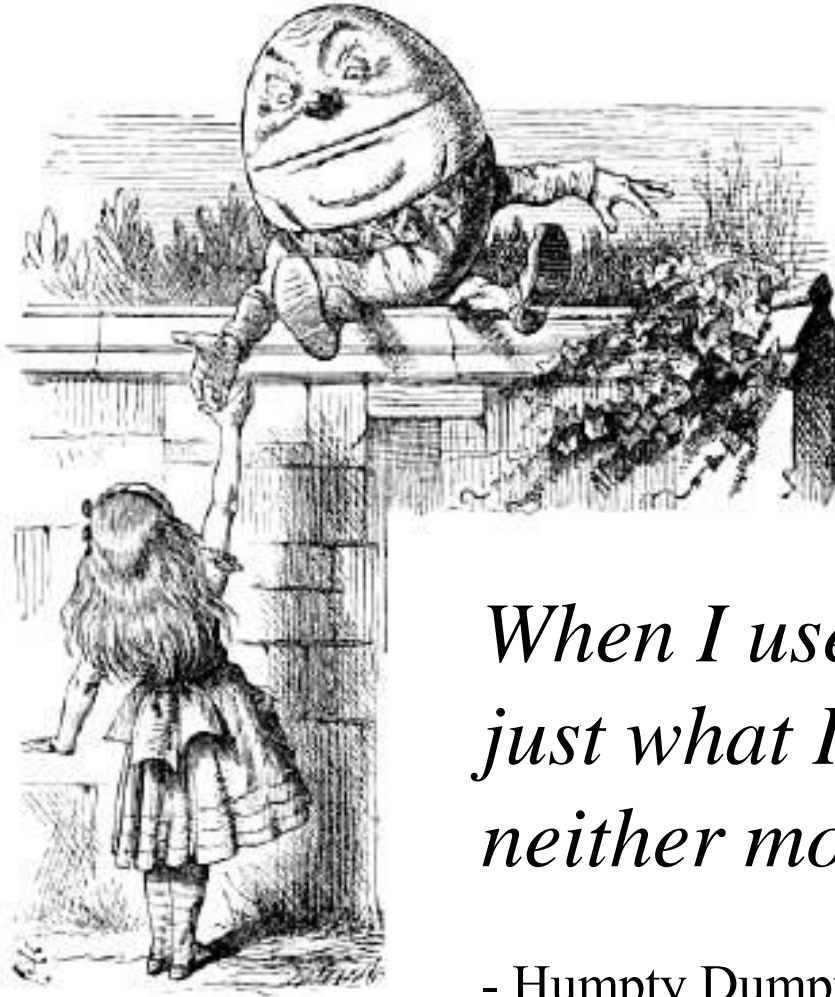


Centralized
master
vocabulary
system ?

Ad hoc,
distributed,
disconnected



Enterprise Vocabulary Management: dream vs. reality



*When I use a word, it means
just what I choose it to mean —
neither more nor less.*

- Humpty Dumpty in Lewis Carroll's *Through the Looking Glass*

Enterprise Vocabulary Management: dream vs. reality

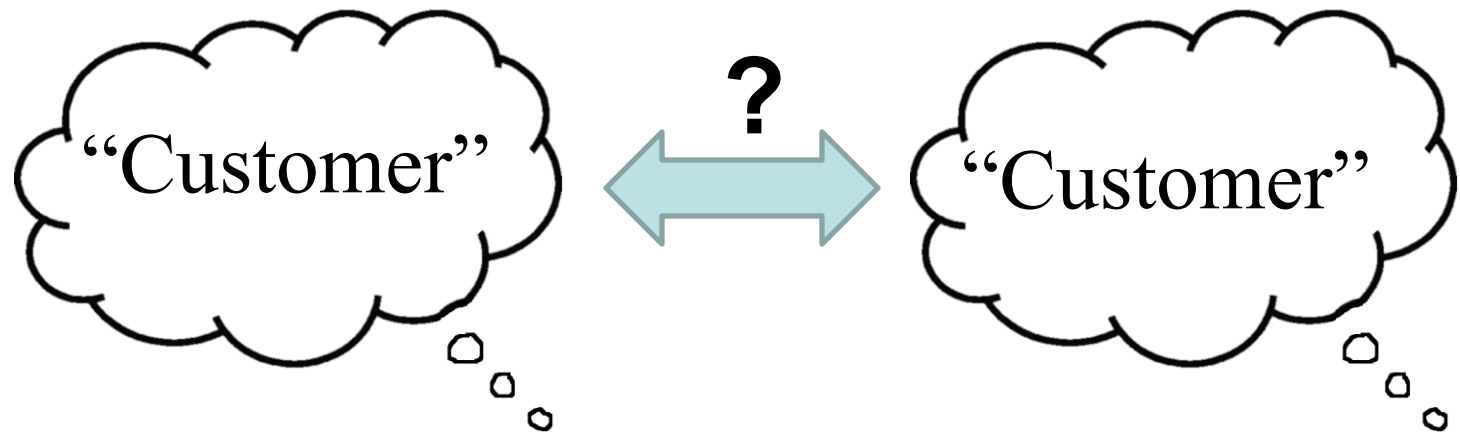


Repair Department



Marketing Department

Enterprise Vocabulary Management: dream vs. reality



Repair Department



Marketing Department

Networking distributed vocabularies

<http://myCompany.com/vocab/repairs/c8927432>

preferred label: " customer "

broader term: <http://myCompany.com/vocab/repairs/c93402934>

Concept

<http://myCompany.com/vocab/marketing/c93402934>

preferred label: "customer"

Concept

Networking distributed vocabularies

<http://myCompany.com/vocab/repairs/c8927432>

preferred label: " customer "

broader term: <http://myCompany.com/vocab/repairs/c93402934>

Concept

<http://myCompany.com/vocab/marketing/c93402934>

preferred label: "customer"

Concept

Repairs department and marketing department vocabularies can be stored in the same datastore on the same server... *or different datastores on different servers stored with different products.*

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



Other semweb tech to use

- Data conversion tools
- Ontology editors
- SPARQL
 - Query, update, transform
 - SPARQL Rules
- Linked Data
- OWL



SPARQL Rules and AGROVOC

- SKOS spec lists six rules not implemented by ontology
- e.g. same term can't be preferred and alternative term in the same language
- Expressed with SPARQL Rules
- Violated by AGROVOC over 1600 times
- e.g. Slovak word for “Buds” (“púèiky”)

Getting started

- Look at existing SKOS thesauri
- Small pilot project...
- ...or projects.
- Existing thesaurus editor vendor SKOS support?
- Free tools, commercial tools, trial software

SKOS Advantages

- Good multi-language support
- Extensible
- Scalable
- Commercial and open source tools available
- RDF basis: more tools
- Traction



TopBraid EVN

- Built-in and custom rules
- Intuitive graphical user interface
- Multi-user support
- Editing and review of multiple working copies
- Controlled role-based access
- Change tracking
- Support of all relevant W3C standards
- Import/Export and web services integration
- Easy extensibility of features



Questions?

Idiosyncratic
vocabularies



TopBraid
Enterprise Vocabulary Net™

Vocabulary
web

*“A wonderful harmony is created when we
join together the seemingly unconnected.”*

- Heraclitus

bducharme@topquadrant.com
info@topquadrant.com