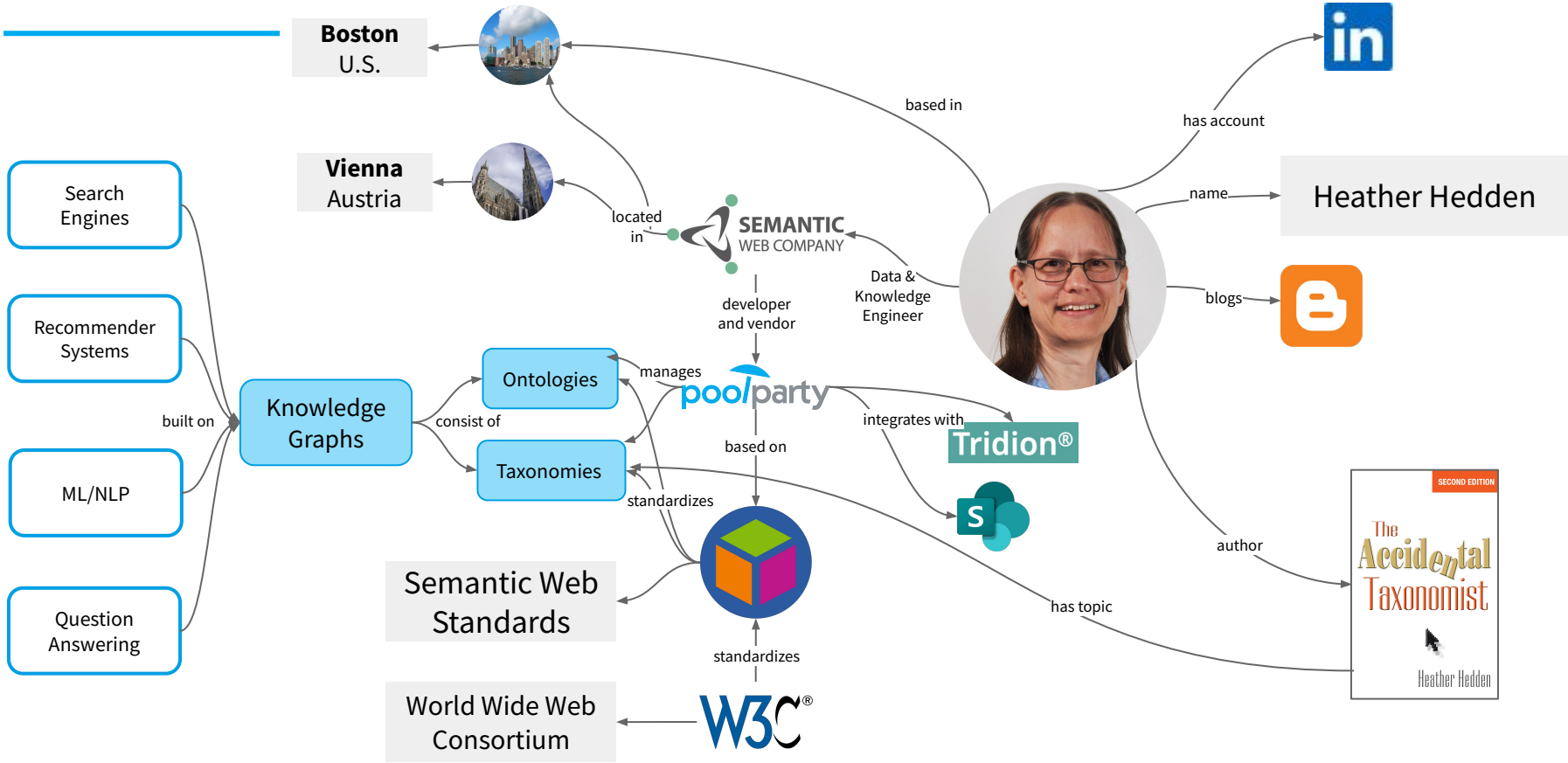
Taxonomy-Driven Ontology Design

Knowledge Graph Conference
May 5, 2022

Heather Hedden
Data and Knowledge Engineer,
Semantic Web Company

Click the Graph—get in contact with us!

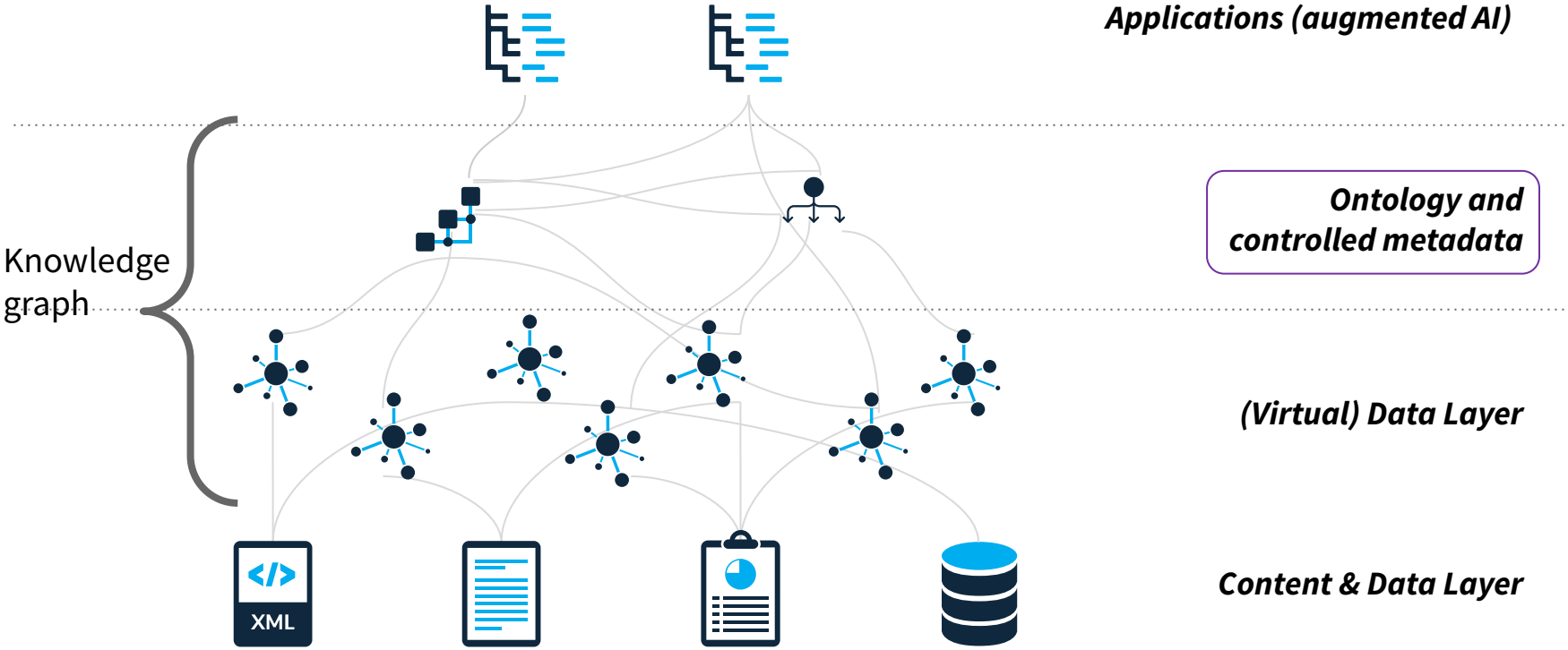


Outline

- Why taxonomies and ontologies
- Standards and features of ontologies and taxonomies
- Ontology types and approaches
- Semantically enriching a taxonomy to become an ontology

Why taxonomies and ontologies

Why Taxonomies and Ontologies

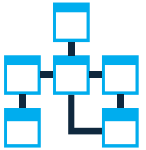


Why Taxonomies and Ontologies

What you can do with a taxonomy

Defined as: “a collection of controlled vocabulary terms organized into a hierarchical structure.”
- ANSI/NISO Z39.19 (r2010)

- Tagging: index content consistently so that retrieval is comprehensive and accurate
- Normalization: Bring together different names, localizations, languages for concepts
- Standard search: find content about…. (search string matches taxonomy concepts)
- Topic browse: explore subjects arranged in a hierarchy and then content on the subject
- Faceted (filtering/refining) search: find content meeting a combination of basic criteria
- Discovery: find other content tagged with same concepts as tagged to found content; explore broader, narrower, and (sometimes) related taxonomy topics
- Content curation: create feeds or alerts based on pre-set search terms
- Metadata management: for retrieval, identification, comparison, analysis, etc.



Why Taxonomies and Ontologies

What you cannot do with a taxonomy alone, but can with an added ontology

Defined as: “a set of representational primitives with which to model a domain of knowledge or discourse.” - Tom Gruber

- Modeling complex interrelationships (e.g. in product approval or supply chain processes) and also connect to content
- Complex multi-part searches: e.g. find contacts in a specific location, who are employed by companies which belong to certain industries
- Exploring explicit relationships between concepts (not just broader, narrower, related)
- Search across datasets, not just search for content
- Search on more specific criteria that vary based on category (class)
- Visualization of concepts and semantic relationships
- Reasoning and inferencing

Standards and features of taxonomies and ontologies

Standards: W3C Recommendations

For both ontologies and taxonomies/controlled vocabularies:

RDF (Resource Description Framework) www.w3.org/TR/rdf11-concepts

- "A standard model for data interchange on the Web"
- Requires the use of URIs and information modelled as subject – predicate – object triples.

For ontologies:

RDFS (RDF-Schema) www.w3.org/2001/sw/wiki/RDFS

- "A general-purpose language for representing simple RDF vocabularies on the Web"
- Goes beyond RDF to designate classes and properties of RDF resources.

OWL (Web Ontology Language) www.w3.org/OWL

- "A Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things"
- Based on RDF and RDFS; OWL is W3C's attempt to extend RDFS.

For taxonomies/controlled vocabularies:

SKOS (Simple Knowledge Organization System) www.w3.org/TR/skos-reference (2009)

- "A common data model for sharing and linking knowledge organization systems via the Web"
- Encoded using XML and RDF for publication and use of vocabularies as linked data

Standards and Features

OWL-Defined Ontology Components

Entities – subjects or objects of properties (domains and ranges)

- Classes
  - Named sets of concepts that share characteristics and relations
  - May group subclasses or individuals (instances of the class)
- Individuals (instances)
  - Members or instances of a class.

Properties – predicates about individuals (instances), defined for classes

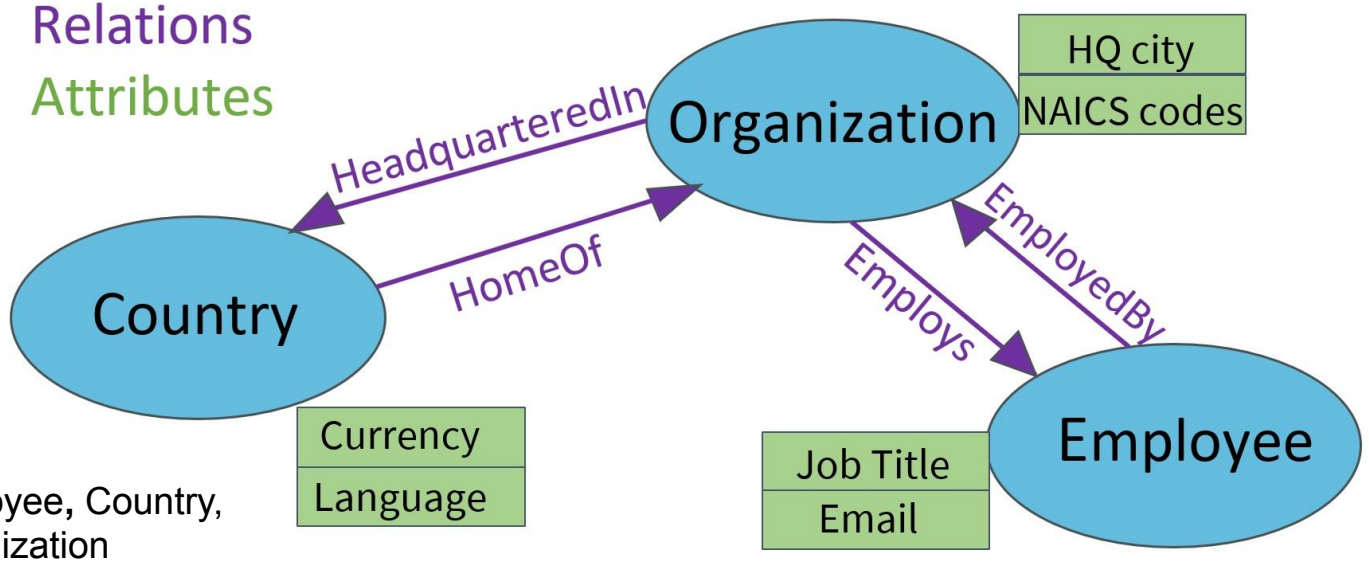
- Object properties
  - Relations between individuals
  - May be directed (single direction), symmetric, or with an inverse (different in each direction)
- Datatype properties
  - Attributes or characteristics of individuals
  - The object of a datatype property is a *value*.

Literals – values of attributes, with just a *lexical form* and a *datatype.*

https://www.w3.org/TR/2012/REC-owl2-primer-20121211/

Standards and Features

Classes
Relations
Attributes



Classes
Employee, Country, Organization

Relations: HeadquarteredIn < > HomeOf
EmployedBy < > Employs

Attributes: Email address, Job title, HQ city, NAICS codes, Currency, Language

Standards and Features

OWL modeling features also include:

- Class disjointness
- Complex classes
- Property hierarchies
- Property characteristics
- Property restrictions
- Property cardinality restrictions
- Domain and range restrictions
- Equality and inequality of individuals
- Enumeration of individuals

Standards and Features

"The SKOS data model views a knowledge organization system as a **concept scheme** comprising a set of concepts." (www.w3.org/TR/skos-reference)

SKOS elements

Concept Scheme & Collection	Concepts	Labels & Notation	Documentation	Semantic Relations	Mapping Relations
ConceptScheme	Concept	prefLabel	scopeNote	broader	exactMatch
inScheme	hasTopConcept	altLabel	definition	narrower	closeMatch
Collection	topConceptOf	hiddenLabel	example	related	broaderMatch
orderedCollection		notation	changeNote		narrowerMatch
member			editorialNote		relatedMatch
memberList			historyNote		

Standards and Features

Taxonomy features of ontologies

Ontologies	Taxonomies/Controlled vocabularies
Subclass hierarchies	Hierarchical generic-specific relationships
Class disjointness	Unique concepts
Instances	Named entities
Annotations	Documentation notes

Taxonomy features lacking in ontologies

Ontologies	Taxonomies/Controlled vocabularies
	Multiple labels for concepts: alternative, hidden, and other languages
	Hierarchical whole-part relationships
	Various types of notes and definitions
	Top concept branches; collections; mapping across concept schemes

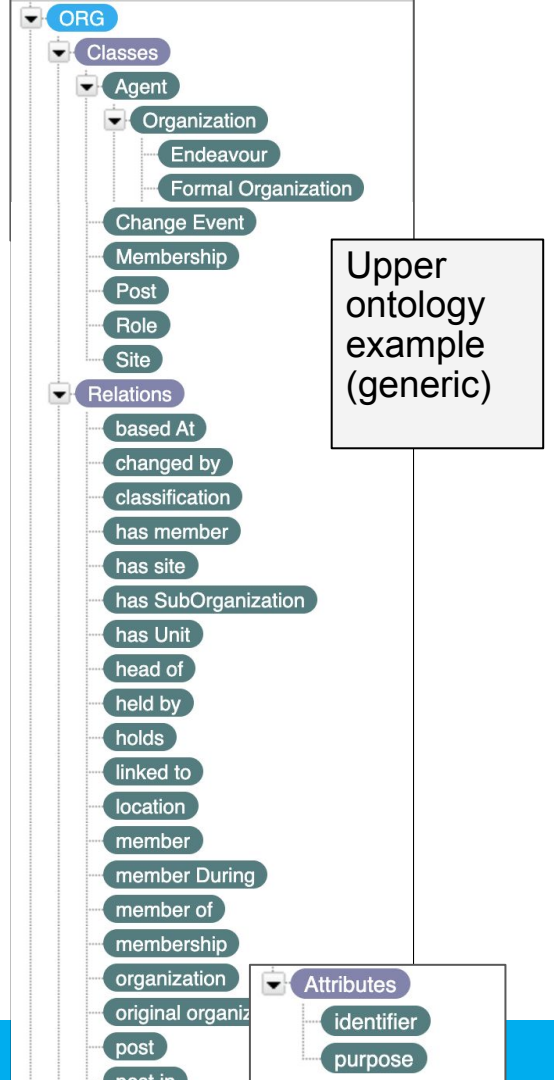
Ontology types and approaches

Ontology Types

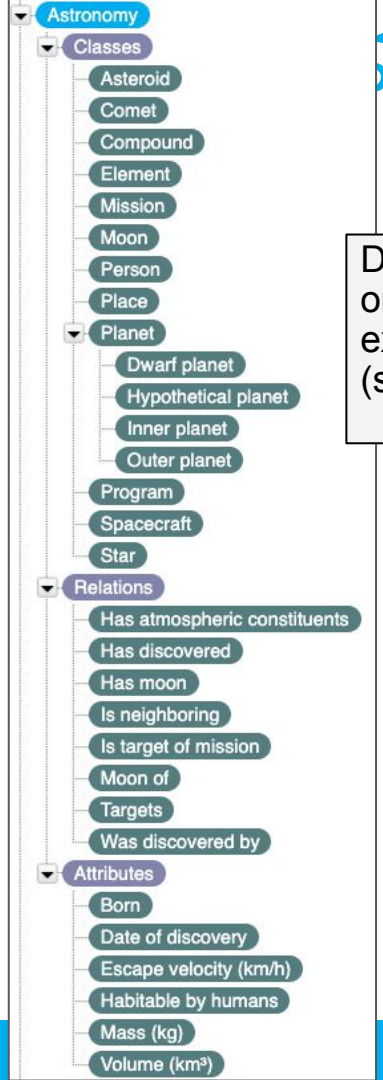
Ontology Types

- An upper ontology is not always smaller than a domain ontology, but it is more generic.
- An upper ontology is intended to be extended to define a domain ontology.
- A domain ontology may also be extended or applied to more specific instances.

Screenshots from Poolparty



Upper ontology example (generic)



Domain ontology example (specific)

Ontology Types and Approaches

Ontology Definition

“Ontology” may refer to

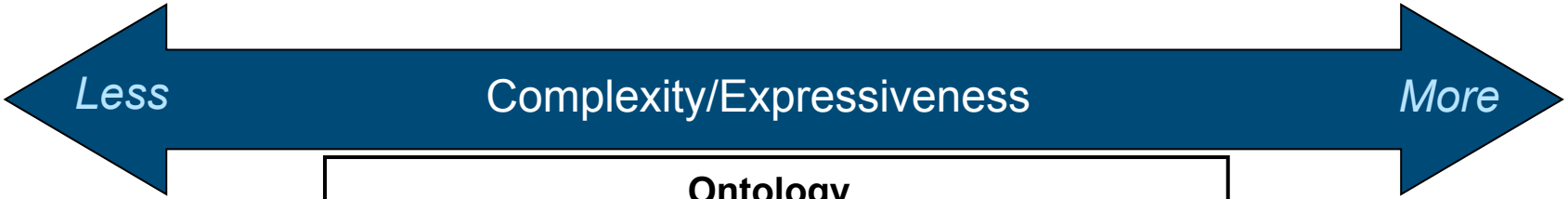
1. A generic model (upper or domain), *or*
2. A combination of the model with all the entities
    - all subclasses and individuals (as called in ontologies), or
    - all taxonomy concepts and named entities in controlled vocabularies

Regardless of the definition,
a functional ontology requires the combination of the model with the entities.

Ontologies and Controlled Vocabularies

Term List	Name Authority	Taxonomy	Thesaurus	Ontology
Ambiguity control	Ambiguity control Synonym control (Attributes)	Ambiguity control (Synonym control) Hierarchical relationships	Ambiguity control Synonym control Hierarchical relationship Associative relationships	Model + entities Semantic relationships Classes Attributes

Ontologies and Controlled Vocabularies

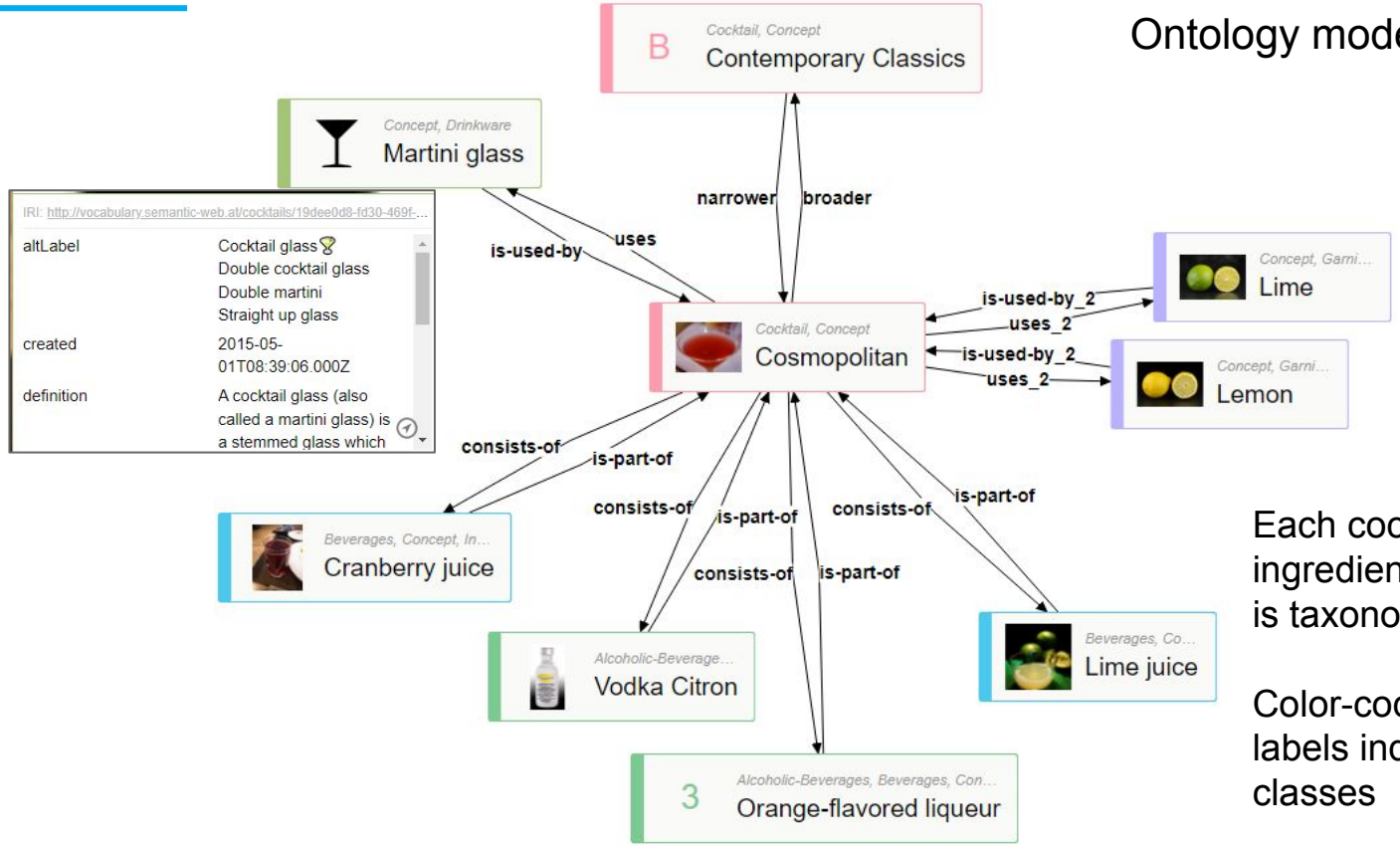


	Ontology		
Term List	**Name Authority**	**Taxonomy**	**Thesaurus**
Ambiguity control	Ambiguity control Synonym control (Attributes)	Ambiguity control (Synonym control) Hierarchical relationships	Ambiguity control Synonym control Hierarchical relationship Associative relationships

Generic model

The ontology is a semantic layer

Ontologies and Controlled Vocabularies



Ontology model + entities

Each cocktail, ingredient, glass type, is taxonomy concept.

Color-coding and italics labels indicate the classes

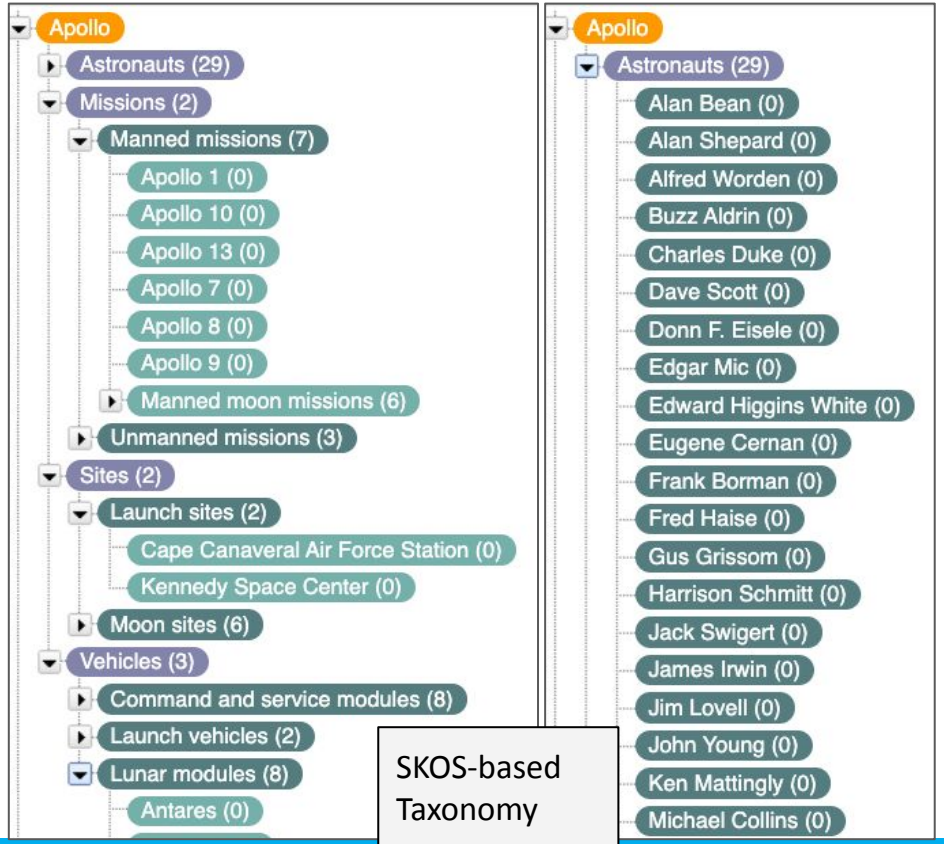
Ontology Approaches

Two approaches to developing domain ontologies:

1. Single knowledge organization system
   - The ontology comprises all entities: classes, subclasses, and instances.
   - The ontology is as detailed as any taxonomy, but has the addition of semantic relations and attributes.
   - Classes have multiple levels of subclasses.
   - Has individuals for unique named entity instances.
   - Modeled in dedicated ontology software, such as Protege.

2. Combination of a taxonomy with an ontology model
   - The ontology comprises classes and subclasses to the extent needed to describe the generic characteristics of the model.
   - The ontology does not include all possible levels of hierarchy, nor any instances.
   - The ontology is a model that is applied as a semantic layer to a taxonomy or multiple taxonomies and/or thesaurus and name authorities.
   - More the hierarchy resides in the SKOS taxonomy.
   - Modeled in a combination taxonomy/ontology software, such as PoolParty.

Ontology Approaches

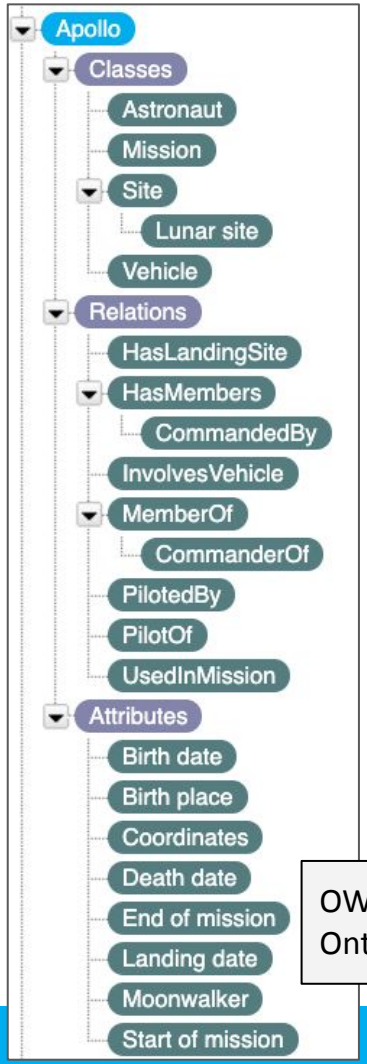


SKOS-based Taxonomy

2. Taxonomy + ontology layer

All concepts (whether class-like or individual instances) are maintained in the same hierarchy.

Screenshots from Poolparty



OWL-based Ontology

OWL classes, relations, and attributes are managed separately.

Semantically enriching a taxonomy to become an ontology

Enriching a Taxonomy

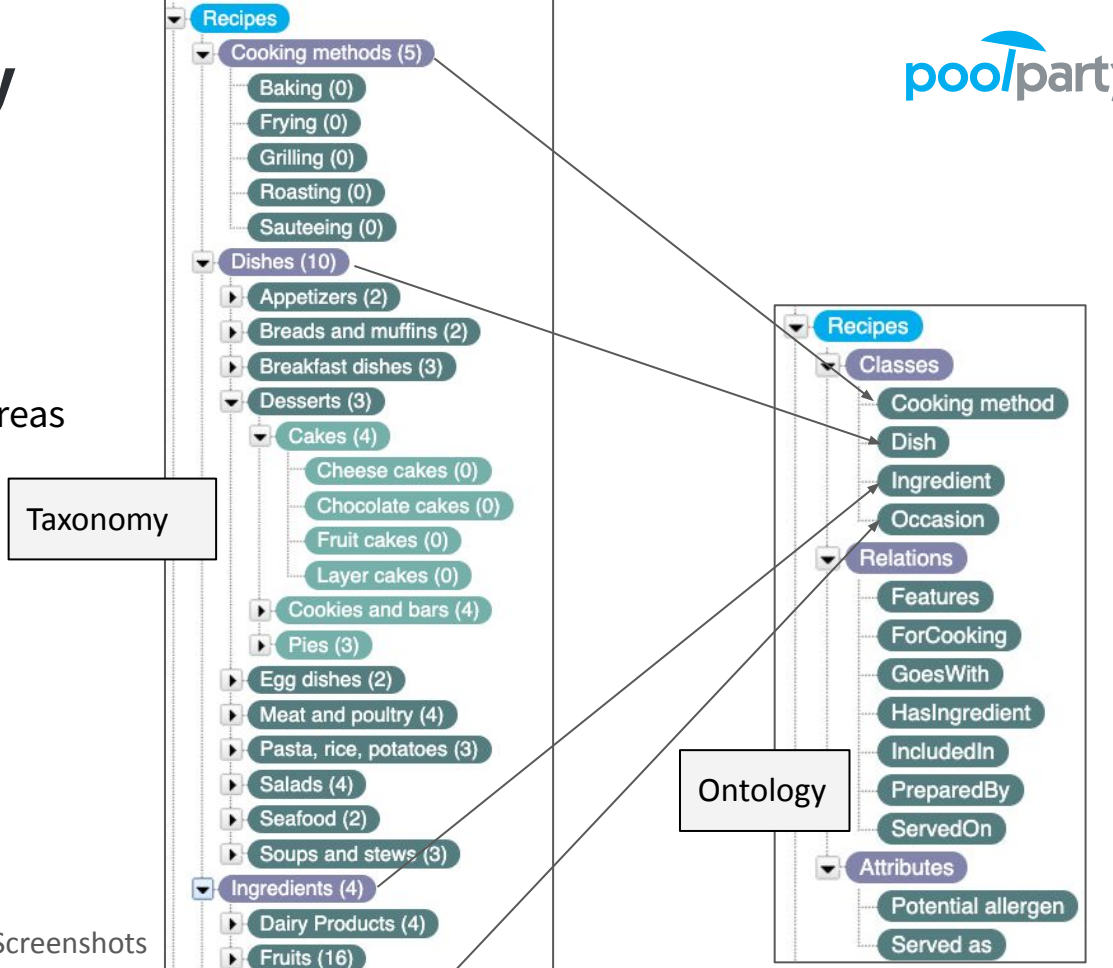
Taxonomy + ontology layer

The ontology tends to be smaller and simpler.

Taxonomy may be based on SKOS, whereas ontology is based on OWL.

Domain ontology example:

Classes - 4
Relations - 6
Attributes - 2

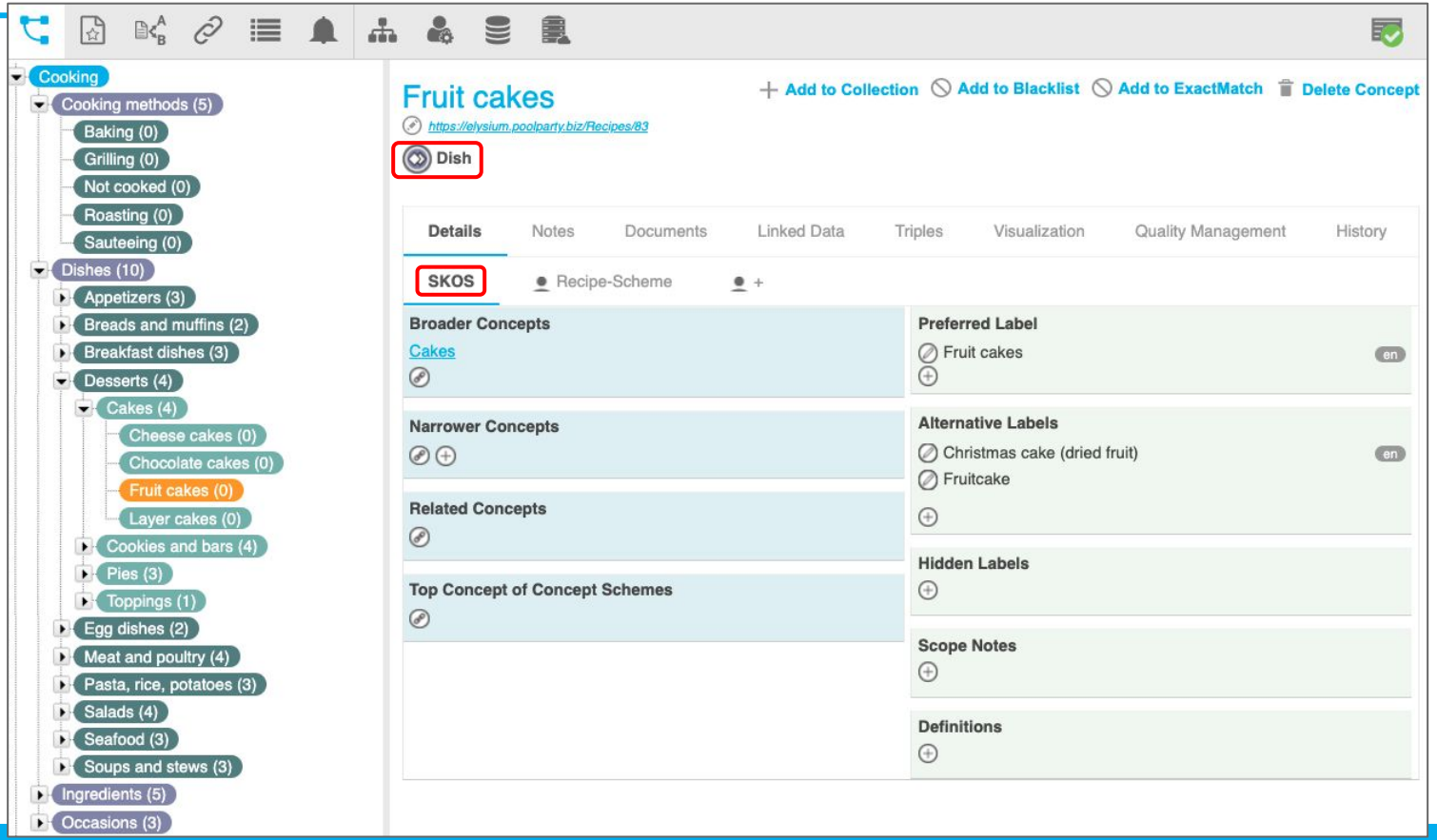


Screenshots from Poolparty

Applying Ontologies to Taxonomies

Concepts have both:

SKOS relationships and properties



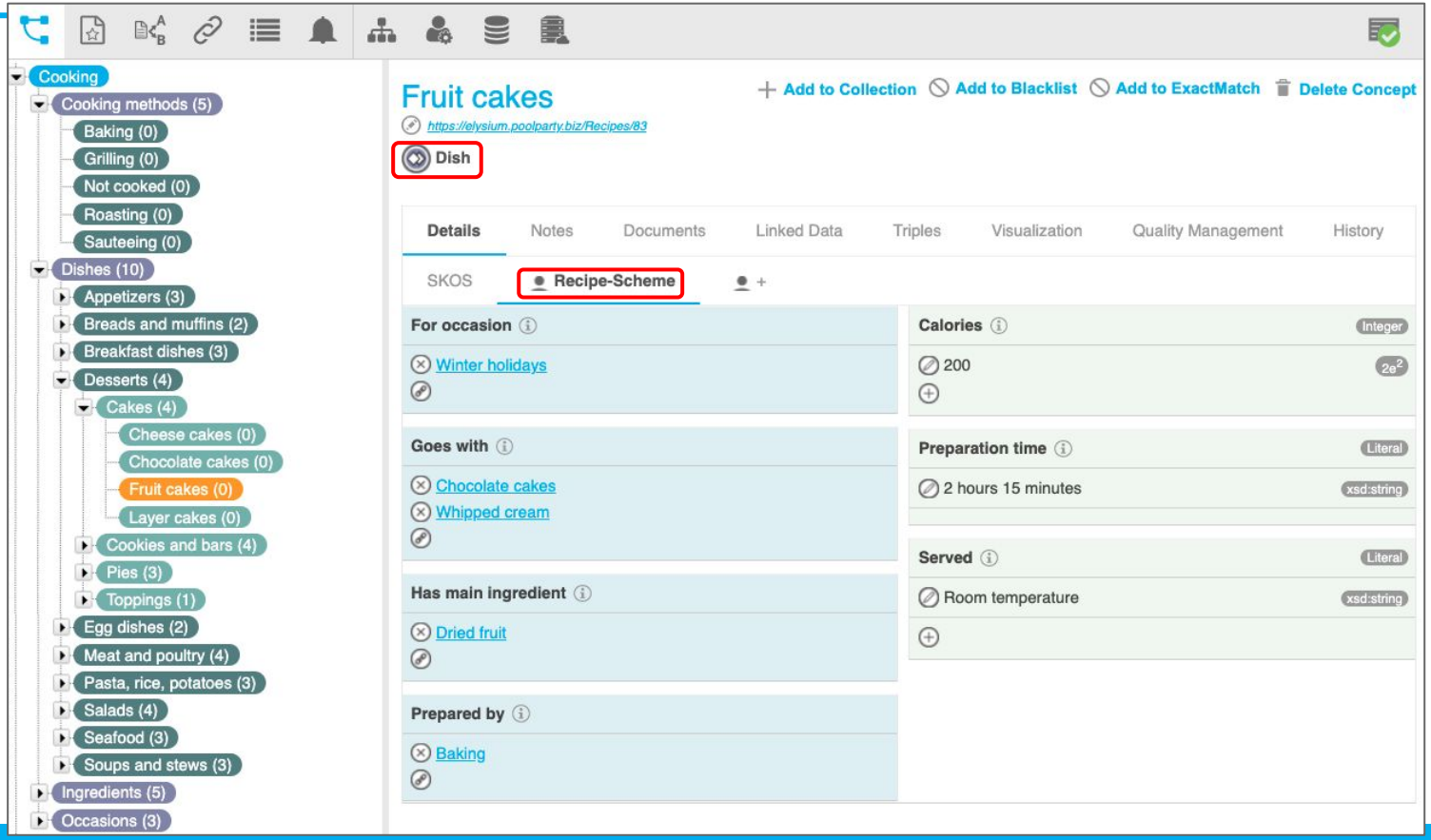
Screenshots from Poolparty

Applying Ontologies to Taxonomies

Concepts have both:

SKOS relationships and properties *and*

OWL-based semantic relationships and attributes from an ontology-based custom scheme.



Screenshots from Poolparty

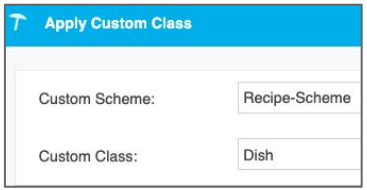
Extending a Taxonomy to become an Ontology

Extending SKOS concepts to be part of an ontology

- Ontology class labels correspond/match the SKOS concept scheme or concept labels to which they will be applied
  - The ontology “layer” is not an upper hierarchical layer, but an overlay to the higher levels of the SKOS project.
  - Tip: consider using singular for ontology class names and plural for SKOS concept names
- There is no dilemma in determining if an entity is an individual/instance or a class, since the ontology layer comprises only classes.

Extending a Taxonomy to become an Ontology

Extending SKOS taxonomy concepts to be part of an ontology



1. An ontology or “custom scheme” is applied/linked to the taxonomy project or to a specific concept scheme of a project.
2. Classes are applied to the levels of concept schemes, top concepts, or broad-level concepts, as appropriate. (Most often to concept schemes.)
3. Class properties (relations and attribute types) are then inherited by all narrower concepts.
4. Relevant relations and attributes are available for all concepts in the taxonomy, based on the class assigned to their broader concept or concept scheme. (SKOS InScheme designation can be used outside a hierarchy.)
5. Instantiating a relation between a pair of specific concepts or adding values to a concept’s attributes may done manually or through importing data.

Extending a Taxonomy to become an Ontology

Adding relations between specific concepts and values to attributes through importing a formatted spreadsheet

Analyze data existing spreadsheets:
Decide if a data in a column shall be an attribute or new class linked by a custom relation.

Concept Class (URI) Relation (URI) Relation (URI) Attribute (URI)

A	B	C	D	E
concept	**type**	https://k2.poolparty.biz/Recipes/mainIngredient	https://k2.poolparty.biz/Recipes/preparedBy	https://k2.poolparty.biz/Recipes/calories
Apple pie	https://k2.poolparty.biz/Recipes/Dish	Apples	Baking	290
Cheese omlette	https://k2.poolparty.biz/Recipes/Dish	Eggs	Frying	280
Guacamole	https://k2.poolparty.biz/Recipes/Dish	Avacados	Not cooked	105

A	B	C	D	E	F
concept	**type**	https://k2.poolparty.biz/Jobs/belongsTo	https://k2.poolparty.biz/Jobs/has-HQ-in	https://k2.poolparty.biz/Jobs/Address	https://k2.poolparty.biz/Jobs/Phone-numbe
Volkswagen	https://k2.poolparty.biz/Jobs/Employer	Automotive	Germany	Dieselstraße 28, 38446 Wolfsburg, Germany	49-5361-90
Accenture	https://k2.poolparty.biz/Jobs/Employer	Information technology and services	Ireland	Unit 1, Grand Canal Square, Grand Canal Quay,	353 1 407 6000
BASF	https://k2.poolparty.biz/Jobs/Employer	Chemicals	Germany	Carl-Bosch-Straße 38, 67056 Ludwigshafen/Rh	49 621 600
Tata Consultancy Services	https://k2.poolparty.biz/Jobs/Employer	Information technology and services	India	TCS House, Raveline Street, Mumbai 400001, M	91-22-6778 9999
Nestlé	https://k2.poolparty.biz/Jobs/Employer	Food and beverages	Switzerland	Av. Nestlé 55, 1800 Vevey, Switzerland	41 21 924 1111
Royal Dutch Shell	https://k2.poolparty.biz/Jobs/Employer	Oil and gas	Netherlands	Carel van Bylandtlaan 16, 2596 HR The Hague,	31 70 3779111
Toyota Motor	https://k2.poolparty.biz/Jobs/Employer	Automotive	Japan	1 Toyota-Cho, Toyota City, Aichi Prefecture, Jap	81 565 282121
Apple Inc	https://k2.poolparty.biz/Jobs/Employer	Computer hardware	United States	1 Infinite Loop Cupertino, CA 95014	1 408 996-1010
Samsung Electronics	https://k2.poolparty.biz/Jobs/Employer	Consumer electronics	South Korea	416, Maetan 3-Dong, Yeongtong-Gu, Suwon, G	82 31 2001114

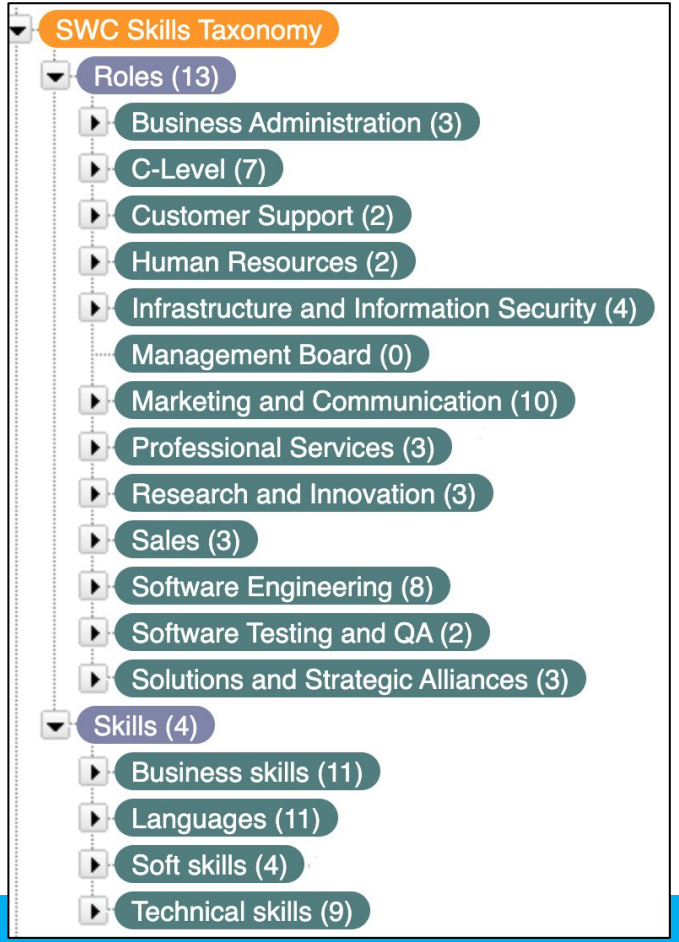
Extending a Taxonomy to become an Ontology

Ontology Design from a Taxonomy

SKOS-based taxonomies are suited for ontologies.

- Common RDF basis
- Combined taxonomy/ontology software is available.
- Concept scheme structure encourages class-based modeling.

  There could be a little as two classes, e.g. **Roles** and **Skills**.



Extending a Taxonomy to become an Ontology

Benefits of combining a high-level ontology as a semantic layer with a taxonomy

- Makes use of existing taxonomies, even multiple taxonomies
- Easier to model the ontology
  - Existing taxonomies and controlled vocabularies provide a basis for knowledge modeling.
  - No need to distinguish between sub-classes and individuals.
- Supports expert specialization
  - Taxonomists develop and maintain taxonomies.
  - Ontologists develop and maintain the ontology.
- More flexible and adaptable
  - The taxonomy changes more frequently than does the ontology.
  - Taxonomies can easily be added.
- Different purposes served
  - The ontology is for modeling, reasoning, and analysis.
  - The taxonomy is for tagging and information/data retrieval.

Questions/Contact

Heather Hedden
Data and Knowledge Engineer
Semantic Web Company Inc.
One Boston Place, Suite 2600
Boston, MA 02108
857-400-0183
heather.hedden@semantic-web.com
www.linkedin.com/in/hedden
accidental-taxonomist.blogspot.com

Semantic Web Company www.semantic-web.com

PoolParty software www.poolparty.biz