From Taxonomy to Ontology

by

Heather Hedden

Documentation Research and Training Centre
Indian Statistical Institute
Bangalore Centre

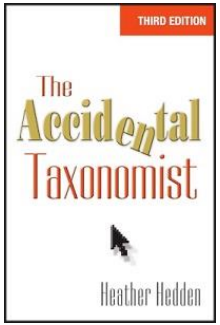
14 January 2025

About Heather Hedden

- Independent taxonomy consultant, Hedden Information Management
- Instructor of online and corporate taxonomy courses and workshops
- Previously a taxonomy consultant in consulting firms, Enterprise Knowledge and PPC. Also, a contract consultant for others.
- Former taxonomy-related roles at Semantic Web Company, Gale/Cengage, Viziant, and First Wind.
- Author of *The Accidental Taxonomist, 3rd ed.* (2022, Information Today, Inc.)

Hedden Information Management

Making information findable



Outline

- Introduction and why combine taxonomies and ontologies
- Definitions and features of taxonomies and ontologies
- Semantic Web standards for taxonomies and ontologies
- Approaches to combining a taxonomy and ontology
- Combining taxonomies and ontologies: examples in tools

Introduction and Why Combine Taxonomies and Ontologies?

Taxonomies and Ontologies

Taxonomies and ontologies both

- Aim to organize and access knowledge
- Have names derived from other disciplines (biology and philosophy)
- Have varied definitions

Traditional differences

- Have had different uses
- Have different origins
- Have been parts of different disciplines with different professionals (library/information science and computer/data science)

Distinctions between taxonomies and ontologies have more recently been blurred, as they have increasingly been combined, and are supported in the same software.

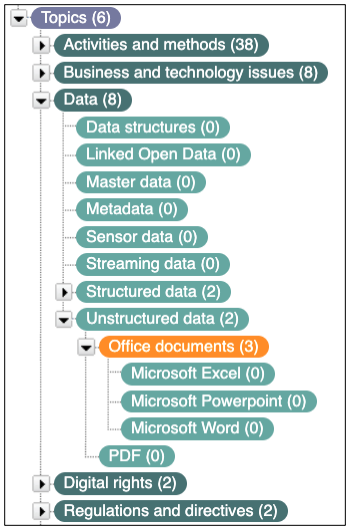
Background to Taxonomies

Origins of modern information taxonomies

1. Classification systems (since Linnaean classification 1758)
2. Subject heading schemes (since 1898) and thesauri (1960s)

classification system

taxonomy



Safety goggles *(May Subd Geog)*
UF Goggles, Safety
Safety glasses
BT Eye—Protection
Eyeglasses, Protective
Personal protective equipment
Safety harness (Aeronautics) *(May Subd Geog)*
UF Safety belts (Aeronautics)
BT Aeronautics—Safety appliances
Safety hats *(May Subd Geog)*
[T55.3.S23 (Industrial safety)]
UF Safety helmets
BT Hats
Helmets
Personal protective equipment
NT Bicycle helmets
Fire helmets
Flying helmets
Motorcycle helmets
Safety helmets
USE Safety hats
Safety in advertising *(May Subd Geog)*
BT Advertising

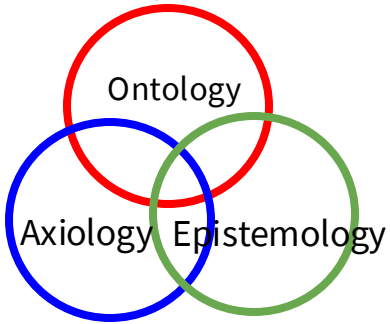
subject heading scheme

Modern information taxonomies arose in 1990s with the user interface capabilities of the Web and the development of information architecture.

Background to Ontologies

Traditional meaning from philosophy:

A branch of metaphysics concerned with the nature and relations of being



Modern computer/data science definition:

"An ontology defines a set of representational primitives with which to model a domain of knowledge or discourse.

The representational primitives are typically classes (or sets), attributes (or properties), and relationships (or relations among class members).

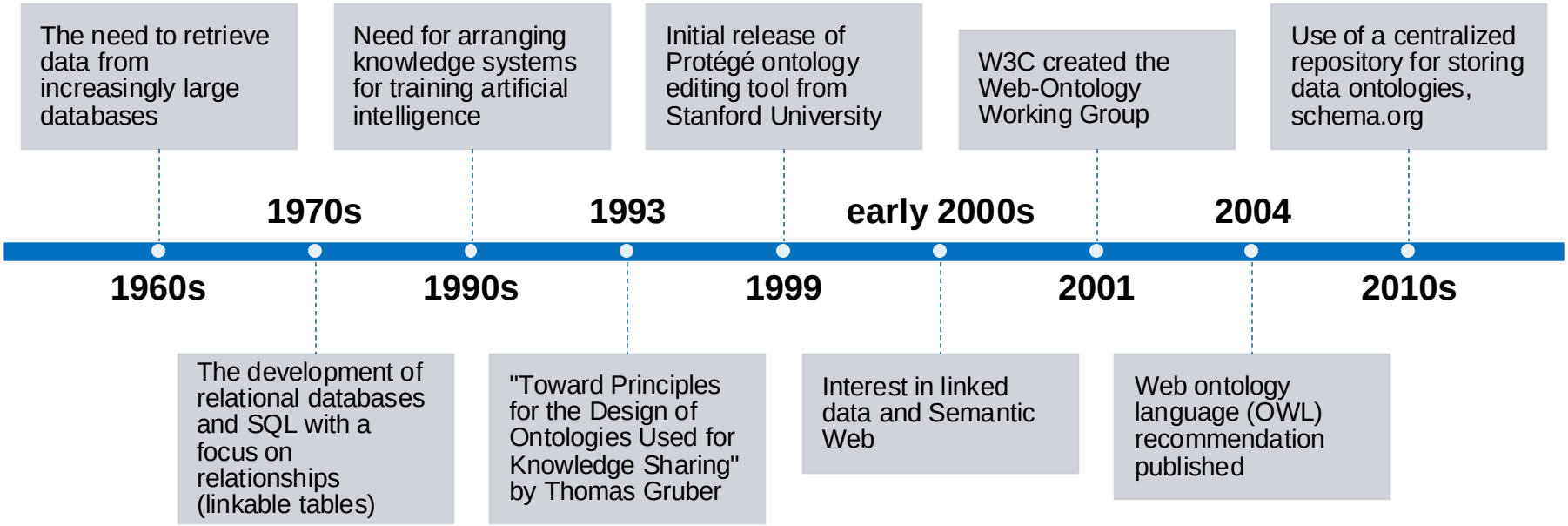
The definitions of the representational primitives include information about their **meaning and constraints on their logically consistent application**.

In the context of database systems, ontology can be viewed as a level of abstraction of data models, analogous to hierarchical and relational models, but intended for modeling knowledge about individuals, their attributes, and their relationships to other individuals."

https://tomgruber.org/writing/definition-of-ontology.pdf

Gruber, Tom (2016) Ontology. In: Liu L., Özsu M. (eds) Encyclopedia of Database Systems. Springer, New York, NY, based on his original definition in 1993.

Background to Ontologies: History of Ontologies



Influencing ontology origins:

- Need to manage and understand data
- Support for computers and artificial intelligence
- Application to the Semantic Web

Why Combine Taxonomies and Ontologies

What you can do with a taxonomy

- Search: find content about…. (search string matches taxonomy concepts)
- Topic browse: explore subjects arranged in a hierarchy linked to content on the subject
- Faceted (filtering/refining) search: find content meeting a combination of terms/criteria
- Tagging: index content consistently so that retrieval is comprehensive and accurate
- Discovery: find other content tagged with same concepts as tagged to found content; explore broader, narrower, and related taxonomy topics
- Content curation: set up feeds or alerts based on user profile (personalization) and taxonomy concepts tagged
- Metadata management: for content identification, comparison, analysis, etc.

Why Combine Taxonomies and Ontologies

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

- Complex multi-part searches: “chained” queries beyond just “about a subject.” (e.g. contacts in a location, employed by companies, which belong to certain industries)
- Search on data: obtain specific data, and not just content/full documents
- Search on more specific criteria: searches that vary based on category (class)
- Knowledge graph or semantic layer creation: to connect data and content from different repositories and to search, analyze, and infer across all sources

What else you can do with an ontology

- Modeling complex interrelationships: to understand operations and systems better (e.g. in product approval or supply chain processes)
- Exploring explicit relationships: between concepts (not just broader, narrower, related)
- Visualization: graph representations of concepts and semantic relationships
- Reasoning and inferencing: based on logic, to gain new knowledge

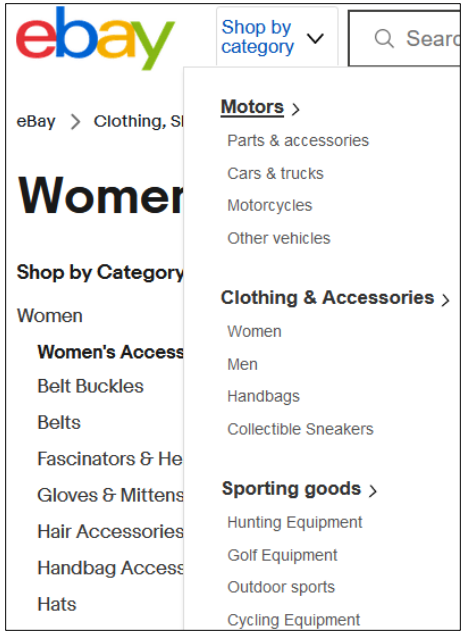
Why Combine Taxonomies and Ontologies

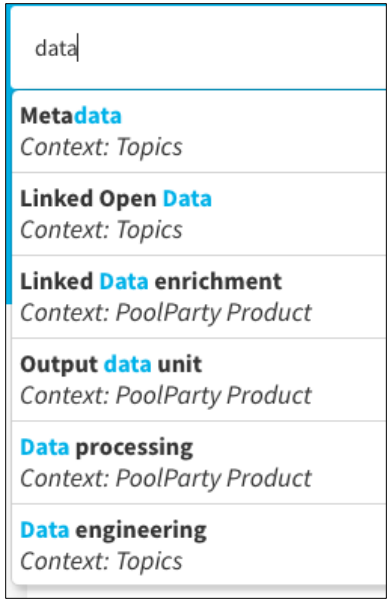
What you can do with a taxonomy and not with an ontology

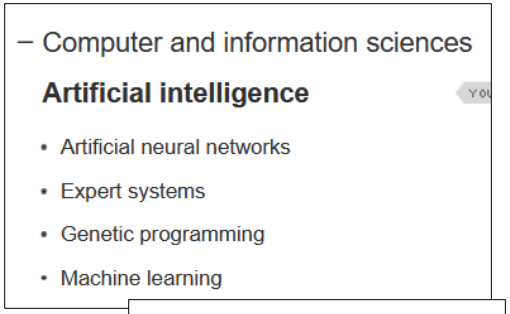
- Search with synonyms: utilize alternative labels to match to search strings
- Tag with synonyms: utilize alternative labels to match to text strings
- Support multiple languages: support access to users of different languages and tag content in different language
- Include definitions and scope notes: as standardized types of fields
- Map/link to other taxonomies: to enable access to other content tagged with other taxonomies
- Support metadata as filters: aligning concept schemes to metadata properties and facets
- Browse hierarchies in a frontend UI: browsable display of concepts linked directly to content, unlike ontology visualizations
- Implement easily in content systems: CMSs, DAM systems, Intranets, websites, etc.

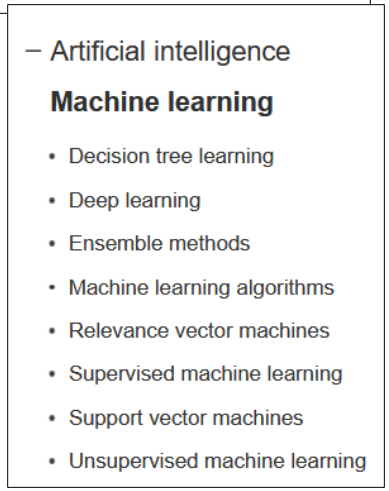
Definitions and Features of Taxonomies and Ontologies

Introduction to Taxonomies









Taxonomies help people find information

Introduction to Taxonomies

Taxonomy type examples

Hierarchical Taxonomy Example

Leisure and culture
. Arts and entertainment venues
. . Museums and galleries
. Children's activities
. Culture and creativity
. . Architecture
. . Crafts
. . Heritage
. . Literature
. . Music
. . Performing arts
. . Visual arts
. Entertainment and events
. Gambling and lotteries
. Hobbies and interests
. Parks and gardens
. Sports and recreation
. . Team sports
. . . Cricket
. . . Football
. . . Rugby
. . Water sports
. . Winter sports
. Sports and recreation facilities
. Tourism
. . Passports and visas
. Young people's activities

Faceted Taxonomy Example

Career Level
- Student
- Entry Level
- Experienced
- Manager
- Director
- Executive

Function
- Customer Service & Support
- Delivery
- Engineering
- Finance
- General Management
- Legal & Regulatory Affairs
- Marketing & Advertising

[more]

Industry
- Agriculture
- Apparel & Fashion
- Automotive
- Aviation & Aerospace
- Banking
- Biotechnology
- Broadcast Media
- Chemicals

[more]

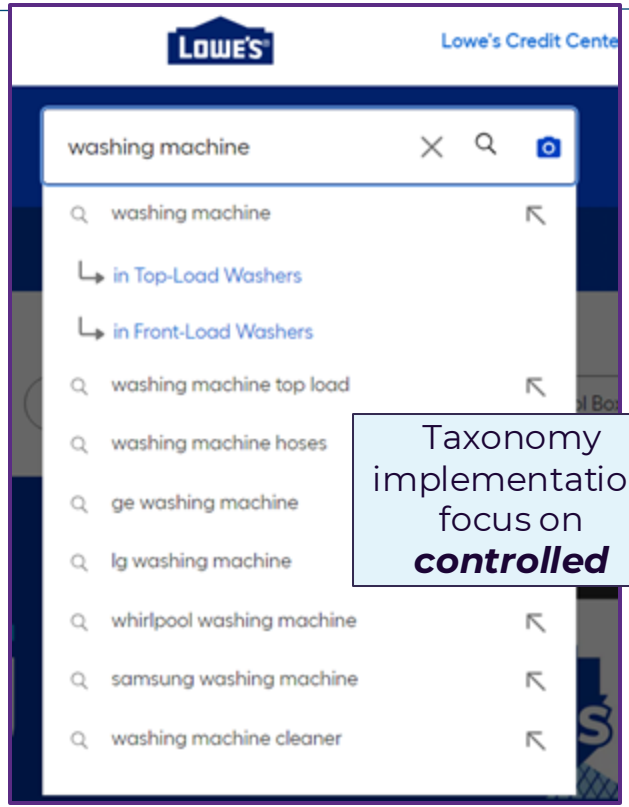
Taxonomy Features

A taxonomy is:

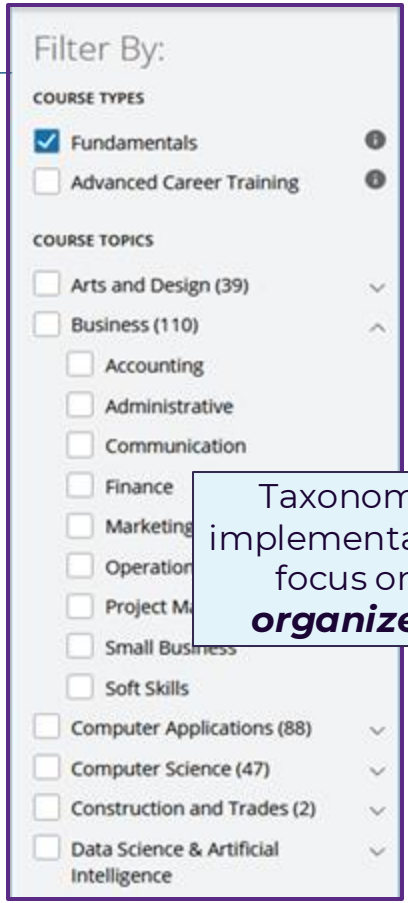
Controlled and ***organized***

1. Controlled:
A kind of controlled vocabulary or knowledge organization system,
based on unambiguous concepts, not just words:
things, not strings

2. Organized:
Concepts are arranged in a structure of hierarchies, categories, or facets to organize them.

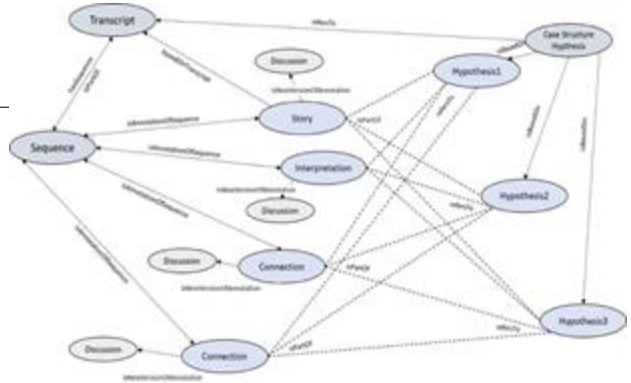


Taxonomy implementation focus on ***controlled***



Taxonomy implementation focus on ***organized***

Introduction to Ontologies

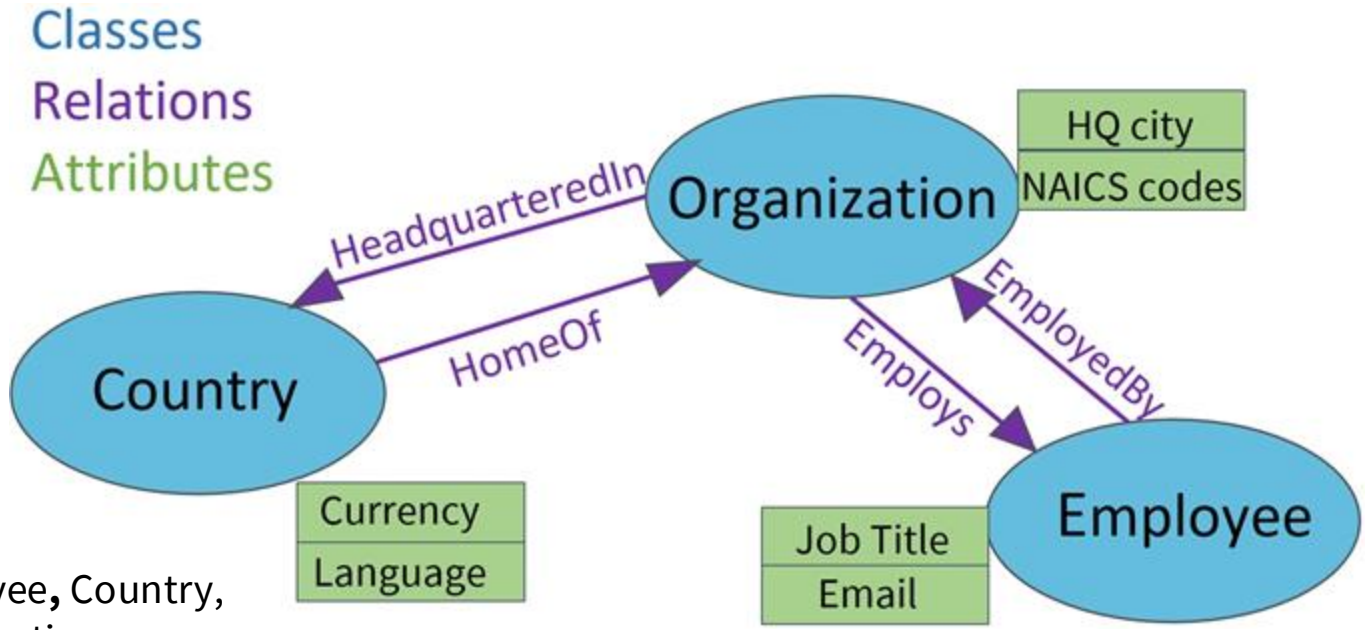


Definitions of an ontology

- A form of knowledge representation
  - Not just knowledge organization
- A formal naming and definition of the **types**, **properties** and **interrelationships** of entities in a particular domain.
  - Classes, custom attributes, and semantic relationships
- A set of precise descriptive statements about a particular domain.
  - Statements as *subject-predicate-object* are expressed as triples.

- A more abstract layer in describing a knowledge organization system
  - Overlays and connects to a taxonomy or other controlled vocabulary to add semantics

Ontology Features



Classes

Employee, Country, Organization

Relations: HeadquarteredIn < > HomeOf
EmployedBy < > Employs

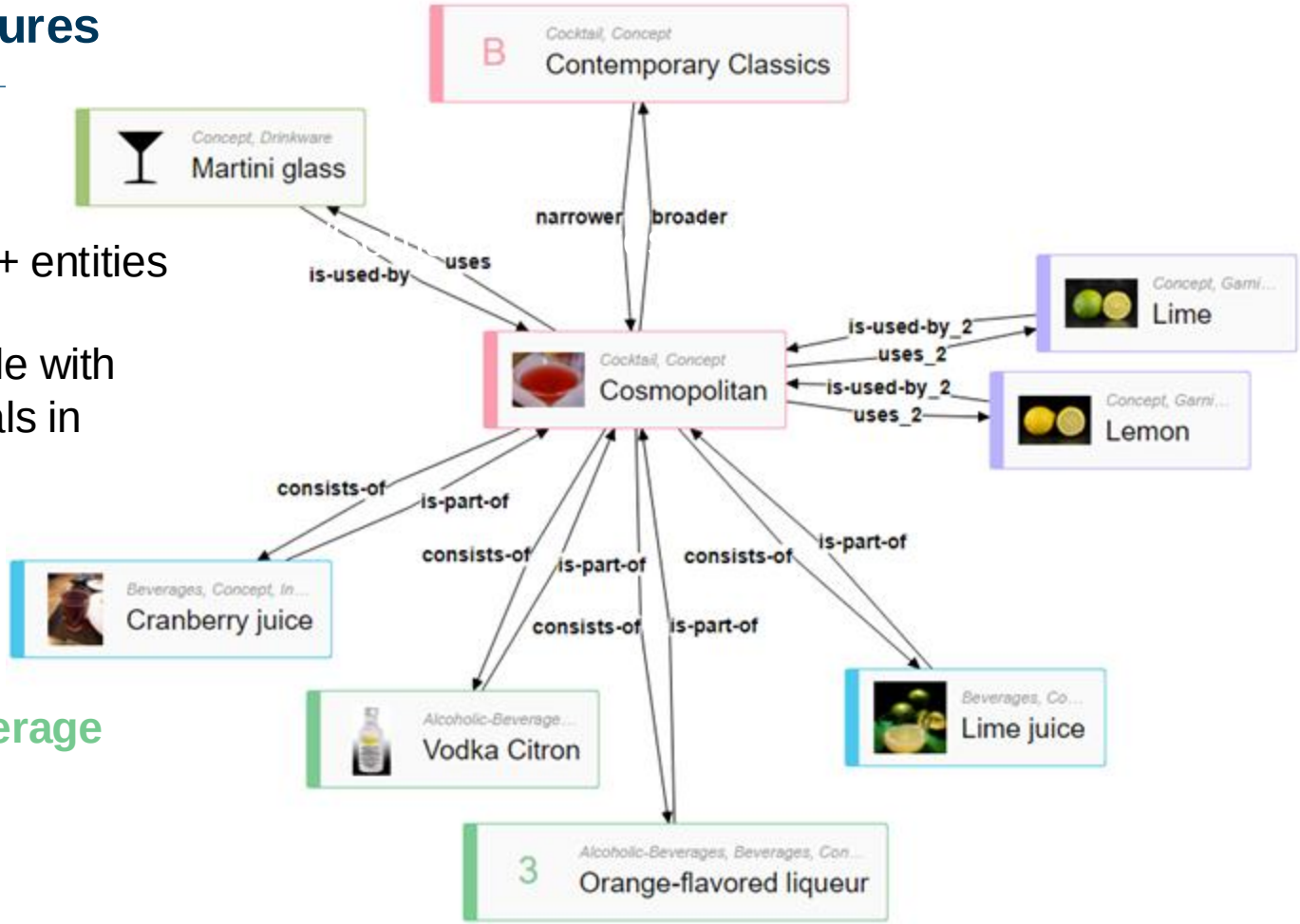
Attributes: Email address, Job title, HQ city, NAICS codes, Currency, Language

Ontology Features

Ontology model + entities

Ontology example with specific individuals in each class:

- Cocktail
- Garnish
- Drinkware
- Beverage
- Alcoholic Beverage



Definitions of Ontologies

Ontology definition questions

Is it

- A knowledge *model*, an *abstract layer*, a form of knowledge representation, that describes the classes, relationship types, and attribute types in a domain,

or

- A knowledge organization system, that includes both:
  the classes, relationship types, and attribute types *and*
  the specific concepts, entities/individuals, and their specific attributes

A model + individuals is often called an “ontology,” becaues ontology management software supports the inclusion of individuals, and ontologies are increasingly domain-specific in enterprises. This might be better called a “knowledge graph.”

Knowledge Organization System Types

Less — Complexity/Semantic expressiveness — *More*

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	Ambiguity control Model + entities Semantic relationships Classes Attributes

Knowledge Organization System Types

Less Complexity/Semantic expressiveness *More*

Ontology (model/layer)			
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

Ontology as a semantic model

Semantic Web Standards for Taxonomies and Ontologies

Semantic Web Standards for Taxonomies and Ontologies

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 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

For ontologies:

RDFS (RDF-Schema) www.w3.org/TR/rdf11-concepts

- "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/TR/owl2-overview

- "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 an extension of RDFS.

Semantic Web Standards for Taxonomies: SKOS

SKOS (Simple Knowledge Organization System) 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	broadMatch
orderedCollection		notation	changeNote	broaderTransitive	narrowMatch
member			editorialNote	narrowerTransitive	relatedMatch
memberList			historyNote		

Human-readable form example: **Preferred label**
Machine-readable form example: **skos:prefLabel**

Semantic Web Standards for Ontologies: OWL

OWL-Defined Ontology Components

Entities – subjects (domains) or objects (ranges) of properties, within RDF triples

- **Classes** (in SKOS: concepts)
  - Named sets of concepts that share characteristics and relations
  - May contain subclasses or individuals (instances of the class)
- **Individuals** (in SKOS: concepts)
  - Members or instances of a class. Unique named entities.

Properties – predicates about individuals (instances)

- **Object properties** (in SKOS: relations)
  - Relations between individuals
  - May be directed (single direction), symmetric, or with an inverse (different in each direction)
- **Datatype properties** (not in SKOS, except for notes)
  - 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/

Approaches to Combining a Taxonomy and Ontology

Taxonomy + Ontology Creation Approaches

Options for creating the semantic model + individual instances type of an ontology:

1. Use a dedicated ontology tool, e.g. Protégé, or hand-coding OWL to
   - create the ontology and build out the full detailed taxonomic hierarchy (of classes, subclasses) as an extension of the ontology
   - create all individuals associated with the ontology as members of classes

2. Use taxonomy/ontology combined tool to
   - create a taxonomy
   - create an ontology to go with the taxonomy
   - link them, with the ontology providing semantic enrichment

Ontology + Taxonomy Creation Approaches

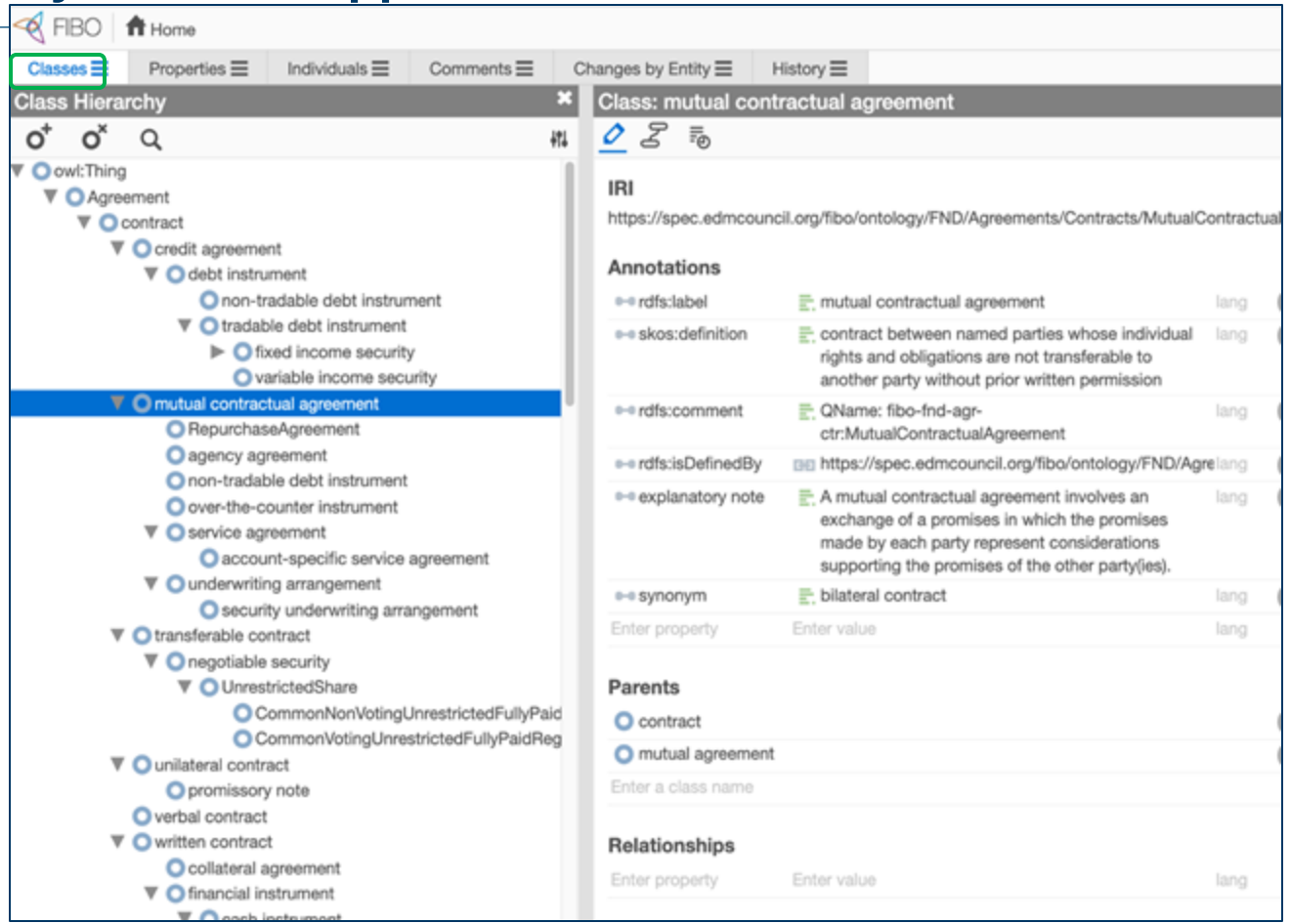
Method 1:

Using a dedicated ontology tool: Protégé

Example of FIBO Financial Industry Business Ontology

Classes and subclesses in a taxonomy hierarchy

protégé



Ontology + Taxonomy Creation Approaches

Method 1 Issues

- Follows ontology standards, but not thesaurus/taxonomy best practices
- Top-down taxonomy design focus; less bottom-up
- Hierarchies focus on subject domain classification; less on supporting users and UX
- Designed by engineers and ontologists, not taxonomists or information architects

Not recommended approach to extending a taxonomy into an ontology:

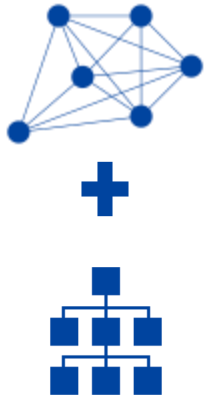
Importing taxonomies into an OWL-based ontology or dedicated ontology tool.

- All taxonomy hierarchies get converted to class-subclass hierarchies.
  - Importing taxonomies into ontologies will incorrectly treat:
    - whole-part taxonomy relations (e.g. geographic) as class-subclass relations
    - generic-instance (named entity) taxonomy relations as class-subclass relations, not class-instance affiliations
- Alternative labels can import, but as “Annotation” properties, and may not function as alternative labels for tagging and search.

Applying Ontologies to Taxonomies

Method 2: Adding an ontology as a semantic model to a SKOS taxonomy

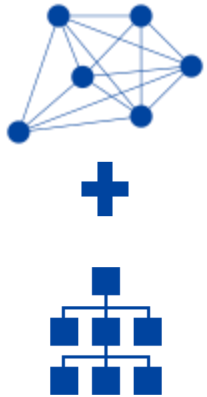
- Starts with existing taxonomies.
- Taxonomies based on SKOS can integrate with RDFS and OWL ontology resources, as all are based on RDF.
- Ontology class labels correspond/match the SKOS concept schemes or concept labels to which they will be applied.
  - The ontology “layer” is not an upper hierarchical layer, but an **overlay** to the higher levels of the SKOS taxonomy.
- Combined taxonomy/ontology management software supports this approach.



Applying Ontologies to Taxonomies

Benefits of the method of adding an ontology as a semantic model to a taxonomy

- Brings together *multiple* existing controlled vocabularies, term sets, named entity lists
- Easier to model the ontology
  - Existing taxonomies provide a basis for knowledge modeling.
  - No need to distinguish between sub-classes and individuals.
- Supports expert specialization
  - Taxonomists develop and maintain taxonomies.
  - Ontologists (if needed) 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, analysis, and data queries
  - The taxonomy is for tagging and information retrieval.
- More practical

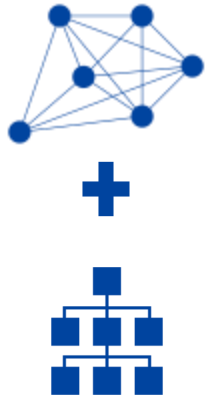


"The main problem SKOS solves is that many of these OWL projects become too involved and the organization loses interest before it ever sees any value. The less formal, concept-oriented SKOS is more practical for 95% of organizations." *Posted on Reddit September 27, 2024 [Best Ontology Development Environment Tool?]*

Applying Ontologies to Taxonomies

Tips on how to design an ontology based on existing taxonomies

- High level classes in the ontology are based on individual taxonomies, controlled vocabularies, terms lists.
- SKOS concept schemes (and to a lesser extent top concepts) are good candidates for classes.
- The OWL-based ontology “semantic model” should include classes and subclasses only for which there is a use case for distinct properties (attributes and semantic relations).
    - Deeper levels of a “subclass” hierarchy reside in the SKOS taxonomy concepts and their broader/narrower relations.

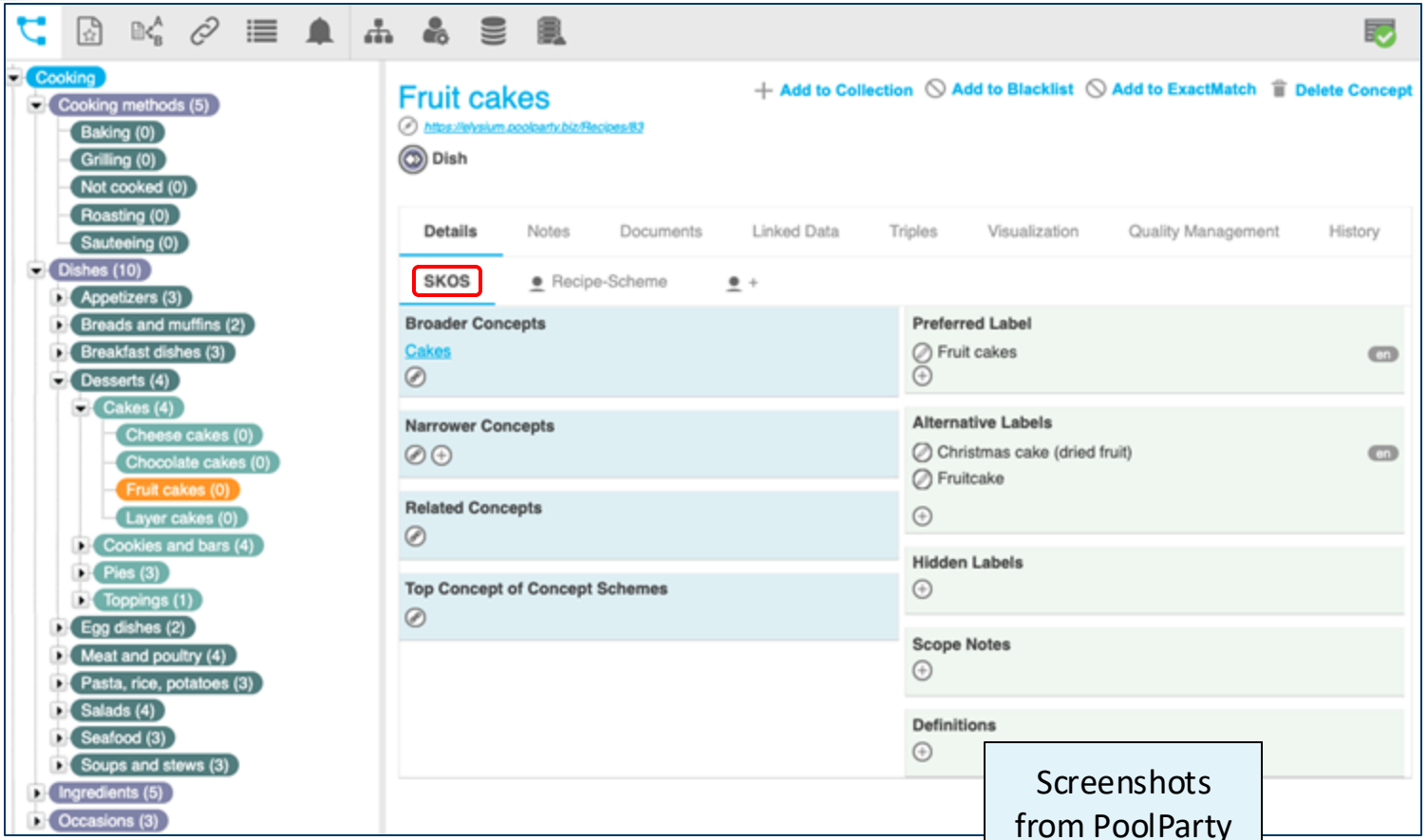


Combining Taxonomies and Ontologies: Examples in Tools

Combining Taxonomies and Ontologies in Software Tools

Concepts have both:

SKOS relationships and properties



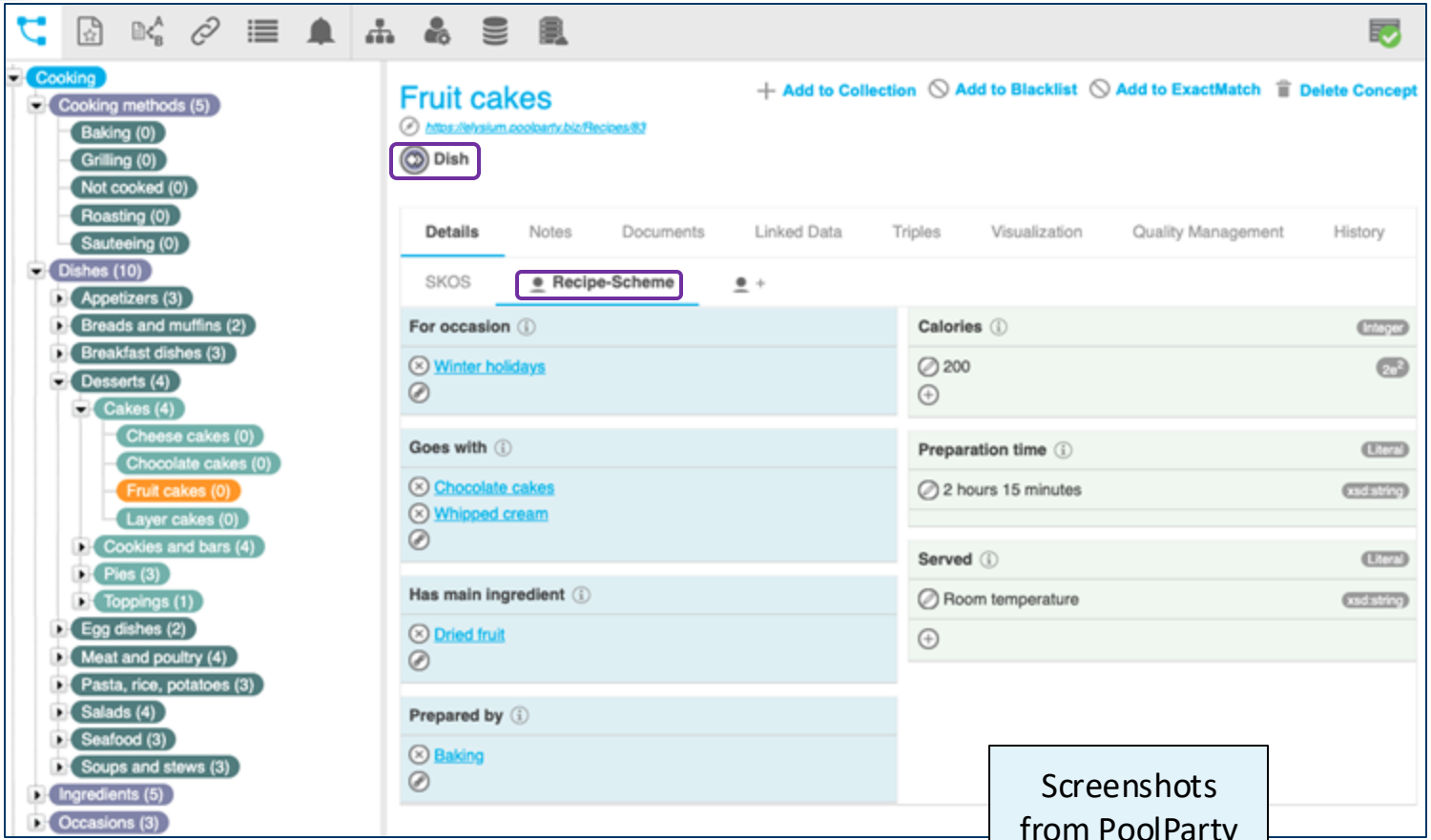
Screenshots from PoolParty

Combining Taxonomies and Ontologies in Software Tools

Concepts have both:

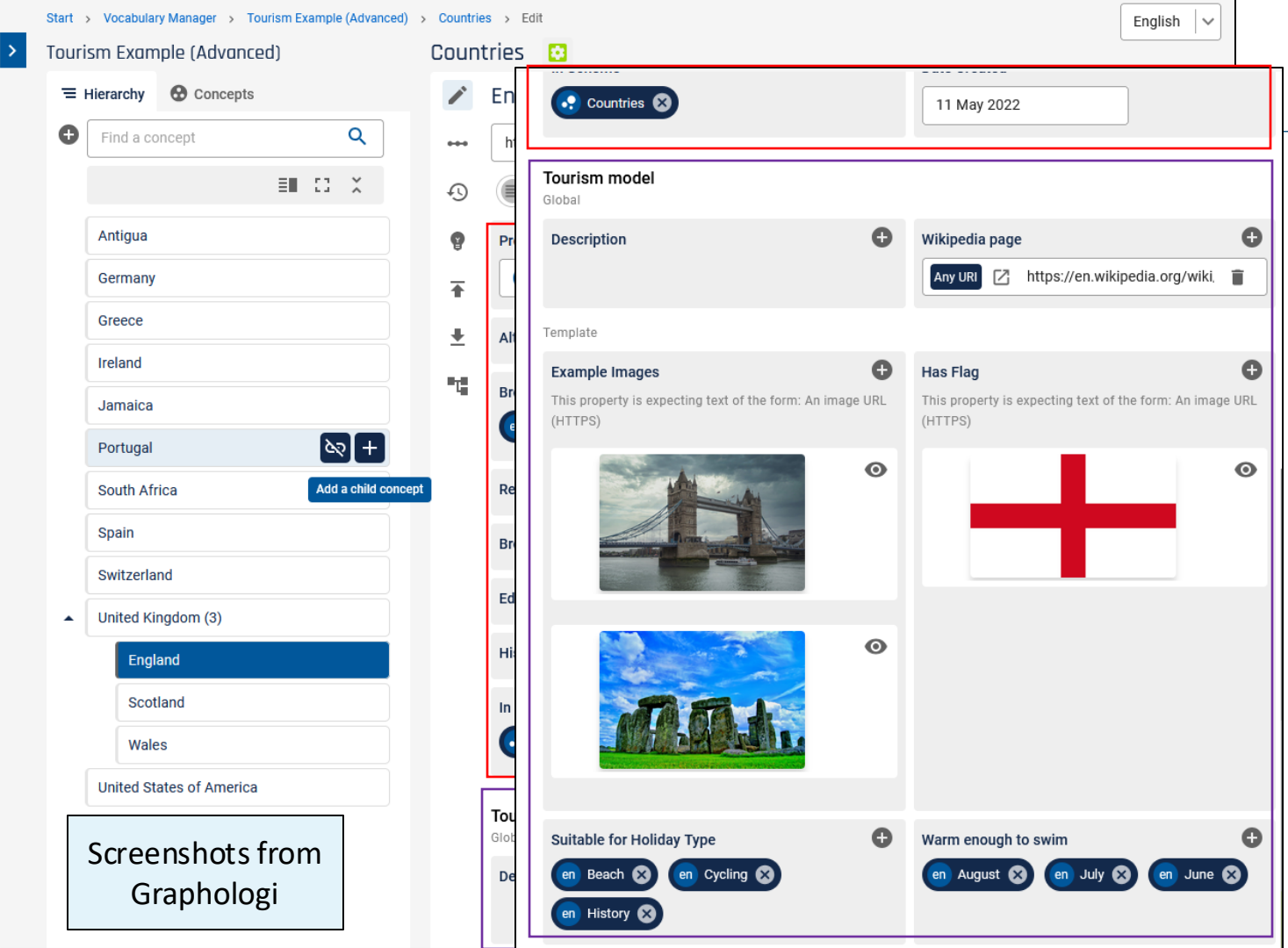
SKOS relationships and properties *and*

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



Screenshots from Poolparty

Screenshots from PoolParty



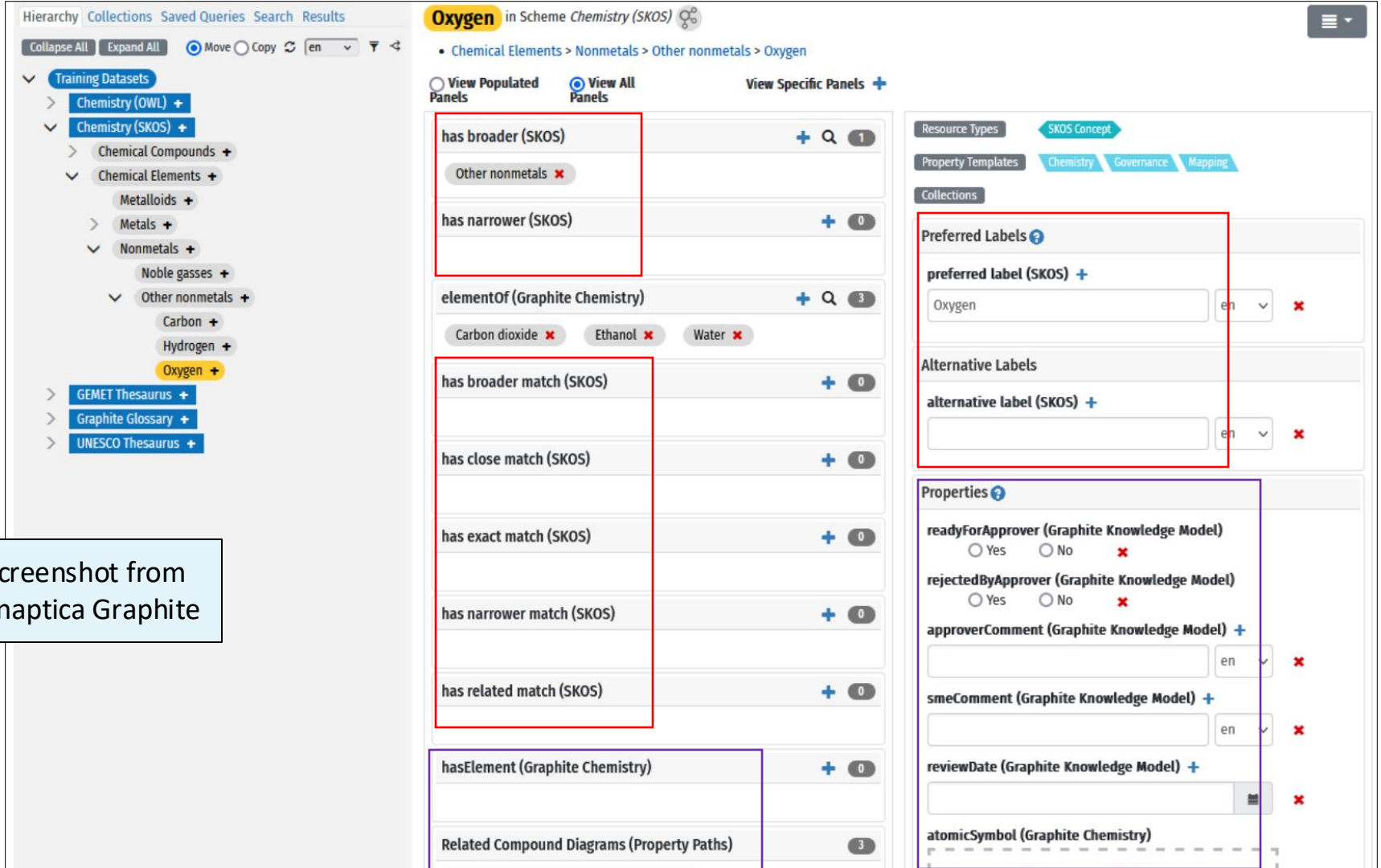
Screenshots from Graphologi

Concepts have both:

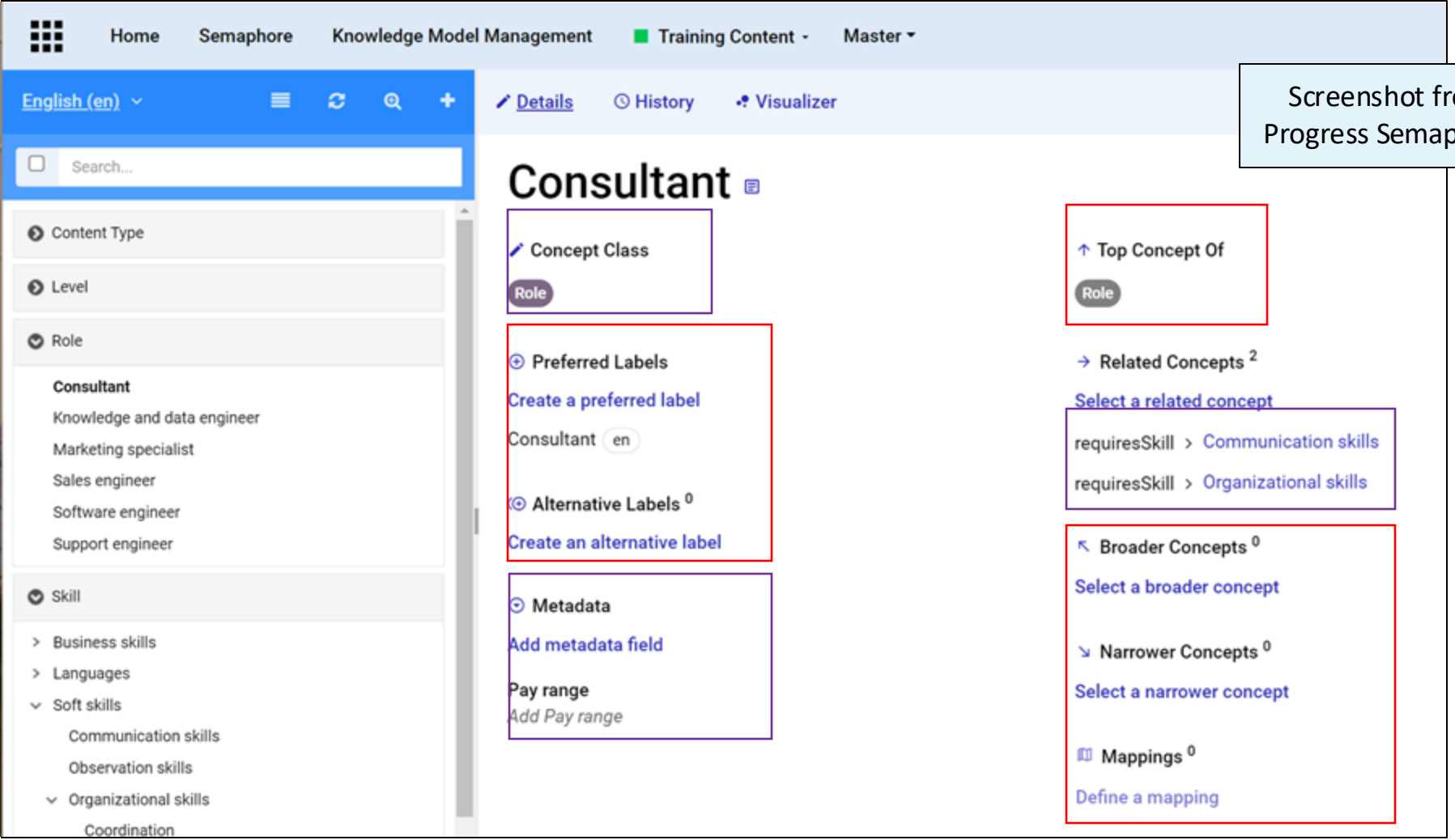
SKOS relationships and properties

and

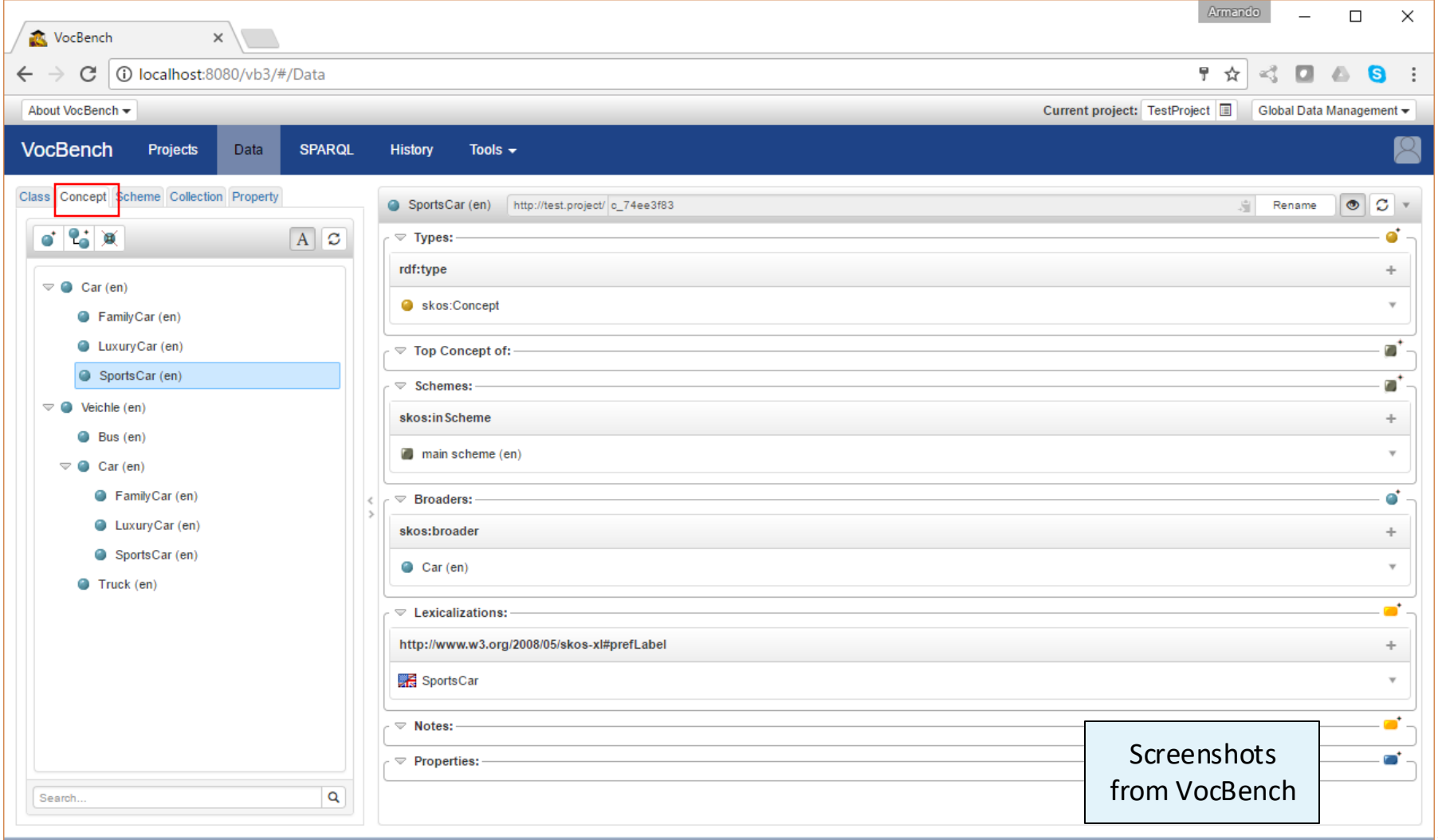
RDFS-based semantic relationships and attributes from an ontology-based custom scheme.

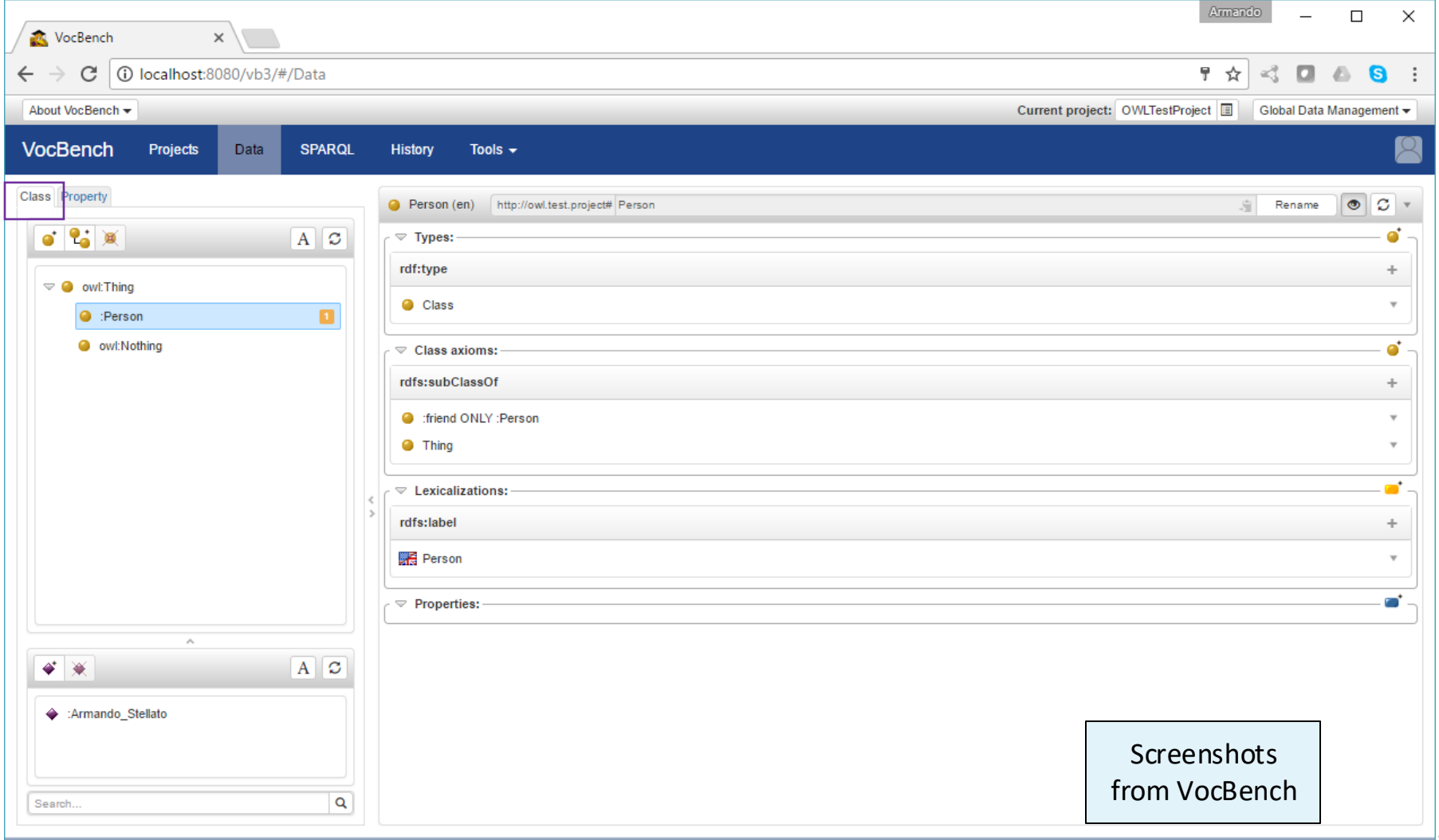


Screenshot from Synaptica Graphite



Screenshot from Progress Semaphore


VocBench
Armando
localhost:8080/vb3/#/Data
About VocBench
Current project: TestProject
Global Data Management
VocBench
Projects
Data
SPARQL
History
Tools
Class
Concept
Scheme
Collection
Property
Car (en)
FamilyCar (en)
LuxuryCar (en)
SportsCar (en)
Veichle (en)
Bus (en)
Car (en)
FamilyCar (en)
LuxuryCar (en)
SportsCar (en)
Truck (en)
Search...
SportsCar (en)
http://test.project/ c_74ee3f83
Rename
Types:
rdf:type
skos:Concept
Top Concept of:
Schemes:
skos:inScheme
main scheme (en)
Broaders:
skos:broader
Car (en)
Lexicalizations:
http://www.w3.org/2008/05/skos-xl#prefLabel
SportsCar
Notes:
Properties:
Screenshots from VocBench
© ART Group, 2016


VocBench
Armando
localhost:8080/vb3/#/Data
About VocBench
Current project: OWLTestProject
Global Data Management
VocBench
Projects
Data
SPARQL
History
Tools
Class
Property
owl:Thing
:Person
owl:Nothing
:Armando_Stellato
Search...
Person (en)
http://owl.test.project#
Person
Rename
Types:
rdf:type
Class
Class axioms:
rdfs:subClassOf
:friend ONLY :Person
Thing
Lexicalizations:
rdfs:label
Person
Properties:
Screenshots from VocBench
© ART Group, 2016

Conclusions

Why connect taxonomies and ontologies or extend a taxonomy to an ontology?

- Provides more uses than either a taxonomy or ontology alone
- Brings together the strengths of each: semantic and linguistic expressiveness
- Supports knowledge graphs and semantic layers and their benefits
- Makes use of existing taxonomies
- Conforms to both taxonomy and ontology standards for interoperability

When to connect taxonomies and ontologies

- When you have identified a business use case
- When you have resources skilled in taxonomies, ontologies, and implementations
- When you have a combined taxonomy/ontology management tool

How to connect a taxonomy and ontology

- Develop an ontology based on existing taxonomies/term lists, and add the ontology “layer”
- Use a software tool dedicated to combined taxonomy/ontology management (based on RDF, SKOS, RDFS and OWL)

Resources on Taxonomy / Ontology Combinations

Posts from The Accidental Taxonomist Blog

- "Taxonomies and Ontologies as Semantic Models," August 18, 2024
- "Taxonomies vs. Ontologies," January 31, 2023

Presentations

- "The Role of Taxonomy and Ontology in Semantic Layers" webinar, April 16, 2024
- (*first half hour of 1-hour video recording link*)
  "Taxonomy-Driven Ontology Design" The Knowledge Graph Conference, May 8, 2023.
  (*1.5-hour video recording link*)

Articles by others

- "Combined usage of SKOS and OWL: an experimentation on the Digital Europa Thesaurus," by Nicolas Figay, April 21, 2024
- "Editing schemas, ontologies, and SKOS taxonomies with VocBench," by Bob Ducharme, August 25, 2024

Upcoming Taxonomy Presentations and Workshops

- "Connecting Users to Content Through Taxonomies" full-day pre-conference workshop Information Architecture Conference, Philadelphia, April 29, 2025
- Virtual Bite-Sized Taxonomy Boot Camp London (3 sessions/hours each time) March 12, June 18, and October 8, 2025

Any time:

- "Controlled Vocabularies and Taxonomies" training course of four 75-minute video recordings, HS Events
- "Taxonomies and Controlled Vocabularies" training course of 5 web-based lessons and exercises, Hedden Information Management

Questions/Contact

Heather Hedden
Taxonomy Consultant
Hedden Information Management
Carlisle, MA USA
heather@hedden.net
www.hedden-information.com
accidental-taxonomist.blogspot.com
www.linkedin.com/in/hedden

Hedden Information Management

Making information findable