

Collaborative Ontology Development

Tania Tudorache

Protégé Short Course

Stanford University

October 11, 2017

Collaborative Ontology Development

- **Collaboration**: a process in which a community of users contribute to the development of one or more ontologies
- Ontologies vary quite a lot: from simple taxonomies to fully fledged OWL
- Usually the community is made of (some) ontology experts and (several) domain experts with varying expertise
- Common concern: how to get the domain experts to contribute without overloading them with representation details
- Think how you author a Word document with your colleagues: same and more challenges apply to ontologies!

Lorem Ipsum

[Lorem Ipsum](#)

[Introduction](#)

[2. Methods](#)

[2.1. Lorem Method](#)

[2.2. Ipsum Method](#)

Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed vitae faucibus enim, non viverra turpis. Aenean semper hendrerit diam vel ultrices. Maecenas porttitor ante lectus, a porta justo pellentesque eu. Donec elementum justo suscipit viverra hendrerit. Donec at egestas purus. Interdum et malesuada fames ac ante ipsum primis in faucibus. Morbi hendrerit massa ligula, non egestas justo vestibulum sit amet. Suspendisse et rhoncus massa. Integer ultrices tempus erat. Quisque eu mauris mi. Donec eu consectetur sem. Morbi libero ipsum, aliquet vitae ex id, suscipit sollicitudin orci. Suspendisse egestas dolor tellus, in malesuada diam condimentum eget. Morbi imperdiet sapien non risus auctor, eget fringilla sem maximus. Cras commodo ultricies metus id tempus.

2. Methods

Definition: Duis ut lorem venenatis, interdum est sit amet, facilisis nunc. Curabitur dapibus, neque ut tempor semper, nunc ex hendrerit sapien, eu hendrerit ex dolor et est. Integer sit amet mauris aliquam leo vehicula tristique sit amet nec nunc. Nulla commodo est quis tellus dictum vehicula. Nunc et luctus libero. Phasellus finibus semper maximus. Fusce sit amet libero sem. Suspendisse potenti. Donec lorem tortor, venenatis et sem ac, volutpat lacinia dolor. Suspendisse tristique metus non semper posuere. Proin at suscipit elit. Nam vel augue turpis. Proin sed ex ultricies, tincidunt ligula ac, placerat tellus.

2.1. Lorem Method

Nam porttitor mollis enim, quis hendrerit ipsum elementum in. Integer odio dui, hendrerit eu consectetur at, ullamcorper eget leo. Nulla facilisi. Fusce nisl felis, lacinia vel dictum ut, rutrum eget neque. Nam pellentesque ipsum sed lacus sollicitudin placerat. Morbi in nunc id lacus maximus luctus. Nullam eget neque et nulla convallis accumsan. Pellentesque elit ligula, venenatis sed odio et, viverra venenatis tortor.



Google Docs

Contributors:



Roles:

- writer
- reviewer
- proof-reader

Structure

Style

Workflow

Quality Assurance

Publishing

2 parts: Editing and Publishing

Editing



- “Internal” development team
- Tools for editing ontologies
- Short development cycles

What your team does

Publishing



- External release of ontologies for public/production use
- Tools for supporting versioning, public feedback and reviewing
- Set release cycles

How others use your ontology

Use cases of collaborative development

Use case: The NCI Thesaurus collaborative development process

NCI Thesaurus: Reference ontology for cancer biology, translational science, and clinical oncology



WebProtégé NCI Thesaurus

Classes Properties Individuals Notes and Discussions Changes By Entity Project Dashboard

Classes

Create Delete Watch Branch Search: Type search

- owl:Thing
 - Abnormal Cell
 - Activity
 - Anatomic Structure, System, or Substance
 - Biochemical Pathway
 - Biological Process
 - Chemotherapy Regimen or Agent Combination
 - Conceptual Entity
 - Diagnostic or Prognostic Factor
 - DIRECTED-BINARY-RELATION
 - Disease, Disorder or Finding
 - Drug, Food, Chemical or Biomedical Material
 - Experimental Organism Anatomical Concept
 - Experimental Organism Diagnosis
 - Gene 2**
 - Gene Product
 - Manufactured Object
 - Molecular Abnormality
 - NCI Administrative Concept
 - Organism
 - PAL-CONSTRAINT
 - Property or Attribute

Class description for Gene

Display name: Gene
IRI: <http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C16612>

Annotations		
rdfs:label	Gene	
ALT_DEFINITION	<ncicp:ComplexDefinition><ncicp:def-definition>The functional and physical unit of heredity passed from parent to offspring. Genes are pieces of DNA, and most genes contain the information for making a specific protein.</ncicp:def-definition><ncicp:def-source>NCI-GLOSS</ncicp:def-source></ncicp:ComplexDefinition>	
code	C16612	
DEFINITION	<ncicp:ComplexDefinition><ncicp:def-definition>A functional unit of heredity which occupies a specific position on a particular chromosome and serves as the template for a product that contributes to a phenotype or a biological function.</ncicp:def-definition><ncicp:def-source>NCI</ncicp:def-source></ncicp:ComplexDefinition>	
DesignNote	The gene as a functional unit consists of a discrete segment of a giant DNA molecule containing the purine (adenine and guanine) and pyrimidine (cytosine and thymine) bases in the ordered and correct sequence that encodes a specific functional product (i.e., a protein or RNA molecule).	

Use case: The NCI Thesaurus collaborative development process (cont.)

NCI Thesaurus: Reference ontology for cancer biology, translational science, and clinical oncology

Classes

te Watch Branch Search: Oncogene_TIM

- Oncogenes, G-Proteins
 - Oncogene TIM
- RAS Family Oncogene
- Oncogenes, Growth Factor
- Oncogenes, Nuclear Protein
- Oncogenes, Protein-Kinase
- PCA3 Gene
- SPRY4-IT1 Gene
- STL Gene
- Susceptibility/Resistance Gene
- TCL6 Gene
- TRERF1 Gene
- Tumor Promoter Induced Gene

Conditions for Oncogene TIM

Equivalent To

('Oncogenes, G-Proteins'
and Allele_In_Chromosomal_Location some 7q33-q35
and Gene_Plays_Role_In_Process some 'Signal Transduction'
and Gene_Found_In_Organism some Human)

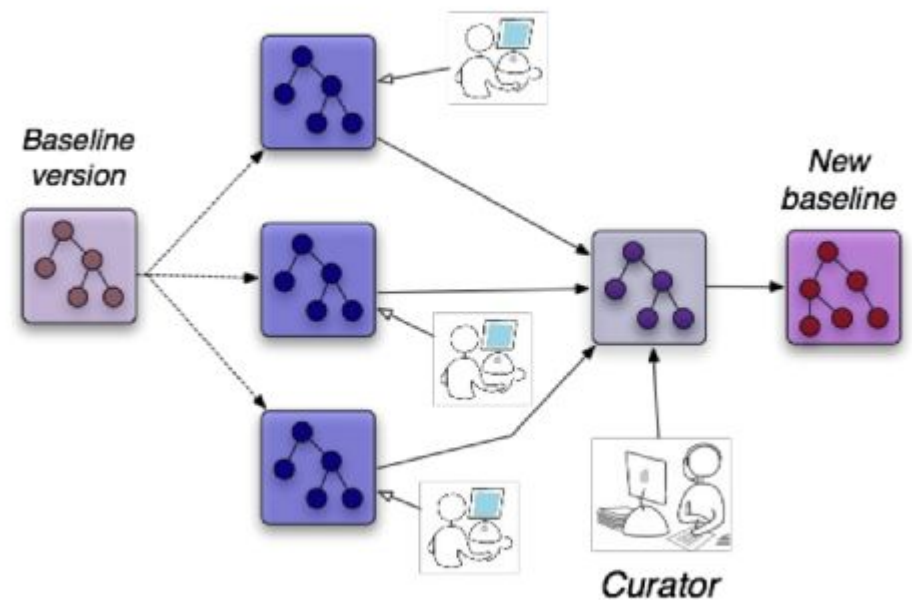
SubClass Of

SubClass Of Ancestor Class

Gene
Gene_Plays_Role_In_Process some Tumorigenesis
Oncogene
Gene_Plays_Role_In_Process some 'Signal Transduction'
'Cancer Gene'
Gene_Found_In_Organism some Human
Gene_Plays_Role_In_Process some Tumorigenesis

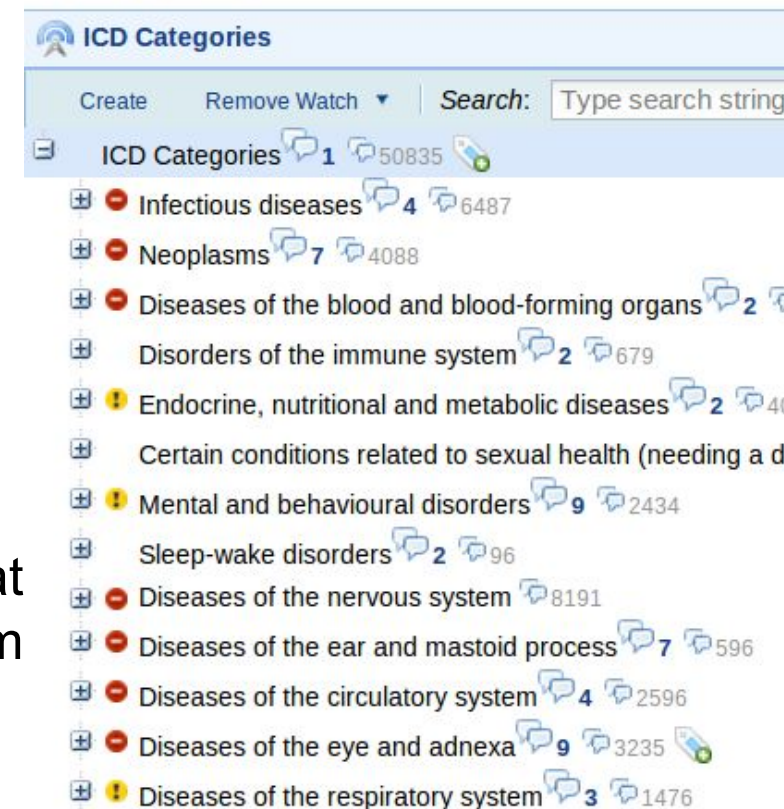
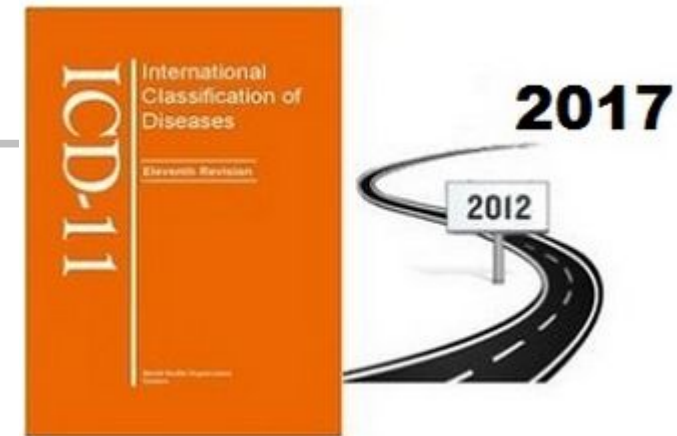
Use case: The NCI Thesaurus collaborative development process (cont.)

- Simultaneous editing in Protégé clients
- Custom UI for restricting user input and enforcing business rules
- Development cycle begins after baseline
- ~20 full-time editors making changes; 1 lead editor approves the changes, and assigns new tasks
- Released version on NCI website and BioPortal



Use case: WHO's ICD

- ICD – International Classification of Diseases
- Developed by the World Health Organization (WHO)
- Current revision in use: ICD-10, development work on ICD-11
- Over 10.000 categories used for coding, billing, statistics, policy making all over the world
- Collaborative and international effort
- Current version: published as books
- Goal for the new version: use a more formal representation and published in electronic format use Web-based collaboration and social platform for editing

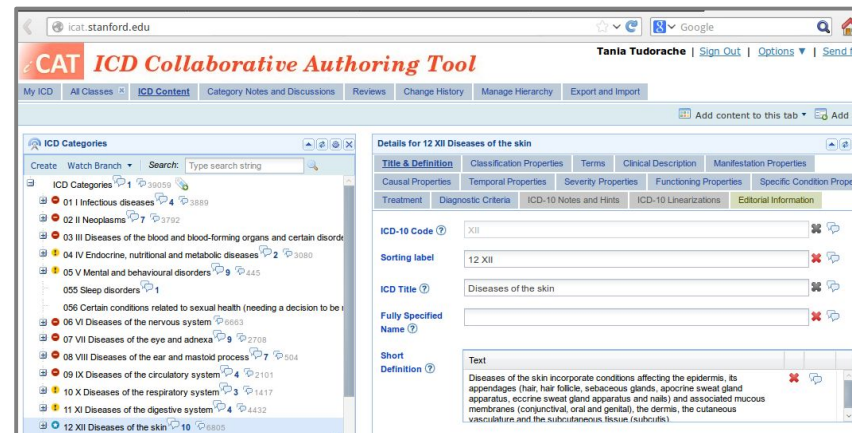


Construction of ICD-10: Revision Process in the 20th Century

- 8 Annual Revision Conferences (1982 - 89)
 - 17 – 58 countries participated
 - 1- 5 person delegations
 - Mainly health statisticians
- Manual curation
 - List exchange
 - Index was done later
- "Decibel" Method of discussion
- Output: Paper Copy
- Work in English only
- Limited testing in the field



ICD-11 process today



- ICD-11 is an OWL ontology edited collaboratively in WebProtégé
- Over 250 domain experts from around the world
- Organized in groups, which edit different parts of the ontology

ICD-11 process today (cont.)

- Each night a snapshot of the commonly edited ontology is published in a public platform to encourage feedback from the larger community
<http://apps.who.int/classifications/icd11/browse/f/en>
- Editorial workflow
- Centrally overseen by WHO
- Peer-reviewed process for the content and structure
- WebProtégé used as the collaborative ontology development platform

Collaboration infrastructures and more use cases

Collaboration Infrastructures: Wikis

- Wikis are well known; Wikipedia
- Semantic Wikis – add semantic extensions to the wiki platforms
- Assign a wiki page to an entity in the ontology
- Usually focused on filling in a knowledge base, less on the classes
- Export/import RDF

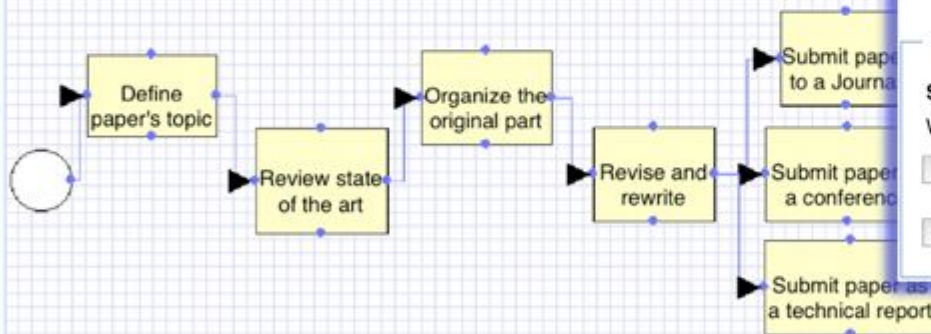
Semantic Wiki: MoKi

Domain Model IsA Browser



Write a paper

Export to eRDF format (for Oryx) Only this process | Sub processes in separate file(Zip Archive) | Sub processes in sa



Description: The process of writing a scientific paper

Required concept: Publication, ResearchTopic, Event, Person

Modify Concept: Workshop

Annotations

Description: An educational seminar or series of meetings emphasizing interaction and exchange of information among a usually small number of participants

Synonyms:

Hierarchical Structure

Workshop is a:

Workshop is part of:

Properties

Subject	Property	Object(s)
Workshop	participant	<input type="text" value="Person"/>
<input type="button" value="Remove"/>		
<input type="button" value="Add another"/>		

The challenge with wikis

BOWiki syntax

Examples

1 on page Apoptosis: `[[OType:Category]]`

2 on page Apoptosis: `[[CC-isa::Biological_process]]`

3 on page HvSUT2: `[[Realizes:: function = Sugar_transporter_activity;
process = Glucose_transport]]`

4 on page Realizes: `[[has-argument:: name = function; type =
OType:Function_category]]`

OWL abstract syntax

`Individual(Apoptosis, type(Category))`

`Individual(Apoptosis value(CC-isa Biological_process))`

`Individual(Realizes-0 type(Realizes))`

`Individual(Realizes-0 value(Realizes-subject HvSUT2))`

`Individual(Realizes-0 value(Realizes-process Glucose_transport))`

`SubClassOf(Realizes gfo:Relator))`


`ObjectProperty(Realizes-function domain(Realizes)
range(Function_category))`




Function Name:	<input type="text" value="Sugar_transporter_activity"/>	<input type="button" value="Load"/>
Functional Item:	<input type="text" value="[[Sugar_transporter]]"/>	
Goal:	<input type="text" value="Sugar located_in B"/>	
Requirements:	<input type="text" value="Sugar present AND Sugar located_in A"/>	
	<input type="button" value="Store"/>	<input type="button" value="Update"/>







Project management platforms, content management platforms or issue trackers

- Some collaboration projects reuse Web-based platforms built for software and other types of projects
- Example: GitHub, SourceForge, GForge, RedMine, WordPress
- Issue trackers: Mostly for managing change proposals or reviews
- Content Management platforms: tagging articles with existing vocabularies (WordPress has some semantic extensions)


Gene Ontology Issue Tracker in GitHub

 geneontology / go-ontology



 Watch 27  Star 7  Fork 1



 Code  Issues 364  Pull requests 0  Projects 1  Pulse  Graphs




Labels Milestones New issue



 364 Open ✓ 12,403 Closed




Author ▾ Labels ▾ Milestones ▾ Assignee ▾ Sort ▾




 maintenance of blood-brain barrier and regulation terms
#12767 opened 2 days ago by tberardini  1

 Review MF term GO:0050827 'toxin receptor binding'
#12766 opened 2 days ago by vanaukenk 

 ZFIN GO file load is failing blocker
#12765 opened 3 days ago by paolaroncaglia   3

 basal pole of outer hair cell
#12764 opened 3 days ago by slaulederkind 

 erroneous links involving spindle pole body terms parent relationship query PomBase
#12761 opened 6 days ago by mah11   4

 Add is_a link between mitochondrial tRNA modification and mitochondrial tRNA processing
dictyBase Other term-related request parent relationship query
#12757 opened 10 days ago by pfey03   2

Gene Ontology - TermGenie

▼ Step 2: Templates *involved_in* (1)

Once you have selected the ontology, the available term generation patterns can be selected from a menu.

Select Template

Template: ***involved_in*** ([More](#), [Less](#))

Description: processes involved in other processes

Hint: *[part] involved in [whole]*

Required

part

biological_process

whole

biological_process

Optional
Name

Definition

DefX_Ref

epithelial cell proliferation

renal system development

([More](#), [Less](#))

After selecting and filling templates, click on the 'Verify Input'-Button below to start the next step.



Please wait.

Verifying your request and generating terms on the server.

Other collaboration processes

- Use source control repositories (e.g., SVN)
 - Text based mechanisms
 - Hard to merge local copies in the shared copy
- Locking mechanisms (lock parts of an ontology for editing)
- Use specialized (domain dependent) ontology repositories, e.g., BioPortal

Textual diffs can be tricky



Textual diff: 588 lines, including the “only” change:

```
648c648,655
< <owl:someValuesFrom rdf:resource="http://www.co-ode.org/ontologies/pizza/2005/10/18/pizza.owl#CheeseTopping"/>
...
> <owl:someValuesFrom>
>   <owl:Class>
>     <owl:unionOf rdf:parseType="Collection">
>       <rdf:Description rdf:about="http://www.co-ode.org/ontologies/pizza/2005/10/18/pizza.owl#CheeseTopping"/>
>       <rdf:Description rdf:about="http://www.co-ode.org/ontologies/pizza/2005/10/18/pizza.owl#VegetableTopping"/>
>     </owl:unionOf>
>   </owl:Class>
> </owl:someValuesFrom>
```


Protégé diff support:

Tools menu → Compare Ontologies

The screenshot shows the Protégé ontology editor interface. The 'Tools' menu is highlighted with a red circle. The 'Ontology Differences' dialog is open, showing a comparison between two versions of the 'pizza' ontology. The dialog has a 'Find' search bar and a table with three columns: 'Description', 'Baseline Axiom', and 'New Axiom'. The 'Baseline Axiom' and 'New Axiom' columns are circled in red.

File Edit View Reasoner **Tools** Refactor Window Help

← → pizza (http://www.co-ode.org/ontologies/pizza/2005/10/18/pizza.owl)

Data Properties Annotation Properties Individuals OWLViz DL Query OntoGraf Ontology Differences SPARQL Query Rules

Active Ontology Entities Classes

Class hierarchy Class hierarchy (inferred)

Class hierarchy: CheeseyPizza Annotations: CheeseyPizza

Annotations Usage

Thing

- DIRECTED-BINARY-RELATIONSHIP
- DomainConcept
 - Country
 - IceCream
 - Pizza
 - CheeseyPizza
 - InterestingPizza
 - MeatyPizza
 - NamedPizza
 - NonVegetarianPizza
 - RealItalianPizza
 - SpicyPizza

Ontology Differences

Find

Created: defaultLanguage
Created: versionInfo
Modified: CheeseyPizza

Source and target entities aligned because they have the same IRI

Description	Baseline Axiom	New Axiom
Definition changed	CheeseyPizza EquivalentTo Pizza and (hasTopping some CheeseTopping)	CheeseyPizza EquivalentTo Pizza and (hasTopping some (CheeseTopping or VegetableTopping))

See demo screencast at: <https://www.youtube.com/watch?v=JzMNDfy4jcg>

Collaboration support in WebProtégé

WebProtégé – quick overview

- **Free, open source collaborative ontology development environment for the Web**
- *Google docs* for ontologies; over 30,000 ontologies submitted or created by users
- OWL 2 ontologies
- A default *simple editing interface*
- *Full change tracking* and revision history
- *Collaboration* tools such as, sharing and permissions, threaded notes and discussions, watches and email notifications
- *Customizable user interface*
- *Customizable Web forms* for application/domain specific editing
- *Multiple formats* for upload and download of ontologies (supported formats: RDF/XML, Turtle, OWL/XML, OBO, and others)

WebProtégé – simplified editing interface

The screenshot displays the WebProtégé interface with the following components:

- Top Navigation Bar:** Home | BFO
- Menu Bar:** Classes (selected), Properties, Individuals, Comments, Changes by Entity, History
- Class Hierarchy Panel (Left):**
 - owl:Thing
 - entity
 - continuant
 - generically dependent continuant
 - independent continuant
 - immaterial entity
 - continuant fiat boundary
 - one-dimensional continuant fiat boundary** (selected)
 - two-dimensional continuant fiat boundary
 - zero-dimensional continuant fiat boundary
 - site
 - spatial region
 - material entity
 - specifically dependent continuant
 - occurent

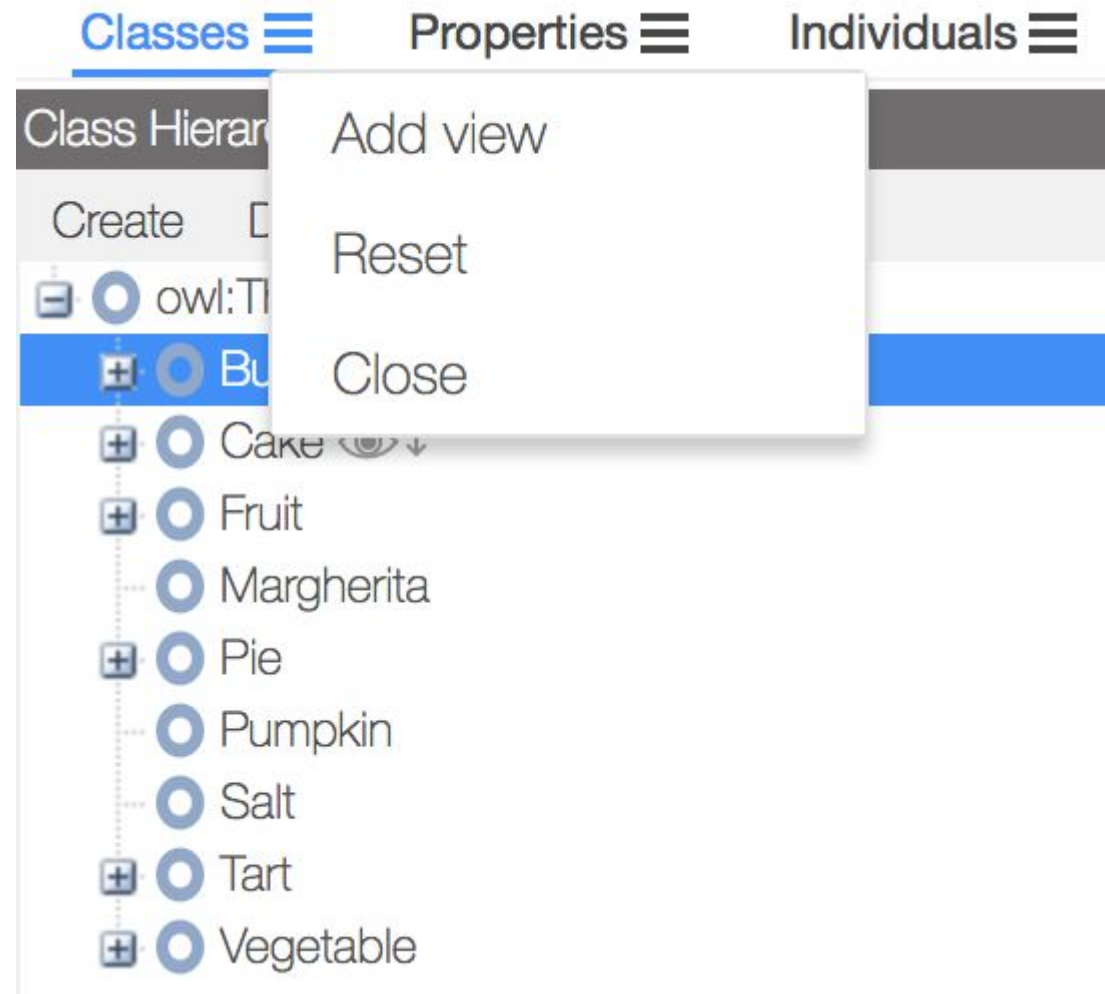
- Class Details Panel (Right):**
- Class:** one-dimensional continuant fiat boundary
- IRI:** http://purl.obolibrary.org/obo/BFO_0000142
- Annotations:**

Property	Value	Language	Action
rdfs:label	one-dimensional continuant fiat boundary	en	✕
BFO CLIF specification label	OneDimensionalContinuantFiatBoundary	lang	✕
BFO OWL specification label	1d-cf-boundary	lang	✕
elucidation	a one-dimensional continuant fiat boundary is a continuous fiat line whose location is defined in relation to some material entity. (axiom label in BFO2 Reference: [032-001])	en	✕
example of usage	The Equator	en	✕
example of usage	all geopolitical boundaries	en	✕
example of usage	all lines of latitude and longitude	en	✕
example of usage	the line separating the outer surface of the mucosa of the lower lip from the outer surface of the skin of the chin.	en	✕
example of usage	the median sulcus of your tongue	en	✕
has associated axiom(fol)	(iff (OneDimensionalContinuantFiatBoundary a) (and (ContinuantFiatBoundary a) (exists (b) (and (OneDimensionalSpatialRegion b) (forall (t) (locatedInAt a b t)))))) // axiom label in BFO2 CLIF: [032-001]	lang	✕
rdfs:isDefinedBy	http://purl.obolibrary.org/obo/bfo.owl	lang	✕
- Input Fields:**
 - Enter property
 - Enter value
 - lang

<http://webprotege.stanford.edu>

Customizable Interface

- Add new tab
- Add views to tabs
- Layout is saved and restored in the next session



Collaboration Features in WebProtégé

- Sharing ontologies
- Simultaneous editing
- Change tracking
- Threaded discussions for ontology entities and changes (notes, discussions, proposals, reviews)
- Watching ontology entities and branches and notifications
- Upload and sharing of ontologies
- Download/revert any revision of the ontology
- Access policies
- User interface customization

Uploading and Sharing an ontology

Sharing Settings

☒ Link sharing enabled (Sign-in Required)

Anyone with the link can view

Share with specific people:

ttania4

manage



Matthew Hor

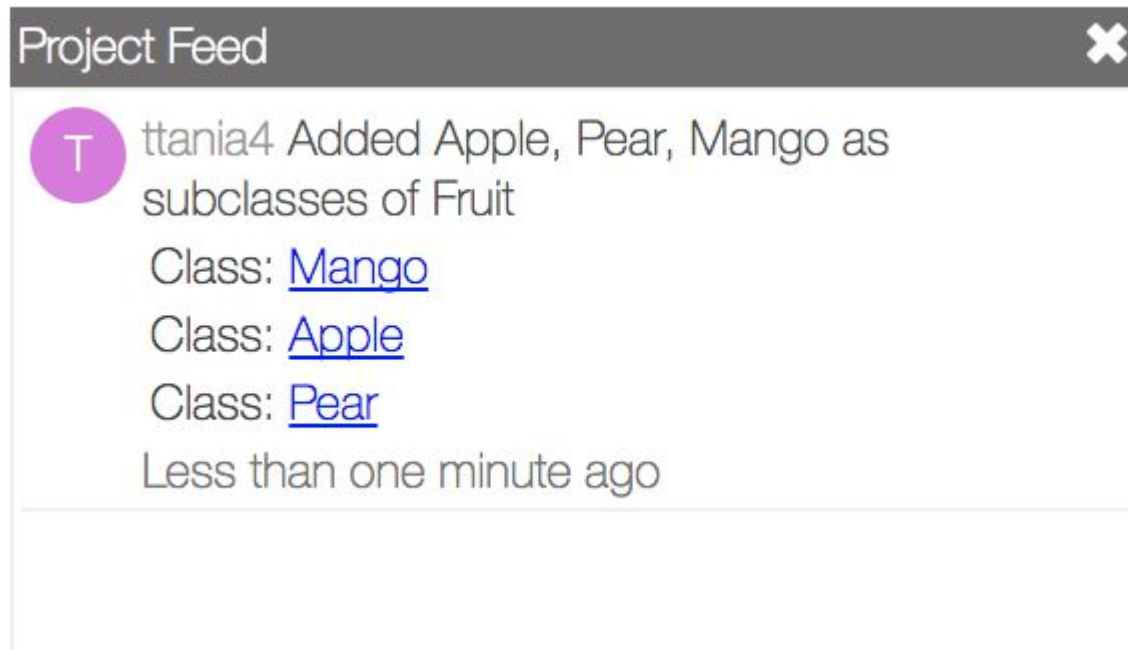
view

Matthew Horridge

CANCEL

APPLY

Project feed: see what other online users are changing in the ontology



Feed updates as new changes are happening

Notes and Discussions

Classes

Properties

Individuals

Comments

Changes by Entity

History

Add tab

Commented entities

Sort by Entity

ApplePie from mobile phone

1 Last comment **3rd november 2015** by M Horridge

1 unresolved

Beet Root

1 Last comment **3rd november 2015** by Csongor Nyulas

1 unresolved

Carrot Cake

1 Last comment **3rd november 2015** by M Horridge

1 unresolved

Pumpkin Pie

1 Last comment **3rd november 2015** by M Horridge

resolved

Comments: Beet Root

Start new thread

CN Csongor Nyulas

3rd November 2015

Error This suppose to be under Root Vegetable

Resolve

Reply

Change tracking

Classes

Properties

Individuals

Comments

Changes by Entity

History

Add tab

Class Hierarchy



Create Delete Watch Search

- owl:Thing
 - Butter
 - Cake
 - Fruit
 - Margherita**
 - Pie
 - Pumpkin
 - Salt
 - Tart
 - Vegetable
 - Beet Root (1)
 - Broccoli
 - Carrot
 - Root Vegetable

Entity Changes



Changes on Tue, 3 Nov 2015

Deleted class: Pizza

R 59

MH M Horridge authored 1 changes 3rd November 2015

[root-ontology] Margherita **SubClassOf** RBiLILU7dWsdXCu1Bbv8l1e

Edited class

R 56

J johardi authored 2 changes 3rd November 2015

[root-ontology] Margherita rdfs:label "Margarita"

[root-ontology] Margherita rdfs:label "Margherita"

Created Margarita as a subclass of Pizza

R 46

J johardi authored 3 changes 3rd November 2015

[root-ontology] **Class:** Margherita

[root-ontology] Margherita rdfs:label "Margarita"

[root-ontology] Margherita **SubClassOf** RBiLILU7dWsdXCu1Bbv8l1e

Watching entities and branches

Class Hierarchy

Create Delete Watch Search

owl:Thing

Butter

Cake

Banana Cake

Carrot Cake (1)

Cupcake

Diplomat Cake

Fruit Cake

Ginger Bread Cake

Pineapple Cake

Sponge Cake

Swiss Roll

Upside Down Cake

Fruit

Margherita

Pie

Pumpkin

Salt

Class: Cake

IRI

<http://webprotege.stanford.edu/RDf13wGdBHcvnNRsu82S1DX>

Annotations

rdfs:label Cake lang

Classes

owl:Thing

Relations

Select the type of watch


☐ None

☐ Entity

☒ Branch

Ok

History

 Home | **Cake** Project ttania4 ▾ Help ▾

Classes ▮ Properties ▮ Individuals ▮ Comments ▮ Changes by Entity ▮ History ▮ Add tab

Project History ✕

Refresh ⌵

● Changes on Tue, 10 Oct 2017

Added Apple, Pear, Mango as subclasses of Fruit R 63 ▾

T

ttania4 authored 9 changes 13 minutes ago

⊕ [root-ontology] **Class:** Apple

⊕ [root-ontology] Apple rdfs:label "Apple"

⊕ [root-ontology] Apple **SubClassOf** Fruit

⊕ [root-ontology] **Class:** Mango

⊕ [root-ontology] Mango rdfs:label "Mango"

⊕ [root-ontology] Mango **SubClassOf** Fruit

⊕ [root-ontology] **Class:** Pear

⊕ [root-ontology] Pear rdfs:label "Pear"

⊕ [root-ontology] Pear **SubClassOf** Fruit

Revert changes in revision 63

Download revision 63

● Changes on Thu, 11 May 2017

Created individuals R 62 ▾

CN

Csongor Nyulas authored 6 changes 11th May 2017

⊕ [root-ontology] **NamedIndividual:** bananacake1

⊕ [root-ontology] bananacake1 rdfs:label "bananacake1"

⊕ [root-ontology] bananacake1 **Type** 'Banana Cake'

⊕ [root-ontology] **NamedIndividual:** bananacake2

Resources

- Online WebProtégé server: <http://webprotege.stanford.edu>
- Mailing list: <http://protege.stanford.edu/support.php>
- WebProtege on GitHub (sources and issue tracker):
<https://github.com/protegeproject/webprotege/>
- WebProtégé documentation: <http://protegewiki.stanford.edu/wiki/WebProtege>
- WebProtégé simplified user interface: "Simplified OWL Ontology Editing for the Web: Is WebProtégé Enough?" Horridge, Matthew, et al., The Semantic Web–ISWC 2013. Springer Berlin Heidelberg, 2013. 200-215.
- WebProtégé paper: "WebProtégé: A Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web", Tania Tudorache, Csongor Nyulas, Natalya F. Noy, Mark A. Musen, Semantic Web Journal (SWJ) 4 (Number 1 / 2013), 89 - 99
- WebProtégé in use: "Will Semantic Web Technologies Work for the Development of ICD-11?", T. Tudorache, S. M. Falconer, C. I. Nyulas, N. F. Noy, M. A. Musen. The 9th International Semantic Web Conference, ISWC 2010 (In-Use track), Shanghai, China, Springer. Published in 2010.
http://bmir.stanford.edu/file_asset/index.php/1646/BMIR-2010-1427.pdf
- Other References: <http://protegewiki.stanford.edu/wiki/WebProtege#References>