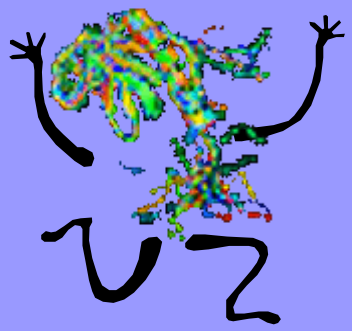
Introduction to Ontologies and Ontology Engineering

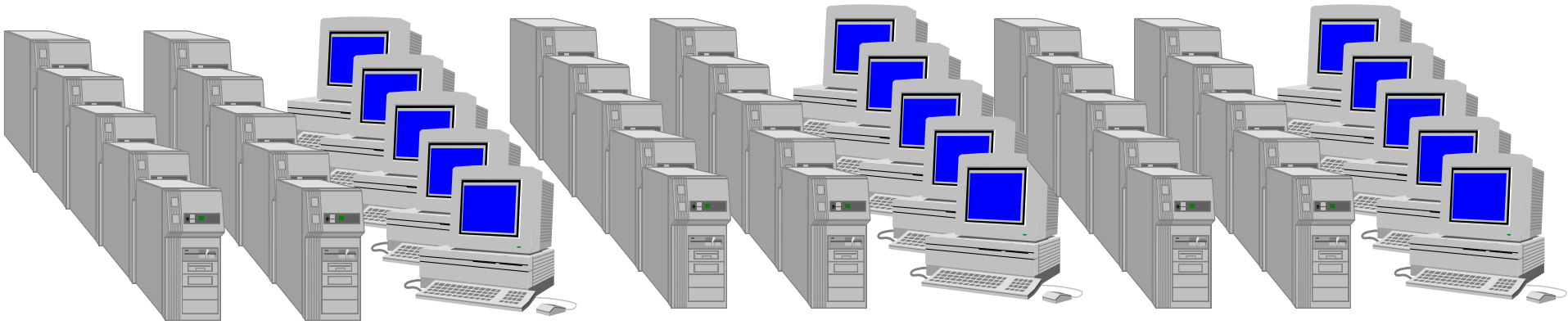
Patrick Lambrix
Linköping University

My current research interests

- Ontology engineering
  - Theory, methods, tools for ontology alignment, completion, debugging; organization of workshops and competitions
  - Development of ontologies for different areas (materials science, animal health surveillance, crime investigation, ice hockey)
- Sports analytics

GET THAT PROTEIN!

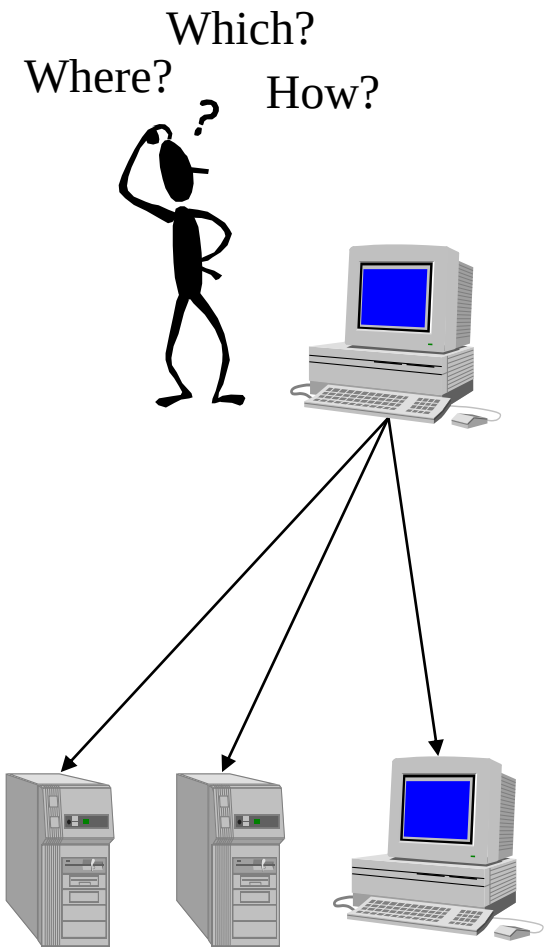




Locating relevant information

Vision: Web services

- Databases and tools (service providers) announce their service capabilities
- Users request services which may be based on task descriptions
- Service matchers find relevant services (composition) based on user needs and user preferences, negotiate service delivery, and deliver results to user



Findable
Accessible

Retrieving relevant information

Vision:

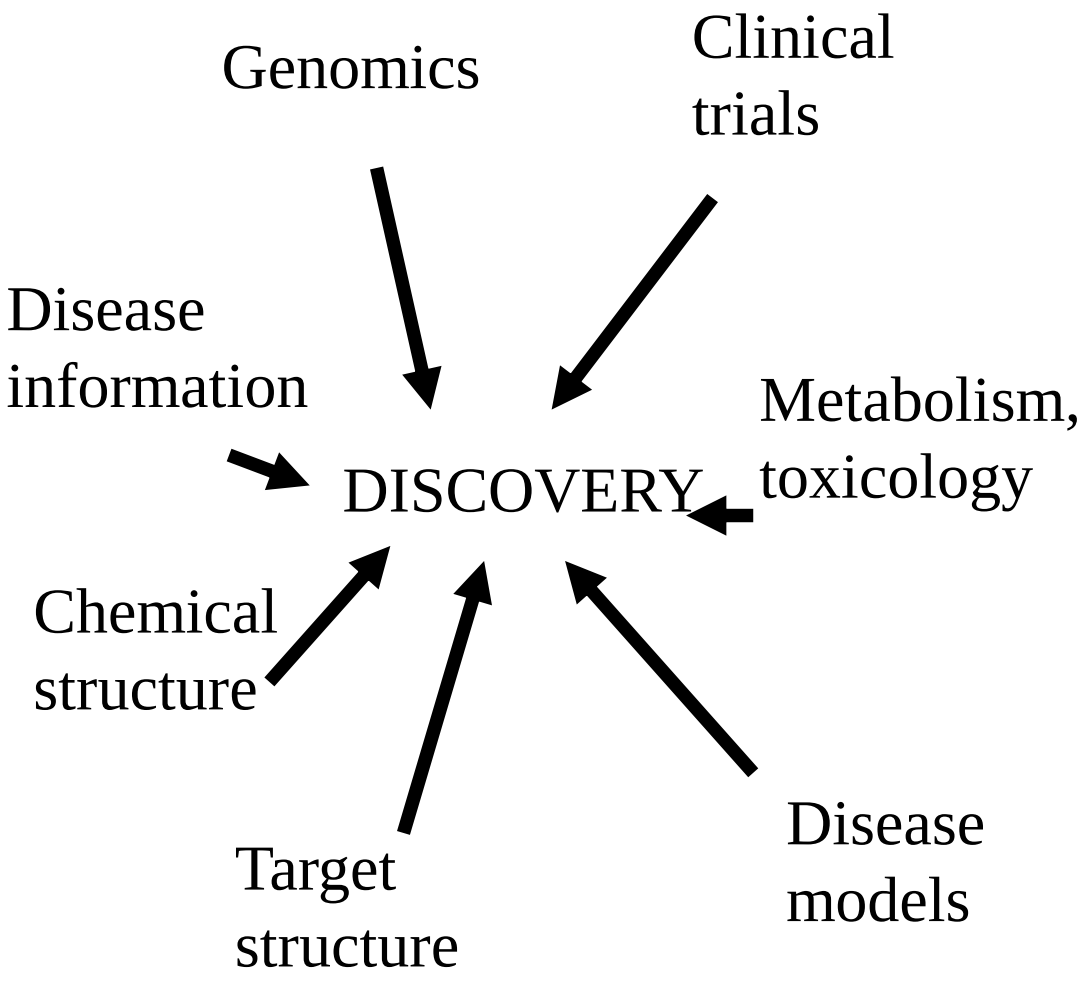
Based on the meaning of the query:
- only relevant information is retrieved
- all relevant information is retrieved

Findable
Accessible

Integrating information

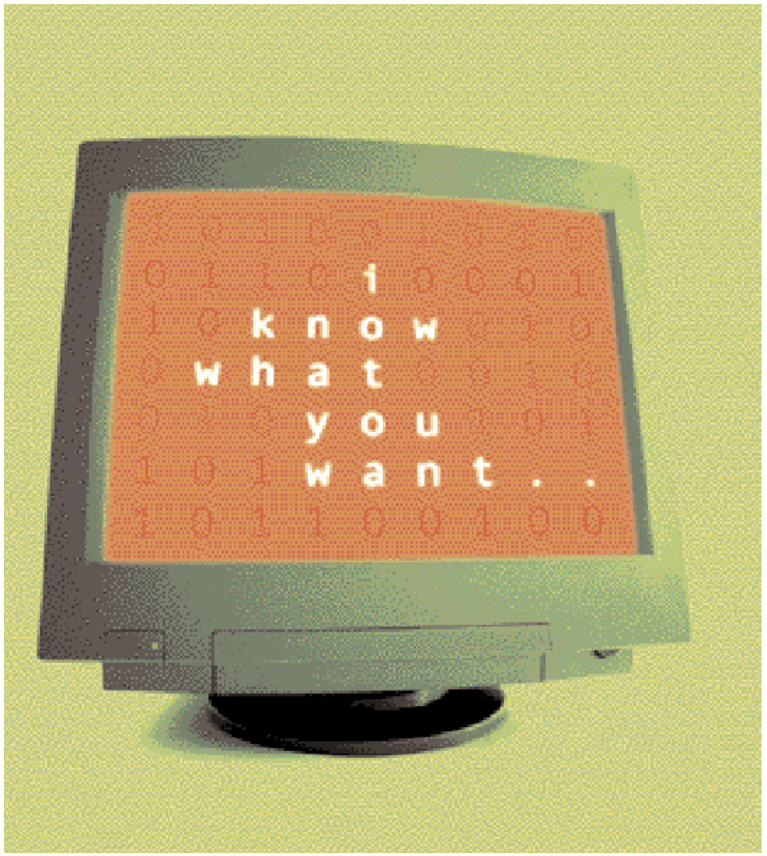
Vision:

Integrate data sources that are heterogeneous in content, data quality, data models, access methods, terminology



Interoperable
Reusable

Scientific American, May 2001:



THE SEMANTIC WEB

A new form of Web content
that is meaningful to computers
will unleash a revolution of new abilities

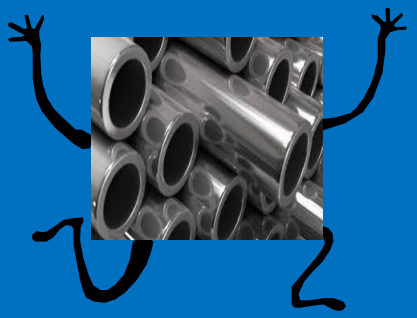
by
TIM BERNERS-LEE,
JAMES HENDLER and
ORA LASSILA

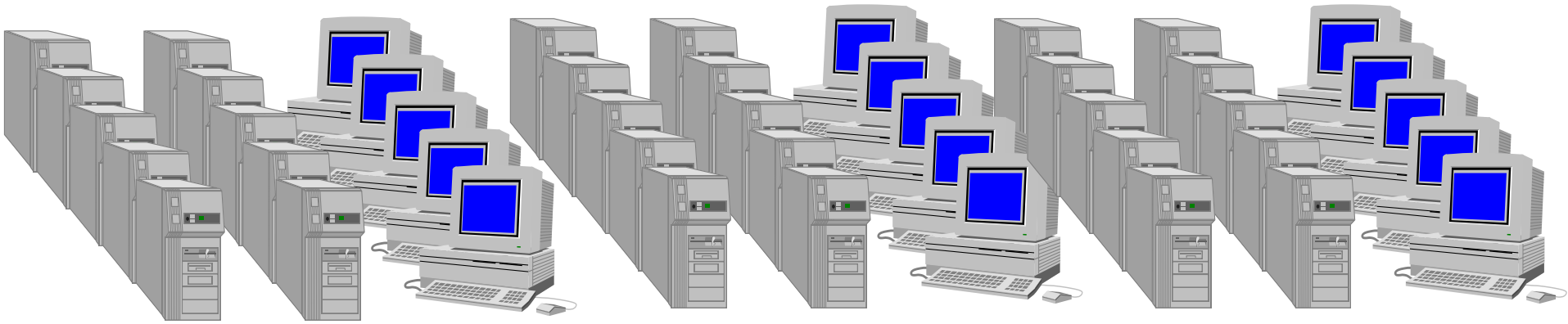
- First step towards the vision:
  adding semantic annotation to web resources

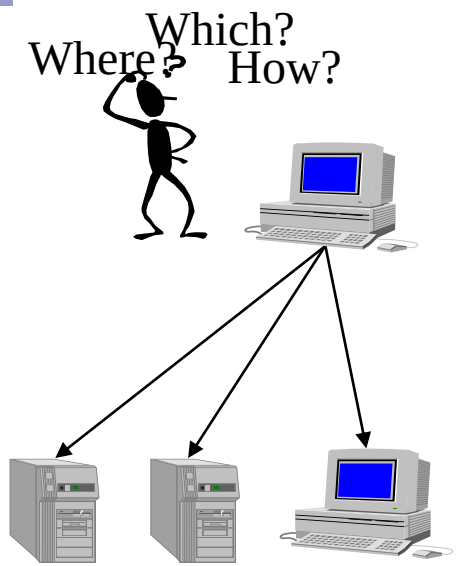
Semantic annotations based on ontologies

- Locating information
  - Web service descriptions use ontologies
  - Users use ontologies when formulating requests
  - Service matchers find services based on meaning
- Retrieving relevant information
  - Reduce non-relevant information (precision)
  - Find more relevant information (recall)
- Integrating information
  - Relating similar entities in different databases

GET THAT MATERIAL!







Cristal structure

Electronic structure

Materials design

Thermodynamic properties

→Standardized terminology

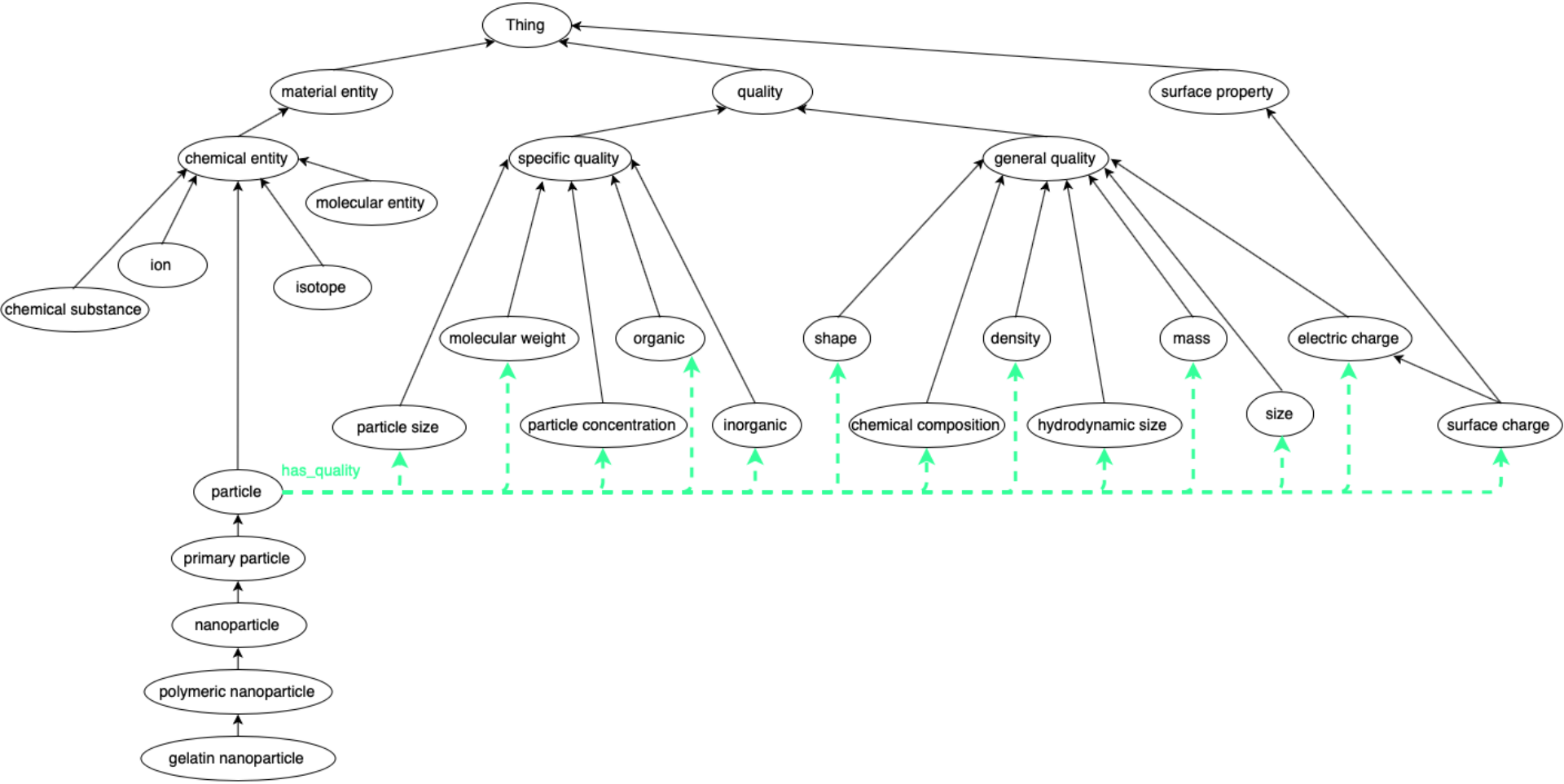
→Relationships between terms

Ontologies

"*Ontologies define the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms and relations to define extensions to the vocabulary.*"

(Neches, Fikes, Finin, Gruber, Senator, Swartout, 1991)

Nanoparticle Ontology



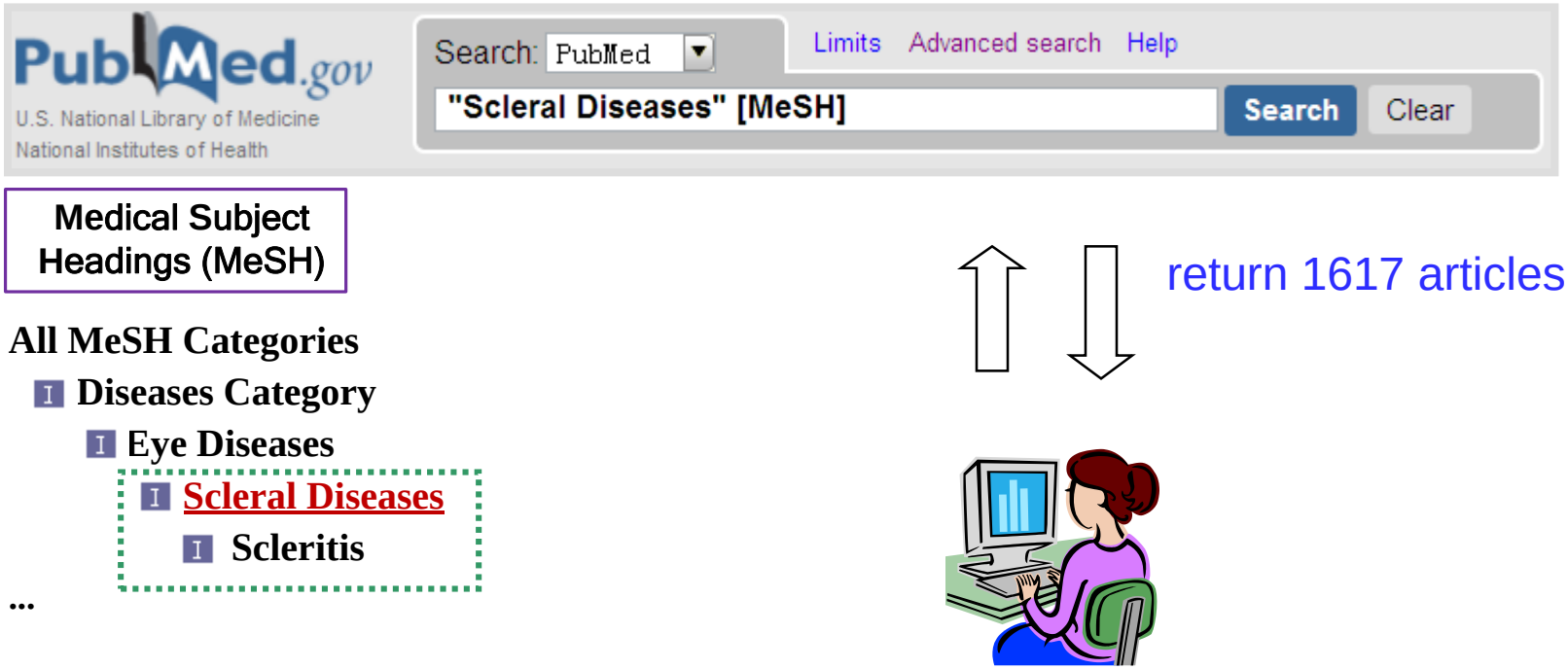
Ontologies used …

- for communication between people and organizations
- for enabling knowledge reuse and sharing
- as basis for interoperability between systems
- as repository of information
- as query model for information sources

Key technology for the Semantic Web

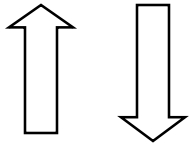
Ontology-based querying

- With ontology.



Ontology-based querying

- Without ontology.

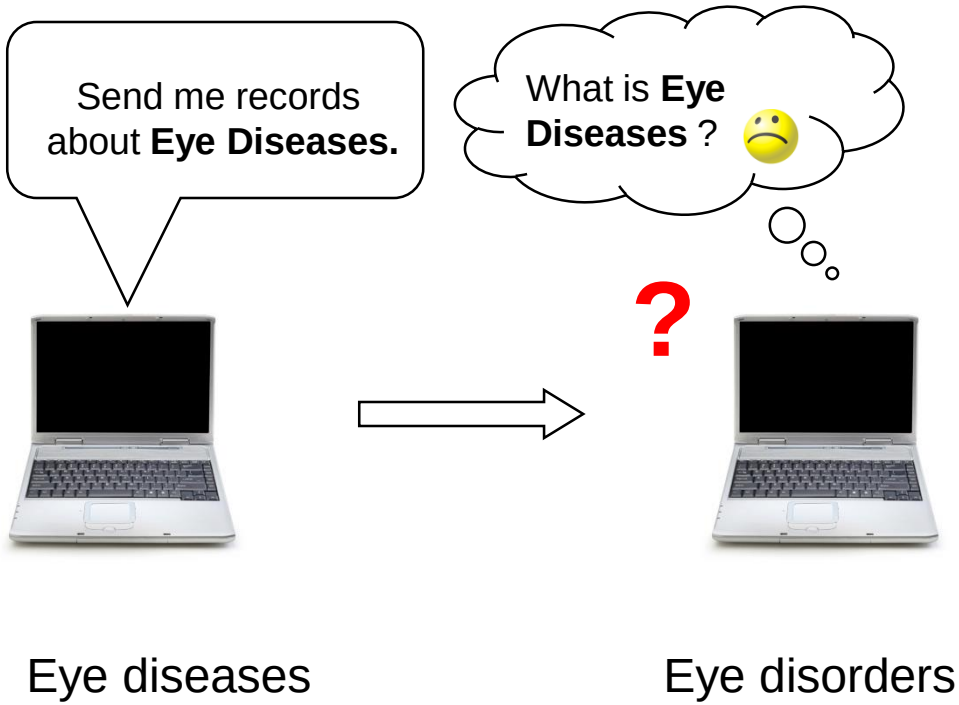


return 1617 articles

return 695 articles

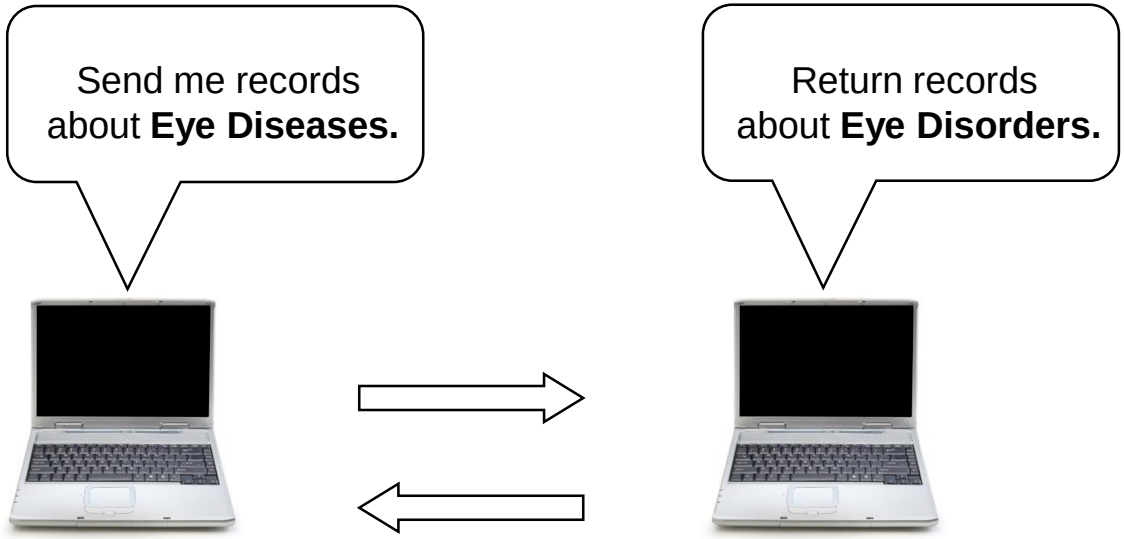
57% results are missed !

Integration of data sources



Integration of data sources

Eye diseases = Eye disorders



Components

- concepts

  - represent a set or class of entities in a domain

    Football-player

  - organized in taxonomies
    (hierarchies based on e.g. *is-a* or *is-part-of*)

    Football-player is-a *Athlete*

- instances

  Eden Hazard

  - often not represented in an ontology
    (instantiated ontology)

Components

- relations

  R: C1 x C2 x … x Cn

  hasMember (between Sports-Team and Athlete)

  hasMember (between Real Madrid and Eden Hazard)

Components

- axioms

  ‘facts that are always true’

The members of a football team are always football players.

Different kinds of ontologies

- Controlled vocabularies
  - Concepts
- Taxonomies
  - Concepts, is-a
- Thesauri
  - Concepts, predefined relations
- Data models (e.g. EER, UML)
  - Concepts, relations, axioms
- Logics
  - Concepts, relations, axioms

(Description) Logics

- Formal languages
- Reasoning services

- OWL-DL / OWL2 are based on description logics

$\mathcal{AL}$

R atomic role, A atomic concept

$C,D \rightarrow A$ | (atomic concept)

T | (universal concept, top) *owl:thing*

$\bot$ | (bottom concept) *owl:nothing*

$\neg A$ | (atomic negation) *owl:complementOf*

$C \cap D$ | (conjunction) *owl:intersectionOf*

$\forall R.C$ | (value restriction) *owl:allValuesFrom*

$\exists R.T$ (limited existential quantification) *owl:someValuesFrom*

$\mathcal{AL}[\mathcal{X}]$

$\mathcal{C}$ $\neg$C (concept negation) *owl:complementOf*

$\mathcal{U}$ C U D (disjunction) *owl:unionOf*

$\mathcal{E}$ $\exists$R.C (existential quantification)

owl:someValuesFrom

$\mathcal{N}$ ≥ n R, ≤ n R (number restriction)

owl:maxCardinality, owl:minCardinality

$\mathcal{Q}$ ≥ n R.C, ≤ n R.C (qualified number restriction)

owl:maxQualifiedCardinality,owl:minQualifiedCardinality

$\mathcal{AL}$ (Semantics)

An interpretation $\mathcal{I}$ consists of a non-empty set $\Delta^{\mathcal{I}}$ (the domain of the interpretation) and an interpretation function $\cdot^{\mathcal{I}}$ which assigns to every atomic concept A a set $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ and to every atomic role R a binary relation $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$.

The interpretation function is extended to concept definitions using inductive definitions.

$\mathcal{AL}$ (Semantics)

universal concept: $T^{\mathcal{I}} = \Delta^{\mathcal{I}}$

bottom concept: $\bot^{\mathcal{I}} = \emptyset$

atomic negation: $(\neg A)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus A^{\mathcal{I}}$

conjunction: $(C \cap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$

value restriction: $(\forall R.C)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \forall b.(a,b) \in R^{\mathcal{I}} \rightarrow b \in C^{\mathcal{I}}\}$

limited existential quantification: $(\exists R.T)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \exists b.(a,b) \in R^{\mathcal{I}}\}$

$\mathcal{ALC}$ (Semantics)

$(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$

$(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$

$(\geq n\ R)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \#\{b \in \Delta^{\mathcal{I}} \mid (a,b) \in R^{\mathcal{I}}\} \geq n\}$

$(\leq n\ R)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \#\{b \in \Delta^{\mathcal{I}} \mid (a,b) \in R^{\mathcal{I}}\} \leq n\}$

$(\exists R.C)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \exists b \in \Delta^{\mathcal{I}} : (a,b) \in R^{\mathcal{I}} \wedge b \in C^{\mathcal{I}}\}$

Semantics

Individual i

$i^{\mathcal{I}} \in \Delta^{\mathcal{I}}$

Unique Name Assumption:

if $i_1 \neq i_2$ then $i_1^{\mathcal{I}} \neq i_2^{\mathcal{I}}$

Concepts and relations

Team

(Team)

Concept/class

¬Team

(**not** Team)

Relation/role/property

Team ∩ ≥ 10 hasMember

(Team **and at least** 10 members)

Team ∩ ≤ 10 hasMember

(Team **and at most** 10 members)

Concepts and relations

Team $\cap \forall$ hasMember.Football-player

(Team **and all** members **are** football players)

Team $\cap \exists$ hasMember.Football-player

(Team **and there is a** member **that is a** football player)

Axioms

$C \subseteq D$ $(R \subseteq S)$

rdfs:subClassOf / rdfs:subPropertyOf

Football-player $\subseteq$ Athlete

(**Every** football player **is a**n athlete)

$C = D$ $(R = S)$

owl:equivalentClass / owl:equivalentProperty

(disjoint C D)

owl:disjointWith

Axioms

C(a) *a rdf:type C*

Team(Real Madrid)

(Real Madrid **is an instance of** Team)

R(a,b) *a R b*

hasMember(Real Madrid , Eden Hazard)

(Real Madrid has member Eden Hazard)

Example

Teams have at least two members, while large teams have at least 10 members. Sports teams are teams which have only athletes as members. A football team is a team which has at least 11 members and all the members are football players. Football players are athletes. Real Madrid is a football team that has Eden Hazard as a member.

Example

Team $\subseteq \geq 2$ hasMember

Large-Team = Team $\cap \geq 10$ hasMember

Sports-team = Team $\cap \forall$ hasMember.Athlete

Football-Team = Team $\cap \geq 11$ hasMember

$\cap \forall$ hasMember.Football-player

Football-player $\subseteq$ Athlete

Football-Team(Real Madrid)

hasMember(Real Madrid,Eden Hazard)

Example

Every team has at least 2 members
Every large team is a team and has at least 10 members
Every sports team is a team and has only athletes as members
Every football team is a team and has at least 11 members
and has only football players as members

Example

Every team has at least 2 members
Every large team is a team and has at least 10 members
Every sports team is a team and has only athletes as members
Every football team is a team and has at least 11 members
and has only football players as members

Reasoning:

Every football team is a large team
Every football team is a sports team

Example

Real Madrid is an instance of football team

Real Madrid has member Eden Hazard

Example

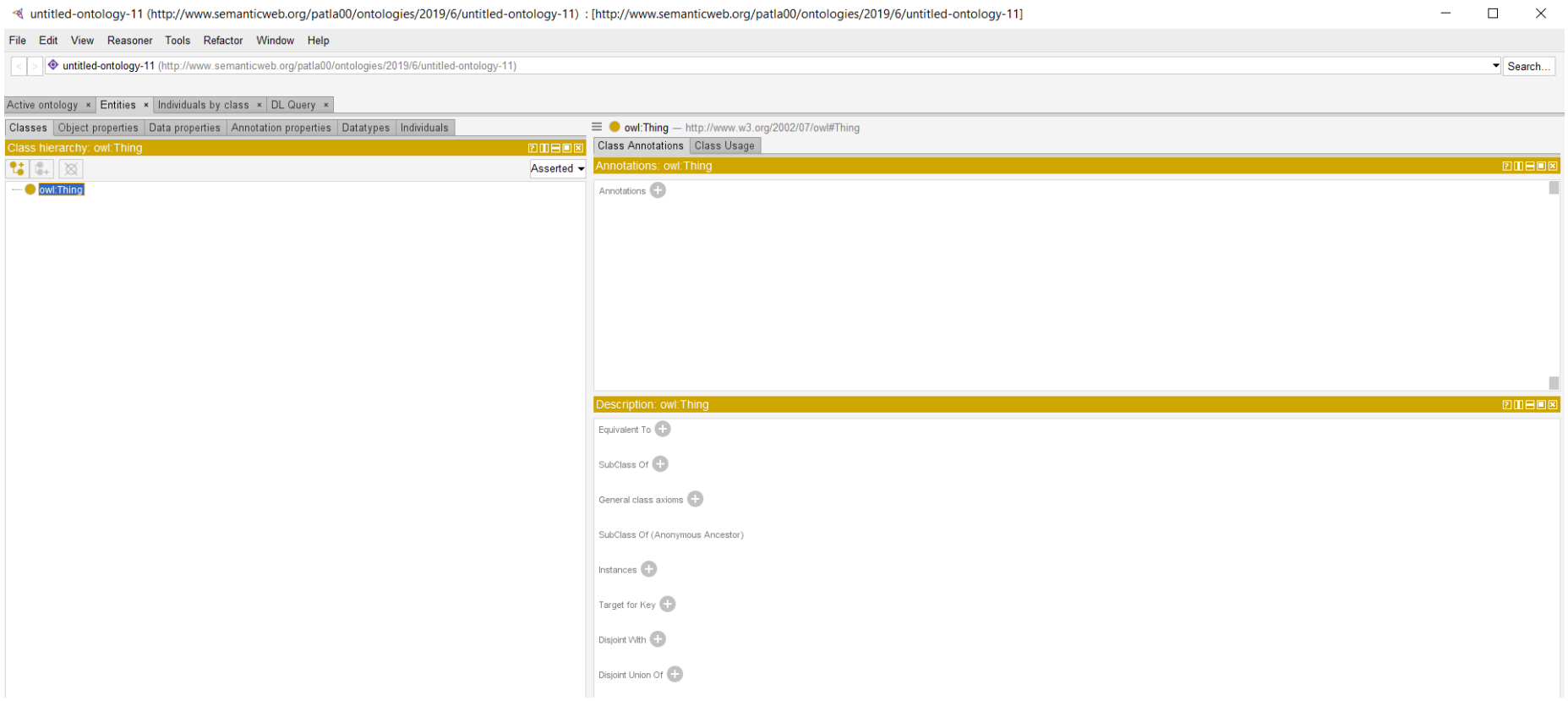
Reasoning:

Real Madrid is an instance of football team
Real Madrid is an instance of large team
Real Madrid is an instance of team
Real Madrid is an instance of sports team
Real Madrid has at least 11 members
All members in Real Madrid are football players
All members in Real Madrid are athletes

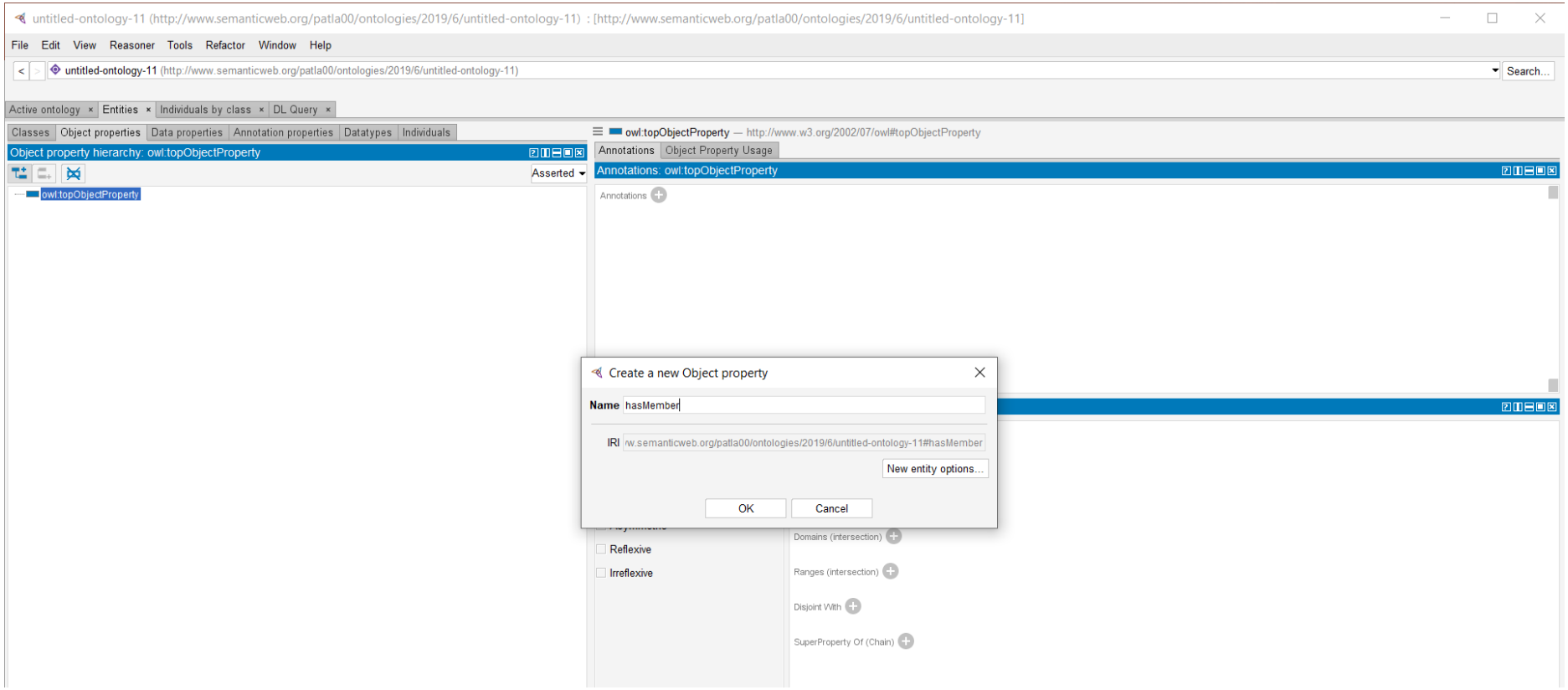
Real Madrid has member Eden Hazard

Eden Hazard is an instance of football player
Eden Hazard is an instance of athlete

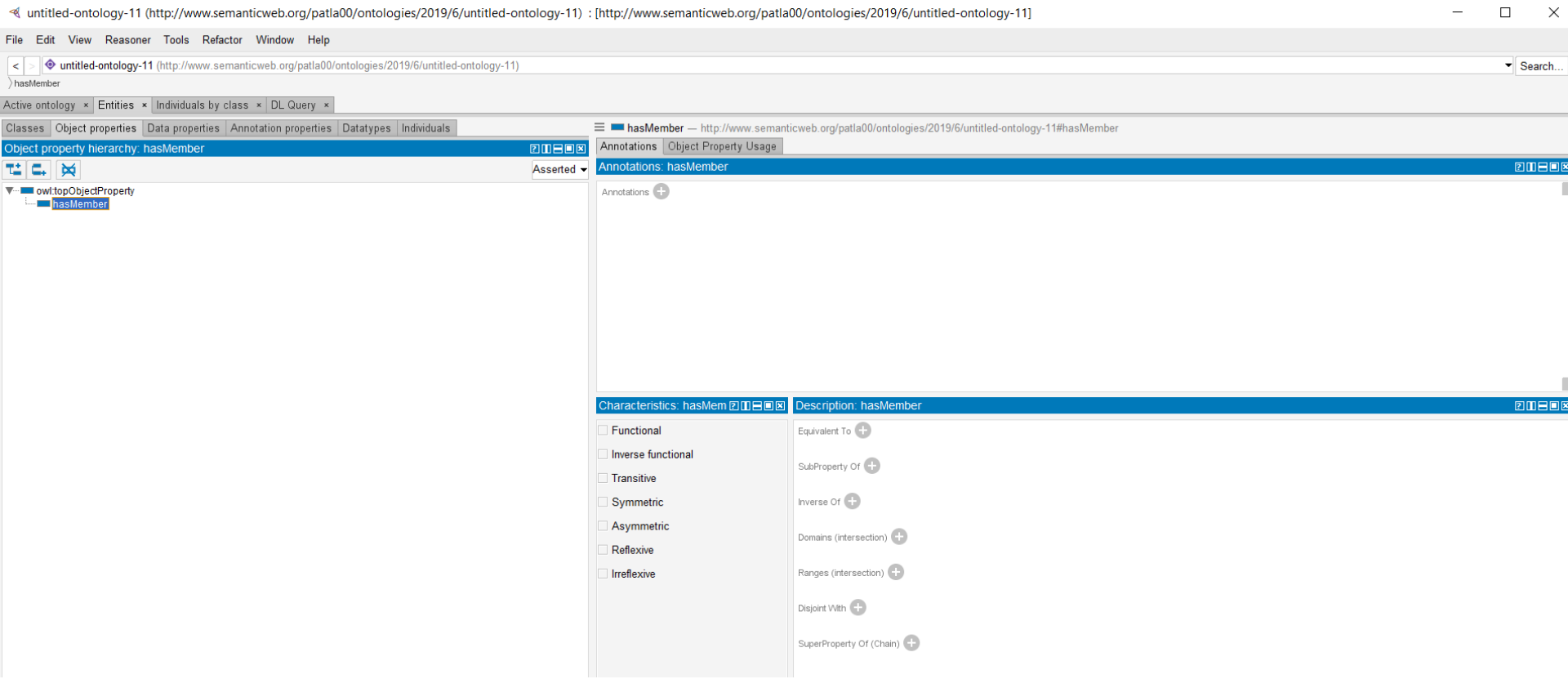
Start


untitled-ontology-11 (http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-11) : [http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-11]
File Edit View Reasoner Tools Refactor Window Help
untitled-ontology-11 (http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-11)
Search...
Active ontology
Entities
Individuals by class
DL Query
Classes
Object properties
Data properties
Annotation properties
Datatypes
Individuals
Class hierarchy: owl:Thing
Asserted
owl:Thing
owl:Thing — http://www.w3.org/2002/07/owl#Thing
Class Annotations
Class Usage
Annotations: owl:Thing
Annotations
Description: owl:Thing
Equivalent To
SubClass Of
General class axioms
SubClass Of (Anonymous Ancestor)
Instances
Target for Key
Disjoint With
Disjoint Union Of

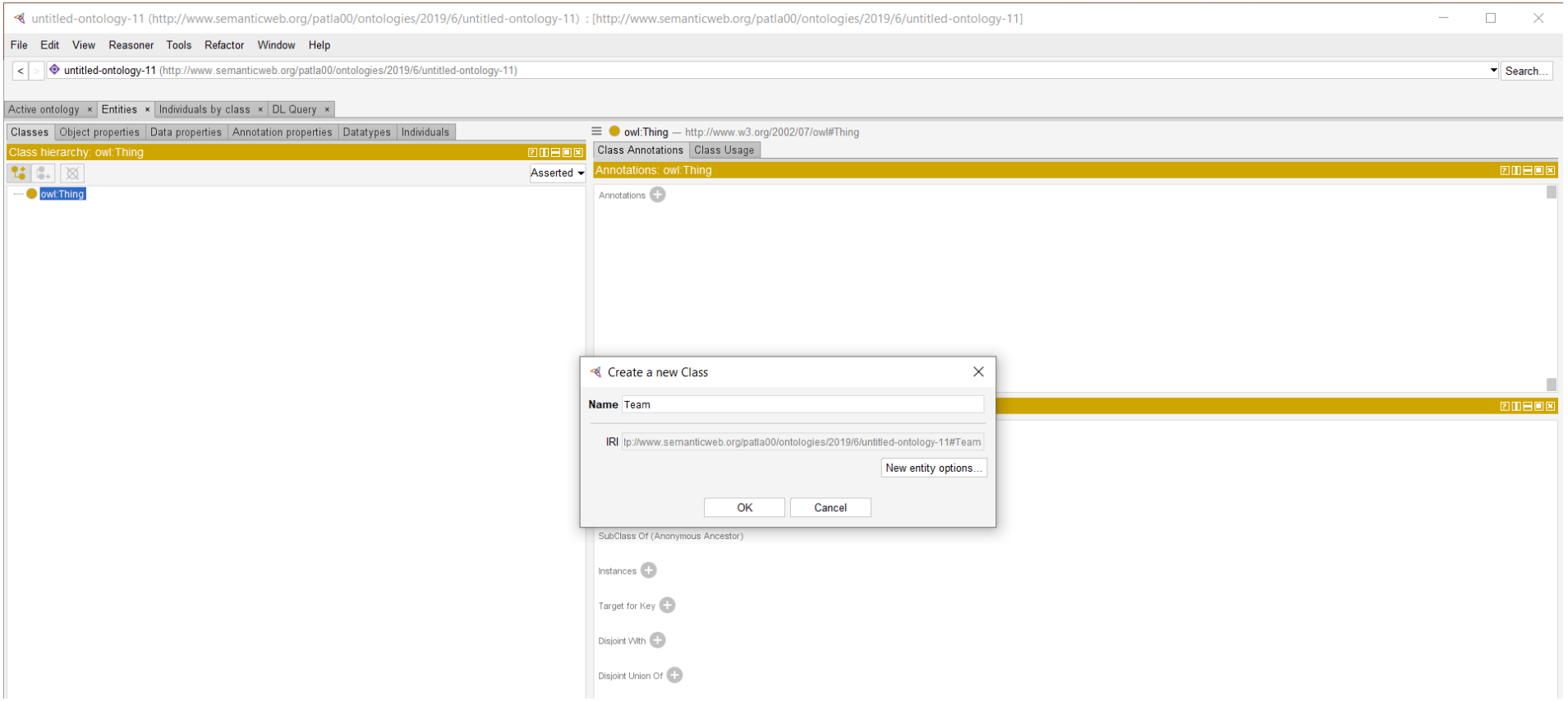
Create relation *hasMember*



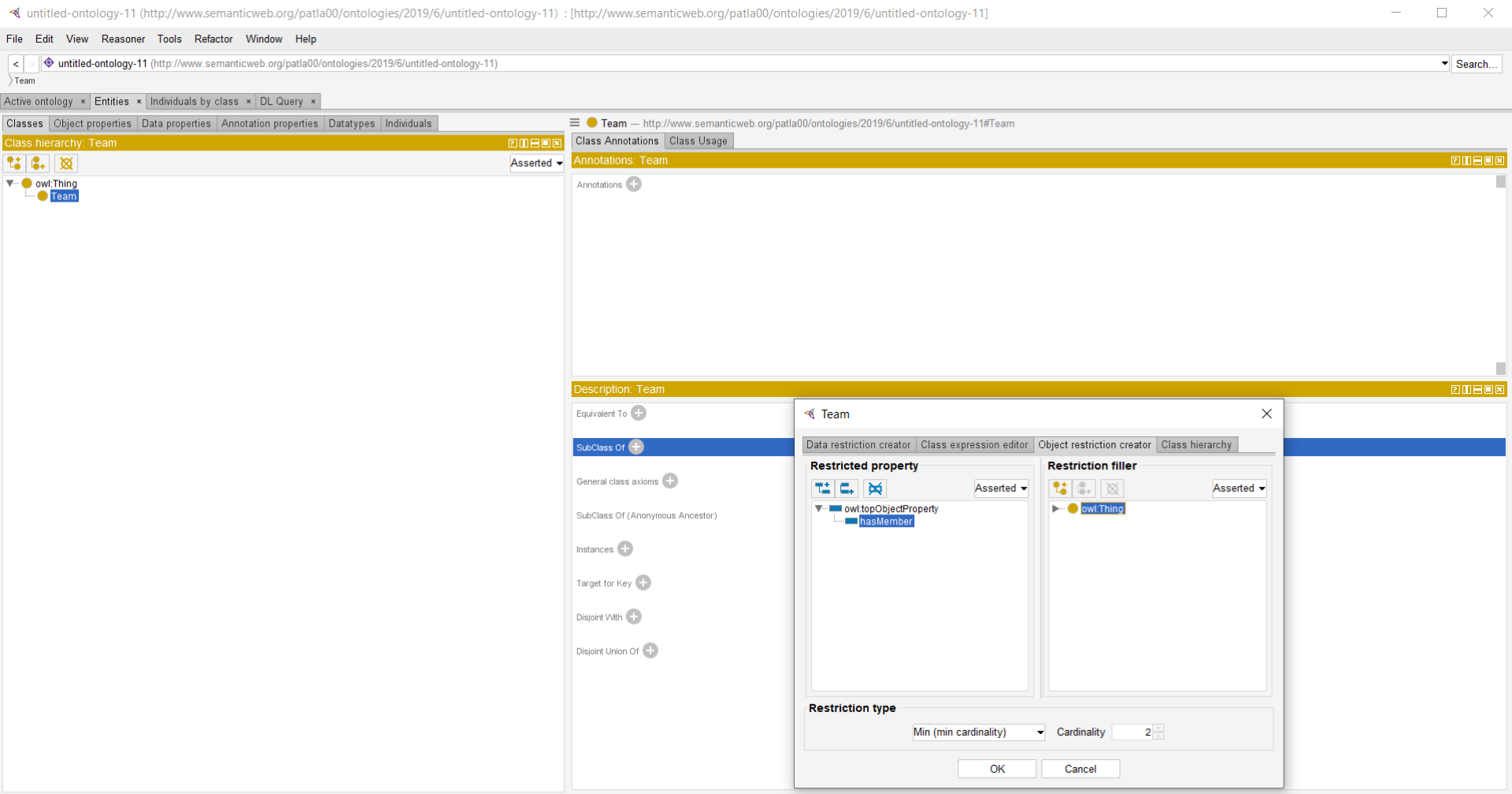
Create relation *hasMember*



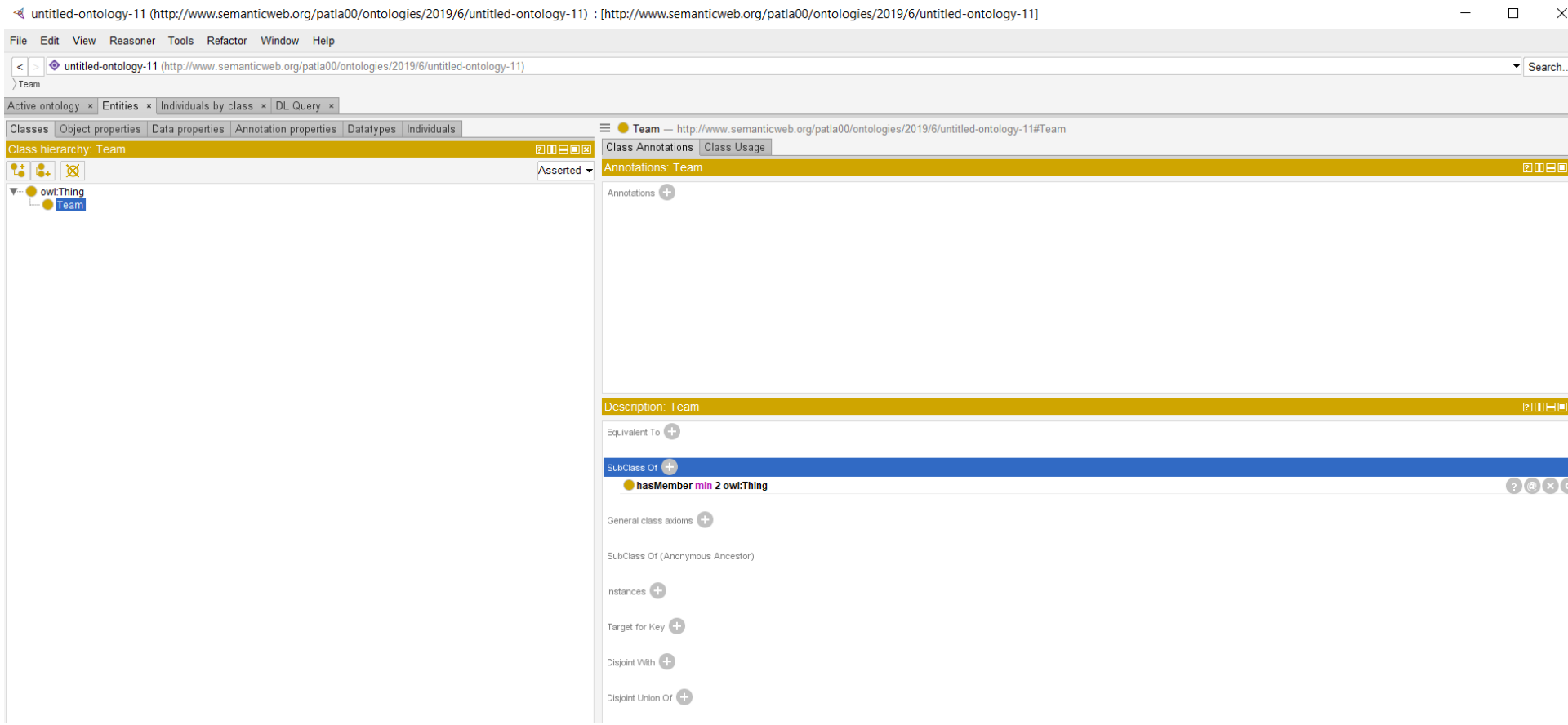
Create concept *Team*



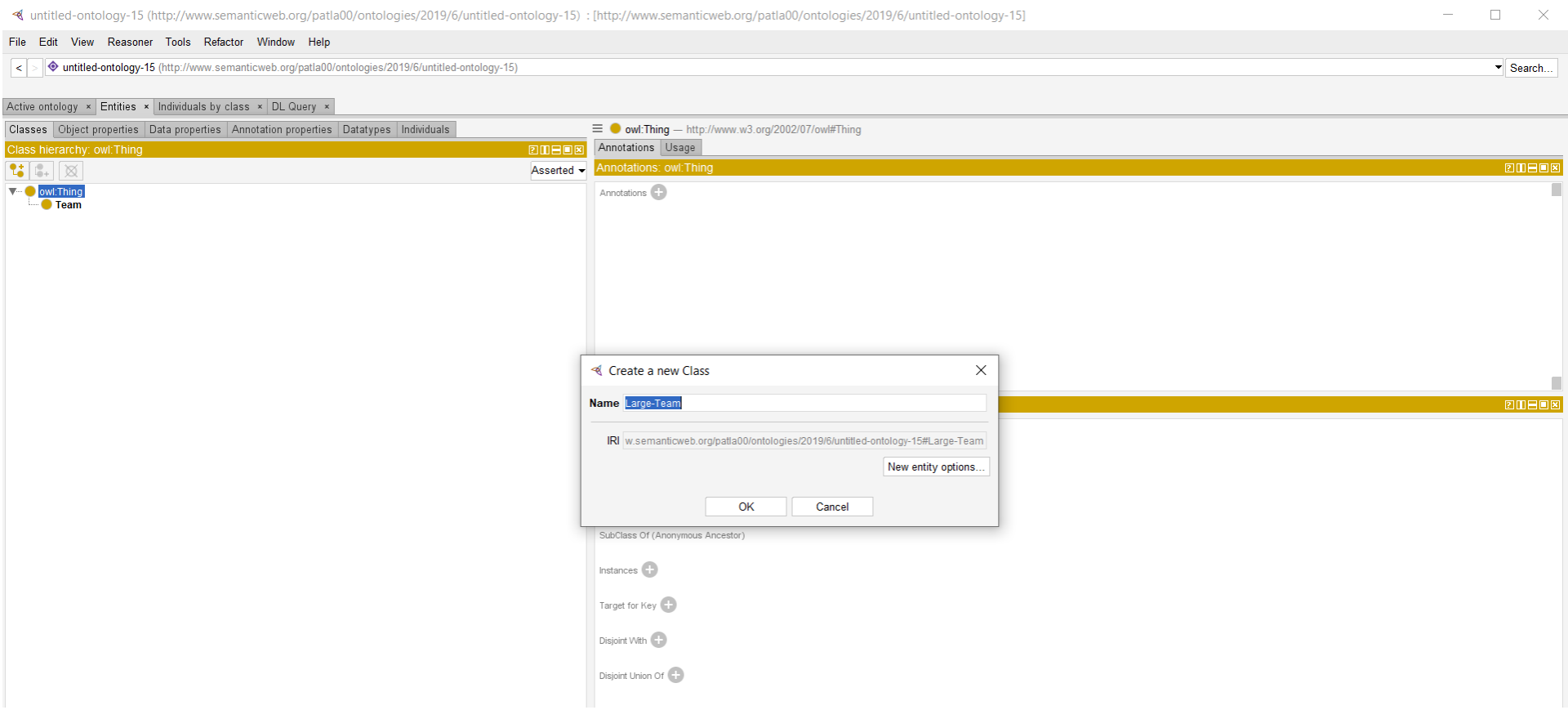
Edit concept *Team* using object restriction creator



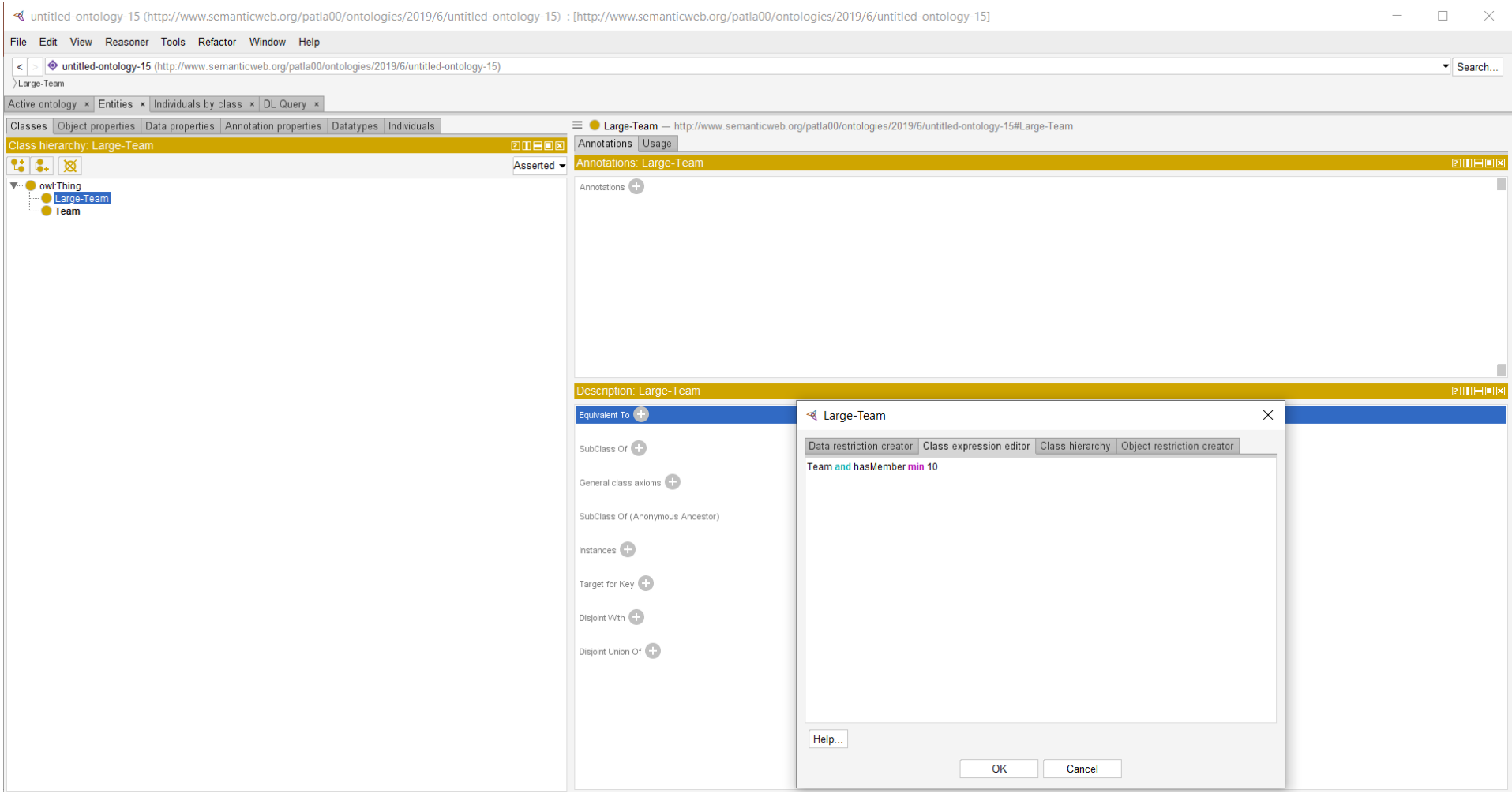
Edit concept *Team* using object restriction creator



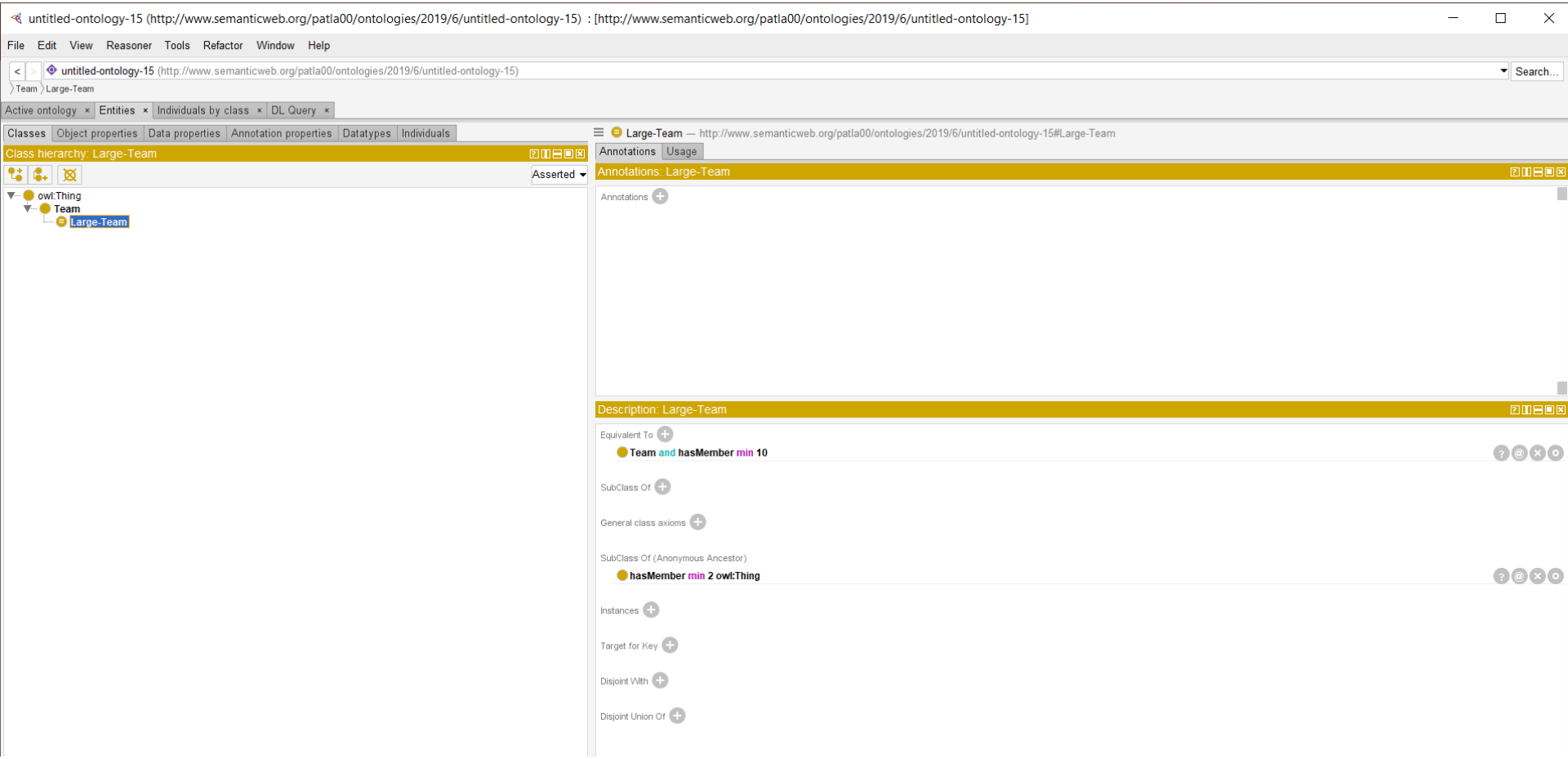
Create concept *Large-Team*



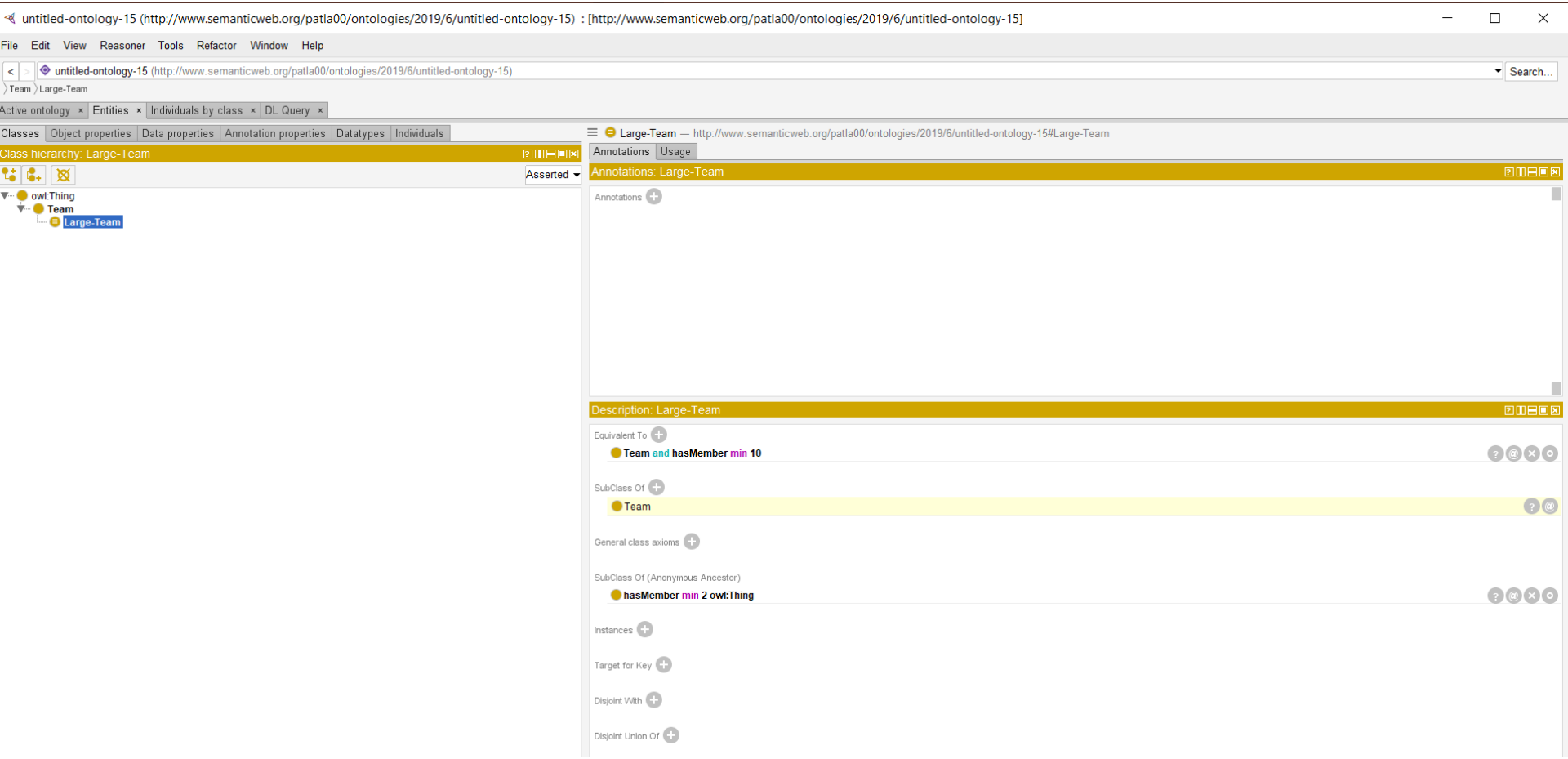
Edit concept *Large-Team* using class expression editor



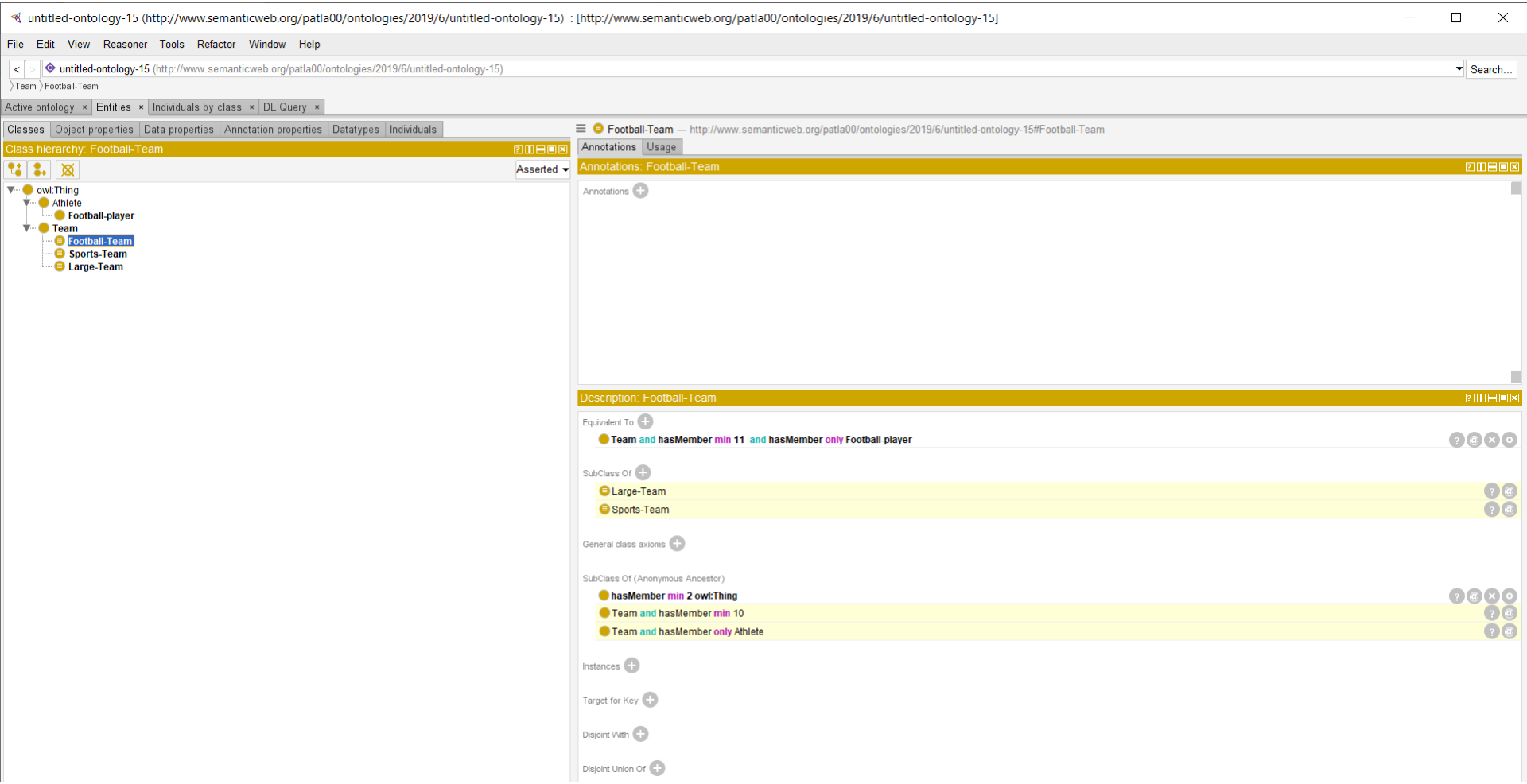
Edit concept *Large-Team* using class expression editor



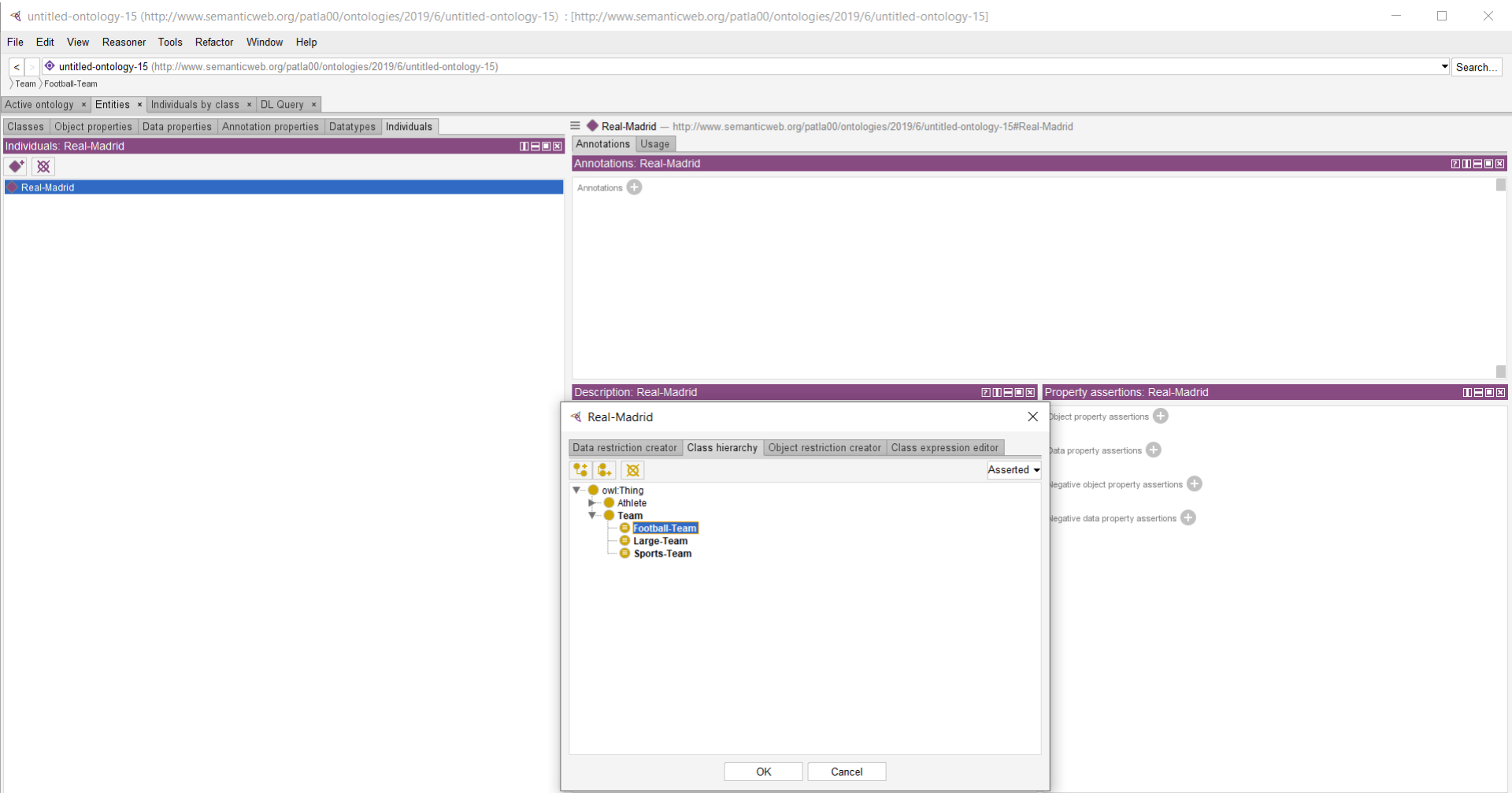
Reasoning


untitled-ontology-15 (http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-15) : [http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-15]
File Edit View Reasoner Tools Refactor Window Help
untitled-ontology-15 (http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-15)
Search...
Team Large-Team
Active ontology Entities Individuals by class DL Query
Classes Object properties Data properties Annotation properties Datatypes Individuals
Class hierarchy: Large-Team
Asserted
owl:Thing
Team
Large-Team
Large-Team — http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-15#Large-Team
Annotations Usage
Annotations: Large-Team
Annotations
Description: Large-Team
Equivalent To
Team and hasMember min 10
SubClass Of
Team
General class axioms
SubClass Of (Anonymous Ancestor)
hasMember min 2 owl:Thing
Instances
Target for Key
Disjoint With
Disjoint Union Of

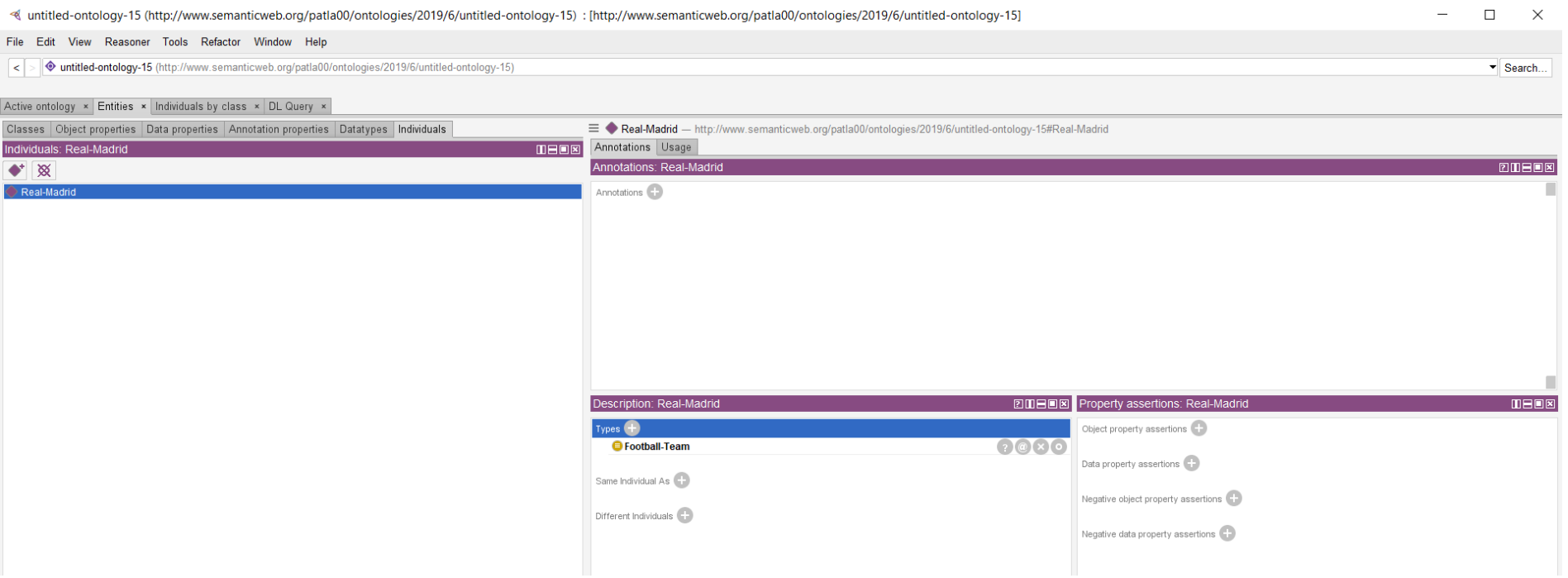
After adding all other concepts and reasoning



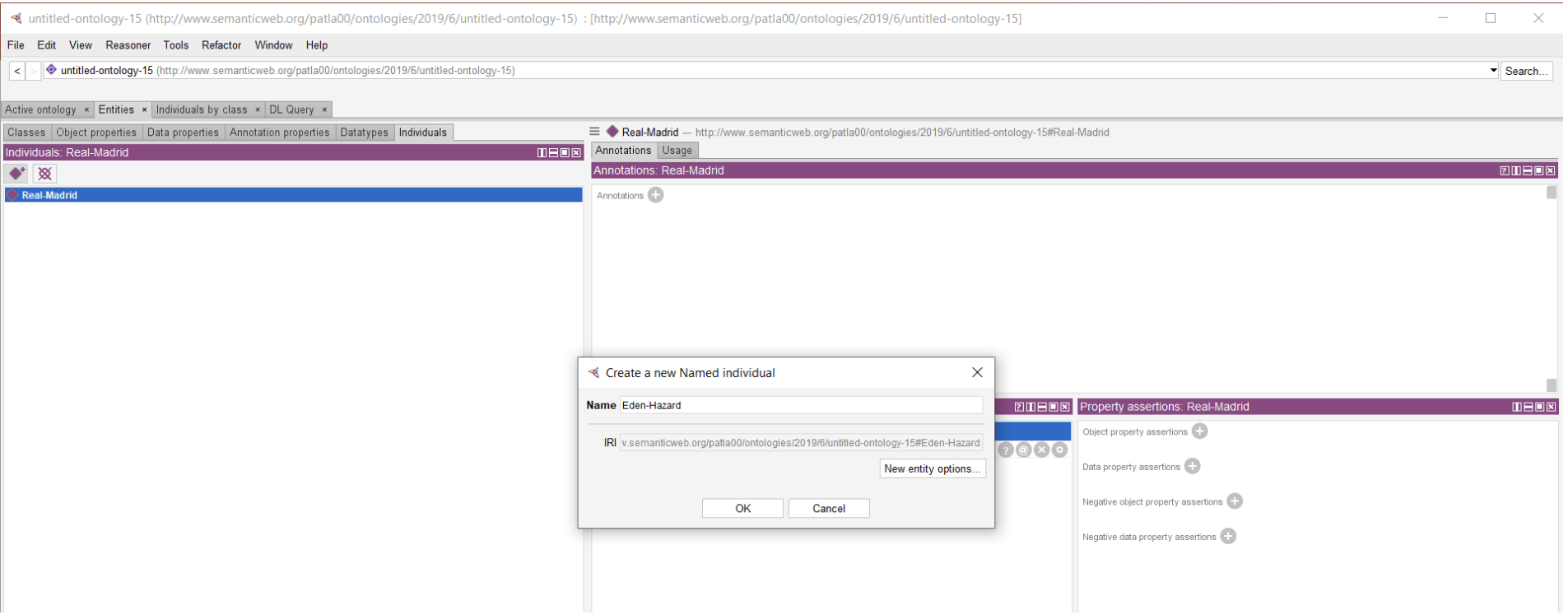
Creating individual *Real-Madrid* and asserting it is a *Football-Team*



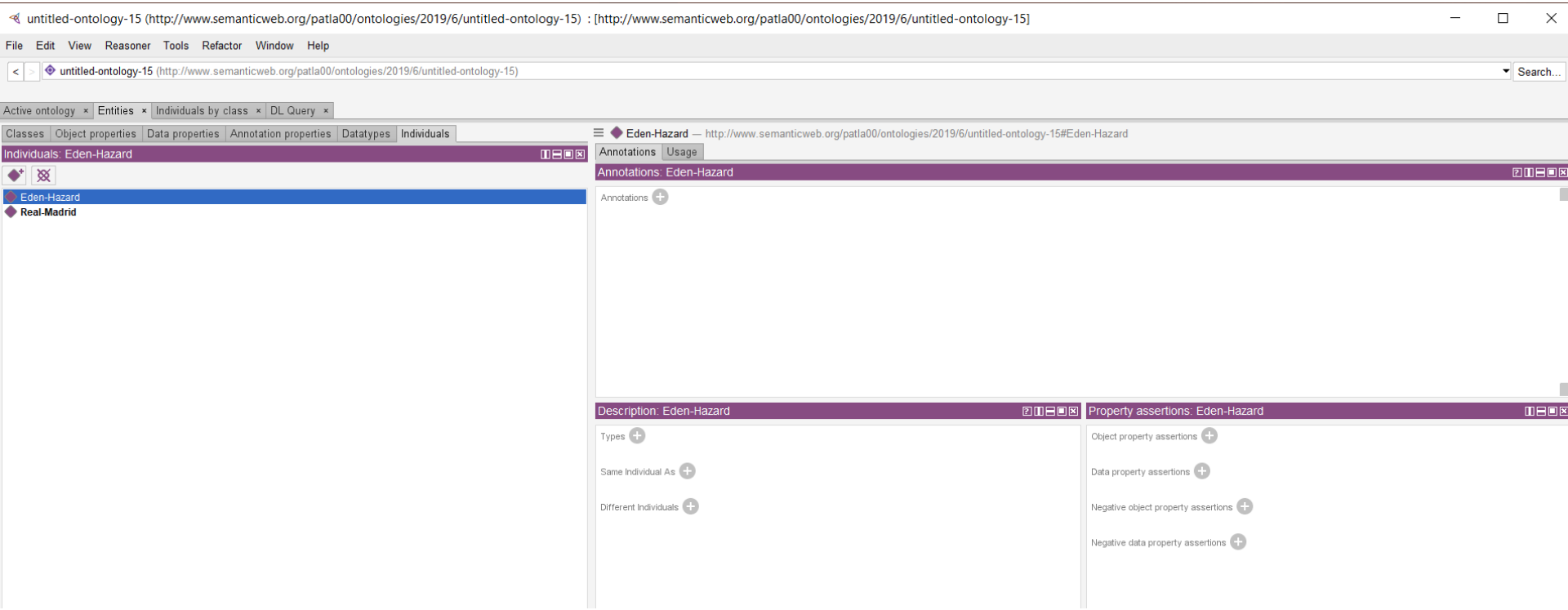
Creating individual *Real-Madrid* and asserting it is a *Football-Team*



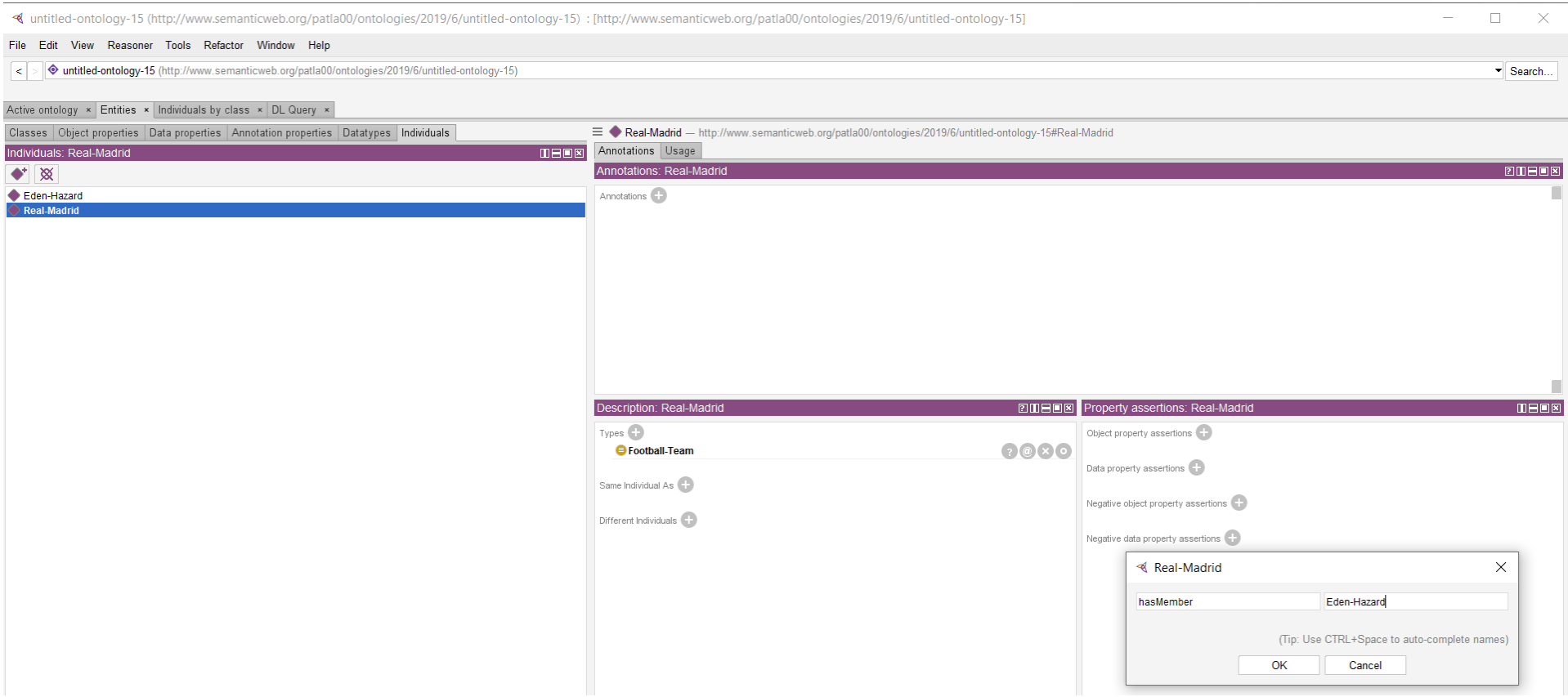
Creating individual *Eden-Hazard*



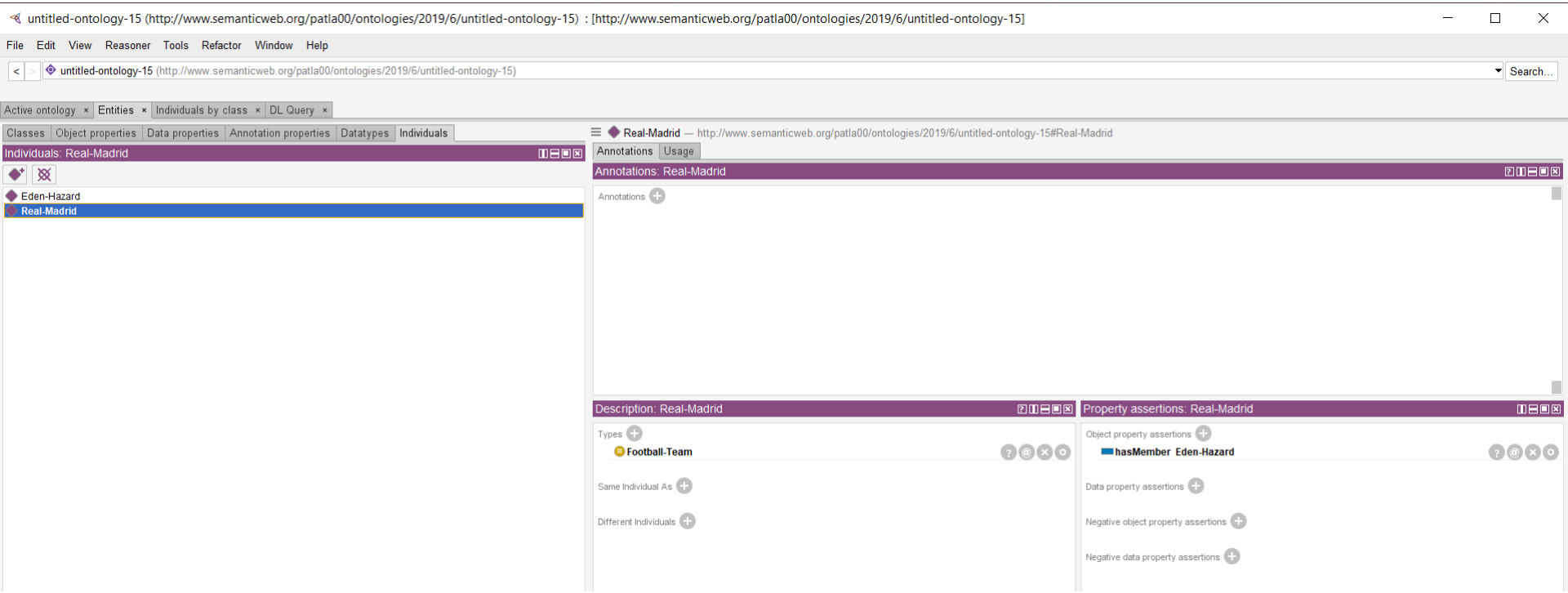
Creating individual *Eden-Hazard*



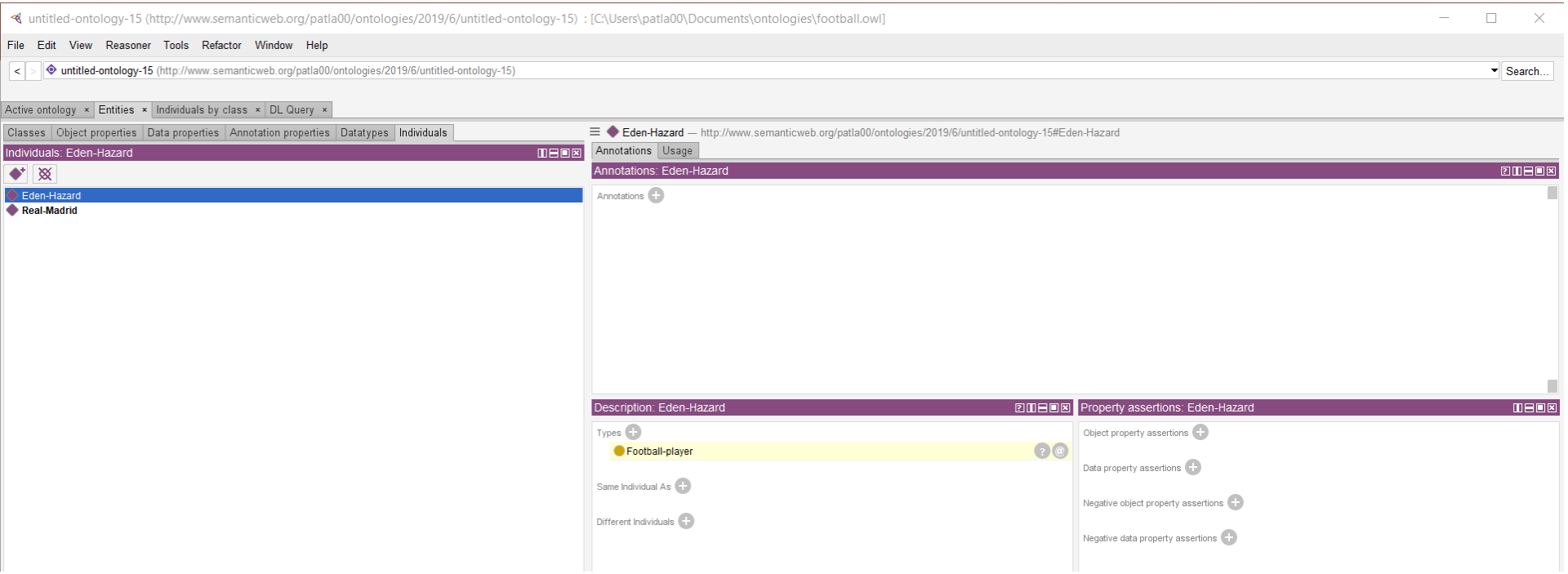
Asserting the relation *hasMember* between *Real-Madrid* and *Eden-Hazard*



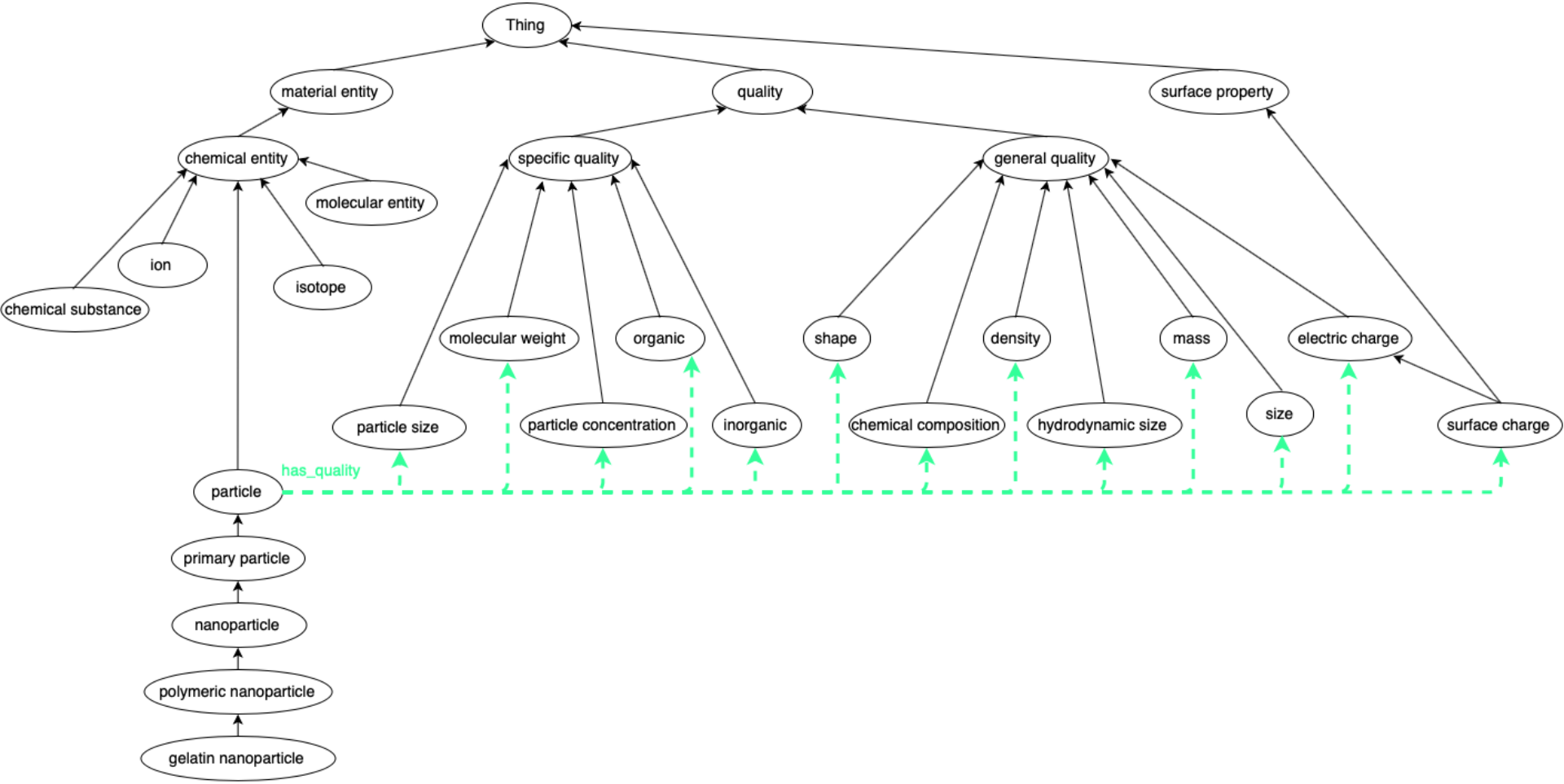
Asserting the relation *hasMember* between *Real-Madrid* and *Eden-Hazard*

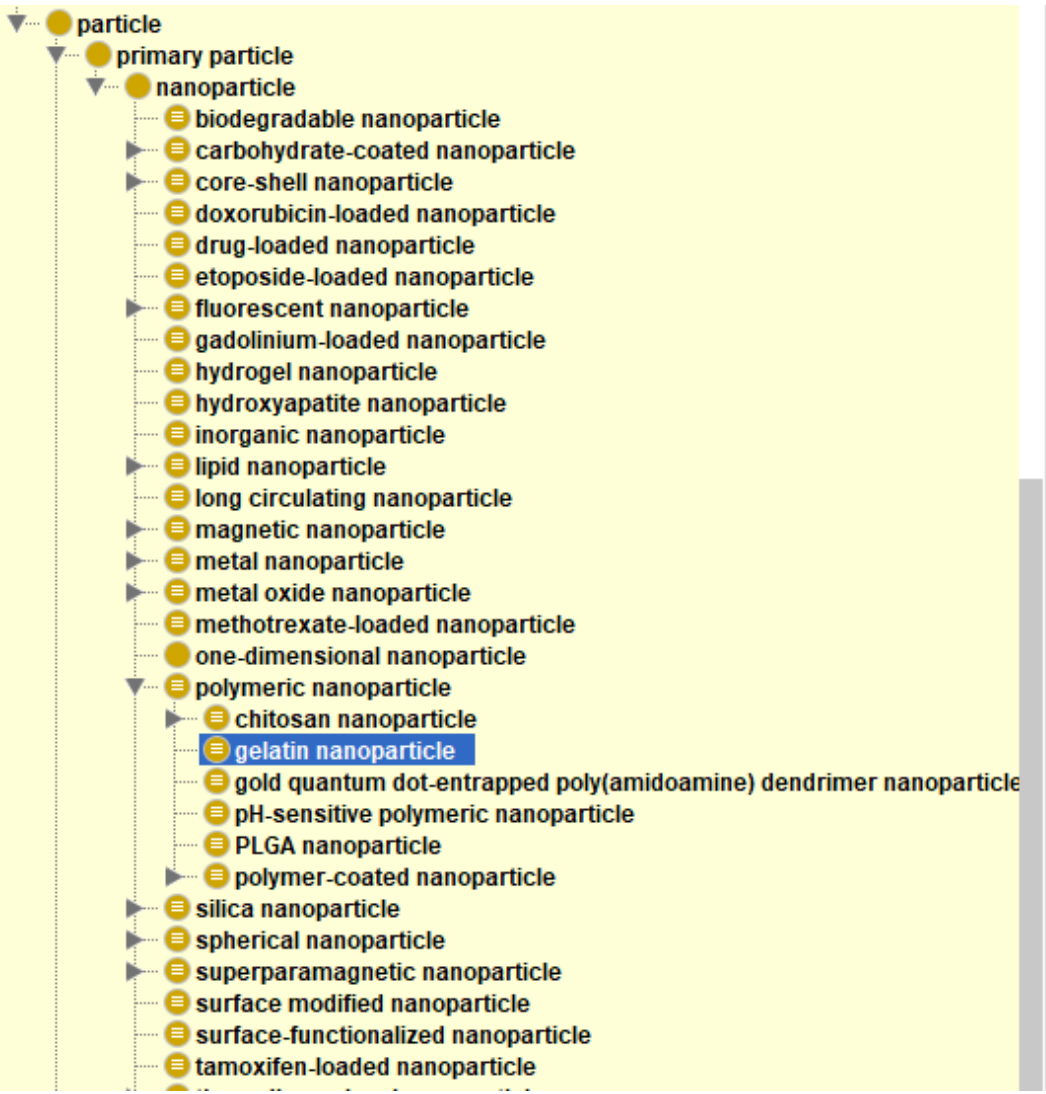


Reasoning


untitled-ontology-15 (http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-15) : [C:\Users\patla00\Documents\ontologies\football.owl]
File Edit View Reasoner Tools Refactor Window Help
untitled-ontology-15 (http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-15)
Search...
Active ontology
Entities
Individuals by class
DL Query
Classes Object properties Data properties Annotation properties Datatypes Individuals
Individuals: Eden-Hazard
Eden-Hazard
Real-Madrid
Eden-Hazard — http://www.semanticweb.org/patla00/ontologies/2019/6/untitled-ontology-15#Eden-Hazard
Annotations Usage
Annotations: Eden-Hazard
Annotations
Description: Eden-Hazard
Types
Football-player
Same Individual As
Different Individuals
Property assertions: Eden-Hazard
Object property assertions
Data property assertions
Negative object property assertions
Negative data property assertions

Nanoparticle Ontology





Description: gelatin nanoparticle

Equivalent To

- nanoparticle and (has_component_part some gelatin)

SubClass Of

- 'polymeric nanoparticle'

General class axioms

SubClass Of (Anonymous Ancestor)

- has_part some 'surface of nanoparticle'
- has_quality some 'particle size'
- has_quality some shape
- has_part some (coat or core or shell)
- has_quality some 'particle concentration'
- has_quality some 'chemical composition'
- has_quality some density
- has_quality some mass
- has_quality some ('electric charge' or 'surface charge')
- has_quality some size
- has_quality some (inorganic or organic)
- has_quality some 'hydrodynamic size'
- has_quality some 'molecular weight'
- nanoparticle and (has_component_part some polymer)

```
<!-- http://purl.bioontology.org/ontology/npo#NPO_1555 -->
<owl:Class rdf:about="#NPO_1555">
    <rdfs:label rdf:datatype="&xsd;string">gelatin nanoparticle</rdfs:label>
    <owl:equivalentClass>
        <owl:Class>
            <owl:intersectionOf rdf:parseType="Collection">
                <rdf:Description rdf:about="#NPO_707"/> <!-- nanoparticle -->
                <owl:Restriction>
                    <owl:onProperty rdf:resource="#has_component_part"/>
                    <owl:someValuesFrom rdf:resource="#NPO_1554"/> <!-- gelatin -->
                </owl:Restriction>
            </owl:intersectionOf>
        </owl:Class>
    </owl:equivalentClass>
    <rdfs:subClassOf rdf:resource="#NPO_1375"/> <!-- polymeric nanoparticle -->
</owl:Class>
```

1

```
<!-- http://purl.bioontology.org/ontology/npo#NPO_1375 -->
<owl:Class rdf:about="#NPO_1375">
    <rdfs:label rdf:datatype="&xsd;string">polymeric nanoparticle</rdfs:label>
    <owl:equivalentClass>
        <owl:Class>
            <owl:intersectionOf rdf:parseType="Collection">
                <rdf:Description rdf:about="#NPO_707"/> <!-- nanoparticle -->
                <owl:Restriction>
                    <owl:onProperty rdf:resource="#has_component_part"/>
                    <owl:someValuesFrom rdf:resource="&CHEBI;CHEBI_33839"/> <!-- polymer -->
                </owl:Restriction>
            </owl:intersectionOf>
        </owl:Class>
    </owl:equivalentClass>
    <rdfs:subClassOf rdf:resource="#NPO_707"/>
</owl:Class>
```

2

Defining ontologies is not so easy ...

The Celestial Emporium of Benevolent Knowledge, Borges

"On those remote pages it is written that animals are divided into:

a. those that belong to the Emperor

b. embalmed ones

c. those that are trained

d. suckling pigs

e. mermaids

f. fabulous ones

g. stray dogs

h. those that are included in this classification

i. those that tremble as if they were mad

j. innumerable ones

k. those drawn with a very fine camel's hair brush

l. others

m. those that have just broken a flower vase

n. those that resemble flies from a distance"

Slide from talk by C. Goble

Defining ontologies is not so easy ...

Dyirbal classification of objects in the universe

- Bayi: men, kangaroos, possums, bats, most snakes, most fishes, some birds, most insects, the moon, storms, rainbows, boomerangs, some spears, etc.
- Balan: women, anything connected with water or fire, bandicoots, dogs, platypus, echidna, some snakes, some fishes, most birds, fireflies, scorpions, crickets, the stars, shields, some spears, some trees, etc.
- Balam: all edible fruit and the plants that bear them, tubers, ferns, honey, cigarettes, wine, cake.
- Bala: parts of the body, meat, bees, wind, yamsticks, some spears, most trees, grass, mud, stones, noises, language, etc.

Slide from talk by C. Goble

Difficulties

- Knowledge engineering
  - Is-a vs part-of
  - Concept vs individual
  - Synonyms as concepts, unconnected terms, cycles in the hierarchy, domain and range problems,missing disjointness … http://oops.linkeddata.es/catalogue.jsp

→ Domain expert + knowledge engineer

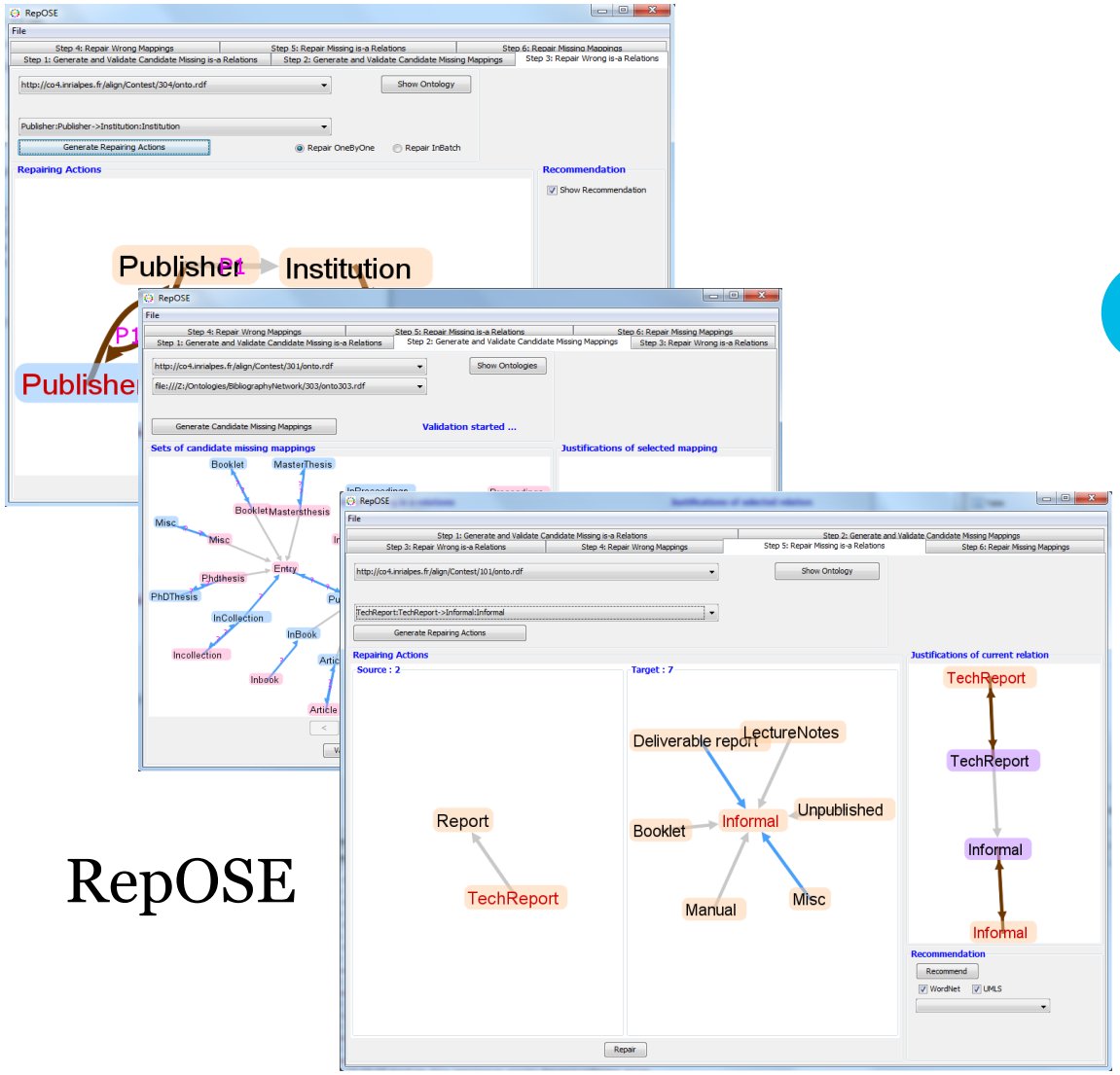
Difficulties

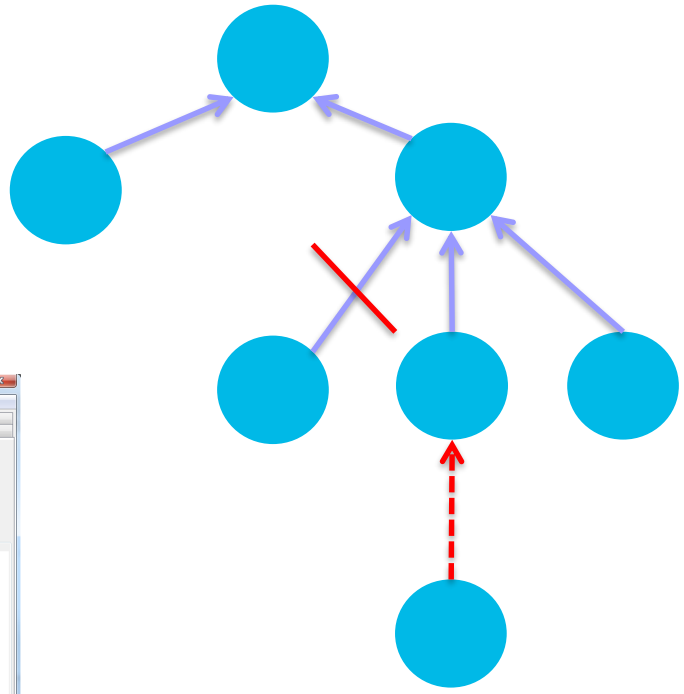
- Completeness
  - Missing information
- Correctness
  - Incoherence
  - Inconsistency

Ontology tools

- Ontology development tools
- Ontology completion and debugging tools
- Ontology alignment tools
- Ontology evaluation tools
- Ontology-based annotation tools
- Ontology storage and querying tools
- Ontology learning tools

Debugging and Completing Ontologies



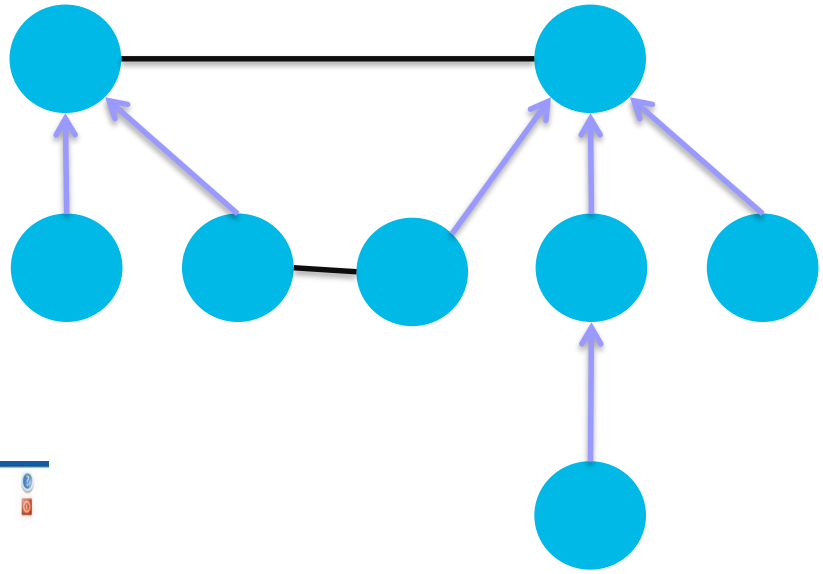


RepOSE

Ontology tools

- Ontology development tools
- Ontology completion and debugging tools
- Ontology alignment tools
- Ontology evaluation tools
- Ontology-based annotation tools
- Ontology storage and querying tools
- Ontology learning tools

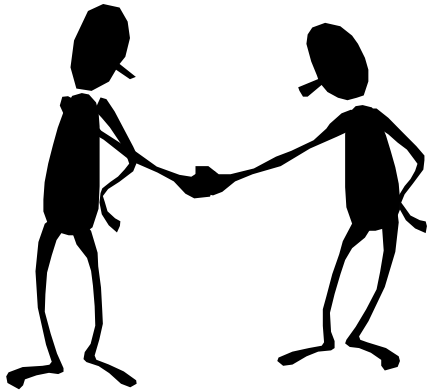
Aligning Ontologies

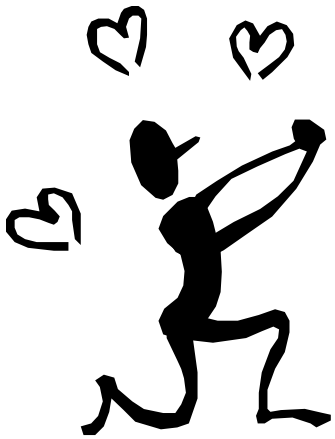


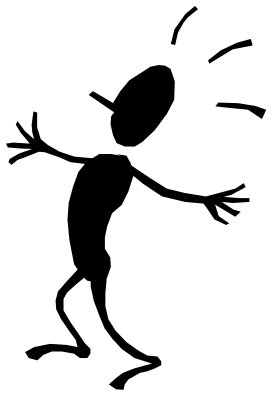
SAMBO

Ontology tools

- Ontology development tools
- Ontology completion and debugging tools
- Ontology alignment tools
- Ontology evaluation tools
- Ontology-based annotation tools
- Ontology storage and querying tools
- Ontology learning tools







Further reading

Ontologies

Staab and Studer, *Handbook on Ontologies*, 2nd ed, Springer, 2009.

Lambrix, Strömbäck, Tan, Information integration in bioinformatics with ontologies and standards, in Bry, Maluszynski (eds), *Semantic Techniques for the Web: The REWERSE perspective*, chapter 8, 343-376, 2009.

Further reading

Description logics

Baader, Calvanese, McGuinness, Nardi, Patel-Schneider. *The Description Logic Handbook*, Cambridge University Press, 2003.

Donini, Lenzerini, Nardi, Schaerf, Reasoning in description logics. *Principles of knowledge representation*, CSLI publications, pp 191-236. 1996.

dl.kr.org

www.w3.org

Further reading

OWL

https://www.w3.org/TR/owl-guide/

OWL2

https://www.w3.org/TR/owl2-primer/

Further reading

Tools

- Development
  - Protégé https://protege.stanford.edu/
- Pitfalls
  - OOPS http://oops.linkeddata.es/
- Debugging and completion
  - Repose https://www.ida.liu.se/~patla00/research/RepOSE/
- Ontology alignment
  - http://www.ontologymatching.org/
  - http://oaei.ontologymatching.org/