Multilingualism in the Web of Data

Elena Montiel-Ponsoda

Ontology Engineering Group, Universidad Politécnica de Madrid

Monnet is supported by the European Union under Grant No. 248458

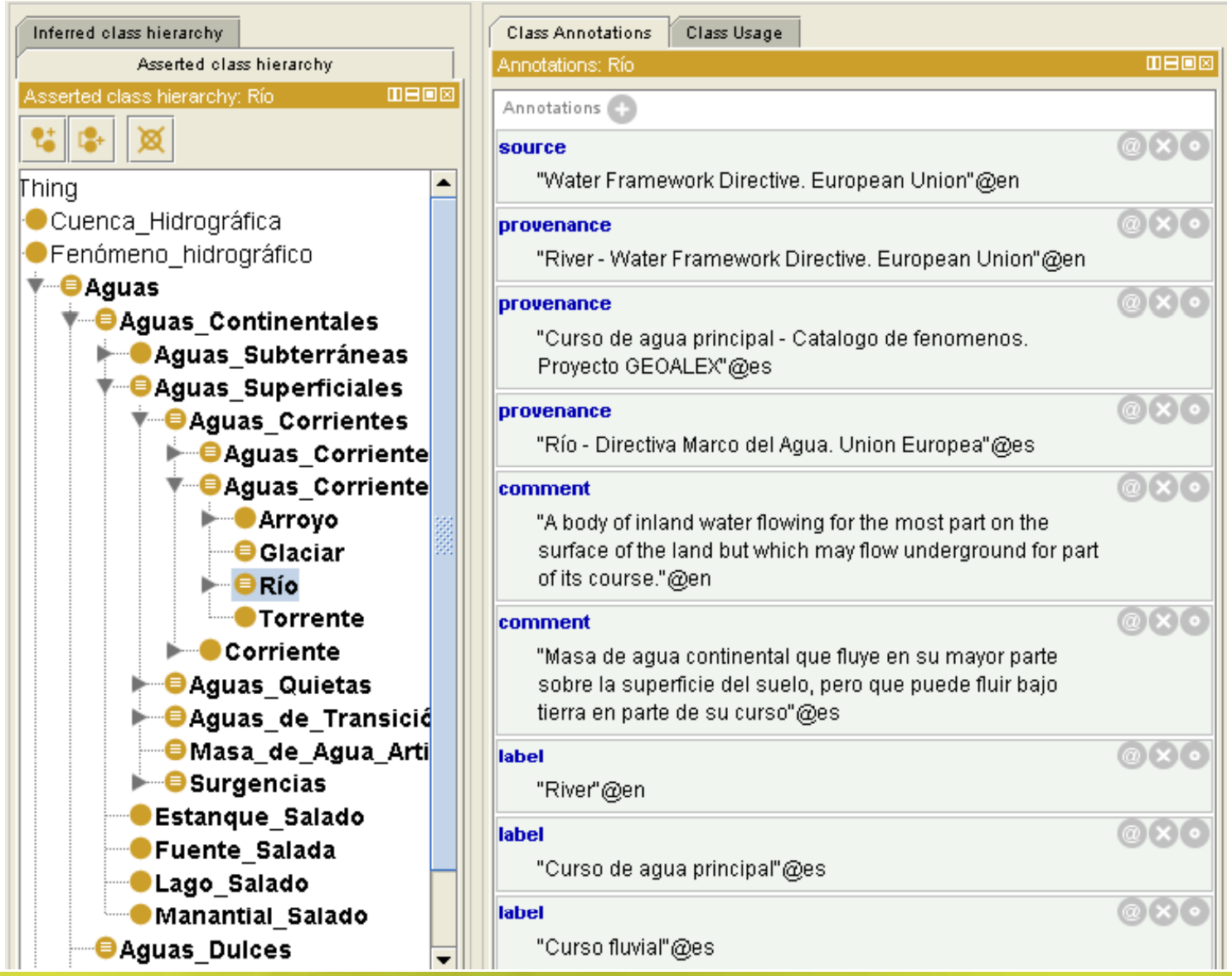
Outline

1. Available representation models
   - RDF(S), OWL
   - SKOS & SKOS-XL
     - Is this enough? What is the evolution?
2. Latest trends
   - *lemon*: representing **rich linguistic descriptions** (lexical, terminological, etc.) **relative** to ontologies and linked data
3. Ontology localization

RDF(S), OWL

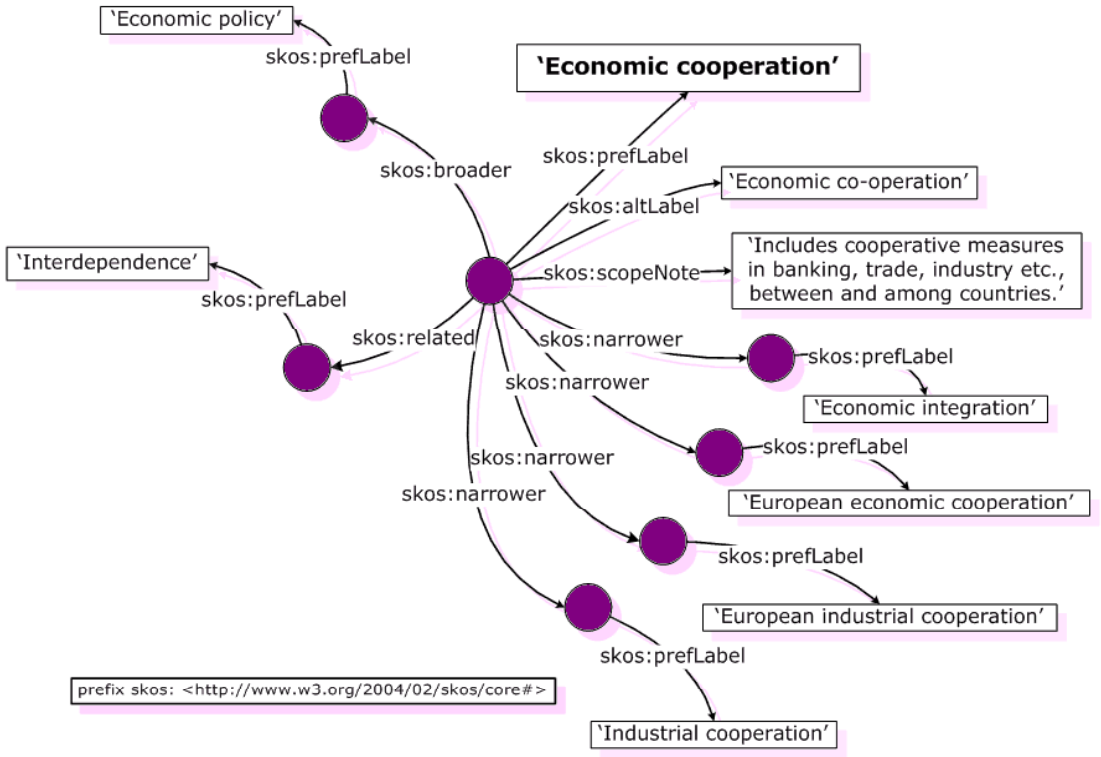
ifrs:FinancialAssets —rdfs:label→ “financial assets”@en

RDF(S), OWL Example



SKOS

- SKOS—Simple Knowledge Organization System— model for expressing the basic structure and content of concept schemes such as **thesauri**, **classification schemes**, **subject heading lists**, **taxonomies**, **folksonomies**, and other similar types of controlled vocabulary[1].



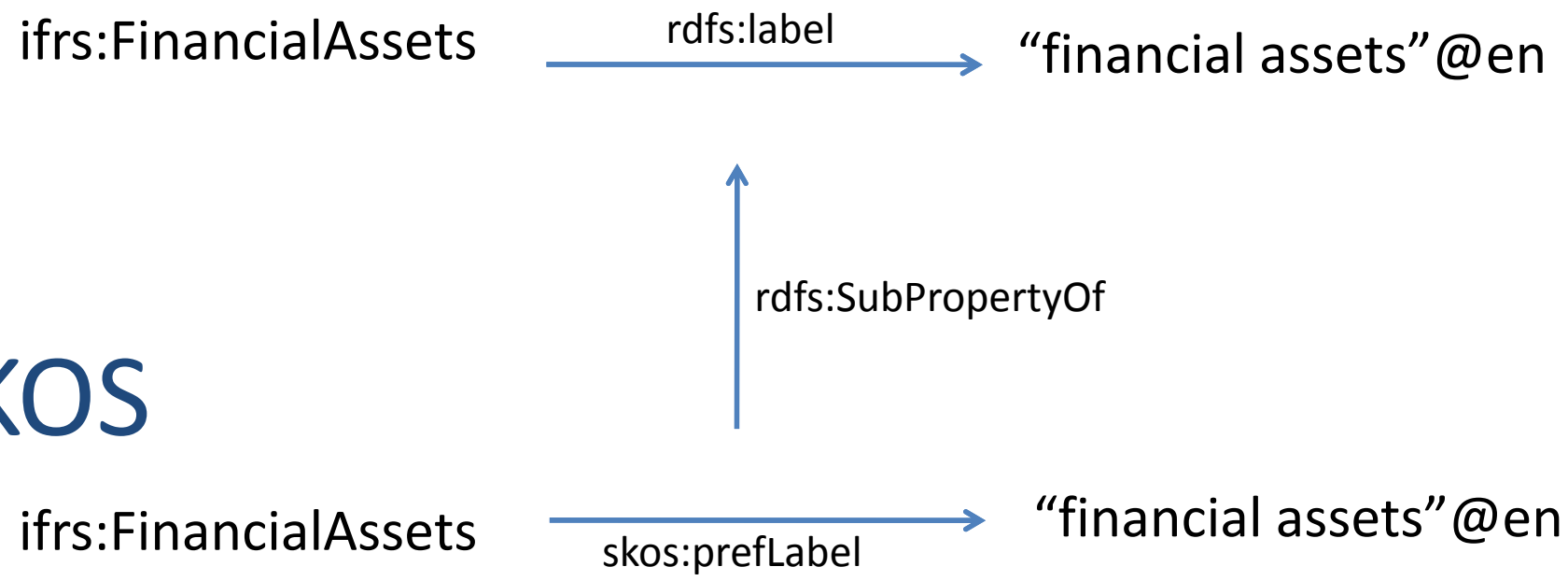
2

1 http://www.w3.org/2004/02/skos/specs

2 http://www.mkbergman.com/date/2007/05/

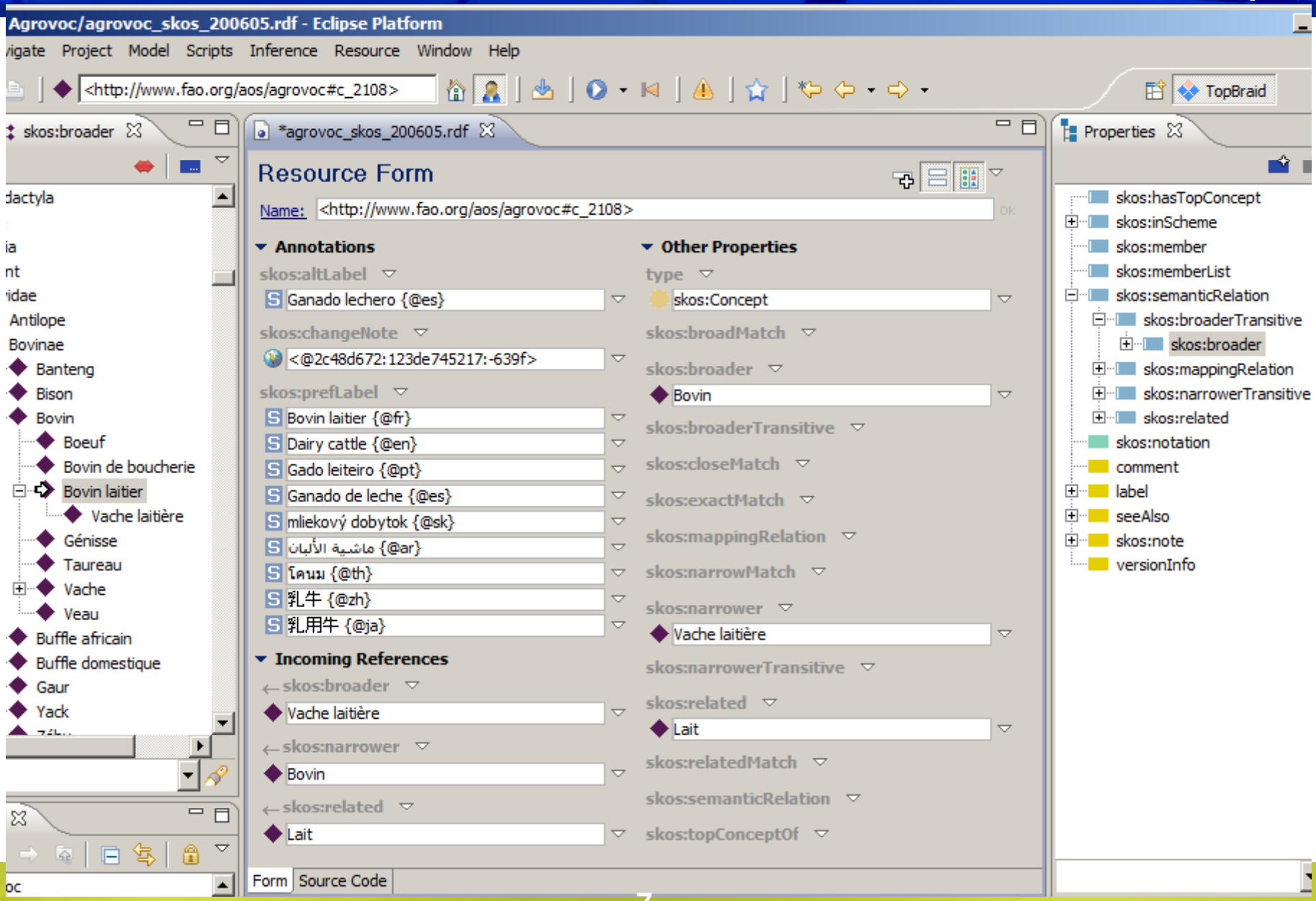
SKOS Labels

RDF(S), OWL



SKOS labels: prefLabel, altLabel & hiddenLabel.

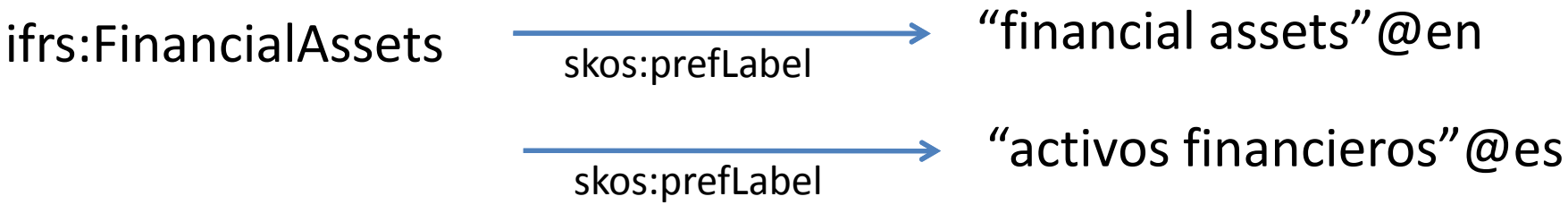
SKOS Example



Multilingual Labels in SKOS

- SKOS enables a simple form of **multilingual labeling**.

```
ex:asset rdf:type skos:Concept;
skos:prefLabel “financial assets"@en;
skos:prefLabel "activos financieros"@es.
```

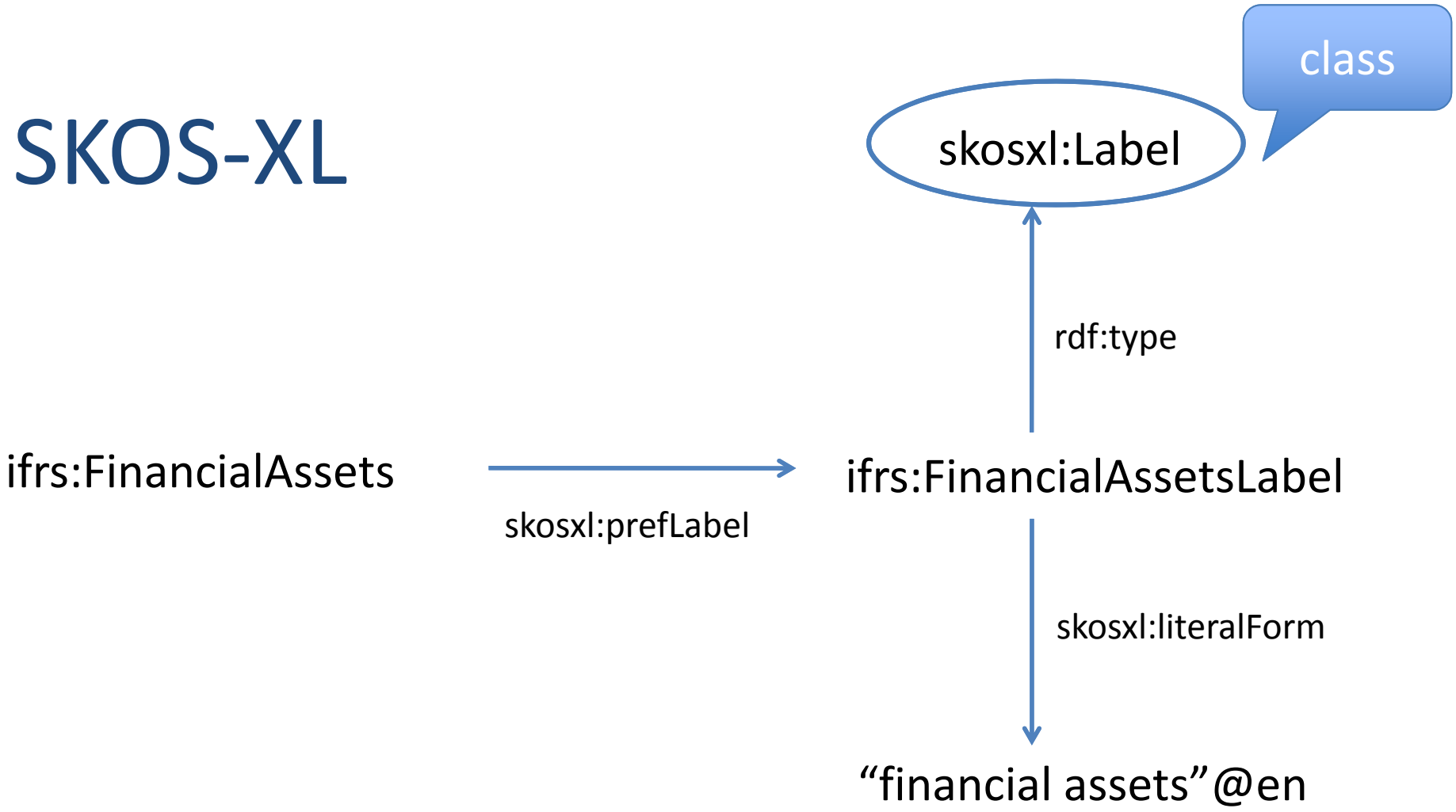


How can we create **explicit links** between labels?

Say that one is *translation* of the other, *acronym*, etc.?

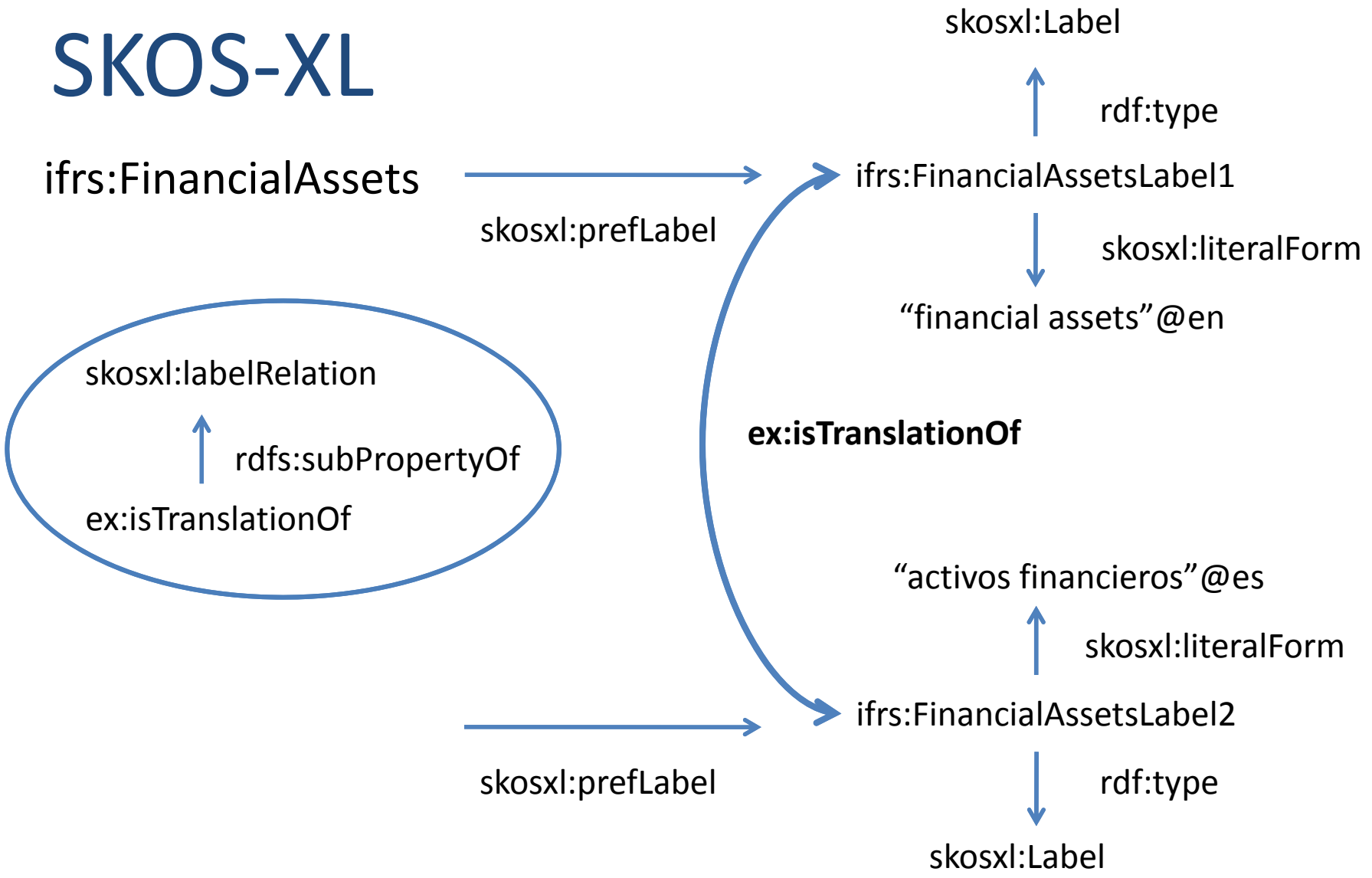
One step forward…

SKOS-XL



Explicit relations in SKOS-XL

SKOS-XL



What is the evolution?

SKOS-XXXXXXXXL ???

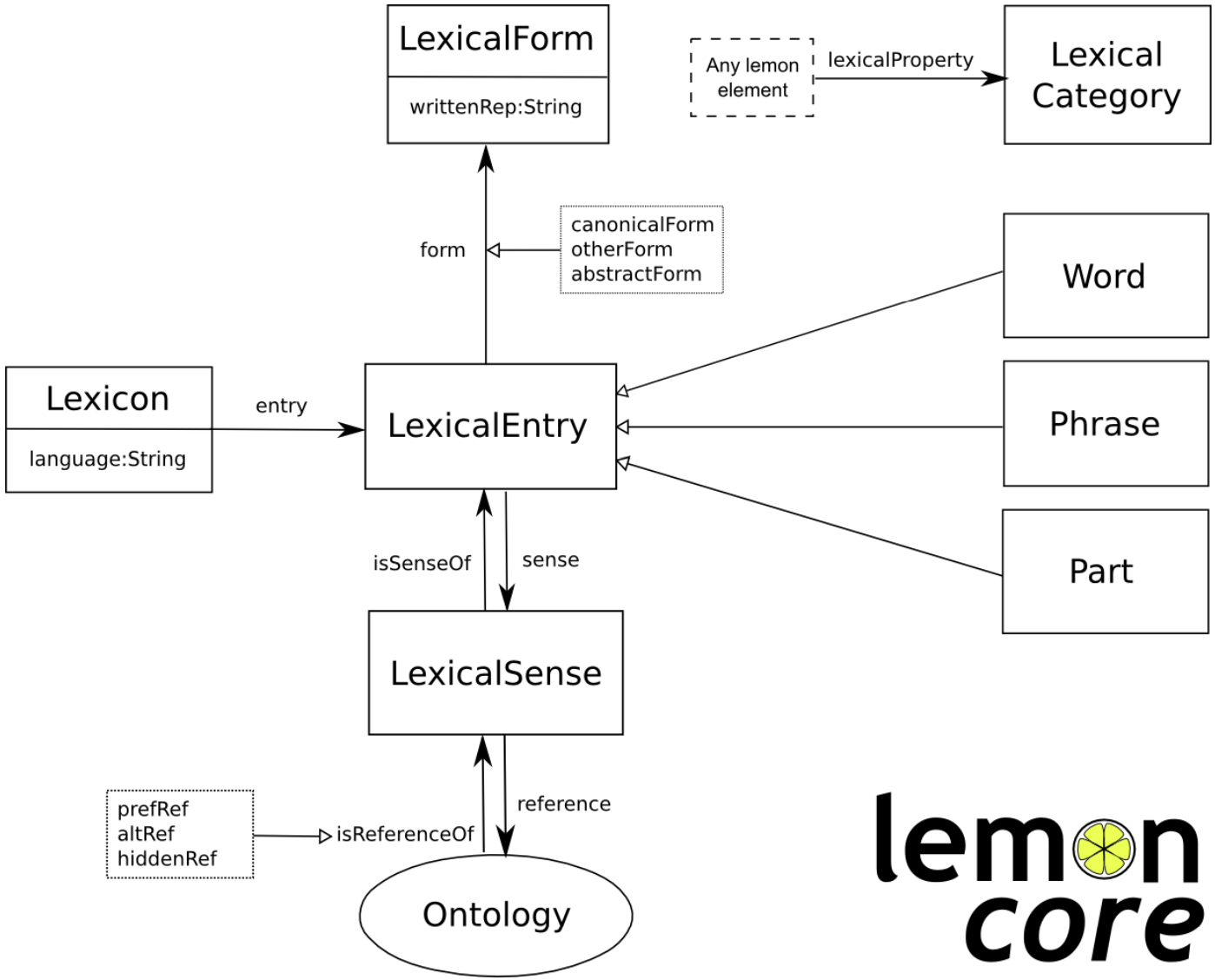

lemon

lemon

lemon, RDF-based lexicon model for ontologies

- Main features:
  - Semantics by reference
  - Rich multilingual lexical and terminological description of ontology elements
  - Concise (i.e., trade off between complexity and expressivity)
  - Descriptive not prescriptive (i.e., uses data categories)
  - Modular and extensible

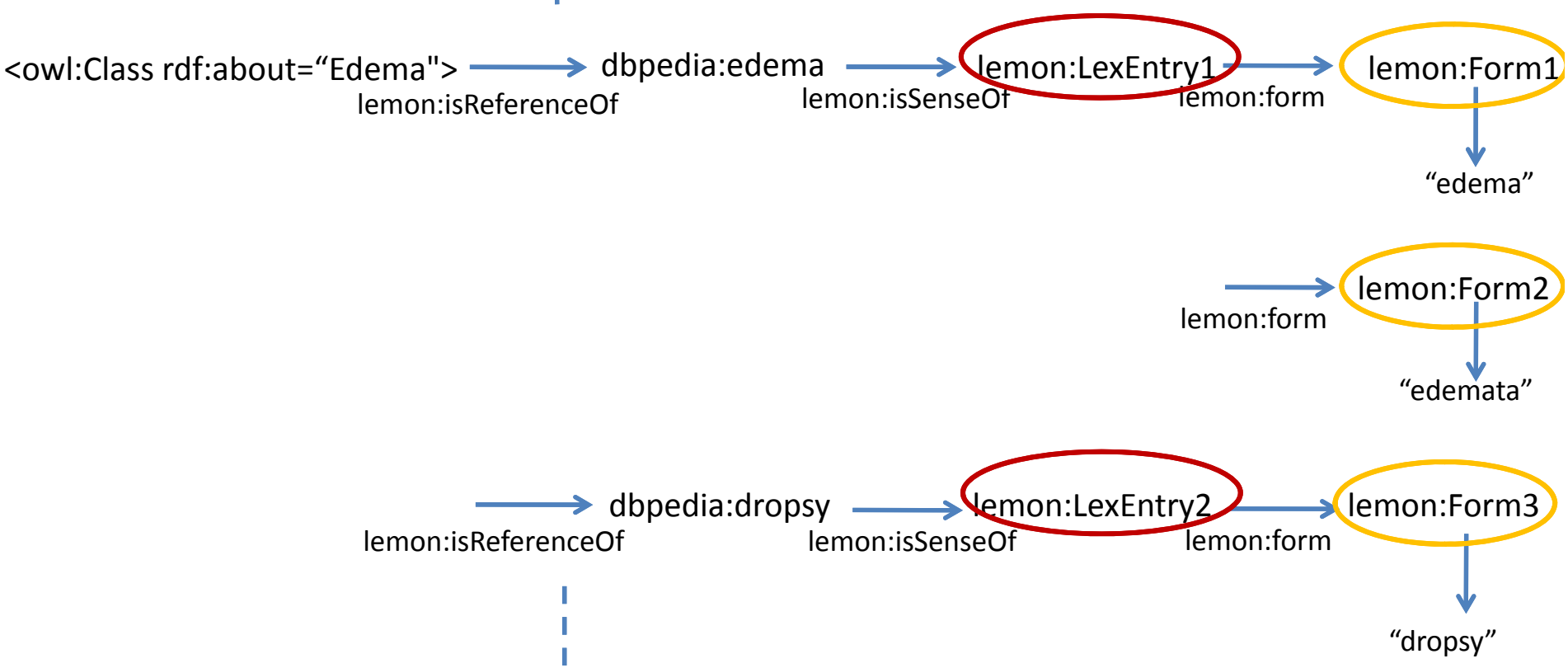
The Model


LexicalForm
writtenRep:String
Any lemon element
lexicalProperty
Lexical Category
form
canonicalForm
otherForm
abstractForm
Word
Lexicon
language:String
entry
LexicalEntry
Phrase
Part
isSenseOf
sense
LexicalSense
prefRef
altRef
hiddenRef
isReferenceOf
reference
Ontology
lemon core

Lexical variants

Ontology

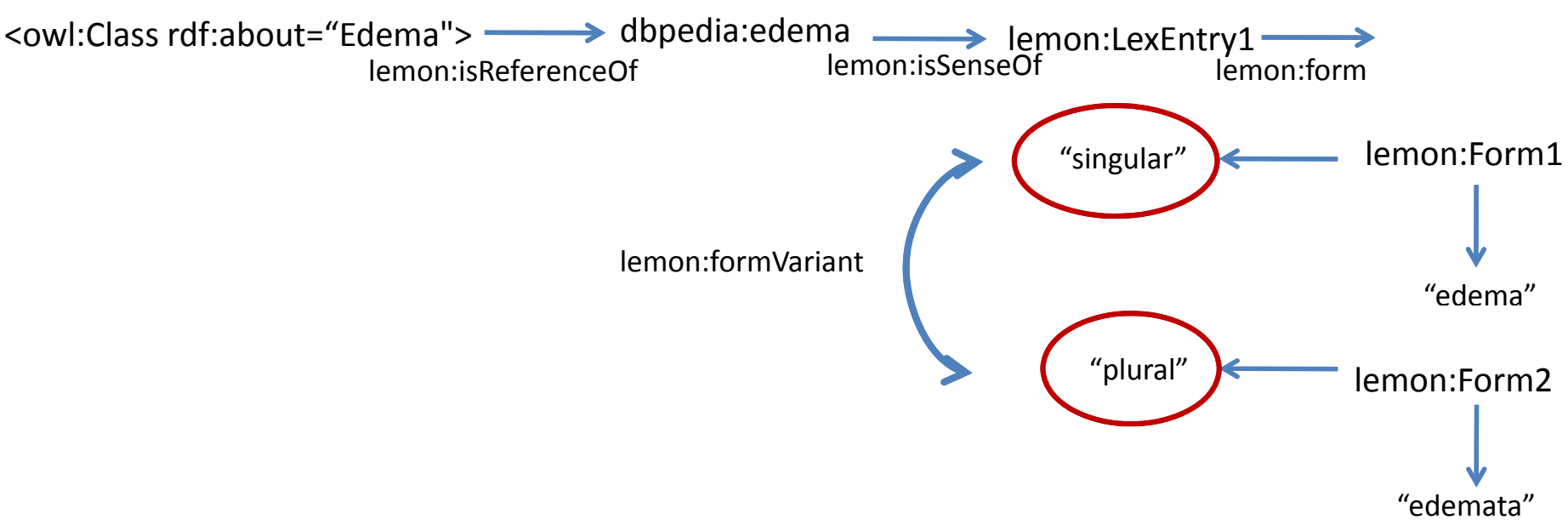
lemon



Form variants

Ontology

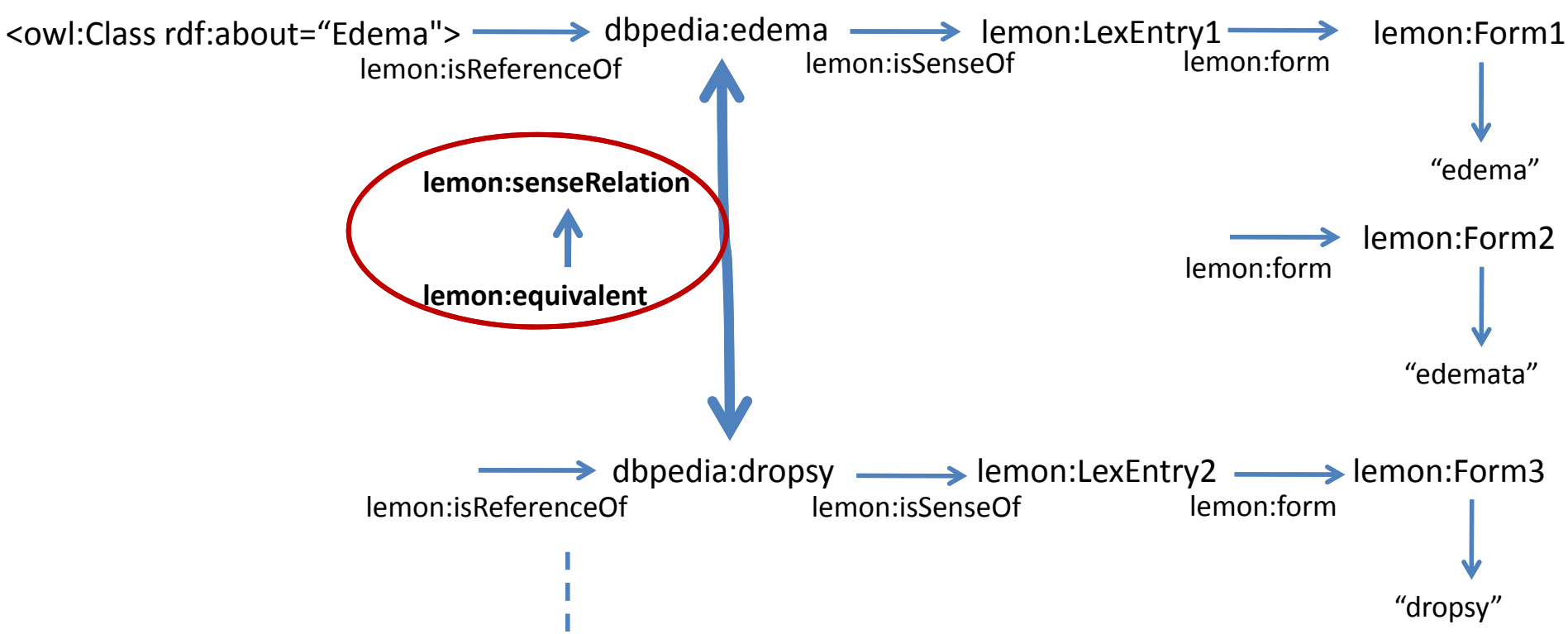
lemon

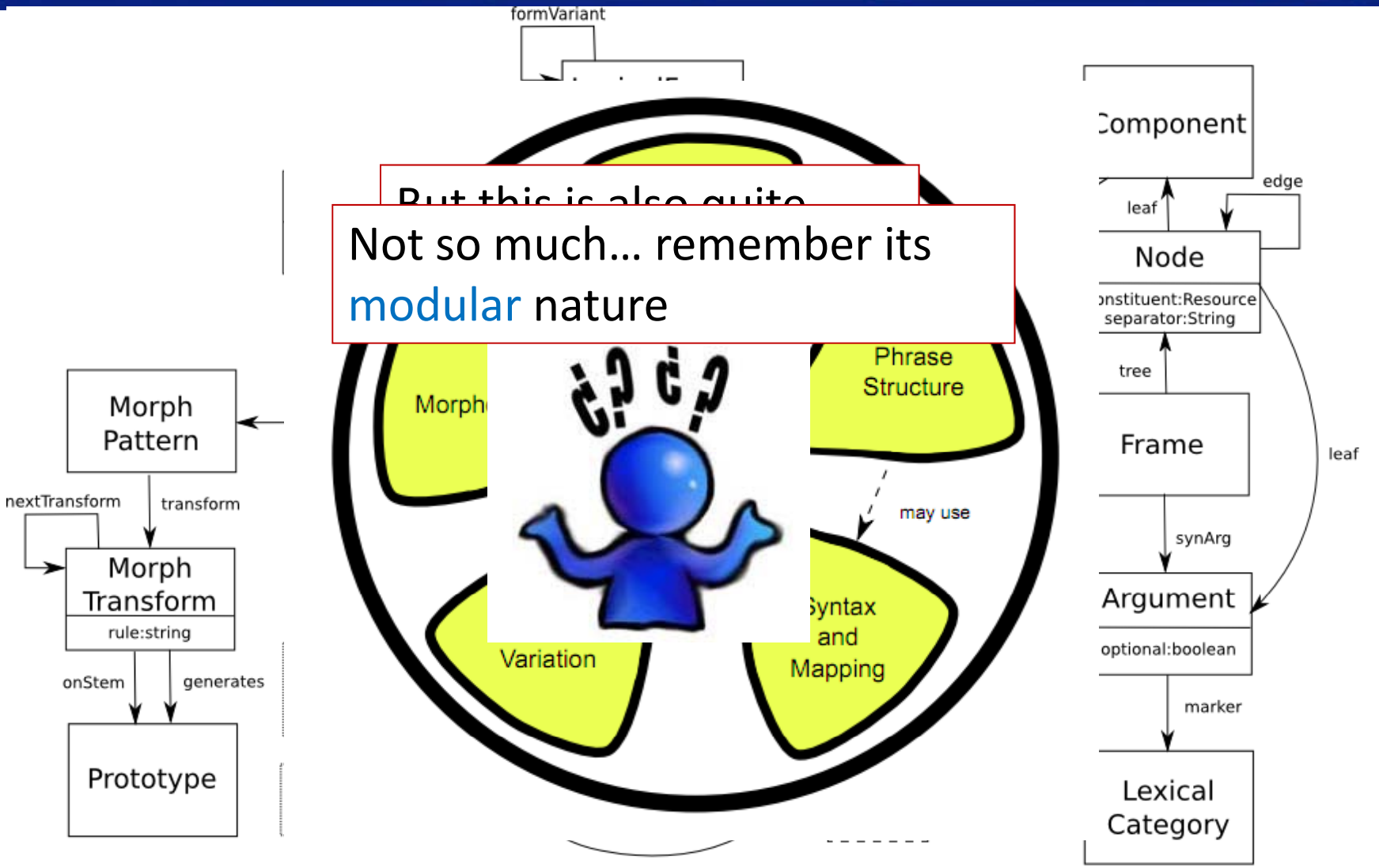


Sense relation

Ontology

lemon





* LexicalEntry has three subclasses: Word, Phrase, Part
‡ definition and example are stated as nodes with a value
⋆ condition has subproperties propertyDomain and propertyRange
† decomposition and element may also be used with Frames and Arguments resp.

lemon

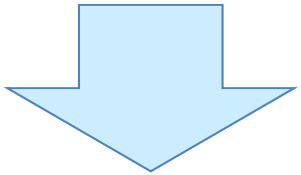
Standardization initiative

- OntoLex W3C Community Group

www.w3.org/**community/ontolex/**

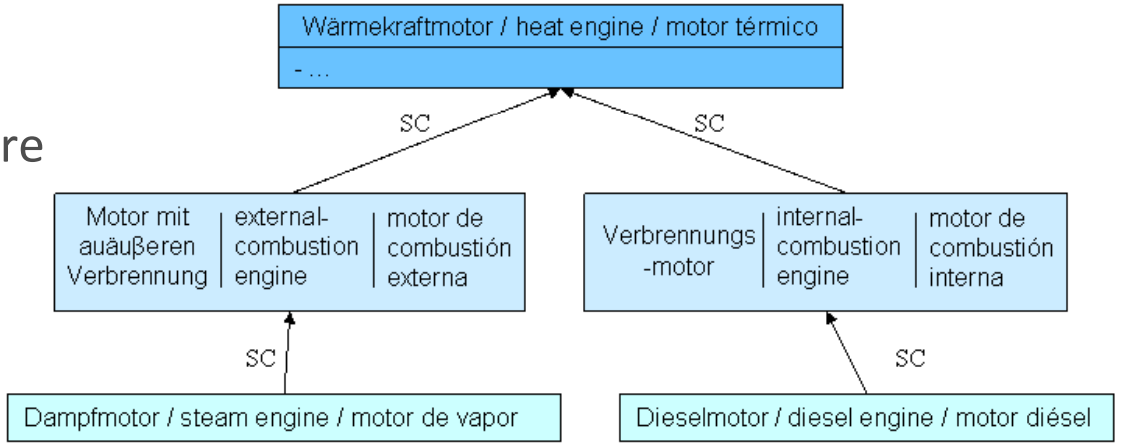
Ontology Translation

Ontology Localization



Multilingual Ontology

(an ontololgy in which labels are documented in multiple NLs)



but…

Does a 1 to 1 correspondence between always exist?

Typology of conceptualizations

Types of domains

Internationalized or standardized domains

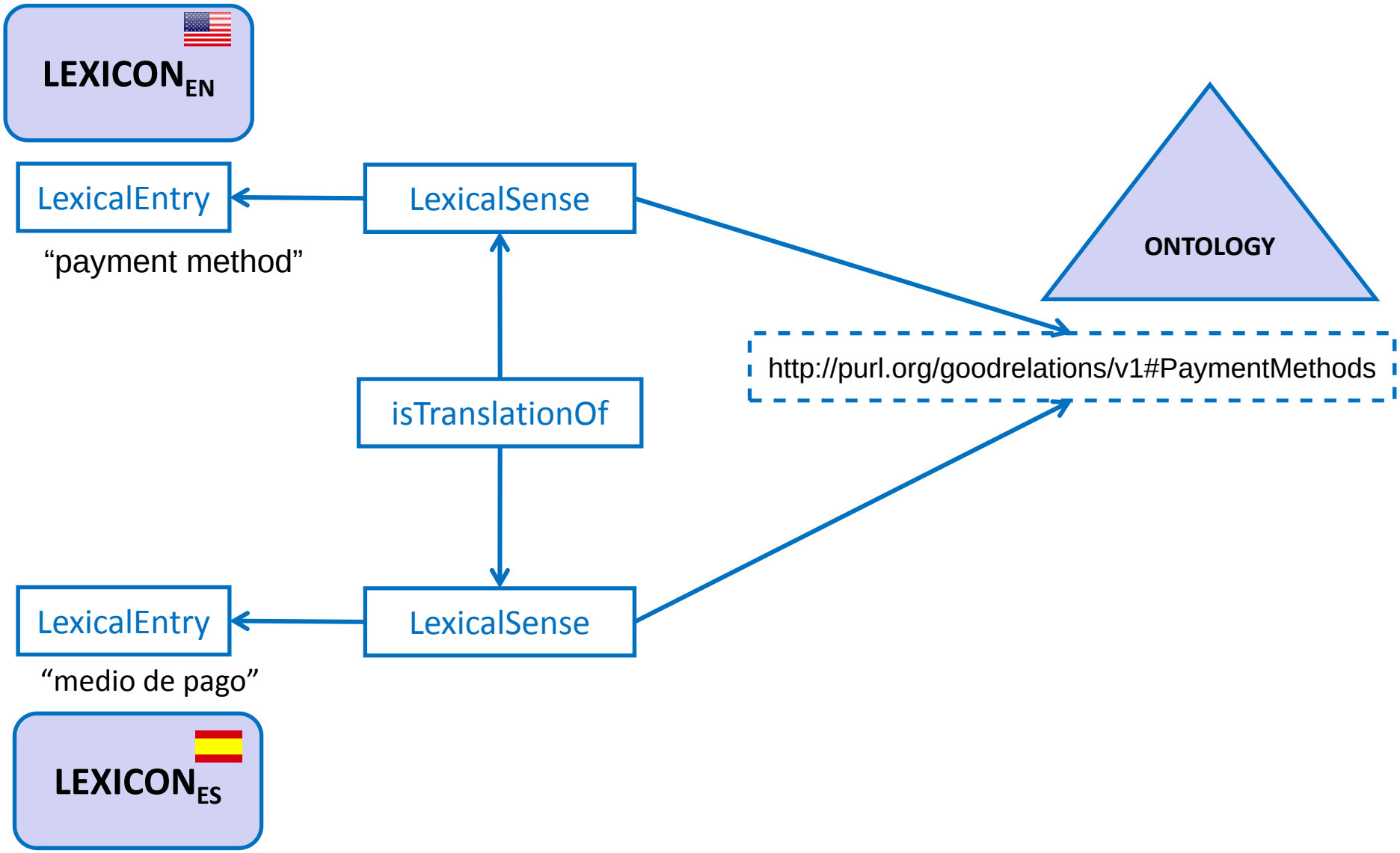
Culturally influenced domains

Types of conceptualizations

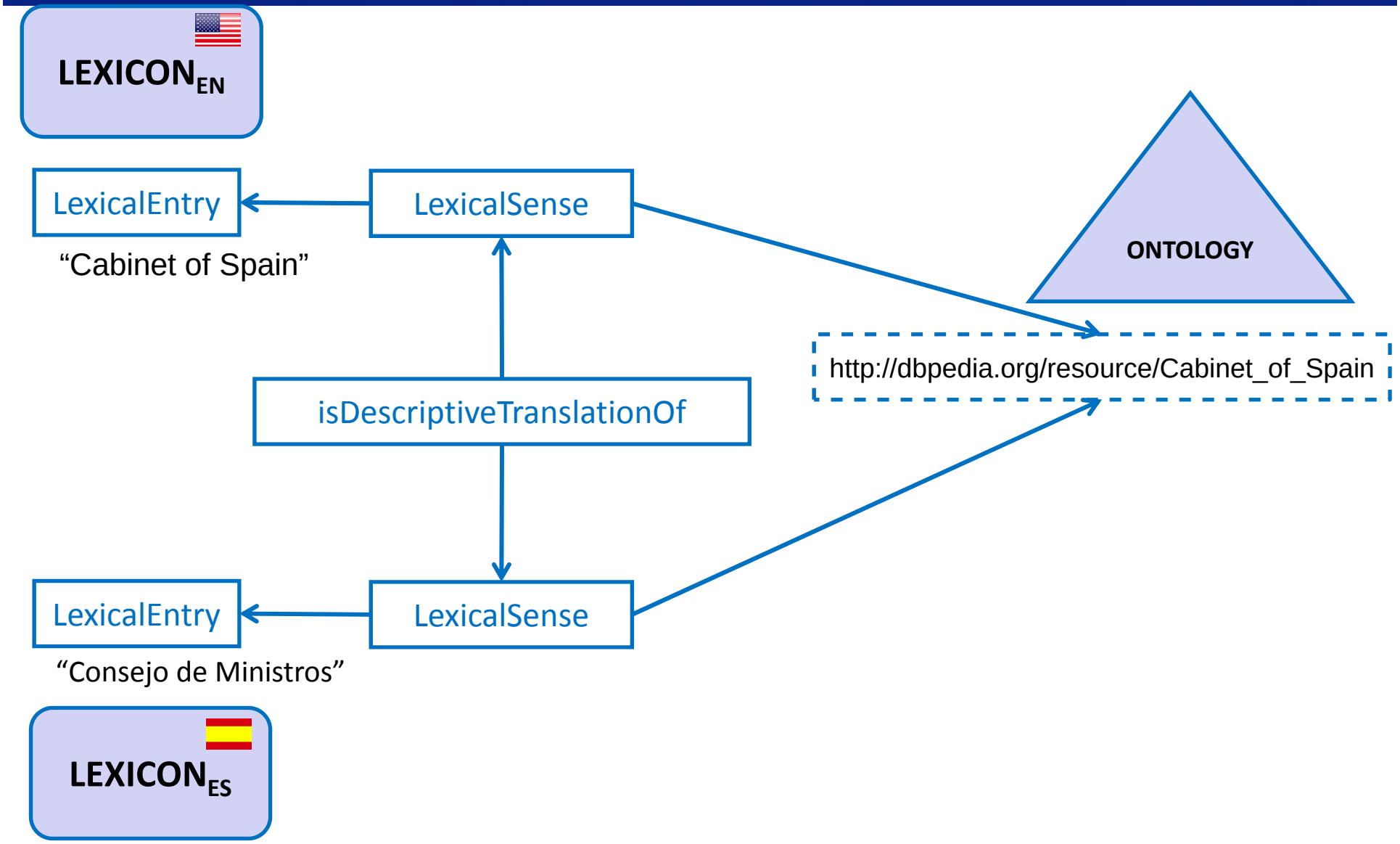
Conceptualizations **shared** among the languages represented in the ontology

Conceptualizations that represent **mismatches** between cultures and languages

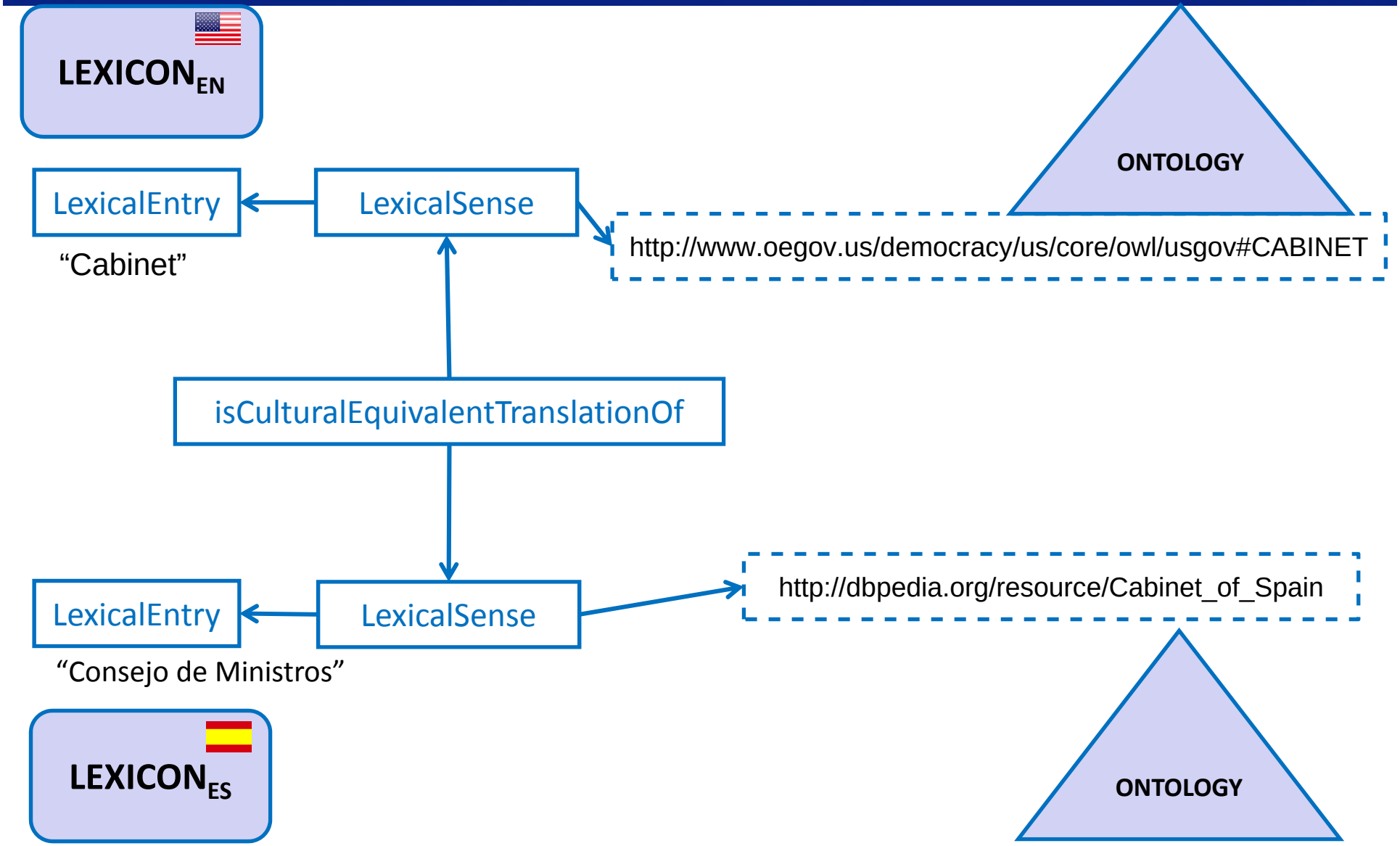
lemon translation module



lemon translation module



lemon translation module



Multilingualism in the Web of Data

Elena Montiel-Ponsoda

Ontology Engineering Group, Universidad Politécnica de Madrid

Monnet is supported by the European Union under Grant No. 248458