



Representing multilingual lexical and terminological information in RDF vocabularies

BabelData project - TIN2010-17550

Elena Montiel-Ponsoda

Ontology Engineering Group – Universidad Politécnica de Madrid

Event: Vocabulary Day – DCMI 2013

Place: Lisbon, Portugal

Date: September 2013

- Representing lexica (a quick overview)
 - RDF(S), OWL
 - SKOS, SKOS-XL
 - LMF
 - ...
- *lemon*
 - Main features
 - *lemon* core
 - Representing lexical variants, terminological variants and translations
 - Tools
- W3C Ontology Lexica Community Group

RDF(S), OWL

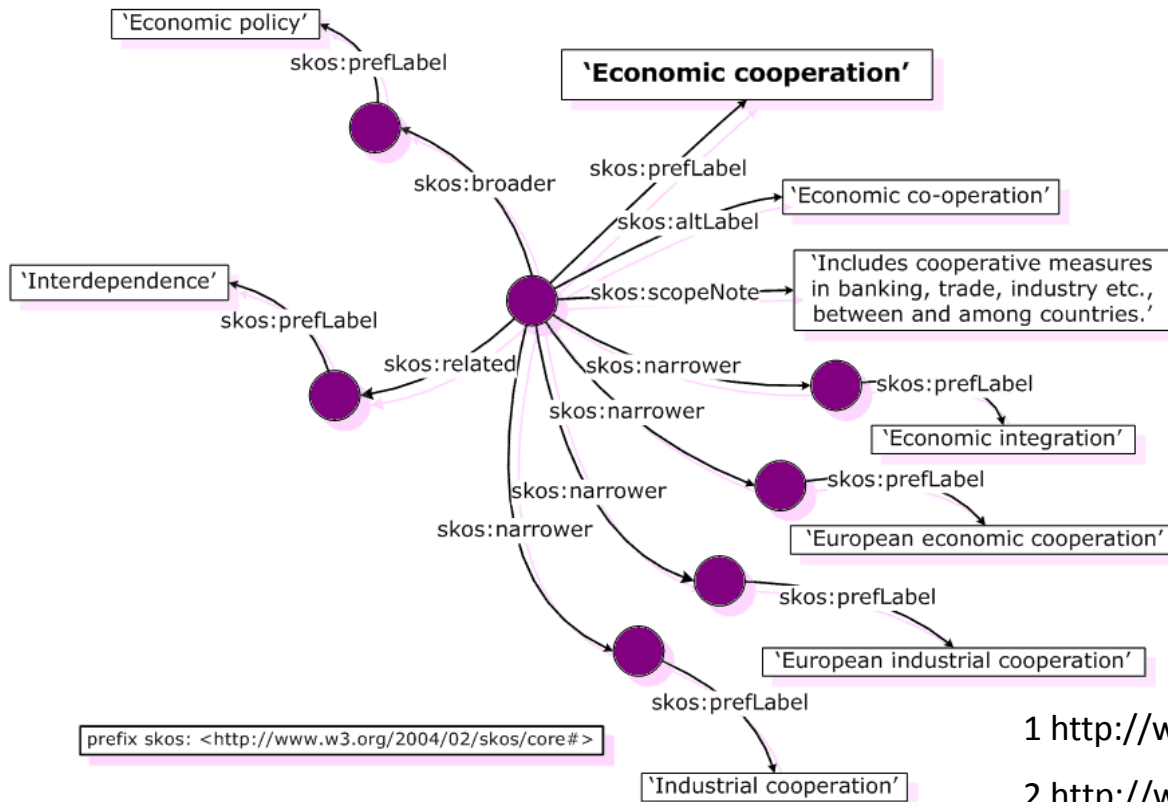
dct:Abstract

rdfs:label



"Abstract"@en

- 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 vocabulary¹.



2

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

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

RDF(S), OWL

dct:creator

rdfs:label

"Creator"@en

SKOS

dct:creator

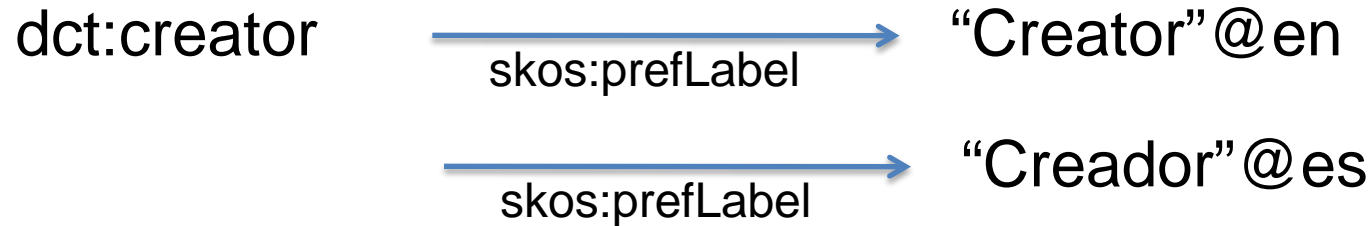
skos:prefLabel

"Creator"@en

rdfs:SubPropertyOf

SKOS labels: prefLabel, altLabel & hiddenLabel.

SKOS enables a simple form of **multilingual labeling**:



```
dct:creator rdf:type rdfs:Property;  
  skos:prefLabel "creator"@en;  
  skos:prefLabel "creador"@es.
```

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

What if we need to define **two preferred labels** in the same language?

SKOS-XL

dct:creator

skosxl:prefLabel

dct:CreatorLabel

rdf:type

skosxl:Label

class

skosxl:literalForm

“Creator”@en

SKOS-XL

dct:creator



skosxl:prefLabel

skosxl:Label



rdf:type

dct:CreatorLabel1



skosxl:literalForm

“Creator”@en

ex:isTranslationOf

“Creador”@es



skosxl:literalForm

dct:Creator2

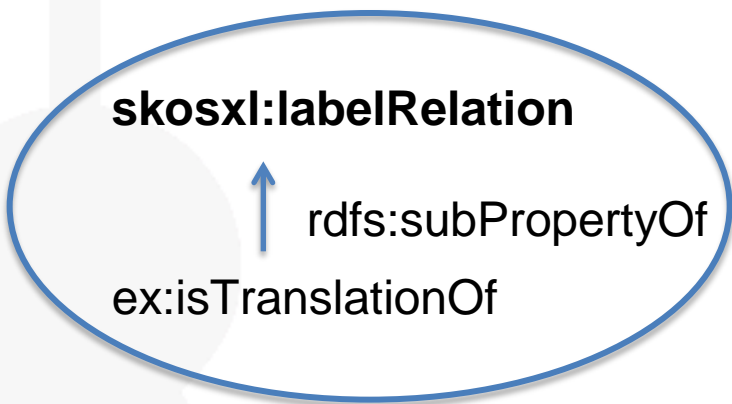


rdf:type

skosxl:Label



skosxl:prefLabel



skosxl:labelRelation



rdfs:subPropertyOf

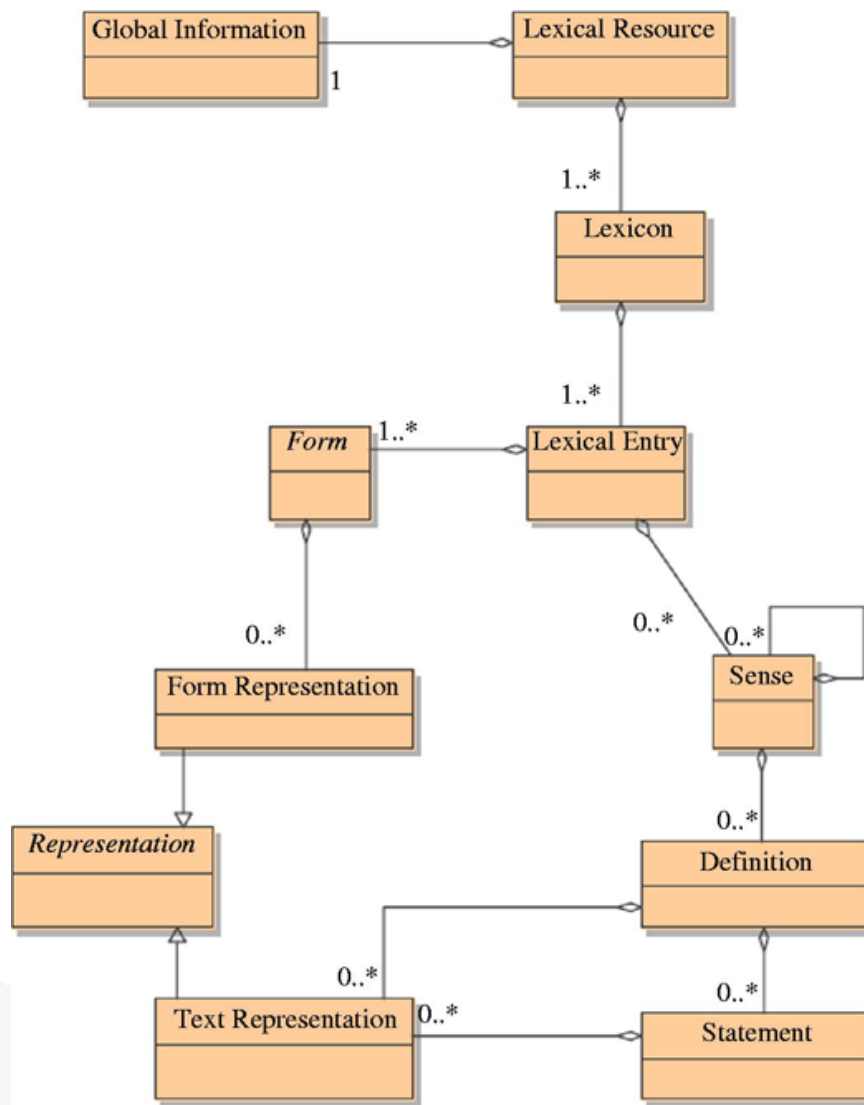
ex:isTranslationOf



What is the
evolution?

LMF

Lexical markup framework



Source: ISO/TC 37/SC 4 N453 (2008)


- SKOS
 - Limited for richer linguistic descriptions
- LMF
 - Conceptualisation not ontology-driven
 - Not RDF-native
 - Some data categories hidden inside literal values
- LingInfo, LIR (Linguistic Information Repository), LexInfo, LexOnto...

lemon

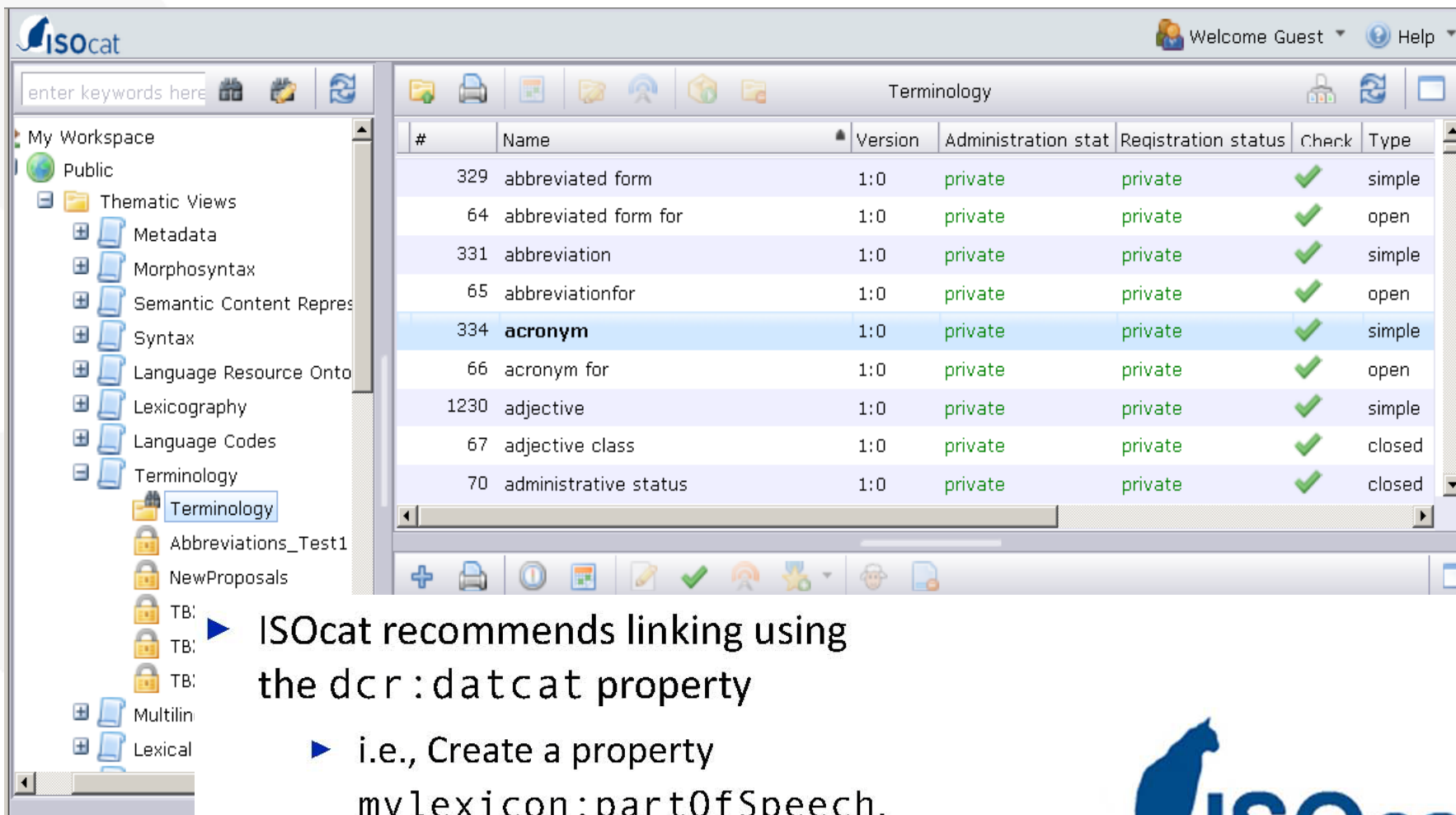
An RDF-based lexicon model for ontologies

- Main features:

lemon 

- Semantics by **reference**
- Rich **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

Details in lemon cookbook: <http://lexinfo.net/lemon-cookbook.pdf>



ISOcat

Welcome Guest Help

enter keywords here

Terminology

#	Name	Version	Administration stat	Registration status	Check	Type
329	abbreviated form	1:0	private	private	✓	simple
64	abbreviated form for	1:0	private	private	✓	open
331	abbreviation	1:0	private	private	✓	simple
65	abbreviationfor	1:0	private	private	✓	open
334	acronym	1:0	private	private	✓	simple
66	acronym for	1:0	private	private	✓	open
1230	adjective	1:0	private	private	✓	simple
67	adjective class	1:0	private	private	✓	closed
70	administrative status	1:0	private	private	✓	closed

My Workspace

- Public
 - Thematic Views
 - Metadata
 - Morphosyntax
 - Semantic Content Repres
 - Syntax
 - Language Resource Onto
 - Lexicography
 - Language Codes
 - Terminology
 - Terminology
 - Abbreviations_Test1
 - NewProposals
 - TB:
 - TB:
 - TB:
 - Multilin
 - Lexical

▶ ISOcat recommends linking using the `dcr:datcat` property

- ▶ i.e., Create a property `mylexicon:partofSpeech`.
- ▶ Add triple relation to ISOcat identifier, e.g., DC-396

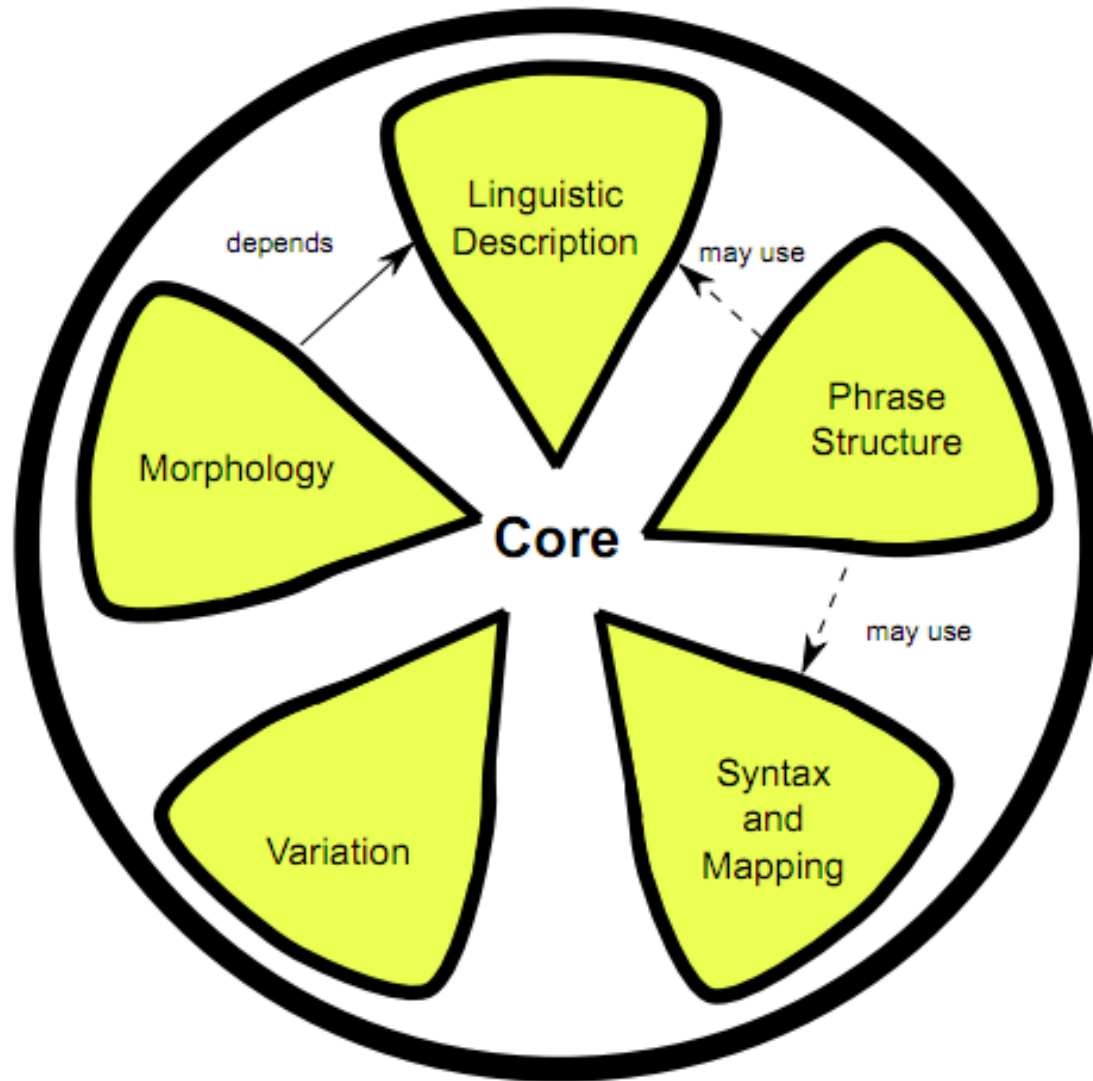


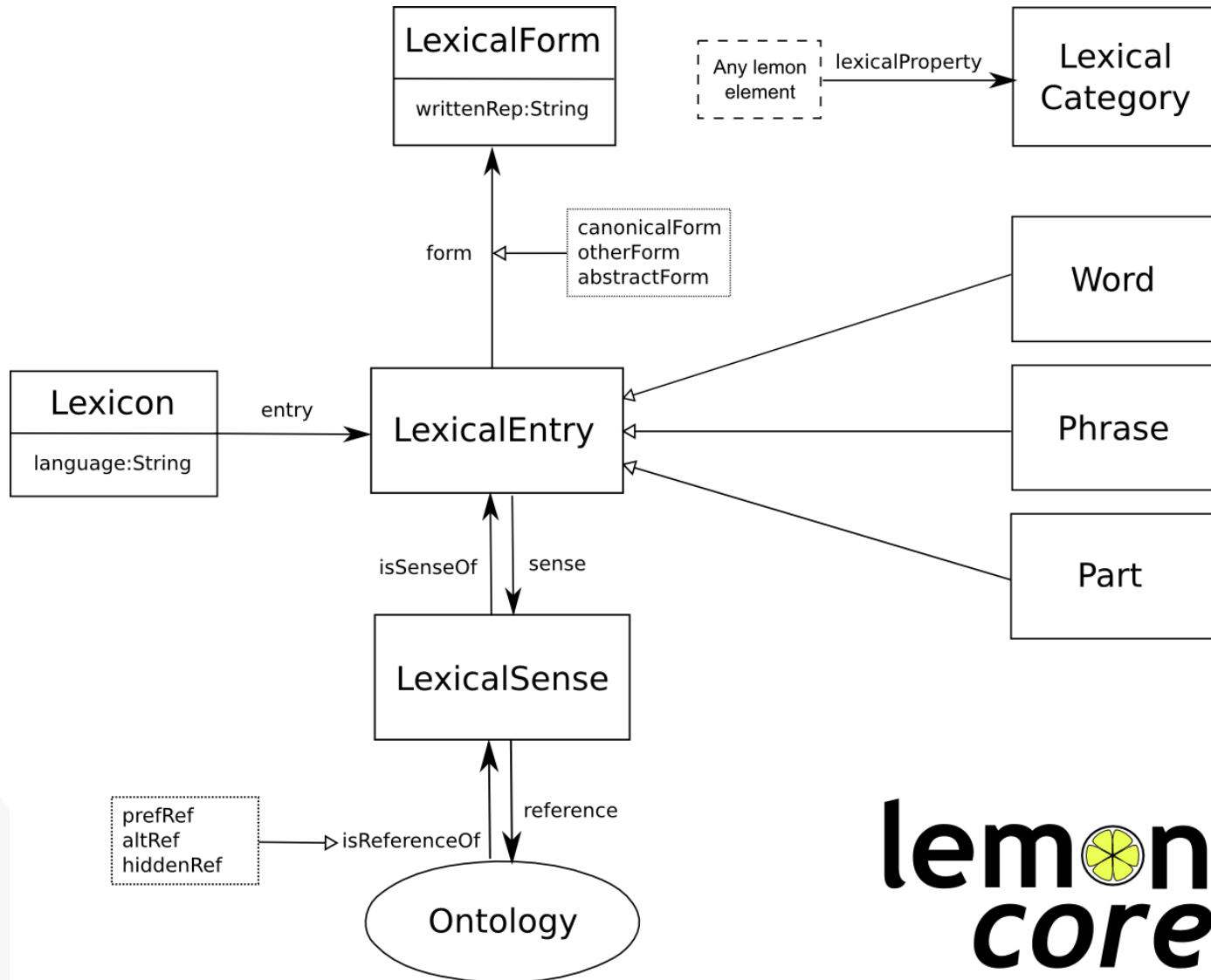
- Main features:



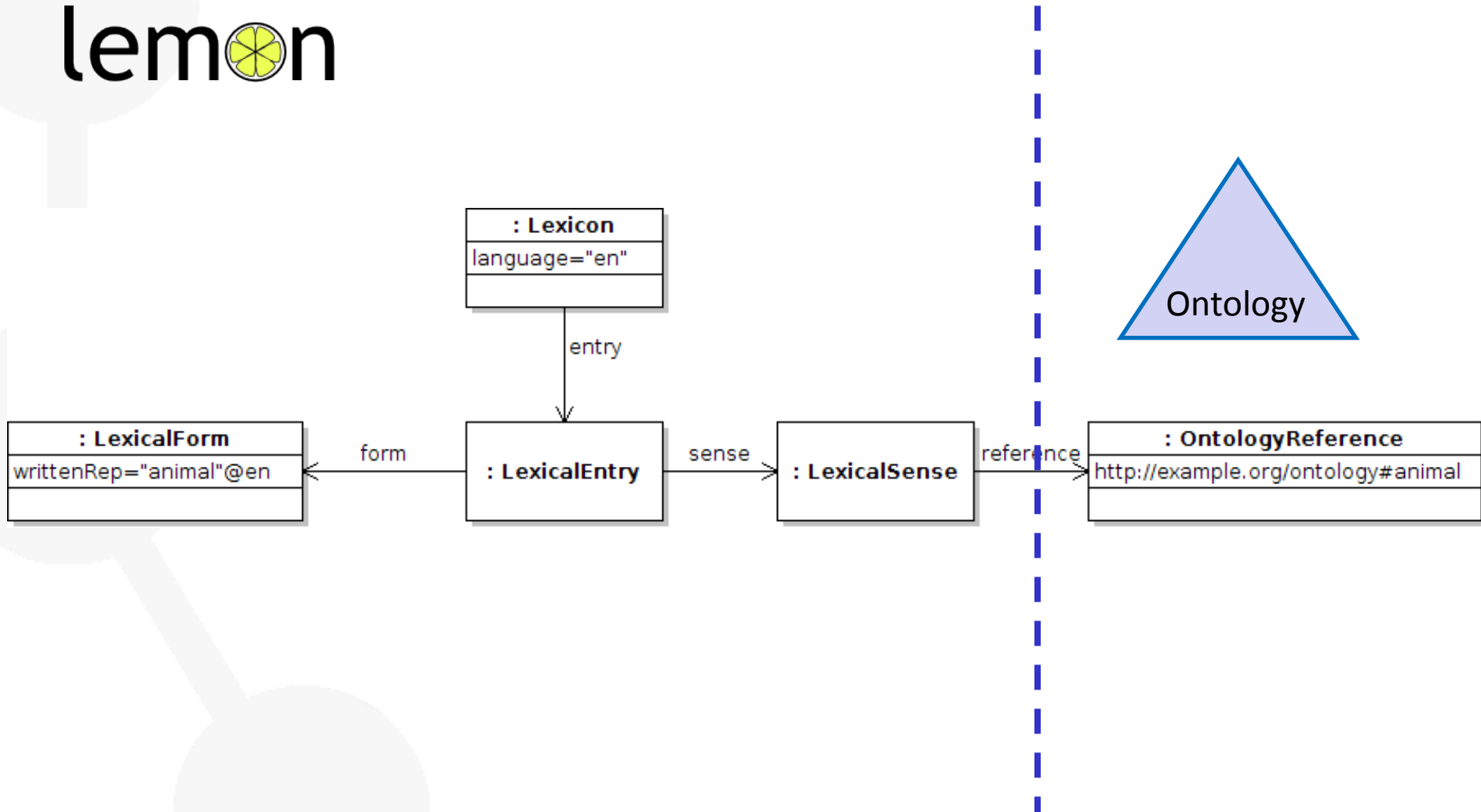
- Semantics by **reference**
- Rich **lexical** and **terminological** description of ontology elements
- **Concise** (i.e., trade off between complexity and expressivity)
- **Descriptive** not prescriptive (i.e., uses data categories)
- **Modular** and extensible →

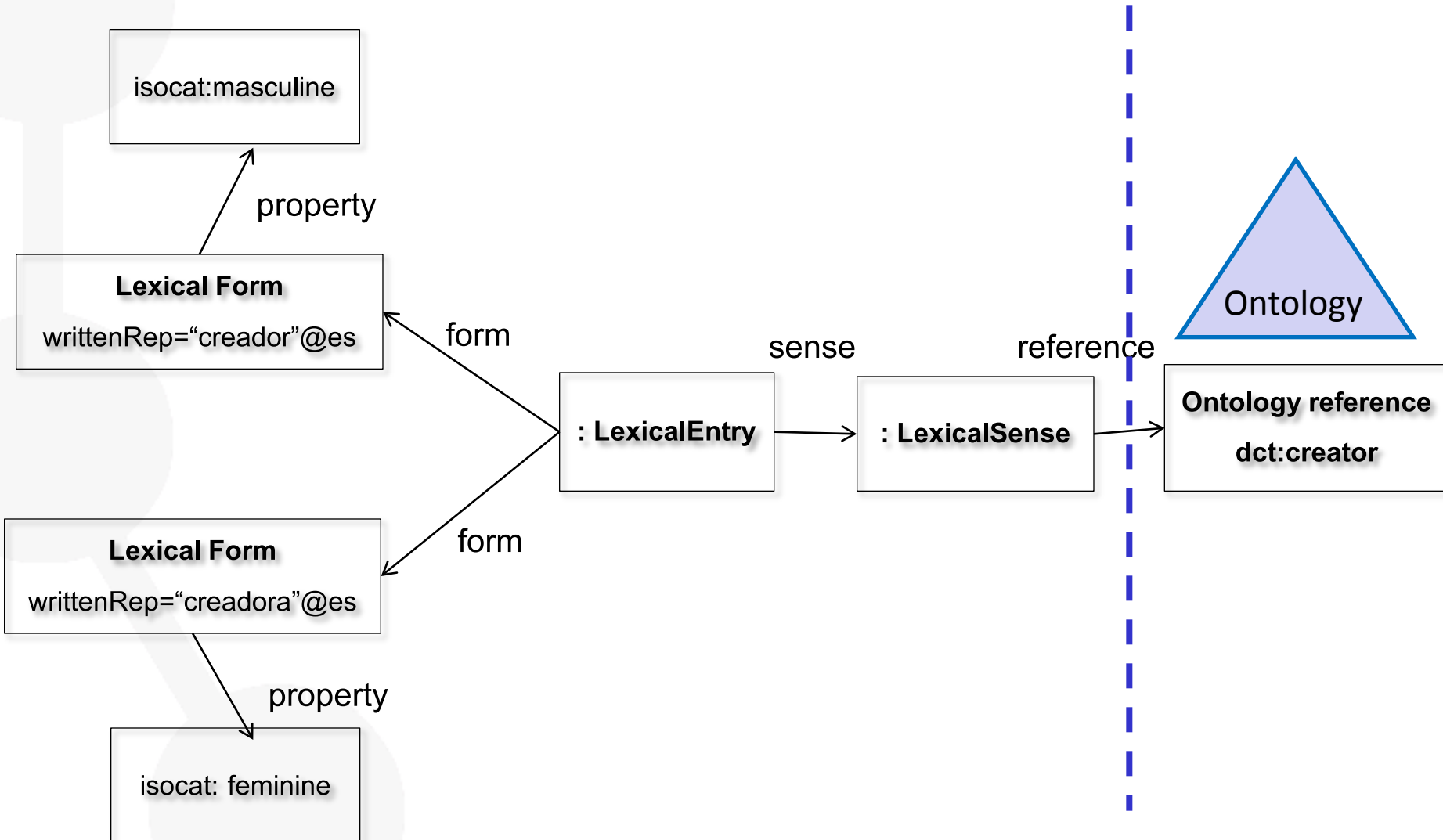
Details in lemon cookbook: <http://lexinfo.net/lemon-cookbook.pdf>



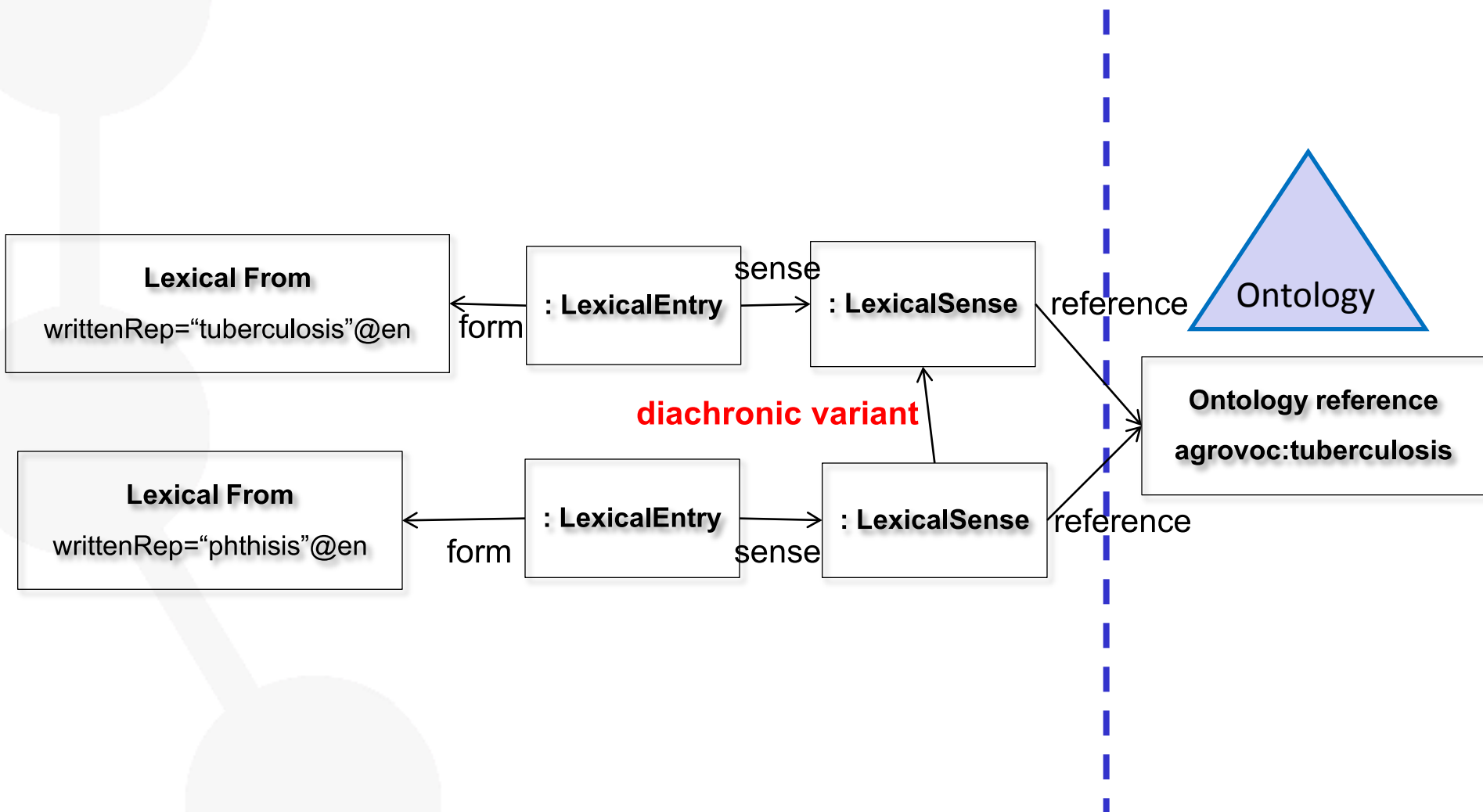


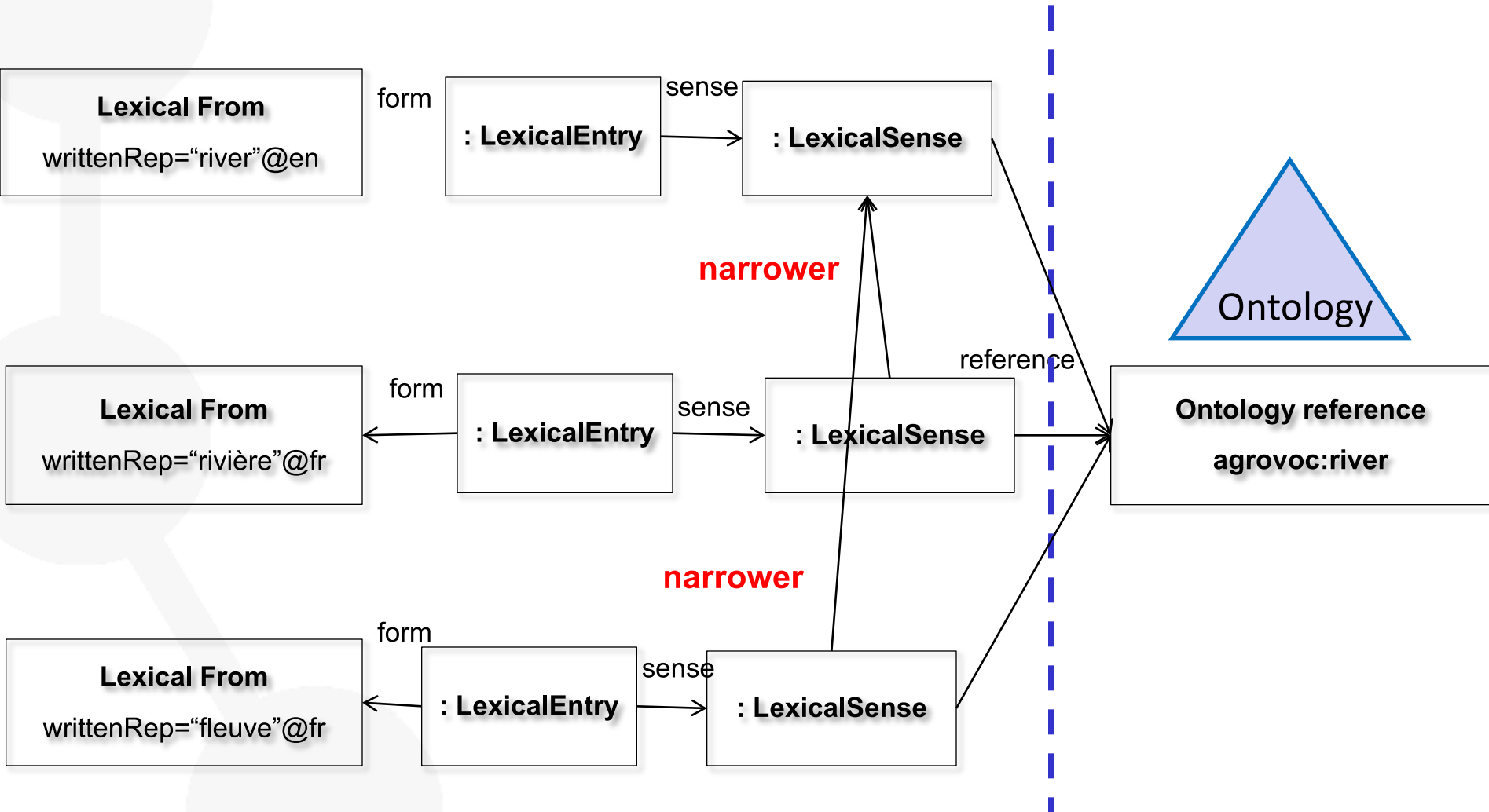
lemon





Terminological variants







Search resources

Home

Learn

Tools and Downloads

Lexica

Model files

API

Source Editor

lemon2gf

Downloads

- [Download the model as RDF/XML](#)
- [Download the model as Turtle](#)
- [View the model as HTML+RDFa](#)

Other Tools

- [Java Lemon API](#)
- [Lemon 2 GF Converter](#)
- [Lemon Source](#)



monnet

author__noun

Instance of: [lemon:LexicalEntry](#)

Property	Value
ns1:partOfSpeech	lexinfo:commonNoun
lemon:canonicalForm	<i>canonicalForm</i>
	rdf:type lemon:Form
	lemon:writtenRep + "author"
lemon:sense	<i>sense</i>
	rdf:type lemon:LexicalSense
	lemon:isA <i>subject</i>
	rdf:type lemon:Argument
	lemon:reference http://dbpedia.org/ontology/Author
lemon:sense	<i>sense1</i>
	rdf:type lemon:LexicalSense
	lemon:objOfProp <i>adpositionalObject</i>
	rdf:type lemon:Argument
	lemon:reference http://dbpedia.org/ontology/writer
	lemon:subjOfProp <i>subject1</i>



Mailing List

Wiki

RSS

Contact Group

[Home](#) / [Ontology-Lexica Community...](#)

Ontology-Lexica Community Group

The mission of the Ontology-Lexicon community group is to: (1) Develop models for the representation of lexica (and machine readable dictionaries) relative to ontologies. These lexicon models are intended to represent lexical entries containing information about how ontology elements (classes, properties, individuals etc.) are realized in multiple languages. In addition, the lexical entries contain appropriate linguistic (syntactic, morphological, semantic and pragmatic) information that constrains the usage of the entry. (2) Demonstrate the added value of representing lexica on the Semantic Web, in particularly focusing on how the use of linked data principles can allow for the re-use of existing linguistic information from resource such as WordNet. (3) Provide best practices for the use of linguistic data categories in combination with lexica. (4) Demonstrate that the creation of such lexica in combination with the semantics contained in ontologies can improve the performance of NLP tools. (5) Bring together people working on standards for representing linguistic information (syntactic, morphological, semantic and pragmatic) building on existing initiatives, and identifying collaboration tracks for the future. (6) Cater for interoperability among existing models to represent and structure linguistic information. (7) Demonstrate the added value of applications relying on the use of the combination of lexica and ontologies.

Reports

Get involved!

Anyone may join this Community Group. All participants in this group have signed the [W3C Community Contributor License Agreement \(CLA\)](#).

[JOIN THIS GROUP](#)

or [learn how to join or request an account](#).

Participants



Paul Buitelaar
Chair



Philipp Cimiano
Chair

- <http://www.w3.org/community/ontolex/>
- Work begun in december 2011
- Monthly telcos
- Chaired by Paul Buitelaar (DERI, Galway) and Philipp Cimiano (University Bielefeld)

- This work has been supported by the BabelData (TIN2010-17550) Spanish national project



Representing multilingual lexical and terminological information in RDF vocabularies

BabelData project - TIN2010-17550

Elena Montiel-Ponsoda

Ontology Engineering Group – Universidad Politécnica de Madrid

Event: Vocabulary Day – DCMI 2013

Place: Lisbon, Portugal

Date: September, 2013