

# INTERCONNECTING AND MANAGING MULTILINGUAL LEXICAL LINKED DATA

Ernesto William De Luca

# Overview

2

- Motivation
- EuroWordNet
- RDF/OWL EuroWordNet
- RDF/OWL LexiRes Tool
- Conclusions

# Overview

3

- Motivation
- EuroWordNet
- RDF/OWL EuroWordNet
- RDF/OWL LexiRes Tool
- Conclusions

# Multilingualism

4

- Language Resources
  - Contain a wide range of linguistic information
  - Vary from simple lists to complex resources
    - lexical resources: word list, machine readable dictionary, thesauri, ontologies, glossaries, etc ...
  - Can be used for language and knowledge engineering
- WordNet/EuroWordNet: Synonym Sets (SynSets)
  - Each representing one constitutional lexicalized concept
- EuroWordNet format
  - defined by the EuroWordNet Database Editor Polaris that uses a proprietary specification.

## Related Work

5

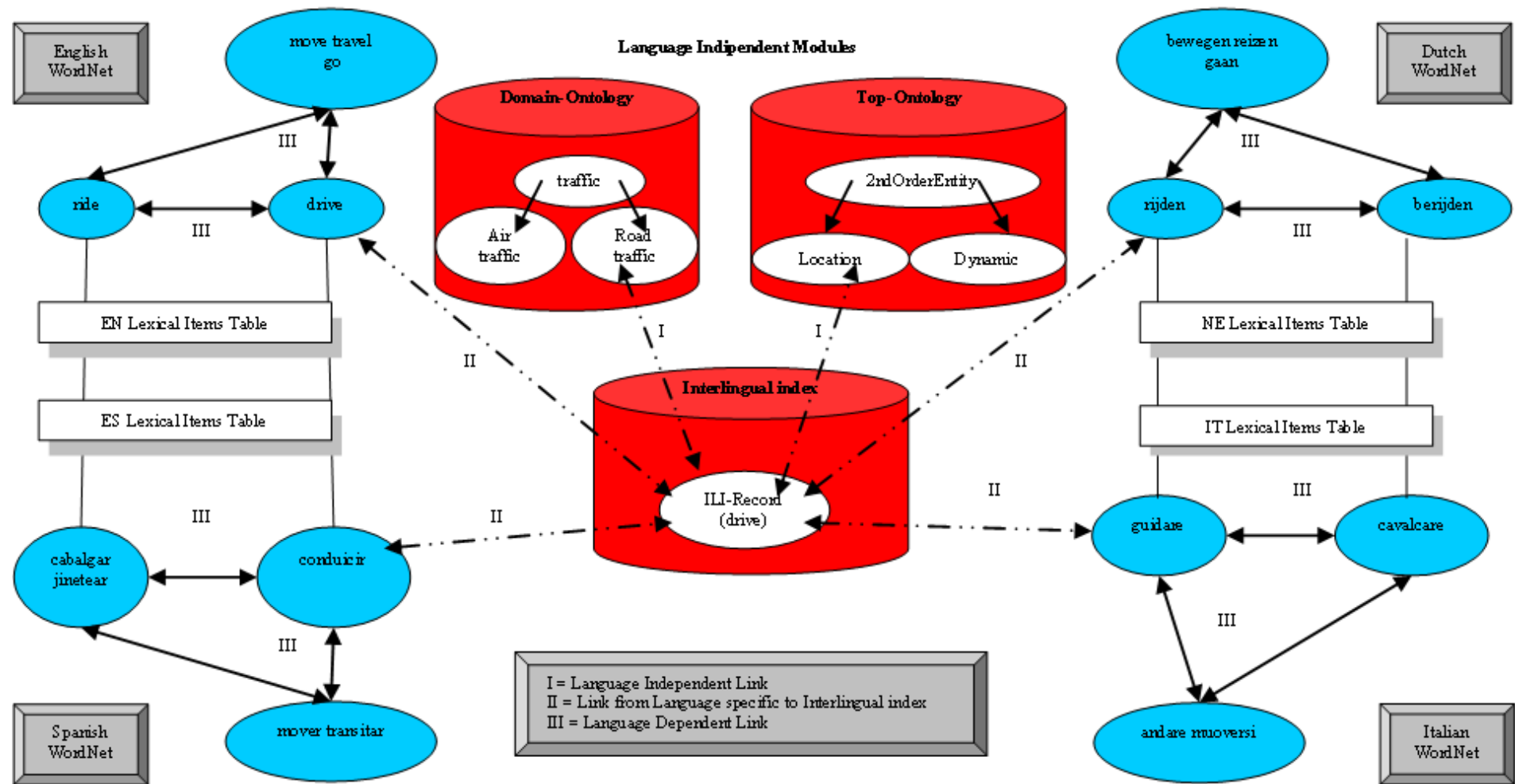
- VisDic
  - Browsing and editing multilingual EuroWordnet information
  - BUT, here users can browse static information on text blocks.
- MultiSemCor Web Interface
  - Bilingual information browsing
  - MultiWordNet-annotated Parallel corpus
    - Browsing of words with their related annotated word senses,
    - BUT, the corpus is very restricted.
    - All accessible information is static.
    - Bilingual search in a closed domain.

# Overview

6

- Motivation
- **EuroWordNet**
- RDF/OWL EuroWordNet
- RDF/OWL LexiRes Tool
- Conclusions

# EuroWordNet Architecture



- General (WordNet) Problems
  - “Expressivity lack” (concepts vs. instances)
    - Solution: Named Entity Recognition Approaches
  - Too fine grained distinction of word senses
    - Solution: Redesign/Merging concepts/SynSets.
  - Not all word senses are covered
    - Solution: WordNet (and OWL) extension



# Overview

9

- Motivation
- EuroWordNet
- **RDF/OWL EuroWordNet**
- RDF/OWL LexiRes Tool
- Conclusions

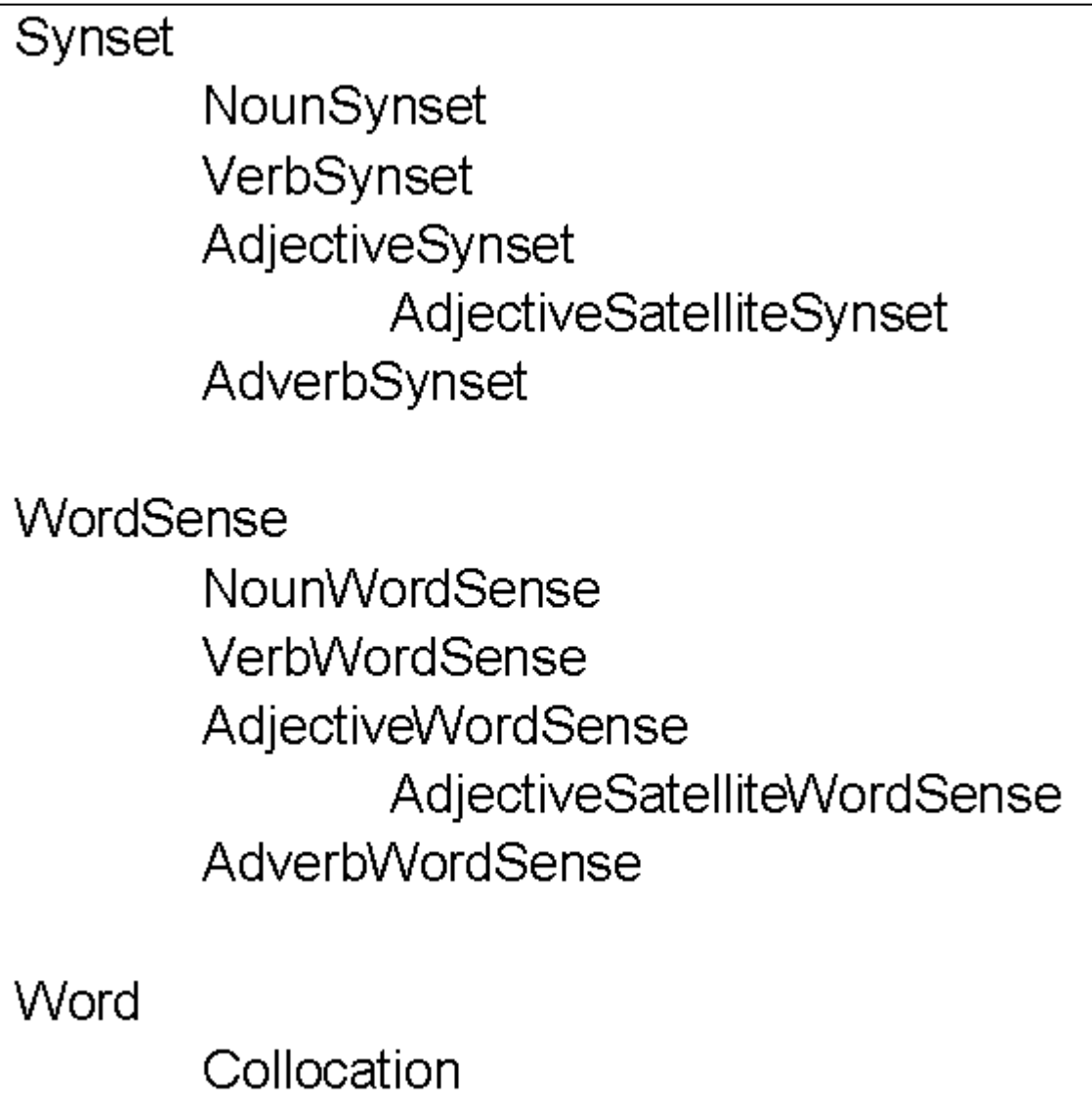
- Conversion of EuroWordNet into RDF/OWL
  - To have a standardized OWL format.
  - To access it with standard OWL query tools.
  - To enrich it with additional domain-specific ontologies.

Ernesto William De Luca, Martin Eul and Andreas Nürnberger. **Converting EuroWordNet in OWL and Extending It with Domain Ontologies.** In: *Proceedings of the Workshop on Lexical-Semantic and Ontological Resources*. In conjunction with GLDV 2007. Tübingen, 2007

- WordNet already been converted into an OWL format from (van Assem et al., 2004) using the OWL-DL sublanguage.

Mark van Assem, Aldo Gangemi, and Guus Schreiber. **Wordnet in rdfs and owl**. Technical report, W3C, 2004.

- WordNet
- OWL
- using



into an  
l., 2004)

Mark  
Guus  
owl. T

d  
nd

Three kinds of properties in the schema:

- Connection of instances of the main classes together.
  - SynSet linked to WordSenses (property containsWordSense)
  - WordSense linked to Word (property word)
- Connection of WordNet relations. Three kinds:
  - between two SynSets (e.g. hyponymOf)
  - between two word senses (e.g. antonymOf)
  - Miscellaneous sets (e.g. gloss)
- Information about entities
  - XML Schema datatypes (e.g. xsd:string).

- Use of URIs to provide some information about the entity meaning, built with pattern similar to:

wn20instances: + synset- + lexical form- + type- + sense number.

- Example Code

```
<wn20schema:NounSynset rdf:about="&wn20instances;synset-  
bank-noun-4" rdfs:label="bank">  
  <wn20schema:synsetId>102690337</wn20schema:synsetId>  
</wn20schema:NounSynset>
```

- Adaptation to the RDF/OWL Schema of WordNet (van Assem et al., 2004) and extension with new relations.
- These steps can be subdivided into:
  - Analysis of the requirements for EuroWordNet
  - Adaptation of WordNet RDF-Schema to EuroWordNet
  - Multilinguality
  - OWL Property Conversion
  - OWL Domain Extension

- To avoid redundancy:
  - Only relations in transitive direction are listed (e.g. hyponymOf and not hypernymOf)
  - Others can be retrieved with the **owl:inverseOf** property implemented in the RDF Schema.
- Instances of all classes and properties (separated in several data files)
  - one for the SynSets
  - one for the WordSenses and Words
  - one for each relation



- van Assem et al. (2004)
  - focused on staying close to the original source (i.e. reflect the original data model without interpretation)
  - direct use in Semantic Web applications, or as a source for modified WordNet versions
- Agreement with their assumptions also for a multilingual task.

# RDF/OWL

## EuroWordNet - Examples

18

```

<ewn20schema:NounSynset rdf:about="&ewn20instances;synset-bank-noun-1"
  rdfs:label="bank">
  <ewn20schema:synsetId>102690337 </ewn20schema:synsetId>
</ewn20schema:NounSynset>
<ewn20schema:Word rdf:about="&ewn20instances;word-bank"
  ewn20schema:lexicalForm="bank"/>
<ewn20schema:NounWordSense rdf:about="&ewn20instances;wordsense-bank-noun-1"
  rdfs:label="bank">
  <ewn20schema:word rdf:resource="&ewn20instances;word-bank"/>
</ewn20schema:NounWordSense>
<rdf:Description rdf:about="&ewn20instances;synset-bank-noun-1">
  <ewn20schema:containsWordSense rdf:resource="&ewn20instances;wordsense-bank-noun-1"/>
  <ewn20schema:containsWordSense rdf:resource="&ewn20instances;wordsense-bank_building-noun-1"/>
</rdf:Description>

```

**eurowordnet-english-synset.rdf**

### e multilingual tool

```

<rdf:Description rdf:about="&ewn20instances;synset-bank-noun-1">
  <ewn20schema:hyponymOf rdf:resource="&ewn20instances;synset-deposit-noun-1"/>
</rdf:Description>

```

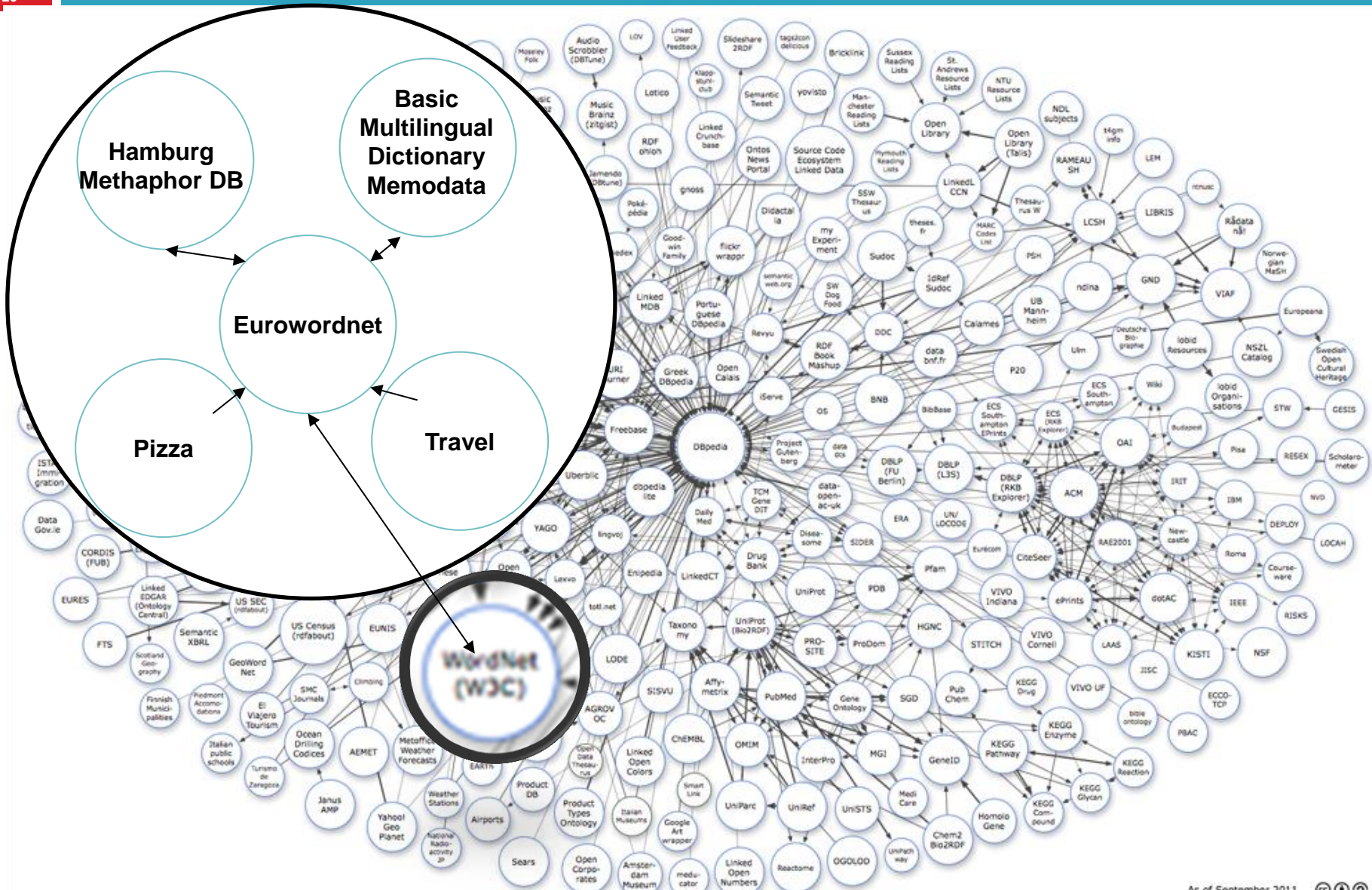
**eurowordnet-english-hyponymOf.rdf**

- WordNet already converted into OWL [van Assem et al., 2004] using the OWL-DL sublanguage.
- Conversion of EuroWordNet into RDF/OWL
  - To have a standardized OWL format.
  - To access it with standard OWL query tools.
  - To enrich it with additional domain-specific ontologies.
- Extensions
  - Adding domain ontologies (Pizza and Travel, EN) [GLDV2007]
  - Adding the Hamburg Metaphor Database (DE, FR) [LREC2008]
  - Adding the Basic Multilingual Dictionary MEMODATA (37655 entries, EN, DE, FR, ES, IT) [IGI-Book2012]

# RDF/OWL

## Lexical Linked Data Cloud

20



But how can we  
manage, interconnect and navigate  
this Lexical Linked Data Cloud?

# Overview

22

- Motivation
- EuroWordNet
- RDF/OWL EuroWordNet
- **RDF/OWL LexiRes Tool**
- Conclusions

## LexiRes Tool - Main Ideas

23

- Navigation of ontology hierarchy for:
  - restructuring it, by manual merging
  - using automatic merging functions
  - organizing the ontology/lexical resource structure
  - visualizing the ontology hierarchy
  
- Support the author in:
  - Using EuroWordNet
  - Exploring the lexical resource ontology hierarchy
  - Disambiguating the word senses of the search word
  - Giving translations of search word in different languages
  - Creating individual lexical collections
  - Adding and deleting meanings
  - Merging meanings
  - Importing OWL ontologies



# RDF/OWL

## LexiRes Tool - Browsing and Merging Word Senses

24

Applet-Ansicht: lexiresRDF.applet.EWN\_Applet.class

File Options Info

Language: english Word: bank Search Logged In Settings

Collection: adminCollection

Show Relation:
 

- Hypernyms
- Hyponyms
- Holonyms
- Meronyms
- Synonyms
- Antonyms
- Other Relations: Additional Relation

Show Translation:
 

- Translations
- german  french
- english  italian
- spanish

Details:
 

- Word: bank
- Part-of-Speech: noun
- Word Sense: 7
- Offset: 5731705
- Synonyms: camber, cant
- Gloss: a slope in the turn of a road

Status: Done.

Applet gestartet



Ernesto William De Luca. **Aggregation and Maintenance of Multilingual Linked Data** (Book Chapter). In *"Semi-Automatic Ontology Development: Processes and Resources"* Book. IGI Global. Paziienza and Stellato (eds.) February, 2012. Copyright © 2012. 340 pages. DOI: 10.4018/978-1-4666-0188-. Hershey, PA 17033, USA

# Overview

26

- Motivation
- EuroWordNet
- RDF/OWL EuroWordNet
- RDF/OWL LexiRes Tool
- **Conclusions**

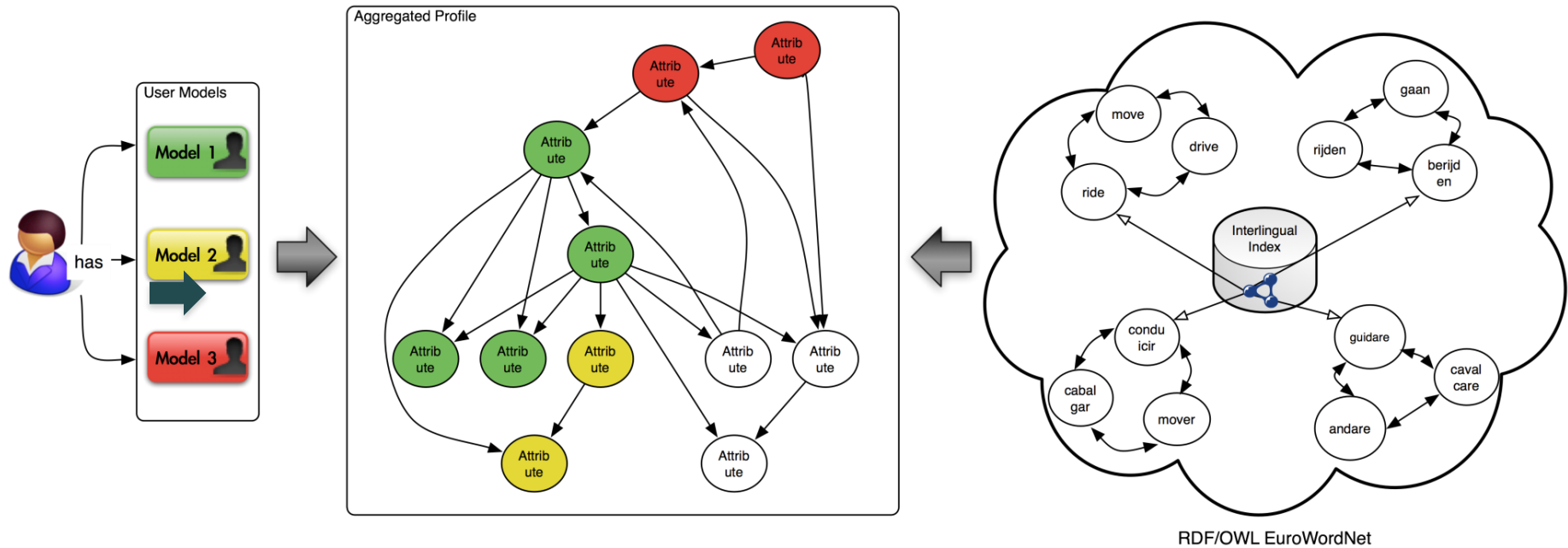
- **Need of tools supporting the user in creating data**
  - RDF/OWL EuroWordNet Browsing
  - Maintenance of Multilingual Resources
  - Customization of Multilingual Resources
  - Addition / deletion of word senses
  - Addition domain-specific ontologies
    - directly under its new hyperonym
  - Manual and Automatic Merging methods
- **Discussion about licenses of LR**

- The RDF/OWL LexiRes Tool
  - RDF/OWL EuroWordNet Browsing
  - Maintenance of Multilingual Resources
  - Customization of Multilingual Resources
  - Addition / deletion of word senses
  - Addition domain-specific ontologies
    - directly under its new hyperonym
  - Manual and Automatic Merging methods

# Conclusions

## Current and Future Work

29



# Conclusions

## Current and Future Work



Prof. Dr.-Ing.  
**Ernesto William De Luca**  
*Informationswissenschaft*

Fachbereich  
Informationswissenschaften  
Friedrich-Ebert-Str. 4  
14467 Potsdam

Tel. +49 331 580-1520  
Fax +49 331 580-1599  
deluca@fh-potsdam.de  
[http://iw.fh-potsdam.de/  
de\\_luca.html](http://iw.fh-potsdam.de/de_luca.html)  
<http://ernestodeluca.de>