



**KYOTO (ICT-211423)**

**Yielding Ontologies for Transition-Based Organization  
Intelligent Content and Semantics**

# **WordNet LMF**

**Monica Monachini – CNR-ILC**



Monica Monachini – 1° KYOTO Workshop –  
Amsterdam 2/3-2-2009



ICT-211423

# Outline

- Background: a KYOTO format for lexical resources
- WordNet-LMF
- The KYOTO Lexical Grid



# KYOTO: the lexical resource perspective

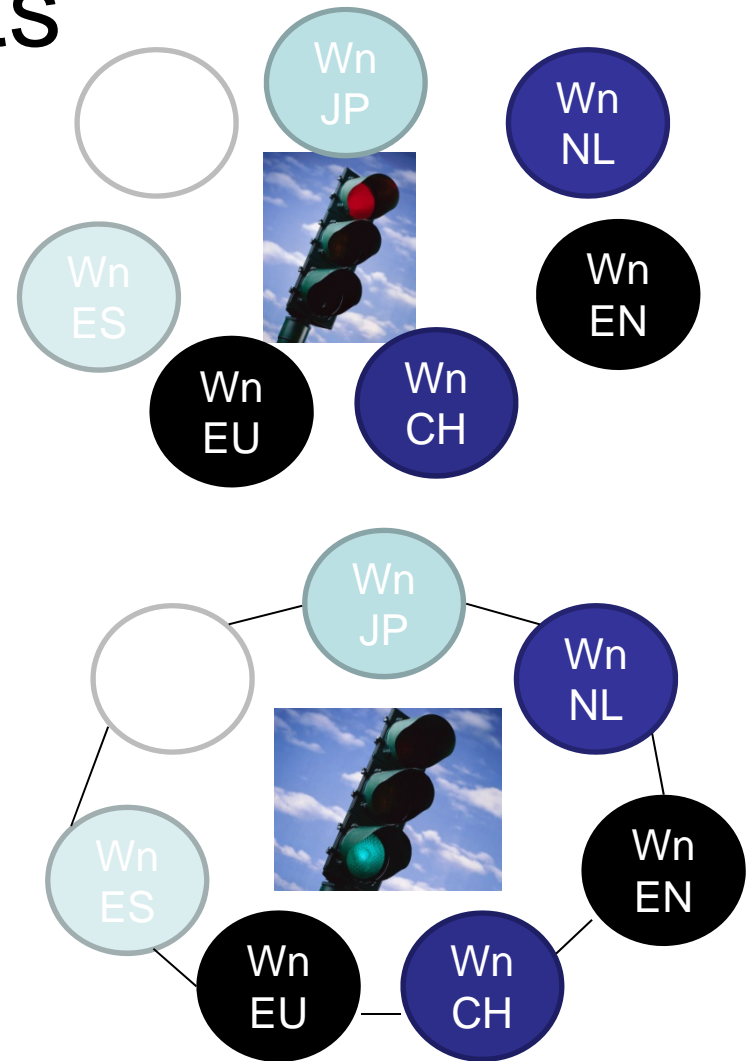
- KYOTO objectives
  - “ ... *facilitating the exchange of information across languages, domains and cultures*”
  - “ ... *allow definition of word meaning in a shared Wiki platform*”
- *from the point of view of linguistic resources ...*
  - needs to share lexical and knowledge bases, both general and domain-related, under the form of lexical repositories and ontologies



# A common representation format for WordNets

## Seven WordNets

- similar but not identical → hampered interoperability
- to be accessed both intra- and inter-linguistically → need to support easier integration
- endow WordNet with a representation format to allow easy access, integration and interoperability among resources



# Standard and Interoperability: SoA

## Achievements

- SubCommittee devoted to standards for linguistic annotation
- ✓ Catalogues of linguistic categories and annotation schemas
- Interest group (ACL) for developing standard annotation of language data
- ✓ Efforts towards interlinked resources
- ✓ Harmonized systems and frameworks
- International conferences/workshops
- EU-funded common resources and technology infrastructure; roadmap for achieving interoperability

## To be achieved

- ✓ Existing standards developed in isolation (not widely accepted)
- ✓ Disagreement concerning theories/linguistic annotation
- ✓ Lack of standard representation format(s)/framework(s)
- ✓ Lack of accessibility





# LMF

- Specifically designed to accommodate as many models of lexical representation as possible
- Its pros:
  - Meta-model: a high-level specification ISO24613
  - Data Category Registry: low-level specifications ISO12620



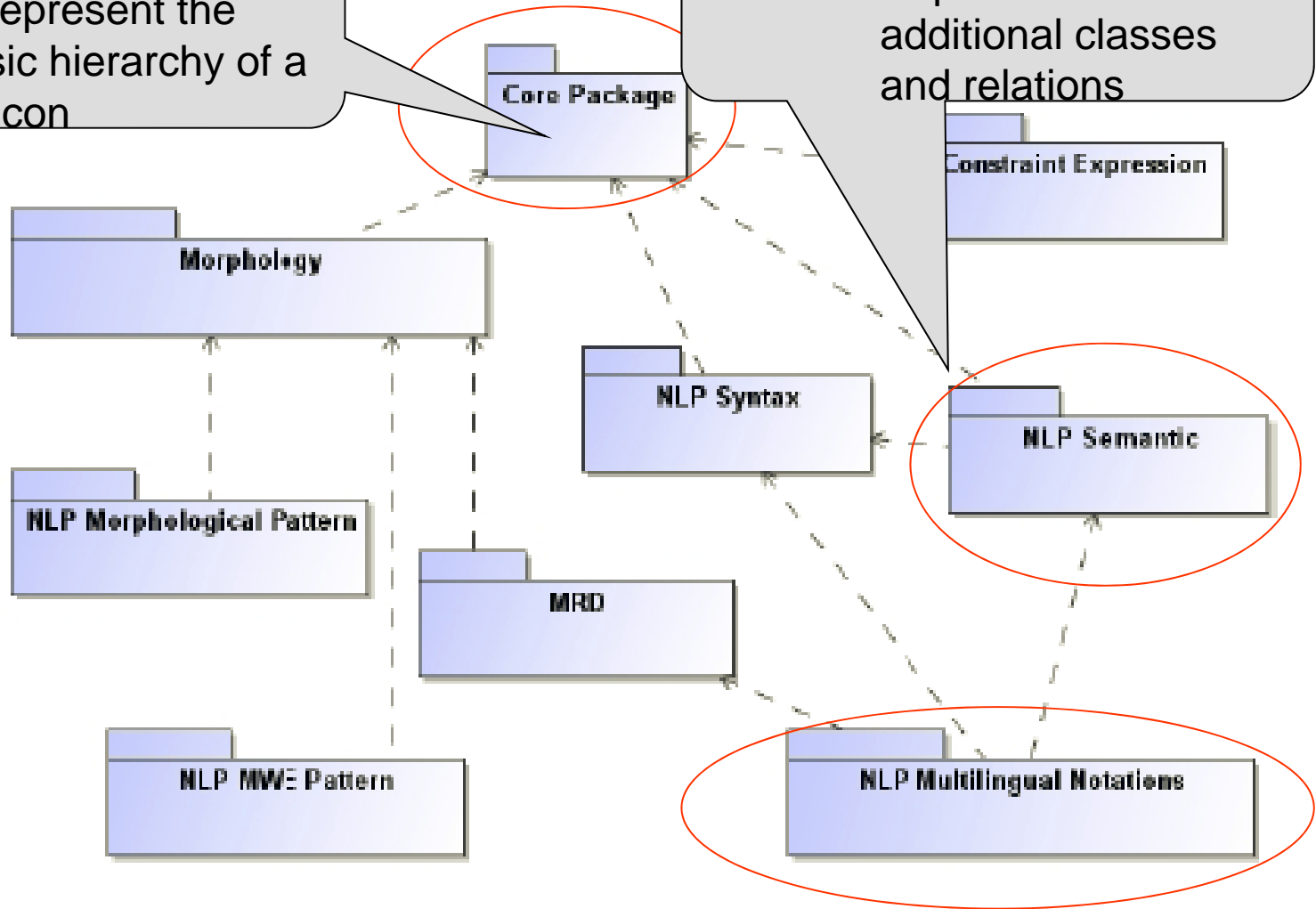
# Main Features

- Not a *monolithic* model rather a *modular* framework
  - LMF library provides the hierarchy of lexical objects (with structural relations among them)
  - Data Category Registry library provides descriptors to encode linguistic information associated to lexical objects (N.B. Data Categories can be also user-defined)



Structural skeleton  
to represent the  
basic hierarchy of a  
lexicon

Components  
required to describe  
additional classes  
and relations







# Centralized WordNet DC Registry

A list of 85 sem.rels as a result of a mapping of the KYOTO WordNet grid

	SWN	IWN
	<b>Inter-WN</b>	
		antonym
		antonym_comp
BE_IN_STATE	be_in_state	be_in_state
	category	
	category_term	
CAUSES	causes	causes
CO_AGENT_INSTRUMENT		co_agent_instrument
CO_AGENT_PATIENT		co_agent_patient
CO_AGENT_RESULT		co_agent_result
CO_INSTRUMENT_AGENT		co_instrument_agent
CO_INSTRUMENT_PATIENT		co_instrument_patient
CO_INSTRUMENT_RESULT		co_instrument_result
CO_PATIENT_AGENT		co_patient_agent
CO_PATIENT_INSTRUMENT		co_patient_instrument
CO_PATIENT_RESULT		co_patient_result
CO_RESULT_AGENT		co_result_agent
CO_RESULT_INSTRUMENT		co_result_instrument
CO_RESULT_PATIENT		co_result_patient
CO_ROLE		co_role
		for_purpose_of
FUZZYNYM	fuzzynym	fuzzynym
	gloss	
	has_derived	derivation
HAS_HOLO_LOCATION	has_holo_location	has_holo_location
HAS_HOLO_MADEOF	has_holo_madeof	has_holo_madeof
HAS_HOLO_MEMBER	has_holo_member	has_holo_member
HAS_HOLO_PART	has_holo_part	has_holo_part
HAS_HOLO_PORTION	has_holo_portion	has_holo_portion
HAS_HOLONYM	has_holonym	has_holonym

**Intra-WN**

EQ\_SYNONYM  
EQ\_NEAR\_SYNONYM  
EQ\_HAS\_HYPERNYM  
EQ\_HAS\_HYPONYM  
EQ\_INVOLVED  
EQ\_ROLE  
EQ\_IS\_CAUSED\_BY  
EQ\_CAUSES  
EQ\_HAS\_HOLONYM  
EQ\_HAS\_MERONYM  
EQ\_HAS\_SUBEVENT  
EQ\_IS\_SUBEVENT\_OF  
EQ\_BE\_IN\_STATE  
EQ\_IS\_STATE\_OF  
EQ\_CO\_ROLE  
EQ\_GENERALIZATION  
EQ\_METONYM  
EQ\_DIATHESIS



2003-2-25/11:31 AM



ICT-211423

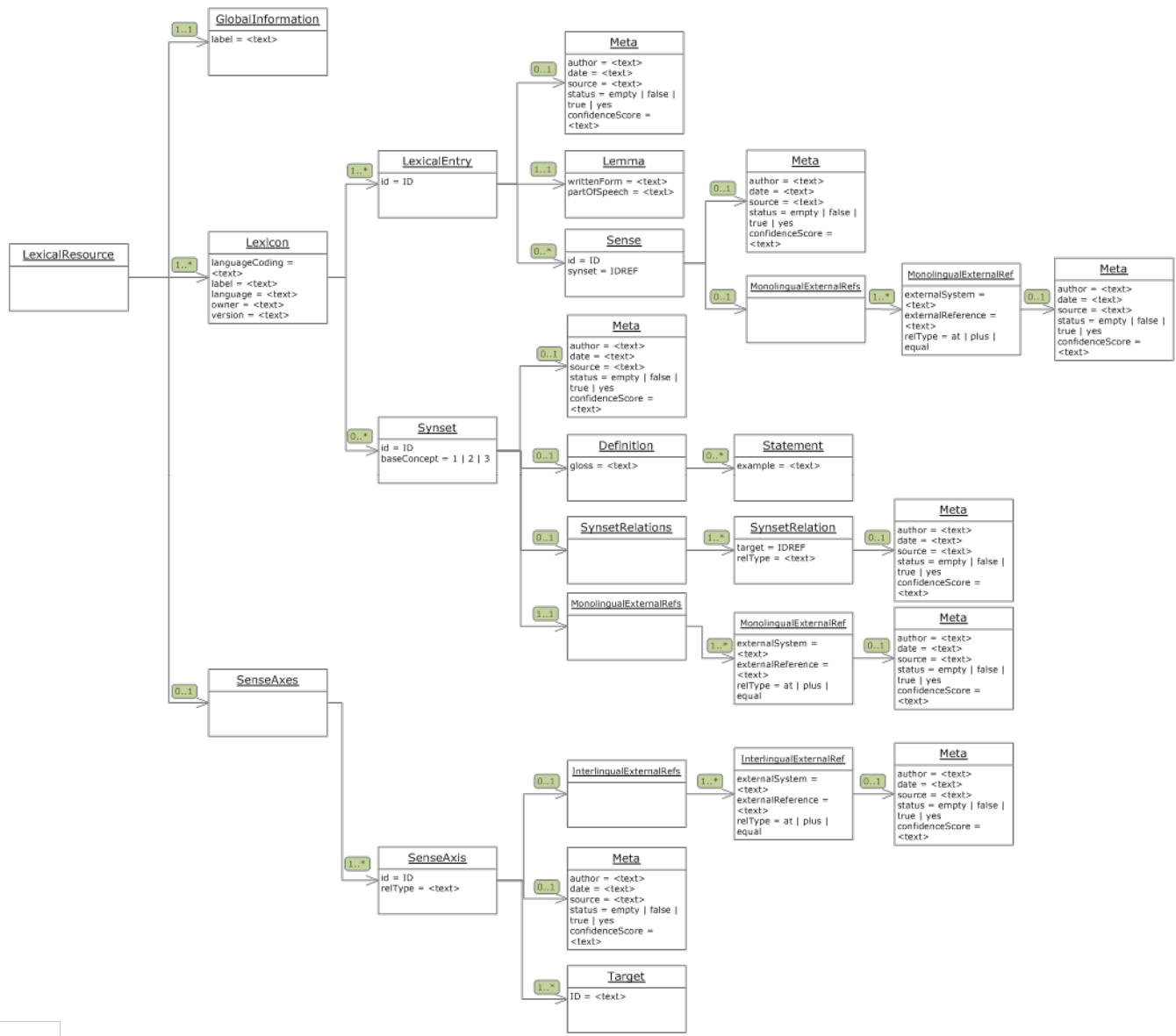
# Principles of WordNet-LMF

Balance between:

- Maintain adherence to architectural principles of LMF
  - Main conceptual building blocks and structural relationships between them maintained
  - The expression of the linguistic info (synset relations) falls in the realm of DCs
- Adapt standard LMF to suit efficiency needs
  - Promote feat-att structures to element attributes
  - Use of bracketing elements



# Kyoto-LMF UML Diagram



# An XML example

```
<SynsetRelation targets="ENG-16-06056130-n">
```

```
<feat att="cs" val="99"/>  
<feat att="status" val="yes"/>  
<feat att="source" val="whatsoever"/>  
<feat att="author" val="german"/>  
<feat att="date" val="2008-05-12"/>
```

```
</SynsetRelation>
```

Pure LMF

```
<SynsetRelation targets="EU-16-06056130-n"
```

```
relType="has_hyperonym" >
```

```
<Meta cs="99"  
status="yes"  
source="whatsoever"  
author="german"  
date="2008-05-12">
```

```
</SynsetRelation>
```

Kyoto-  
WordNet  
LMF



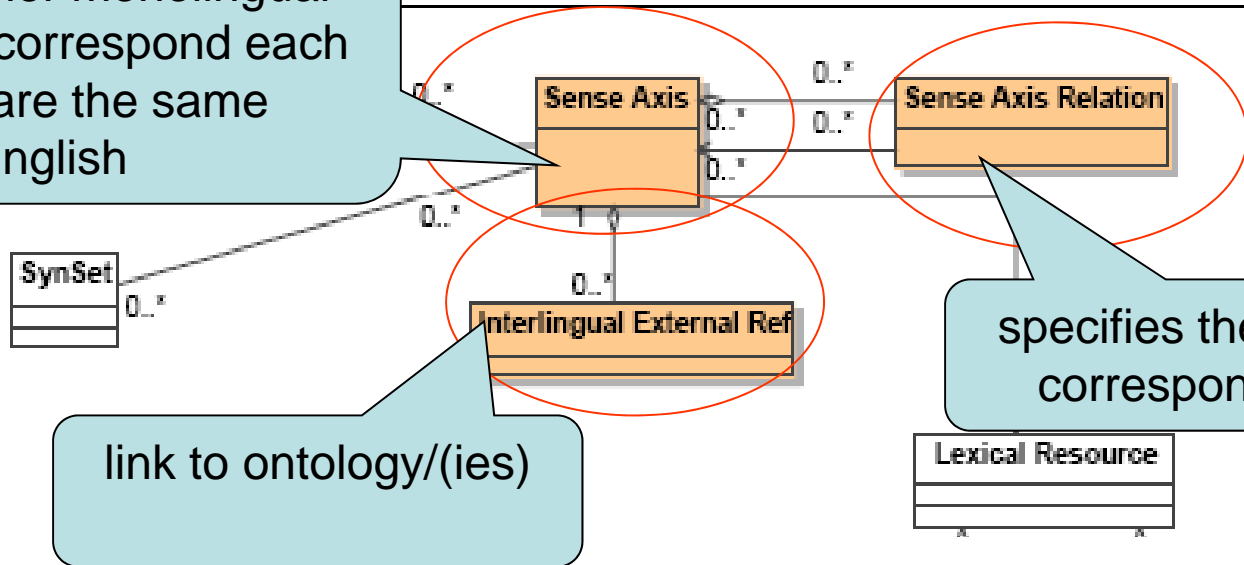
# A complete XML synset

```
<Synset id="eng-30-06645039-n" baseConcept="1">
<Definition gloss="mark of a foot or shoe on a surface">
<Statement example="the police made casts of the footprints in the soft earth outside the
window" />
</Definition>
<SynsetRelations>
<SynsetRelation target="eng-30-06798750-n" relType="has_hyperonym" >
<Meta author="AH" date="2008-07-01" source="Wordnet3.0" status="yes"
confidenceScore="1.0" />
</SynsetRelation>
<SynsetRelation target="eng-30-06645266-n" relType="has_hyponym" >
<Meta author="AH2" date="2008-07-01" source="eng-Wordnet3.0" status="yes"
confidenceScore="1.0" />
</SynsetRelation>
</SynsetRelations>
<MonolingualExternalRefs>
<MonolingualExternalRef externalSystem="Wordnet1.6" externalReference="eng-16-
01234567-n" />
<MonolingualExternalRef externalSystem="SUMO" externalReference="superficialPart"
relType="at"/>
</MonolingualExternalRefs>
```



# The multilingual extension

groups together monolingual synsets that correspond each other and share the same relations to English

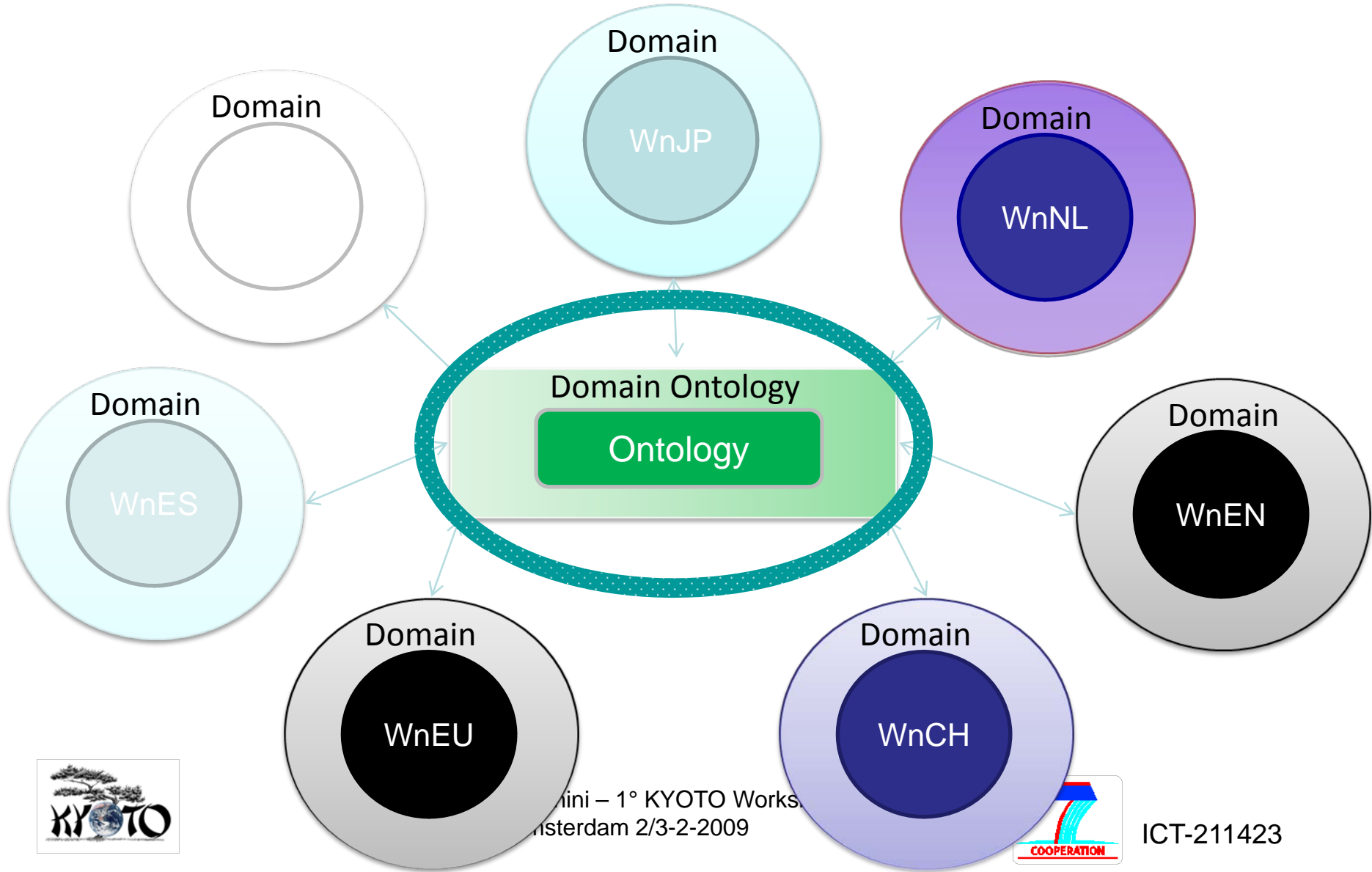


link to ontology/(ies)

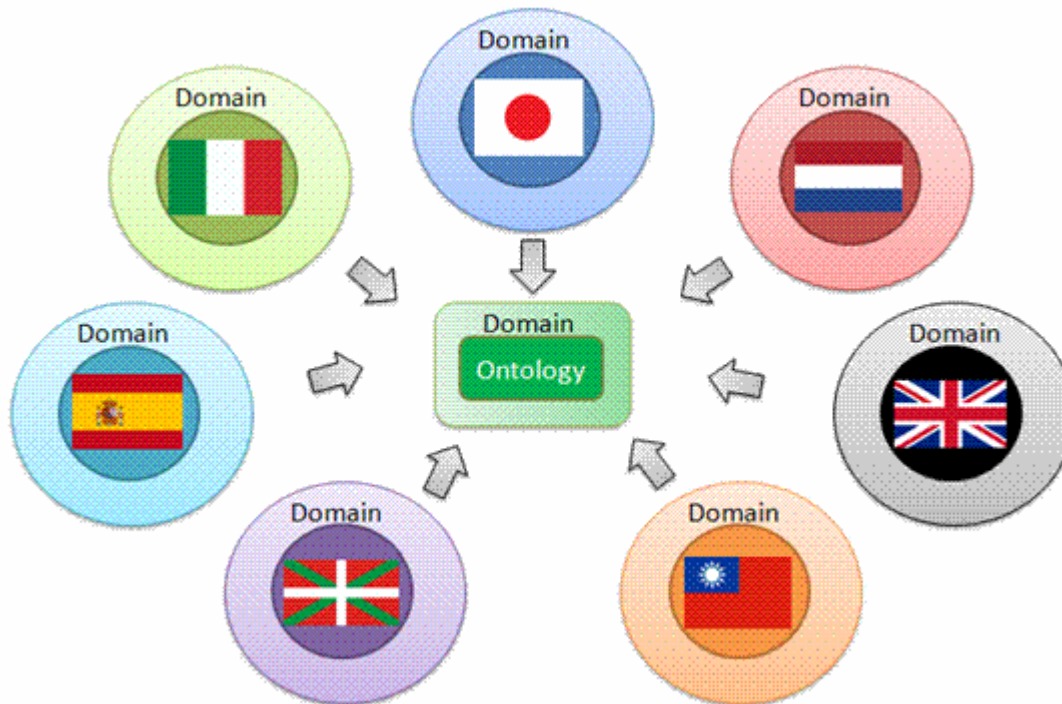
specifies the type of correspondence

```
<SenseAxis id="sa_ita16-spa30-zho30-eng30_001" relType="eq_synonym">  
<Target ID="ita-16-1251-n"/>  
<Target ID="spa-30-09686541-n"/>  
<Target ID="zho-30-05231501-n"/>  
<Target ID="eng-30-13480848-n"/>  
</SenseAxis>
```

# Kyoto Knowledge Base







# Thank you!

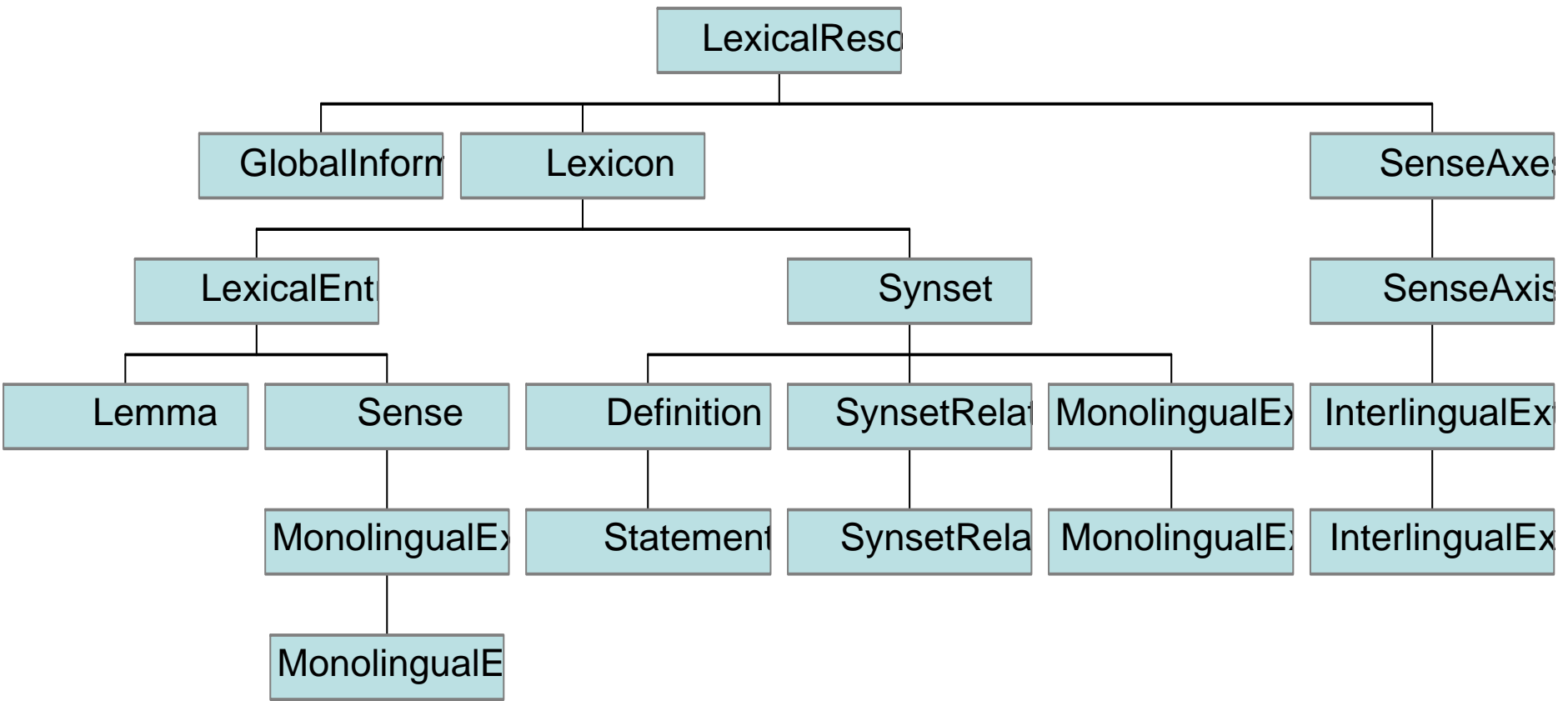


Monica Monachini – 1° KYOTO Workshop –  
Amsterdam 2/3-2-2009



ICT-211423

# General Scheme



# Kyoto Knowledge Base

