

# **XML Schema for Wordnet and Ontology DELIVERABLE NR. 1 /WP NR. 7**

Version 2.0  
1/07/2008

Andrea Marchetti – CNR IIT  
Francesco Ronzano – CNR IIT  
Maurizio Tesconi – CNR IIT  
Claudia Soria – CNR ILC  
Monica Monachini – CNR ILC  
Piek Vossen – VUA/IRION  
Wauter Bosma – VUA



**Knowledge Yielding Ontologies for Transition-based Organization**

**ICT 211423**





<b>Grant Agreement No.</b>	<b>ICT 211423</b>
<b>Project Acronym</b>	<b>KYOTO</b>
<b>Project full title</b>	<b>Knowledge Yielding Ontologies for Transition-based Organization</b>
<b>Technologies</b>	
<b>Funding Scheme</b>	<b>FP7 – ICT</b>
<b>Date latest version Annex I</b>	<b>19-12-2007</b>
<b>Project Coordinator</b>	<b>Prof. Dr. Piek T.J.M. Vossen VU University Amsterdam Tel. + 31 (0) 20 5986466 Fax. + 31 (0) 20 5986500 Email: <a href="mailto:p.vossen@let.vu.nl">p.vossen@let.vu.nl</a></b>
<b>Project website</b>	<b><a href="http://www.kyoto-project.eu/">http://www.kyoto-project.eu/</a></b>
<b>Deliverable Document Number</b>	<b>D7.1</b>
<b>Status</b>	<b>Draft/Final</b>
<b>Security (distribution level)</b>	<b>Public</b>
<b>Contractual date of delivery</b>	<b>May 31, 2008</b>
<b>Actual date of delivery</b>	
<b>Type</b>	<b>report/prototype/software/ontology/ wordnets/data/interface/demo/etc.</b>
<b>WP contributing to the deliverable</b>	<b>WP 7</b>
<b>WP responsible</b>	<b>Andrea Marchetti (CNR-IIT)</b>
<b>Authors</b>	<b>Andrea Marchetti (CNR-IIT) Francesco Ronzano (CNR-IIT) Maurizio Tesconi (CNR-IIT) Claudia Soria (CNR-ILC) Monica Monachini (CNR-ILC) Piek Vossen (VUA/IRION) Wauter Bosma – (VUA)</b>
<b>EC project officer</b>	<b>Werner Janusch</b>
<b>Keywords</b>	<b>XML Schema, Wordnet, Ontology, LMF, TMF, OWL, KIF</b>
<b>Abstract</b>	<b><i>This deliverable describes the XML schema adopted to represent all the data related to the management of the multi-language wordnets and the ontology; they constitute the set of linguistic and semantic resources of KYOTO system.</i></b>

## **Table of Content**

### **1.Introduction**

### **2.XML Schemata for *Wordnets representation***

#### **2.1 The KYOTO LMF Schema**

### **3.XML Schemata for *ontology representation***

### **4.XML Schemata for *candidate terms representation***

### **5.Appendix**

#### **5. 1 Appendix A: KYOTO LMF XML Schema Definition**

#### **5. 2 Appendix B: KYOTO TMF XML Schema Definition**

### **Bibliography**

## **List of Tables**

## **List of Figures**

**Figure 1 – Wordnets XML Schemata and KYOTO architecture**

## Introduction

This deliverable describes the XML schemata adopted to represent all the data related to the management of the multi-language wordnets and the ontology; they constitute the set of linguistic and semantic resources of KYOTO system.

We have chosen XML as data representation format because of the high level of interoperability assured as well as its wide adoption, especially for distributed Web systems like KYOTO. In particular, considering KYOTO architecture, the choice of XML to store and exchange data related to the wordnets and the ontology allows for an easier integration and interaction with external systems; in the meanwhile it provides the definition of a set of constraints over exchanged data thanks to the validation rules defined in the XML Schemata.

Part of the linguistic and the semantic resources of KYOTO needs to be collaboratively enriched and maintained by KYOTO community of users, mainly starting from the integration of a set of candidate terms derived from text processing procedures by so called Tybots (Term Yielding Robots). As a consequence we can define two general groups of data formats:

- *the wordnets and ontology representation formats* (the XML Schemata representing the content of the wordnets and the ontology);
- *the candidate terms representation formats* (the XML Schema collecting all the features of candidate terms to be integrated in the wordnets or in the ontology).

In this deliverable we detail all these XML Schemata. In the first part, we describe the representation formats of the wordnets and the ontology. For both these lexical and semantic resources, we introduce different XML Schemata to satisfy the representational needs of their distinct processing steps inside KYOTO system. While the exporting format of a resource needs to be standard-compliant in order to achieve strong interoperability, its corresponding storage and processing formats have to be optimized in order to assure efficient parsing, validation and retrieval functions.

In order to ease the exchange of data related to the wordnets and to the ontology with external systems, we have decided to conform as much as possible our XML Schemata to the widely adopted standards. We have considered the Lexical Markup Framework model, a common backbone for the creation, the use and the exchange of data between and among different multilingual lexical resources (LMF, ISO/TC37) for wordnets, thus defining the KYOTO LMF Format. For the ontology we considered the Web Ontology Language (OWL) and the Knowledge Interchange Format (KIF) as alternatives for an ontology representation language. In particular, the discussion about the right ontology to exploit in KYOTO is still ongoing (WP6). Likewise, we have still to define the right choice between OWL and KIF, thus choosing the right XML Schemata to adopt for our ontological reference. We pay particular attention on the ways of linking different wordnets of distinct languages as well as the way of connecting wordnets entries with the corresponding ontological ones.

In the second part of this deliverable, we focus our attention on the description of the XML Schema adopted to represent the candidate terms extracted from documents. It is referred to as the KYOTO TMF Format because it is derived from the top level of the Terminological Markup Framework (TMF, ISO16642/2001), an international standard designed in the framework of the ISO initiatives to support the creation and use of computer applications for terminological data.

More details on the XML representations and our choices can be found in the KYOTO working papers and reports:

- TR002, WP02: C.Soria, M. Monachini: "Kyoto-LMF WordNet representation format", Version 3, June 23, 2008
- TR003, WP02: A. Marchetti, F. M. Tesconi, S. Minutoli: "Formalizing Knowledge by Ontologies: OWL and KIF", Version 2, May. 29, 2008
- TR004, WP02: P. Vossen, W. Bosma: "The Representation of Terms", Version 1, June 23, 2008
- Report M. Monachini "TMF and LMF for raw terms", April 24, 2008

This deliverable gives a short overview of the result of the discussion in the working papers.



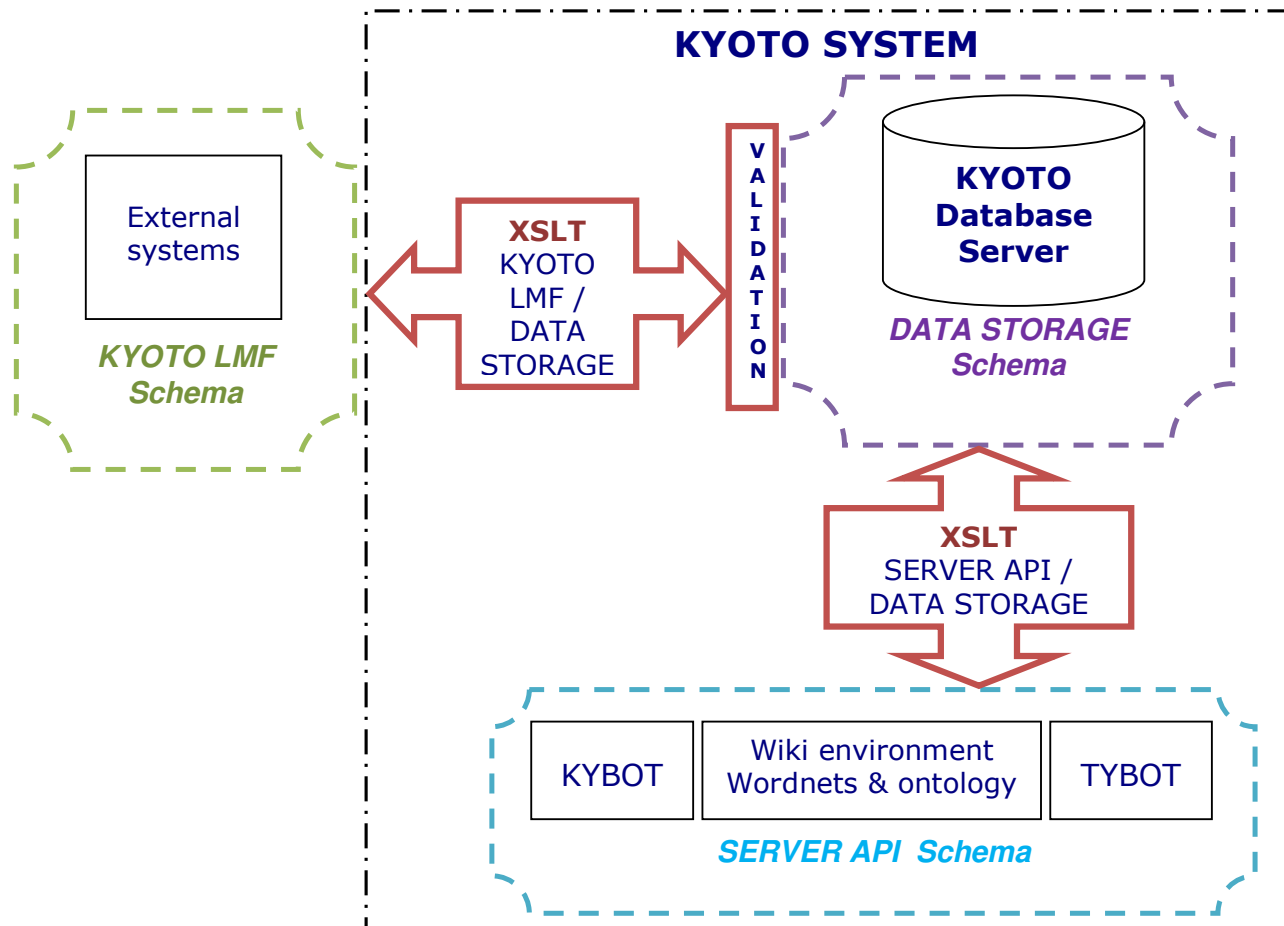
## 1. XML Schemata for *wordnets* representation

The grid of multi-language wordnets represents the core set of lexical resources of KYOTO system. In this section we describe the XML Schemata exploited to interact with it providing some illustration.

As briefly stated in the introductory section of this deliverable, we have decided to provide different wordnet representation Schemata tailored to the representational needs of distinct processing phases. In particular, as shown in Figure 1, we have identified three Schemata to be used in three different contexts:

- *KYOTO LMF Schema*: it defines the format of lexical data to be exchanged with external systems. It has been derived with some minor modification from the Document Type Definition (DTD) of the Lexical Markup Framework, version 16 [1], a common backbone for the creation, the use and the exchange of data between and among different multilingual lexical resources (LMF, ISO/TC37).
- *DATA STORAGE Schema*: it is quite similar to the KYOTO LMF Schema but with some important change mainly concerning the distribution of information among different XML attributes or elements. This is the schema adopted internally to the XML database that is intended to store the lexical resources. These changes, even if in some case they represent relevant deviations from the LMF standard, are necessary in order to improve the efficiency of database information storage and indexing as well as to allow for better data validation.
- *SERVER API Schema*: it defines the format of the data exchanged between the KYOTO Database Server and the other components of KYOTO system, querying the Web Application Program Interface exposed by the same Server. Among the components of KYOTO interacting with the KYOTO Database Server through this API there are the Tybots, the Kybots (Knowledge Yielding Robots) and the Wiki environment for editing domain wordnets and ontology. Thanks to the *SERVER API Schema* we essentially want to standardize the representation of data sent and received during Web interactions with the KYOTO Database Server.

The data transformations from one specific Schema to another are carried out thanks to suitable XSLT stylesheets.



**Figure 1 – Wordnets XML Schemata and KYOTO architecture**

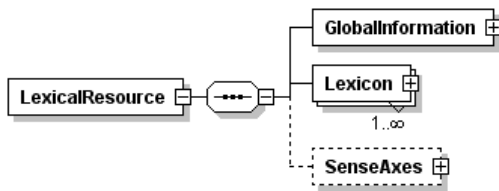
In the following section, we describe the KYOTO LMF Schema. Then we detail all the modifications that need to be done over this Schema in order to transform the data in their storage format. The definition of the DATA STORAGE Schema and of the SERVER API Schema will be developed in the next phase of the project after the architecture of the KYOTO Database Server and its Application Program Interface are completely defined.

## 2.1 The KYOTO LMF Schema

In this section we describe all the elements of KYOTO LMF Schema. The full definition of the schema can be found in appendix

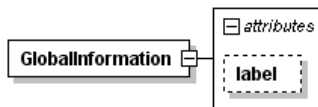
### LexicalResource

LexicalResource is the root element, as in LMF. A lexical resource can contain more than one lexicon, and inter-lingual correspondences are grouped in the section SenseAxes, separated from the lexical resources proper, containing only inter-lexicon correspondences.



### GlobalInformation

This is used to record general information about the lexical resource. The attribute “label” is a free text field.



Attributes					
Name	Type	Use	Default	Fixed	Annotation
label	xs:string				

Example:

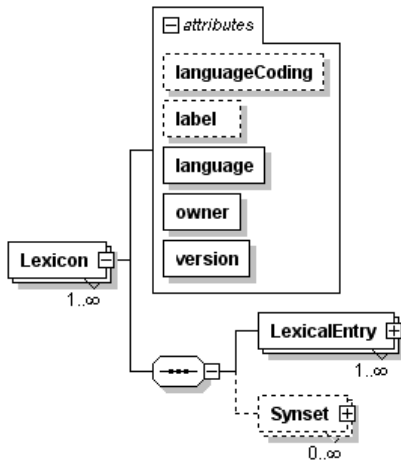
```
<GlobalInformation label="Proposal for Kyoto-internal WordNet representation"/>
```

## Lexicon

A monolingual resource. Attribute 'languageCoding' has "ISO 639-3" as a fixed value. This standard uses 3 lowercase letters to code the language (e.g. eng, nld), to be specified in the 'language' attribute.

Attributes 'owner' and 'version' are used to declare copyright holder and resource version, respectively. 'label' is an optional attribute for recording any additional information that may be needed.

The *Lexicon* element has two child elements, *LexicalEntry* and *Synset*.



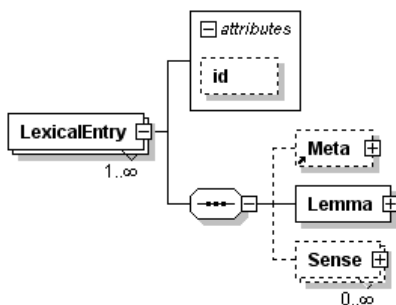
Name	Type	Use	Default	Fixed	Annotation
languageCoding	<b>xs:string</b>			ISO 639-3	
label	<b>xs:string</b>				
language	<a href="#">languageType</a>	required			
owner	<b>xs:string</b>	required			
version	<b>xs:string</b>	required			

Example:

```
<Lexicon languageCoding="ISO 639-3" label="English Wordnet 1.6, Meaning" language="eng"
owner="Princeton" version="16">
```

## LexicalEntry

A container for representing a lexeme in a lexicon. A *LexicalEntry* element can contain one lemma and zero to many different senses. It has one attribute: 'id' (a unique identifier).



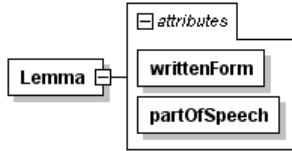
Name	Type	Use	Default	Fixed	Annotation
id	<b>xs:ID</b>				

Example:

```
<LexicalEntry id="Department_of_Justice">
```

## Lemma

It represents a word form chosen by convention to designate the lexical entry. Attribute 'partOfSpeech' is attributed to *Lemma*, in conformance with LMF, and takes as its value the part-of-speech value that is in general specified for a synset.



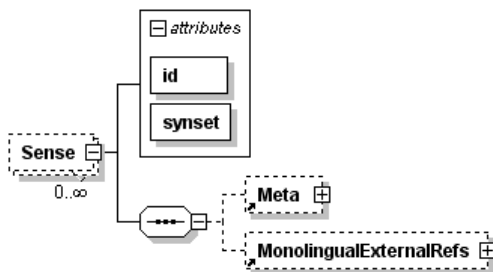
Name	Type	Use
writtenForm	<b>xs:string</b>	required
partOfSpeech	<a href="#">partOfSpeechType</a>	required

Example

```
<Lemma writtenForm="Department_of_Justice" partOfSpeech="n" />
```

## Sense

This element represents one meaning of a lexical entry. For wordnet representation, it represents the variant (or literal) of a synset. The element *Sense* can contain zero to one *Meta* elements and zero to one *MonolingualExternalRefs* elements. Concerning use and meaning of this latter, refer to the corresponding section.



Name	Type	Use	Description
id	<b>xs:ID</b>	required	Id must be specified according to the convention used in WordNet, i.e. word_sense#nr.
synset	<b>xs:IDREF</b>	required	Synset takes as its value the ID of the synset to which the sense belongs

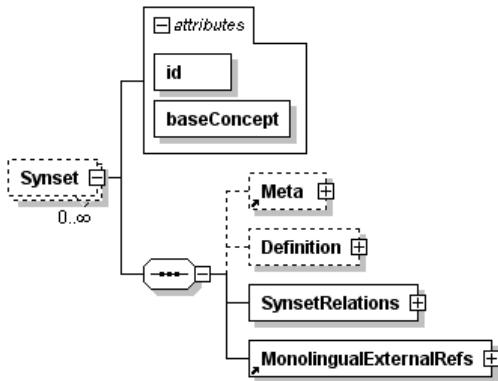
Example:

```

<Sense id="Department_of_Justice_1" synset="eng-16-06060223-n">
  <MonolingualExternalRefs>
    <MonolingualExternalRef externalSystem="Wordnet3.0" externalReference="justice%1:14:00::"/>
  </MonolingualExternalRefs>
</Sense>
  
```

## Synset

This element encodes information about a wordnet synset. A *Synset* element can link senses of different *LexicalEntry* instances within the same part of speech. Synset elements can contain zero to one *Meta*, zero to one *Definition*, one *SynsetRelations* and one *MonolingualExternalRefs* bracketing elements. Concerning use and meaning of this latter, refer to the corresponding section.



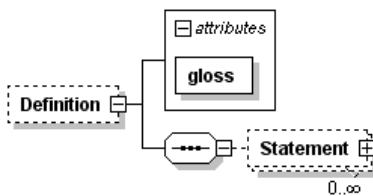
Name	Type	Use	Description
id	<a href="#">synsetIdType</a>	required	The agreed syntax is "language code-version-id-pos tag"
baseConcept	<a href="#">baseConceptType</a>	required	values for the baseConcept attribute will be numerical (1, 2, 3) which correspond to the BaseConcept sets

Example:

```
<Synset id="eng-16-06060223-n" baseConcept="1">
  <Meta author="Piek Vossen" date="2008-05-12"/>
  <Definition gloss="...">
    <Statement example="..."/>
  </Definition>
  <SynsetRelations>...</SynsetRelations>
  <MonolingualExternalRefs>...</MonolingualExternalRefs>
</Synset>
```

## Definition

It allows to represent the gloss associated with each synset by means of the obligatory attribute 'gloss'.



Name	Type	Use
gloss	<b>xs:string</b>	required

Example:

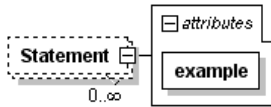
```

<Definition gloss="a motor vehicle with four wheels; usually propelled by an internal combustion
engine">
  <Statement example="he needs a car to get to work"/>
</Definition>

```

## Statement

*Statement* allows to represent examples of use associated with the synset by means of the attribute 'example'.



Name	Type	Use
example	<b>xs:string</b>	required

Example:

```

<Definition gloss="a motor vehicle with four wheels; usually propelled by an internal combustion
engine">
  <Statement example="he needs a car to get to work"/>
</Definition>

```

## SynsetRelations

Bracket element used to collect all synset relations.



Example

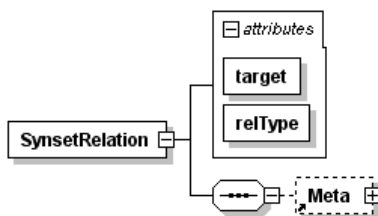
```

<SynsetRelations>
  <SynsetRelation target="eng-16-06056130-n" relType="has_hyperonym"/>
  <SynsetRelation target="eng-16-06060479-n" relType="has_mero_part"/>
  <SynsetRelation target="eng-16-00403152-n" relType="gloss"/>
</SynsetRelations>

```

## SynsetRelation

It codifies a relation between two synsets.



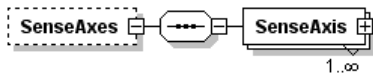
Name	Type	Use	Description
target	<a href="#">synsetIdType</a>	required	synsetId: the ID value of the synset that is target of the relation
relType	<a href="#">internalRelationType</a>	required	synset relation type : such as hyponym, hyperonym

Example:

```
<SynsetRelation target="eng-16-06056130-n" relType="has_hyperonym">
  <Meta author="German Rigau" date="2008-05-12" status="yes" source="whatsoever"
confidenceScore="0.99"/>
</SynsetRelation>
```

### SenseAxes

A bracket element.

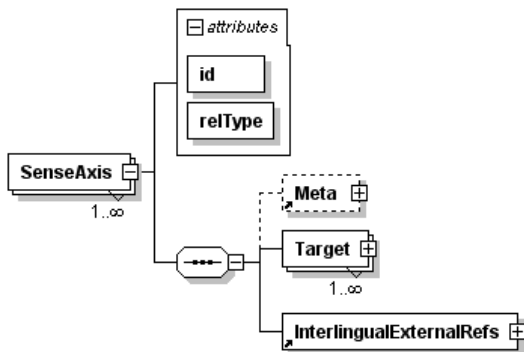


Example

```
<SenseAxes>
  <SenseAxis id="sa_en16-en30_001" relType="equal_synonym">...</SenseAxis>
  <SenseAxis id="sa_en16-en30_002" relType="equal_near_synonym">...</SenseAxis>
</SenseAxes>
```

### SenseAxis

Represents the relationship among different closely related senses in different languages. In wordnet terms, it encodes ILI correspondences. Any SenseAxis element groups together monolingual synsets that correspond one to another by means of a particular type of relation, specified by means of the 'relType' attribute.



Name	Type	Use
id	<b>xs:ID</b>	required
relType	<a href="#">externalRelationType</a>	required

Example



```

<SenseAxis id="sa_en16-en30_001" relType="equal_synonym">
  <Meta author="Monica" date="2008-05-27"/>
  <Target ID="eng-16-06060223-n"/>
  <Target ID="eng-30-08135342-n"/>
  <InterlingualExternalRefs>
    <InterlingualExternalRef externalSystem="SUMO"externalReference="PoliticalProcess" relType="at">
      <Meta author=" Claudia Soria " date="06-06-2008"/>
    </InterlingualExternalRef>
  </InterlingualExternalRefs>
</SenseAxis>

```

## Target

The monolingual synset ID that is referenced by each SenseAxis. One target for each synset.



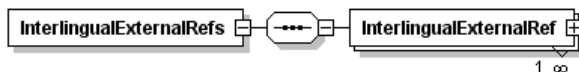
Name	Type	Use
ID	<a href="#">synsetIdType</a>	required

Example

```
<Target ID="eng-16-06060223-n"/>
```

## InterlingualExternalRefs

Bracket element used to collect all InterlingualExternalRef elements.



used by SenseAxis

Example

```

<InterlingualExternalRefs>
  <InterlingualExternalRef externalSystem="SUMO"externalReference="PoliticalProcess" relType="at">
    <Meta author=" Claudia Soria " date="06-06-2008"/>
  </InterlingualExternalRef>
</InterlingualExternalRefs>

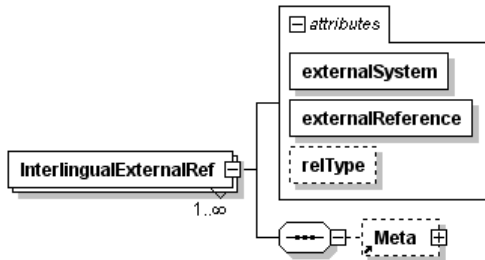
```

## InterlingualExternalRef

It represents a relationship between a SenseAxis instance and an external system, that is an ontology.

The 'externalSystem' and 'externalReference' recommended attributes allow to encode, respectively, the name of the external system and the specific relevant nodes in the given external system.

InterlingualExternalRefs connect the group of synset to an ontological concept.



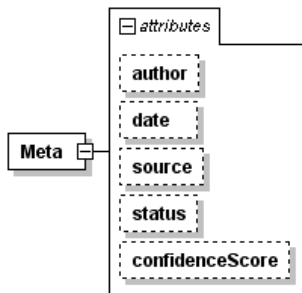
Example

```
<InterlingualExternalRef externalSystem="SUMO" externalReference="PoliticalProcess" relType="at">
  <Meta author="Claudia Soria" date="06-06-2008"/>
</InterlingualExternalRef>
```

Name	Type	Use	Description
externalSystem	<b>xs:string</b>	required	The name of the external system or ontology
externalReference	<b>xs:string</b>	required	A reference to a node of the external system, or a concept of the ontology
relType	<a href="#">ontologyRelationType</a>	optional	Used only if the externalSystem is an ontology

Meta

Meta-data information about the creation (date and author), the source, the status and the confidence score.



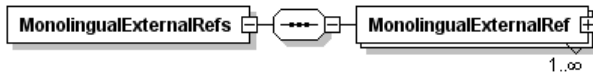
Used by: LexicalEntry, MonolingualExternalRef, Sense, Synset, SynsetRelation

Name	Type	Use	Description
author	<b>xs:string</b>	Optional	
date	<b>xs:string</b>	Optional	
source	<b>xs:string</b>	Optional	expresses the originating database/system. It is typically associated with SynsetRelation elements.
status	<b>statusType</b>	Optional	a key expressing editing status of the parent element. Possible values are empty (=not confirmed), false (wrong to be deleted) or true (confirmed as ok) and sometimes yes (confirmed as ok)
confidenceScore	<b>rateType</b>	Optional	a numeric value indicating the degree of certainty about a given element. Typically, it is specified for SynsetRelation and MonolingualExternalRef elements

```
<Meta author="German Rigau" date="2008-05-12" status="yes" source="whatsoever"
confidenceScore="0.99"/>
```

## MonolingualExternalRefs

A bracket element grouping together references from a monolingual database to one or more external resources.



Used by: Sense, Synset

```
<MonolingualExternalRefs>
  <MonolingualExternalRef externalSystem="Wordnet 3.0" externalReference="justice%1:14:00::"/>
</MonolingualExternalRefs>
```

## MonolingualExternalRef

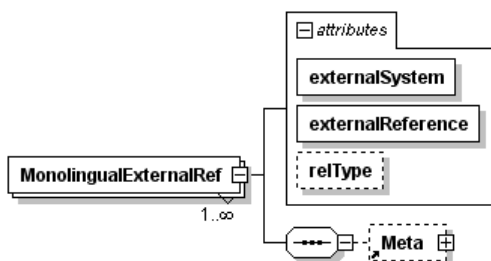
This element can be used to encode any reference or correspondence to an external resource. Its use is defined by slightly different conventions according to the particular section (Sense or Synset) in which it appears.

When occurring inside the *Sense* section, it can be used to express mapping between a sense and its correspondent in another lexical resource. As a recommendation, use *MonolingualExternalRef* to express exact mapping between a sense and its correspondent in another lexical resource of the same language. In the particular case of English WordNet it can also serve as a representational device to express *SenseKey* value.

When occurring inside the *Synset* element, it allows to encode reference to the domain and/or one or more links to an ontological system.

The *MonolingualExternalRef* element has two required attributes, 'externalSystem' and 'externalReference', and the optional attribute 'relType'. The required attributes are used to express, respectively, the name of the external resource and the particular identifier or node. Possible values of the 'externalSystem' attribute are, for instance, 'domain', 'SuperSense', 'SUMO', 'TCO' (= Top Concept Ontology), and 'WordNet3.0' (for recording *SenseKey* values).

The attribute 'relType' serves to specify relations with nodes in SUMO ontology. Possible values are "at", "plus", "equal".



Name	Type	Use	Description
externalSystem	<a href="#">externalSystemType</a>	required	The name of the external resource, for instance, 'domain', 'SuperSense', 'SUMO', 'TCO' (= Top Concept Ontology), and 'WordNet3.0'
externalReference	<b>xs:string</b>	required	the particular identifier or node
relType	<a href="#">ontologyRelationType</a>	optional	Used only if the externalSystem is SUMO

Example:

```
<MonolingualExternalRef externalSystem="Domain" externalReference="administration"/>
<MonolingualExternalRef externalSystem="Domain" externalReference="law"/>
```

```

<MonolingualExternalRef externalSystem="SUMO" externalReference="PoliticalProcess" relType="at"/>
<MonolingualExternalRef externalSystem="TCO" externalReference="Agentive"/>
<MonolingualExternalRef externalSystem="TCO" externalReference="Purpose"/>
<MonolingualExternalRef externalSystem="TCO" externalReference="Social"/>
<MonolingualExternalRef externalSystem="TCO" externalReference="UnboundedEvent"/>

```

**simpleType baseConceptType**

type	restriction of <b>xs:unsignedInt</b>						
used by	attribute <a href="#">LexicalResource/Lexicon/Synset/@baseConcept</a>						
facets	<table border="1"> <tr><td>enumeration</td><td>1</td></tr> <tr><td>enumeration</td><td>2</td></tr> <tr><td>enumeration</td><td>3</td></tr> </table>	enumeration	1	enumeration	2	enumeration	3
enumeration	1						
enumeration	2						
enumeration	3						

*List of base concept sets: 1,2,3*

**simpleType externalRelationType**

type	restriction of <b>xs:string</b>																																												
used by	attribute <a href="#">LexicalResource/SenseAxes/SenseAxis/@relType</a>																																												
facets	<table border="1"> <tr><td>enumeration</td><td>eq_synonym</td></tr> <tr><td>enumeration</td><td>eq_near_synonym</td></tr> <tr><td>enumeration</td><td>eq_has_hyponym</td></tr> <tr><td>enumeration</td><td>eq_has_hyponym</td></tr> <tr><td>enumeration</td><td>eq_involved</td></tr> <tr><td>enumeration</td><td>eq_role</td></tr> <tr><td>enumeration</td><td>eq_is_caused_by</td></tr> <tr><td>enumeration</td><td>eq_causes</td></tr> <tr><td>enumeration</td><td>eq_has_holonym</td></tr> <tr><td>enumeration</td><td>eq_has_meronym</td></tr> <tr><td>enumeration</td><td>eq_has_subevent</td></tr> <tr><td>enumeration</td><td>eq_is_subevent_of</td></tr> <tr><td>enumeration</td><td>eq_be_in_state</td></tr> <tr><td>enumeration</td><td>eq_is_state_of</td></tr> <tr><td>enumeration</td><td>eq_co_role</td></tr> <tr><td>enumeration</td><td>eq_generalization</td></tr> <tr><td>enumeration</td><td>eq_metonym</td></tr> <tr><td>enumeration</td><td>eq_diathesis</td></tr> <tr><td>enumeration</td><td>eq_in_manner</td></tr> <tr><td>enumeration</td><td>eq_has_instance</td></tr> <tr><td>enumeration</td><td>eq_belongs_to_class</td></tr> <tr><td>enumeration</td><td>eq_antonym</td></tr> </table>	enumeration	eq_synonym	enumeration	eq_near_synonym	enumeration	eq_has_hyponym	enumeration	eq_has_hyponym	enumeration	eq_involved	enumeration	eq_role	enumeration	eq_is_caused_by	enumeration	eq_causes	enumeration	eq_has_holonym	enumeration	eq_has_meronym	enumeration	eq_has_subevent	enumeration	eq_is_subevent_of	enumeration	eq_be_in_state	enumeration	eq_is_state_of	enumeration	eq_co_role	enumeration	eq_generalization	enumeration	eq_metonym	enumeration	eq_diathesis	enumeration	eq_in_manner	enumeration	eq_has_instance	enumeration	eq_belongs_to_class	enumeration	eq_antonym
enumeration	eq_synonym																																												
enumeration	eq_near_synonym																																												
enumeration	eq_has_hyponym																																												
enumeration	eq_has_hyponym																																												
enumeration	eq_involved																																												
enumeration	eq_role																																												
enumeration	eq_is_caused_by																																												
enumeration	eq_causes																																												
enumeration	eq_has_holonym																																												
enumeration	eq_has_meronym																																												
enumeration	eq_has_subevent																																												
enumeration	eq_is_subevent_of																																												
enumeration	eq_be_in_state																																												
enumeration	eq_is_state_of																																												
enumeration	eq_co_role																																												
enumeration	eq_generalization																																												
enumeration	eq_metonym																																												
enumeration	eq_diathesis																																												
enumeration	eq_in_manner																																												
enumeration	eq_has_instance																																												
enumeration	eq_belongs_to_class																																												
enumeration	eq_antonym																																												

*List of external relation types*

**simpleType externalSystemType**

type	restriction of <b>xs:NMTOKEN</b>										
used by	attribute <a href="#">MonolingualExternalRefs/MonolingualExternalRef/@externalSystem</a>										
facets	<table border="1"> <tr><td>enumeration</td><td>SUMO</td></tr> <tr><td>enumeration</td><td>Domain</td></tr> <tr><td>enumeration</td><td>TCO</td></tr> <tr><td>enumeration</td><td>SuperSense</td></tr> <tr><td>enumeration</td><td>Wordnet3.0</td></tr> </table>	enumeration	SUMO	enumeration	Domain	enumeration	TCO	enumeration	SuperSense	enumeration	Wordnet3.0
enumeration	SUMO										
enumeration	Domain										
enumeration	TCO										
enumeration	SuperSense										
enumeration	Wordnet3.0										

*List of External Resource*

simpleType **internalRelationType**

type restriction of **xs:string**  
 used by attribute [LexicalResource/Lexicon/Synset/SynsetRelations/SynsetRelation/@relType](#)  
 facets

enumeration	antonym
enumeration	antonym_comp
enumeration	be_in_state
enumeration	category
enumeration	category_term
enumeration	causes
enumeration	co_agent_instrument
enumeration	co_agent_patient
enumeration	co_agent_result
enumeration	co_instrument_agent
enumeration	co_instrument_patient
enumeration	co_instrument_result
enumeration	co_patient_agent
enumeration	co_patient_instrument
enumeration	co_patient_result
enumeration	co_result_agent
enumeration	co_result_instrument
enumeration	co_result_patient
enumeration	co_role
enumeration	for_purpose_of
enumeration	fuzzynym
enumeration	gloss
enumeration	has_derived
enumeration	has_holo_location
enumeration	has_holo_madeof
enumeration	has_holo_member
enumeration	has_holo_part
enumeration	has_holo_portion
enumeration	has_holonym
enumeration	has_hyperonym
enumeration	has_hyponym
enumeration	has_mero_location
enumeration	has_mero_madeof
enumeration	has_mero_member
enumeration	has_mero_part
enumeration	has_mero_portion
enumeration	has_meronym
enumeration	has_pertainym
enumeration	has_subevent
enumeration	has_xpos_hyperonym
enumeration	has_xpos_hyponym
enumeration	in_manner
enumeration	instance
enumeration	involved
enumeration	involved_agent
enumeration	involved_direction
enumeration	involved_instrument
enumeration	involved_location
enumeration	involved_patient
enumeration	involved_result
enumeration	involved_source_direction
enumeration	involved_target_direction
enumeration	is_a_value_of
enumeration	is_caused_by
enumeration	is_derived_from

enumeration	is_subevent_of
enumeration	manner_of
enumeration	near_antonym
enumeration	near_synonym
enumeration	nearest
enumeration	pertains_to
enumeration	region
enumeration	region_term
enumeration	related
enumeration	related_to
enumeration	results_in
enumeration	rgloss
enumeration	role
enumeration	role_agent
enumeration	role_direction
enumeration	role_instrument
enumeration	role_location
enumeration	role_manner
enumeration	role_patient
enumeration	role_result
enumeration	role_source_direction
enumeration	role_target_direction
enumeration	see_also_wn15
enumeration	state_of
enumeration	usage
enumeration	usage_term
enumeration	verb_group
enumeration	xpos_fuzzynym
enumeration	xpos_near_antonym
enumeration	xpos_near_synonym

*List of internal relation types*

#### simpleType languageType

type	restriction of <b>xs:string</b>
used by	attribute <a href="#">LexicalResource/Lexicon/@language</a>
facets	enumeration eus enumeration eng enumeration ita enumeration jpn enumeration nld enumeration spa enumeration zho

*List of all languages code from ISO 639-3: three lower case letters: eng, ita, eus, zho, jpn, spa, nld*

#### simpleType ontologyRelationType

type	restriction of <b>xs:NMTOKEN</b>
used by	<a href="#">MonolingualExternalRefs/MonolingualExternalRef/@relType</a> <a href="#">InterlingualExternalRefs/InterlingualExternalRef/@relType</a>
facets	enumeration equal enumeration plus enumeration at

*List of SUMO relation types: at, plus, equal*

simpleType **partOfSpeechType**

type restriction of **xs:string**  
 used by attribute [LexicalResource/Lexicon/LexicalEntry/Lemma/@partOfSpeech](#)  
 facets enumeration n  
 enumeration v  
 enumeration a  
 enumeration r  
 enumeration s

*List of Pos type: n, v, a (adjective), r (adverb), s (satellite adjective)*

simpleType **synsetIdType**

type restriction of **xs:string**  
 used by [LexicalResource/Lexicon/Synset/@id](#)  
[LexicalResource/SenseAxes/SenseAxis/Target/@ID](#)  
[LexicalResource/Lexicon/Synset/SynsetRelations/SynsetRelation/@target](#)  
 facets pattern [a-z](3)-[dd]-d-[n,v,a,r,s]

*Format of a synset ID: language code-version-id-pos tag*

simpleType **versionType**

type restriction of **xs:unsignedInt**  
 facets

enumeration	15
enumeration	16
enumeration	17
enumeration	18
enumeration	19
enumeration	20
enumeration	21
enumeration	30

*List of wordnet version...*

## 2 XML Schemata for ontology representation

The ontology of KYOTO system provides the language-independent semantic backbone of the whole architecture. The discussion about the best ontology language to adopt and thus the right XML Schema to represent it is still ongoing; this choice is also strictly related to the definition of what specific ontology to exploit and consequently what is the best language to express its contents.

In any way, we have restricted our choice about the ontology representation format to two possible languages: the *Web Ontology Language* (OWL) [2] and the *Knowledge Interchange Format* (KIF) [3].

The *Web Ontology Language* is widely used to describe ontologies over the Web. Its current version is 1.0, standardized at the beginning of 2004 as the outcome of the Web Ontology Working Group of the World Wide Web Consortium (W3C). It is one of the fundamental languages of the Semantic Web, an ongoing effort to enrich the Web with a set of metadata in order to make information directly machine processable. OWL is built on other two core languages of the Semantic Web: the Resource Description Framework (RDF) [4] and RDF Schema (RDFS) [5]; OWL provides additional vocabulary for describing properties and classes of an ontology. It is based on the expressive power of a particular family of description logics. Since September 2007, starting from the great amount of extensions requirements collected, a new W3C Working Group, the OWL Working Group [6], has been constituted in order to formalize a new and enriched OWL version: OWL 2.0. Along with the process of standardization of OWL 1.0, also an XML presentation syntax has been defined [7]; it describes an XML dialects useful to express the information contained in OWL ontologies. Thus a set of XML Schemata for OWL has been specified. If OWL will be chosen as the ontology representation language, these standard XML Schemata can be adopted to make KYOTO system exchange ontology information with other external ones. In the same way as we have done with the representation of Wordnets, we can define also slightly different internal XML Schemata for ontology representation, optimized for storage and indexing of data.

An alternate way to express ontology-related data is represented by the Knowledge Interchange Format (KIF). It is a draft proposed by the American National Standard (dpANS - NCITS.T2/98-004) to describe knowledge among different computer systems so as to facilitate its exchange. It has a declarative semantics and has been defined keeping into consideration the goals of translatability, readability and implementability. It is based on the first order logic and has a specific syntax to express the characteristics of ontological entries. As far as we know, there not exist any XML Schema to express KIF data. Thus if KIF will be chosen as the KYOTO ontology representation language, we will define a proper Schema for its representation. An interesting way to represent KIF expressions through a set of triplets has been adopted in the Cornetto Project [8] (Dutch Stevin Project – STE05039, ending in Summer 2008). Even if it is not possible to represent complex KIF expressions through triplets, the structure of each triplet can be easily expressed in XML format. In this context, we can also adopt the Resource Description Framework, heavily based on the notion of triplet as well as its XML Syntax to represent KIF triplets. In such a situation we can define a sort of ontology or RDF Schema useful to represent the information related to each kind of KIF triplet defined in the Cornetto Project. Among the alternate ontologies that can be exploited in KYOTO system, there is the Suggested Upper Merged Ontology (SUMO) [9]. It has been written exploiting KIF expressions, but also an OWL version of it is available. Probably it is not complete with all the data contained in the KIF version, also because OWL expressivity is not so extensive as those of KIF.

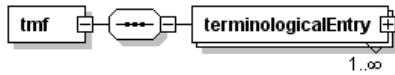


## 3 XML Schemata for *candidate terms representation*

In this section we describe all the elements of KYOTO TMF Schema.

### tmf

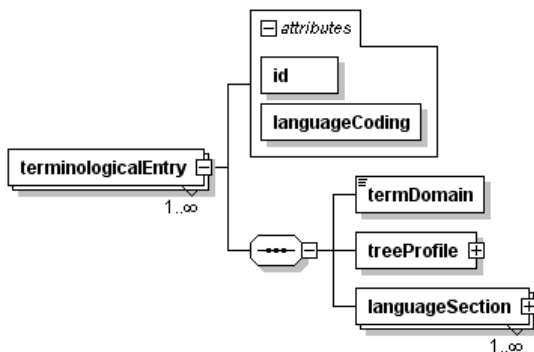
The root is identical to the TMF proposal.



```
<tmf>
  <terminologicalEntry id="t001" languageCoding="ISO-693-3">...</terminologicalEntry>
  <terminologicalEntry id="t002" languageCoding="ISO-693-3">...</terminologicalEntry>
</tmf>
```

### terminologicalEntry

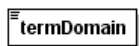
An entry containing information on terminological units.



```
<terminologicalEntry id="t001" languageCoding="ISO-693-3">...</terminologicalEntry>
```

### termDomain

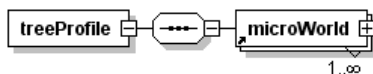
The domain of the term.



```
<termDomain>environment</termDomain>
```

### treeProfile

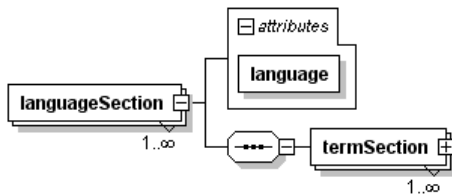
Classification of the complete term tree.



```
<treeProfile>
  <microWorld score="0.88">Topography</microWorld>
  <microWorld score="0.72">Finance</microWorld>
  <microWorld score="0.7">Bio</microWorld>
</treeProfile>
```

## languageSection

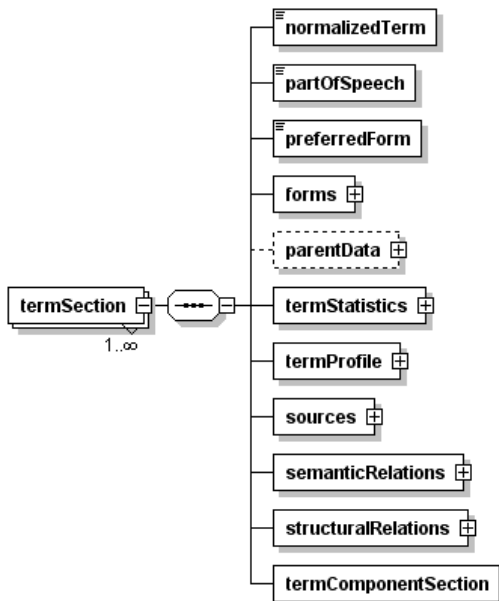
The part of a terminological entry containing information related to one language.



```
<languageSection language="eng">
  <termSection> </termSection>
  <termSection> </termSection>
</languageSection>
```

## termSection

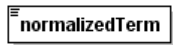
The part of a language section giving information about a term.



```
<termSection>
  <normalizedTerm>climate change</normalizedTerm>
  <partOfSpeech>n</partOfSpeech>
  <preferredForm>Climate Change</preferredForm>
  <forms>...</forms>
  <parentData>...</parentData>
  <termStatistics>...</termStatistics>
  <sources>...</sources>
  <semanticRelations>...</semanticRelations>
  <structuralRelations>...</structuralRelations>
  <termComponentSection>...</termComponentSection>
</termSection>
```

## normalizedTerm

Unifies different variants of the term. It may be omitted if the term identifier is used.



```
<normalizedTerm>climate change</normalizedTerm>
```

## partOfSpeech

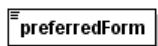
Part of speech of the term, should be the same for all variants.



```
<partOfSpeech>n</partOfSpeech>
```

## preferredForm

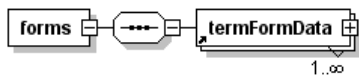
Canonical form for representation purposes. It is used for labelling the term for the user in an interface.



```
<preferredForm>Climate Change</preferredForm>
```

## forms

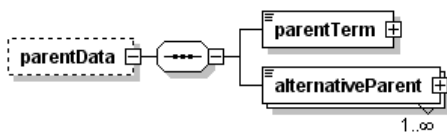
List of term forms with frequency and pointers to positions in SemAF files.



```
<forms>
  <termFormData id="tf_1" frequency="3">
    <termForm>climate change</termForm>
    <spans docId="1234">...</spans>
  </termFormData>
  <termFormData id="tf_2" frequency="6">
    <termForm>climate changes</termForm>
    <spans docId="134">...</spans>
  </termFormData>
</forms>
```

## parentData

Structural parent relation for establishing the term tree. It's optional because the root term has no parent.

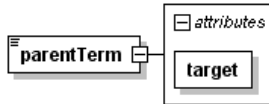


```
<parentData>
  <parentTerm target="t13"/>
  <alternativeParent>...</alternativeParent>
  <alternativeParent>...</alternativeParent>
</parentData>
```

```
</parentData>
```

## parentTerm

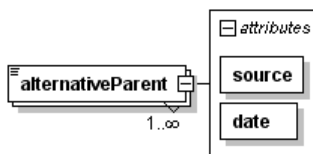
It contains the parent term id into the target attribute.



```
<parentTerm target="t13"/>
```

## alternativeParent

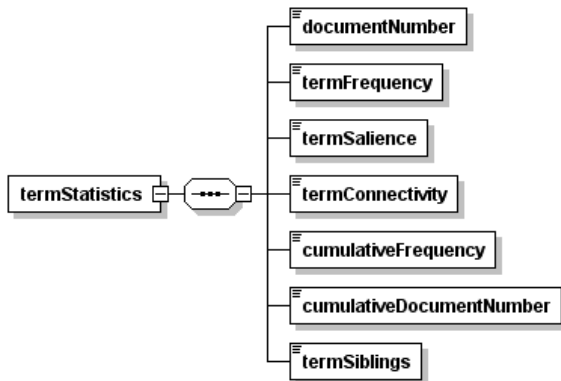
Alternative parents extracted from external sources.



```
<alternativeParent source="http://en.wikipedia.org" date="2008-06-20">Climate change
feedbacks and causes</alternativeParent>
```

## termStatistics

Statistics at the term level.



```
<termStatistics>
  <documentNumber>5</documentNumber>
  <termFrequency>13</termFrequency>
  <termSalience>0.04</termSalience>
  <termConnectivity>13</termConnectivity>
  <cumulativeFrequency>18</cumulativeFrequency>
  <cumulativeDocumentNumber>5</cumulativeDocumentNumber>
  <termSiblings>3</termSiblings>
</termStatistics>
```

## documentNumber

The number of source documents from which the term is derived.

`documentNumber`

```
<documentNumber>5</documentNumber>
```

## termFrequency

The total number of occurrences in source documents.

`termFrequency`

```
<termFrequency>13</termFrequency>
```

## termSaliency

A statistic parameter with a range value between 0 and 1.

`termSaliency`

```
<termSaliency>0.04</termSaliency>
```

## termConnectivity

The number of connections of a term class in the tree.

`termConnectivity`

```
<termConnectivity>13</termConnectivity>
```

## cumulativeFrequency

The sum of termFrequency of all descendants.

`cumulativeFrequency`

```
<cumulativeFrequency>18</cumulativeFrequency>
```

## cumulativeDocumentNumber

The sum of documentNumber of all descendants.

`cumulativeDocumentNumber`

```
<cumulativeDocumentNumber>5</cumulativeDocumentNumber>
```

## termSiblings

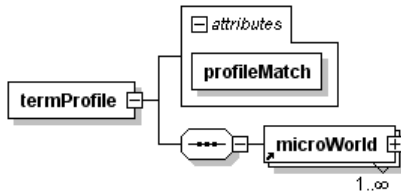
The number of terms that share the same parent.

`termSiblings`

<termSiblings>3</termSiblings>

## termProfile

The termProfile is similar to the treeProfile except that it has an additional attribute profileMatch that indicates the overlap of microWorld values across the termProfile and the treeProfile.

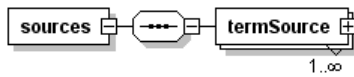


```

<termProfile profileMatch="0.69">
  <microWorld score="0.88">Geography</microWorld>
  <microWorld score="0.75">Finance</microWorld>
  <microWorld score="0.73">Metereology</microWorld>
  <microWorld score="0.7">Society</microWorld>
</termProfile>
  
```

## sources

List of type of sources sections (Toc, Body) from which the term is extracted.

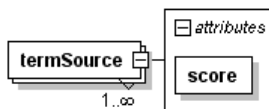


```

<sources>
  <termSource score="0.2">TOC</termSource>
  <termSource score="0.8">BODY</termSource>
</sources>
  
```

## termSource

The source section is a list of controlled values: the table of content (TOC) the body (BODY)...

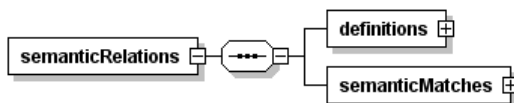


```

<termSource score="0.2">TOC</termSource>
  
```

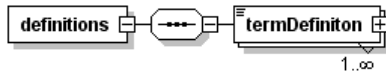
## semanticRelations

List of relations to external semantic resources that can be definitions or semantic matches.



## definitions

List of definitions extracted from external semantic resources.



<definitions>

<termDefiniton source="http://en.wikipedia.org/wiki/Climate\_change" date="2008-06-20">

Climate change is any long-term significant change in the "average weather" that a given region experiences. Average weather may include average temperature, precipitation and wind patterns. It involves changes in the variability or average state of the atmosphere over durations ranging from decades to millions of years. These changes can be caused by dynamic process on Earth, external forces including variations in sunlight intensity, and more recently by human activities.

</termDefiniton>

<termDefiniton source="googleSnippets" date="2008-06-20">

factors such as climate changes affecting our oceans

</termDefiniton>

<termDefiniton source="googleSnippets" date="2008-06-20">

environniental problems such as climate changes or acid rains

</termDefiniton>

<termDefiniton source="googleSnippets" date="2008-06-20">

global environmental issues such as climate changes

</termDefiniton>

<termDefiniton source="googleSnippets" date="2008-06-20">

environmental changes such as climate changes

</termDefiniton>

<termDefiniton source="googleSnippets" date="2008-06-20">

related activities such as climate changes and changes in land use pattern  
explanatory events such as climate changes

</termDefiniton>

<termDefiniton source="googleSnippets" date="2008-06-20">

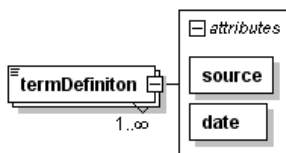
all kinds of other geographical data such as climate changes, plant growth,  
radiation, rainfall, forest fires

</termDefiniton>

</definitions>

## termDefiniton

A term definition extracted from an external semantic resource



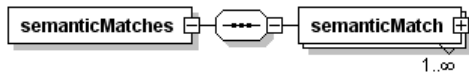
<termDefiniton source="googleSnippets" date="2008-06-20">

all kinds of other geographical data such as climate changes, plant growth,  
radiation, rainfall, forest fires

</termDefiniton>

## semanticMatches

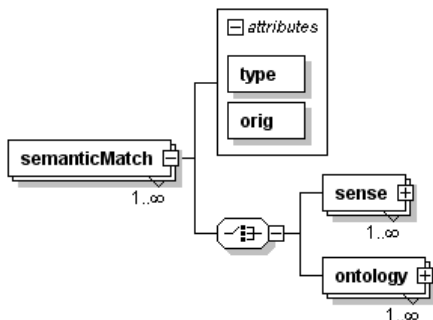
List of mapping to external semantic resource. This semantic layer summarizes the best synset mappings from all occurrences



```
<semanticMatches>
  <semanticMatch type="senseAlt" orig="urn:wordnet1.7">
    <sense source="EHU-WSD1" sensecode="ENG30-00180570-n" weight="0.80" />
    <sense source="EHU-WSD1" sensecode="ENG30-00290564-n" weight="0.30" />
  </semanticMatch>
  <semanticMatch type="ontologyAlt" orig="urn:sumo">
    <ontology source="EHU-WSD1" class="Process" weight="0.65" />
    <ontology source="EHU-WSD1" class="NaturalProcess" weight="0.70" />
  </semanticMatch>
</semanticMatches>
```

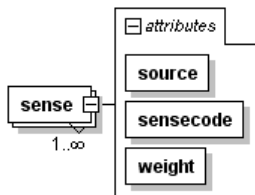
## semanticMatch

A mapping with a sense or an ontological entry. A separate match to the ontology should be added since it may resolve fine-grained ambiguities at the synset level.



```
<semanticMatch type="senseAlt" orig="urn:wordnet1.7">
  <sense source="EHU-WSD1" sensecode="ENG30-00180570-n" weight="0.80" />
  <sense source="EHU-WSD1" sensecode="ENG30-00290564-n" weight="0.30" />
</semanticMatch>
```

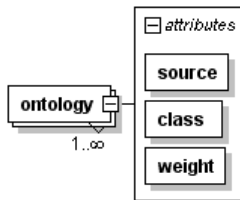
## sense



```
<sense source="EHU-WSD1" sensecode="ENG30-00180570-n" weight="0.80" />
```



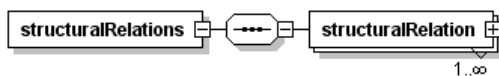
## ontology



```
<ontology source="EHU-WSD1" class="Process" weight="0.65" />
```

## structuralRelations

List of structural relations.



```

<structuralRelations>
  <structuralRelation syntaxRole="leftnp" semanticRole="">
    <syntaxElement />
    <termFormData id="tf_23" frequency="1">
      <termForm>sayan</termForm>
      <!-- dep is a higher level of the layered SEMAF notation (wordSpan) -->
      <deps docId="1824"><dep from="t3" to="t4" /></deps>
      <termContext />
    </termFormData>
  </structuralRelation>
  <structuralRelation syntaxRole="leftnp" semanticRole="">
    <syntaxElement />
    <termFormData id="tf_89" frequency="1">
      <termForm>location</termForm>
      <deps docId="1824"><dep from="t6" to="t7" /></deps>
      <termContext />
    </termFormData>
  </structuralRelation>
  <!-- Based on the term structure tables these can be extended to provide richer data -->
  <!-- PP occurring to the right of the term as NP -->
  <structuralRelation syntaxRole="np_right_pp" semanticRole="LOCATION">
    <syntaxElement>in</syntaxElement>
    <termFormData id="t65" frequency="2">
      <termForm>tropical area</termForm>
      <deps docId="124"><dep from="t6" to="t8" /></deps>
    </termFormData>
    <termFormData id="t68" frequency="1">
      <termForm>marine area</termForm>
      <deps docId="124"><dep from="t10" to="t11" /></deps>
      <termContext>climate change in marine areas</termContext>
    </termFormData>
  </structuralRelation>
  <structuralRelation syntaxRole="np_right_pp" semanticRole="TIME">
    <syntaxElement>from</syntaxElement>
    <termFormData id="t256" frequency="1">
      <termForm>1970</termForm>
      <deps docId="124"><dep from="t6" to="t8" /></deps> chang
  
```

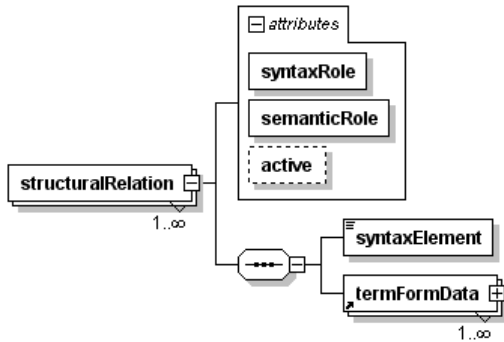
```

    <termContext>climate change from 1970 to 2005</termContext>
  </termFormData>
</structuralRelation>
<!-- NP to the left of the term as PP -->
<structuralRelation syntaxRole="pp_left_np" semanticRole="CAUSE">
  <syntaxElement>Of</syntaxElement>
  <termFormData id="t597" frequency="1">
    <termForm>impact</termForm>
    <deps docId="124"><dep from="t13" to="t15" /></deps>
    <termContext>impact of climate change</termContext>
  </termFormData>
</structuralRelation>
<!-- Term is the subject of the main verb in an ACTIVE sentence -->
<structuralRelation syntaxRole="subj" active="true" semanticRole="PATIENT">
  <syntaxElement />
  <termFormData id="t33" frequency="1">
    <termForm>accelerate</termForm>
    <deps docId="124"><dep from="t20" to="t28" /></deps>
    <termContext>climate change accelerated</termContext>
  </termFormData>
</structuralRelation>
<!-- Term is the subject of the main verb in an ACTIVE sentence -->
<structuralRelation syntaxRole="subj" active="true" semanticRole="AGENT">
  <syntaxElement />
  <termFormData id="t11" frequency="1">
    <termForm>cause</termForm>
    <deps docId="124"><dep from="t56" to="t63" /> </deps>
    <termContext>climate change causes a decline of biodiversity</termContext>
  </termFormData>
</structuralRelation>
<!-- Term is the subject of the main verb in a PASSIVE sentence -->
<structuralRelation syntaxRole="subj" active="false" semanticRole="PATIENT">
  <syntaxElement />
  <termFormData id="t11" frequency="1">
    <termForm>cause</termForm>
    <deps docId="124"><dep from="t356" to="t359" /> </deps>
    <termContext>climate change is caused by an increase in industrial
    activity</termContext>
  </termFormData>
</structuralRelation>
<!-- Term is modified by an adjective or adverb -->
<structuralRelation syntaxRole="mod" semanticRole="ATTRIBUTE">
  <syntaxElement />
  <termFormData id="t111" frequency="1">
    <termForm>rapid</termForm>
    <deps docId="124"><dep from="t356" to="t359" /></deps>
    <termContext>rapid climate changes</termContext>
  </termFormData>
</structuralRelation>
</structuralRelations>

```

## structuralRelation

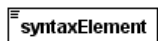
It groups all occurrences of a morpho-syntactic pattern in which a term occurred.



```

<structuralRelation syntaxRole="np_right_pp" semanticRole="LOCATION">
  <syntaxElement>in</syntaxElement>
  <termFormData id="t65" frequency="2">
    <termForm>tropical area</termForm>
    <deps docId="124"><dep from="t6" to="t8" /></deps>
  </termFormData>
  <termFormData id="t68" frequency="1">
    <termForm>marine area</termForm>
    <deps docId="124"><dep from="t10" to="t11" /></deps>
    <termContext>climate change in marine areas</termContext>
  </termFormData>
</structuralRelation>
  
```

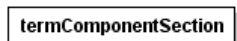
## syntaxElement



```
<syntaxElement>in</syntaxElement>
```

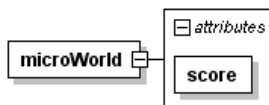
## termComponentSection

The section of a term section giving information about components of a term.



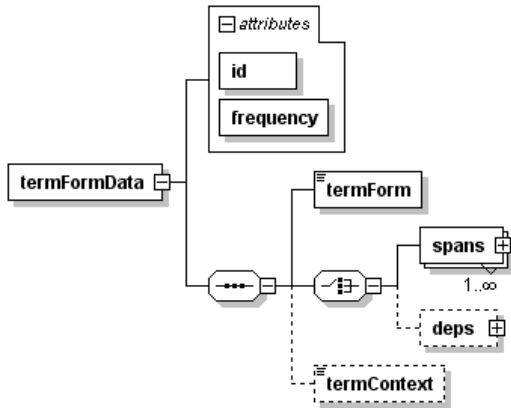
## microWorld

It contains a domain label with a confidence score.



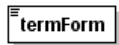
## termFormData

Used for listing all the different appearances of the term in the source documents with pointers to the locations in the SemAF notation.



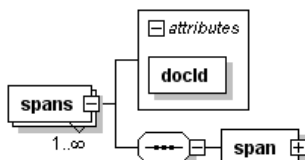
```
<termFormData id="t68" frequency="1">
  <termForm>marine area</termForm>
  <deps docId="124"><dep from="t10" to="t11" /></deps>
  <termContext>climate change in marine areas</termContext>
</termFormData>
```

### termForm



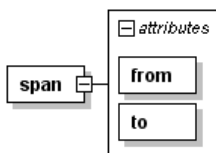
```
<termForm>marine area</termForm>
```

### spans

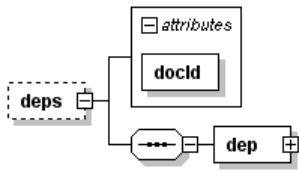


```
<spans docId="7824">
  <span from="w24" to="w43" />
  <span from="w123" to="w125" />
  <span from="w5627" to="w5628" />
</spans>
```

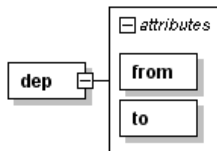
### span



```
<span from="w24" to="w43" />
```

**deps**

```
<deps docId="124">
  <dep from="t10" to="t11" />
</deps>
```

**dep**

```
<dep from="t10" to="t11" />
```

**termContext**

```
<termContext>climate change in marine areas</termContext>
```

## 5 Appendix

### 5.1 A: KYOTO LMF XML Schema Definition

```

<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified">
  <xs:annotation>
    <xs:documentation>KYOTO - LMF REPRESENTATION FORMAT</xs:documentation>
  </xs:annotation>
  <xs:element name="LexicalResource">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="GlobalInformation">
          <xs:complexType>
            <xs:attribute name="label" type="xs:string"/>
          </xs:complexType>
        </xs:element>
        <xs:element name="Lexicon" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="LexicalEntry" maxOccurs="unbounded">
                <xs:complexType>
                  <xs:sequence>
                    <xs:element ref="Meta" minOccurs="0"/>
                    <xs:element name="Lemma">
                      <xs:complexType>
                        <xs:attribute name="writtenForm" type="xs:string" use="required"/>
                        <xs:attribute name="partOfSpeech" type="partOfSpeechType" use="required"/>
                      </xs:complexType>
                    </xs:element>
                    <xs:element name="Sense" minOccurs="0" maxOccurs="unbounded">
                      <xs:complexType>
                        <xs:sequence>
                          <xs:element ref="Meta" minOccurs="0"/>
                          <xs:element ref="MonolingualExternalRefs" minOccurs="0"/>
                        </xs:sequence>
                        <xs:attribute name="id" type="xs:ID" use="required">
                      </xs:attribute>
                      <xs:attribute name="synset" type="xs:IDREF" use="required">
                      </xs:attribute>
                    </xs:complexType>
                  </xs:element>
                </xs:sequence>
                <xs:attribute name="id" type="xs:ID"/>
              </xs:complexType>
            </xs:element>
          <xs:element name="Synset" minOccurs="0" maxOccurs="unbounded">
            <xs:complexType>

```

```

<xs:sequence>
  <xs:element ref="Meta" minOccurs="0"/>
  <xs:element name="Definition" minOccurs="0">
    <xs:annotation>
      <xs:documentation>The gloss of the synset</xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element name="Statement" minOccurs="0"
MaxOccurs="unbounded">
          <xs:annotation>
            <xs:documentation>examples of use</xs:documentation>
          </xs:annotation>
          <xs:complexType>
            <xs:attribute name="example" type="xs:string" use="required"/>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
      <xs:attribute name="gloss" type="xs:string" use="required"/>
    </xs:complexType>
  </xs:element>
  <xs:element name="SynsetRelations">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="SynsetRelation">
          <xs:annotation>
            <xs:documentation>Relation between synset</xs:documentation>
          </xs:annotation>
          <xs:complexType>
            <xs:sequence>
              <xs:element ref="Meta" minOccurs="0"/>
            </xs:sequence>
            <xs:attribute name="target" type="synsetIdType"
use="required">
              <xs:annotation>
                <xs:documentation>synsetId</xs:documentation>
              </xs:annotation>
            </xs:attribute>
            <xs:attribute name="relType" type="internalRelationType"
use="required">
              </xs:attribute>
            </xs:complexType>
          </xs:element>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
    <xs:element ref="MonolingualExternalRefs"/>
  </xs:sequence>
  <xs:attribute name="id" type="synsetIdType" use="required"/>

```

```

        <xs:attribute name="baseConcept" type="baseConceptType" use="required"/>
    </xs:complexType>
</xs:element>
</xs:sequence>
<xs:attribute name="languageCoding" type="xs:string" fixed="ISO 639-3"/>
<xs:attribute name="label" type="xs:string"/>
<xs:attribute name="language" type="languageType" use="required"/>
<xs:attribute name="owner" type="xs:string" use="required"/>
<xs:attribute name="version" type="xs:string" use="required"/>
</xs:complexType>
</xs:element>
<xs:element name="SenseAxes" minOccurs="0">
    <xs:complexType>
        <xs:sequence>
            <xs:element name="SenseAxis" maxOccurs="unbounded">
                <xs:complexType>
                    <xs:sequence>
                        <xs:element ref="Meta" minOccurs="0"/>
                        <xs:element name="Target" maxOccurs="unbounded">
                            <xs:annotation>
                                <xs:documentation>One target for each synset</xs:documentation>
                            </xs:annotation>
                            <xs:complexType>
                                <xs:attribute name="ID" type="synsetIdType" use="required"/>
                            </xs:complexType>
                        </xs:element>
                        <xs:element ref="InterlingualExternalRefs"/>
                    </xs:sequence>
                    <xs:attribute name="id" type="xs:ID" use="required"/>
                    <xs:attribute name="relType" type="externalRelationType" use="required"/>
                </xs:complexType>
            </xs:element>
        </xs:sequence>
    </xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="Meta">
    <xs:complexType>
        <xs:attribute name="author" type="xs:string"/>
        <xs:attribute name="date" type="xs:string"/>
        <xs:attribute name="source" type="xs:string"/>
        <xs:attribute name="status" type="xs:string"/>
        <xs:attribute name="confidenceScore" type="rateType"/>
    </xs:complexType>
</xs:element>
<xs:element name="MonolingualExternalRefs">
    <xs:annotation>

```



```

    <xs:documentation>References to resources of the same language </xs:documentation>
  </xs:annotation>
</xs:complexType>
<xs:sequence>
  <xs:element name="MonolingualExternalRef" maxOccurs="unbounded">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="Meta" minOccurs="0"/>
      </xs:sequence>
      <xs:attribute name="externalSystem" type="externalSystemType" use="required"/>
      <xs:attribute name="externalReference" type="xs:string" use="required"/>
      <xs:attribute name="relType" type="ontologyRelationType" use="optional">
        <xs:annotation>
          <xs:documentation>Used only if the externalSystem is SUMO</xs:documentation>
        </xs:annotation>
      </xs:attribute>
    </xs:complexType>
  </xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="InterlingualExternalRefs">
  <xs:annotation>
    <xs:documentation>References to ontologies</xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:sequence>
      <xs:element name="InterlingualExternalRef" maxOccurs="unbounded">
        <xs:complexType>
          <xs:sequence>
            <xs:element ref="Meta" minOccurs="0"/>
          </xs:sequence>
          <xs:attribute name="externalSystem" type="ontologyType" use="required">
            <xs:annotation>
              <xs:documentation>The name of the ontology</xs:documentation>
            </xs:annotation>
          </xs:attribute>
          <xs:attribute name="externalReference" type="xs:string" use="required">
            </xs:attribute>
          <xs:attribute name="relType" type="ontologyRelationType" use="optional">
            </xs:attribute>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
<!-- ***** TYPES ***** -->
<xs:simpleType name="internalRelationType">
  <xs:annotation>

```

```

    <xs:documentation>List of internal relation types</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:enumeration value="antonym"/>
    <xs:enumeration value="antonym_comp"/>
    <xs:enumeration value="be_in_state"/>
    <xs:enumeration value="category"/>
    <xs:enumeration value="category_term"/>
    <xs:enumeration value="causes"/>
    <xs:enumeration value="co_agent_instrument"/>
    <xs:enumeration value="co_agent_patient"/>
    <xs:enumeration value="co_agent_result"/>
    <xs:enumeration value="co_instrument_agent"/>
    <xs:enumeration value="co_instrument_patient"/>
    <xs:enumeration value="co_instrument_result"/>
    <xs:enumeration value="co_patient_agent"/>
    <xs:enumeration value="co_patient_instrument"/>
    <xs:enumeration value="co_patient_result"/>
    <xs:enumeration value="co_result_agent"/>
    <xs:enumeration value="co_result_instrument"/>
    <xs:enumeration value="co_result_patient"/>
    <xs:enumeration value="co_role"/>
    <xs:enumeration value="for_purpose_of"/>
    <xs:enumeration value="fuzzynym"/>
    <xs:enumeration value="gloss"/>
    <xs:enumeration value="has_derived"/>
    <xs:enumeration value="has_holo_location"/>
    <xs:enumeration value="has_holo_madeof"/>
    <xs:enumeration value="has_holo_member"/>
    <xs:enumeration value="has_holo_part"/>
    <xs:enumeration value="has_holo_portion"/>
    <xs:enumeration value="has_holonym"/>
    <xs:enumeration value="has_hyperonym"/>
    <xs:enumeration value="has_hyponym"/>
    <xs:enumeration value="has_mero_location"/>
    <xs:enumeration value="has_mero_madeof"/>
    <xs:enumeration value="has_mero_member"/>
    <xs:enumeration value="has_mero_part"/>
    <xs:enumeration value="has_mero_portion"/>
    <xs:enumeration value="has_meronym"/>
    <xs:enumeration value="has_pertainym"/>
    <xs:enumeration value="has_subevent"/>
    <xs:enumeration value="has_xpos_hyperonym"/>
    <xs:enumeration value="has_xpos_hyponym"/>
    <xs:enumeration value="in_manner"/>
    <xs:enumeration value="instance"/>
  <xs:enumeration value="involved"/>
    <xs:enumeration value="involved_agent"/>
    <xs:enumeration value="involved_direction"/>

```

```

<xs:enumeration value="involved_instrument"/>
<xs:enumeration value="involved_location"/>
<xs:enumeration value="involved_patient"/>
<xs:enumeration value="involved_result"/>
<xs:enumeration value="involved_source_direction"/>
<xs:enumeration value="involved_target_direction"/>
<xs:enumeration value="is_a_value_of"/>
<xs:enumeration value="is_caused_by"/>
<xs:enumeration value="is_derived_from"/>
<xs:enumeration value="is_subevent_of"/>
<xs:enumeration value="manner_of"/>
<xs:enumeration value="near_antonym"/>
<xs:enumeration value="near_synonym"/>
<xs:enumeration value="nearest"/>
<xs:enumeration value="pertains_to"/>
<xs:enumeration value="region"/>
<xs:enumeration value="region_term"/>
<xs:enumeration value="related"/>
<xs:enumeration value="related_to"/>
<xs:enumeration value="results_in"/>
<xs:enumeration value="rgloss"/>
<xs:enumeration value="role"/>
<xs:enumeration value="role_agent"/>
<xs:enumeration value="role_direction"/>
<xs:enumeration value="role_instrument"/>
<xs:enumeration value="role_location"/>
<xs:enumeration value="role_manner"/>
<xs:enumeration value="role_patient"/>
<xs:enumeration value="role_result"/>
<xs:enumeration value="role_source_direction"/>
<xs:enumeration value="role_target_direction"/>
<xs:enumeration value="see_also_wn15"/>
<xs:enumeration value="state_of"/>
<xs:enumeration value="usage"/>
<xs:enumeration value="usage_term"/>
<xs:enumeration value="verb_group"/>
<xs:enumeration value="xpos_fuzzynym"/>
<xs:enumeration value="xpos_near_antonym"/>
<xs:enumeration value="xpos_near_synonym"/>
</xs:restriction>
</xs:simpleType>
<xs:simpleType name="externalRelationType">
  <xs:annotation>
    <xs:documentation>List of external relation types</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:enumeration value="eq_synonym"/>
    <xs:enumeration value="eq_near_synonym"/>
    <xs:enumeration value="eq_has_hypernym"/>
  </xs:restriction>
</xs:simpleType>

```

```

    <xs:enumeration value="eq_has_hyponym"/>
    <xs:enumeration value="eq_involved"/>
    <xs:enumeration value="eq_role"/>
    <xs:enumeration value="eq_is_caused_by"/>
    <xs:enumeration value="eq_causes"/>
    <xs:enumeration value="eq_has_holonym"/>
    <xs:enumeration value="eq_has_meronym"/>
    <xs:enumeration value="eq_has_subevent"/>
    <xs:enumeration value="eq_is_subevent_of"/>
    <xs:enumeration value="eq_be_in_state"/>
    <xs:enumeration value="eq_is_state_of"/>
    <xs:enumeration value="eq_co_role"/>
    <xs:enumeration value="eq_generalization"/>
    <xs:enumeration value="eq_metonym"/>
    <xs:enumeration value="eq_diathesis"/>
    <xs:enumeration value="eq_in_manner"/>
    <xs:enumeration value="eq_has_instance"/>
    <xs:enumeration value="eq_belongs_to_class"/>
    <xs:enumeration value="eq_antonym"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="languageType">
  <xs:annotation>
    <xs:documentation>List of all languages code from ISO 639-3: three lower case letters: eng, ita,eus, zho, jpn,
    spa, nld</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:enumeration value="eus"/>
    <xs:enumeration value="eng"/>
    <xs:enumeration value="ita"/>
    <xs:enumeration value="jpn"/>
    <xs:enumeration value="nld"/>
    <xs:enumeration value="spa"/>
    <xs:enumeration value="zho"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="partOfSpeechType">
  <xs:annotation>
    <xs:documentation>List of Pos type: n, v, a (adjective), r (adverb), s (satellite adjective)</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:enumeration value="n"/>
    <xs:enumeration value="v"/>
    <xs:enumeration value="a"/>
    <xs:enumeration value="r"/>
    <xs:enumeration value="s"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="versionType">

```

```

<xs:annotation>
  <xs:documentation>List of wordnet versions: 15, 16, 17, 20, 21, 30, ...</xs:documentation>
</xs:annotation>
<xs:restriction base="xs:unsignedInt">
  <xs:enumeration value="15"/>
  <xs:enumeration value="16"/>
  <xs:enumeration value="17"/>
  <xs:enumeration value="18"/>
  <xs:enumeration value="19"/>
  <xs:enumeration value="20"/>
  <xs:enumeration value="21"/>
  <xs:enumeration value="30"/>
</xs:restriction>
</xs:simpleType>
<xs:simpleType name="synsetIdType">
  <xs:annotation>
    <xs:documentation>Format of a synset ID: language code-version-id-pos tag</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:pattern value="[a-z]{3}-[0-9]{2}-[0-9]+--[nvars]"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="baseConceptType">
  <xs:annotation>
    <xs:documentation>List of base concept sets: 1,2,3</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:unsignedInt">
    <xs:enumeration value="1"/>
    <xs:enumeration value="2"/>
    <xs:enumeration value="3"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="ontologyRelationType">
  <xs:annotation>
    <xs:documentation>List of SUMO relation types: at, plus, equal</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:NMTOKEN">
    <xs:enumeration value="equal"/>
    <xs:enumeration value="plus"/>
    <xs:enumeration value="at"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="externalSystemType">
  <xs:annotation>
    <xs:documentation>List of External Resource: Domain, SuperSense, SUMO, TCO,
wordnet3.0</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:NMTOKEN">
    <xs:enumeration value="SUMO"/>

```

```
<xs:enumeration value="Domain"/>
<xs:enumeration value="TCO"/>
<xs:enumeration value="SuperSense"/>
<xs:enumeration value="Wordnet3.0"/>
<xs:enumeration value="Wordnet2.1"/>
<xs:enumeration value="Wordnet2.0"/>
<xs:enumeration value="Wordnet1.7"/>
<xs:enumeration value="Wordnet1.6"/>
<xs:enumeration value="Wordnet1.5"/>
</xs:restriction>
</xs:simpleType>
<xs:simpleType name="ontologyType">
  <xs:annotation>
    <xs:documentation>List of Ontologies: SUMO, DOLCE</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:NMTOKEN">
    <xs:enumeration value="SUMO"/>
    <xs:enumeration value="DOLCE"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="rateType">
  <xs:restriction base="xs:float">
    <xs:maxInclusive value="1"/>
    <xs:minInclusive value="0"/>
  </xs:restriction>
</xs:simpleType>
</xs:schema>
```

## 2 B: KYOTO TMF XML Schema Definition

```

<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified">
  <xs:element name="tmf">
    <xs:annotation>
      <xs:documentation>The root is identical to the TMF proposal </xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element name="terminologicalEntry" maxOccurs="unbounded">
          <xs:annotation>
            <xs:documentation>An entry containing information on terminological units</xs:documentation>
          </xs:annotation>
          <xs:complexType>
            <xs:sequence>
              <xs:element name="termDomain">
                <xs:annotation>
                  <xs:documentation>The domain of the term</xs:documentation>
                </xs:annotation>
                <xs:simpleType>
                  <xs:restriction base="xs:string">
                    <xs:enumeration value="environment"/>
                  </xs:restriction>
                </xs:simpleType>
              </xs:element>
              <xs:element name="treeProfile">
                <xs:annotation>
                  <xs:documentation>Classification of the complete term tree</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                  <xs:sequence>
                    <xs:element ref="microWorld" maxOccurs="unbounded"/>
                  </xs:sequence>
                </xs:complexType>
              </xs:element>
              <xs:element name="languageSection" maxOccurs="unbounded">
                <xs:annotation>
                  <xs:documentation>The part of a terminological entry containing information related to
                    one language</xs:documentation>
                </xs:annotation>
                <xs:complexType>
                  <xs:sequence>
                    <xs:element name="termSection" maxOccurs="unbounded">
                      <xs:annotation>
                        <xs:documentation>The part of a language section giving information about a
                          term</xs:documentation>
                      </xs:annotation>
                    </xs:element>
                  </xs:sequence>
                </xs:complexType>
              </xs:element>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>

```

```

<xs:sequence>
  <xs:element name="normalizedTerm">
    <xs:annotation>
      <xs:documentation>Unifies different variants of the term. It may be
        omitted if the term identifier is used</xs:documentation>
    </xs:annotation>
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="climate change"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:element>
  <xs:element name="partOfSpeech">
    <xs:annotation>
      <xs:documentation>Part of speech of the term, should be the same
        for all variants</xs:documentation>
    </xs:annotation>
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="noun"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:element>
  <xs:element name="preferredForm">
    <xs:annotation>
      <xs:documentation>Canonical form for representation purposes. It is
        used for labelling the term for the user in an
        interface</xs:documentation>
    </xs:annotation>
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="Climate Change"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:element>
  <xs:element name="forms">
    <xs:annotation>
      <xs:documentation>List of forms and pointers to positions in SemAF
        files</xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="termFormData" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="parentData" minOccurs="0">
    <xs:annotation>

```



```

    <xs:documentation>Structural parent relation for establishing the
    term tree. It's optional because the root term has no
    parent.</xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:sequence>
      <xs:element name="parentTerm">
        <xs:annotation>
          <xs:documentation>It contains the parent term id into the
          target attribute</xs:documentation>
        </xs:annotation>
        <xs:complexType mixed="true">
          <xs:attribute name="target" use="required">
            <xs:annotation>
              <xs:documentation>Term id of the
              parent</xs:documentation>
            </xs:annotation>
            <xs:simpleType>
              <xs:restriction base="xs:string">
                <xs:enumeration value="t13"/>
              </xs:restriction>
            </xs:simpleType>
          </xs:attribute>
        </xs:complexType>
      </xs:element>
      <xs:element name="alternativeParent" maxOccurs="unbounded">
        <xs:annotation>
          <xs:documentation>Alternative parents extracted from
          external sources</xs:documentation>
        </xs:annotation>
        <xs:complexType>
          <xs:simpleContent>
            <xs:extension base="xs:string">
              <xs:attribute name="source" type="xs:anyURI"
              use="required"/>
              <xs:attribute name="date" type="xs:date"
              use="required"/>
            </xs:extension>
          </xs:simpleContent>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="termStatistics">
  <xs:annotation>
    <xs:documentation>Statistics at the term level</xs:documentation>
  </xs:annotation>
  <xs:complexType>

```

```

<xs:sequence>
  <xs:element name="documentNumber" type="xs:byte">
    <xs:annotation>
      <xs:documentation>The number of source documents
        from which the term is derived</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="termFrequency" type="xs:integer">
    <xs:annotation>
      <xs:documentation>The total number of occurrences in
        source documents </xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="termSalience" type="rateType">
    <xs:annotation>
      <xs:documentation>A statistic parameter with a range
        value [0,1 ]</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="termConnectivity" type="xs:integer">
    <xs:annotation>
      <xs:documentation>The number of connections of a term
        class in the tree</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="cumulativeFrequency" type="xs:integer">
    <xs:annotation>
      <xs:documentation>The sum of termFrequency of all
        descendants</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="cumulativeDocumentNumber"
    type="xs:integer">
    <xs:annotation>
      <xs:documentation>The sum of documentNumber of all
        descendants</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element name="termSiblings" type="xs:integer">
    <xs:annotation>
      <xs:documentation>The number of terms that share the
        same parent</xs:documentation>
    </xs:annotation>
  </xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="termProfile">
  <xs:annotation>

```

```

    <xs:documentation>The termProfile is similar to the treeProfile
    except that it has an additional attribute profileMatch that indicates
    the overlap of microWorld values across the termProfile and the
    treeProfile</xs:documentation>
  </xs:annotation>
</xs:complexType>
<xs:sequence>
  <xs:element ref="microWorld" maxOccurs="unbounded"/>
</xs:sequence>
<xs:attribute name="profileMatch" use="required">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:enumeration value="0.69"/>
    </xs:restriction>
  </xs:simpleType>
</xs:attribute>
</xs:complexType>
</xs:element>
<xs:element name="sources">
  <xs:annotation>
    <xs:documentation>List of type of sources sections (Toc, Body) from
    which the term is extracted</xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:sequence>
      <xs:element name="termSource" maxOccurs="unbounded">
        <xs:complexType>
          <xs:attribute name="score" type="rateType"
            use="required">
            <xs:annotation>
              <xs:documentation>It indicates the proportion to
              which a term was found in the particular
              section</xs:documentation>
            </xs:annotation>
          </xs:attribute>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="semanticRelations">
  <xs:annotation>
    <xs:documentation>List of relations to external semantic
    resources</xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:sequence>
      <xs:element name="definitions">
        <xs:annotation>

```

```

    <xs:documentation>List of definitions extracted from
    external semantic reources</xs:documentation>
  </xs:annotation>
</xs:complexType>
<xs:sequence>
  <xs:element name="termDefiniton"
    maxOccurs="unbounded">
    <xs:complexType mixed="true">
      <xs:attribute name="source" type="xs:string"
        use="required"/>
      <xs:attribute name="date" type="xs:date"
        use="required"/>
    </xs:complexType>
  </xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="semanticMatches">
  <xs:annotation>
    <xs:documentation>List of mapping to external semantic
    resource</xs:documentation>
  </xs:annotation>
</xs:complexType>
<xs:sequence>
  <xs:element name="semanticMatch"
    maxOccurs="unbounded">
    <xs:complexType>
      <xs:choice>
        <xs:element name="sense"
          maxOccurs="unbounded">
          <xs:complexType>
            <xs:attribute name="source"
              type="xs:string" use="required"/>
            <xs:attribute name="sensecode"
              type="synsetIdType"
              use="required"/>
            <xs:attribute name="weight"
              type="rateType" use="required"/>
          </xs:complexType>
        </xs:element>
        <xs:element name="ontology"
          maxOccurs="unbounded">
          <xs:complexType>
            <xs:attribute name="source"
              type="xs:string" use="required"/>
            <xs:attribute name="class"
              type="xs:string" use="required"/>
            <xs:attribute name="weight"
              type="rateType" use="required"/>
          </xs:complexType>
        </xs:element>
      </xs:choice>
    </xs:complexType>
  </xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>

```

```

        </xs:complexType>
      </xs:element>
    </xs:choice>
    <xs:attribute name="type" use="required">
      <xs:simpleType>
        <xs:restriction base="xs:string">
          <xs:enumeration
            value="ontologyAlt"/>
          <xs:enumeration value="senseAlt"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:attribute>
    <xs:attribute name="orig" use="required">
      <xs:simpleType>
        <xs:restriction base="xs:QName">
          <xs:enumeration
            value="urn:sumo"/>
          <xs:enumeration
            value="urn:wordnet1.7"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:attribute>
  </xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="structuralRelations">
  <xs:annotation>
    <xs:documentation>List of structural relations </xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:sequence>
      <xs:element name="structuralRelation"
        maxOccurs="unbounded">
        <xs:annotation>
          <xs:documentation>It groups all occurrences of a
            morpho-syntactic pattern in which a term
            occurred</xs:documentation>
        </xs:annotation>
        <xs:complexType>
          <xs:sequence>
            <xs:element name="syntaxElement"
              type="xs:string"/>
            <xs:element ref="termFormData"
              maxOccurs="unbounded"/>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>

```

```

        </xs:sequence>
        <xs:attribute name="syntaxRole" type="xs:string"
use="required"/>
        <xs:attribute name="semanticRole" type="xs:string"
use="required"/>
        <xs:attribute name="active" type="xs:boolean"/>
    </xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="termComponentSection">
    <xs:annotation>
        <xs:documentation>The section of a term section giving information
about components of a term</xs:documentation>
    </xs:annotation>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:sequence>
<xs:attribute name="language" use="required">
    <xs:annotation>
        <xs:documentation>The language of the term. The coding is defined by the
languageCoding attribute</xs:documentation>
    </xs:annotation>
</xs:attribute>
</xs:complexType>
</xs:element>
</xs:sequence>
<xs:attribute name="id" use="required">
    <xs:annotation>
        <xs:documentation>The id of the term: used to build the term tree</xs:documentation>
    </xs:annotation>
</xs:attribute>
<xs:attribute name="languageCoding" use="required">
    <xs:annotation>
        <xs:documentation>The language coding is: ISO-639-3</xs:documentation>
    </xs:annotation>
</xs:attribute>
</xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="struct">
    <xs:annotation>
        <xs:documentation>Structural node</xs:documentation>
    </xs:annotation>

```

```

<xs:complexType>
  <xs:attribute name="id" type="xs:ID"/>
  <xs:attribute name="type" type="structuralNodeType" use="required"/>
</xs:complexType>
</xs:element>
<xs:element name="microWorld">
  <xs:annotation>
    <xs:documentation>It contains a domain label with a confidence score</xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:attribute name="score" type="rateType" use="required"/>
  </xs:complexType>
</xs:element>
<xs:element name="termFormData">
  <xs:annotation>
    <xs:documentation>Used for listing all the different appearances of the term in
      the source documents with pointers to the locations in the SemAF notation</xs:documentation>
  </xs:annotation>
  <xs:complexType>
    <xs:sequence>
      <xs:element name="termForm" type="xs:string"/>
      <xs:choice>
        <xs:element name="spans" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="span">
                <xs:complexType>
                  <xs:attribute name="from" type="xs:string" use="required"/>
                  <xs:attribute name="to" type="xs:string" use="required"/>
                </xs:complexType>
              </xs:element>
            </xs:sequence>
            <xs:attribute name="docId" type="xs:string" use="required"/>
          </xs:complexType>
        </xs:element>
        <xs:element name="deps" minOccurs="0">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="dep">
                <xs:complexType>
                  <xs:attribute name="from" use="required"/>
                  <xs:attribute name="to" use="required"/>
                </xs:complexType>
              </xs:element>
            </xs:sequence>
            <xs:attribute name="docId" use="required">
              <xs:simpleType>
                <xs:restriction base="xs:string">
                  <xs:enumeration value="1 24"/>
                </xs:restriction>
              </xs:simpleType>
            </xs:attribute>
          </xs:complexType>
        </xs:element>
      </xs:choice>
    </xs:sequence>
  </xs:complexType>
</xs:element>

```

```

        <xs:enumeration value="1 824"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>
</xs:complexType>
</xs:element>
</xs:choice>
  <xs:element name="termContext" type="xs:string" minOccurs="0"/>
</xs:sequence>
<xs:attribute name="id" type="xs:ID" use="required"/>
<xs:attribute name="frequency" type="xs:integer" use="required"/>
</xs:complexType>
</xs:element>
<xs:simpleType name="structuralNodeType">
  <xs:annotation>
    <xs:documentation>List of sctructural node types: TE, TS, LS, TCS</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:NMTOKEN">
    <xs:enumeration value="TE"/>
    <xs:enumeration value="TS"/>
    <xs:enumeration value="LS"/>
    <xs:enumeration value="TCS"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="rateType">
  <xs:restriction base="xs:float"/>
</xs:simpleType>
<xs:simpleType name="synsetIdType">
  <xs:annotation>
    <xs:documentation>Format of a synset ID: language code-version-id-pos tag</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:pattern value="[a-z]{3}-[0-9]{2}-[0-9]+--[nvars]"/>
  </xs:restriction>
</xs:simpleType>
</xs:schema>

```



## Bibliography

1. Lexical Markup Framework, Web Site - <http://www.lexicalmarkupframework.org/>
2. Web Ontology Language (OWL), W3C Recommendation - <http://www.w3.org/TR/owl-features/>
3. Knowledge Interchange Format (KIF), Specification - <http://logic.stanford.edu/kif/dpans.html>
4. Resource Description Framework, W3C Recommendation - <http://www.w3.org/TR/rdf-primer/>
5. Resource Description Framework Schema, W3C Recommendation - <http://www.w3.org/TR/rdf-schema/>
6. The W3C OWL Working Group, Web site - [http://www.w3.org/2007/OWL/wiki/OWL\\_Working\\_Group](http://www.w3.org/2007/OWL/wiki/OWL_Working_Group)
7. OWL Web Ontology Language XML Presentation Syntax, W3C Note - <http://www.w3.org/TR/owl-xmlsyntax/>
8. The Cornetto Project, Combinatorial and Relational Network as Toolkit for Dutch Language Technology Web site, Facultiet der Letteren, Vrije Universiteit - <http://www.let.vu.nl/onderzoek/projectsites/cornetto/index.html>
9. The Suggested Upper Level Ontology (SUMO), Web site - <http://www.ontologyportal.org/>