Ontology Learning for Semantic Metadata Extraction

Paul Buitelaar

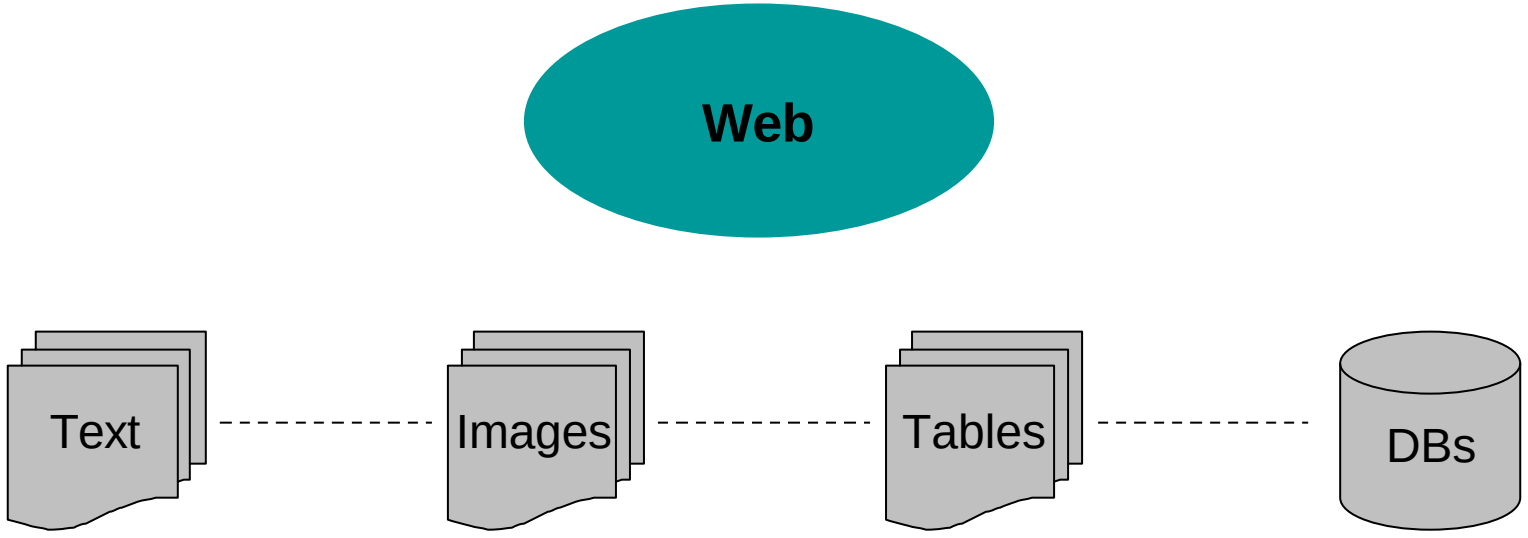
DFKI GmbH, Saarbrücken

Overview

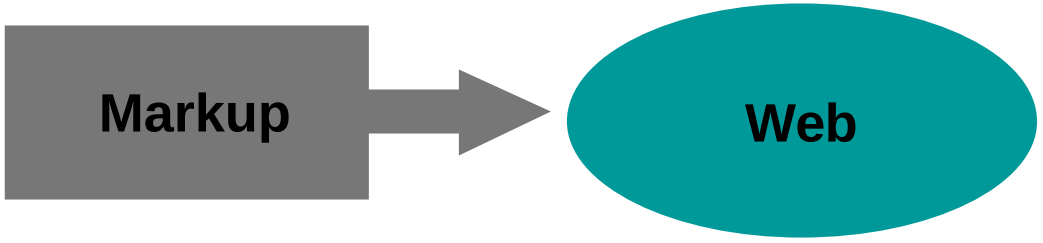
- Metadata and the Semantic Web
  - Semantic Metadata, Knowledge Markup and Ontologies
- Ontology Learning
  - Ontology Learning Layer Cake
- Tools
  - OntoLT – Connecting Linguistic Analysis and Ontology Engineering
  - RelExt – Linguistic Analysis in Relation Extraction
  - ISOLDE – Web-based Ontology Learning

Metadata and the Semantic Web

Web Consists of Non-Interpreted Data



Interpretation through Metadata - Categories



YAHOO! SEARCH Web | Images | Video | Audio | Directory

director

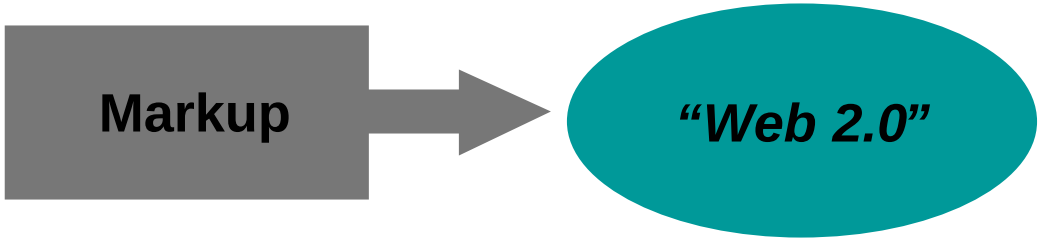
Directory Results

« Complete Directory Results

TOP 10 RELATED DIRECTORY CATEGORIES
out of 1,323.

1. Movie and Film **Directors**
2. Actors > Kim **Director**
3. Fanlistings > **Directors** and Producers
4. Silent Movie Actors and **Directors**
5. Clint Eastwood Films as **Director**
6. Animation > **Directors** and Animators
7. **Director** of National Intelligence
8. Theater > **Directors**
9. Talent and Crew > **Directors**
10. Classic Hollywood **Directors**

Interpretation through Metadata – User Tags



flickr

You aren't signed in Sign In Help

Home Learn More Sign Up! Explore Search everyone's photos Search

Explore / Tags /

Hot tags

In the last 24 hours
utatathursdaywalk11, gnomedex, virtualtata06guest, apachecon, fourthofjuly, contextwatcher guadec2006, mobilife, thursday, wimbledon, ixus, effects, coolpix, valparaiso, grd, nikond70s, cbdflickr, srilanka, fz30, nycpb

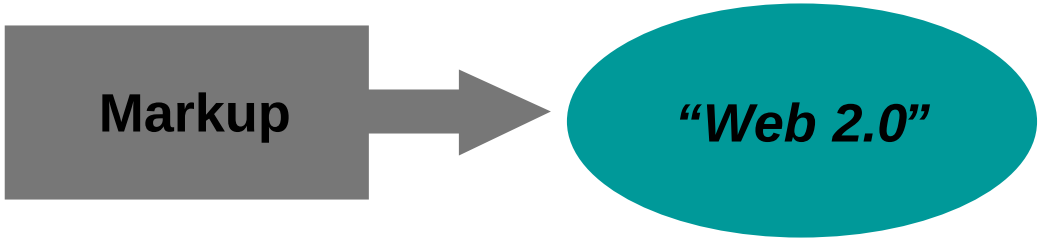
Over the last week
mermaidparade, intonation, sfpride, ala2006, comfest, utatathursdaywalk11 o2wirelessfestival, bloggercon, pab2006, bloggerconiv, olympicisland, sfpride2006, juhannus, mermaidparade2006, dykemarch, achtelfinale, isummit06, barcampsf, greektheatre, wirelessfestival

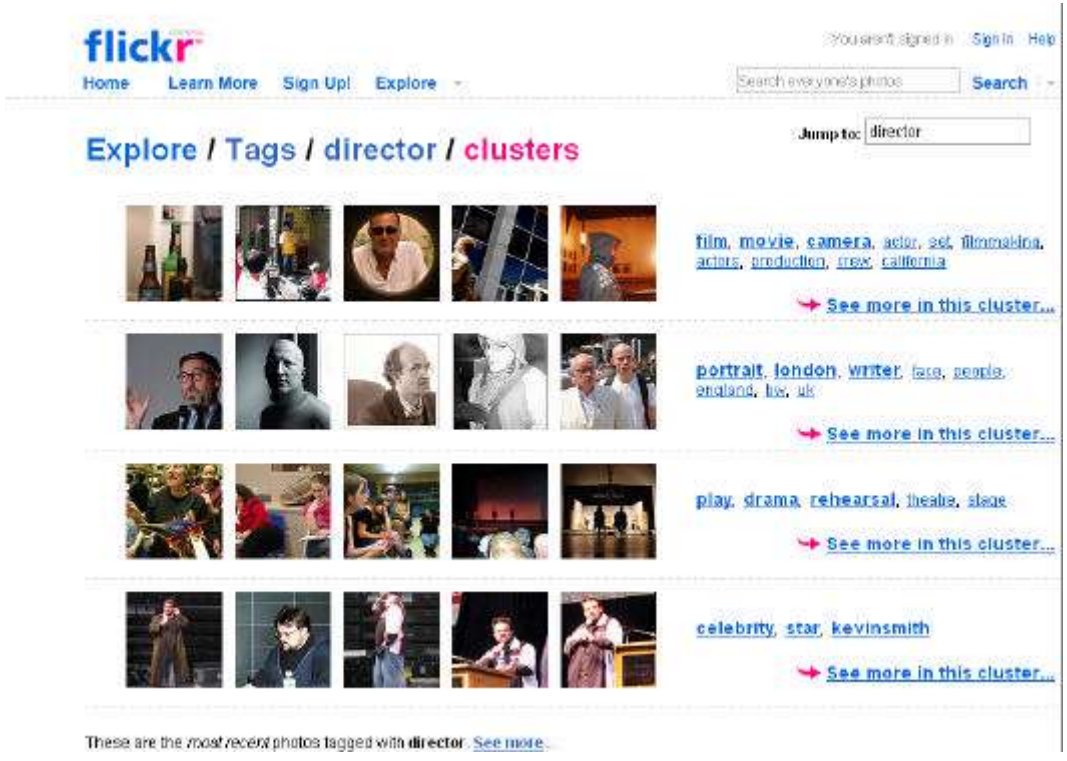
Jump to: GO

All time most popular tags

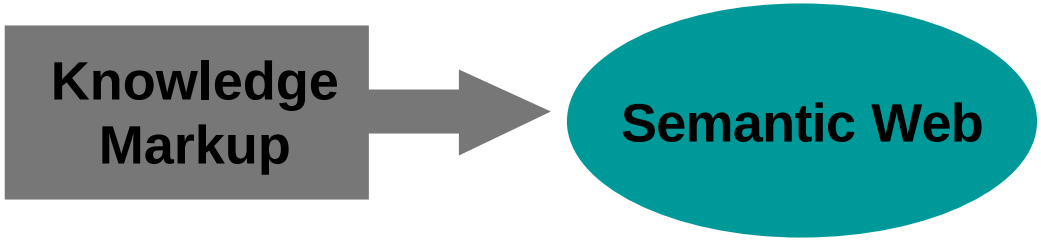
africa amsterdam animal animals april architecture art australia baby barcelona beach berlin birthday black blackandwhite blue boston bw california cameraphone camping canada canon car cat cats chicago china christmas church city clouds color concert day dc dog dogs england europe family festival film florida flower flowers food france friends fun

Interpretation through Metadata – User Tags



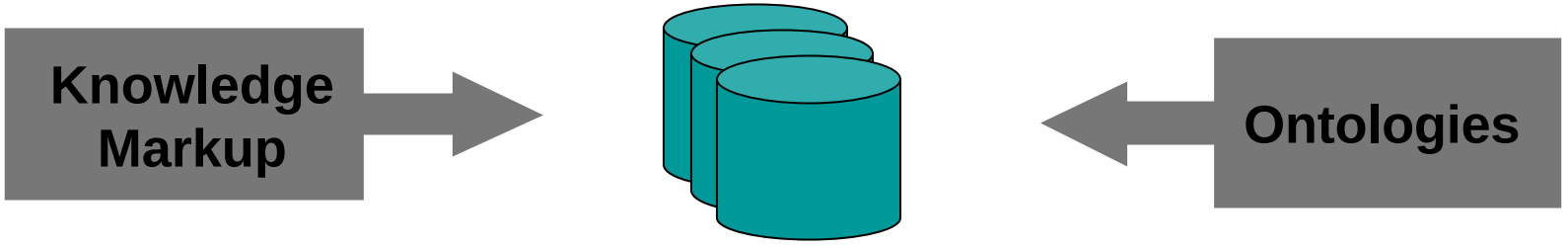


Formal Interpretation (Knowledge Markup)

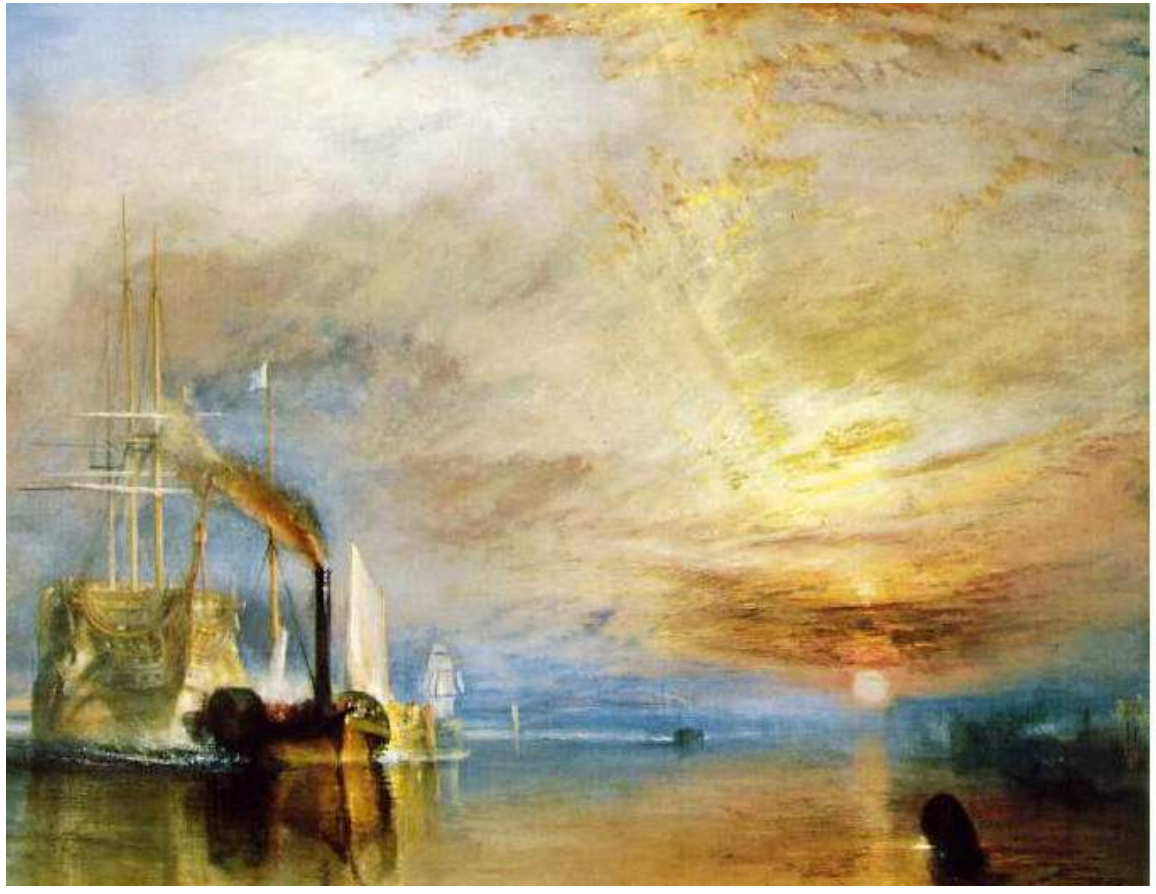


Based on Ontologies

Turns the Web into a Knowledge Base



Allows for Ontology-based Retrieval



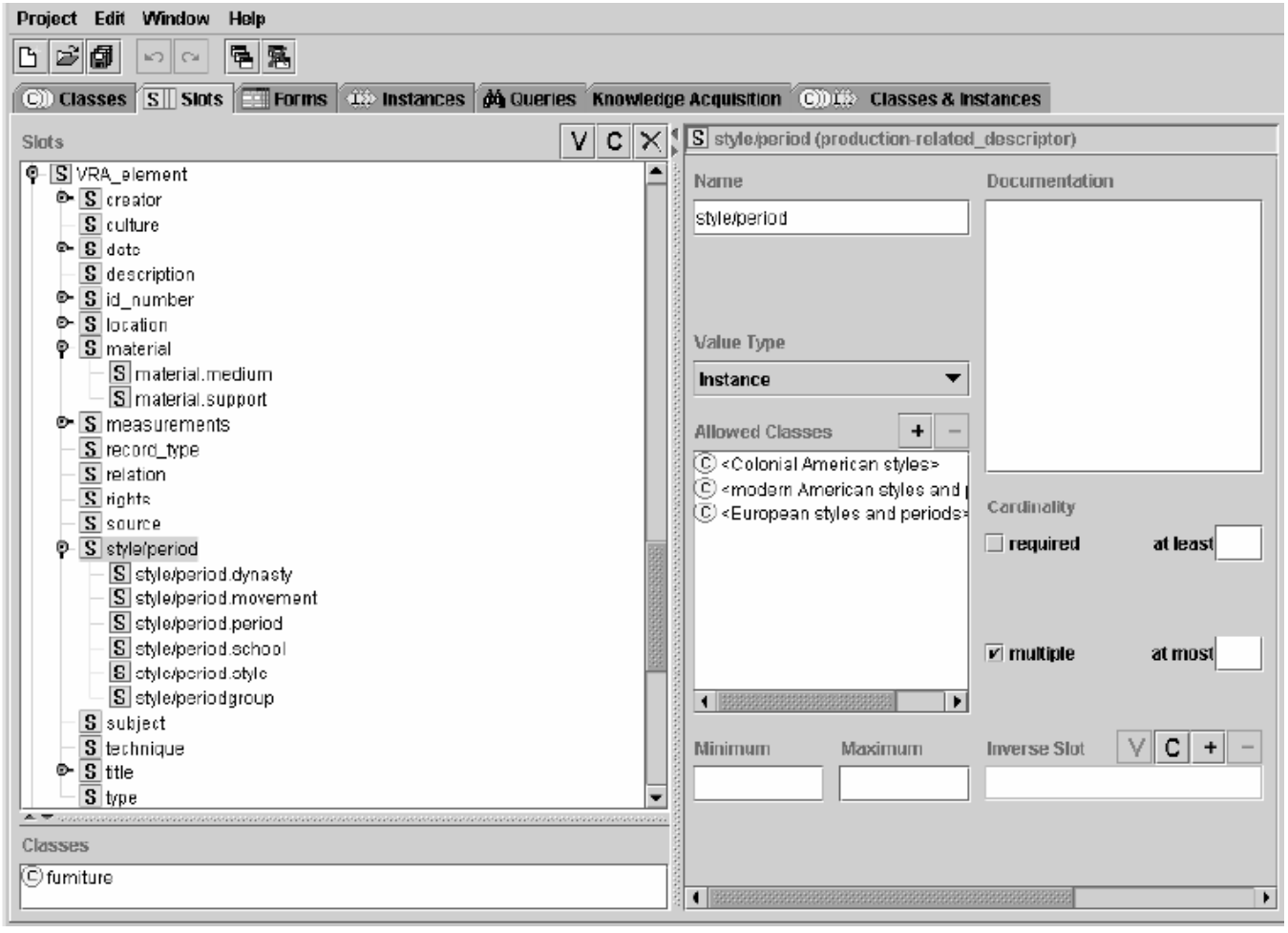
vra Creator	TURNER, Joseph Mallard William;
vra Date	1838;
vra Current Site	National Gallery, London;
vra Material.Medium	aat:oil paint;
vra Material.Support	aat:unprimed canvas;
vra Measurements.Dimensions	91 x 122 cm;
vra Title	The Fighting "Temeraire" tugged to her last berth to be broken up;
rdf type	ec:Artchive Work;

Value:

artchive:turner/temeraire.jpg; vra relation.depicts

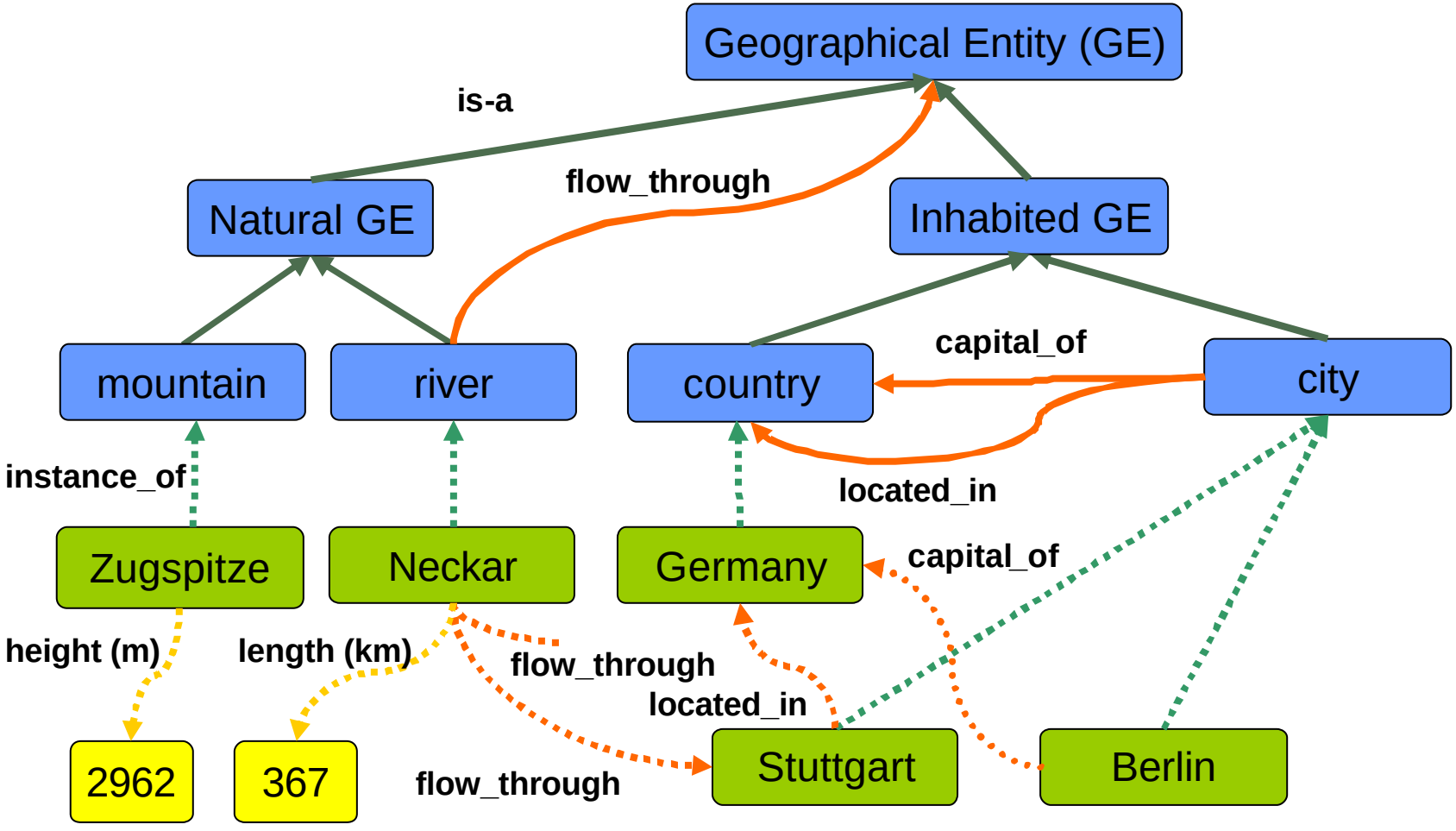
MultimediaN Project (CWI, Amsterdam) http://e-culture.multimedian.nl/demo/search

Ontology Development Needed



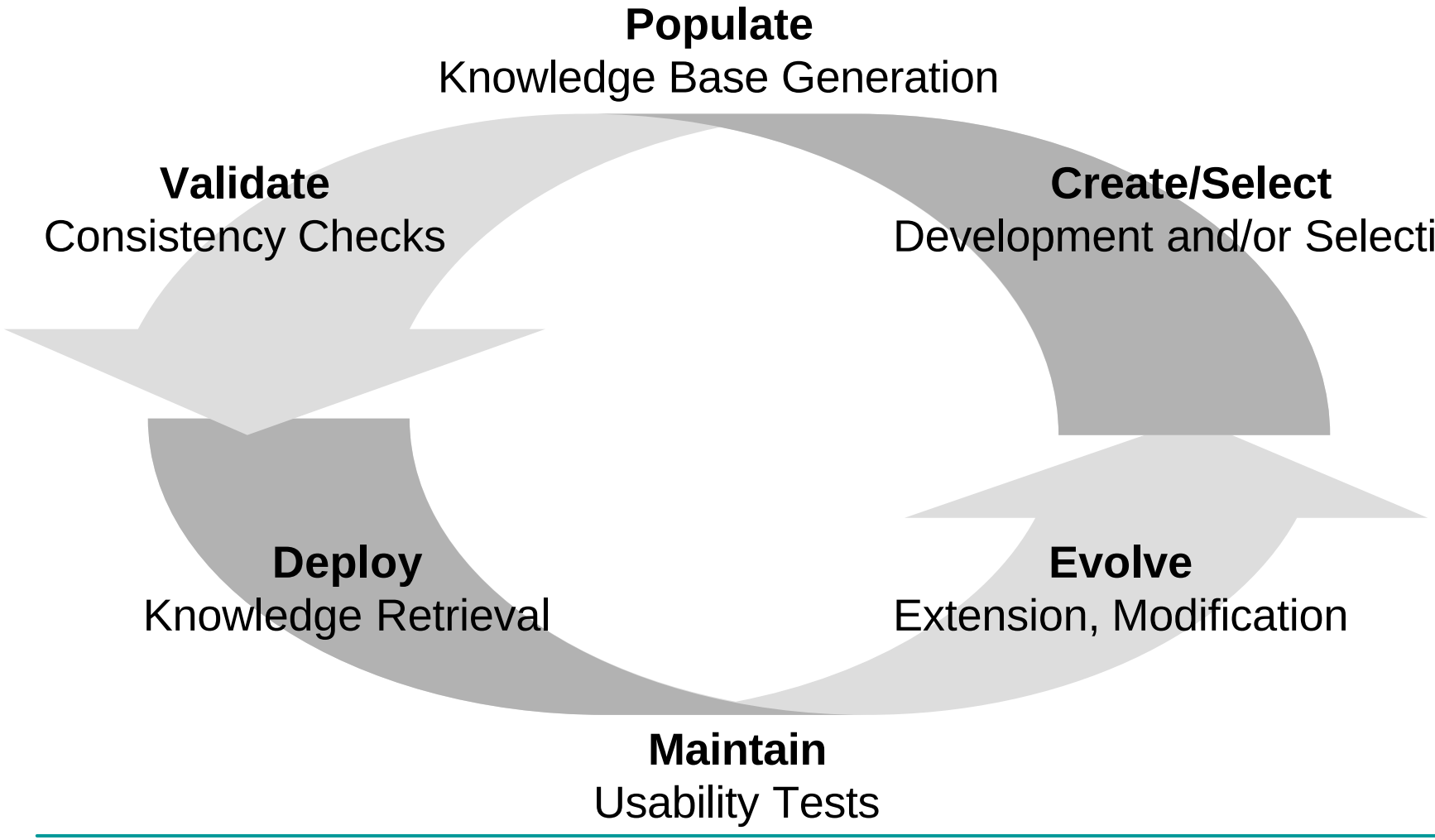
Ontology Learning

Ontology Example - Geography



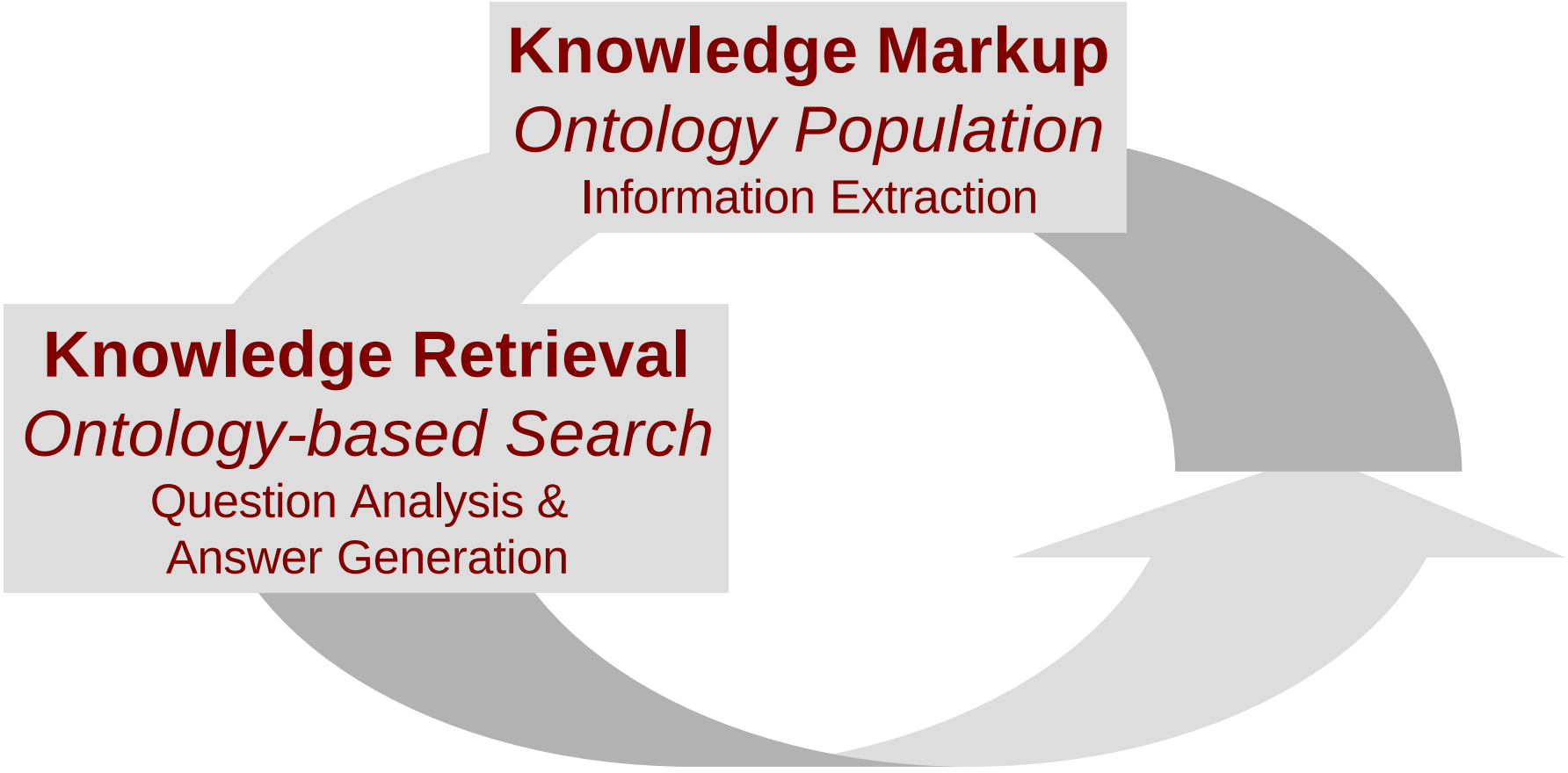
Design: Philipp Cimiano

Ontology Life Cycle

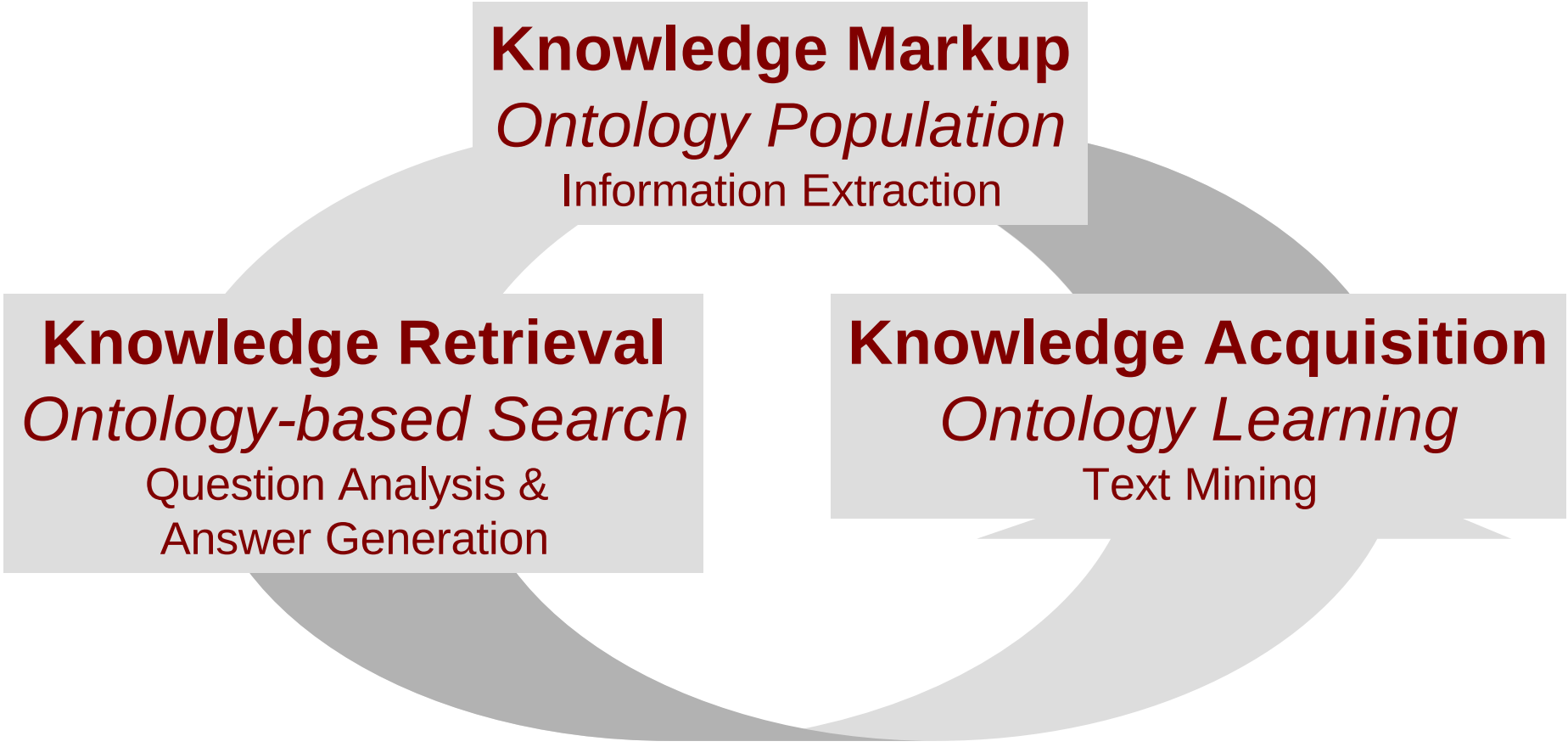


NLP in the Ontology Life Cycle

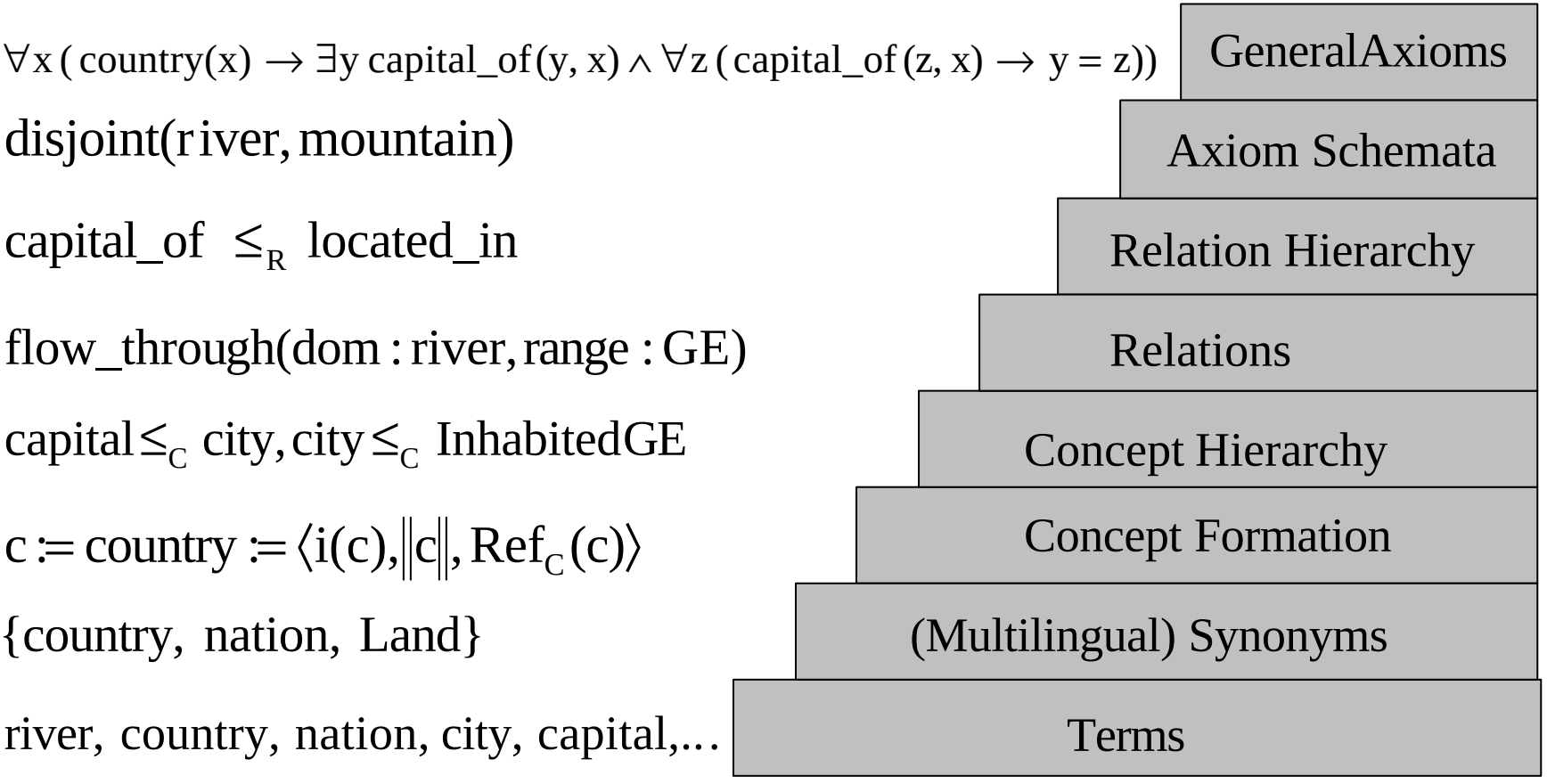
NLP in the Ontology Life Cycle



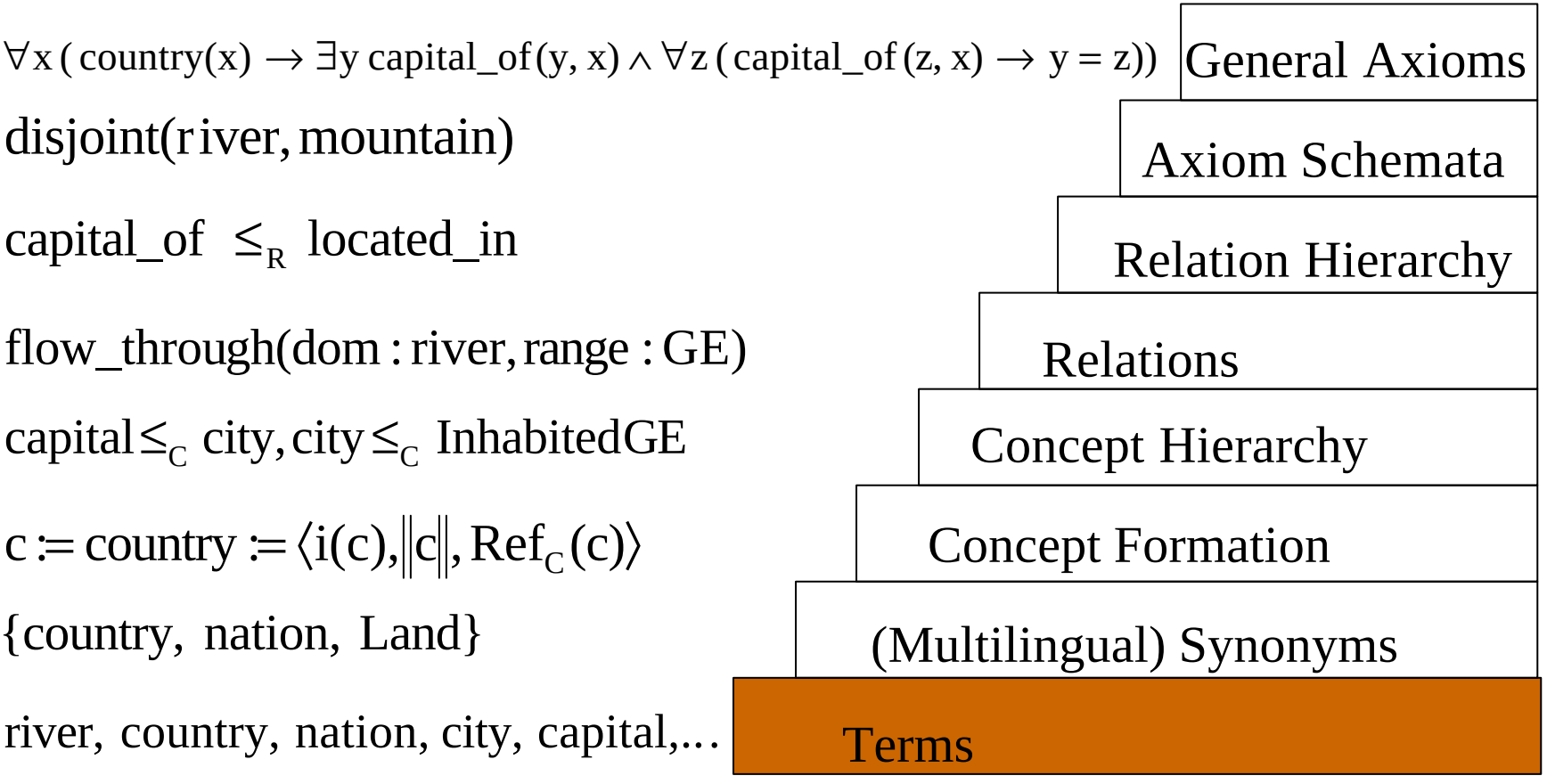
NLP in the Ontology Life Cycle



Ontology Learning Layer Cake



Ontology Learning Layer Cake



Terms

Terms are at the basis of the ontology learning process

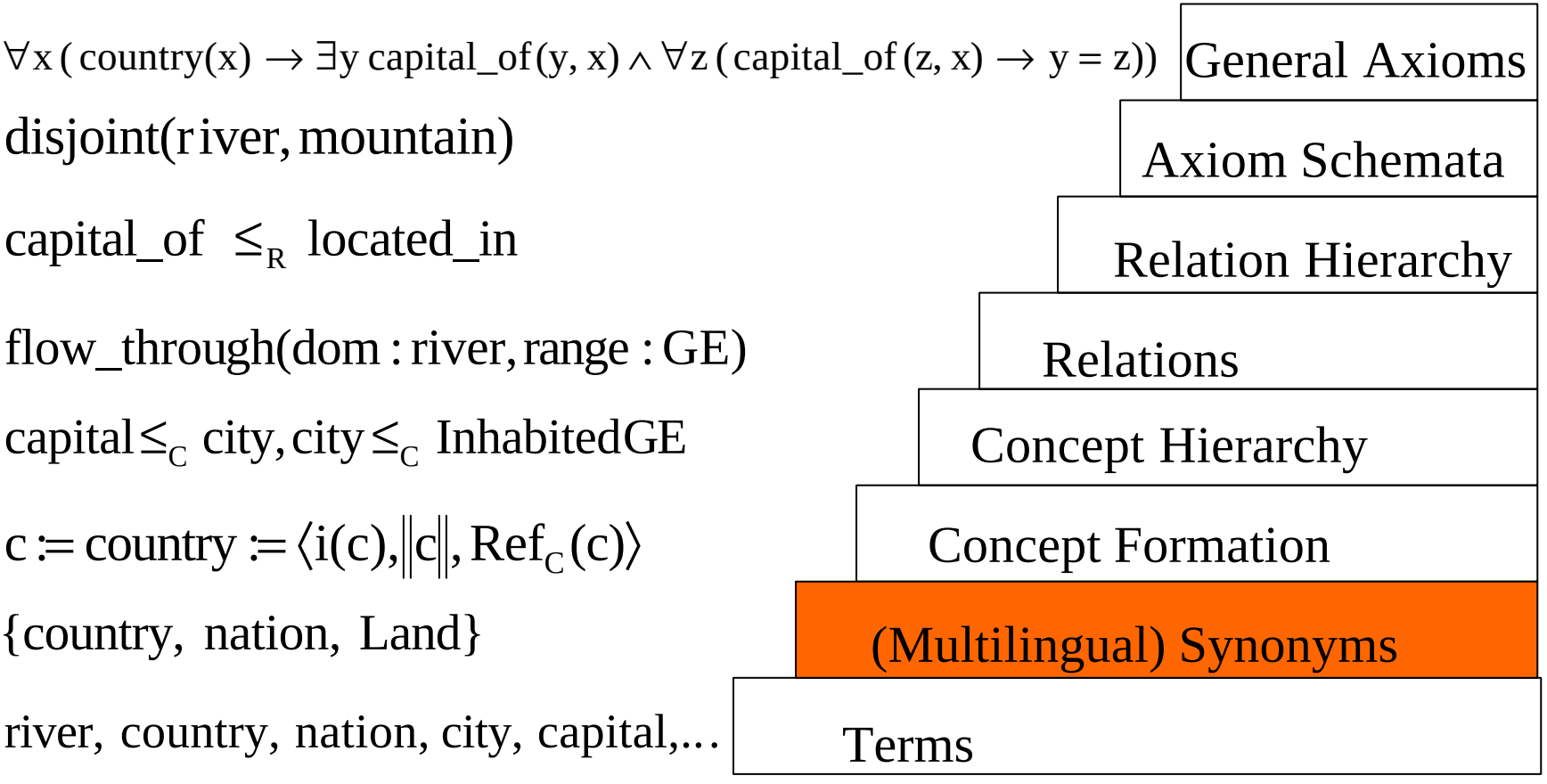
- Terms express more or less complex semantic units
- But what is a term?

Huge Selection of Top Brand Computer Terminals Available for Immediate Delivery
Because Vecmar carries such a large inventory of high-quality computer terminals, including: ADDS terminals, Boundless terminals, DEC terminals, HP terminals, IBM terminals, LINK terminals, NCR terminals and Wyse terminals, your order can often ship same day. Every computer terminal shipped to you is protected with careful packing, including thick boxes. All of our shipping options - including international - are available through major carriers.

- Extracted term candidates (phrases)

  - computer
  - terminal
  - computer terminal
  - ? high-quality computer terminal
  - ? top brand computer terminal
  - ? HP terminal, DEC terminal, …

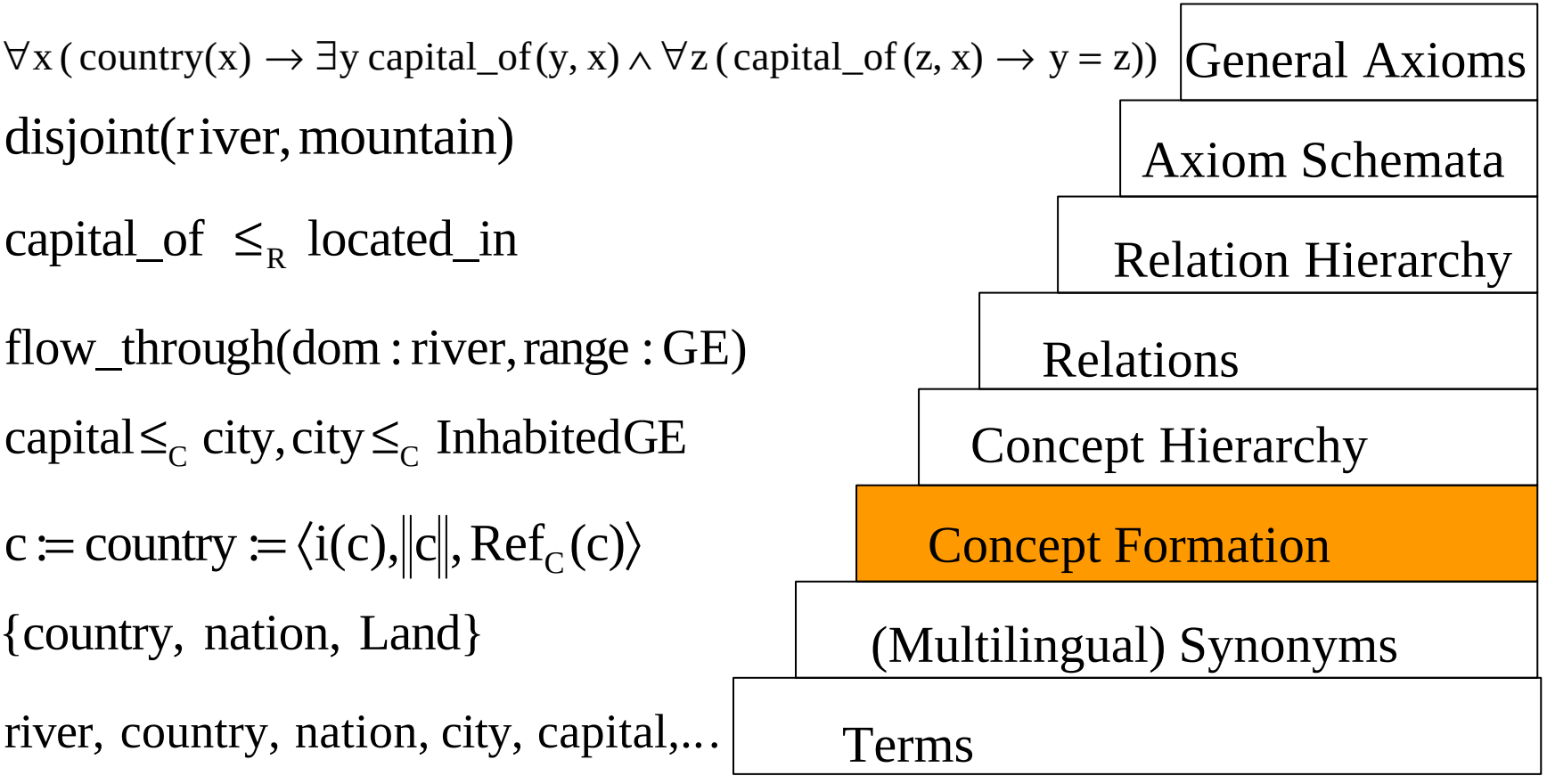
Ontology Learning Layer Cake



(Multilingual) Synonyms

- Next step in ontology learning is to identify terms that share (some) semantics, i.e., potentially refer to the same concept
- Synonyms (Within Languages)
  - '100% synonyms' don't exist – only term pairs with *similar* meanings
  - Examples from http://thesaurus.com
    - `terminal – video display – input device`
    - `graphics terminal - video display unit - screen`
- Translations (Between Languages)
  - '100% translations' don't exist - only multilingual term pairs with *similar* meanings
  - Examples from http://dict.leo.org
    - `input device` (English) – `Eingabegerät` (German)
    - Back to English: `input device, input unit, signal conditioning device`
    - `video display unit` (English) – `Videosichtgerät` (German)

Ontology Learning Layer Cake



Concept Formation

A term may indicate a concept, if we can define its

- Intension
  - (in)formal definition of the set of objects that this concept describes
    - *a disease is an impairment of health or a condition of abnormal functioning*
- Extension
  - a set of objects (instances) that the definition of this concept describes
    - *influenza, cancer, heart disease, …*
- Lexical Realizations
  - the term itself and its multilingual synonyms
    - *disease, illness, Krankheit, maladie, …*

Concept Formation

A term may indicate a concept, if we can define its

- Intension
  - (in)formal definition of the set of objects that this concept describes
    - *a disease is an impairment of health or a condition of abnormal functioning*
- Extension
  - a set of objects (instances) that the definition of this concept describes
    - *influenza, cancer, heart disease, …*

  Discussion: what is an instance? - '*heart disease*' or '*my uncle's heart disease*'
- Lexical Realizations
  - the term itself and its multilingual synonyms
    - *disease, illness, Krankheit, maladie, …*

Concept Formation

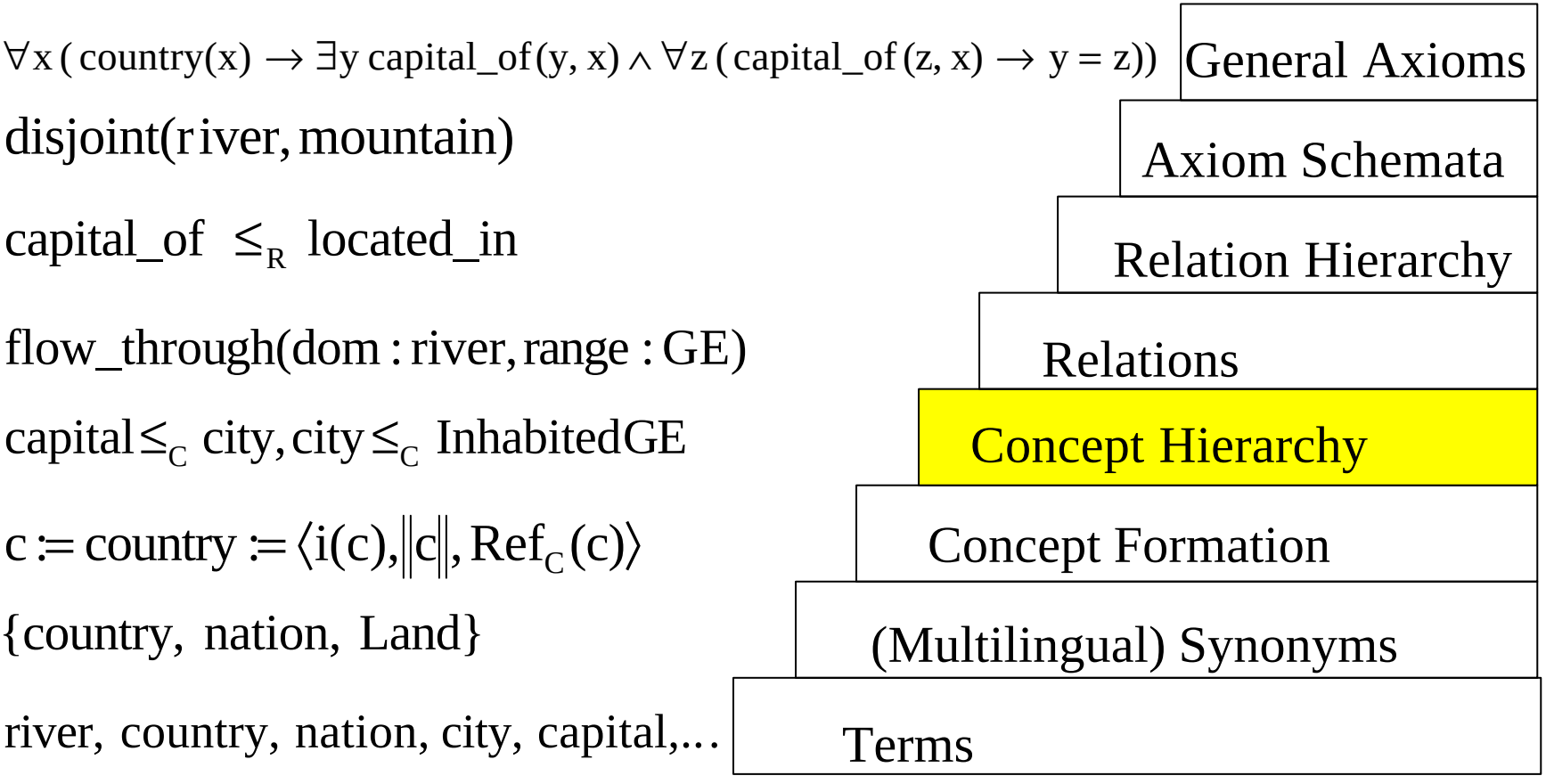
A term may indicate a concept, if we can define its

- Intension
  - (in)formal definition of the set of objects that this concept describes
    - *a disease is an impairment of health or a condition of abnormal functioning*
- Extension
  - a set of objects (instances) that the definition of this concept describes
    - *influenza, cancer, heart disease, …*

  Discussion: what is an instance? - '*heart disease*' or '*my uncle's heart disease*'
- Lexical Realizations
  - the term itself and its multilingual synonyms
    - *disease, illness, Krankheit, maladie, …*

  Discussion: synonyms vs. instances – '*disease*', '*heart disease*', '*cancer*', …

Ontology Learning Layer Cake



Concept Hierarchy

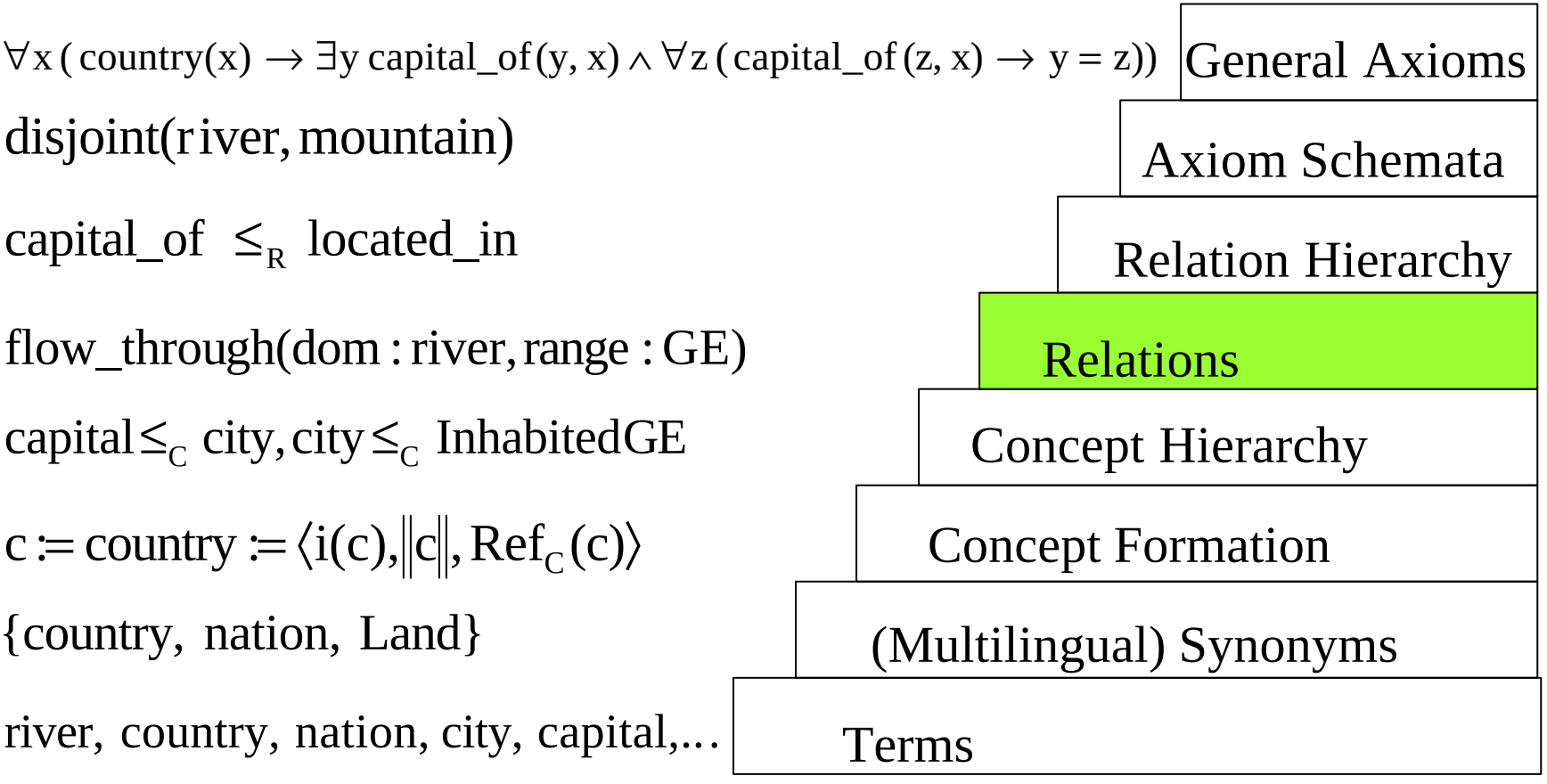
- Clustering
  - Collect context information and represent it as a vector
  - Compute similarity among vectors wrt. a measure

- Lexico-Syntactic Patterns
  - *"A sea bass **is** a fish" > bass ISA fish*
  - *"Vehicles **such as** cars, trucks and bikes" > car ISA vehicle*
  - *"Publications, **especially** books > book ISA publication*

- Linguistic Structure
  - Modifiers that narrow down the meaning of the head noun

"international credit card" > international credit card ISA credit card

Ontology Learning Layer Cake



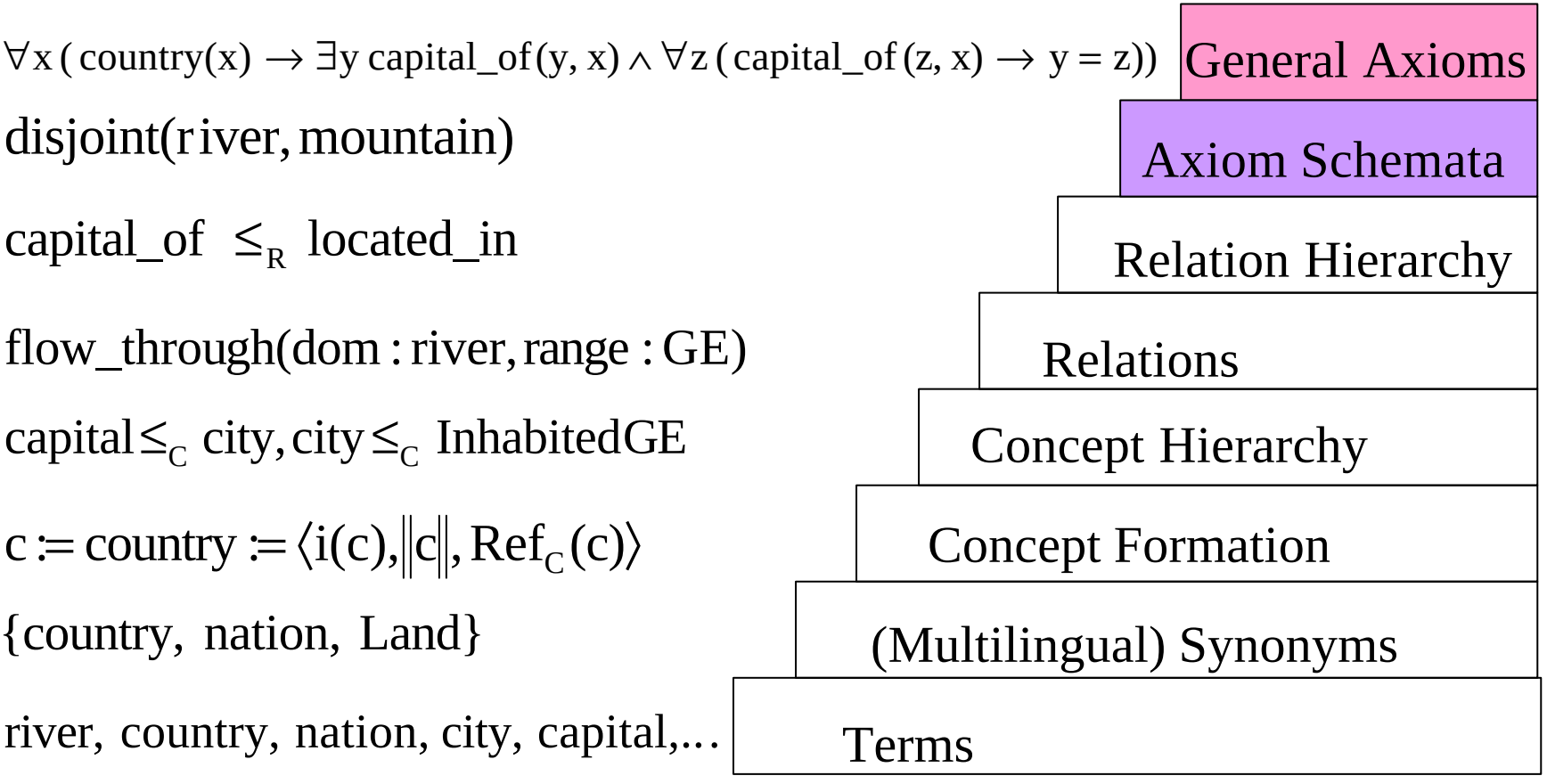
Types of Relations/Attributes

- Part-of [Charniak et al. 98]
  - X consists of Y
- Causation [Girju 02]
  - X leads to Y
- Attributes [Poesio and Almuhareb 05]
  - Qualities (e.g. *color of a car*)
  - Parts (e.g. *hood of a car*)
  - Related-Objects (e.g. *the track of the deer*)
  - Activities (e.g. *the repairing of the car*)
  - Related-Agents (e.g. *the driver of the car*)
  - Non-Attributes (e.g. *the majority of the deer*)

Methods

- Association Rules, e.g. [Maedche/Staab00]
  - Finds correlating strings/concepts, e.g. *hotel – beach*
  - Unnamed relations
- Linguistic Structure, e.g. [Schutz/Buitelaar05]
  - Maps predicate-argument structure to ontological relation
  - Named relations

Ontology Learning Layer Cake



Axiom Schemata & General Axioms

- DIRT (Discovery of Inference Rules from Text) [Lin et al. 01]
  - *„X solves Y“ > Y is solved by X, X resolves Y, X finds a solution to Y, X tries to solve Y, Y deals with X, Y is resolved by X, X addresses Y, X seeks a solution to Y, X do something about Y, …*
- Discovering Disjointness Axioms on the basis of coordination [Haase and Völker 05]
  - *“men and women” > disjoint(man,woman)*

Tools

Tools – An Overview

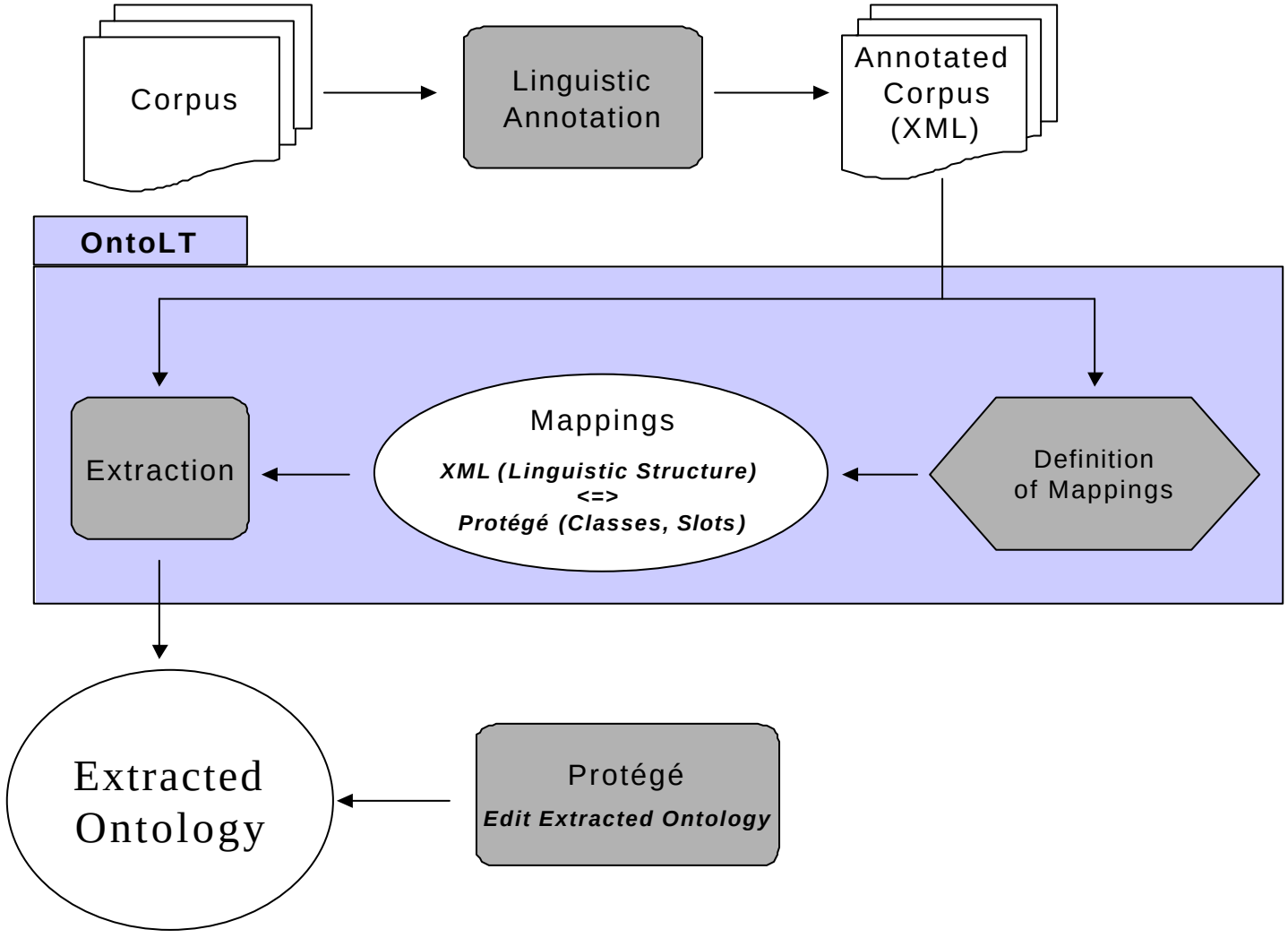
Organization	System	Ontology Learning Layers							
		Terms	*Synonyms*	*Concept Formation*	*Concept Hierarchy*	*Relations*	*Relation Hierarchy*	*Axioms Schemata*	*General Axioms*
AIFB, Univ. Karlsruhe	*Text2Onto*	X	clusters	intension	X	X			
	AEON							X	
Amir Kabir Univ. Tehran	*HASTI*	X			X	X			X
CNTS, Univ. Antwerpen	*OntoBasis*		clusters	clusters		?			
DFKI	*OntoLT / RelExt*	X			X	X			
Economic Univ. Prague	*TextToOnto ++*					labels			
ISI, USC	*CBC*		clusters	clusters					
	DIRT								X
Keio Univ.	*DODDLE*					X			
NRC-CNRC	*PMI-IR*		X						
Univ. de Paris-Sud	*ASIUM / Mo'k*		clusters	clusters	X	X			
Univ. di Roma	*OntoLearn*	X	X	intension	X	X			
Univ. of Salford	*ATRACT*	X	clusters	clusters					
Univ. Zürich	*AirKM*	X			X				

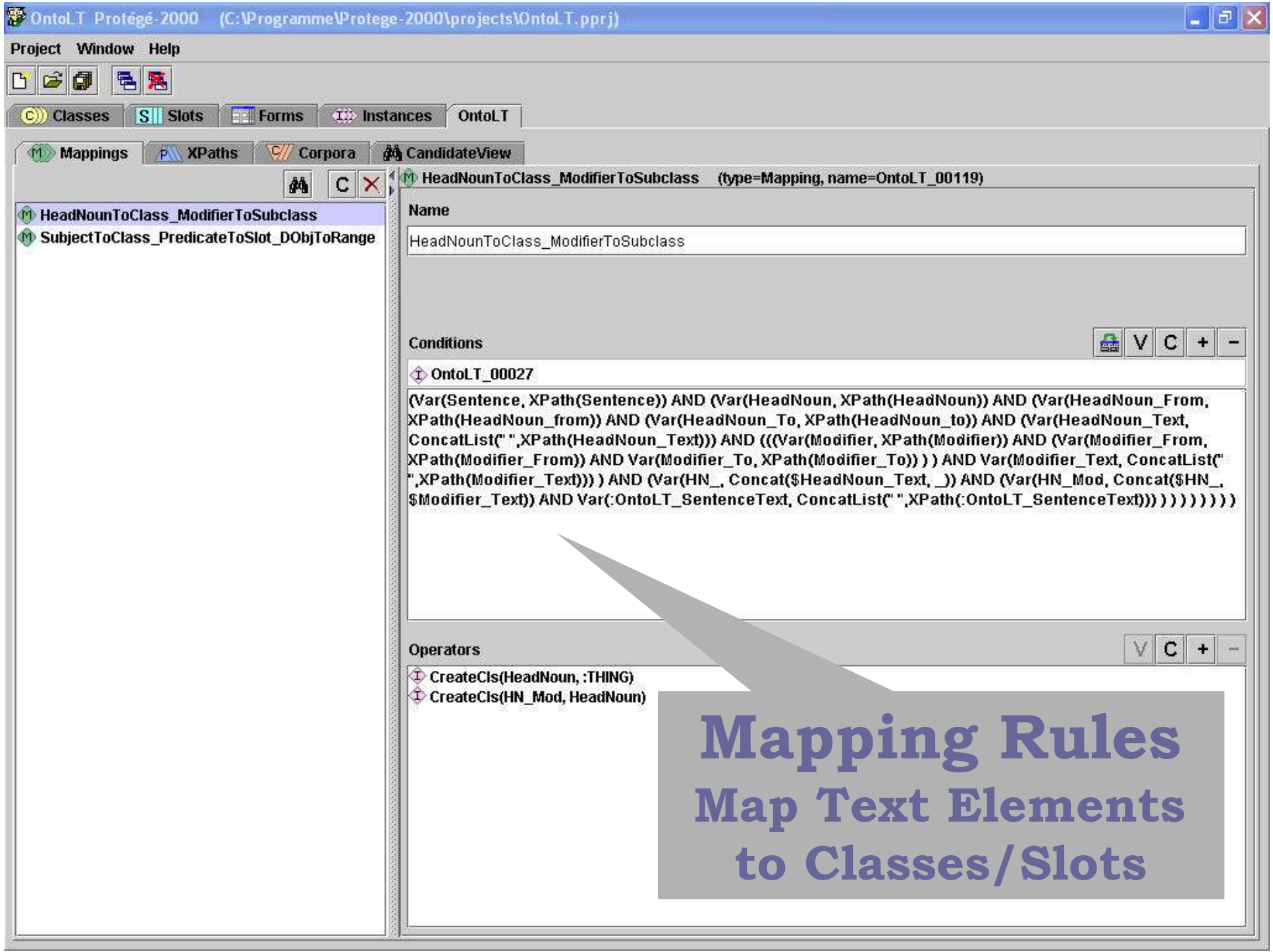
OntoLT

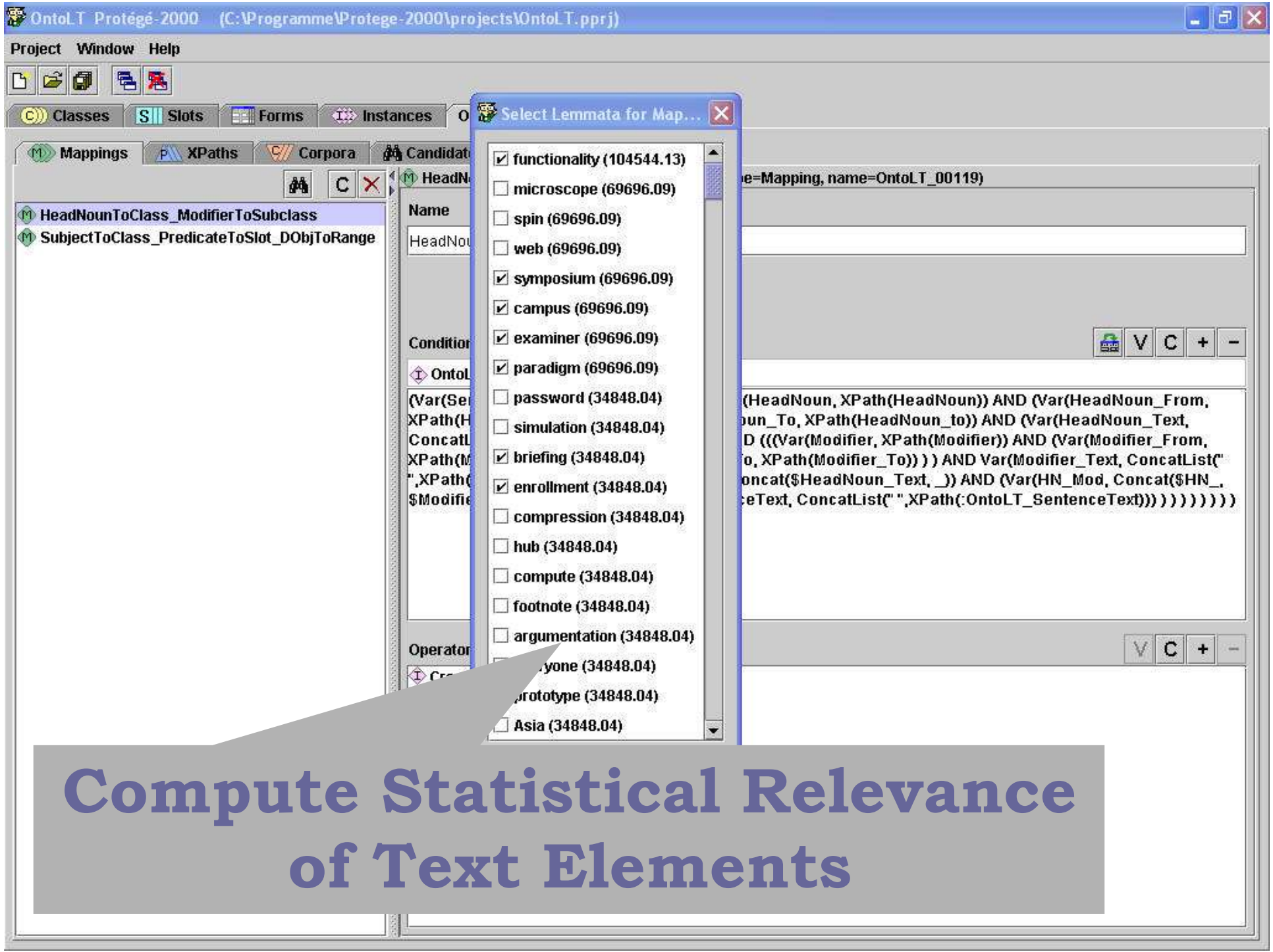
Joint work with Michael Sintek, Daniel Olejnik

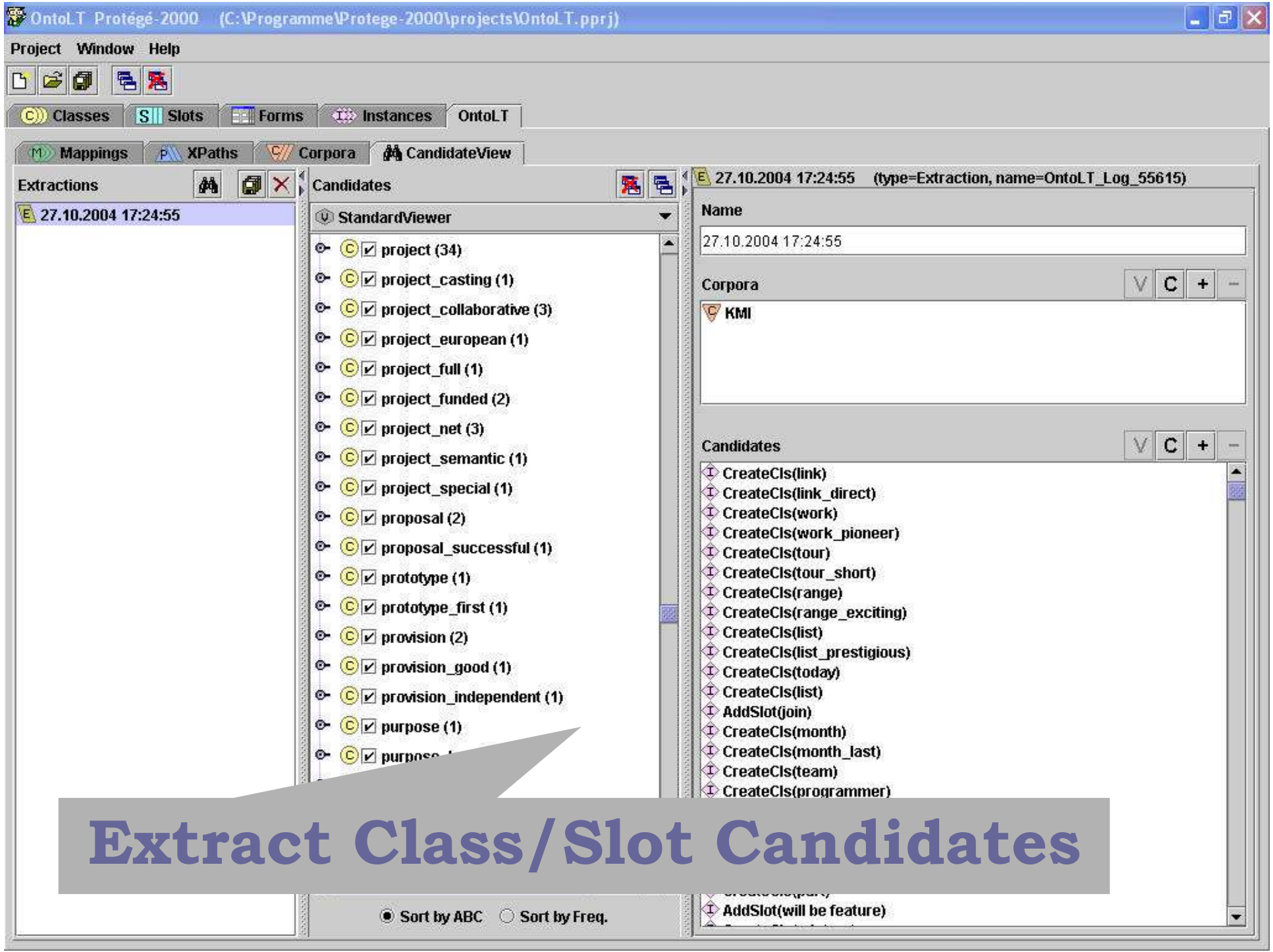
- Methods
  - Term extraction by statistical methods ($?^2$)
  - Term and relation extraction by mapping linguistic structure to ontological structure
- Open source (Java, Protégé plugin)
- Link: *http://olp.dfki.de/OntoLT/OntoLT.htm*

Architecture

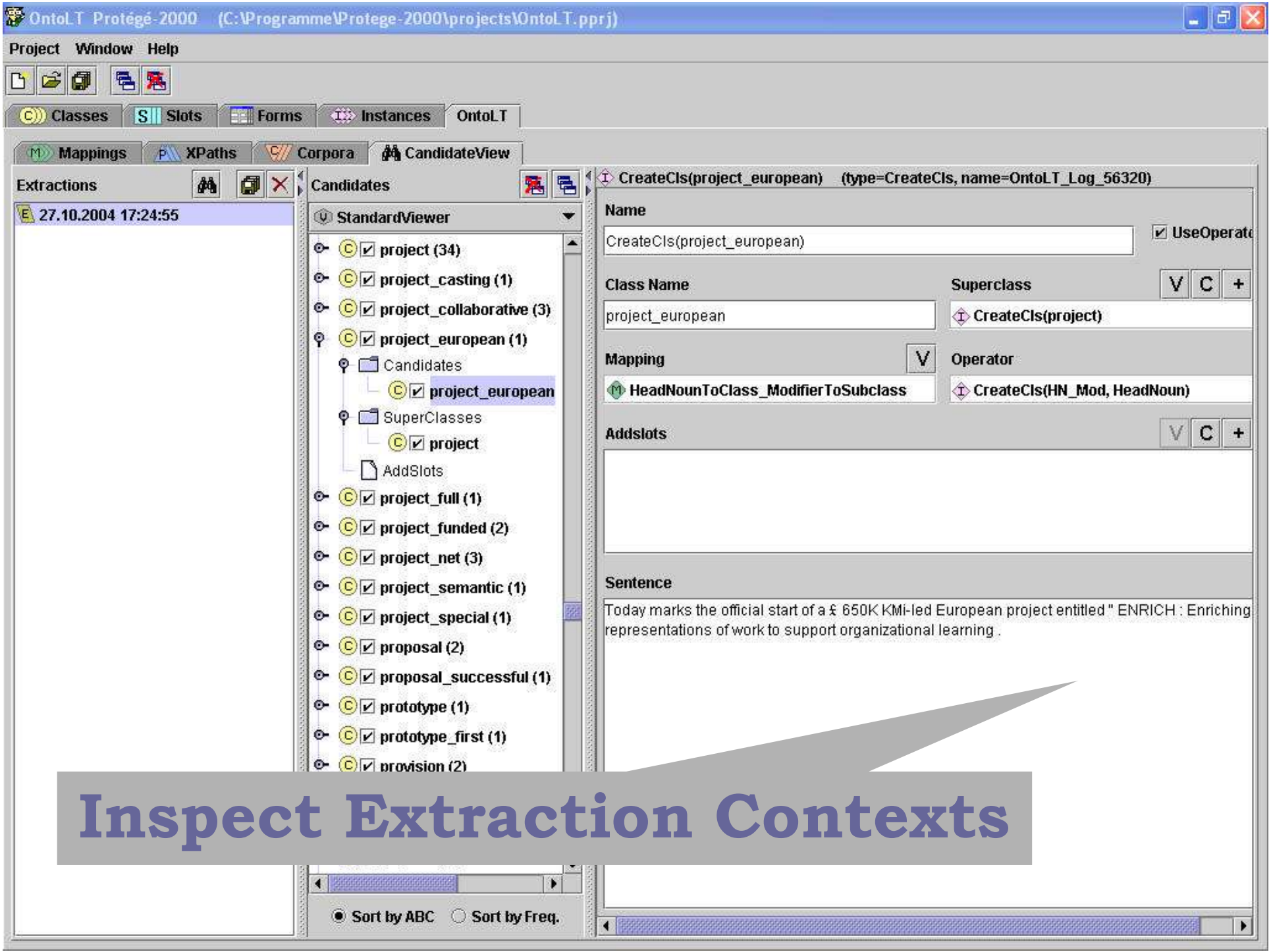


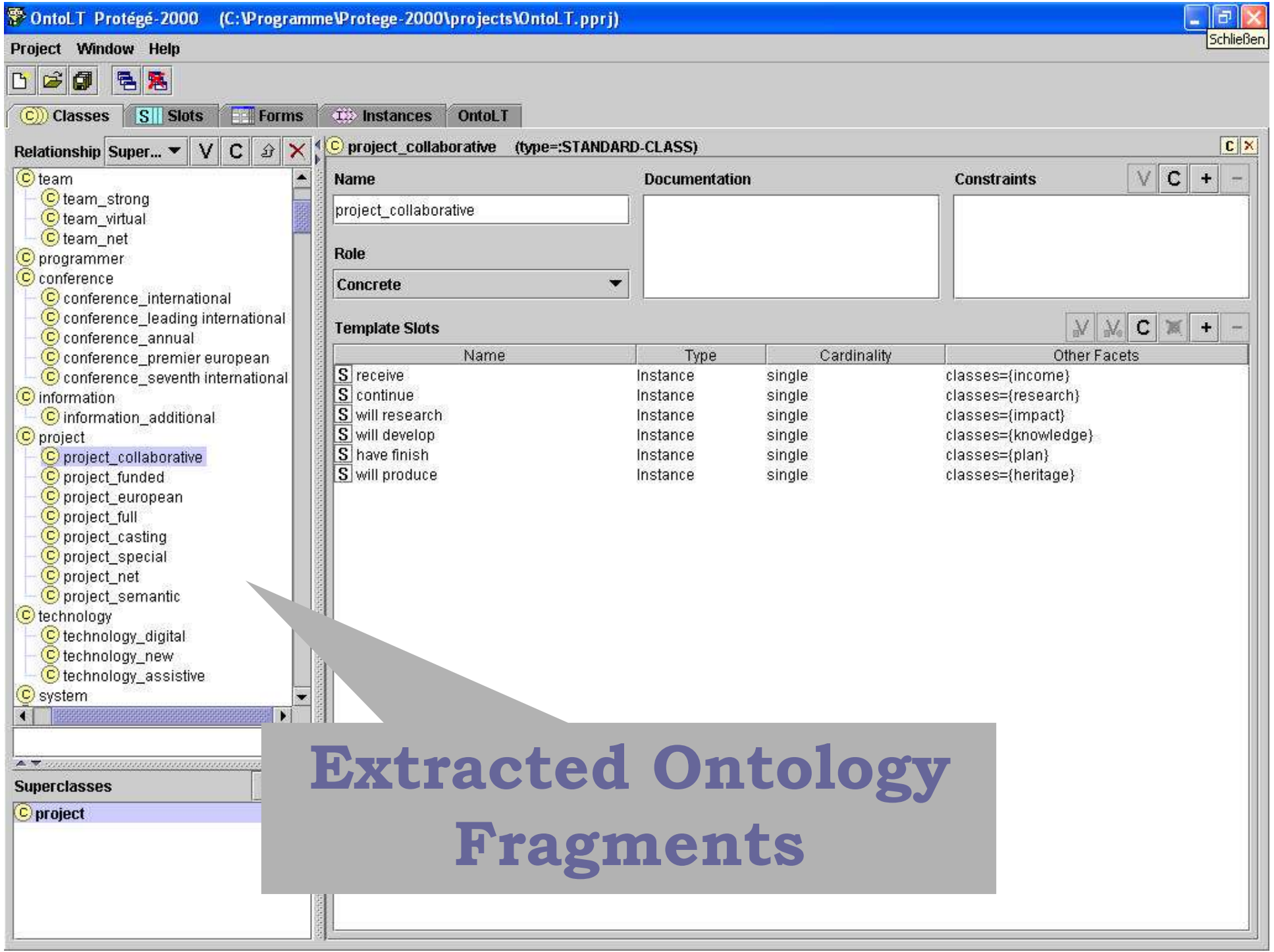

OntoLT Protégé-2000 (C:\Programme\Protege-2000\projects\OntoLT.pprj)
Project Window Help
Classes Slots Forms Instances OntoLT
Mappings XPaths Corpora CandidateView
HeadNounToClass_ModifierToSubclass
SubjectToClass_PredicateToSlot_DObjToRange
HeadNounToClass_ModifierToSubclass (type=Mapping, name=OntoLT_00119)
Name
HeadNounToClass_ModifierToSubclass
Conditions
V C + −
OntoLT_00027
(Var(Sentence, XPath(Sentence)) AND (Var(HeadNoun, XPath(HeadNoun)) AND (Var(HeadNoun_From, XPath(HeadNoun_from)) AND (Var(HeadNoun_To, XPath(HeadNoun_to)) AND (Var(HeadNoun_Text, ConcatList(" ",XPath(HeadNoun_Text))) AND (((Var(Modifier, XPath(Modifier)) AND (Var(Modifier_From, XPath(Modifier_From)) AND Var(Modifier_To, XPath(Modifier_To)))) AND Var(Modifier_Text, ConcatList(" ",XPath(Modifier_Text)))) AND (Var(HN_, Concat($HeadNoun_Text, _)) AND (Var(HN_Mod, Concat($HN_, $Modifier_Text)) AND Var(:OntoLT_SentenceText, ConcatList(" ",XPath(:OntoLT_SentenceText))))))))))))
Operators
V C + −
CreateCls(HeadNoun, :THING)
CreateCls(HN_Mod, HeadNoun)
Mapping Rules
Map Text Elements
to Classes/Slots


OntoLT Protégé-2000 (C:\Programme\Protege-2000\projects\OntoLT.pprj)
Project Window Help
Classes Slots Forms Instances
Mappings XPaths Corpora
HeadNounToClass_ModifierToSubclass
SubjectToClass_PredicateToSlot_DObjToRange
Name
Select Lemmata for Map...
functionality (104544.13)
microscope (69696.09)
spin (69696.09)
web (69696.09)
symposium (69696.09)
campus (69696.09)
examiner (69696.09)
paradigm (69696.09)
password (34848.04)
simulation (34848.04)
briefing (34848.04)
enrollment (34848.04)
compression (34848.04)
hub (34848.04)
compute (34848.04)
footnote (34848.04)
argumentation (34848.04)
Asia (34848.04)
Compute Statistical Relevance of Text Elements



Extract Class/Slot Candidates


OntoLT Protégé-2000 (C:\Programme\Protege-2000\projects\OntoLT.pprj)
Project Window Help
Classes Slots Forms Instances OntoLT
Mappings XPaths Corpora CandidateView
Extractions
27.10.2004 17:24:55
Candidates
StandardViewer
project (34)
project_casting (1)
project_collaborative (3)
project_european (1)
Candidates
project_european
SuperClasses
project
AddSlots
project_full (1)
project_funded (2)
project_net (3)
project_semantic (1)
project_special (1)
proposal (2)
proposal_successful (1)
prototype (1)
prototype_first (1)
provision (2)
Sort by ABC
Sort by Freq.
CreateCls(project_european) (type=CreateCls, name=OntoLT_Log_56320)
Name
CreateCls(project_european)
UseOperat
Class Name
project_european
Superclass
CreateCls(project)
Mapping
HeadNounToClass_ModifierToSubclass
Operator
CreateCls(HN_Mod, HeadNoun)
Addslots
Sentence
Today marks the official start of a £ 650K KMi-led European project entitled " ENRICH : Enriching representations of work to support organizational learning .
Inspect Extraction Contexts


OntoLT Protégé-2000 (C:\Programme\Protege-2000\projects\OntoLT.pprj)
Schließen
Project Window Help
Classes Slots Forms Instances OntoLT
Relationship Super...
team
team_strong
team_virtual
team_net
programmer
conference
conference_international
conference_leading international
conference_annual
conference_premier european
conference_seventh international
information
information_additional
project
project_collaborative
project_funded
project_european
project_full
project_casting
project_special
project_net
project_semantic
technology
technology_digital
technology_new
technology_assistive
system
Superclasses
project
project_collaborative (type=:STANDARD-CLASS)
Name
project_collaborative
Documentation
Constraints
Role
Concrete
Template Slots
Name
Type
Cardinality
Other Facets
receive Instance single classes={income}
continue Instance single classes={research}
will research Instance single classes={impact}
will develop Instance single classes={knowledge}
have finish Instance single classes={plan}
will produce Instance single classes={heritage}
Extracted Ontology Fragments

RelExt

Joint work with Alexander Schutz

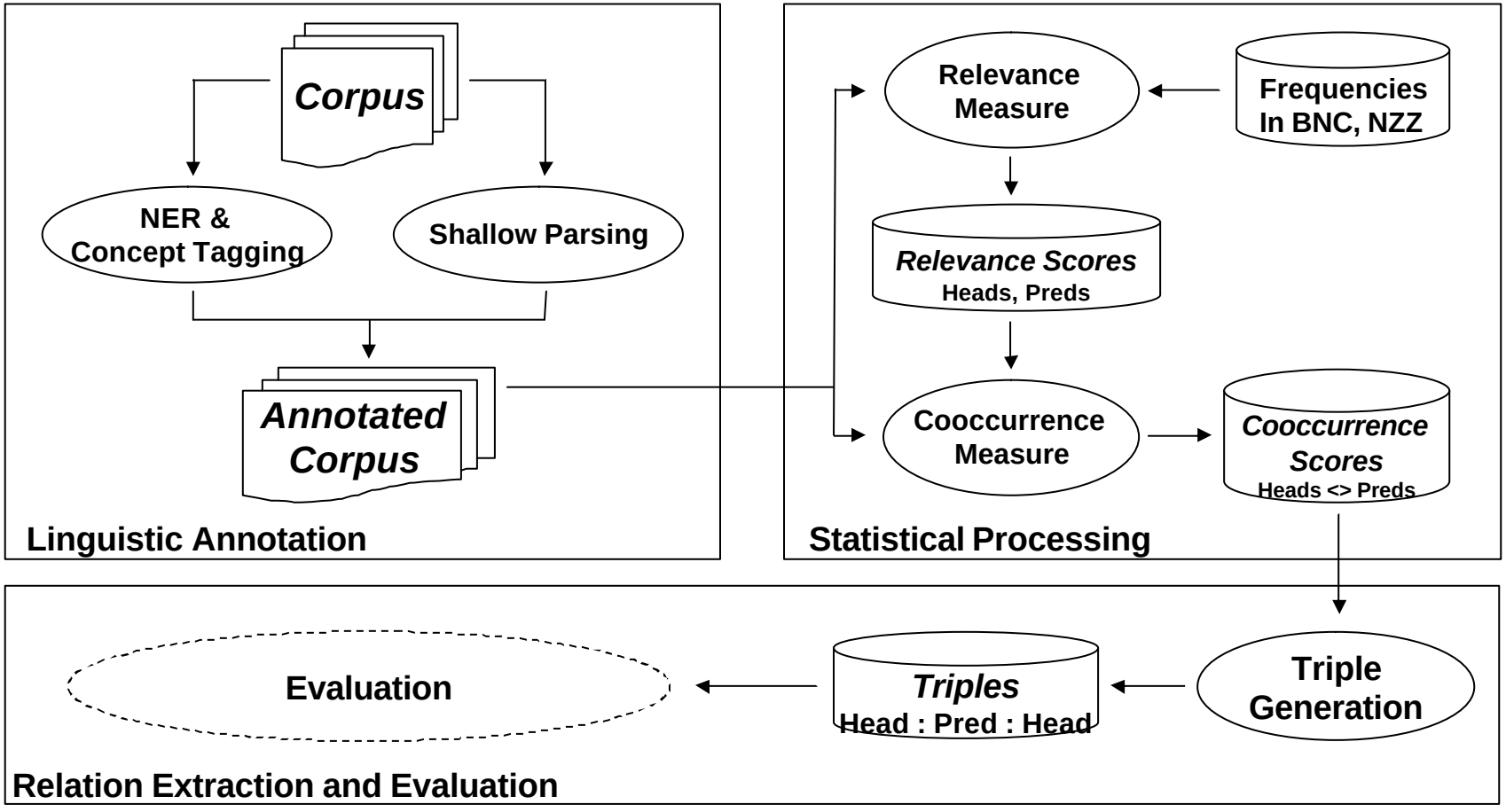
- Context ...
  - SmartWeb SportEvents Ontology
    - ~ 400 Classes in the Soccer Domain
    - ~ 60 Relations, e.g. *hasName, atMinute*, …
- ... and Motivations
  - Extend SmartWeb SportEvents Ontology with Events

    Example:

    "Ballack shoots the ball in the net."

    Relation: Shoot (**Domain:** FootballPlayer **Range:** BallObject)

System Architecture



Linguistic Annotation

- Shallow Parsing
  - Part-of-Speech Tagging

    Fussballspieler (soccer player) : Noun
  - Morphological Analysis

    Fussballspieler : Fussball - Spieler
  - Dependency Structure Analysis

    The team *won* *the second match.*

    SUBJECT PREDICATE DIRECT_OBJECT
- Named-Entity Recognition

  Michael Ballack : FootballPlayer
- Concept Tagging

  Ball (ball), Leder (leather) : BallObject

Relevance Ranking

Rank	$?^2$	Headnoun	Frequency
1	125245.24	Ball *(ball)*	6849
2	121888.52	Tor *(goal)*	7767
3	95003.21	Meter *(meters)*	5967
4	64157.18	Schuss *(shot / drive)*	3575
5	57185.76	Ecke *(corner)*	3132
6	45474.96	Strafraum *(penalty area)*	2298
7	34668.11	Freistoss *(freekick)*	1752
8	30017.75	Leder *(leather / ball)*	1561
9	27989.09	Flanke *(cross)*	1479
10	27414.66	Pfosten *(post)*	1457

Top-10 Head-Nouns before and after mapping to Ontology Classes

Rank	$?^2$	Concept Label	Frequency
1	565510.99	FOOTBALLPLAYER	28494
2	162137.82	GOALOBJECT	8188
3	143528.88	BALLOBJECT	7249
4	138535.44	GOALKEEPER	6887
5	70814.86	SHOT	3578
6	49018.16	TEAM	2477
7	45474.96	PENALTYAREA	2298
8	34668.11	FREEKICK	1752
9	29324.54	WING	1482
10	28829.78	POST	1457

Top-10 Predicates

Rank	$?^2$	Predicate	Frequency
1	27167.41	flanken *(to cross)*	1373
2	22045.39	klaeren *(to clear)*	1435
3	21908.37	schiessen *(to shot)*	1503
4	20439.09	koepfen *(to head)*	1033
5	16342.99	lassen *(to let / to leave)*	826
6	9563.41	ziehen *(to pull / to drag)*	1548
7	9468.57	passen *(to pass / to play)*	814
8	7752.84	spielen *(to play / to pass)*	1559
9	7653.68	lenken *(to divert)*	537
10	7637.45	parieren *(to parry / to save)*	405

Co-Occurrence Analysis

Rank	$?^2$	Predicate	Frequency
1	27167.41	flanken *(to cross)*	1373
2	22045.39	klaeren *(to clear)*	1435
3	21908.37	schiessen *(to shot)*	1503
4	20439.09	koepfen *(to head)*	1033

flanken	SUBJ:FOOTBALLPLAYER	*"Klasnic"*
flanken	DOBJ:FOOTBALLPLAYER	*"Klose"*
flanken_in	PP_ADJ	*"Zuschauer" (audience)*
beschimpfen (to insult)	SUBJ:FOOTBALLPLAYER	"Klasnic"

Rank	$?^2$	Concept Label	Frequency
1	565510.99	FOOTBALLPLAYER	28494
2	162137.82	GOALOBJECT	8188
3	143528.88	BALLOBJECT	7249
4	138535.44	GOALKEEPER	6887

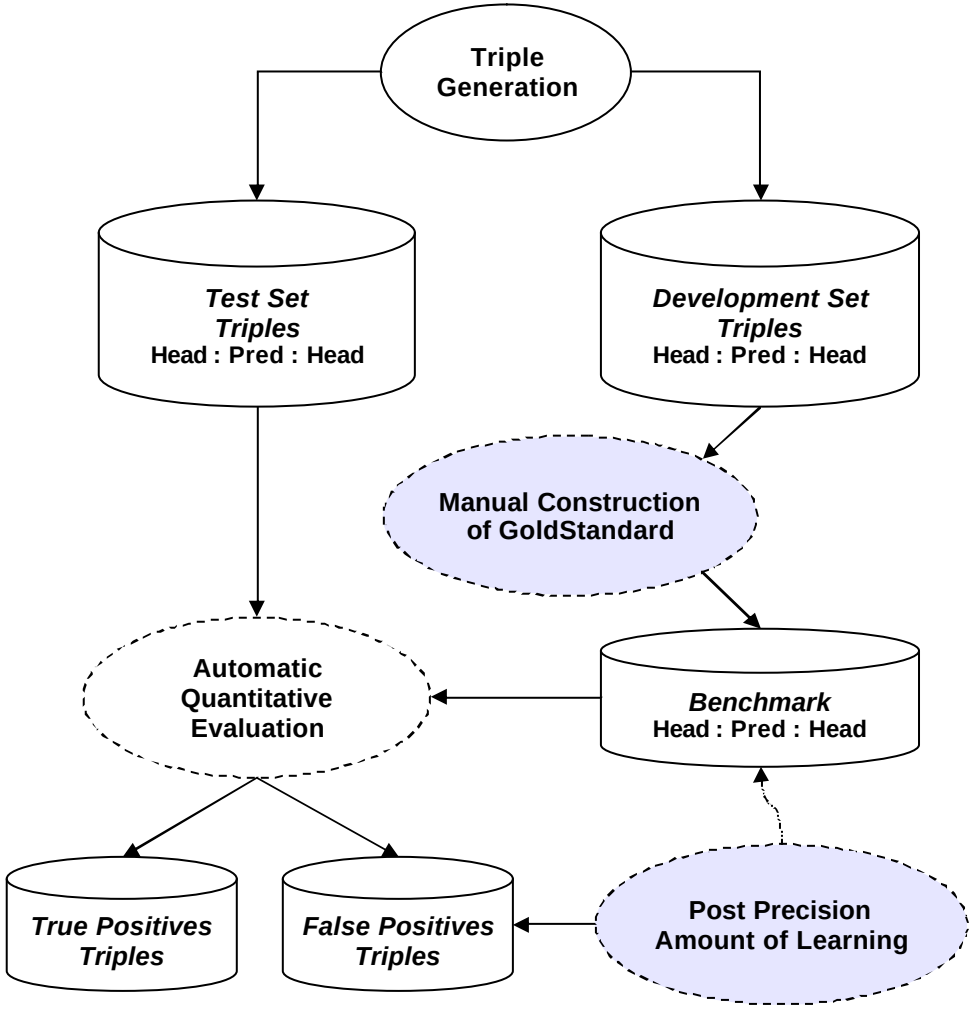
Co-Occurrence Analysis

Rank	$?^2$	Predicate	Frequency
1	27167.41	flanken *(to cross)*	1373
2	22045.39	klaeren *(to clear)*	1435
3	21908.37	schiessen *(to shot)*	1503
4	20439.09	koepfen *(to head)*	1033

flanken	SUBJ:FOOTBALLPLAYER	*"Klasnic"*
flanken	DOBJ:FOOTBALLPLAYER	*"Klose"*
flanken_in	PP_ADJ	*"Zuschauer" (audience)*
beschimpfen (to insult)	SUBJ:FOOTBALLPLAYER	"Klasnic"

Rank	$?^2$	Concept Label	Frequency
1	565510.99	FOOTBALLPLAYER	28494
2	162137.82	GOALOBJECT	8188
3	143528.88	BALLOBJECT	7249
4	138535.44	GOALKEEPER	6887

Evaluation



Gold Standard Construction

- 192 Triples from Development Set presented to 3 Domain Experts
- Classification for each Triple:

  <good, bad, undecided>
- Gold Standard: 38 Triples

Experiment – Evaluation Results

	Corpus	# of Triples Evaluated	Recall	a priori Precision		a posteriori Precision	
				percentage	true positives	percentage	true positives
	1	38	15,80%	8,60%	6	20,00%	14
GS size	2	38	23,70%	13,40%	9	23,90%	16
	3	38	15,80%	8,60%	6	20,00%	14
	Average over Samples		18,43%	10,20%		21,30%	
	1	95	39,50%	12,70%	15	24,60%	29
all Triples	2	84	34,20%	11,90%	13	23,90%	26
	3	92	34,20%	11,10%	13	23,10%	27
	Average over Samples		35,97%	11,90%		23,87%	

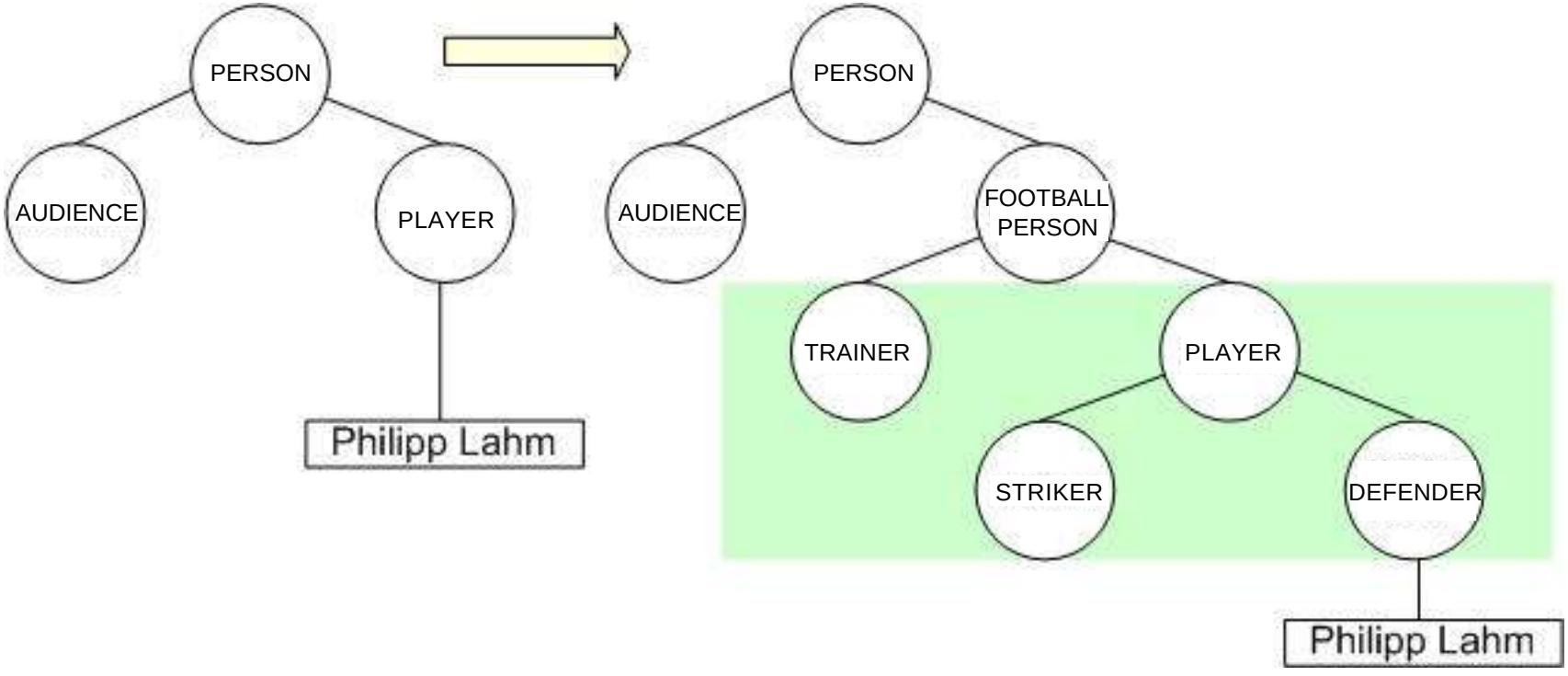
ISOLDE (*Information System for Ontology learning and Domain Exploration*)

Joint work with Nicolas Weber

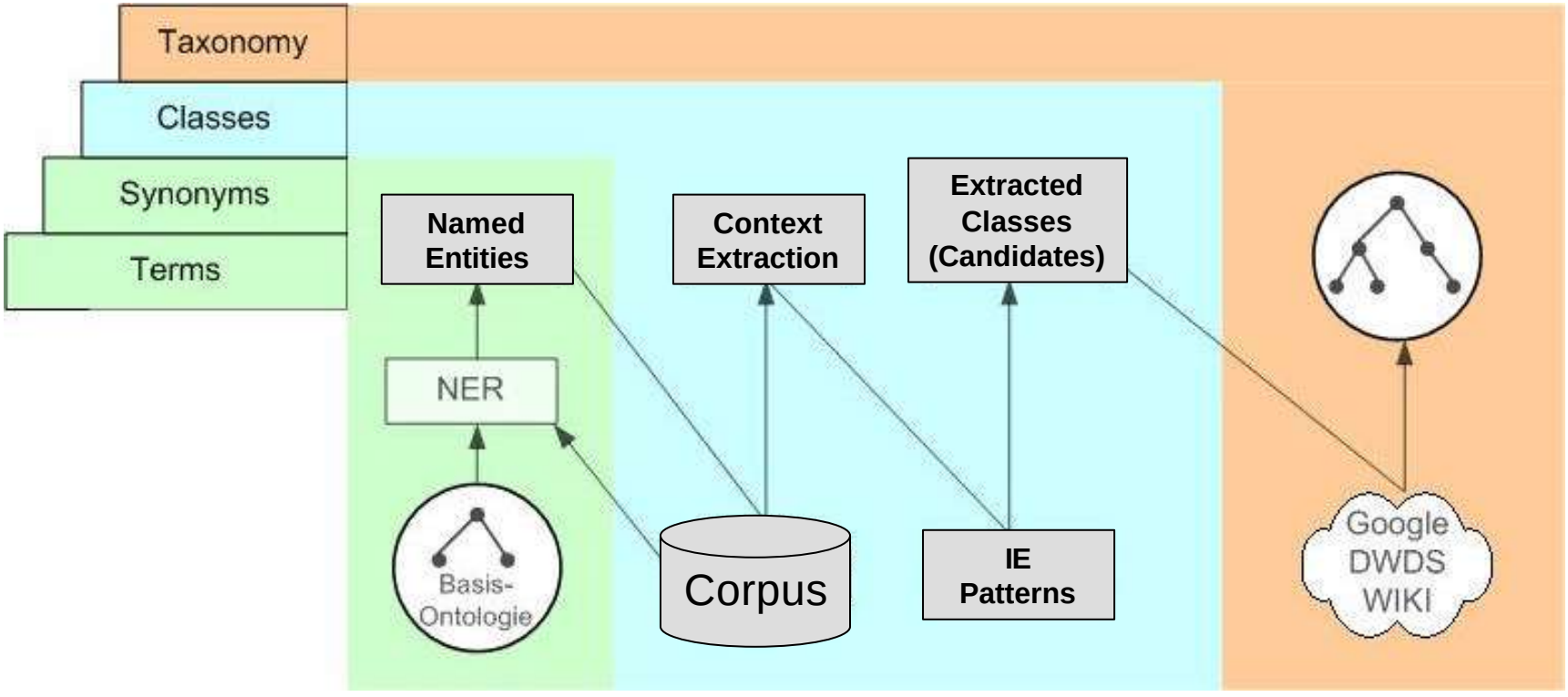
Extend an existing ontology (taxonomy) by:

1. Deriving class candidates from terms in the context of Named-Entities in annotated corpora
2. Deriving additional knowledge for these classes from web resources

Basic Idea


PERSON
AUDIENCE
PLAYER
Philipp Lahm
PERSON
AUDIENCE
FOOTBALL PERSON
TRAINER
PLAYER
STRIKER
DEFENDER
Philipp Lahm

Architecture – OL Layer Cake



NER, Contexts, IE Patterns

„... obwohl ein Trainer wie Klinsmann da anders vorgehen würde ...“
(... although a trainer like Klinsmann would act differently ...)

Named-Entity Recognition

„... obwohl, ein Trainer wie <PERSON> Klinsmann </>...“

Context Extraction / IE Patterns

{NP} [wie] NE

Class Candidate Extraction

Trainer

Web Resources - DWDS

DWDS – Digitales Wörterbuch der Deutschen Sprache
(Digital Dictionary for the German Language)

Web Resources - Wiktionary

Wikiwörterbuch
Wiktionary
['wɪkʃənərɪ], *n*
Das freie Wörterbuch
a wiki-based Open Content dictionary

Navigation
- Hauptseite
- Themenportale
- Zufälliger Eintrag
- Inhaltsverzeichnis

Mitarbeit
- Wiktionary-Portal
- Fehlende Einträge
- Letzte Änderungen

Hilfe
- Fragen?
- Finanzielle Hilfe
- Nutzungshinweise

Suche
Eintrag Suche

Anmelden | Benutzerkonto erstellen

Wörterbucheintrag | Diskussion | Seite bearbeiten | Versionen/Autoren

Torwart

Torwart (Deutsch) [bearbeiten]

Substantiv, *m* [bearbeiten]

Silbentrennung: Torwart, Plural:
Aussprache:
IPA: [...], Plural: [...]
Hörbeispiele: -, Plural: -

Bedeutungen:
[1] Derjenige Fußballspieler, dessen Aufgabe es ist, gegnerische Tore zu vermeiden und der hierfür als einziger Spieler auch seine Hände einsetzen darf.

Abkürzungen:
Herkunft:
aus Tor und Wart

Synonyme:
Tormann, Torhüter, Keeper

Gegenwörter:
Feldspieler, Verteidiger, Abwehrspieler, Mittelfeldspieler, Stürmer, Offensivspieler

Oberbegriffe:
Sport, Fußball, Fußballspieler

Kasus	Singular	Plural
Nominativ	der Torwart	die Torleute
Genitiv	des Torwarts	der Torleute
Dativ	dem Torwart	den Torleuten
Akkusativ	den Torwart	die Torleute

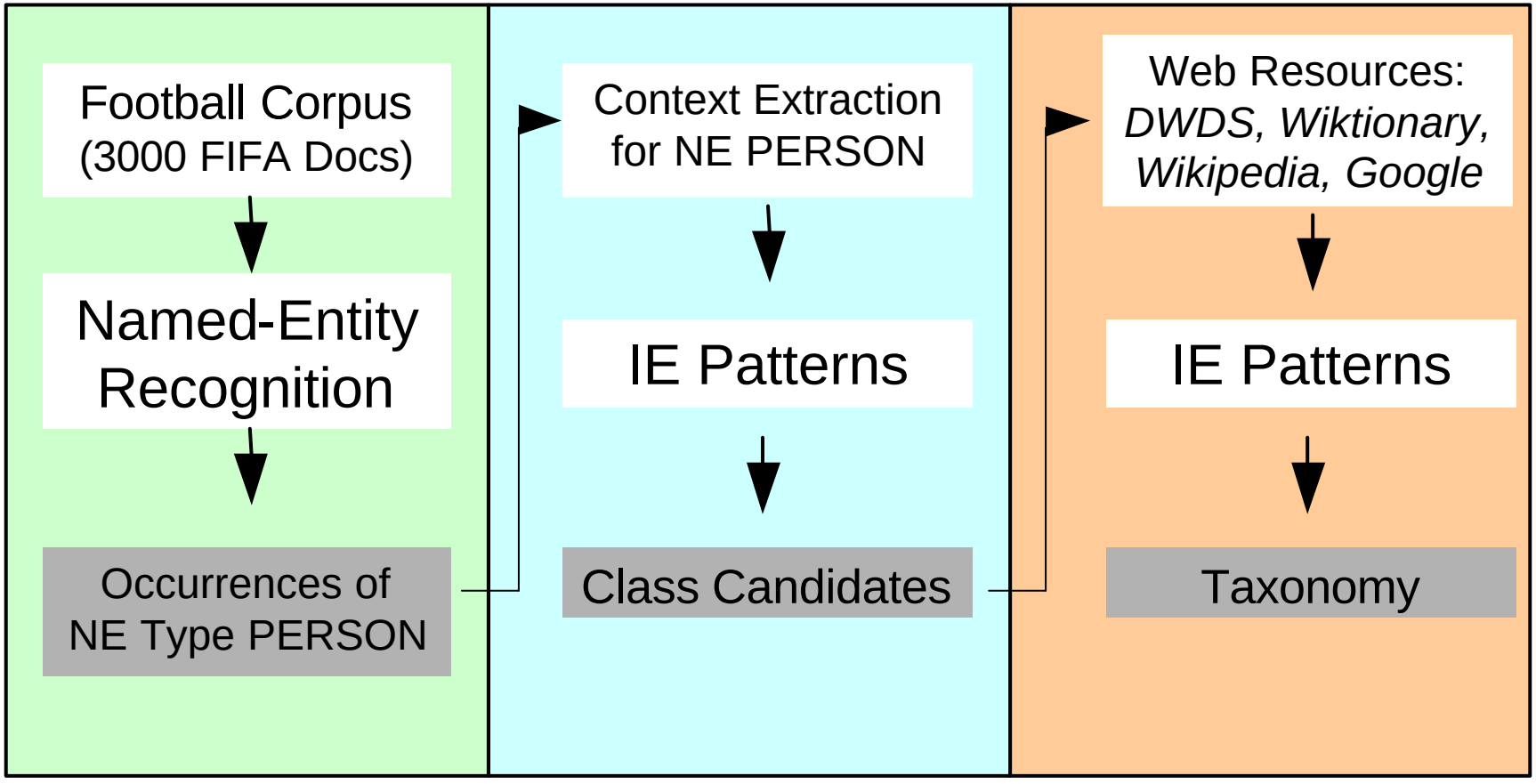
Web Resources - Wikipedia

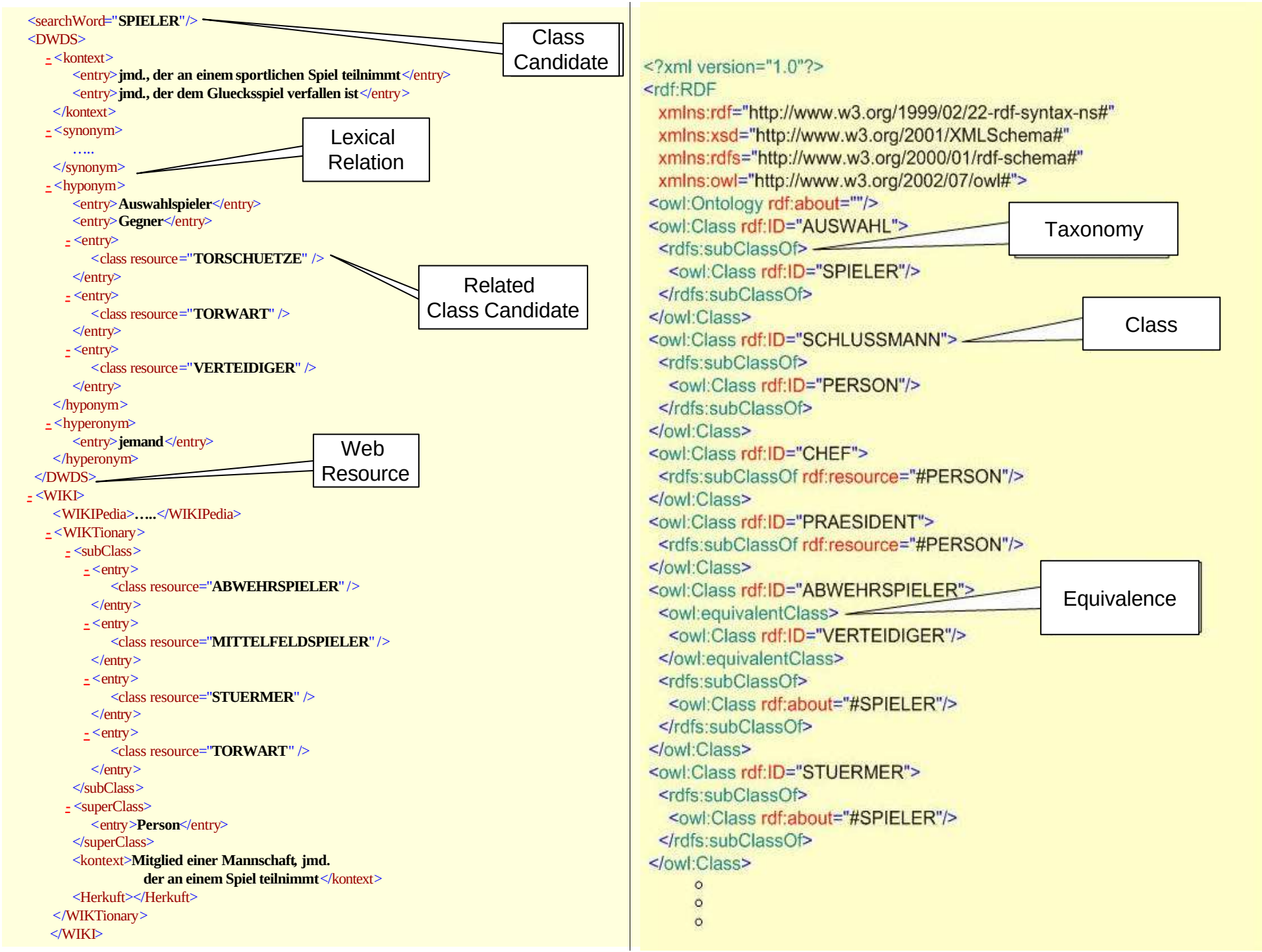
Der **Torwart** (Torhüter, Tormann, Keeper, Schweiz. Goalie) **ist ein Mitspieler** einer Mannschaftssportart. Er ist der defensivste Spieler seiner Mannschaft und seine Hauptaufgabe besteht darin zu verhindern, dass das Spielgerät (z.B. ein Ball) ins Tor der eigenen Mannschaft ....

> **Torwart ist ein Mitspieler**

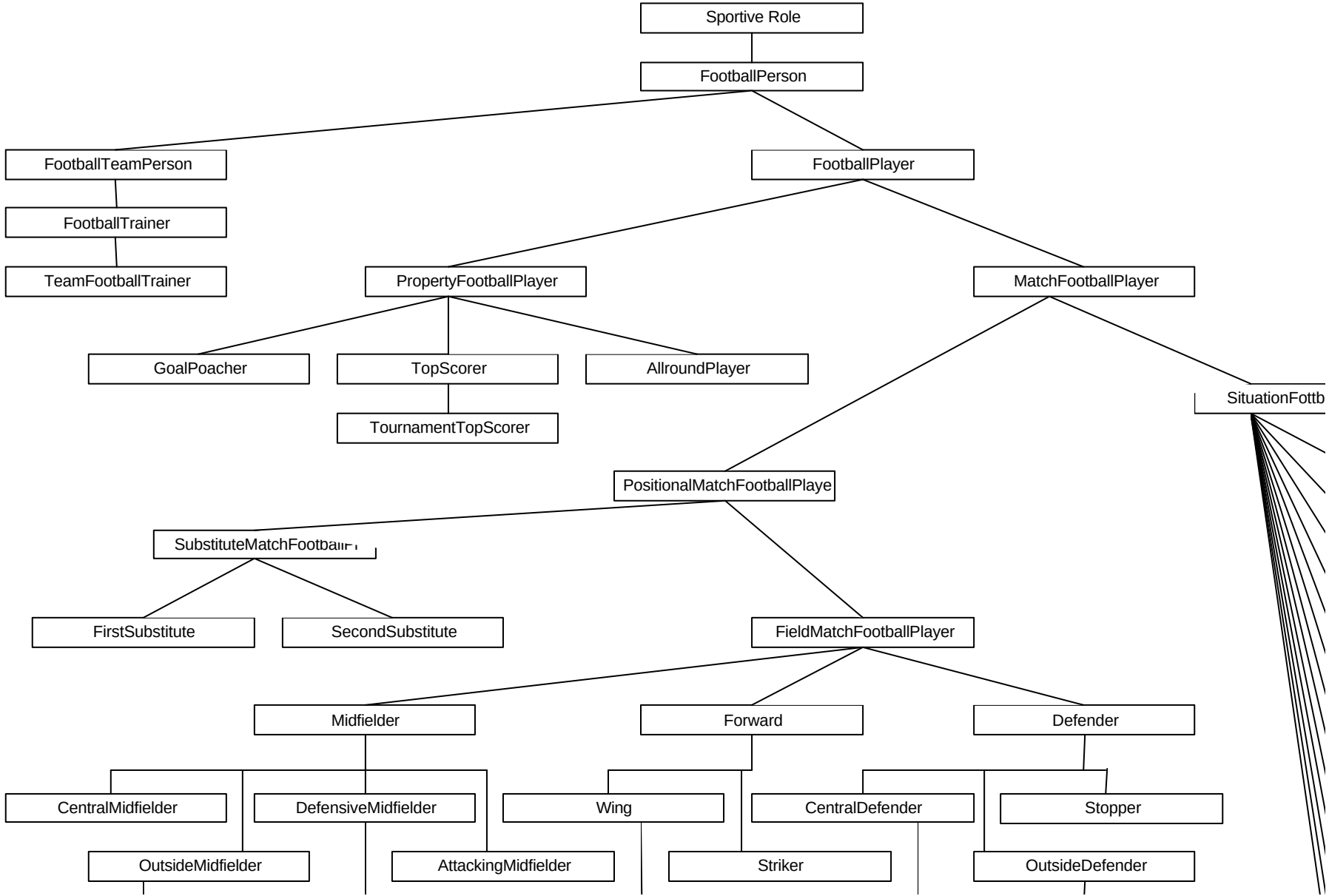
> ***SubClassOf (Torwart, Spieler)***

Experiment

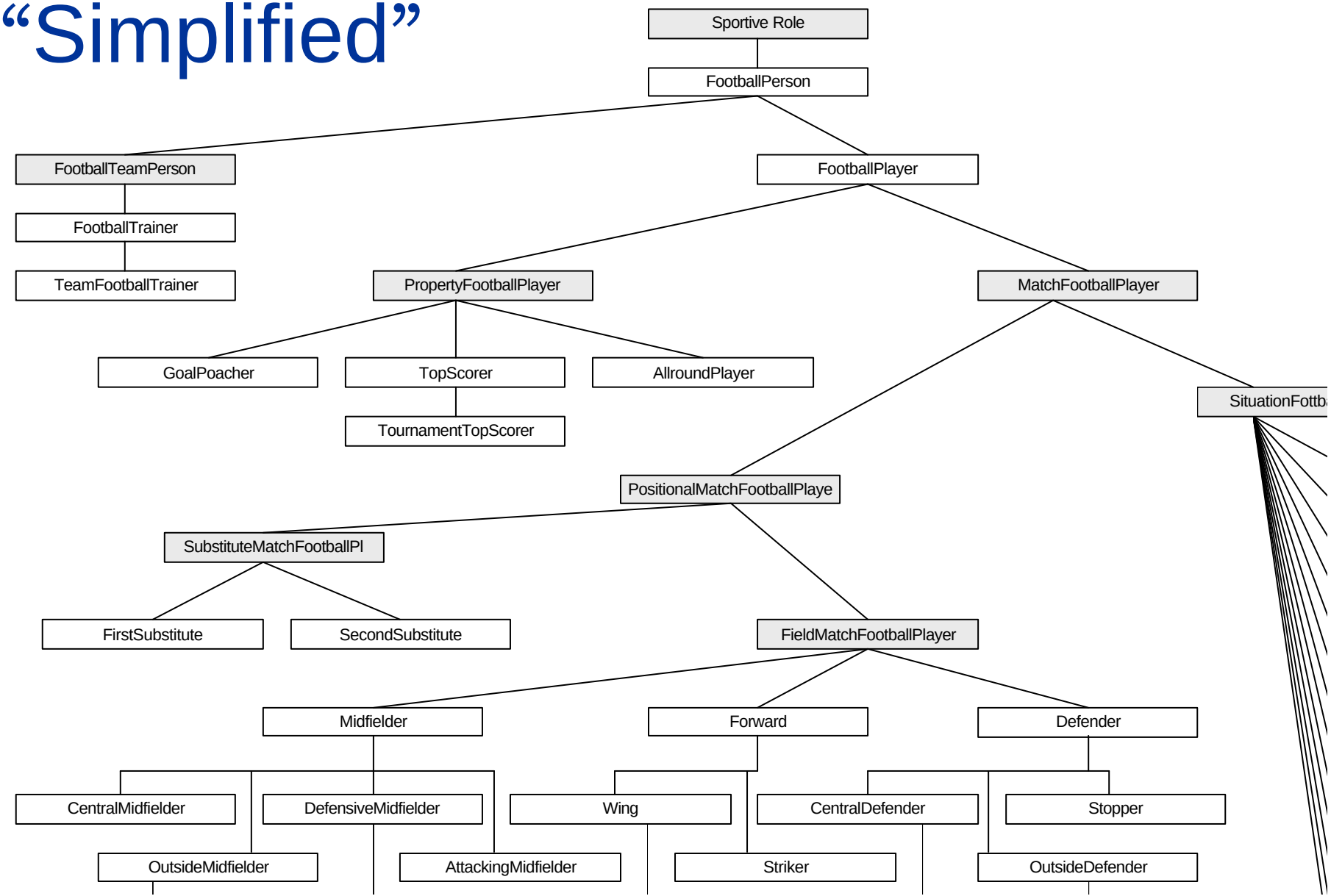

Football Corpus
(3000 FIFA Docs)
Named-Entity
Recognition
Occurrences of
NE Type PERSON
Context Extraction
for NE PERSON
IE Patterns
Class Candidates
Web Resources:
DWDS, Wiktionary,
Wikipedia, Google
IE Patterns
Taxonomy


searchWord="SPIELER"/>
<DWDS>
<kontext>
<entry>jmd., der an einem sportlichen Spiel teilnimmt</entry>
<entry>jmd., der dem Gluecksspiel verfallen ist</entry>
</kontext>
<synonym>
…..
</synonym>
<hyponym>
<entry>Auswahlspieler</entry>
<entry>Gegner</entry>
<entry>
<class resource="TORSCHUETZE" />
</entry>
<entry>
<class resource="TORWART" />
</entry>
<entry>
<class resource="VERTEIDIGER" />
</entry>
</hyponym>
<hyperonym>
<entry>jemand</entry>
</hyperonym>
</DWDS>
<WIKI>
<WIKIPedia>…..</WIKIPedia>
<WIKTionary>
<subClass>
<entry>
<class resource="ABWEHRSPIELER" />
</entry>
<entry>
<class resource="MITTELFELDSPIELER" />
</entry>
<entry>
<class resource="STUERMER" />
</entry>
<entry>
<class resource="TORWART" />
</entry>
</subClass>
<superClass>
<entry>Person</entry>
</superClass>
<kontext>Mitglied einer Mannschaft, jmd.
der an einem Spiel teilnimmt</kontext>
<Herkuft></Herkuft>
</WIKTionary>
</WIKI>
Class Candidate
Lexical Relation
Related Class Candidate
Web Resource
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#">
<owl:Ontology rdf:about=""/>
<owl:Class rdf:ID="AUSWAHL">
<rdfs:subClassOf>
<owl:Class rdf:ID="SPIELER"/>
</rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="SCHLUSSMANN">
<rdfs:subClassOf>
<owl:Class rdf:ID="PERSON"/>
</rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="CHEF">
<rdfs:subClassOf rdf:resource="#PERSON"/>
</owl:Class>
<owl:Class rdf:ID="PRAESIDENT">
<rdfs:subClassOf rdf:resource="#PERSON"/>
</owl:Class>
<owl:Class rdf:ID="ABWEHRSPIELER">
<owl:equivalentClass>
<owl:Class rdf:ID="VERTEIDIGER"/>
</owl:equivalentClass>
<rdfs:subClassOf>
<owl:Class rdf:about="#SPIELER"/>
</rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:ID="STUERMER">
<rdfs:subClassOf>
<owl:Class rdf:about="#SPIELER"/>
</rdfs:subClassOf>
</owl:Class>
Taxonomy
Class
Equivalence

Evaluation - SmartWeb Ontology



Evaluation - SmartWeb Ontology “Simplified”



Evaluation – Extracted Classes and Relations

Classes	**45**

Relations	**32**
DWDS	*17*
Wiktionary	*10*
Google	*3*
Wikipedia	*2*

Evaluation – Precision and Recall

	total	true positives	RECALL	PRECISION
Classes				
SportEvent	50	23	46,0%	31,9%
SportEvent (simple)	***43***	***23***	***53.4%***	***35,3%***
Relations				
SportEvent	226	24	10,6%	10,4%
SportEvent (simple)	***107***	***24***	***22,4%***	***21,6%***

Thanks for your Attention!

paulb@dfki.de

http://dfki.de/~paulb/