

# WP4. Knowledge Engineering



Milen Chechev

Third Project Meeting of MOLTO  
University Of Helsinki, 31.08.2011

# Work done

- November 2010 - D4.1 – Knowledge representation infrastructure.
- May 2011 - D4.2 – Aligned semantic models and instance bases
- September 2011 (coming soon) - D4.3 Grammar Ontology Interoperability
- Prototype – <http://molto.ontotext.com>

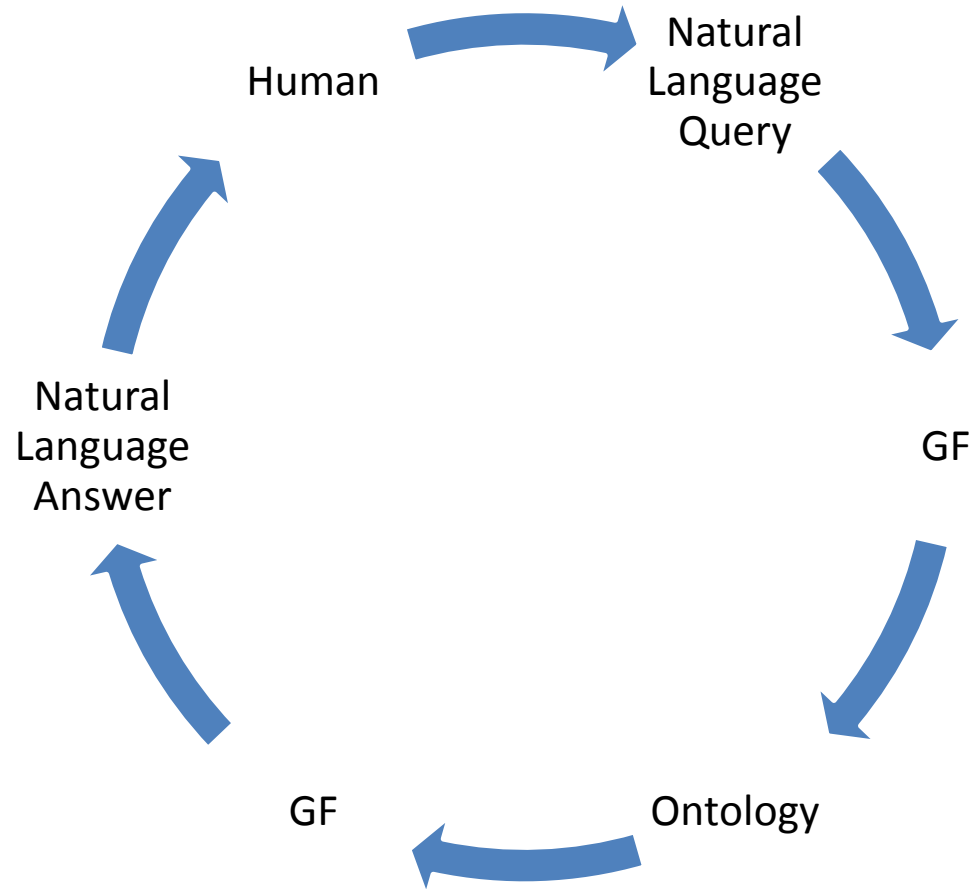
## D4.1 – Knowledge Representation Infrastructure

- OWLIM - a semantic repository that stores all structured data such as ontologies, background knowledge, etc., and provides SPARQL query mechanism and reasoning;
- RDFDB - an API that provides a remote access to the stored structured data via JMS;
- PROTON Ontology - a light-weight upper-level ontology, which defines about 300 classes and 100 properties, covering most of the upper-level concepts, necessary for semantic annotation, indexing and retrieval;
- KRI Web UI a UI that accesses OWLIM through the RDFDB layer. The web UI gives the user the possibility to browse the ontologies and the database, to execute SPARQL queries, etc.

## D4.2 – Data Models and Alignment

- Linked Open Data
- Reason-able view
  - DBpedia - the RDF-ized version of Wikipedia, describing more than 3.5 million things.
  - Geonames - a geographic database that covers 6 million of the most significant geographical features on Earth.
  - PROTON - an upper-level ontology, 542 entity classes and 183 properties.

# D4.3 Grammar-Ontology Interoperability



# Grammar-Ontology Interoperability (details)

- Controlled language
- Auto complete
- Steps:
  - NL to GF Abstract representation (all samples are preprocessed offline)
  - Semi automated transformation from GF Abstract representation to SPARQL (through predefined rules)
  - The data is retrieved from the Owlrim semantic repository
  - Semi automated transformation from the results to GF Abstract representation
  - GF Abstract representation to NL

# Prototype

MOLTO Natural Language Query | SPARQL Query | Contact

An application for viewing datasets of the [MOLTO](#) project.

MOLTO is funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement FP7-ICT-247914.

**Natural Language Query**

Execute

**Examples:**

- [where is "New York"](#)
- [all people that work as "project manager" at "Google"](#)
- [what are the subregions of "United States"](#)
- [give me all organizations with their locations](#)
- [all people](#)

MOLTO Natural Language Query | SPARQL Query | Contact

An application for viewing datasets of the [MOLTO](#) project.

MOLTO is funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement FP7-ICT-247914.

© 2009-2011 Ontotext AD. All rights reserved.

**Natural Language Query**

<input type="text" value="give me"/>	<span style="background-color: #800000; color: white; padding: 2px 10px; border-radius: 3px;">Execute</span>
give me all	
give me information	
give me locations	
give me organizations	
give me people	
give me persons	
give me regions	
give me "United States"	464154 documents
give me "The Associated Press"	240937 documents
give me "Washington"	153691 documents

# Prototype(2)



Natural Language Queries | SPARQL | RelFinder | Contact

An application for viewing datasets of the [MOLTO](#) project.



MOLTO is funded by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement FP7-ICT-247914.

## Search

Results for "give me information about 'United States'" (7) (Sparql: [SELECT DISTINCT...](#))

United States is a country.

location	location_label	type
<a href="http://www.ontotext.com/kim/2006/05/w...">http://www.ontotext.com/kim/2006/05/w...</a>	United States	<a href="http://proton.semanticweb.org/2006/05...">http://proton.semanticweb.org/2006/05...</a>
<a href="http://www.ontotext.com/kim/2006/05/w...">http://www.ontotext.com/kim/2006/05/w...</a>	United States	<a href="#">psys:Entity</a>
<a href="http://www.ontotext.com/kim/2006/05/w...">http://www.ontotext.com/kim/2006/05/w...</a>	United States	<a href="#">ptop:Object</a>
<a href="http://www.ontotext.com/kim/2006/05/w...">http://www.ontotext.com/kim/2006/05/w...</a>	United States	<a href="#">ptop:Location</a>
<a href="http://www.ontotext.com/kim/2006/05/w...">http://www.ontotext.com/kim/2006/05/w...</a>	United States	<a href="#">pext:Country</a>
<a href="http://www.ontotext.com/kim/2006/05/w...">http://www.ontotext.com/kim/2006/05/w...</a>	United States	<a href="#">pext:PoliticalRegion</a>
<a href="http://www.ontotext.com/kim/2006/05/w...">http://www.ontotext.com/kim/2006/05/w...</a>	United States	<a href="#">pext:PopulatedPlace</a>



# Future work

- ontology-grammar interoperability in the specific domain of the museum and patent use cases ( WP7, WP8)
- Combine free text search and rdf search ( needed for WP7)
- Improvements in the user interface

**Thank you for your attention!**

**Questions**

**?**