SEMANTIC DEFINITION AND MATCHING FOR NATIONAL SPATIAL DATA INFRASTRUCTURE

Gülten KARA¹, Deniztan ULUTAŞ², Çetin CÖMERT¹

gispir@ktu.edu.tr, deniztanulutas@ktu.edu.tr, ccomert@ktu.edu.tr

¹Karadeniz Technical University, Engineering Faculty, Geomatics Engineering, Trabzon, TURKEY
²Computer Science Department, Semantic Web Lab. University of Texas at Dallas, USA

http://www.harita.ktu.edu.tr/
Outline

• Introduction
• Ontology
• Why Semantic?
• Semantic Definition Projects and Studies
• Semantic Web Requirements
• Methodology for Semantic Definition
• Semantic Web Languages
• Implementation Architecture
• Transformation Problems-Tools
• Matching Results
• Future Works and Conclusion
Introduction

In Turkey, the establishment of National Spatial Data Infrastructure (NSDI) is on the agenda.

The technologies used for technological infrastructure of any SDI are “Syntactict Web” technologies.

In the near feature, the current technologies will be replaced by “Semantic Web” technologies.
Ontology

An ontology represents a set of concepts within a domain, the relationships between these concepts and the constraints on the properties. In accordance with RDF or OWL, ontologies are made up two main components. Ontologies can assist in communication between people and computers.
Why Semantic?

In the Information Age, the importance given to knowledge is rapidly increasing. But, information sources are vastly varied and gradually increase. It is difficult to find information requested. Therefore, the information on the Web is expressed that is understandable, interpretable and usable by computers to provide for finding the more quickly and easily.
Problem?

- To make semantic definition of the participators of NSDI

- To implement the semantic matching between INSPIRE and GCM schemas
Semantic Definition - Projects

- **FinnONTO Project (2003-2012)** - Semantic Computing Research Group (SeCo)


- ...
Semantic Definition- Academic Studies

- Schade (2009). Ontology-Driven Translation of Geospatial Data
- ...
Semantic Web Requirements

- The semantic definition of data and services.
- The semantic annotation of data and services.
- The semantic matching.
Methodology for Semantic Definition

1. The creation of syntactic definitions
2. The selection of transformation tools
3. The transformation the semantic web languages of organization schema
4. The selecting the suitable ontologies (Upper and domain)
5. The determining references between application and upper level ontologies
Use Case

INSPIRE TN Road Schema-GML Application Schema (*.xsd)
General Command of Mapping (GCM- Road Schema-UML Diagram (*.vsd))
For the semantic matching, the semantic definition is required.
The General Command of Mapping-GCM
(The National Mapping Agency of Turkey)

• Why GCM?
  - Syntactic definition studies
  - One of the most main organization in NSDI Standards of geographic domain in Turkey

- GCM FACC for 1/25000 scale map
- Large Scale Map and Map Information Production Regulation-FACC for 1/5000 scale map

- Currently, there is no common model in Turkey
The Creation of Syntactic Definition

- Syntactic Definition
  - Feature and attribute definitions
  - Database Schema, XML Schema
GCM Road Schema
Semantic Web Languages

- RDF (Resource Description Framework)
- RDFS (RDF Schema)
- DAML+OIL (DARPA Agent Markup Language) + (Ontology Interface Layer)
- **OWL (Web Ontology Language)**
- WSML (Web Service Modeling Language)
Selection the Semantic Web Language

The selection Semantik Web Language

-Current tools
WSMT – WSML, Protégé – OWL

-Expressivity providing of semantic web language
Concepts, relationships between concepts and constraints on them the ability to express of each semantic web language is different. But, semantic web languages use different logic languages.

OWL uses disjoint classes in addition to RDFS
The Selection Transformation Tools

- TopBraidComposer, XSD2OWL, UML2OWL, Protege-UMLBackend, ArgoUML, XMLSpy, Umodel...
The Selection Transformation Tools

INSPIRE TN Schema (GML Application Schema)
XML Schema-OWL, DirectTranslation
TopBraidComposer-commercial

GCM Road Schema (Microsoft Visio)
XML-UML-OWL, Indirect Translation
Microsoft Visio Professional 2010-commercial
Visual Paradigm for UML 9.0-commercial
UML2OWL-open source
Transformation to the Semantic Web Language

Microsoft Visio

Protege
Proposed Ontology Classification
Selecting suitable ontology

- How do I find the suitable ontology?
  Ontology search engines and ontology libraries
  Swoogle, Watson, Sindice, SWSE, Protege Ontology Library, DAML
  Ontology Library, etc.

Currently, a standard tool does not exist for ontology selection

Upper level ontologies: DOLCE, SUMO, Cyc

We select DOLCE ontology because its dimension smaller others and it have extensions.
Semantic Definition

Semantic definition of an organization data is that organization data is coded as understood by computers.
Semantic Annotation

Semantic annotation is formal statement establishing a link between concepts in ontology and features in a data source.
Semantic Matching

- If schemas thought as graph structure, “Semantic Matching” can be perceived as concepts of two graph nodes comparing semantically for determining the similarities between them.

- Graph nodes may be concepts and attributes of concepts.

- Semantic matching compares schema elements (concepts or attributes) semantically according to a common schema, like ontologies.
## Syntactic-Semantic Matching

The existing datasets have different definitions of features.

### Syntactic Matching

<table>
<thead>
<tr>
<th>Item</th>
<th>Length</th>
<th>Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highway</td>
<td>1000 km</td>
<td>20 m</td>
</tr>
<tr>
<td>Pathway</td>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>

### Semantic Matching

<table>
<thead>
<tr>
<th>Type</th>
<th>Length</th>
<th>Breadth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highway</td>
<td>1000 km</td>
<td>20 m</td>
</tr>
<tr>
<td>Pathway</td>
<td>....</td>
<td>....</td>
</tr>
</tbody>
</table>
In our schema matching scenario, we used S-MATCH software.

It is a generic semantic schema matching tool.

It takes two schemas (XML, OWL..) and returns semantic relations between the nodes of the schemas using WordNet lexical database as an external resource (Background Knowledgebase).
Why S-MATCH...

• open-source
• match schemas not only element level schema matching but also structure level schema matching
• match schemas using both semantic and syntactic techniques
• make use of a background knowledge base when schema matching
• Has extendable or changeable background knowledgebase as our domain needs.
• userfriendly
Implementation.

- We implemented the semantic matching between GCM-Road Ontology and INSPIRE-TN Ontology with S-Match.

- S-Match uses WordNet for
  - Concept meaning
  - Relationship between concepts
  - Glosses of concepts
Implementation Architecture

INSPIRE TN Ontology

GCM Road Ontology

S-MATCH

Match Results

WORDNET
Transformation Problems-Tools

• Classes and attributes include Turkish characters and «/» During the XML to UML (manually)

• Software dependant data types (esriFieldTypeInteger…) are not transformed. During the UML to OWL

• UML2OWL is not correctly transform subtype relationship.
Matching Results-Language

GCM Road Ontology

1. GCM Road Ontology(Tur)-INSPIRE TN Ontology: Classes are matched only Thing class (Wordnet does not Turkish language)
GCM Road Ontology(Eng)-INSPIRE TN Ontology: GCM Road ontology is translated into English language.
First hierarchy includes Subtype relationship. GCM Road > INSPIRE Road
Second hierarchy does not include Subtype relationship. INSPIRE Road > GCM Road
Future Works and Conclusion

• We proposed a methodology for semantic definition of an organization data in NSDI

• We are planning to extend WordNet that contains Turkish spatial concepts, attributes and relations.

• We are performed semantic matching in the class level. We plan to make works about mapping of schema attributes.
Future Works and Conclusion

• Tools should be developed for the transformation of software dependant data types (esriFieldTypeInteger...),

for the correction of names of classes and attributes

for transformation from syntactic definition to semantic definition
Thank you for your attention.