Projects Presentation

A Practical Example

Project Structure

- Projects are divided into three Tasks of increasing complexity
  - Knowledge Engineering
  - Data Linking
  - Semantic Web Application

Knowledge Engineering

In practical terms, developing an ontology includes:

- defining classes in the ontology,
- arranging the classes in a taxonomic (subclass–superclass) hierarchy,
- defining slots and describing allowed values for these slots,
- filling in the values for slots for instances.

An Iterative Process (1)

STEP 1 Determine the domain and scope of the ontology:

- What is the domain that the ontology will cover?
    - Wine -> Wine Food Combinations -> Red Wines
- For what we are going to use the ontology?
    - Query Answering, Data Integration, Reasoning
- For what types of questions the information in the ontology should provide answers?
    - Competency Questions

Competency Questions

Frame the scope by formulating a number of questions the ontology is intended to answer.

This basically means choosing the level of detail.

This questions have to be precise yet not exhaustive at this point.

- What are the wines produced in Trentino that are well paired with Fish?

An Iterative Process (2)

STEP 2 Consider reusing existing ontologies:

- People -> FOAF
- Sensors/IoT -> SSN
- Social Interaction -> SiOC
- General -> schema.org, DBPedia Ontology

An Iterative Process (3)

STEP 3 Enumerate important terms in the ontology

- What are the terms we would like to talk about?
    - brainstorming about the central aspect of our ontology
- What properties do those terms have?
    - try to connect the elements by means of verbs
- What would we like to say about those terms?
    - is something missing? Values?

An Iterative Process (4/5)

STEP 4/5 Define the classes/properties and the class hierarchy

Risk here is over-design.

Follow Jim Handler's advice:

"Little Semantics Goes The Long Way!"

An Iterative Process (6)

STEP 4/5 Define data types and cardinality

Cardinality of a property prescribes how many instances could observed as subject of that property

Data Type are used of connecting instances and data, e.g., number.

To be used carefully because they impact performance of the reasoning process.

An Iterative Process (7)

STEP 7 Create Instances

For the knowledge engineering project it is sufficient to manually add some instances and show they respect the competency questions.

Data Linking

- convert/annotate a given dataset into RDF using the the ontology previously developed

- translate the competency questions into SPARQL queries and provide evidence of the results

- publish the dataset as Linked Data

Conversion

- Mappings
- R2RML* allows you to move from CSV/JSON/MySQL to RDF Data
- Software: ONTOP, RML, CARML
    - examples with Ontop and CARML

Querying

Converting the competency questions into SPARQL queries, e.g,

What are the wines produced in Trentino that are well paired with Fish?

```
SELECT ?w
WHERE {
    ?w :suggestedFood :Fish ;
        :producedBy <ABC> .
    <ABC> :locatedIn <Trentino>  . }
```

Publishing

- Data dump
  - data are saved [compressed] and made available remotely
  - A description using VoID/DCat is mandatory
- SPARQL Endpoint
  - data are stored in a Triplestore, e.g., Virtuoso or Jena Fuseki and provisioned as SPARQL endpoint.
  - A description using SPARQL-SD is mandatory

Publishing

- Would that be enough?
- For the project yes. An actual publication pipeline does not stop here
  - Register your dataset to data catalogs like datahub
  - List the dataset/endpoint in the LOD Cloud
  - Maintain the dataset and apply Data Quality techniques

Semantic Web Application

- have fun!