Introduction

The GOSPL (Grounding Ontologies with Social Processes and Natural Language) wiki aims to enable communities to develop and maintain a representation of their world. This is essential for facilitating the uptake of the Linked Open Data (LOD) [1] initiative, which aims to annotate and expose datasets on the World Wide Web. Thus, the community as a whole needs to reach an agreement on the meaning of annotations associated with the (legacy) data. However, the LOD initiative relies on RDF(S) [2] and URI mechanisms to represent these annotations, which is not IT laymen-friendly.

As a result, the GOSPL application is based on the DOGMA framework [3], which is an ontology engineering approach grounded in natural language. Furthermore, GOSPL relies on Social Web technologies to allow every member of a community to express their own knowledge, and to support provenance by tracking changes to the common agreement on this knowledge.

Grounding in Natural Language

RDF(S) provides constructs for naming and accessing resources on the Web by both machines and humans. These constructs use URIs to achieve this, but URIs are not always the easiest thing for humans to understand even though constructs such as rdfs:label and rdfs:comment are available, but are not imposed. IT laymen may be reluctant to learn a new paradigm and feel more comfortable expressing knowledge in natural language.

DOGMA has its groundings in natural language by storing its conceptualizations as lexons [3]. Lexons are formally described as a 5-tuple $\langle \gamma, \text{head term}, \text{role}, \text{co-role}, \text{tail term} \rangle$, where $\gamma$ is an abstract context identifier used to group lexons that are logically related to each other, head term plays role on tail term and tail term plays co-role on head term. GOSPL uses the same lexons as the elementary building blocks for constructing ontologies.

Visualization

Compared to existing wiki-based ontology engineering tools, GOSPL provides a graphical modeling language to represent concepts and the relations between them in Natural Language using NORM-Trees [4], which are undirected trees constructed by concatenating lexons. We make a distinction between the formal model and natural language description of the model for both viewing and editing the model.

Provenance

In GOSPL, we track changes on two levels: one for the model and a second for the textual representation of the annotations from the lexons. The advantage of the tool is that the agreement is reached through other tools such as DOGMA Studio Workbench [5], which are used to verify changes and manipulate the lexons. Each page represents a concept and contains its relations to other concepts. These concepts can, in turn, have their own page and can be easily accessed via a hyperlink. The lexons are easily transformed into RDF(S) by generating classes, properties and their respective hierarchies from the lexons. The advantage of the tool is that the agreement is reached based on natural language and that this shared agreement is then transformed into RDF(S) to annotate datasets, enabling a community to publish their data on the Web.

References


Acknowledgements

The research leading to these results has received funding from the EC’s FP7 program under grant agreement number 216287 (TAS³ - Trusted Architecture for Securely Shared Services).