

Thesis: A label service for the Semantic Web

Background

Entities on the Semantic Web are uniquely referred to by using URIs (Uniform resource identifiers). This URI is not necessarily meaningful to humans. Therefore entities need to have labels in order to be exposable to humans in a meaningful way. These labels can then be used for exploring the data, i.e., for displaying the entities in a linked data browser or other front-end applications, but also to support keyword-based or natural-language based search over the Web of Data. However, the retrieval of labels is a burden for applications:

1. There are multiple properties (about 36) that are used for labeling entities. Applications need to know which properties can be used as labeling properties.
2. Where to find a label for a given URI may not be known by an application. When performing a look-up on a URI the content of the retrieved RDF document may contain labels. But also in documents published anywhere on the web a label can be defined due to the distributed nature of the Semantic Web.
3. Retrieving labels for a set of URIs from multiple data sets requires search in multiple datasets thus being a time-intensive task.
4. For many URIs labels are missing but labels can sometimes be generated automatically or semi-automatically. These new labels need to be published and be accessible by applications.

Goals

The challenge in this thesis is to collect requirements and implement a service for the Semantic Web that will serve as a provider of labels for URIs. Instead of that applications are dealing with the above mentioned problems on their own, they can request this service. Therefore requirements for a data model that regards provenance of the labels (who added the labels, how were labels generated, ...) needs to be elicited. Interfaces that serve the needs of various applications that request labels need to be identified. The system needs to be deployed and integrated into the Semantic Web. This will simplify the development of applications for the Semantic Web which will lead to an increase in uptake of the Semantic Web.

Requirements

Knowledge of semantic technologies as well as Java and PHP skills are required.

For further information, please contact:

Basil Ell
basil.ell@kit.edu
Forschungsgruppe Wissensmanagement
Institut AIFB