Guidance, please! Towards a framework for RDF-based constraint languages.

Thomas Bosch
GESIS – Leibniz Institute
for the Social Sciences, Germany
thomas.bosch@gesis.org

Kai Eckert
Stuttgart Media University, Germany
eckert@hdm-stuttgart.de

Abstract

In the context of the DCMI RDF Application Profile task group and the W3C Data Shapes Working Group solutions for the proper formulation of constraints and validation of RDF data on these constraints are developed. Several approaches and constraint languages exist but there is no clear favorite and none of the languages is able to meet all requirements raised by data practitioners. To support the work, a comprehensive, community-driven database has been created where case studies, use cases, requirements and solutions are collected. Based on this database, we published by today 81 types of constraints that are required by various stakeholders for data applications. We generally use this collection of constraint types to gain a better understanding of the expressiveness of existing solutions and gaps that still need to be filled. Regarding the implementation of constraint languages, we already proposed to use high-level languages to describe the constraints, but map them to SPARQL queries in order to execute the actual validation; we demonstrated this approach for Description Set Profiles. In this paper, we generalize from the experience of implementing Description Set Profiles by introducing an abstraction layer that is able to describe any constraint type in a way that is more or less straight-forwardly transformable to SPARQL queries. It provides a basic terminology and classification system for RDF constraints to foster discussions on RDF validation. We demonstrate that using another layer on top of SPARQL helps to implement validation consistently accross constraint languages and simplifies the actual implementation of new languages.