

Chapter 4

ONTOLOGY REASONING WITH LARGE DATA REPOSITORIES

Stijn Heymans¹, Li Ma², Darko Anicic¹, Zhilei Ma³, Nathalie Steinmetz¹, Yue Pan², Jing Mei², Achille Fokoue⁴, Aditya Kalyanpur⁴, Aaron Kershenbaum⁴, Edith Schonberg⁴, Kavitha Srinivas⁴, Cristina Feier¹, Graham Hench¹, Branimir Wetzstein³, Uwe Keller¹

¹Digital Enterprise Research Institute (DERI), University of Innsbruck, Technikerstrasse 21a, 6020 Innsbruck, {stijn.heyman|darko.anicic|nathalie.steinmetz|cristina.feier|graham.hench|uwe.keller}@deri.at; ²IBM China Research Lab, Building 19 Zhongguancun Software Park, Beijing 100094, China, {malli|panyue|meijing}@cn.ibm.com; ³Institute of Architecture of Application Systems (IAAS), University of Stuttgart, {zhilei.ma|branimir.wetzstein}@iaas.uni-stuttgart.de; ⁴IBM Watson Research Center, P.O.Box 704, Yorktown Heights, NY 10598, USA, {achille|adityakal|aaronk|ediths|ksrinivs}@us.ibm.com

Abstract: Reasoning with large amounts of data together with ontological knowledge is becoming a pertinent issue. In this chapter, we will give an overview of well-known ontology repositories, including native stores and database based stores, and highlight strengths and limitations of each store. We take Minerva as an example to analyze ontology storage in databases in depth, as well as to discuss efficient indexes for scaling up ontology repositories. We then discuss a scalable reasoning method for handling expressive ontologies, as well as summarize other similar approaches. We will subsequently delve into the details of one particular ontology language based on Description Logics called WSML-DL and show that reasoning with this language can be done by a transformation from WSML-DL to OWL DL and support all main DL-specific reasoning tasks. Finally, we illustrate reasoning and its relevance by showing a reasoning example in a practical business context by presenting the Semantic Business Process Repository (SBPR) for systematic management of semantic business process models. As part of this, we analyze the main requirements on a such a repository. We then compare different approaches for storage mechanisms for this purpose and show how a RDBMS in combination with the IRIS inference engine provides a suitable solution that deals well with the expressiveness of the query language and the required reasoning capabilities even for large amounts of instance data.

Keywords: business repository; IRIS; OWL DL; Reasoning with large datasets; Semantic Business Process Management; WSML DL