Integrating Open Data:
(How) Can Description Logics Help me?

Axel Polleres

Vienna University of Economics and Business, Vienna, Austria

At last year's DL workshop Alon Halevy told us about Web tables and how Google makes sense of tabular data on the Web together with Web knowledge graphs [2]. Somewhat surprinsing, a still more unconquered area for Web data extraction seems to be the realm of Open Data: rather than extracting structured data from the surface Web, another emerging source of data on the Web are lots of structured data sets being published openly on various Open Data Portals (e.g. http://www.publicdata.eu/, http://data.gov.gr/ http://data.gov.uk/, http://www.data.gov/, http://data.gv.at/, http://open.wien.at/, just to name a few). However, despite already offering structured data, these Open Data portals often offer only limited search functionality, and intergrating and using Open Data from these portals involves various challenges, such as data quality problems [3], heterogeneity within metadata descriptions, dynamics, or lack of semantic descriptions of the data. Driven by a practical use case, the Open City Data pipeline project [1], we will report on experiences and obstacles for collecting and integrating Open Data across various data sets. We wil discuss how both methods from knowledge representation and reasoning as well as from statistics and data mining can be used to tackle some issues we encountered.

References

1. Bischof, S., Martin, C., Polleres, A., Schneider, P.: Open City Data Pipeline: Collecting, Integrating, and Predicting Open City Data. In: 4th Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data (Know@LOD). Portoroz, Slovenia (May 2015), http://www.polleres.net/publications/bisc-etal-2015KnowLOD.pdf
2. Halevy, A.: Structured data on the web (or, a personal journey away from and back to ontologies). In: Informal Proceedings of the 27th International Workshop on Description Logics. CEUR Workshop Proceedings, vol. 1193 (2014)
3. Umbrich, J., Neumaier, S., Polleres, A.: Towards assessing the quality evolution of open data portals. In: ODQ2015: Open Data Quality: from Theory to Practice Workshop. Munich, Germany (Mar 2015), http://polleres.net/publications/umbr-etal-2015ODQ.pdf