

# Knowledge Representation in the Big Data Age

Pascal Hitzler, Kno.e.sis Center, Wright State University

Position statement for the NSF Workshop on Research Challenges and Opportunities in KR  
Arlington, VA, February 2013

Enormous amounts of data are now at our fingertips. It is produced with breathtaking speed, while at the same time it differs widely in formats, perspectives, usefulness, reliability. One of the foremost research challenges in Computer Science is how to make non-trivial use of this wealth of data, and it seems obvious that the development of impact methods and tools will require cooperative efforts which span many sub-disciplines of Computer Science, while working hand-in-hand with application areas. Facing this challenge, KR needs to redefine itself as an application-oriented discipline. The following are three corresponding challenges which seem to be of central importance.

*Closing the gap between deduction and induction.* Current KR methods work at their best if data and knowledge bases are hand-crafted or hand-curated and adhere to a single modeling perspective. But this approach does not scale up to Big Data requirements, neither in terms of volume nor in terms of noise present in the data, which may come from errors or from different modeling perspectives. At the same time, inductive methods like those based on statistical machine learning or data mining are good at filtering out noise, generally perform better under higher data volume, and are designed to detect higher-level features in data—however they lack deductive reasoning abilities and largely remain black boxes. A successful dealing with Big Data necessitates that we have the best of both worlds available, which means that we have to establish methods which bridge the gap between logic-based knowledge representation and inductive learning and mining methods. This includes the development of applications which cross between these fields, as well as novel theoretical underpinnings for such an integration.

*Developing light-weight KR for education and practice.* KR research has resulted in a plethora of different logics for knowledge modeling, and many of them are well-motivated by formal or informal considerations. However, the landscape of logics is unwieldy and difficult to navigate even for the expert, and even the KR community itself is essentially split into sub-communities which concern themselves only with certain aspects of the KR landscape. We require a consolidation of the field which maps out main approaches, preferably in a modular fashion, in such a way that KR methods become more easily accessible and easy to apply for the non-expert. This will include cross-paradigm studies, but also a boiling down of complicated logics to practical requirements, which are often orthogonal to what the available deep theories can deliver.

*Facing the educational challenge.* The entry barrier to KR is very high. Anticipating higher demands for knowledge engineers in the near future, we need to design courses and curricula which educate in key skills required for applying KR methods in practice. Sound theoretical underpinnings are required, but the focus of such an education needs to be on the transfer of KR methods to application scenarios, starting with a clear understanding of knowledge modeling in different KR paradigms, and exposure to application domains and their peculiarities. A key challenge for developing and delivering such courses and curricula is to provide bridges for a transfer of theoretical aspects of KR to practice-oriented applications.

## References

- [1] P. Hitzler, Towards Reasoning Pragmatics. In: K. Janowicz, M. Raubal, S. Levashkin, GeoSpatial Semantics, Proc. GeoS2009. LNCS 5892, Springer, 2009, 9-25.
- [2] P. Hitzler, F. van Harmelen, A reasonable Semantic Web. Semantic Web 1(1-2), 39-44, 2010.
- [3] K. Janowicz, P. Hitzler, Key Ingredients For Your Next Semantics Elevator Talk. In: S. Castano et al., Advances in Conceptual Modeling, Proc. ER2012 Workshops. LNCS 7518, Springer, 2012, 213-220.
- [4] K. Janowicz, P. Hitzler, The Digital Earth as Knowledge Engine. Semantic Web 3(3), 213-221, 2012.
- [5] A. Krisnadhi, K. Sengupta, P. Hitzler, Local Closed World Semantics: Keep it simple, stupid! In: R. Rosati et al., Proc. DL 2011, CEUR Workshops Vol. 745, 268-278.