

KIELD 2010

1st EKAW Workshop on Knowledge Injection into and Extraction from Linked Data

Foreword

The rapid growth of the Linked Data (LD) cloud, in parallel with on-the-fly design of relevant vocabularies, presents new opportunities for traditional research disciplines such as Knowledge Modelling and Knowledge Discovery from Data. Most notably:

- Although the popular vocabularies reflect today needs, they sometimes lack deeper ontological reflections. State-of-the-art knowledge modelling, especially the pattern-based ontology design principles, could help connect Linked Data vocabularies to more sophisticated models, while keeping themselves simple. Furthermore, collaborative ontology design methodologies could find their way to the process of vocabulary design, currently undertaken by VoCamp communities and stand-alone groups.
- The linked data themselves represent a large and growing resource woven from numerous components. Empirical knowledge discovery in linked data, carried out by machine learning and data mining algorithms, could reveal interesting patterns on frequently used structures, which could then be fuelled back to the vocabulary design. Existing techniques for mining network-structured data, such as graph databases or web links, are likely to require adaptation so as to take account of links that are typed according to semantically rich and heterogeneous schemata.

This complex of research challenges was the main incentive for organizing the 1st Workshop on Knowledge Injection into and Extraction from Linked Data (KIELD 2010), collocated with the 17th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2010), in Lisbon, Portugal.

KIELD 2010 aimed at being an interdisciplinary event, of interest for both researchers and practitioners in all three areas: linked data, knowledge modeling and knowledge discovery. Furthermore, it also assumed space for ongoing and pioneering research activities, which would still be too preliminary as conference publications, although extremely hot as topics in the semantic technology field.

One of the highlights of the half-day event, held in the afternoon of October 15, 2010, following the main EKAW conference, was the keynote talk by Prof. Martin Hepp, “Ontology Engineering for Linked Data: What Makes for a Good Ontology?”, which discussed the impact of ontology design choices and ontology quality criteria on the overall impact of the linked data initiatives.

The workshop received six submissions, which were all carefully reviewed by at least 2 (and mostly 3) reviewers. Five of the submissions satisfied the quality standards for being accepted as full papers. There was also an additional call for short late-breaking news, leading to two submissions (which did not undergo a full review but were checked for relevance by the workshop organizers).

Two of the contributed talks focused (following a similar direction as the keynote) on the need to ‘inject’ more knowledge into the linked data vocabularies. Nuzzoese et al. dealt with expliciting the semantics of (especially, relational) data when putting them to RDF, through meta-modeling such data in OWL; i.e., the linked data are thus ‘injected’ with knowledge already when the resource is being built, i.e. ‘a priori’. In contrast, Vacura&Svátek analyzed some implicit assumptions of vocabularies (specifically for FOAF) and suggested to make them explicit when the given vocabulary is imported as upper-level into a more specific ontology; this corresponds to ‘a posteriori injection’ of knowledge into linked data (at the reuse time of the vocabulary).

Three of the talks focused on the possibility to ‘extract’ useful knowledge from linked data, or to intertwine linked data with other data resources in order to increase the quality of these resources (which, presumably, has the potential of ‘injecting’ the new knowledge back into the original linked data resources). Markotschi&Völker presented a new online game with a purpose, combining the wisdom of crowds with linked data in order to build richer ontological descriptions of concepts. Drăgan et al. showed how semantic desktop data can be published as linked data, via unifying local and web identifiers of entities. Finally, Valle et al. presented a case study in transferring a database of tenders to linked data while exploiting existing LD resources such as DBpedia and Geonames.

The workshop chairs are grateful to all people who contributed to the event, from the PC members, through the presenters (most notably, to the keynote speaker), to all participants. A special thank is due to the local organizers, for their support.

Lisbon, October 15, 2010

Valentina Presutti
François Scharffe
Vojtěch Svátek

Program Committee Members

- Eva Blomqvist, STLab ISTC-CNR, Italy
- Ciro Cattuto, ISI Foundation, Italy
- Claudia d'Amato, University of Bari, Italy
- Mathieu d'Aquin, KMI Open University, UK
- Nicola Fanizzi, University of Bari, Italy
- Aldo Gangemi, STLab ISTC-CNR, Italy
- Alfio Gliozzo, STLab ISTC-CNR, Italy
- Marko Grobelnik, Jozef Stefan Institute, Slovenia
- Tom Heath, Talis, UK
- Luigi Iannone, Manchester University, UK
- David Jensen, University of Massachusetts Amherst, USA
- Agnieszka Lawrynowicz, Poznan University of Technology, Poland
- Pascal Poncelet, LIRMM, Universit Montpellier 2, France
- Marko Rodriguez, AT&T, and Vrije Universiteit Brussels, Belgium
- Steffen Staab, University of Koblenz-Landau, Germany

Additional Reviewers

Alessandro Adamou, STLab ISTC-CNR, Italy

Enrico Daga, STLab ISTC-CNR, Italy

Andrea Giovanni Nuzzolese, STLab ISTC-CNR, Italy

Workshop Homepage

<http://ontologydesignpatterns.org/wiki/0dp:KIELD2010>