CALL FOR PAPERS
Journal of Web Semantics
Special Issue on Dealing with the Messiness of the Web of Data

submission: February 2011
to appear: January 2012

(Guest editors: Stefan Schlobach, Craig A. Knoblock)

Description

Research on the Semantic Web, which is now in its second decade, has had a tremendous success in encouraging people to publish data on the Web in structured, linked, and standardized ways. The success of what has now become the Web of Data can be read from the sheer number of triples available within the Linked-Open Data, Linked Life Data and Open-Government initiatives. However, this growth in data makes many of the established assumptions inappropriate and offers a number of new research challenges.

In stark contrast to early Semantic Web applications that dealt with small, hand-crafted ontologies and data-sets, the new Web of Data comes with a plethora of contradicting world-views and contains incomplete, inconsistent, incorrect, fast-changing and opinionated information. This information not only comes from academic sources and trustworthy institutions, but is often community built, scraped or translated.

In short: the Web of Data is messy, and methods to deal with this messiness are paramount for its future.

For this special issue we seek articles describing foundational and theoretical work as well as technological solutions for dealing with the messiness of the Web of Data. More specifically, we expect submissions on (but not restricted to) the following topics in the context of the Web of Data:

* Knowledge Representation in the presence of messy
  * Context and multi-dimensionality
  * Ontology and data versioning
  * Enforcing and encouraging conventions
  * Representation of uncertain, incomplete and inconsistent data
  * Emergent semantics and self-organizing behaviour

* Querying and reasoning over the messy Web of Data
  * Schemaless querying and integration
  * Dataspaces for the Web of Data
  * Federated querying
  * Reasoning over uncertain, incomplete and inconsistent data
  * Quantitative and statistical methods

* Data integration
  * Identify resolution and record linkage
* Ontology Alignment
* Bridging structured and unstructured data
* Knowledge extraction from noisy data

---------------------------------

Important Dates

We will aim at an efficient publication cycle in order to guarantee prompt availability of the published results. We will review papers on a rolling basis as they are submitted and explicitly encourage submissions well before the submission deadline. Submit papers online at the journal's Elsevier Web site.

Submission deadline: 1 February 2011
Author notification: 15 June 2011
Revisions submitted: 1 August 2011
Final decisions: 15 September 2011
Publication: 1 January 2012

---------------------------------

Instructions for submission:
The submission website for this journal is located at:
http://ees.elsevier.com/jws
To ensure that all manuscripts are correctly identified for inclusion into the special issue you are editing, it is important that authors select "S.I.: Messiness of the Web of Data" when they reach the "Article Type" step in the submission process.

-------------

Guest Editors / Contact (*)
Stefan Schlobach (*) - Vrije Universiteit Amsterdam - schlobac@few.vu.nl
Craig A. Knoblock - University of Southern California - knoblock@isi.edu

-------------

Some example problems that would be interesting for this special issue:
1) Similarity Search: Often users are interested in finding similar resources on the WoD. For example, find cities like Amsterdam or compare universities across Europe. Here, users may not be able to specifically identify the desired overlap. Instead, it is up to the query answering system to identify the overlap and supply reasonable answers.

2) Schemaless Query: One of the positive things about the WoD is the ability for data providers and consumers to use their preferred schema. However, this makes it difficult to query new data sources. Users must discover which schema is used. Furthermore, it makes queries across data sources even more difficult because mappings between vocabularies must be available. We believe that approximation can help alleviate this problem by finding answers "close enough" to the posed query's schema.

3) Robust Query: Misspellings, misuse of vocabulary, violations of schema constraints, all these are part of daily life on the WoD. Today, technologies either skip over such data or must contain
workarounds to deal with it. A systematic approach to dealing with these issues using approximation techniques, would provide a more usable WoD.

4) Aggregated Search Results: answers to more sophisticated queries do not reside all within one triple store. Only by aggregating facts from multiple stores can answers be provided. While federation can virtually provide a single triple store, it has limitations in terms of the consistency required across the underlying triple stores. We believe that approximation can provide a mechanism to enable more robust aggregated search results and federation.

5) Robust extraction: most data that is useful for the Web of Data is not build using Semantic technology but stems from traditional databases. Often this data is translated, or scraped from Web Services or even html pages. Linking this information in well-understood and Semantically correct are crucial for the WoD.