SIEMENS

“Information Rich” = “Messy Data” … Let’s deal with it!

- Messy data is inevitable
  - My use case: Making sense of Structured Data on the Web
  - Assumptions: Loads of data from heterogeneous providers, no way they all agree on a joint schema, no way they provide consistent information
- Be Tolerant
  - Analysis and Repair are key (our analyses show that lots of otherwise useful data is not readable for consumption [HUH+12]
- Reasoning is needed:
  - No agreement upfront on schemas, but “linkage” between used vocabularies can be fruitfully exploited, e.g. a la [HHP09],
  - We also need reasoning & a query language that can take confidence/trust into account [HBPS11], [ZLPS12]
  - But: Keep it Simple, e.g. a la OWL LD [GHKP12]

Conclusion: Information-rich programming needs to allow for consuming and making sense of inconsistent, partial, and wrong information!

SIEMENS

References

[HUH+12] Aidan Hogan, Jürgen Umbrich, Andreas Harth, Richard Cyganiak, Axel Polleres, and Stefan Decker. An empirical survey of linked data conformance. *Journal of Web Semantics (JWS)*, accepted, 2012. To appear.
[GHKP12] Birte Glimm, Adian Hogan, Markus Krötzsch, and Axel Polleres. Owl: Yet to arrive on the web of data? In *WWW2012 Workshop on Linked Data on the Web (LDOW2012)*, Lyon, France, April 2012.
[ZLPS12] Antoine Zimmermann, Nuno Lopes, Axel Polleres, and Umberto Straccia. A general framework for representing, reasoning and querying with annotated semantic web data. *Journal of Web Semantics (JWS)*, 12:72-95, March 2012.
[HBPS11] Aidan Hogan, Piero Bonatti, Axel Polleres, and Luigi Sauro. Robust and scalable linked data reasoning incorporating provenance and trust annotations. *Journal of Web Semantics (JWS)*, 9(2):165-201, 2011.
[HHP09] Aidan Hogan, Andreas Harth, and Axel Polleres. Scalable authoritative owl reasoning for the web. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5(2):49-90, 2009.