# Sebastian Hellmann

Job title: Director of Knowledge Integration and Linked Data Technologies

Organisation: University of Leipzig

Twitter: @dbpedia

Abstract:
**Title: DBpedia - A Global Open Knowledge Network**

In the last 10 years DBpedia has developed one of the most successful knowledge graphs projects with a thriving community. After the foundation of the DBpedia Association in 2014, there has been a three year long discussion about the new strategy and identity of DBpedia to further push Open Data as well as inter-organisational collaboration and curation of data. In the coming months DBpedia will change significantly: (1) from a single source dataset (Wikipedia/Wikidata) to a multi-source dataset with transparent provenance ("a knowledge graph of knowledge graphs") with frequent bi-weekly updates.

The newly created DBpedia Platform was inspired by principles of the Apache Software Foundation and Github and patches systemic flaws of a fully decentralized Web of Data. Data behaves fundamentally different than oil, but is currently more than often transferred via pipelines with tremendous disadvantages. Furthermore, the Web of Data was designed according to the WWW, but incentive models for links in content also behaves very differently than in data and we might even have to rethink the original ideas of Tim Berners-Lee to reach a hierarchical-decentralized data network with full-duplex channels, where "copy" works different.

Joining the DBpedia Platform has clear cut-out benefits for the individual data providers (so called adopters). The mission of DBpedia and its incubator model is to improve its sources over time with automatic test-driven quality assessment (based on the SHACL standard, which was created by DBpedia's CTO), data comparison with all other included sources to spot differences and global curation of links between the sources. Furthermore, feedback given either by professional curators or data users at any node in the network will be relayed to any other relevant node allowing web wide collaboration on data quality without violating data governance principles.


Bio:

Sebastian Hellmann has completed his PhD thesis under the guidance of Jens Lehmann and Sören Auer at the University of Leipzig in 2014 on the transformation of NLP tool output to RDF. Sebastian is a senior member of the "Agile Knowledge Engineering and Semantic Web" AKSW research center, which currently has 50 researchers (PhDs and senior researchers) focusing on semantic technology research – often in combination with other areas such as machine learning, databases, and natural language processing. Sebastian is head of the "Knowledge Integration and Language Technologies (KILT)" Competence Center at InfAI. He also is the executive director and board member of the non-profit DBpedia Association. Sebastian is contributor to various open-source projects and communities such as DBpedia, NLP2RDF, DL-Learner and OWLG and wrote code in Java, PHP, JavaScript, Scala, C & C++, MatLab, Prolog, Smodels, but now does everything in Bash and Zsh since he discovered the Ubuntu Terminal. Sebastian is the author of over 80 peer-reviewed scientific publications (h-index of 21 and over 4300 citations [Google Scholar] (https://scholar.google.com/citations?user=caLrIhoAAAAJ)) and a not-yet-deleted Wikipedia article about Knowledge Extraction. Currently, he is project manager for Leipzig University and InfAI of the EU H2020 Projects ALIGNED and FREME and the BMWi funded project Smart Data Web. Before that, he was also involved in other funded projects such as FREME (EU H2020), LIDER (EU FP7), BIG and LOD2. Sebastian was chair at the Open Knowledge Conference in 2011, the Workshop on Linked Data in Linguistics 2012, the Linked Data Cup 2012, the Multilingual Linked Data for Enterprises (MLODE) 2012 workshop, the NLP & DBpedia Workshop 2014, the SEMANTiCS 2014, 2015 and 2016 conference as well as the KEKI Workshop 2016. In 2012, we held a hackathon at MLODE 2012 bootstrapping an initial version of the Linguistic Linked Open Data cloud image, which led to the LIDER project and linguistic-lod.org that now publishes regular updates (thanks to John McCrae).