A Hitchhiker's Guide to Ontology

Fabian M. Suchanek
Télécom ParisTech University, Paris, France

Abstract

In this talk, I will present our recent work in the area of knowledge bases. It covers 4 areas of research around ontologies and knowledge bases: The first area is the construction of the YAGO knowledge base. YAGO is now mulitlingual, and has grown into a larger project at the Max Planck Institute for Informatics and Télécom ParisTech. The second area is the alignment of knowledge bases. This includes the alignment of classes, instances, and relations across knowledge bases. The third area is rule mining. Our project finds semantic correlations in the form of Horn rules in the knowledge base. I will also talk about watermarking approaches to trace the provenance of ontological data. Finally, I will show applications of the knowledge base for mining news corpora.

ACM Classification
H.1 [**Information Systems**]: Models and Principles

Keywords
Knowledge base, knowledge extraction, text understanding

Short Bio

Fabian M. Suchanek is an associate professor at the Telecom ParisTech University in Paris. He obtained his PhD at the Max-Planck Institute for Informatics under the supervision of Gerhard Weikum. In his thesis, Fabian developed inter alia the YAGO-Ontology, one of the largest public ontologies, which earned him a honorable mention of the SIGMOD dissertation award. Fabian was a postdoc at Microsoft Research in Silicon Valley (reporting to Rakesh Agrawal) and at INRIA Saclay/France (reporting to Serge Abiteboul). He continued as the leader of the Otto Hahn Research Group "Ontologies" at the Max-Planck Institute for Informatics in Germany. Fabian taught classes on the Semantic Web, Information Extraction and Knowledge Representation in France, in Germany, and in Senegal. With his students, he works on information extraction, rule mining, ontology matching, and other topics related to large knowledge bases. He has published around 40 scientific articles, among others at ISWC, VLDB, SIGMOD, WWW, CIKM, ICDE, and SIGIR, and his work has been cited more than 3500 times.