

# The YAGO Knowledge Base

Fabian M. Suchanek

January 4, 2016

A knowledge base (KB) is a computer-processable collection of knowledge. KBs typically store *entities*, such as people, companies, movies, and cities, as well as *facts* about these entities, such as which company produced which movie, or which person was born in which city. KBs find applications in question answering, automated translation, or information retrieval. Most notably, several major players in the Web search business now use KBs, including Google (with the “Knowledge Vault”), Microsoft (with “Satori”), and Amazon (with “Evi”). The Semantic Web unites thousands of publicly available commercial and academic KBs.

In this talk, I will present YAGO<sup>1</sup>, one of the oldest and largest KBs on the Semantic Web, with more than 10 million entities and 120 million facts about them [7, 8, 4, 1, 6]. The project is driven jointly by Télécom ParisTech and the Max Planck Institute for Informatics in Germany. YAGO is constructed automatically from Wikipedia, WordNet, GeoNames, and other resources. I will explain how the knowledge is extracted from multilingual Wikipedias, cleaned, and distilled into a KB of 95% accuracy. All in all, the original YAGO paper has been cited more than 1600 times.

I will also briefly mention the applications of YAGO. We use it for mining news articles of *Le Monde* to extract trends in time [5]. We also use it to mine rules such as  $marriedTo(x, y) \wedge hasChild(x, z) \Rightarrow hasChild(y, z)$  [2, 3].

## References

- [1] Joanna Asia Biega, Erdal Kuzey, and Fabian M. Suchanek. Inside YAGO2s: A Transparent Information Extraction Architecture . In *WWW demo track*, 2013. <https://suchanek.name/work/publications/www2013demo.pdf>.
- [2] Luis Galárraga, Christina Teflioudi, Katja Hose, and Fabian M. Suchanek. AMIE: Association Rule Mining under Incomplete Evidence in Ontological Knowledge Bases . In *WWW*, 2013. <https://suchanek.name/work/publications/www2013.pdf>.

---

<sup>1</sup><http://www.yago-knowledge.org/>

- [3] Luis Galárraga, Christina Teflioudi, Katja Hose, and Fabian M. Suchanek. Fast Rule Mining in Ontological Knowledge Bases with AMIE+ . In *VLDBJ*, 2015. <https://suchanek.name/work/publications/vldb2015.pdf>.
- [4] Johannes Hoffart, Fabian M. Suchanek, Klaus Berberich, and Gerhard Weikum. YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia . In *Artificial Intelligence* , 2013. <https://suchanek.name/work/publications/aaaij.pdf>.
- [5] Thomas Huet, Joanna Asia Biega, and Fabian M. Suchanek. Mining History with Le Monde . In *AKBC workshop*, 2013. <https://suchanek.name/work/publications/akbc2013lemonde.pdf>.
- [6] Farzaneh Mahdisoltani, Joanna Asia Biega, and Fabian M. Suchanek. YAGO3: A Knowledge Base from Multilingual Wikipedias . In *CIDR*, 2015. <https://suchanek.name/work/publications/cidr2015.pdf>.
- [7] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago - A Core of Semantic Knowledge . In *WWW*, 2007. <https://suchanek.name/work/publications/www2007.pdf>.
- [8] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago - A Large Ontology from Wikipedia and WordNet . In *Elsevier Journal of Web Semantics* , 2008. <https://suchanek.name/work/publications/jws2008.pdf>.