

17.09.2013

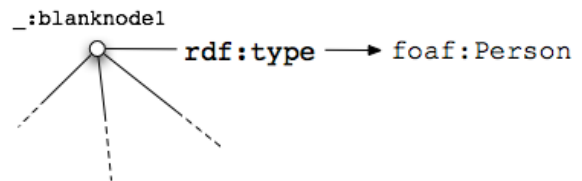
Bachelor-/Master-/Diplomarbeit „Das Klassen-Problem in DBpedia – Versuch einer Konsolidierung“ zu vergeben!

Um was geht es?



DBpedia [1] ist eine der größten Wissensbasen derzeit. Sie wird automatisch aus Wikipedia-Fakten gewonnen. Oft ist jedoch zu gegebenen Entitäten keine Klasse angegeben (fehlendes Prädikat *rdf:type*) oder die Klassenzugehörigkeit kann zwischen ähnlichen Instanzen heterogen ausfallen (Beispiel: *Device* vs. *Smartphone*).

Tools wie Tipalo erlauben es, Entitäten, die in Wikipedia – und damit auch in DBpedia – vorkommen, mit einem Typ zu versehen und damit das „Klassenzugehörigkeitsproblem“ zu lösen. Im Falle von Tipalo wird etwa der Anfangssatz des Wikipedia-Artikels analysiert.



In der Abschlussarbeit soll ein Überblick über bestehende Verfahren/Tools geschaffen werden, die zu gegebenen Instanzen in einer Wissensbasis (wie DBpedia) die Relation *rdf:type* hinzufügen bzw. Inkonsistenzen auflösen, indem sie Klassen vereinheitlichen. Ferner ist anhand einer ausgewählten Domäne wie Smartphones in DBpedia der Einsatz und Nutzen (Effektivität) eines oder mehrerer solcher Tools zu evaluieren. Dies besteht u.a. in einem Vergleich zwischen den Klassen in DBpedia und den Klassen, wie sie durch das Tool gefunden werden.

Was sollten Sie mitbringen?

- Interesse an semantischen Technologien (Vorkenntnisse im Bereich von *Linked Data* von Vorteil, aber nicht notwendig).

[1] <http://dbpedia.org>

Kontaktperson:
Michael Färber
michael.farber@kit.edu
Tel.: 0721/608 479 46