

Thema für Masterarbeit  
**„Balloon: Integration von Apache Giraph“**

Die Veröffentlichung von strukturierten und semantischen Daten mittels RDF in der Linked-Data-Community nimmt stetig zu. Aus dieser Aktivität heraus entstand die Linked Open Data Cloud, welche ein loser Zusammenschluss von verteilten Informationssystemen ist.

Das Projekt „Balloon“ extrahiert und indexiert wichtige Beziehungen als Grundlage für eine schematische Sicht die Daten. Das Ergebnis ist ein lokaler Teilgraph, welcher für weitere Analysen und Operationen genutzt werden kann (z.B. automatische Föderation von SPARQL Anfragen).



**Apache Giraph** ist ein hoch skalierbares verteilbares Graph-System und bildet eine open-source Implementierung zu Google Pregel. Es wird unter anderem von Facebook zur sozialen Graph-Analyse genutzt und basiert auf einem „Bulk Synchronous Parallel (BSP)“-Modell (Map-Reduce für Graphen).

**Arbeit:**

Ziel dieser Arbeit soll sein, Apache Giraph auf eine mögliche Integration in Balloon zu prüfen und gegebenenfalls zu implementieren. Des Weiteren sollen Graph-Analysen auf vorhandene Daten durchgeführt werden und Algorithmen wie z.B. kürzester Weg zwischen zwei Ressourcen angeboten werden.

**Anmerkung:**

Die Masterarbeit wird im Rahmen des CODE-Projektes geschehen. Die Vision von CODE ist ein web-basiertes kommerzielles Ökosystem für semantisches Wissen in der Forschung.

**Kontakt:**

Kai Schlegel (M.Sc.) - Lehrstuhl für Verteilte Informationssysteme  
Innstr. 43, 94032 Passau (Room 248 ITZ)  
Tel.: +49 851 509 3063  
[schlegel@dimis.fim.uni-passau.de](mailto:schlegel@dimis.fim.uni-passau.de)

Weitere Informationen zum Projekt unter <http://www.code-research.de/>

