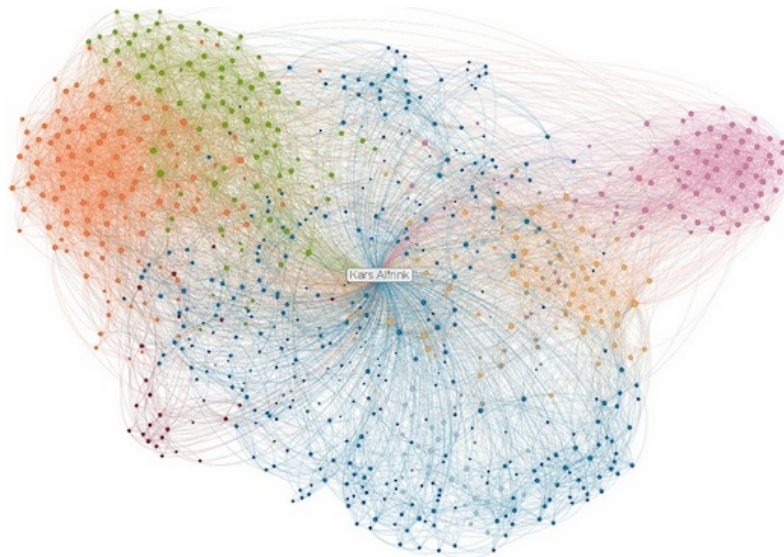


# Ein Framework für Graphdatenbanken auf Basis von Graphtransformationen und Multi-Core Architekturen

## Projekthintergrund

Graphdatenbanken gewinnen, z.B. in Form von sozialen und wissenschaftlichen Netzwerken, Mediendatenbanken und Inferenzsystemen, an Popularität. Im Gegensatz zu relationalen Datenbanken unterliegen die Inhalte von graphbasierten Datenbanken einer Graphstruktur und bieten spezialisierte Graphalgorithmen an, um Suchanfragen zu vereinfachen. Die Suchanfragen selbst können dabei durch Teilgraphen beschrieben werden nach denen in der graphbasierten Datenbank zur Beantwortung der Suchanfrage gesucht wird. Die Inhalte von graphbasierten Datenbanken müssen dabei in Form von Updates aktuell gehalten werden. Suchanfragen und Updates können dabei durch sogenannte Graphtransformationen umgesetzt werden.



**Abbildung 1 Freundschaftsbeziehungen in einem Sozialen Netzwerk<sup>1</sup>**

Suchanfragen und Updates operieren auf beliebig großen graphbasierten Daten und können beliebig nebenläufig auftreten. Da Updates, im Gegensatz zu Suchanfragen, Seiteneffekte auf die Inhalte der graphbasierten Datenbank haben, müssen Konsistenz und Atomarität sichergestellt werden was eine parallele Abarbeitung von Suchanfragen und Updates erschwert.

In diesem Bachelorprojekt wird zusammen mit dem SAP Innovation Center basierend auf der Theorie von Graphtransformationen eine Graphdatenbank konzipiert und prototypisch implementiert. Dabei soll besonderer Wert auf die Parallelisierung im Hinblick auf aktuelle Multi-Core Systeme gelegt werden während andere Aspekte wie z.B. Persistenz für den Prototypen nicht betrachtet werden sollen.

## Projektgegenstand

Ziel dieses Bachelorprojekts ist die Konzeption und Implementierung einer graphbasierten Datenbank. Dafür ist eine detaillierte Erhebung von Anforderungen hinsichtlich eines Framework für Graphdatenbanken auf Basis von Graphtransformationkonzepten und Multi-Core Architekturen erforderlich. Diese Anforderungen sind u.a. a) effiziente Datenstrukturen, die nebenläufigen Zugriff erlauben, b) ein Eventsystem für effizientes Änderungsmanagement, c) die Kompatibilität mit

---

<sup>1</sup> <https://suifaijohnmak.wordpress.com/2012/11/10/>

bestehenden Modell-basierten Softwarewerkzeugen, d) die Austauschbarkeit des verwendeten Graphtransformationssystems (z.B. Henshin<sup>2</sup>, Story Diagram Interpreter<sup>3</sup>) sowie e) die Konfigurierbarkeit des Frameworks zu Evaluierungszwecken.

Im weiteren Verlauf des Projektes sollen diese Anforderungen konzeptionell erfasst und abschließend in Form eines Framework auf Basis des Eclipse Modeling Framework<sup>4</sup> (EMF) implementiert werden.

Das Konzept und dessen Implementierung soll auf Basis großer Beispiele in verschiedenen Konfigurationen evaluiert werden. Dazu ist eine vorherige Datenbeschaffung und Datenaufbereitung notwendig. Außerdem soll das implementierte Framework mit existierenden Lösungen zu verglichen werden.

## Umsetzung

Die ersten Schritte im Bachelorprojekt umfassen die Einarbeitung in Graphtransformationen sowie in EMF und das Vergegenwärtigen der Skalierbarkeitsprobleme von EMF in der Big Data Domäne. Dazu sollen Daten für große Beispiele aufbereitet werden, die später als Evaluierungsgrundlage dienen. Mit den gewonnenen Erkenntnissen sollen detaillierte Anforderungen erhoben werden, die konzeptionell umgesetzt werden. Abschließend wird das Konzept implementiert und mit Hilfe der aufbereiteten Daten evaluiert.

## Projektumfeld

Das Projekt wird in Zusammenarbeit mit dem SAP Innovation Center in Potsdam stattfinden. Derzeit wird am Innovation Center die Skalierbarkeit von Graphdatenbanksystemen basierend auf In-Memory Technologien anhand von standardisierten Benchmarks untersucht. Das entwickelte Framework soll in Bezug auf Ausdruckstärke und Skalierbarkeit mit SAP-eigenen Ansätzen verglichen werden.

## Organisation

In der Seminarphase werden durch die teilnehmenden Studenten die Grundlagen zu ausgewählten Themen erarbeitet und präsentiert (1. Meilenstein). Die Ergebnisse der Seminarphase bilden die Grundlage für die folgende Anforderungserhebung, die in einem Anforderungsdokument zusammengetragen wird (2. Meilenstein). Das Anforderungsdokument wird in einem Antrittsvortrag vor dem Projektpartner vorgestellt. Aus den Anforderungen soll ein Konzept abgeleitet werden, dass in einem Designdokument beschrieben wird (3. Meilenstein). Die Umsetzung der Konzepte wird in Form der Bachelorarbeiten beschrieben und evaluiert (4. Meilenstein). Abschließend werden die Ergebnisse des Bachelorprojektes in Form eines Abschlussvortrags vor dem Projektpartner präsentiert (5. Meilenstein). Bei gutem Gelingen wird eine Anstellung als studentische Hilfskraft in Aussicht gestellt.

## Teilnehmer und Projektbeginn

Bis zu 6 Teilnehmer können in diesem Projekt mitarbeiten. Projektbeginn ist der 1.10.2014.

## Informationen

Für ausführliche Informationen zu dem Projekt stehen Prof. Holger Giese (A-2.5, holger.giese@hpi.uni-potsdam.de), Thomas Beyhl (A-2.11, [thomas.beyhl@hpi.uni-potsdam.de](mailto:thomas.beyhl@hpi.uni-potsdam.de)) und Johannes Dyck (A-2.8, [johannes.dyck@hpi.uni-potsdam.de](mailto:johannes.dyck@hpi.uni-potsdam.de)) zur Verfügung. Ansprechpartner seitens des SAP Innovation Centers ist Dr. Christian Krause (SAP AG, christian.krause01@sap.com).



<sup>2</sup> <http://www.eclipse.org/henshin/>

<sup>3</sup> <http://www.hpi.uni-potsdam.de/giese/public/mdelab/mdelab-projects/story-diagram-tools/>

<sup>4</sup> <http://www.eclipse.org/modeling/emf/>