

TH Köln · Gustav-Heinemann-Ufer 54 · 50968 Köln · Germany

Bearbeitszeit: 12 bis 16 Wochen 1. Prüfer: Prof. Dr. Andreas Behrend

Beginn der Bearbeitung: ab sofort 2. Prüfer: M.Sc. Leonard Traeger

Abgabetermin: nach Vorgespräch und Absprache

Köln, 01.01.2025

Thema der Projektarbeit

Entwicklung einer Web-Anwendung zur Visualisierung von Relationalen Datenbanken Schemata als Graph mit Verlinkungen für die Datenintegration.

Vorbemerkung

Große Datenmengen aus unterschiedlichen Organisationen kollaborativ abzufragen stellt aus Data Science Perspektive eine große Herausforderung dar. Eine integrative Datensicht kann zahlreiche Vorteile in unterschiedlichen Bereichen wie der Marktanalyse, dem Gesundheitswesen sowie der Zusammenarbeit im öffentlichen Sektor bieten. Dabei kann die Datenintegration von Machine Learning (ML) Methoden unterstützt werden um Ähnlichkeiten als Verlinkungen zwischen unterschiedlichen Datenherkünften automatisch abzuleiten. Um dem Volumen von Big Data gerecht zu werden, ist es unerlässlich, den automatisierten Integrationsprozess visuell und interaktiv aufzubereiten.

Aufgabenstellung

Im Rahmen dieser Projektarbeit soll eine moderne Web-Anwendung zur Visualisierung von Verlinkungen zwischen Graphen entwickelt werden. Hierfür sind zunächst relationale Datenbanken als Graph mit den Knoten und Kanten zwischen Schema, Tabelle, und Attribut zu visualisieren. Durch die Konfiguration von Parametern der ML stützenden Verlinkungsmethoden sind dann Kanten zwischen den Knoten der unterschiedlichen Datenbanken anzuzeigen.

Rahmenbedingungen

- Entwicklung einer Web-Anwendung zur Visualisierung von relationalen Schemata als interaktive Graphen mittels D3.js o.ä. Bibliothek (Beispiel: <https://observablehq.com/@d3/force-directed-graph/2>).
- HTML, CSS und JavaScript Kenntnisse und Vorerfahrungen in modernen Frontend Frameworks wie React, Vue.js, Angular o.ä. wünschenswert.
- Die darzustellenden Knoten und Kanten werden samt ML generierten Werten als Flat-File (.csv) bereitgestellt, sodass eine aufwendige Back-End Anbindung zunächst entfällt. Perspektivisch können Vektordatenbanken als Back-End angebunden werden oder ein eingebettetes Aufrufen von ML Python-Skripten implementiert werden.

Aussicht

Die Aufgabenstellung ist zunächst forschungsgetrieben, wird (wenn nötig) eng betreut und schafft erste Einblicke und Teilnahme am wissenschaftlichen Arbeiten im Data Science und Big Data Management Bereich.

Titel Vorname Name

ggf. Funktion/Lehrgebiet/Orga-Einheit

+49 221-8275-Durchwahl

vorname.nachname@th-koeln.de

Room

Straße

PLZ Ort

Germany

Technische Hochschule Köln

University of Applied Sciences

Mailing address:

Gustav-Heinemann-Ufer 54

50968 Köln

Germany

President's Office:

Claudiusstraße 1

50678 Köln

Germany

www.th-koeln.de

Tax ID: 214/5817/3402

VAT ID: DE 122653679

Bank account:

Sparkasse KölnBonn

IBAN DE34 3705 0198 1900 7098 56

BIC COLSDE33