

Bachelor-/ Masterarbeit: Erweiterung einer LLM-basierten Pipeline zur semantischen Modellierung durch ein Graph-RAG-Modul mit Ontologieintegration

Hintergrund und Zielsetzung

Large Language Models (LLMs) zeigen großes Potenzial in der automatisierten semantischen Modellierung strukturierter Daten. Um ihre Leistung weiter zu steigern, können sie mit externem Wissen ergänzt werden. Der sogenannte Graph-RAG (Retrieval-Augmented Generation mit Knowledge Graphs)-Ansatz erlaubt es, domänenspezifische Ontologien oder Wissensgraphen gezielt in die Modellgenerierung einzubeziehen.

Ziel der Arbeit ist es, eine bestehende Pipeline zur LLM-basierten semantischen Modellierung um ein Graph-RAG-Modul zu erweitern. Dafür sollen relevante Ontologien (z. B. VC-SLAM sowie eine Ontologie aus dem Mobilitätsdatensatz) in einen Graphen eingebettet werden, auf den das LLM kontextuell zugreifen kann. Anschließend soll untersucht werden, wie sich die zusätzliche Wissensbasis auf die Generierung semantischer Modelle auswirkt.

Aufgabenstellung und Vorgehen

- Einarbeitung in die bestehende Pipeline und Konzepterstellung für Graph-RAG
- Auswahl und Transformation der Ontologien (z. B. VC-SLAM, Mobilität)
- Aufbau eines RDF- oder Property-Graphen (z. B. mit RDFLib, Neo4j)
- Implementierung eines Retrieval-Moduls für das RAG
- Promptgestaltung zur Integration von Graph-Wissen in die LLM-Anfragen
- Evaluation der semantischen Modelle mit und ohne Graph-RAG
- Analyse und Dokumentation der Ergebnisse

Anforderungen

- Grundkenntnisse in Python
- Interesse an semantischen Technologien (RDF, Ontologien, Knowledge Graphs)
- Erste Erfahrungen mit LLMs und Prompt Engineering von Vorteil
- Interesse an Retrieval-Architekturen und Wissensintegration

Art der Arbeit

Bachelor-/ Masterarbeit

Ansprechpartner

Sebastian Chmielewski | **E-Mail:** chmielewski@uni-wuppertal.de