

# All You Need is Graph-Based RAG

## Grounding Retrieval Augmented Generation with Knowledge Graphs

Graduate



Philipp Josef Vetter

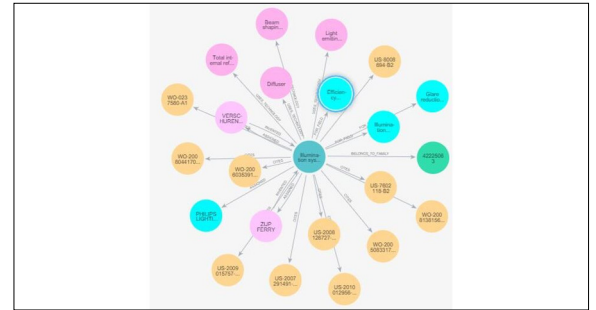
**Introduction:** This master thesis proposes and evaluates a retrieval-augmented generation (RAG) model that uses technology patents as a source of information. Extracting and interpreting complex information from unstructured text poses significant challenges for Large Language Models (LLMs), even when using advanced techniques like Retrieval Augmented Generation (RAG) with vector databases. To address this, the thesis presents a novel hybrid approach to RAG, which combines Knowledge Graph and vector database based techniques.

**Approach / Technology:** The thesis consists of two parts: the first part extracts knowledge from patent claims using a large language model and integrates this knowledge into a graph database; the second part develops a RAG model that combines vector similarity and graph queries to retrieve relevant subgraphs from the knowledge graph and use them to answer questions.

**Result:** The thesis demonstrates that the hybrid method can enhance question-answer systems, improving the accuracy and contextual relevance of information extraction from technology patents. It also demonstrates that LLMs can assist in the construction of knowledge graphs.

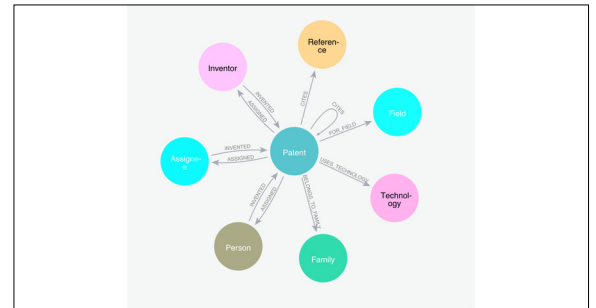
**Example of a subgraph of a patent.**

Source\_: Author



**A knowledge graph ontology.**

Own presentation



Advisor

Dr. Shao Jü Woo

Co-Examiner

Dr. Peter Reiter

Subject Area

Data Science