Infineon Technologies AG is looking for an interested and qualified student to conduct his/her

Master Thesis

on the topic

Exploring New Approaches Using Large Language Models for Knowledge Retrieval in Supply Chain Management

Description:

Infineon's supply chain domain is a rich repository of diverse data, guidelines, onboarding materials, process documentation, and answers to frequently asked questions. Enabling efficient knowledge retrieval within this environment is essential to enhance decision-making, reduce response times, and improve operational efficiency.

Traditional retrieval methods—such as Retrieval-Augmented Generation (RAG) and vector database-based systems—have dominated knowledge management for some time. However, with advancements in artificial intelligence, particularly large language models (LLMs), the potential for enhanced knowledge retrieval has expanded significantly. Models now offer larger context windows, improved reasoning capabilities, and tool-handling features like advanced search, paving the way for innovative solutions that go beyond conventional approaches.

Infineon Technologies AG is one of the largest European semiconductor manufacturers that develops semiconductors and systems for automotive, green industrial power, connected secure systems and power and sensor systems. The products are developed to make life easier, safer and greener with the technology that achieves more, consumes less, and is accessible to everyone.

This thesis focuses on exploring and evaluating state-of-the-art methods for knowledge retrieval that use large language models in the context of supply chain management at Infineon. The goal is to establish a scalable approach for creating a central entry point for supply chain-related queries. By investigating novel methodologies and leveraging the latest advancements, the project aims to overcome existing challenges and fully realize the capabilities of modern AI.

Key objectives include investigating methods for efficient data-fetching, preprocessing, and context definition, while building and tailoring tools specifically designed to enhance AI agent performance. The project also aims to evaluate how standardized protocols like Model Context Protocols (MCP) can ensure scalability and seamless integration, demonstrating the capabilities of AI agents in handling real-world, data-intensive scenarios.

Specifically, the thesis should include the following objectives:

- Development of a framework for leveraging large language models in supply chain knowledge retrieval.
- Design and implementation of a proof-of-concept system that serves as a central entry point for answering supply chain questions.
- Exploration and documentation of alternative retrieval mechanisms, including reasoning-based approaches and tool-assisted queries using LLMs.
- Detailed analysis of the performance, scalability, and limitations of the system compared to classical methods.
- Recommendations for adopting these approaches in real-world supply chain environments at Infineon.

Deliverables

The thesis aims to deliver a research-driven framework for applying advanced LLM methods to knowledge retrieval. It includes a proof-of-concept system showcasing effective retrieval, reasoning,

and tool usage, accompanied by detailed documentation of methods, implementation, and evaluation. The work will also feature a performance analysis comparing the new approach with classical techniques like RAG and vector databases, culminating in final recommendations for implementing LLM-based solutions in supply chain knowledge management.

Requirements:

The thesis is suitable for Management & Technology or FIM students with a major in the area of supply chain management. Candidates must have a strong analytical background, be able to work independently and must show absolute reliability and an interest in the methodological foundation of Agentic AI and its practical use cases in solving real-world problems. Very good MS-Office skills (Word, Excel, PowerPoint) as well as good English communication and presentation skills are mandatory. Basic knowledge about Python programming is needed. Experiences with LLM is a clear plus and will be preferred.

The thesis has to be conducted in **English**. During the thesis, employment with Infineon Technologies AG is provided.

Begin:	as soon as possible
Mentor (Infineon):	Dr. Marco Ratusny
Supervisor (TUM):	tba, preliminarily Heiko Hoppe

Any interested student, please send by email your application together with your curriculum vitae and transcripts of records to **Heiko Hoppe (lehre@osm.wi.tum.de)**