Chair of Spacecraft Systems
TUM School of Engineering and Design
Technical University of Munich

Interdisciplinary Project

A Retrieval Augmented Generation system for the selection of spacecraft components from a database in LLM-based spacecraft design assistants

Start date: 21.10.2024

Duration: 6 months (semester-long)

Topic

At the Chair of Spacecraft Systems, one of the research areas focuses on integrating generative AI into the initial stages of space missions and spacecraft design (referred to as Phases 0 and A in ECSS terminology). The objective is to accelerate and cost-optimize these early design phases, ultimately aiming to democratize spacecraft design.

Early investigation in this research area has led to the creation of an initial spacecraft design assistant, which integrates generative AI -it is based on a Large Language Model- to address requirements and technical specifications generation ranging from high-level mission statements to spacecraft subsystem parameters. [1]

An additional component that brings a lot of value to the design assistant, and which this IDP centers on building, is Retrieval Augmented Generation (RAG). RAG capabilities allow searching for information or data that is relevant for the LLM to answer a user prompt, find it, and then use it to generate a better response. During this IDP, the student is expected to research on the current state-of-the-art of RAG techniques, build a comprehensive local database of spacecraft systems and components, and a system that can select data from it and in connection with an LLM, generate appropriate information. The student will explore the topics of generative AI in general, and RAG, Natural Language Processing, tokenization, word embeddings, vector databases, and LLM chains, in particular.

We offer various courses that can help the student become acquainted with space systems: Space Mission Design (M.Sc. lecture and exercise course), Spacecraft Electronics (M.Sc. practical course), Introduction to Spaceflight (B.Sc. lecture and exercise course), or Systems Engineering Fundamentals (M.Sc. lecture and exercise course).

Tasks

- Conduct a systematic literature review on Retrieval Augmented Generation methodologies, as well as Large Language Models in general. Identify relevant techniques and best practices, as well as the hardware and software architecture required. (1 month)
- Become familiar with the Python environment, the call of pre-trained generic LLM from Python (e.g. through the OpenAI API), the scraping of websites, downloading of information, treatment and processing of text, conversion word embeddings, and integration into a vector, among other relevant techniques, and libraries required for the project. (1 month)
- Develop a RAG prototype using given data on spacecraft components and a pre-trained generic LLM, that can, on the prompts of users, find and select appropriate spacecraft systems and components that fit the requirements or characteristics of a particular spacecraft design. (1 month)

- Develop a set of tests and accuracy evaluation metrics to validate the developed RAG system. (1 month)
- Test the RAG system, debug, and incorporate any required changes to improve its accuracy and performance. Establish a Gitlab repository where the final version of the system's code can be uploaded and generate a small Wiki inside the repository to document the functionalities and usage of the system. (1 month)
- Document project methodologies, findings, and recommendations in a structured manner. Write a comprehensive report (10-20 pages) summarizing the project's objectives, methods, and results. Prepare a presentation and deliver it to the Chair. (1 month)

Requirements

- Proficiency in Python or Rust programming language.
- Command of database systems.
- Basic knowledge of generative AI and Large Language Models.
- Capability to systematically find, evaluate, and synthesize existing literature to inform the project effectively.
- Interest in concepts related to generative AI systems and LLMs.
- Interest in concepts related to space systems engineering.
- Good command of the English language, scientific writing, and presentation of results.

Expectations

- Work autonomously and professionally, administering time properly and autonomously assessing the advancement of the tasks outlined by the plan, according to the timeline.
- Find relevant literature on the area of the work, analyze it critically, and integrate it into the work whenever possible.
- Understand the Engineering trade-offs of the decisions, and document them adequately.
- Communicate regularly with the advisor and supervisor, and promptly when in need of feedback. Be present in the chair weekly to facilitate communication and work progress and attend at least one weekly meeting.
- Pinpoint the scientific or research questions on the area of work, for future publication or work in the frame of a future Master's thesis, for instance.
- Following our internal evaluation, the candidate may be funded, on a merit basis, to present the results of the work at a world-class international conference.

Supervisor

Ramón García Alarcia
Phone: +49 89 289 – 55752
E-Mail: ramon.garcia-alarcia@tum.de

References

[1] Ramon Maria Garcia Alarcia, Alessandro Golkar: Architecture of a generative design tool for spacecraft and user front-end implementation through a chatbot smart design assistant. IAC 2023 Congress Proceedings, 74th International Astronautical Congress, 2023 Baku, Azerbaijan