


Irakaslegoak aurkeztutako GRADU AMAIERAKO LANAREN proposamena		BILBOKO INGENIARITZA ESKOLA ESCUELA DE INGENIERÍA DE BILBAO
Propuesta de TRABAJO FIN DE GRADO por parte del profesorado		


GEHIENEZKO LUZERA: ORRI BAT (2 ALDEAK) / LONGITUD MÁXIMA 1 HOJA (2 CARAS)

DATU OROKORRAK / DATOS GENERALES
Lanaren izenburua / Título del trabajo: LLMs para mejorar la gestión documental en industria
Tutorea (gehienez bi) / Tutor-a (máx. 2): Mikel Egaña Aranguren, Unai Lopez Novoa
e-mail: mikel.egana, unai.lopez @ehu.eus
Hizkuntza / Idioma: Castellano, Euskera, Inglés
Izena emateko aurrebaldintzak / Requisitos para apuntarse: Se recomienda haber cursado la asignatura Administración de Bases de Datos.

DESKRIBAPEN ZEAHTUTA / DESCRIPCION DETALLADA
<p>NOTA: <i>Esta propuesta es una colaboración con el Centro de Fabricación Avanzada Aeronáutica (CFAA)¹, un centro de investigación parte de UPV/EHU, y es posible realizar prácticas remuneradas para el desarrollo del TFG.</i></p> <p>Las empresas industriales actuales están obligadas a gestionar grandes cantidades de documentación en su día a día. Esto viene provocado, en gran medida, por la documentación asociada a las máquinas de fabricación - fresadoras o tornos generalmente vienen acompañadas de manuales de uso y programación que provee el fabricante, certificados de auditorías hechos por terceros o memorias de proyectos en las que son usadas.</p> <p>Aunque la gran mayoría de documentación que se gestiona hoy en día es digital, sigue siendo muy laborioso llevar un control exhaustivo de la misma o incluso consultarla en busca de información concreta. Por este motivo, en este TFG se propone desarrollar una herramienta que permita a las empresas ahorrar tiempo en estas labores.</p> <p>Esta herramienta permitirá centralizar la documentación asociada a ciertas máquinas industriales y que un empleado pueda realizar consultas sobre su contenido de forma conversacional. Para ello, se usará una combinación modelos grandes de lenguaje (o LLMs, como ChatGPT o Copilot) y técnicas de generación aumentada por recuperación (o RAG, por sus siglas en inglés).</p> <p>Estas 2 técnicas se combinan por el siguiente motivo: un LLM, sin utilizar RAG, responderá a cuestiones planteadas usando como base los datos con los que ha sido entrenado. Entrenar un LLM puede ser costoso e ineficiente y, por ello, RAG permite tener una base de conocimiento adicional a la que realizar consultas y con las que enriquecer las respuestas de un LLM, sin necesidad de entrenarlo².</p> <p>Esta herramienta se presentará a sus usuarios con una interfaz sencilla con el objetivo de que perfiles sin grandes conocimientos técnicos puedan hacer uso de la misma. Inicialmente, se plantea que esta interfaz sea web para poder acceder desde diferentes tipos de dispositivo.</p> <p>Este trabajo se realizará en colaboración con CFAA, que cuenta con varias máquinas en su sede del Parque Tecnológico de Zamudio. Sus máquinas están orientadas a investigación en procesos de fabricación y se dispone de documentación para todas ellas.</p>

¹ CFAA: <https://cfaa.eus/>

² ¿Qué es RAG?: <https://aws.amazon.com/es/what-is/retrieval-augmented-generation/>

Irakaslegoak aurkeztutako GRADU AMAIERAKO LANAREN proposamena		BILBOKO INGENIARITZA ESKOLA ESCUELA DE INGENIERÍA DE BILBAO
Propuesta de TRABAJO FIN DE GRADO por parte del profesorado		

El trabajo se realizará en las siguientes fases:

- 1) .. *Estudio de las técnicas RAG*: Aprendizaje sobre el uso de RAG en diferentes dominios y combinación con LLMs. Se seguirá documentación y tutoriales online, p.e.³.
- 2) .. *Definición del caso de uso en CFAA*: Elección de la(s) máquina(s) a utilizar y la documentación asociada con la que construir un primer prototipo. Elección de la técnica RAG a aplicar con la documentación elegida.
- 3) .. *Mejora de las técnicas RAG*: Considerar y aplicar variaciones sobre la técnica RAG para mejorar su precisión. Se pueden incluir otras técnicas si procede.
- 4) .. *Desarrollo de prototipo completo*: Crear aplicación completa con interfaz web que utilice las técnicas desarrolladas en los puntos anteriores.
- 5) .. *Pruebas de usuario*: Definición y realización de pruebas con usuarios finales.

GrAL-AREN HELBURUAK / OBJETIVOS DEL TFG

Para que el TFG sea satisfactorio, es obligatorio:

- ... Elegir y poner en práctica, al menos, 1 técnica RAG que permita realizar consultas sobre la documentación de máquinas industriales de forma conversacional.
- ... Desarrollar un prototipo funcional que presente una interfaz web y que permita interactuar con la documentación mediante técnicas LLM+RAG.

GrAL-AREN EKARPENAK / CONTRIBUCIONES DEL TFG

- ... Propuesta de 1 o varias técnicas para facilitar la consulta de documentación industrial.
- ... Prototipo de herramienta usable con interfaz web.

ERABILIKO DIREN TRESNAK / HERRAMIENTAS A USAR

- ... Librerías y frameworks de código abierto, p.e. LangChain⁴.
- ... Documentación de 1 o varias máquinas de fabricación, provisto por CFAA.

³ Build a Retrieval Augmented Generation (RAG): <https://python.langchain.com/docs/tutorials/rag/>

⁴ LangChain: <https://www.langchain.com/langchain>