

Document Management using Large Language Model

Graduate



Momoko Wymann



Andrew Willi

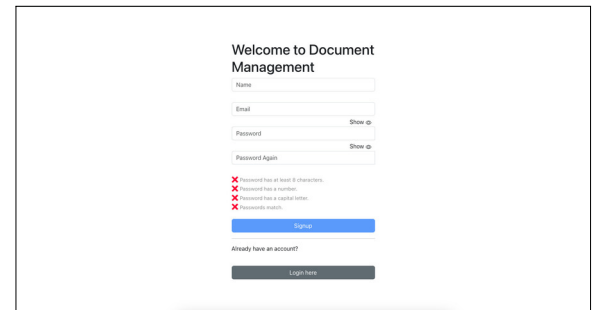
Initial Situation: The management and retrieval of documents can be challenging due to the unstructured nature of their content. Traditional search methods, which rely on document names, are often inefficient and ineffective. With the rise of Transformers based on Large Language Models (LLMs), which significantly enhance our everyday tasks, there is an opportunity to improve the search and management of these documents. The integration of LLMs can transform unstructured data into structured metadata, making documents more accessible and organized.

Objective: Within the scope of this project we aim to build a Single Page Application of this type. We created a prototype application, that is able to store, read and process PDF documents with unstructured data, and generating metadata using a LLM. This metadata improves document search and management, as well as streamline repetitive tasks such as sending an email reminder to a customer. The prototype is designed to be easily expandable to facilitate the continuous development and implementation of new features.

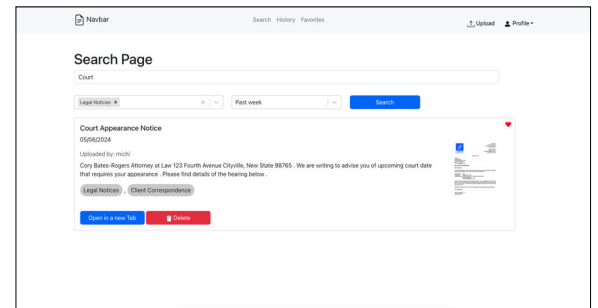
Conclusion: In this project, we developed a prototype application utilizing React and TypeScript for the frontend, while using Node.js with TypeScript for the backend. We integrated external services such as Huggingface and Zapier. The application enables users to upload PDF documents, extract a suitable title, summary and tags using a Large Language Model (LLM), and search for documents based on this metadata. Additionally, it includes a feature for sending email reminders to customers for specific files. Designed with scalability in mind, the application allows for easy addition of new features in the future. Such as integrating various file type like pictures,

which could utilize the LLM. This proof of concept demonstrates the potential of using LLMs to enhance document management and retrieval.

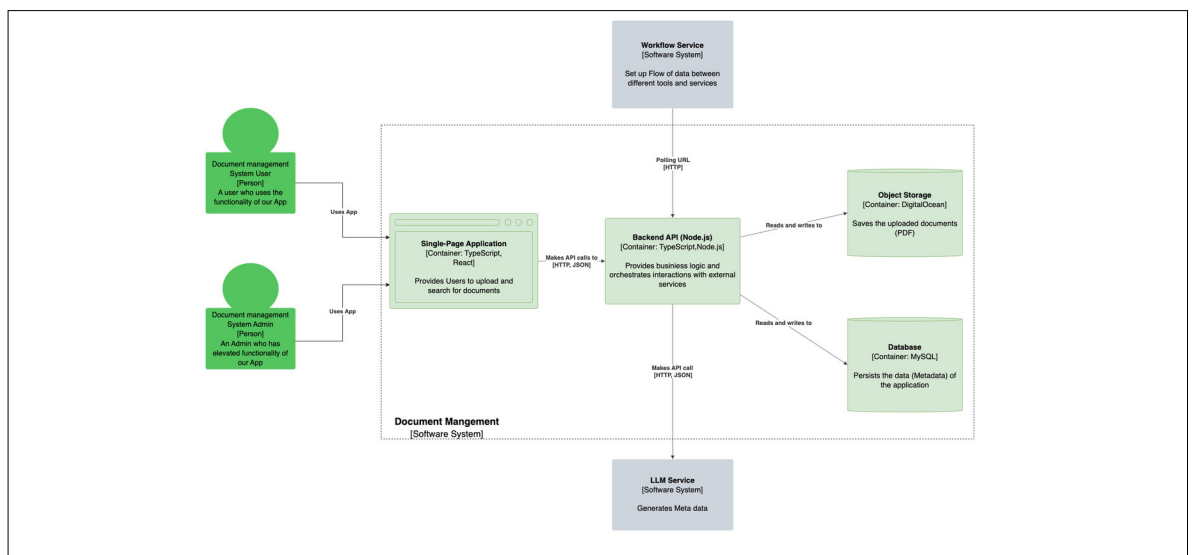
Register Page Own presentation



Search Page Own presentation



Container Diagram Own presentation



Advisor

Prof. Frank Koch

Co-Examiner

Prof. Hansjörg Huser,
Menzingen, ZG

Subject Area

Application Design,
Artificial Intelligence,
Software

Project Partner

AdaptIT GmbH, Jona,
SG

