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 impove 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 presentment

Name Enal Personed Personed Age Remond	Name Enal Passend Passend Passend Again Passend Again Passend Again Passend Again Passend Ref Alexanon Passend Passen Passend Ref Alexanon Passend Passen Passend Passen Passend Passen Passend Passend Passen Passend Passen Passend	Management	Intent
End Sour o Person of Again Boar o Person of Again Boar o Xhousen for a source of the events, Xhousen of the accession, Xhousen of the accession inc. Space	End Size of Si	Name	
Paravol Par	Bite g Passed Passed Agent Sharen Falser I hard E detaction. Sharen Falser I hard E detaction. Bite Agent	Email	
Passed Passed April 1 Passed Passed April	Passed Passed Again Passed Again When the A data data data When the A scalar data data When the A scalar data data When the A scalar data data Passed Again Passed		Show @
Done © Personni dagan X Personni da Calundi Antorporo, X Personni da calundi antorporo, X Personni danali: Sprane	Done © Personni Again X *summer fail and interactions X *summer fails and set of the set	Password	
Passend Agan X Passendra ta Indiana, X Passendra ta Indiana, X Passendra ta Indiana, X Passendra tanak K Passendra tanak Sigana	Passer digin Weiner dia di ante di herinen, Weiner di ante di herinen, Weiner di ante di herinen, Weiner di ante di herinen, Manda herine di ante di di Ante di herine di ante di Ante di ante di ante di Ante di herine di ante di Ante di di di Ante di ante di ante di di Ante di		Show @
X Present has at least totautins. X Present has a runder. X Present has a calleline. X Present has a calleline. X Present has a calleline. Signal	X homodra a load () ansatu Monodra a scatch anno. X homodra a scatch anno. Laguest	Password Again	
Signup	Signup Already have an account?	Password has at least 8 characters. Password has a number. Password has a capital letter. Passwords match.	
	Already have an account?	Signup	

Search Page Own presentment

Court			
Legal Notices ×	× v Past week	✓ Search	
Court Appearance N	tice		•
05/06/2024		5	
Cory Bates-Rogers Attorn	y at Law 123 Fourth Avenue Cityville, New State 98765 . We are s	writing to advise you of upcoming court date	<u> </u>
that requires your appear	sce . Please find details of the hearing below .	10.	bit:
Legal Notices , Client	forrespondence	100	
Open in a new Tab	Delote		



Container Diagram Own presentment

Advisor Prof. Frank Koch

Co-Examiner Prof. Hansjörg Huser, Menzingen, ZG

Subject Area Application Design, Artificial Intelligence, Software

Project Partner

AdaptIT GmbH, Jona, SG

