

Vodafone & Google Hackfest

NewsLetter.AI

Chat with your newsletters using Gen AI.

Cloud Code Crafters

Sep 18, 2024

Agenda

NewsLetter.AI

1 The Pitch

2 The Team

3 About NewsLetter.AI

4 Uniqueness of NewsLetter.AI

5 Explanations for complex terms

6 Demo, Questions & Contacts

The Pitch

NewsLetter.AI

The Problem	The Solution	Value Proposition
<p>Company newsletters can flood employees with too much information every week.</p> <p>It is hard to read everything or find past newsletters.</p> <hr/>	<p>Newsletter.AI is a chatbot that uses retrieval-augmented generation (RAG) to process and search newsletters.</p> <p>Employees ask questions and receive answers from the chatbot about Newsletters.</p> <hr/>	<p>Newsletter.AI lets you easily chat with past and present newsletters using Generative AI.</p> <p>It lets you find important info, even if you do not use exact words.</p> <hr/>
<p>A tool that lets employees "chat" with current and past newsletters would be really helpful.</p>	<p>Tech Stack</p> <ul style="list-style-type: none">• Google Gemini LLM, Cloud Run• Streamlit, LangChain• HuggingFace, Facebook AI	<p>All Documents Supported</p> <p>Newsletter.AI processes any document type, including non-newsletters, with no code changes required.</p>

The Team



Kunal Pathak

AI/ML



Benjamin Bathen

Cloud



Dhirain Kirti Jain

DevOps



Rafael Krysciak

Python



Frank Mieves

Python

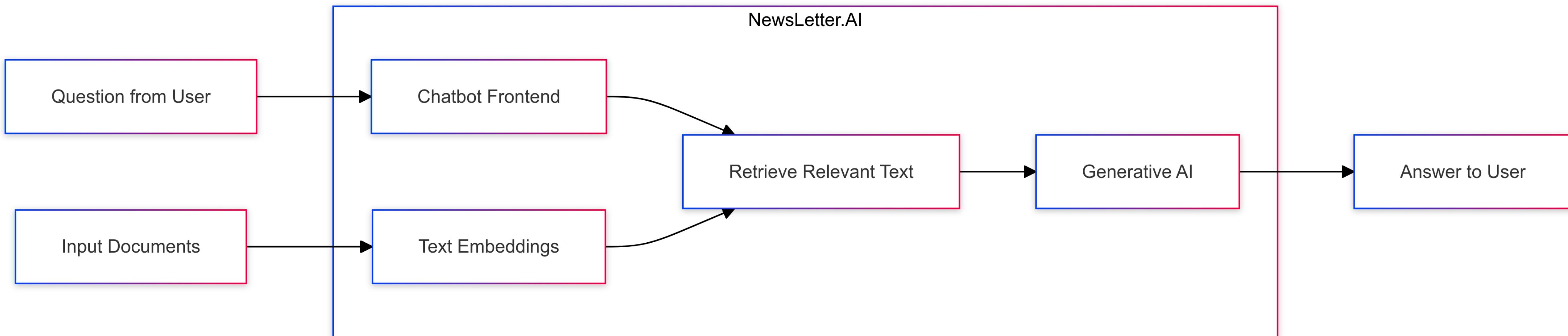


Arushi Arushi

Python

How NewsLetter.AI works

Simple, secure and scalable



[Link to the Chatbot](#)

[Link to Code](#)

Uniqueness of NewsLetter.AI

[Link to detailed list](#)

01 Language Model Flexibility Select models based on speed or quality.	02 Source Transparency Indicates the source document from which the answer was sourced.	03 Smart Indexing Adaptive indexing based on size of input documents.	04 Efficient Caching Multi-tier cache to boost retrieval times for queries.
05 Dynamic Cache Updates Auto-refresh caches when content changes (files added/edited/removed).	06 Automated Deployment CI/CD pipeline triggered for auto-deployment for each code commit.	07 Scalable Environment Scale on-demand based on workload with Google Cloud Run.	08 Model Integration Choose any Model from Model Garden without any restrictions.

Simple Explanations

Breakdown of complex terms used in this presentation.

1

Prompt Engineering

"Prompts" are questions you ask a Gen AI model. To get better answers, you must "engineer" or "refine" them. Details [here](#).

2

Fine-tuning

Fine-tuning makes an AI model better at specific tasks. [Link](#) to further explanation.

3

Retrieval-Augmented Generation

RAG combines retrieving relevant documents and generating responses. More details [here](#).

4

Vector Database

Stores data to easily find similar items. Handy for recommendations like those on Amazon. Image [link](#).

Demo & Questions

Thank You