

Understanding Generative AI Applications - Architecture & Impact

Instructor

Dipanjan Sarkar

AI Advisor at Analytics Vidhya

Google Developer Expert - ML & Cloud Champion Innovator

Published Author



Outline

- Standard components of a Generative AI application
- Standard End-to-End Generative AI App Architecture
- No-code based Generative AI Apps
- Code-based Generative AI Apps
- Impact of End-to-end Generative AI Apps

Standard components of a Generative AI Application



A frontend user interface (UI) to accept user prompts



A database to store embeddings or results (optional)



A backend to handle logic and processing



Custom data or documents (optional)

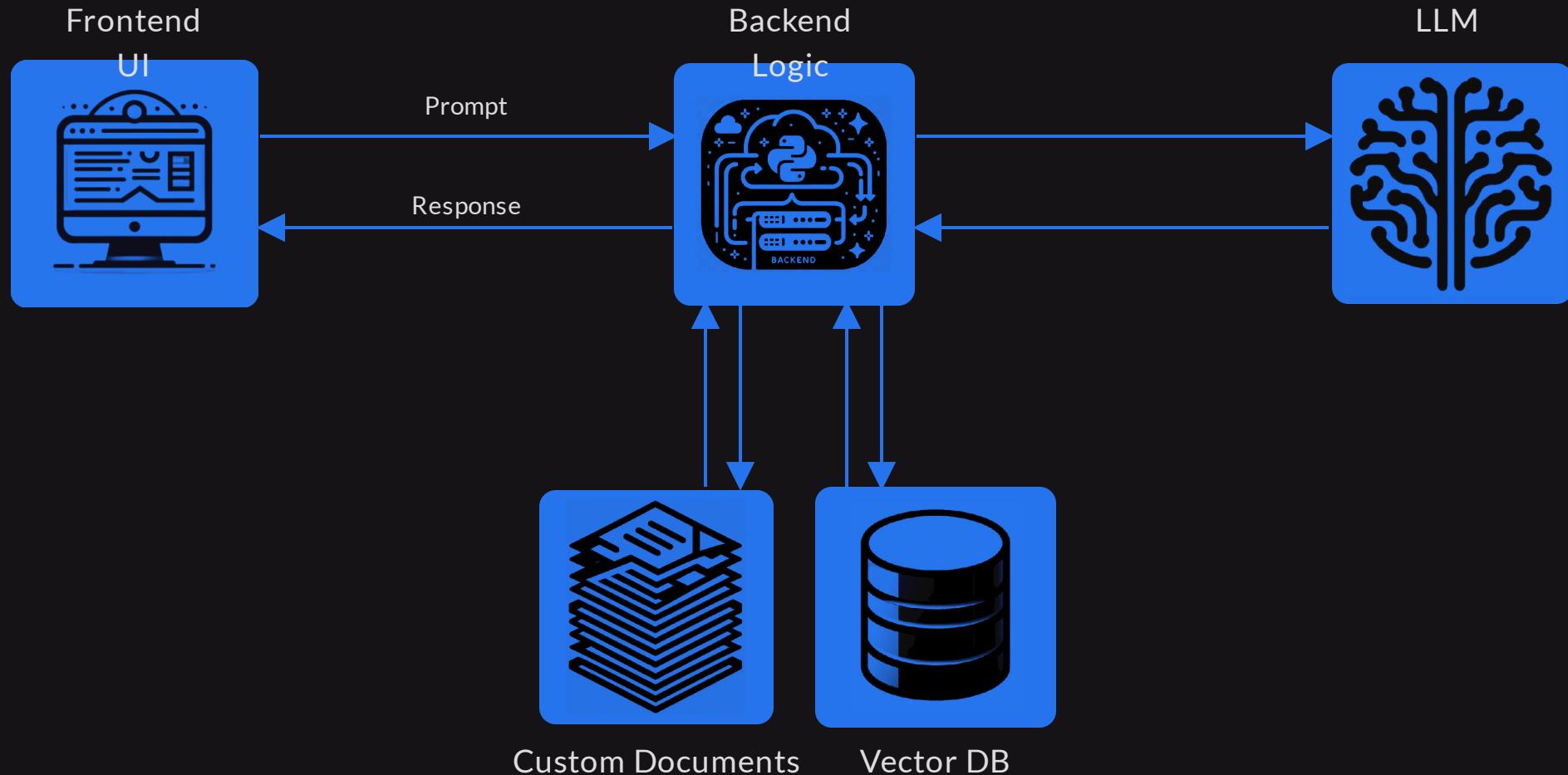


A large language model (LLM) to generate responses



A server to deploy and monitor the app (to be covered in LLMOps)

Standard End-to-End Generative AI App Architecture



No-code based Generative AI Apps



ChatGPT

- Platforms like ChatGPT and HuggingChat allow you to create no-code-focused apps



HuggingChat

No-code based Generative AI Apps

- Frontend customization is minimal
- Backend customization options are limited to fixed model choices and prompts
- Building quick proof of concepts (PoCs) is easy even without learning to code

Code based Generative AI Apps

- LLM APIs from HuggingFace and OpenAI allow for the use of various models



HuggingFace



OpenAI

Code based Generative AI Apps

- Frontend customization is flexible and detailed
- Backend customization can be done using Python and libraries like LangChain and Llamalndex.
- Lot more flexibility, control, and customization
- It is not necessary to write huge codebases to create quick proof of concepts (PoCs).
- Knowledge of frontend languages like JavaScript is not necessary

Impact of end-to-end Generative AI Apps

- Can build working proof of concepts (PoCs) and prototypes effectively with just a single language (Python) or platform (ChatGPT)



python



ChatGPT

Impact of end-to-end Generative AI Apps

- Knowledge of frontend languages like JavaScript is not necessary
- Complete apps with backend and UI can be easily demoed to stakeholders.
- Easy signoffs and approvals to transition from PoC to full application and system
- Simplify application use for end-users to boost adoption

Thank You
