Amir Suleimenov K12247291
PROJECT IN INTELIGENT INFORMATION SYSTEMS
W2024
CHATBOT AI(LLM) ASSISTANT

Project Goal:

Develop a intelligent LLM chatbot assistant prototype

Measure and collect performance based on human conversations about onboarding process

System(Chatbot) Goal:
Assist web portal users(npo managers, volunteers):
Chatbot should consult user, provide information, and help user to use web platform (Onboarding process)

Abstract:
Chatbots have several types and complexity. They started from ELIZA, , developed by Joseph Weizenbaum. ELIZA simulated conversation using pattern-matching techniques and scripts but lacked real understanding, only creating the appearance of comprehension (Weizenbaum, 1966). Based on this idea was grown family of rule based chatbots.

Rule-based chatbot are basic and easiest in implementation chatbots. It is made to interact with users by previously determined guidelines and requirements. These solutions initiate with pre-programmed reactions upon identifying specified phrases or patterns within user input. Developers manually create and design a rule-based chatbot's rules, which specify how the bot will react to different user inputs. And that mean chatbot conversation always have strict restrictions on chat flow algoritm. User can't interact in specific way or get/ receive information in other order. (ChatInsight 2024). In recent years, the development of neural networks and deep learning led to the creation of large language models (LLMs) such as OpenAI's GPT series. These models, trained on vast datasets and fine-tuned for natural language understanding, can engage in coherent, context-aware conversations across a wide range of topics (Brown et al., 2020). Unlike earlier models, LLMs utilize transformer architectures, allowing for superior handling of context and generation of nuanced, human-like responses (Vaswani et al., 2017). This adaptability makes LLM chatbots particularly well-suited for tasks like volunteer onboarding, where they can guide new users by answering questions, collecting relevant information, and recommending opportunities tailored to each user's profile and preferences.

Brief description:
Chatbots have evolved significantly over the years, transitioning from basic rule-based systems to sophisticated AI-powered conversational agents. Their theoretical application for websites has expanded beyond simple user interaction, enabling a new level of engagement and personalization. In the context of modern web applications, a chatbot serves not only as a means of communication but as a core component of the user experience, offering tailored services, guiding users through tasks, and providing dynamic content based on individual needs and preferences.

The theoretical purpose of a chatbot for a website revolves around streamlining interaction, offering assistance, and improving accessibility. Websites often require systems that can handle common user inquiries, support complex workflows, and assist users in navigating

the platform efficiently. A well-designed chatbot can act as a virtual assistant, guiding users through product information, troubleshooting steps, or even assisting with tasks like filling out forms, providing status updates, or offering personalized recommendations.

One of the primary purposes of a chatbot on a website is to provide real-time customer support. Instead of relying on static FAQs or waiting for a human representative to become available, users can get immediate answers to their questions. The chatbot can interact in a conversational manner, making the experience feel more personal and responsive. It can address a wide range of inquiries, from basic questions about products or services to more complex issues requiring the chatbot to access and integrate user-specific data, like purchase history or account status.

JP Morgan workers were encouraged to use LLM Suite for "writing, generating ideas, solving problems using Excel, summarizing documents," among other things, according to an email sent by the bank. At JPMorgan, the chatbot could augment the work being done with "a hybrid of human and AI analysts, very similar to how the intelligence community works," Igor Jablokov, founder and CEO of AI startup Pryon told Fortune. LLM Suite is not JPMorgan's only AI chatbot. The bank also has two other tools known as Connect Coach and SpectrumGPT that are specific to business tasks, rather than a general purpose tool like LLM Suite. (Confino 2024)

In this project, we document the design and development of a chatbot based on these LLM capabilities. Project contains the web site, which is using django framework and store authorization user profiles in Mongo database with rendering chat page. Chatbot is powered by configured Llama LLM model, which running through Ollama instance. Llama 3 was chosen as it's significance lies in its balance of performance, cost-efficiency, and open accessibility. In Meta experimental evaluation discovered that Llama3 model performs on par with leading language models such as GPT-4. (Llama Team 2024). his project not only highlights the potential of LLMs in chatbots to enhance user interactions, but also pushes the boundaries of what is possible with AI-driven communication.

Requirements

Summary System Description

(Prototype for) AI-assisted chatbot for Austrian volunteers web portal.

The system based on modelfile configuration shall help to users by generating answers for users input. As An Ollama Modelfile is a configuration file that defines and manages models on the Ollama platform. Create new models or modify and adjust existing models through model files to cope with some special application scenarios.(GPU-mart)The system shall have authorization before accesing chatbot

Stakeholder Identification
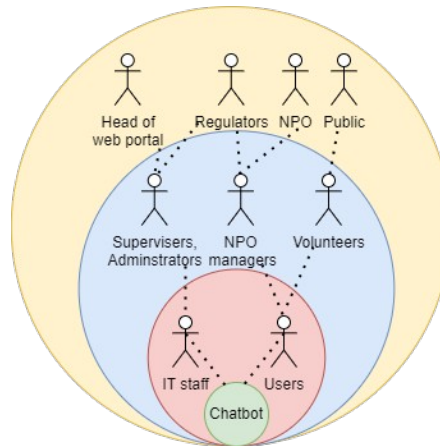
Onion diagram

🟩 System:

The system represents the product or solution that is being delivered by the project (Olson 2013)

- Chatbot

🟥 Business System:

This layer represents the Business System and entails not just the final product, but those stakeholders who interact directly with it, like operators. (ConceptDraw)

- IT staff(web master, server technician etc)

- Users

🟦 Web application:

Those stakeholders who are functional "beneficiares" according Ian Alexander. These are the other stakeholders within the organization who may not interact directly with the solution who benefit from it. (Alexander 2003)

- Volunteers

- NPO managers

- Supervisers, Administrators

🟨 Public Environment:

Those stakeholders who are wider environment in which the organization operates. This layer is populated with stakeholders who are outside the firm but who are still important. As public or various pseudo-governmental organizations has a crucial stake in most projects. (Alexander 2003)

- Head of web-portal environment

- Non Profit Organization

- Public / Citizens

- Regulators

    - AI Act of European AI Office (AI Act 2024

    - The Federal Ministry of Social Affairs, Health, Care and Consumer Protection

    - Municipal/ Local offices and representatives of European AI Office and Federal Ministry

<u>**Using the EARS template, document the functional requirements**</u>

Instructions / Guidelines

- Req<ID> <optional preconditions> <optional trigger> the <system name> shall <system response>
- Document as many requirements as you desire
  - Not all of them need to be implemented in this course

List of functional requirements

<u>Authentification</u>

FR100: The system shall allow user create account (registrate) for the purpose of providing conversation for every user

FR101: The system shall allow user to his login (authentificate) by entering his password

FR102: The system shall allow user change password based on his account

FR103: The system shall use Mongo db as DBMS for easier integration with other parts of web portal

<u>Chatbot</u>

FR200: The LLM-powered chatbot shall respond to user inputs in natural language.

FR201: The chatbot should generate an answer for every message from user input.

FR202: The system shall be able to understand the user's intention from their input.

FR203: The bot shall extract specific information (entities) from the user's message, such as dates, names, or task information.

FR204: The bot shall have context management and conversation flow

> **SFR2040:** The system shall maintain the context of a conversation across multiple turns, remembering prior user inputs and responses.

> **SFR2041:** The system should handle ongoing conversations, where it can maintain state and carry information across multiple exchanges.

> **SFR2042:** The chatbot should continuously update its internal state to reflect new information from the conversation and respond in a relevant manner based on the current state.

> **SFR2043:** The system shall decide the next action (asking for more information, providing an answer, executing a task) based on the current state.

FR205: The system shall handle misunderstandings and have adaptability

> **SFR2050:** The system shall be able to handle misunderstandings, detect when it cannot understand, and ask clarifying questions.

> **SFR2050(2nd ver):** The chatbot should recognize when it cannot fully understand a user's input and ask clarifying questions to resolve ambiguities.

SFR2051: The system shall adapt to individual users by learning preferences over time.

List of nonfunctional requirements

Chatbot

NFR 200: The LLM-powered chatbot's architecture shall allow for updates and bug fixes to be deployed without downtime.

NFR 201: The system should handle a large number of concurrent users and scale efficiently without performance degradation.

Actors and Agents

Primary roles (who use the system in their daily activities) in product system

#	Primary Actor	Type	Description
1	User of chatbot	Human	Responsible for 1. autentification process before starting chat 2. entering messages and questions for defining exact topic of user interest
2	Chatbot Processing Model	AI	Provides core logic for answering on messages in natural language processing 1. Configured as chatbot assistant for Austrian volunteers portal 2. Identifying individual chat and conversation history
3	IT staff	Human	Responsible for 1. Maintenance technical equipment such as server equipment 2. Problems and bug solving 3. Help users troubles and questions with interaction to chatbot instance

Prototyping Scenario

Have conversation as chatbot with predefined role and knowledge specialization

It contains Use case according template, Architecture Diagram and Screenshot examples according requirements template

Use Case

ID	UC100-Auth-Login
Description	User logs into app
Actors	Volunteer/NPO manager as User of chatbot
Stakeholders:	NPO managers, Volunteers
Pre-Conditions	The system has access to the intranet via the device's WLAN access point
Success end condition:	The user is logged into the app and into the chatbot in the background
Failure end condition:	The user is not logged in and is being asked to contact support to verify credentials & permissions

Main Success Scenario

1	The user enters his user info (user id, password)
2	The user submits the user info
3	The systems checks whether the user is permissioned to use the chatbot app
4	The system creates a session for the user for submitting API requests LLM engine
5	The app launches and displays the main screen/menu

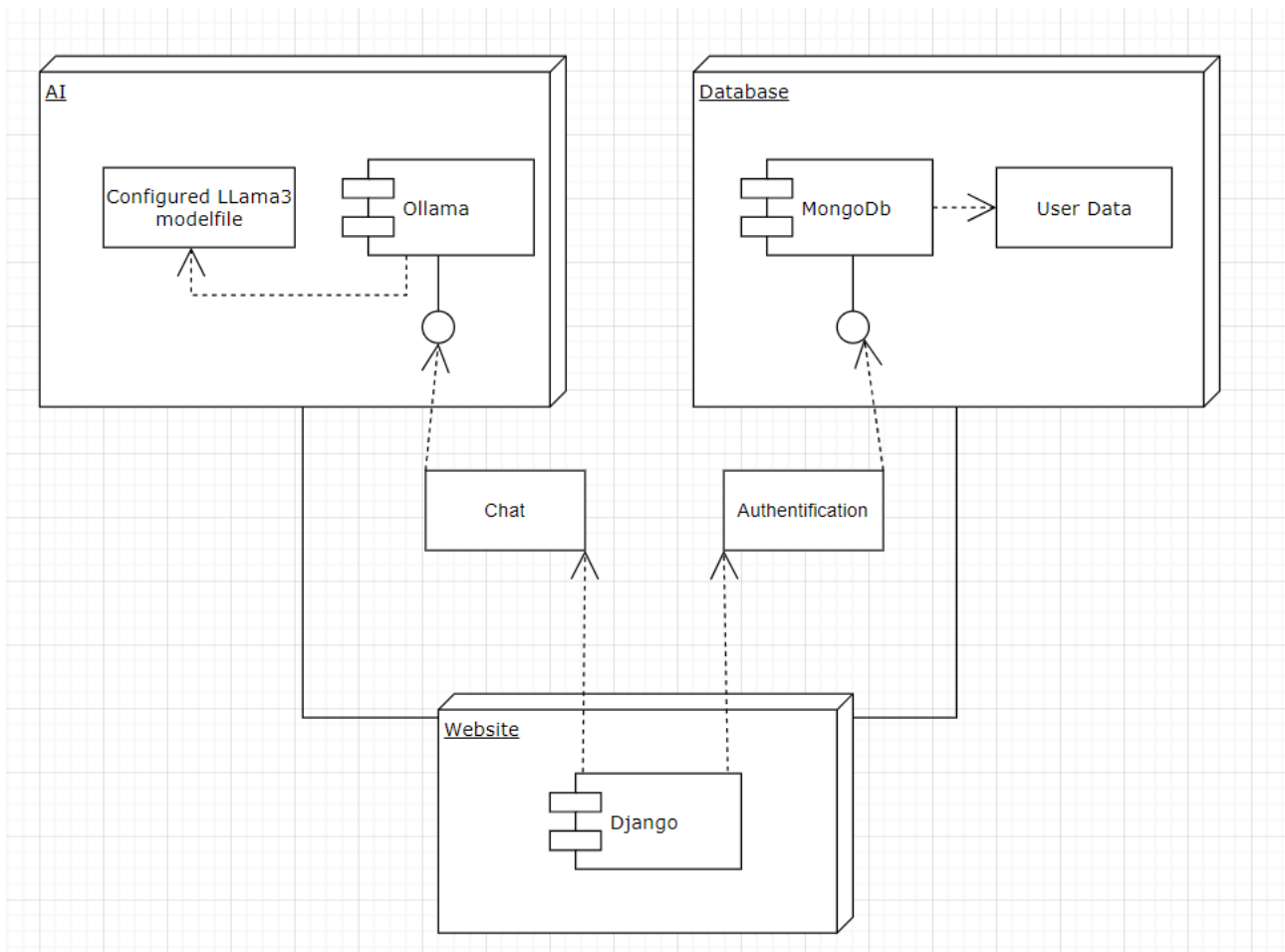## Alternative Scenarios

Exception Scenario

3.A1.1	Checking the user credentials results in an authentication error (unknown/invalid user)
3.A1.2	The system asks the user the verify credentials and to try to login again
3.A2.1	Checking the user credentials results in an authorisation error (user is known but not permissioned to use the app)
3.A2.2	The system informs that permission to use the app is missing and to contact support Before trying to log in again
4.A3.1	The system is unable to create a session with the chatbot system API (either authentication, authorisation or general technical error)
4.A3.2	The system asks the user to verify network connection and permissions Inside the web application system before trying to

log in again

Base architecture Diagram

AI

Configured LLama3 modelfile
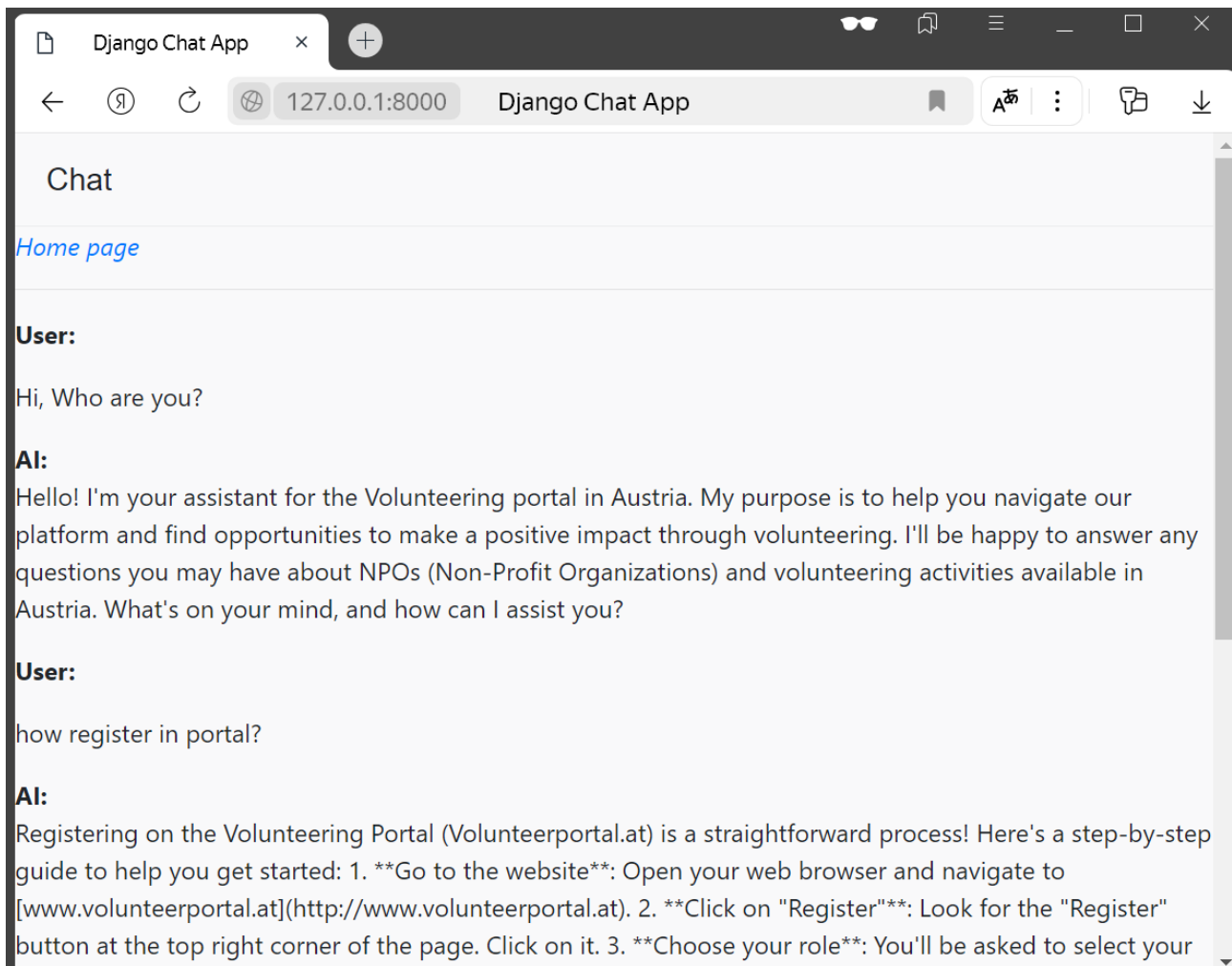
Ollama

Database

MongoDb

User Data

Chat

Authentification

Website

Django

main page image

login page image

chat image

So now chatbot can answer questions and even try to generate close to reality answer when tries to answer as example register procedure.

Future steps.

For next steps of developing it could be adding difference on behavior between volunteer and npo manager like concentrating on different topics and themes.

To create task creation procedure I faced with some restrictions of LLMs as variability of answering of same question. So to create working pocedure there is need of rule based approach as in usual chatbot with button and questions with strict restrictions of rules like format of enter and etc to future conversion to JSON or XML format for sharing it to other services.

References

1. AI Office of the European Commission, "The AI Office: What is it, and how does it work?" https://artificialintelligenceact.eu/the-ai-office-summary/, – published 21 March 2024

2. Weizenbaum, J. (1966). "ELIZA - A Computer Program for the Study of Natural Language Communication Between Man and Machine." *Communications of the ACM*, 9(1), 36-45.

3. Tiffany, ChatInsight "A Comprehensive Guide to Build a Rule-Based Chatbot" https://www.chatinsight.ai/chatbots/rule-based-chatbot/ , - published 20 February 2024 –

4. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amodei, D. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.

5. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. arXiv preprint arXiv:1706.03762

6. Llama Team, AI @ Meta. The Llama 3 Herd of Models. https://scontent-vie1-1.xx.fbcdn.net/v/t39.2365-6/463020162_522238820565582_8192401983671993921_n.pdf?_nc_cat=108&ccb=1-7&_nc_sid=3c67a6&_nc_ohc=ZGqgxjTvDMYQ7kNvgEk4Tt9&_nc_zt=14&_nc_ht=scontent-vie1-1.xx&_nc_gid=AJtJhR8dBRBkApUnavDR2xw&oh=00_AYBLgff_1hfIRxucPP9FKolsU3_AzlpL_ffuA-bAmcWeYw&oe=67270659. published July 23, 2024.

7. GPU-mart, How to Customize LLM Models with Ollama's Modelfile https://www.gpu-mart.com/blog/custom-llm-models-with-ollama-modelfile

8. Paolo Confino, Fortune https://fortune.com/2024/07/26/jpmorgan-chatbot-llm-suite-chatgpt-ai-machine-learning-wall-street-banking-finance/ published 26 July 2024

9. David Olson, Bawiki "Stakeholder Onion Diagram What is it?" http://www.bawiki.com/wiki/Stakeholder-Onion-Diagram.html published 2013

10. ConceptDraw "Stakeholder Onion Diagrams" https://www.conceptdraw.com/examples/management-stakeholder-onion-diagrams

11. Ian Alexander "Stakeholders – Who is Your System For?" https://www.scenarioplus.org.uk/papers/stakeholders/stakeholders.htm in Computing & Control Engineering, Vol 14, Issue 1, pp 22-26, April 2003