IGNIS – Vulnerability Management Assistant

Khairul Amirin Syahrean| Email: C00265680@setu.ie | Supervisor: Richard Butler

BSc (Hons) Cybercrime and IT Security

Department of Computing & Networking, South East Technological University, Carlow

Ollscoil Teicneolaíochta an Oirdheiscirt
South East Technological University

INTRODUCTION

With the advent of **ChatGPT**, AI has been at the forefront of technological development, streamlining and minimizing overhead for various businesses and fields.

Introducing Ignis, an AI-powered personal assistant that assists cybersecurity members in vulnerability management of a company's IT security. It is a user-friendly personal chatbot allowing users to obtain insightful information on threats or vulnerabilities of the company's network environment at a moment's notice. Chatbot workflow would be tailored individually depending on the user's needs and records it in the database for future sessions.

The purpose of this project is to showcase how AI can be integrated into vulnerability management of IT systems for businesses and organizations. Organizations in 2023 are turning towards AI tools to keep an eye on and fight against cyber-attacks and cybercrimes more effectively. Lately, researchers and cybersecurity experts have ramped up efforts exploring all the latest AI methods for cybersecurity[1]. Ignis aims to be one of the pioneers of AI applications that can be used for production within an organization.

RESEARCH QUESTIONS

How can AI and Large Language Models be implemented to maximize vulnerability management and IT Security?

How efficient can GPT parse and categorize security information and vulnerabilities of a network environment?

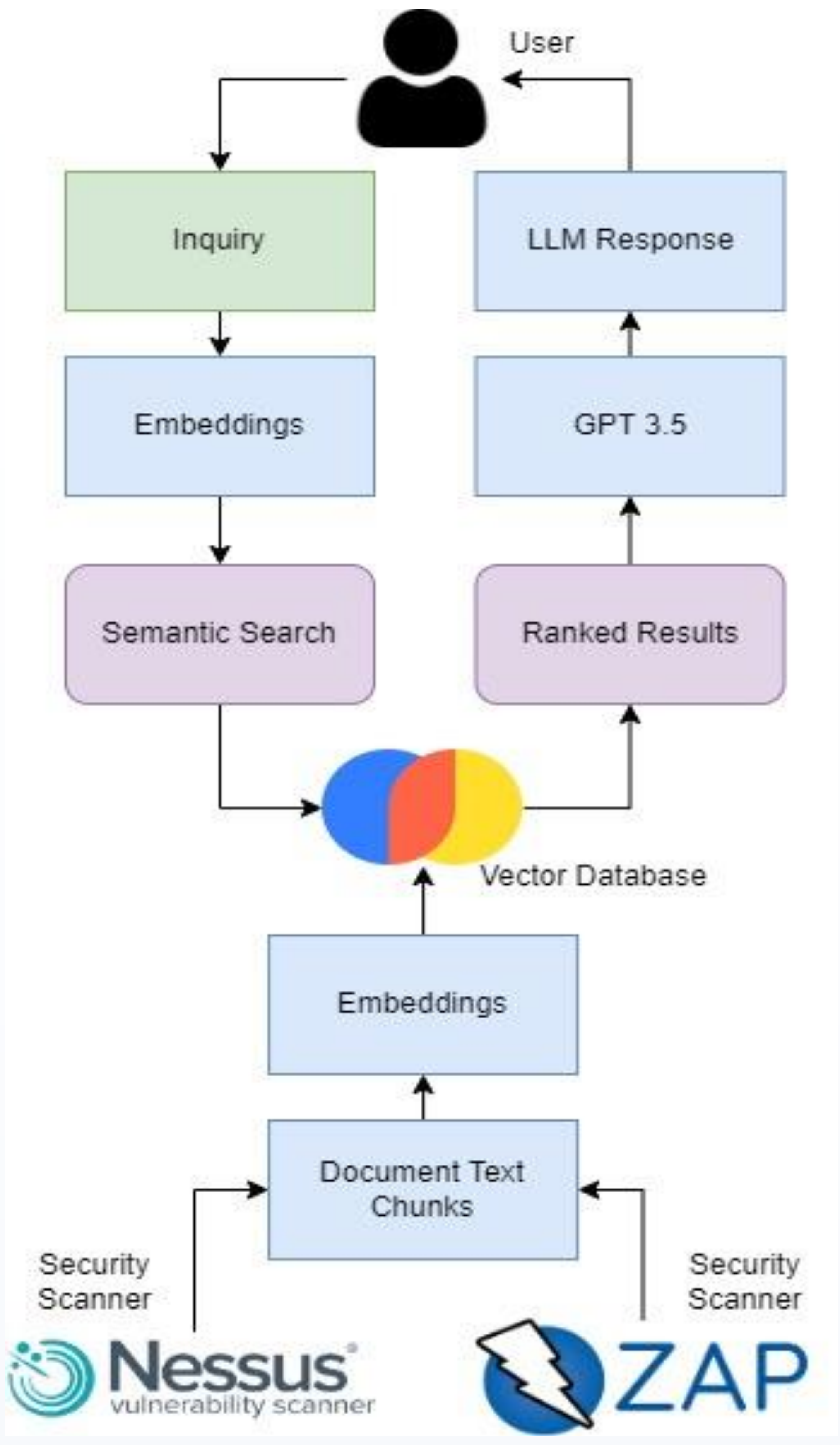
TECHNOLOGIES

OpenAI, PyCharm, python, Nessus vulnerability scanner, LangChain, Streamlit, ZAP

METHODOLOGY

- ❑ For the project we are analyzing a wide range of scan reports to determine the LLMs ability to process and output useful data.
- ❑ Most data is obtained by simulating a network environment with Virtual Machines and configured so that security scanners would detect vulnerabilities and output reports in PDF or CSV format.
- ❑ For every test run conducted, the output is recorded and examined. We would need at least 10 answers so that the fine-tuning would make a significant difference to the LLM's output.
- ❑ Using LangChain, our scan reports would be segregated into chunks of text. Text chunks are then converted into vector embeddings, each with a special numerical value only discernable by machines and stored into the database.
- ❑ The main knowledge base is stored locally in ChromaDB, a vector database. A new way to store data, vector databases excel at similarity searches using algorithms unique to each database[2]. Other vector databases such as Pinecone, FAISS, Qdrant, etc. have their own proprietary algorithms to index embeddings and conduct semantic searches.

ARCHITECTURE DIAGRAM



1. User would ask the application through a chatbot interface.
2. Application conducts semantic search between query and entries with vector database.
3. Any selected embeddings are sent to GPT 3.5, processed and output as response.

CONCLUSION

By the end of development, Ignis should be able to assist users by providing a holistic overview of vulnerability management.

The application can process typical scan reports and store information in the local database. It can process user queries, search the database and respond within 6-8 seconds. There is a consideration to upload mitigation knowledge base from credible sources (MITRE ATTACK) to further supplement remediation process.

Response is fairly accurate and almost human-like, however extended use of the application causes the LLM hallucinates information that is false or irrelevant. LLM parameters must be scrutinized, and more fine-tuning is needed.

NEXT STEPS

- ❑ Develop incident triaging mechanism, notifying response teams for remediation
- ❑ Explore other LLMs that provide better reasoning, consistency or processing speed.
- ❑ Research best practices when considering putting vector database in the cloud
- ❑ Improve on data quality for fine tuning the LLM and avoid bias.

REFERENCES

1. Wiafe, I., Koranteng, F.N., Obeng, E.N., Assyne, N., Wiafe, A. and Gulliver, S.R. 2020. Artificial intelligence for cybersecurity: a systematic mapping of literature. IEEE Access, 8, pp.146598-146612.
2. Cloudflare (2023). What is a vector database? [online] Cloudflare. Available at: https://www.cloudflare.com/en-gb/learning/ai/what-is-vector-database/