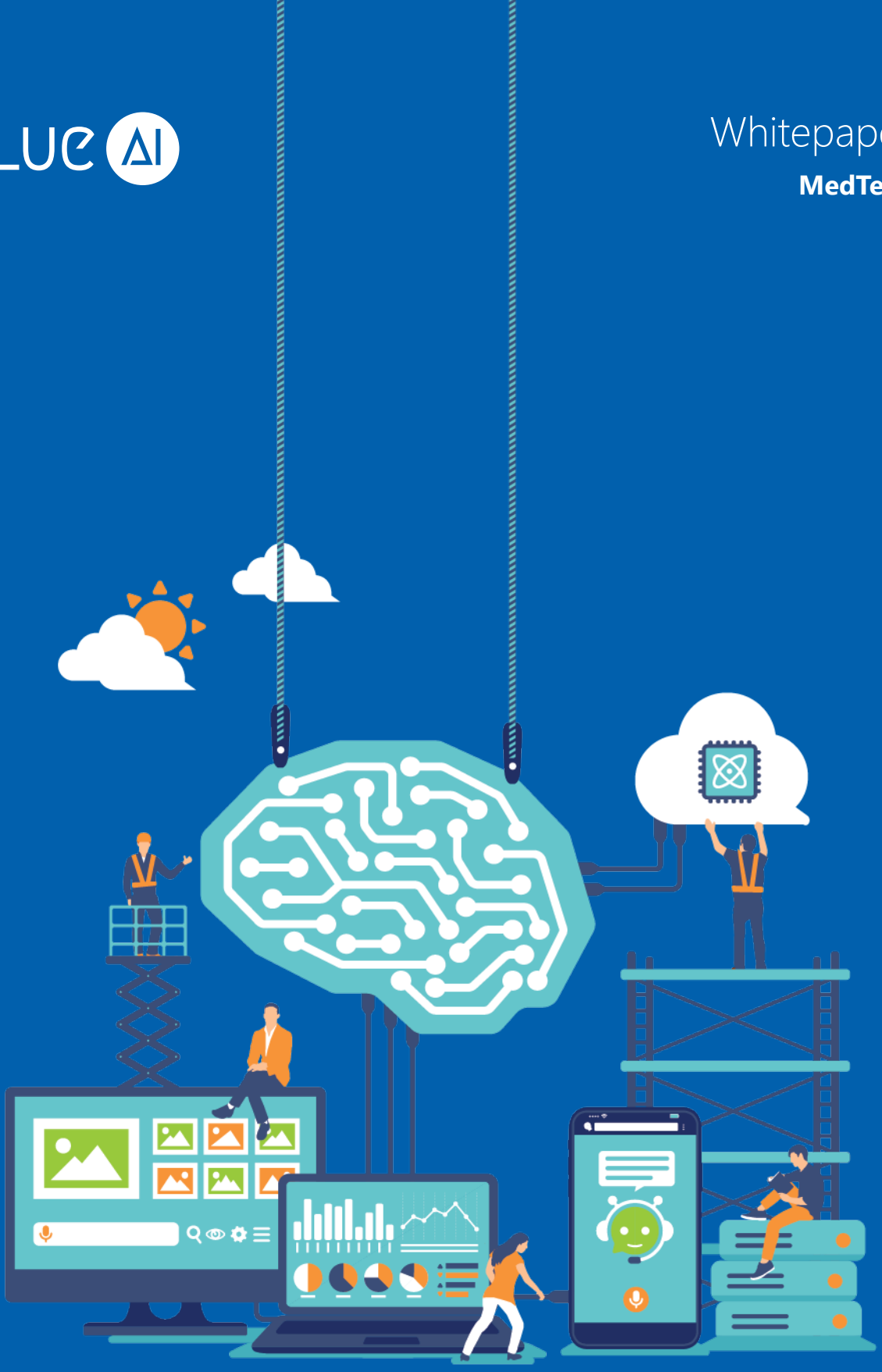
THEBLUE AI

Whitepaper

MedTech



theBlue.ai Optimizing MedTech processes with generative AI for security questionnaires

theBlue.ai Optimizing MedTech processes with generative AI for security questionnaires

TheBlue.ai and apoQlar partnered to streamline the process of filling out extensive security questionnaires required by hospitals interested in apoQlar's advanced AR and AI-based medical solutions. By leveraging Generative AI and Retrieval-Augmented Generation (RAG) technology, the project significantly reduced the time and effort needed to complete these questionnaires, enhancing accuracy and efficiency. This whitepaper details the challenges faced, the innovative solutions implemented, and the substantial benefits realized from this collaboration.

Medtech companies must fill out security questionnaires to ensure their products meet stringent security and compliance standards required by hospitals. These questionnaires protect patient data, ensure regulatory compliance, manage integration risks, build trust, and evaluate cybersecurity resilience. However, the detailed nature of these documents, requiring input from various departments, posed significant challenges for apoQlar, leading to delays and resource strains.

Challenge

The security questionnaires presented several key challenges:

- **Length and Detail:** Comprised of tens or hundreds of questions, these questionnaires required detailed responses about technical security measures and internal policies.
- **Cross-Departmental Involvement:** Answering these questions necessitated input from various departments, each with unique schedules and priorities.
- **Time-Consuming:** The need to consult extensive documentation and the involvement of multiple stakeholders made the process lengthy and cumbersome, often taking several weeks.

Solution

ApoQlar teamed up with the AI experts from theBlue.ai to address these challenges by implementing an innovative solution based on Generative AI enhanced by Retrieval-Augmented Generation (RAG). This solution integrated all company policies stored as PDF files and the technical documentation from Confluence into a virtual assistant named Zippy. Zippy could accurately answer specific questions by referencing exact sources, including document names and page numbers.

The solution was built on Microsoft Azure, ensuring secure data processing. Azure OpenAI Services were used for creating document embeddings, utilizing the latest models from the GPT family. ChromaDB was employed as a vector database for efficient data retrieval.

Results

The new solution significantly improved efficiency and accuracy in completing security questionnaires. Previously, the process took about a month and required 8-10 people. With Zippy, it can now be done in less than a week with fewer people, mainly for verifying answers. This allows employees to focus on their core tasks, boosting productivity. The accuracy of responses was enhanced, as answers were derived from up-to-date sources. Additionally, the feedback mechanism encouraged continuous documentation updates, improving the virtual assistant's reliability.

> "The inefficiency in handling these questionnaires was not just a productivity issue; it also impacted our ability to swiftly onboard new clients. We needed a smarter, faster way to manage this process."
>
> **Sirko Pelzl, CEO of apoQlar**

The Role and Importance of Security Questionnaires

Security questionnaires are widely used across various industries, including healthcare, finance, and technology, to assess the security and compliance posture of potential vendors and partners. These questionnaires are essential for several reasons:

- **Data Protection:** Ensuring that sensitive information, such as patient data in healthcare or financial data in banking, is secure and protected from breaches and unauthorized access.
- **Regulatory Compliance:** Verifying adherence to industry-specific regulations such as HIPAA and GDPR in healthcare, PCI DSS in finance, and various cybersecurity standards like ISO/IEC 27001.
- **Risk Management:** Identifying potential vulnerabilities and risks associated with integrating new technologies or services into existing systems.
- **Vendor Assurance:** Building trust between organizations by demonstrating that vendors have implemented adequate security measures and are committed to protecting sensitive data.

In the healthcare industry, security questionnaires help hospitals ensure that MedTech solutions comply with stringent regulations and standards designed to protect patient data and maintain the integrity of medical devices. These questionnaires typically include detailed questions about data encryption, access controls, incident response, and vulnerability management.

The detailed nature of these questionnaires often necessitates input from various departments within an organization, including IT, legal, and compliance. This cross-departmental involvement can lead to coordination challenges and delays, making the process time-consuming and resource-intensive.

Implementation Details

Architecture Overview

To effectively combine proprietary company data with Large Language Models, theBlue.ai employed a Retrieval-Augmented Generation (RAG) architecture. This architecture allows efficient utilization of company data to generate accurate responses to user queries. The solution comprises an intuitive web application built on Streamlit with a Python backend, Azure OpenAI Services for language models (LLMs), ChromaDB as a vector store for effective data retrieval, preprocessing scripts, and LangFuse for monitoring the solution. The entire application is deployed on Microsoft Azure, ensuring secure and compliant data processing.

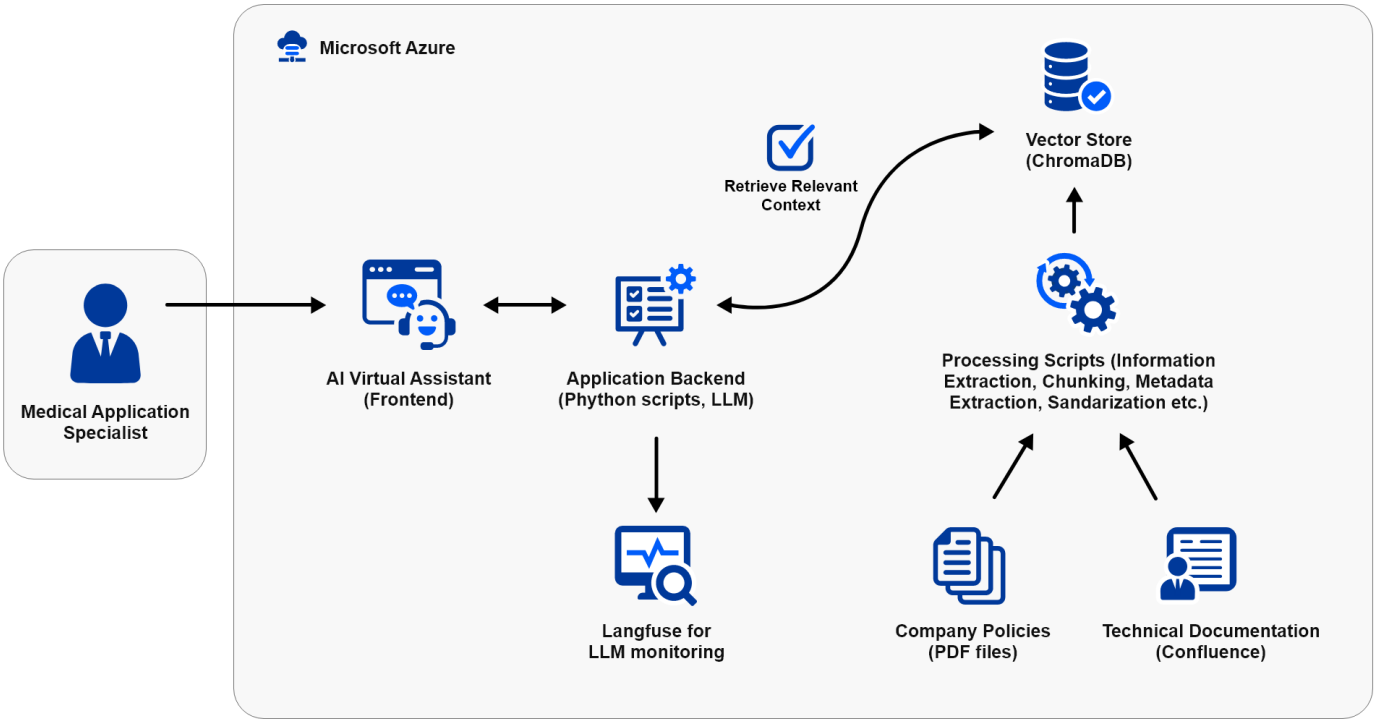


Figure. The image illustrates the integration of proprietary company data with Large Language Models (LLMs) using a Retrieval Augmented Generation (RAG) architecture.

Creating an effective RAG-based system involves optimizing various steps, particularly data processing and architecture. Our project handled data from Confluence and various static files. Using Langchain plugins, we integrated Confluence data with additional metadata smoothly. However, PDFs posed challenges due to diverse formatting and structures. To ensure scalability, we developed custom scripts for data extraction, focusing on creating coherent text segments while also extracting additional metadata, such as document pages. This metadata enables users to verify the exact sources used in generating answers and to ask follow-up questions. After extensive experimentation, we employed separate chunking strategies tailored to different data sources to maintain the structural context of the information. Each chunk retained its metadata, enhancing search capabilities. We used Ada embeddings from Azure OpenAI Services to optimize retrieval results in our RAG architecture.

Prompt Engineering
Prompt engineering was crucial for solution performance. It included, for example, crafting system prompts that provide the virtual assistant with the appropriate task context and ways of including retrieved chunks, while also ensuring that the answers are based solely on relevant context to minimize hallucinations. Guardrails were implemented to keep the model focused on relevant topics. Function-calling mechanisms redirected conversation flows to appropriate sources, enhancing the system's relevance and accuracy.
We explored additional techniques, such as the HyDE method, which creates artificial answers to enhance retrieval for complex questions. This feature is available as an optional parameter, providing flexibility in handling challenging queries.

> "Through the integration of Retrieval-Augmented Generation, we've enabled apoQlar to complete security questionnaires faster and with greater precision. This is a strong demonstration of AI's potential to optimize complex business processes."
>
> **Agata Chudzinska, CTO of theBlue.ai**

LLM Monitoring
To monitor the solution's performance, we implemented LangFuse. This tool is used for prompt versioning and tracking all model traces for debugging purposes. LangFuse captures all relevant steps, including chat history, retrieved text chunks, responses, function calls, errors, costs, and API latency, ensuring comprehensive monitoring and analysis.

Application and Continuous Improvement
The application, built on Streamlit with a Python backend, features a chat interface for user interaction. An advanced options panel allows users to enable additional prompt techniques like HyDE, switch LLMs (with GPT-4o from Azure OpenAI Services as the default), or adjust the number of text chunks retrieved per query. Collecting user feedback on specific conversation pieces was crucial for continuous improvement, integrated into LangFuse for seamless adaptation based on user input.
Based on user feedback, we have also implemented an additional application feature enabling users to get more detailed answers to specific topics. This functionality extends queries to neighboring chunks based on the metadata of relevant chunks used in retrieval, allowing for more comprehensive exploration of complex subjects.

Benefits and ROI

The implementation of the AI-driven solution by theBlue.ai and apoQlar has brought significant benefits, both tangible and intangible, to the company's operations. The benefits can be seen in the following areas:

Time Efficiency: The time required to complete security questionnaires was reduced dramatically. Previously, the process took about a month and involved 8-10 people. With Zippy, the process can be completed in just one week with far fewer people involved.

Cost Savings: By significantly reducing the time spent on filling out questionnaires, the company has saved on labor costs. Assuming an average hourly wage of $50 for the employees involved, the cost savings are substantial. Saving 120 hours per questionnaire at $50 per hour equates to $6,000 saved per questionnaire. With an average of 15 questionnaires per year, this amounts to $90,000 in annual savings.

Improved Accuracy: The AI assistant, Zippy, ensures that responses are accurate and consistent, reducing the risk of errors that could lead to compliance issues or reputational damage.

Enhanced Productivity: Employees are now able to focus on their core responsibilities rather than spending excessive time on documentation. This has led to increased productivity across various departments.

Faster Client Onboarding: The ability to complete security questionnaires quickly and accurately has accelerated the client onboarding process, enhancing customer satisfaction and potentially increasing sales. The average onboarding time has been reduced from 6 weeks to 2 weeks, allowing for faster deployment of apoQlar's solutions in hospitals.

> "Managing the completion of security questionnaires is no longer a logistical nightmare. The new system is easy to manage and ensures our responses are accurate and comprehensive."
>
> **Maciej Antoszczuk, Tech Product Owner**

Summary

The collaboration between theBlue.ai and apoQlar has transformed the process of completing security questionnaires for hospitals. By leveraging advanced AI technologies, the project streamlined what was previously a cumbersome process, freeing up valuable resources and ensuring accuracy. This innovative approach sets a new standard for handling complex documentation requirements in the healthcare and MedTech industries, paving the way for future advancements.

About theBlue.ai

theBlue.ai is a leading company based in Hamburg, specializing in the development of custom and scalable AI solutions. With extensive expertise in generative AI, theBlue.ai provides tailored solutions that help businesses accelerate their digital transformation and efficiently implement innovative projects. From the concept phase to full implementation, theBlue.ai supports its clients in maximizing the potential of AI. The company has extensive experience and collaborates closely with renowned partners and clients across various sectors, including healthcare, pharmaceuticals, and more, to deliver sustainable and effective AI projects.

About apoQlar

apoQlar GmbH is a healthcare technology provider specializing in mixed and augmented reality. Their software platform, VSI HoloMedicine®, uses Microsoft HoloLens to convert medical images, clinical workflows, and medical education into interactive 3D mixed reality environments. This technology allows surgeons to visualize anatomical structures during surgical planning, enhancing precision, saving time, and reducing post-operative procedures. apoQlar's work is advancing the fields of medical practice, education, and collaboration.

Impressum

theBlue.ai GmbH
Raboisen 32
20095 Hamburg
Germany
Website: https://theblue.ai
E-Mail: contact@theblue.ai
Tel: +49 (0)40 280 56 248

THEBLUE AI