SICKLE CELL DISEASE (SCD) PHYSICIAN CHATBOT

STUDENTS: EMILY LEE, TIANLI (MARCUS) REN, QIZHUO SHEN, TEJASHREE KHEDKAR, AURELIA AUDREY KISANAGA, JIAQING (ALLEN) LIU, YONG LU

The Purpose of SCD Chatbot

- Sickle Cell Disease (SCD) is a rare genetic condition, and many doctors (also in emergency departments) lack the specialized knowledge necessary for effective treatment.
- SCD patients come from marginalized communities and often require emergency care, but the current healthcare system offers fragmented and inconsistent support for these patients.
- SCD chatbot provides immediate SCD-related expertise to doctors when they are treating patients with this disease, aiming to close the gaps in current care practices.

SCD Chatbot Overview

- The Chatbot App is developed with React Native for full iOS and Android compatibility, giving the healthcare providers a rich user experience with sophisticated designs designed in Figma.
- A web-based admin portal restricts data and admin access to authorized users, supporting secure and reliable medical data management.
- Employs data models specifically tailored to SCD protocol data for clinical analysis and treatment optimization.
- Utilize advanced analytical techniques to examine emotional, behavioral, and cultural data, thereby informing personalized care strategies aligned with treatment protocols.

App Interface & App Flow

- User opts to sign-up using relevant name, email, and password, or sign-in with an existing account using email and password only.
- Successful authentication will proceed to the chatbot where physicians can ask relevant questions to be answered.
- **Question mark** (?) : on the bottom left corner of texting screen as FAQs regarding SCD.
- **Profile page**: on the top right corner of the chat page to access the user's full name, email, and medical ID of the physician. This is also where the user can sign out from the account.
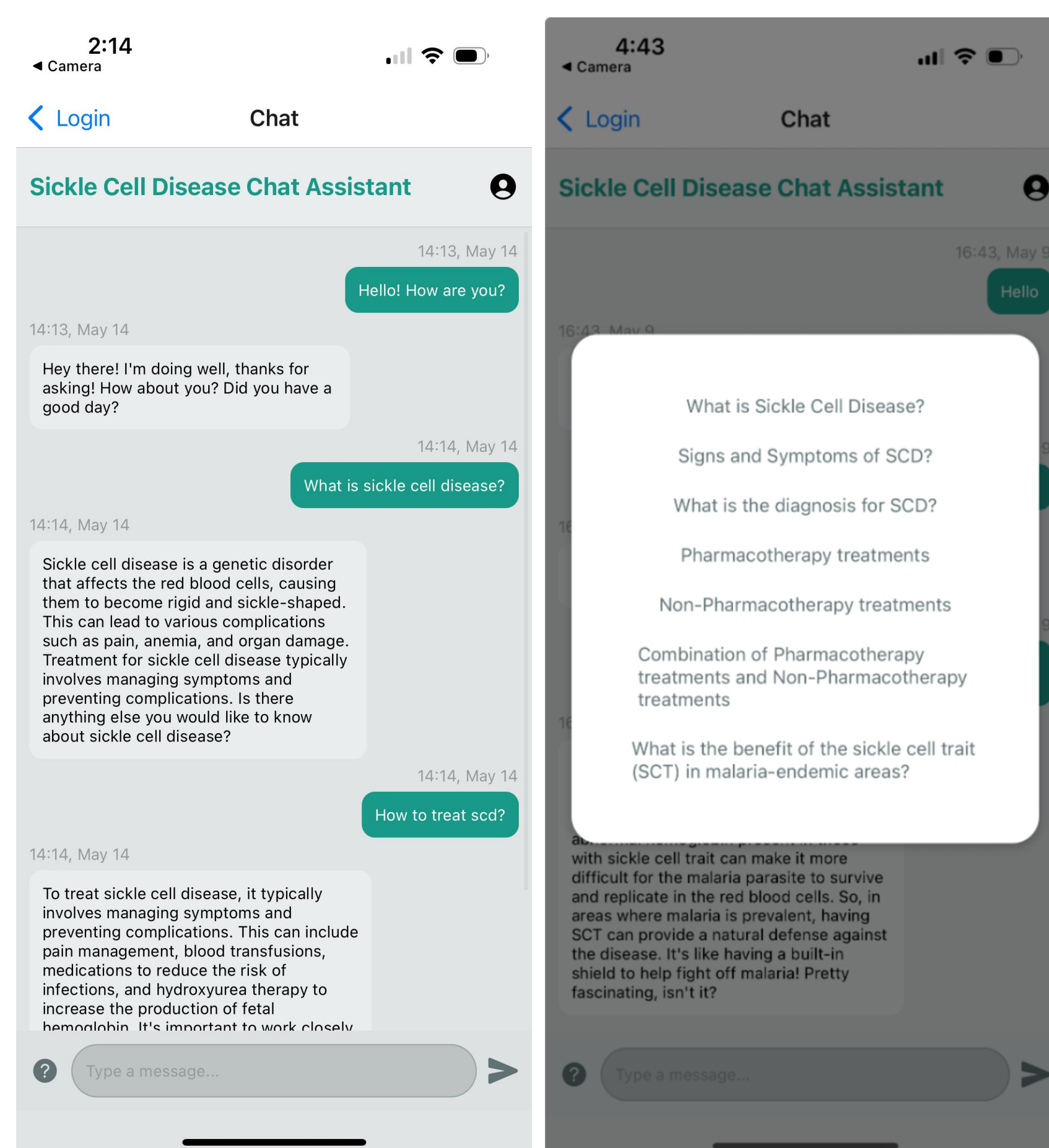

Technical Design & Implementation
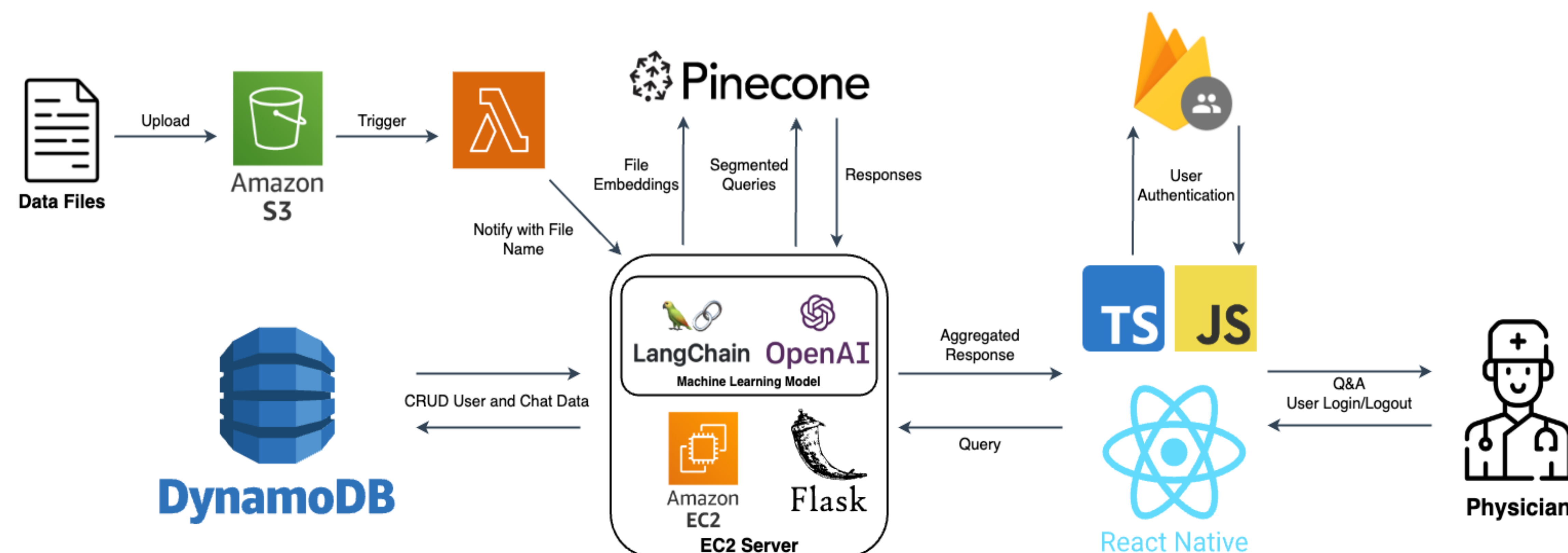
- The React Native front-end provides compatibility on both Android and iOS platforms and enables rapid iteration.
- Firebase Authentication by Google ensures security through OAuth 2.0.
- The machine learning model utilizes Langchain and the OpenAI API.
- An Amazon Lambda function is set up to respond to data file uploads.
- Amazon DynamoDB handles user and chat data due to its scalability, managed service, and low-latency performance.
- Amazon S3 stores raw data files securely, with ingress logging enabled for added protection.
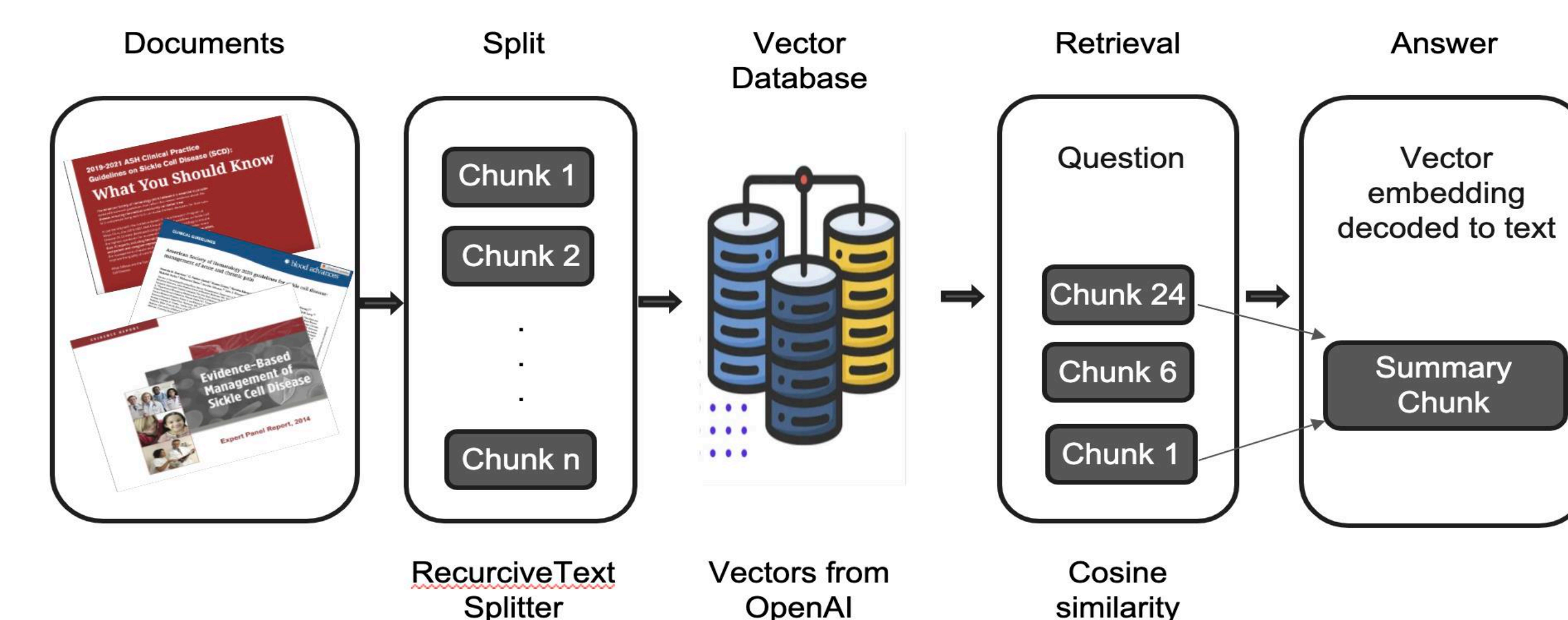
RAG Agents Flowchart
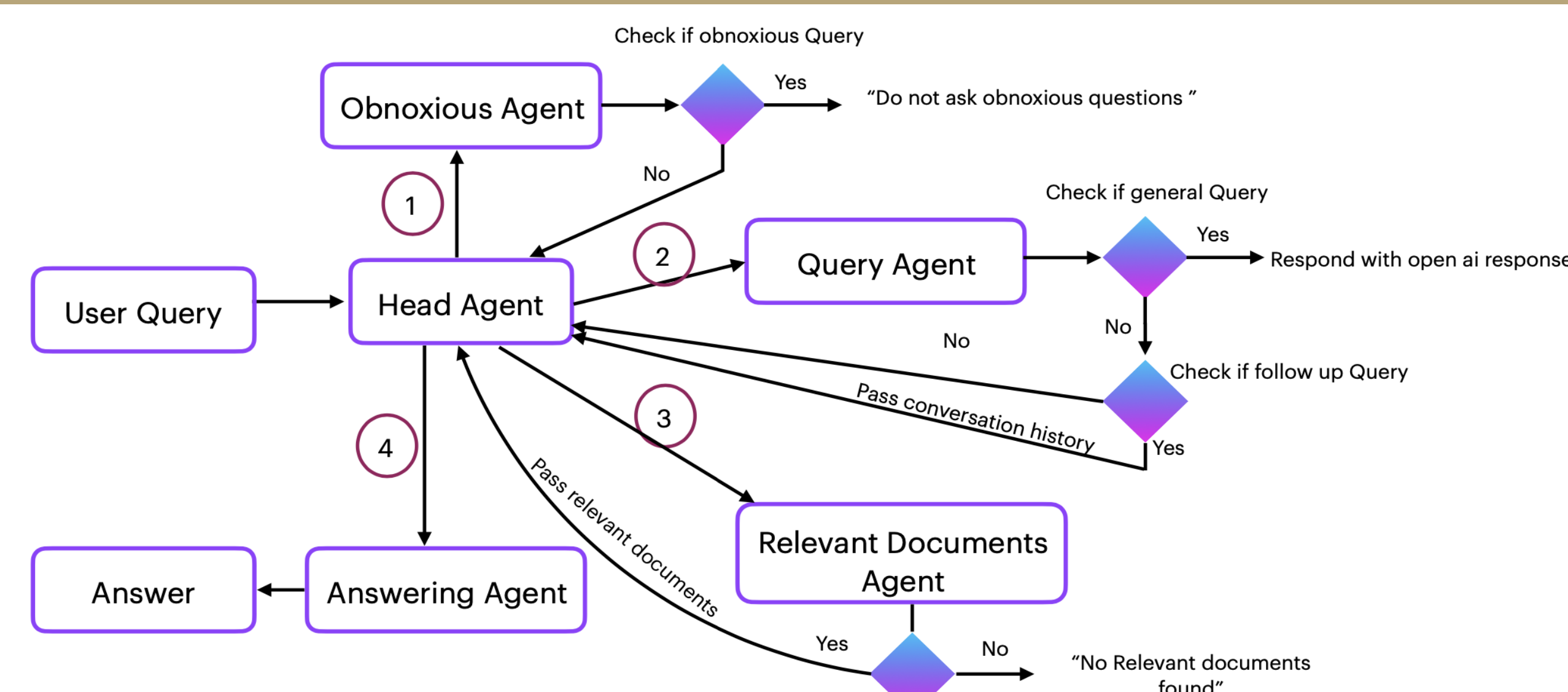
- **User Query**: Where queries are submitted.
- **Head Agent**: Coordinates query flow between agents.
- **Obnoxious Agent**: Screens for inappropriate queries.
- **Query Agent**: Determines query type.
- **Relevant Documents Agent**: Finds relevant documents, reports if none found.
- **Answering Agent**: Constructs final response from other agents' input.

LLM Technical Design

- Documents are meticulously curated and uploaded to S3 for a reliable chatbot.
- A Recursive Text Splitter, with Amazon Lambda and Langchain, divides documents into chunks for optimal vectors.
- GPT-3.5 Turbo converts chunks into embeddings.
- Pinecone Vector Database stores embeddings for fast retrieval.
- Queries are matched to document chunks for a concise summary.

Future Works

- Chat History : Support multiple search filters such as date, symptoms, or treatment plans for rapid retrieval of critical information.
- Voice-Typing help: It offers voice-to-text conversion and supports multiple languages, ensuring usability across diverse linguistic backgrounds.
- Feedback Button: Users will be able to provide feedback on the response of the chatbot to further improve datasets and machine learning in the future.

Acknowledgement, and References

[1] ASH, "Ash clinical practice guidelines on Sickle Cell Disease," Clinical Practice Guidelines on Sickle Cell Disease - Hematology.org, May 9, 2024.
[2] NIH, "EVIDENCE-BASED MANAGEMENT OF SICKLE CELL DISEASE: EXPERT PANEL REPORT, 2014"
[3] Y. Li et al., "Chatdoctor: A medical chat model fine-tuned on a large language model meta-ai (llama) using medical domain knowledge," arXiv.org, https://arxiv.org/abs/2303.14070, May 13, 2024

ADVISERS: John Canevari, Karthik Kumar, Shabana Motlani, Folashele Olabisi, Anubhav Srivastava, Karthik Mohan, Megha Chandra Nandyala
SPONSOR: Novo Nordisk