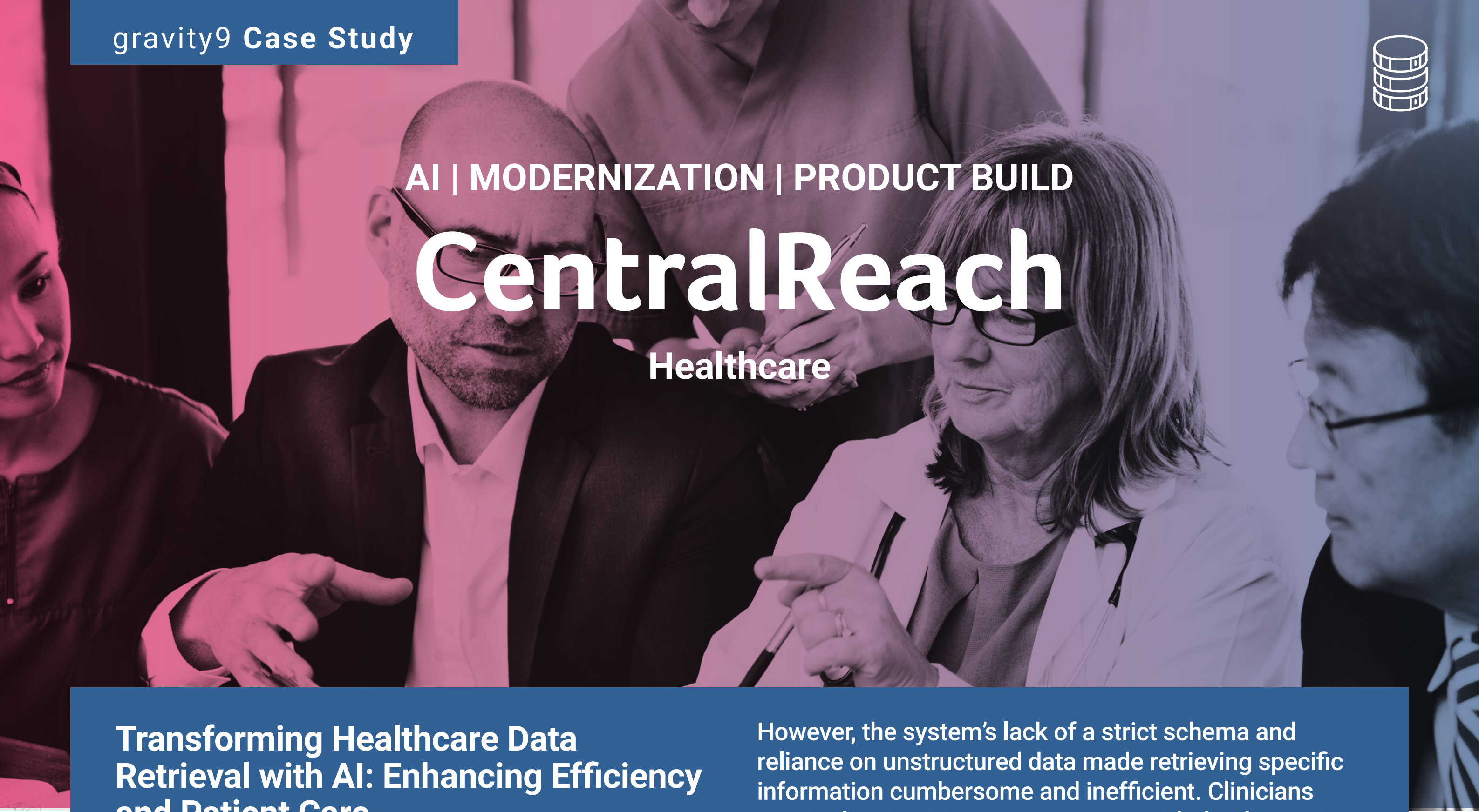

CentralReach

Healthcare

Transforming Healthcare Data Retrieval with AI: Enhancing Efficiency and Patient Care

Our client, CentralReach, a leading AI-powered healthcare software provider, specializes in Applied Behavior Analysis (ABA) and multidisciplinary therapies for autism spectrum disorder (ASD) care and faced significant challenges managing complex, unstructured patient data. Their Learning Tree system stored diverse information including demographics, medical history, treatment progress, and associated documents.

However, the system's lack of a strict schema and reliance on unstructured data made retrieving specific information cumbersome and inefficient. Clinicians required an intuitive way to interact with the data using natural language queries to support evidence-based practice. Recognizing the need for a more efficient solution, CentralReach sought an AI-powered system capable of handling complex, unstructured data and enabling seamless interaction through natural language queries.

Utilized Technology Stack

- Database: MongoDB
- Backend: FastAPI
- AI: LlamaIndex, MongoDB Atlas

Review of Challenges

The hierarchical nature of the Learning Tree system introduced complexities in data retrieval. Extracting specific patient information required significant manual effort due to the lack of standardized formats across different data points. Clinicians struggled with inefficiencies, spending hours sifting through unstructured data to generate reports and track patient progress. Additionally, the presence of diverse data types—JSON objects, PDFs, and scans stored in AWS S3—added another layer of complexity, making it difficult to implement a unified retrieval system. These challenges underscored the need for a scalable AI-powered solution to improve data accessibility and streamline workflows.

Our Solution

Our AI-powered system introduced an intelligent approach to retrieving and managing Learning Tree data. The solution was built around two key components: a Data Ingestion and Enrichment Pipeline and an Agentic Retrieval-Augmented Generation (RAG) Service using LlamaIndex and MongoDB.

The Data Ingestion and Enrichment Pipeline structured raw Learning Tree data to enhance searchability through taxonomy processing, standardizing data classification using Learning Tree Standard Reference Markdown files. The Learning Tree Analysis extracted hierarchical relationships and metadata, while data enrichment summarized nodes with Large Language Models (LLMs) and categorized them into structured formats.

> The expertise of the professionals on this project has been instrumental in helping us accelerate our progress and unlock new possibilities in generative AI

Our Approach

The high-level approach of our solution centers on leveraging advanced AI technologies to enhance data retrieval and response quality for clinicians. We designed an approach that integrates natural language processing (NLP) with database querying to enable seamless interaction with complex Learning Tree data stored in MongoDB. The system architecture consists of two main components: the Data Ingestion and Enrichment pipeline and the Agentic Service, which together transform raw data into a structured knowledge base and provide AI-driven query capabilities.

The Data Ingestion and Enrichment pipeline is a critical component that processes raw data to prepare it for efficient querying. This pipeline includes several key stages: Taxonomy Processing, Learning Tree Analysis, and the Enrichment Pipeline. The taxonomy processing step builds a hierarchical structure of categories using Learning Tree Standard Reference Markdown files, ensuring consistent classification of nodes. The Learning Tree Analyzer extracts data from Learning Trees, analyzes their structure, and adds hierarchical relationships as metadata. The Enrichment Pipeline further enhances the data by summarizing nodes using large language models (LLMs), categorizing them into predefined taxonomy categories, and extracting content from diverse file formats stored in AWS S3 buckets. This process ensures that the data is well-structured, categorized, and ready for advanced retrieval techniques.

Our system employs a hybrid approach to data retrieval, combining vector-based semantic search with traditional keyword-based search using MongoDB Atlas functionalities. We leveraged MongoDB's full-text search, vector search, and hybrid search capabilities to retrieve relevant nodes from the Learning Tree.

To facilitate this process, we developed an expression-based retrieval tool that serves as a high-level interface for users. This tool utilizes large language models (LLMs) to translate natural language queries into intermediate domain-specific language (DSL) expressions, which are then directly converted into MongoDB queries. This approach ensures that user queries are processed efficiently and accurately, providing optimal results for both semantic and lexical search types.

The Agentic RAG Service is designed to fill progress report sections guided by user queries leveraging the retrieval tools mentioned above. It uses AI-driven workflows to refine user queries, retrieve relevant data, and generate answers. Key components of this service include query refinement workflows, reasoning and response generation workflows, and citation workflows. The system also employs LangFuse for monitoring and logging all agentic interactions. This enables labeling responses in order to test and evaluate the system.

This AI solution serves as the foundation for a retrieval-based document intelligence system, empowering healthcare professionals. Clinicians can now obtain precise answers to their queries without manually sifting through extensive Learning Tree data, allowing them to focus more on what matters instead of the details of the Learning Tree format. The system provides a chat interface for searching patient data, making it easier to obtain accurate and relevant responses. Finally, tracing tools enable evaluating the system behavior, enhancing solution robustness.

Subsequent Outcomes

The implementation of this solution resulted in significant improvements. Clinicians could now retrieve patient information quickly, reducing report generation time from days to minutes. Structured, categorized data allowed for seamless navigation and interaction, while AI-assisted search improved the relevance of retrieved data, reducing errors in treatment reporting. By automating data extraction and retrieval, clinicians could focus more on patient care rather than administrative tasks. AI-driven monitoring and evaluation tools ensured ongoing system improvements, creating a robust and reliable retrieval solution.

Client Feedback

The MongoDB MAAP program has been a game-changer for our AI initiatives. Gaining access to industry-leading AI experts and thought leaders, such as gravity9, has provided us with invaluable insights and strategic support. The expertise of the professionals on this project has been instrumental in helping us accelerate our progress and unlock new possibilities in generative AI.

> "It surfaces all that information and puts it right at your fingertips. You don't have to hunt and peck, or worry if someone filled out the right field or not"
>
> David Stevens,
> Head of AI at CentralReach.

Visit our Insights page for more articles about emerging technology trends, the IT industry, interviews, and more!