



Microsoft Innovation Challenge

Edu Echo



Interactive Voice-Based Learning Assistant

**Presented by Antoinette Davis, Natasha
Williams, and Kharee Smith**



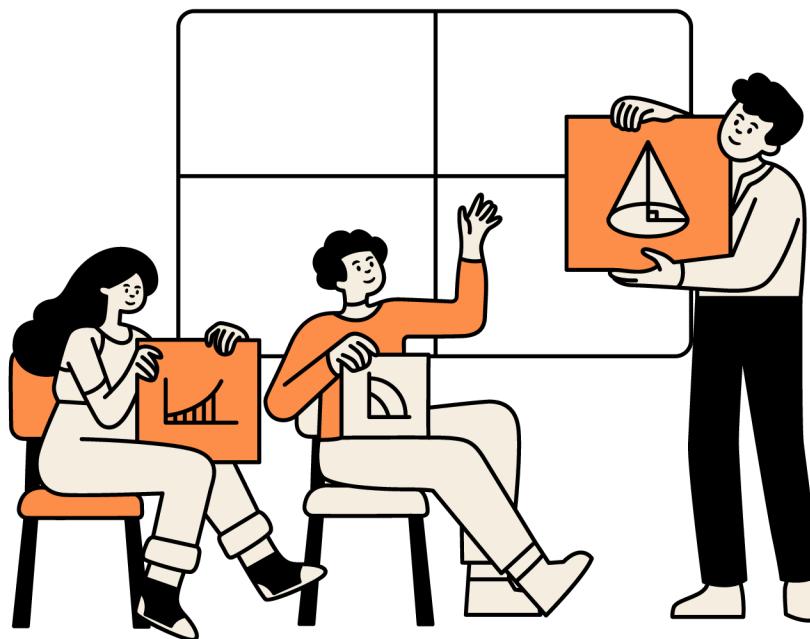
The Challenge: VoiceRAG

Towards more natural application user interfaces



- The new Azure OpenAI gpt-4o-realtime-preview model opens the door for even more natural application user interfaces with its speech-to-speech capability.
- How do you implement retrieval-augmented generation (RAG), the prevailing pattern for combining language models with your own data, in a system that uses audio for input and output?

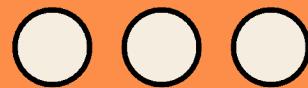
The Solution: Edu Echo: Your Interactive Voice-Based Learning Assistant



Edu Echo is a voice-first education platform designed to help elementary and middle school students excel in STEM subjects.

Students can ask questions like:

- "How do I multiply multi-digit numbers?"
- "Can you help me understand verb tenses for my 4th-grade language arts class?"
- "Please explain context clues in simple terms."



Edu Echo Features

Using Azure OpenAI GPT-4o Realtime, it combines conversational interfaces Retrieval-Augmented Generation (RAG) to deliver personalized, step-by-step explanations and answers to academic queries.

Interactive Voice Tutoring

Step-by-Step Problem Solving

Personalized Learning Experience

Focused Subject Support

Multi-Language Support

Demo

EDU ECHO



Talk to a tutor!

Components

Azure OpenAI GPT-4o Realtime API

Processes voice input and generates conversational responses.

Azure AI Search



Retrieves relevant content from the stored documents.

Uses semantic ranking to return the most relevant document.

Azure OpenAI Text Embedding API

Generates embeddings (numerical representations of text) for the stored documents.

Helpful for the AI Search.

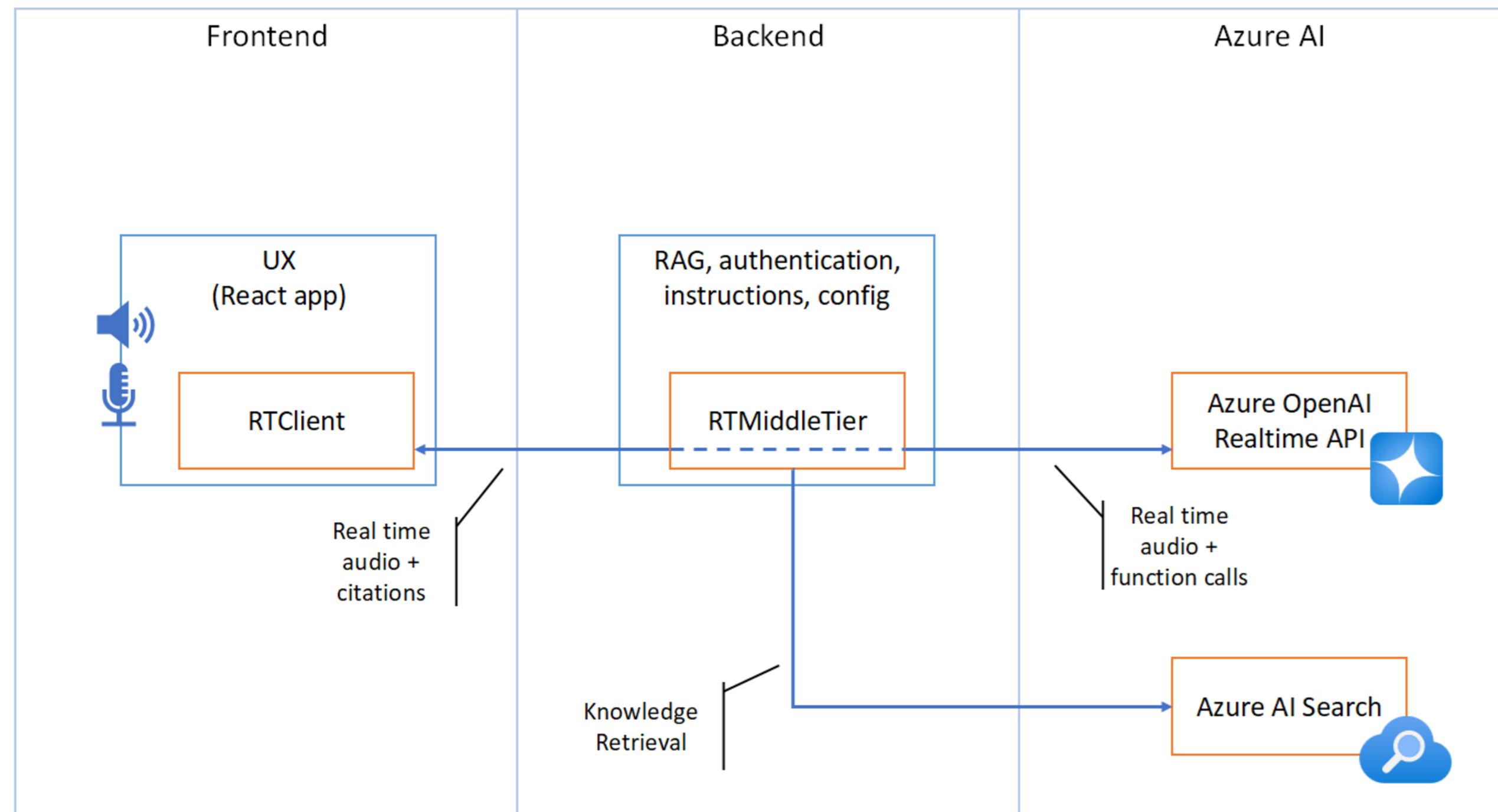


Azure Storage

Stores educational documents in BLOB storage for retrieval.

Architecture

VoiceRAG: app pattern for RAG with real-time audio-enabled models



Key Learnings



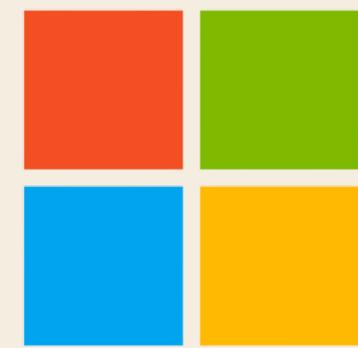
- Real-time voice demonstrates amazing potential in educational applications.
- Optimizing the RAG pipeline is crucial for having contextually relevant answers.
- Building the solution required iterative testing to fine-tune the voice interface and RAG pipeline.
- Multi-language support and cloud-based deployment make it feasible to expand to international markets.

Future State

- Gamified Learning Approach (XP, Levels, Achievements, Quests, etc.)
- User Profiles (Student, Admin, Parent, etc.)
- More Subjects (Science, Social Studies), Subtopics, and Grades (3, 7, 8, etc.)
- Personalized Learning Plans
- Image Understanding and Visual Guidance
- Web Search (Additional Grounding)
- More Accessibility Features
- AI Tutoring Agents



Thank You!



Microsoft

