Names :
Razi Mograbi
Dolev Seri

GlobalCommunication

GlobalCommunication is a near real-time communication system created to seamlessly bridge language barriers through integrated speech-to-speech translation app with voice cloning.

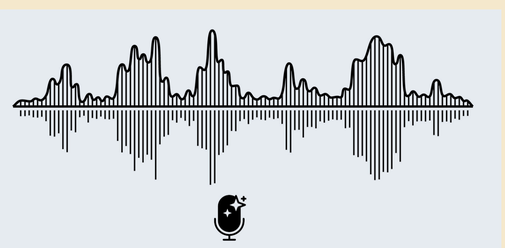
Advisors :
Dr. Dan Lemberg
Mrs. Elena Kramer

BACKGROUND

- Communication Gap
  - Language barriers often prevent efficient voice communication between people in different languages.
  - Current solutions typically provide audio output that sounds artificial and robotic.
- Why Our System?
  - Our system centralizes AI-based speech-to-text, translation, and text-to-speech for near real-time cross-lingual calling

REQUIREMENTS

- Develop an android mobile app.
- Ensure Speech-to-Speech translation with voice cloning in supported languages.
- Near real-time latency



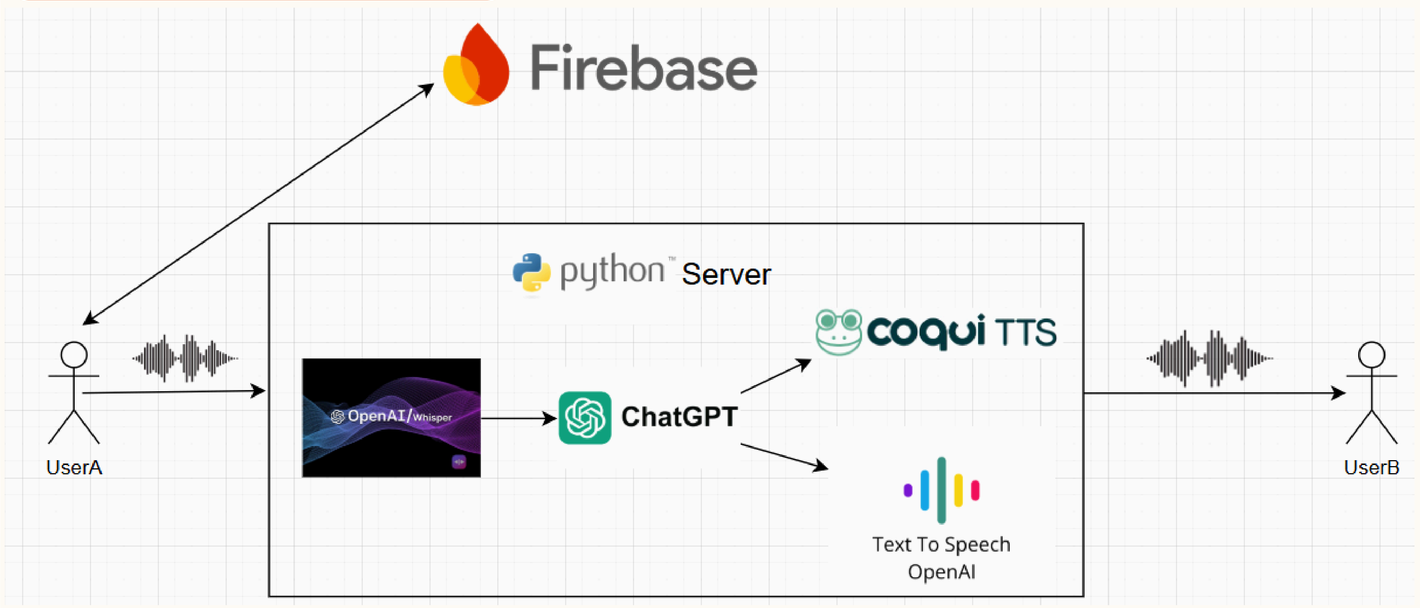
ARCHITECTURE & TOOLS

- Android Client - with android studio
- Firebase integration for user management - user auth and db.
- FastAPI Server with WebSockets
- Redis for Metadata & Ordering
- Chunking & Buffering Logic
- AI Processing Chain:
  - Whisper (Speech-to-Text)
  - ChatGPT 4o-mini (Translation)
  - XTTS / OpenAI TTS (Text-to-Speech)

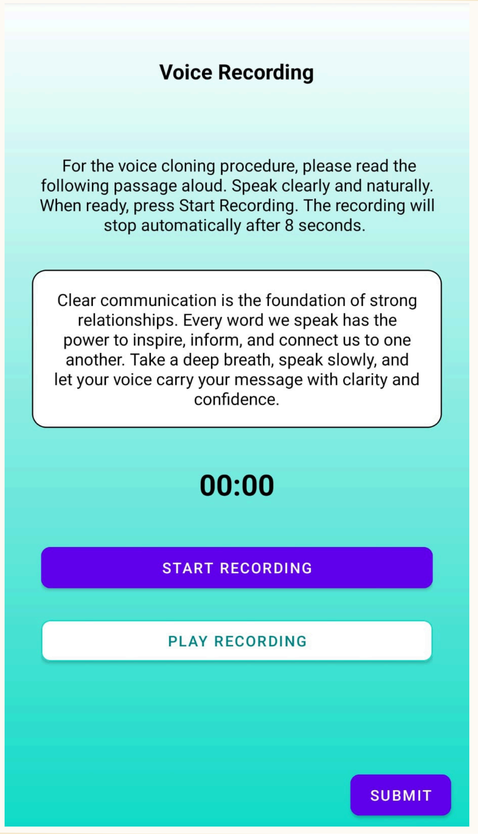
RESULTS & ACHIEVEMENTS

- User-Friendly Android Mobile App
- Seamless Two-Person Voice Calling
- Cross-Language Compatibility
- Voice Cloning for Natural Conversation (Supported Languages Only)
- Low-Latency Speech Translation (Achieves ~5s delay with regular TTS, ~5–6s with GPU voice cloning, and ~7–10s on CPU)

DIAGRAM & APP



Android A → FastAPI Server + Redis → Whisper → ChatGPT → XTTS→ Android B.



CHALLENGES

- Packet Fragmentation
- Language Coverage: XTTS is limited to certain languages
- Data Ordering: Ensuring correct sequencing of audio chunks
- User Management without server overhead.

CONCLUSION & FUTURE WORK

- Project Highlights: A working system for cross-lingual calls, successfully integrating AI-based translation and voice cloning.

- **Potential Improvements:**

1. Expand XTTS Language Support for broader voice cloning.
2. Enable Group Calls for multi-participant scenarios.
3. Instead of API’s we can use models that run on GPU’s
4. Use the openAI beta real-time models.(New)