

Interacción Humano-Computador

PC3 - Aporte Personal

Nombres y Apellidos: Guido Anthony Chipana Calderon

Código: 20221428H

Aporte en la Implementación del Asistente de Voz Conversacional

Mi contribución principal en este proyecto fue el diseño y desarrollo integral del asistente de voz conversacional que simula un psicólogo, el núcleo interactivo de la aplicación de apoyo emocional. A continuación, se detalla la arquitectura y las tecnologías implementadas para dar vida a esta funcionalidad.

1. Captura y Síntesis de Voz (Interfaz del Navegador)

Para crear una interacción fluida y natural, utilicé dos APIs nativas de los navegadores web, evitando la necesidad de librerías externas para las funciones básicas de voz.

Reconocimiento de Voz (SpeechRecognition API):

- Implementé esta API para capturar la voz del usuario a través del micrófono.
- La voz se transcribe a texto en tiempo real directamente en el navegador.
- El componente SessionView.tsx gestiona el estado de la escucha (activado/desactivado) y procesa el texto resultante.
- Se incluyó manejo de errores y permisos, contemplando casos donde el usuario no autoriza el uso del micrófono.

Síntesis de Voz (SpeechSynthesis API):

- Para que el asistente pudiera "hablar", utilicé esta API para convertir las respuestas de texto generadas por la IA en audio.
- La función speak en SessionView.tsx se encarga de vocalizar los mensajes del asistente.
- Se implementó una lógica para dividir textos largos en fragmentos más cortos, asegurando una reproducción continua y sin cortes en todos los navegadores.

2. Integración con el Servicio de Inteligencia Artificial

El "cerebro" del asistente reside en el aiService.ts, un servicio que diseñé para ser modular y robusto.

Prompt Engineering (Definición de la Personalidad):

- Creeé un SYSTEM_PROMPT cuidadosamente diseñado que define la personalidad del asistente. Se le instruye para ser un asistente terapéutico virtual empático, profesional y comprensivo.
- Este prompt guía a la IA para que escuche activamente, valide emociones, haga preguntas abiertas y mantenga un tono cercano, asegurando que las interacciones sean siempre seguras y de apoyo.

Conexión con Múltiples Proveedores de IA:

- Aunque el proyecto utiliza Groq por defecto (con el modelo Llama 3) por su rapidez y bajo costo, la arquitectura permite cambiar fácilmente a otros proveedores como OpenAI (GPT-4o mini), Anthropic (Claude 4.5 Sonnet) o Google (Gemini 1.5 Flash) si en algún momento necesitamos que sea más robusto.
- El servicio gestiona el historial de la conversación, enviando el contexto necesario en cada petición para que el asistente pueda mantener conversaciones coherentes.

Manejo de Errores y Fallbacks:

En caso de que la API de la IA falle, implementé una respuesta de respaldo predefinida para que la conversación no se interrumpa bruscamente.

3. Interfaz de Usuario y Flujo de la Conversación

La interfaz de la sesión (SessionView.tsx) fue construida para ser intuitiva y centrarse en la conversación.

- **Visualización de la Conversación:** Un diseño de chat clásico que muestra los mensajes del usuario y del asistente.
- **Entrada Dual:** Aunque el foco es la voz, añadí un campo de texto como método de entrada alternativo, garantizando la accesibilidad para usuarios que no pueden o no desean usar el micrófono.
- **Retroalimentación Visual:** La interfaz proporciona feedback en tiempo real, mostrando cuándo el asistente está "escuchando", "procesando" o "hablando".

4. Funcionalidades de Apoyo y Seguridad

Consciente de la naturaleza sensible de la aplicación, integré funcionalidades clave para la seguridad del usuario.

- **Detección de Crisis:** Implementé una función detectCrisis en aiService.ts que analiza los mensajes del usuario en busca de palabras clave relacionadas con suicidio o autolesión. Si se detecta una crisis, la aplicación muestra automáticamente un modal (CrisisModal) con recursos de ayuda profesional inmediata.
- **Sugerencia de Ejercicios:** El sistema puede identificar cuándo la IA sugiere un ejercicio de respiración. Al detectarlo, la interfaz ofrece al usuario la posibilidad de iniciar un ejercicio guiado de respiración 4-7-8

(BreathingExercise.tsx), conectando la conversación con una herramienta práctica de manejo de la ansiedad.

En resumen, mi aporte consistió en construir un sistema de conversación completo que no solo entiende y responde al usuario, sino que lo hace de una manera empática, segura y terapéutica, integrando tecnologías de voz, inteligencia artificial y una interfaz de usuario cuidadosamente diseñada para el bienestar emocional.