

GPT assistant

Nombre: César Luis Briceño Ojeda

Correo institucional: Cbricenoojeda42@correo.unicordoba.edu.co

Usuario: cesarbriceno

¿Cuál es el objetivo principal de la aplicación? ¿Cuál es el problema o necesidad que busca resolver?

El objetivo principal de la aplicación es proporcionar una interfaz para convertir voz en texto, procesar esa información con ChatGPT y luego convertir la respuesta en texto a voz. El problema que busca resolver es permitir una interacción de voz natural con la tecnología, facilitando la comunicación oral y generando respuestas habladas en tiempo real, en pocas palabras, crear un asistente personalizado.

¿Quiénes son los usuarios de la aplicación? ¿Cuáles son sus características y necesidades específicas?

Los usuarios de la aplicación pueden ser personas que deseen interactuar con ChatGPT de manera oral en lugar de texto escrito. Sus características pueden incluir la capacidad de hablar y escuchar, y su necesidad principal es poder comunicarse con el modelo de ChatGPT de forma eficiente y obtener respuestas habladas en lugar de texto escrito para resolver todas sus preguntas e inquietudes.

¿Qué funcionalidades se requieren para cumplir con el objetivo de la aplicación? ¿Cuáles son las más importantes?

- Conversión de voz a texto mediante una API de reconocimiento de voz.
- Procesamiento del texto obtenido utilizando el modelo de ChatGPT.
- Síntesis de voz a partir de la respuesta generada por ChatGPT utilizando una API de síntesis de voz.

¿Cómo debe ser la interfaz de usuario? ¿Qué elementos gráficos y de diseño son importantes para los usuarios?

1. Botón de grabación de voz: Un botón prominente en la pantalla principal de la aplicación que permite al usuario iniciar y detener la grabación de voz. Puede tener una representación visual clara, como un ícono de micrófono.

2. Visualización del texto en tiempo real: Mientras el usuario habla, el texto convertido de la voz se muestra en tiempo real en la pantalla. Esto permite al usuario ver lo que se está transcribiendo y verificar la precisión de la conversión.
3. Botón de procesamiento: Después de que el usuario haya terminado de hablar, un botón de procesamiento permite al usuario enviar el texto transcrito a ChatGPT para su procesamiento y generación de una respuesta.
4. Visualización de la respuesta: Una vez que ChatGPT ha generado una respuesta, se muestra en la pantalla en formato de texto. Esto permite al usuario leer y comprender la respuesta generada.
5. Botón de reproducción de voz: Junto a la respuesta generada por ChatGPT, se coloca un botón de reproducción de voz. Al hacer clic en él, se reproduce la respuesta en formato de voz, permitiendo al usuario escuchar la respuesta generada.

¿Qué plataformas y dispositivos deben ser compatibles con la aplicación? ¿Deben ser móviles, de escritorio, o ambos?

La aplicación puede ser compatible con una variedad de plataformas y dispositivos, incluyendo dispositivos móviles (como smartphones y tablets) y computadoras de escritorio, ya que la aplicación será web.

¿Qué requisitos de seguridad y privacidad son necesarios para la aplicación?

Es importante considerar la seguridad y privacidad de los datos de voz y texto durante la transmisión y el procesamiento. Se deberá realizar una encriptación de datos y cumplir con las regulaciones de privacidad aplicables.

¿Cuáles son los plazos y presupuestos para el desarrollo de la aplicación?

El presupuesto para el proyecto es de \$0 ya que se trabajará una prueba beta y se utilizarán APIs con pruebas gratuitas. Por otro lado, los tiempos y plazos estarían programados para finalizar al ir terminando 3 corte.

¿Hay alguna tecnología o herramienta específica que deba ser utilizada en el desarrollo de la aplicación?

Para la conversión de voz a texto, podemos utilizar APIs como Google Cloud Speech-to-Text o IBM Watson Speech to Text. Para el procesamiento de texto con ChatGPT, se puede utilizar la API de OpenAI. Y para la síntesis de voz, opciones como Google Cloud Text-to-Speech o Amazon Polly pueden ser utilizadas. Además, se requerirán tecnologías para desarrollar la interfaz de usuario y gestionar la comunicación entre las diferentes APIs.