



Prueba técnica

Puesto: Desarrollador con enfoque a Inteligencia Artificial

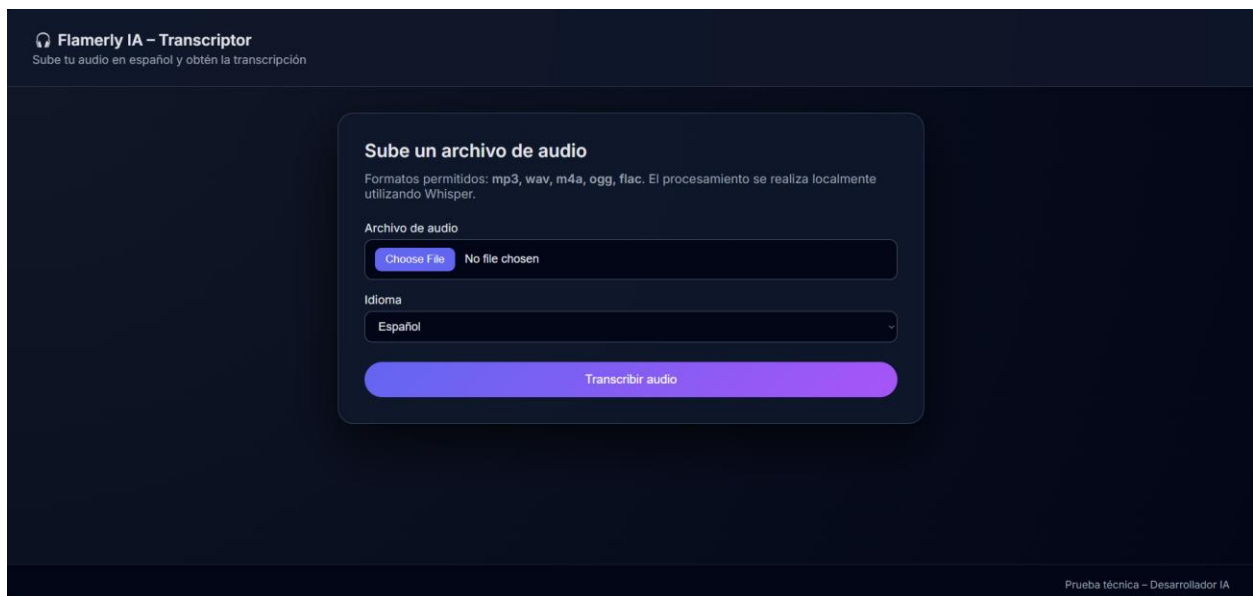
Mauricio Julio Rodrigo Lemus Guzmán

Contexto

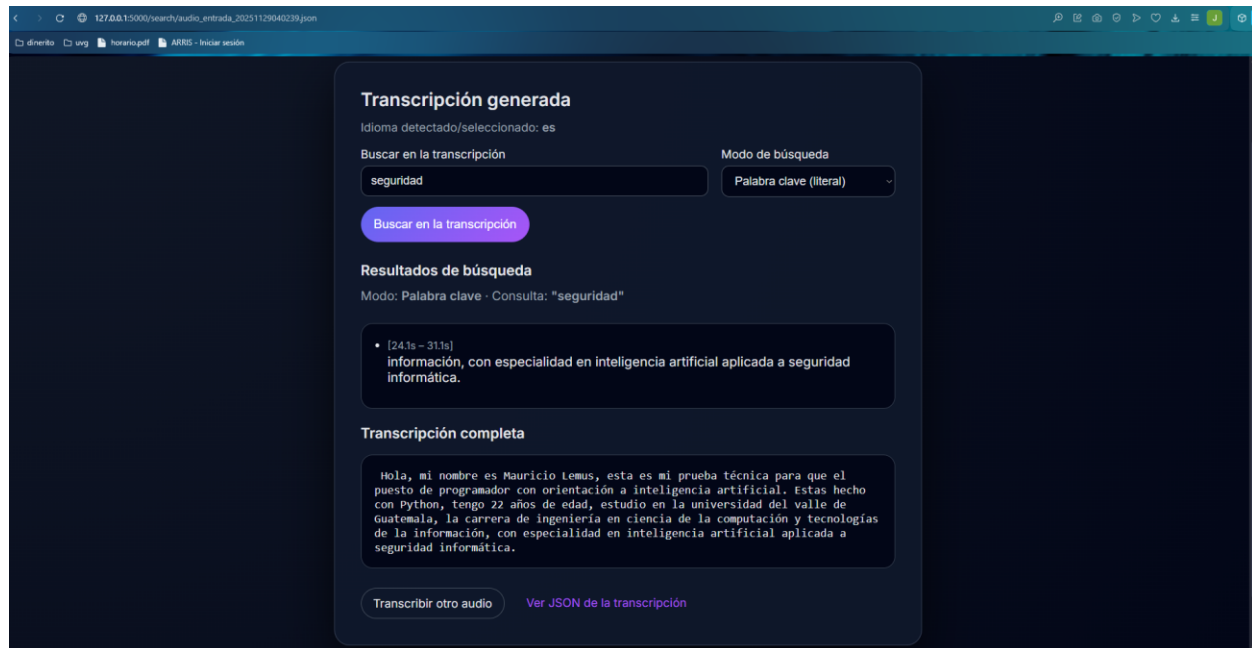
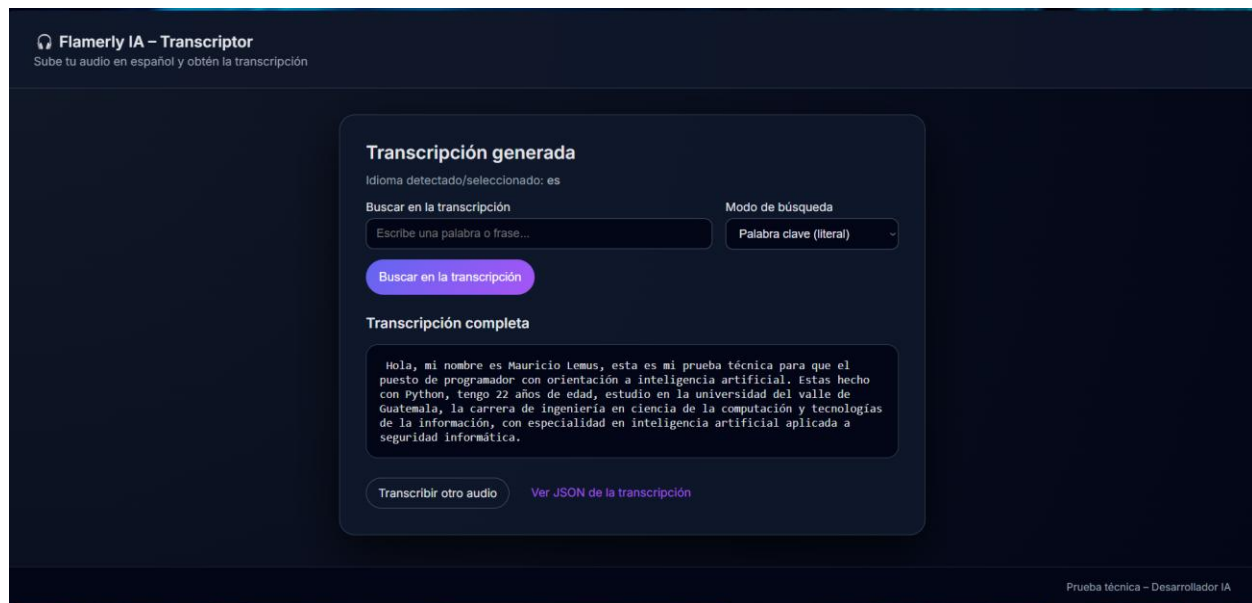
El proyecto desarrollado consiste en un sistema completo de transcripción de audio en español, diseñado para evaluar la capacidad de integrar modelos de reconocimiento automático de voz (ASR) dentro de una aplicación funcional. La solución implementa Whisper como motor de transcripción local, permitiendo procesar audios sin depender de servicios externos, y almacena los resultados en formato JSON para facilitar su manipulación. Se incorporaron dos métodos de búsqueda dentro de la transcripción: una búsqueda literal por palabra clave y una búsqueda semántica basada en embeddings, esta última como valor agregado opcional.

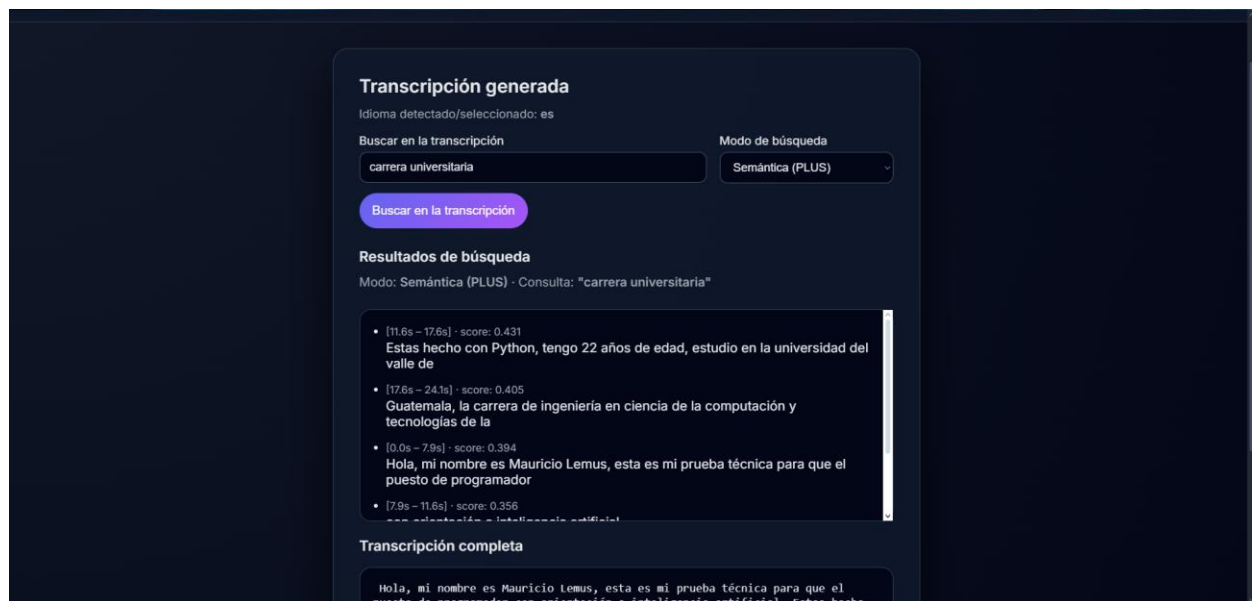
El sistema fue complementado con una interfaz web sencilla que permite al usuario subir un archivo, visualizar la transcripción generada y ejecutar consultas sobre ella, cumpliendo con los requerimientos técnicos establecidos y demostrando un enfoque práctico, modular y escalable

Funcionamiento del sistema (Interfaz web)



The screenshot displays the web interface for 'Flamerly IA - Transcriptor'. The header includes the logo and the tagline 'Sube tu audio en español y obtén la transcripción'. The main content area features a central form titled 'Sube un archivo de audio'. This form contains instructions on permitted audio formats (mp3, wav, m4a, ogg, flac) and a note about local processing using Whisper. It includes a file selection area with a 'Choose File' button and a 'No file chosen' status. Below this is a language dropdown menu currently set to 'Español'. A large, prominent blue button labeled 'Transcribir audio' is positioned at the bottom of the form. The footer of the page identifies it as a 'Prueba técnica - Desarrollador IA'.





Conclusión

La implementación desarrollada demuestra la integración exitosa de un sistema completo de transcripción automática de voz en español, complementado con capacidades de búsqueda literal y búsqueda semántica. A través del uso de Whisper para el procesamiento del audio y un modelo de embeddings para la recuperación contextual, se logró construir una solución robusta, funcional y autónoma que no depende de configuraciones adicionales en el entorno del usuario.

La interfaz web facilita la interacción con el sistema, permitiendo cargar audios, visualizar transcripciones y realizar consultas de manera intuitiva. Este proyecto no solo cumple con los requerimientos establecidos en la prueba técnica, sino que también incorpora mejoras opcionales que fortalecen su alcance y valor técnico, evidenciando un entendimiento sólido de los componentes involucrados y una capacidad clara para diseñar soluciones prácticas y escalables.