

Programación y Robótica

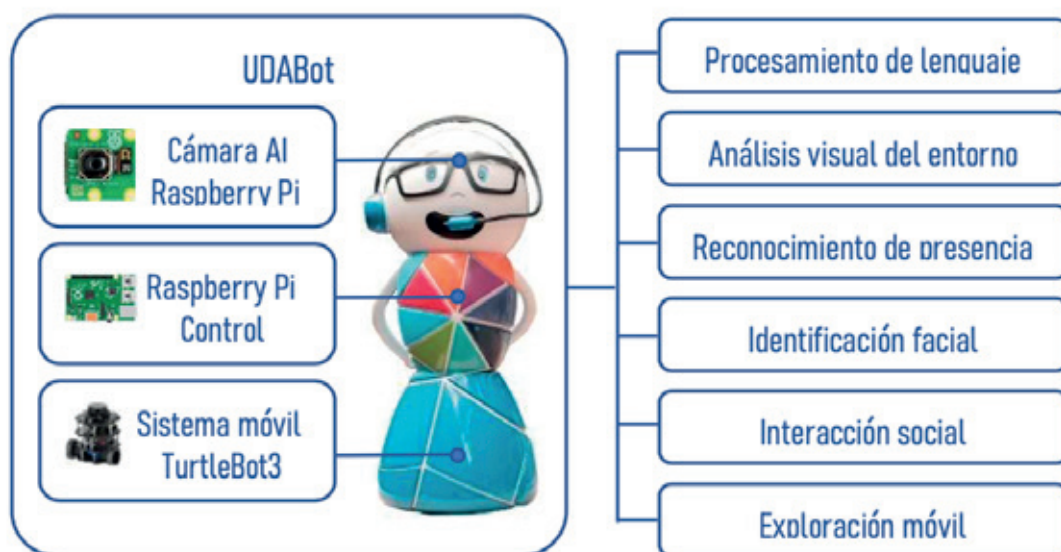
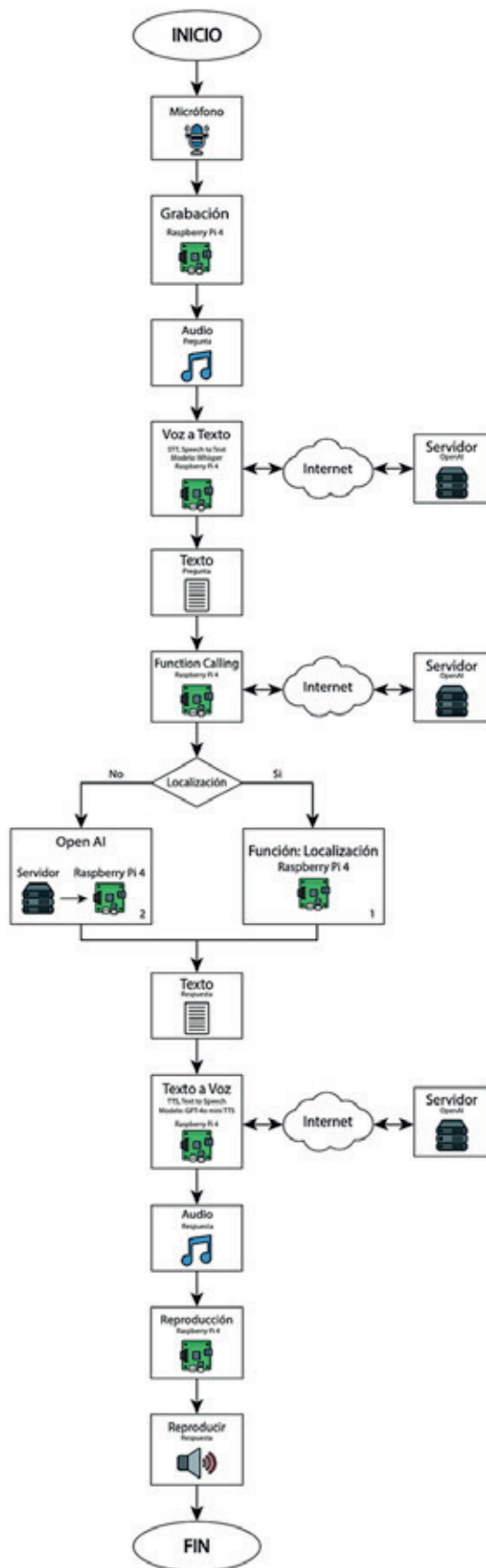


Fig.1. Funcionalidades del UDAbot Raspberry Pi Control - Chatbot



En la **Fig. 2** se observa el diagrama de flujo completo del sistema de chatbot integrado en el Raspberry Pi, donde se especifica la estructura y los pasos a seguir para un correcto funcionamiento del robot social.

Fig. 2.- Diagrama de flujo del funcionamiento del sistema de Chatbot.

A. Grabación

Esta etapa representa el inicio del sistema para la interacción del robot social con el usuario, donde se captura una señal de audio a través de un arreglo de micrófonos, tal

como se observa en la Fig. 3. Ésta registra de forma precisa la voz o los comandos del usuario, el cual mediante un Raspberry Pi genera un archivo de audio que será utilizado para las siguientes etapas.



Fig. 3.- Diagrama del funcionamiento de la etapa de Grabación.

B. Voz a texto (STT, Speech To Text)

Para esta fase se convierte el archivo de audio almacenado anteriormente a una variable de tipo texto, tal como se observa en la Fig. 4. El proceso ocurre mediante métodos de reconocimiento de voz, para ello se utiliza el modelo conocido como Whisper, implementado por

la empresa OpenAI, el mismo que genera la transcripción del audio original a un archivo de texto. Para el desarrollo de este método, el procedimiento descrito anteriormente tiene que ser procesado en el servidor de OpenAI y no de manera local en el Raspberry Pi 4 por las limitaciones de procesamiento.

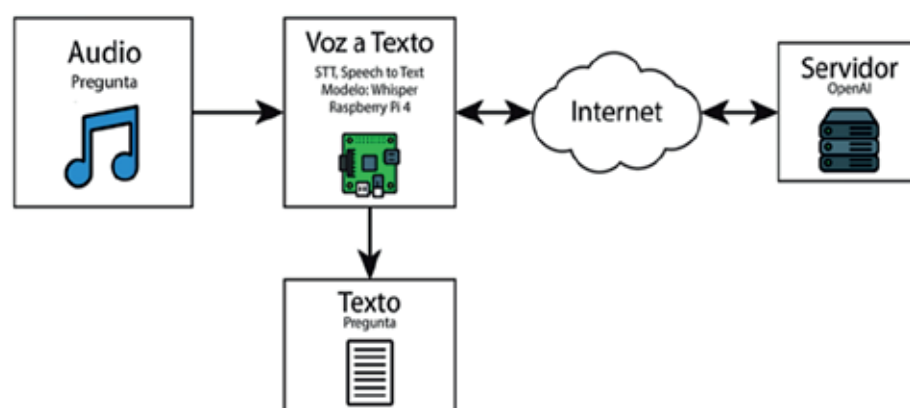


Fig. 4. Diagrama del funcionamiento de la etapa STT.

C. Function Calling

En esta etapa, se integra la función conocida como Function Calling, que se observa en la Fig. 5, proporcionada por la API de OpenAI. Esta característica permite que el

modelo de lenguaje interactúe de manera dinámica con respecto a la ejecución de tareas específicas basadas en las consultas que realizó el usuario.

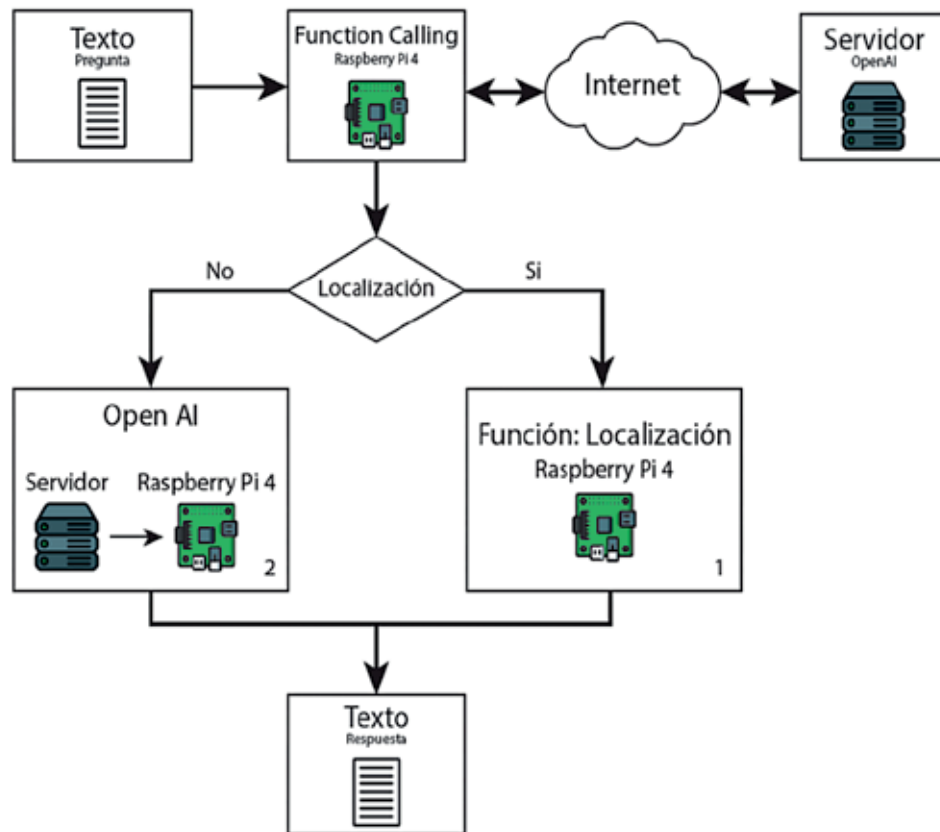


Fig. 5. Diagrama del funcionamiento de la etapa STT.

D. Texto a Voz (TTS, Text To Speech)

En esta fase, el modelo previamente ya generó una respuesta en formato texto, pero para que el robot social tenga una mejor interacción con el usuario que realizó la pregunta, la respuesta debe reproducirse de forma audible, como se puede evidenciar en la Fig. 6, por lo tanto, el archivo de texto generado

anteriormente tiene que ser procesado por los servidores de OpenAI, para utilizar el modelo GPT-4o mini TTS, donde se transforma a voz generada por inteligencia artificial, pero con una voz natural y expresiva. Esta respuesta audible se guarda en formato de audio en el Raspberry Pi 4 y puede personalizarse en aspectos como el tono, ritmo y cadencia

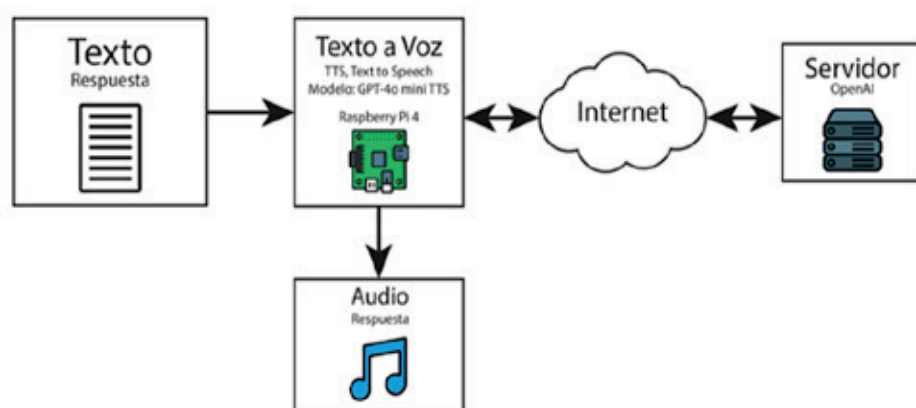


Fig. 6. Diagrama del funcionamiento de la etapa TTS.

E. Reproducción

En la Fig. 7 se detalla la etapa que emite el audio sintetizado a través de unos altavoces conectados directamente con el Raspberry Pi, donde se garantiza que la respuesta en forma de voz se reproduzca de manera inmediata y con

claridad. Aquí se confirma que el usuario que realizó la pregunta al UDABot, puede recibir de forma efectiva la información procesada en todas las etapas previas en forma de voz, completando la comunicación bidireccional del sistema completo.

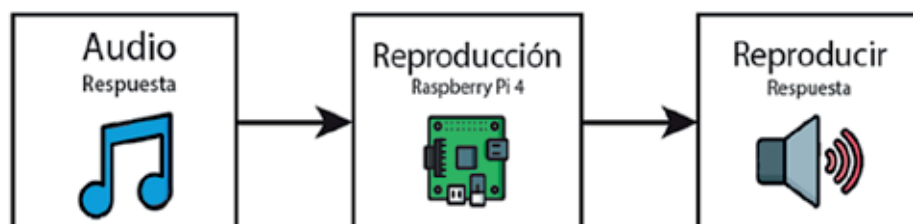


Fig. 7. Diagrama del funcionamiento de la etapa de Reproducción.

Sistema Móvil - Turtlebot3

Como se observa en la Fig. 8. Las tareas de navegación del UDABot se fundamentan en el movimiento controlado de un robot Turtlebot 3 Burger. En un miniordenador Raspberry Pi, con un sistema operativo Linux Ubuntu y sobre el cual se corre ROS (Robot Operating System), que es el framework de código abierto más usado por la comunidad mundial dedicada al desarrollo y creación de robots. Este framework se potencia con el uso de varias librerías y herramientas que permiten trabajar de forma modular con sensores y actuadores. En el UDABot se incluyeron

sensores para la navegación autónoma, principalmente una cámara de profundidad, complementada con unidades de medición inercial (IMU) integradas a su placa de control OPEN CR. Para el movimiento se seleccionaron ruedas Mecanum, debido a su omnidireccionalidad. El algoritmo programado para la detección de objetos, se basa en el uso de la cámara de profundidad, la cual informa la distancia a la que se encuentran los diferentes objetos del entorno por el que circula el robot social. Actualmente, el programa establece que si hay un objeto a 50 cm de distancia de frente, el UDABot se acercará reduciendo su velocidad; y en caso de llegar a una distancia mínima crítica el UDABot se detendrá automáticamente para evitar una colisión. La IMU integrada en la placa OPEN CR se encarga de dar la información de odometría del robot para monitorear el movimiento y estimar la posición del UDABot durante su navegación autónoma.

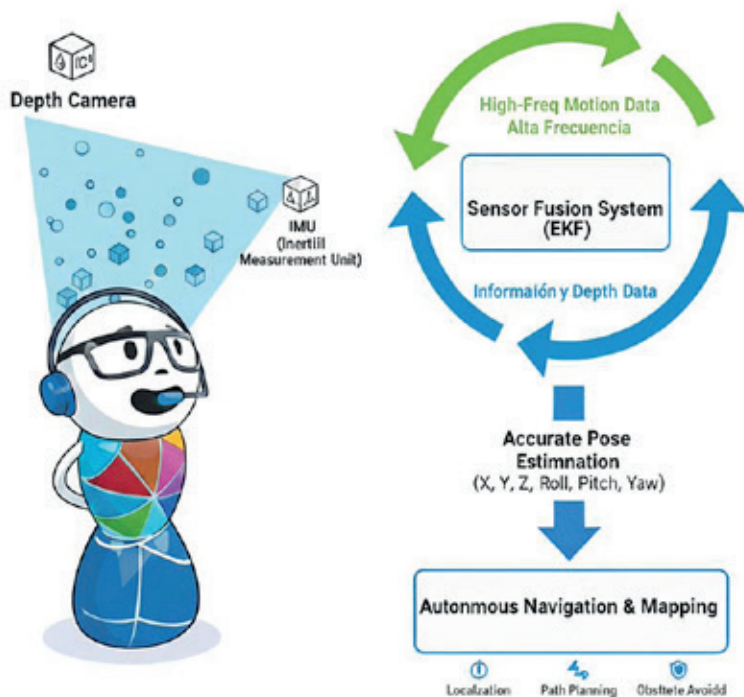


Fig. 8. Sistema Móvil - Turtlebot3.