

## Introduction

The ACSO - Research Center in Computer Architecture, Intelligent Systems and Robotics at the State University of Bahia - UNEB, Brazil, introduces BILL (the Bot, Intelligent, Large capacity and Low cost). BILL is part of the RoboCup@Home league of RoboCup Competition, which aims to promote research on assistive robot technology and the development of fully autonomous robots capable of interacting with a domestic environment.

## Bill - Bot Intelligent Large capacity Low cost

Bill is a service robot capable of communicating with humans through natural language processing, recognizing people and objects, and navigating environments. In its third generation Bill, codename Estranho, had its architecture wholly redesigned and was updated to ROS 2.

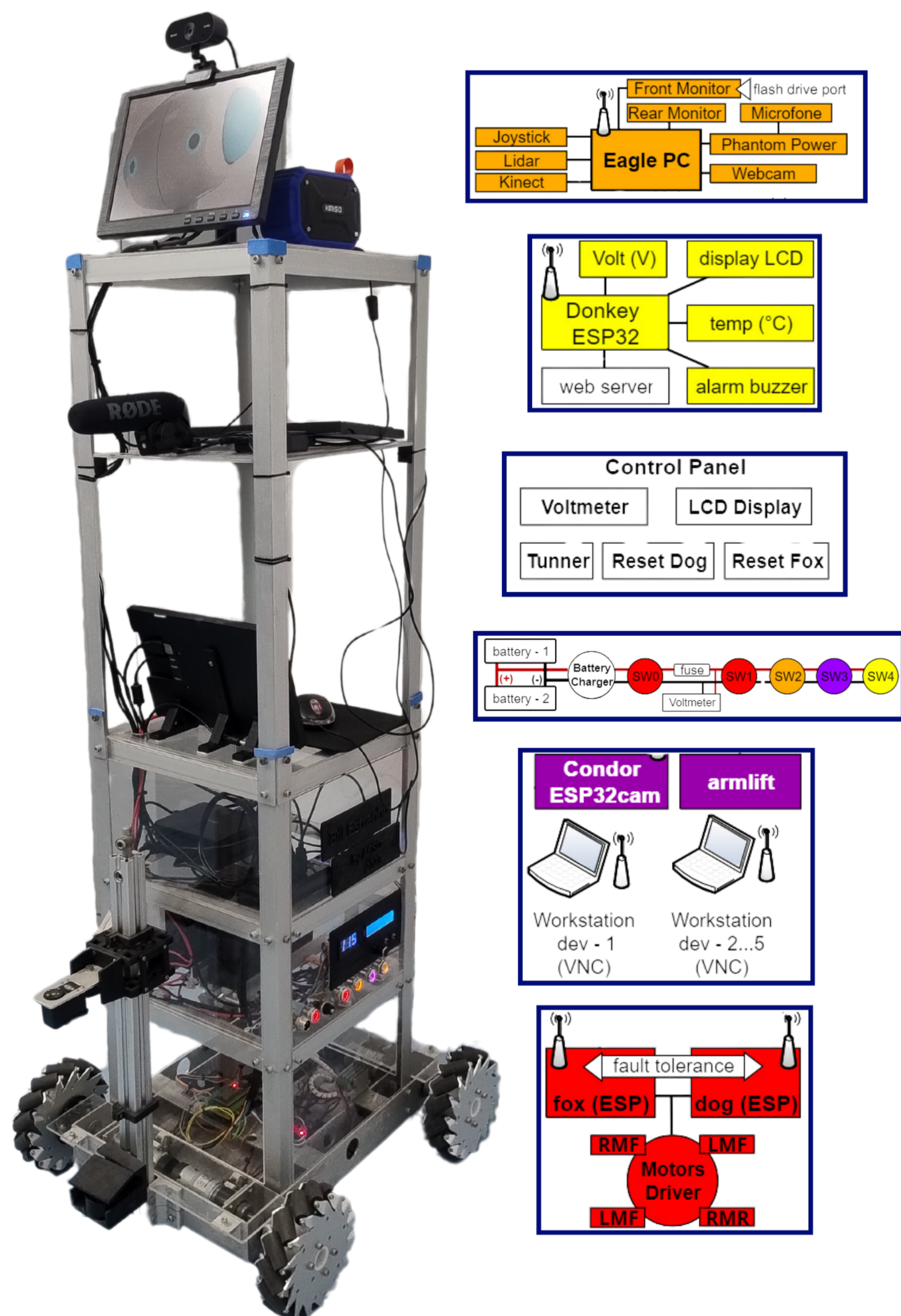


Figura 1: BILL Estranho

Bill's main components include:

- Two monitors, one representing Bill's face and another used for input and output via console;
- A microphone and speakers;
- Lidar and webcam;
- Additional manual control performed via a joystick using a mobile application;
- A gateway and a computer as part of his body;
- Access for installing new software performed remotely using VNC software.

## Bill's Capacities

- Object Detection, using YOLOv8;
- People Recognition, using OpenCV library for facial detection and classification, and the Dlib algorithm for recognition;
- Motion, using 4 Mecanum wheels, which allow it to move in any direction and turn around its own axis with ease;
- Navigation, using encoders, ROS 2 modules, and a 360° laser scanner for accurate robot control, mapping, and obstacle avoidance;
- Speech Recognition and Voice, using Google Speech Recognition software to enable voice interaction and ChatGPTs (version 2) model for understanding and generating responses in natural language.

## Bill's Natural Language Process with ChatGPT

Our speech recognition system employs advanced machine learning techniques, specifically the GPT-2 model for understanding and generating responses in natural language. Developed by OpenAI, GPT-2 is a transformer-based language model for generating coherent and contextually relevant text. For our purposes, GPT-2 has been fine-tuned with a custom dataset consisting of various command phrases and natural language interactions relevant to BILL.

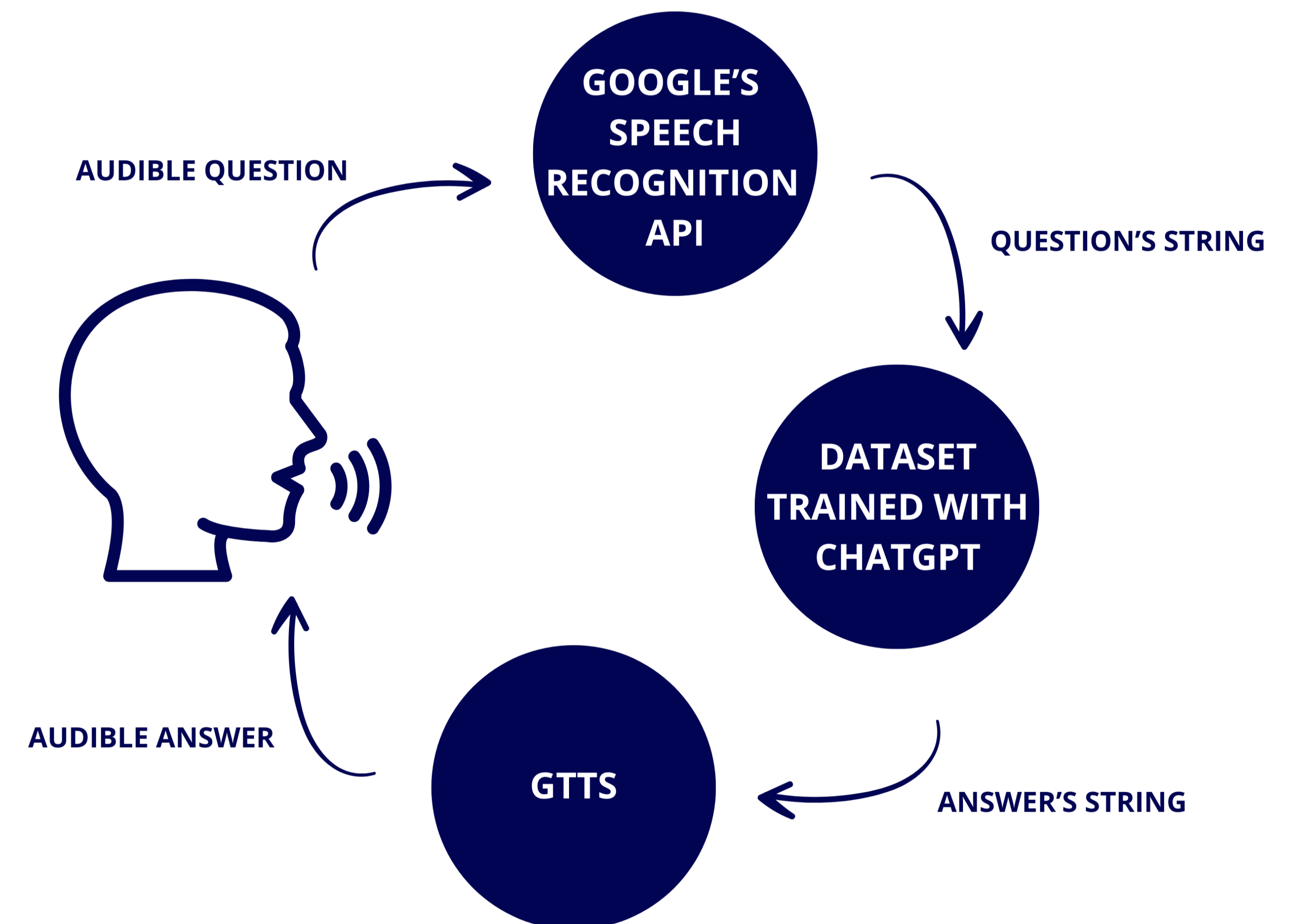


Figura 2: Bill's Natural Language Process using ChatGPT

During inference, the system captures voice commands through BILL's microphones. The captured audio is pre-processed to enhance clarity and reduce background noise. The processed audio is then transcribed into text using Google's Speech Recognition API. The transcribed text is fed into the fine-tuned GPT-2 model, which interprets the command and generates an appropriate response or action.

## Conclusion

Bill Estranho presents significant progress, specially in ROS 2 integration and speech recognition. The combination of AI and speech recognition of natural language gives Bills a versatile communication tool that improves its iteration possibilities.

## Partners: