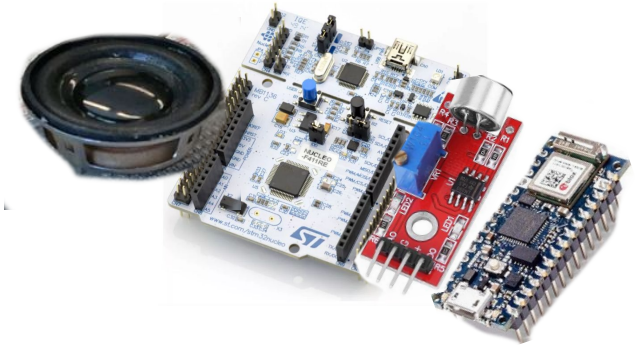
Embedded AI Recording-and-Playback Speech Device

Colin's Group

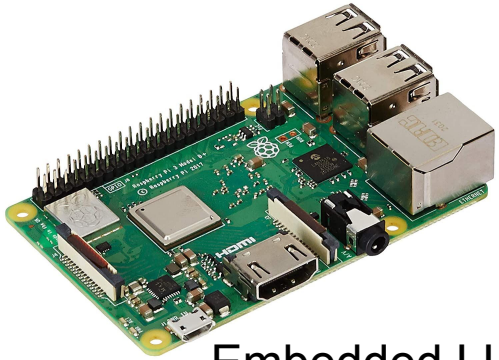
Kevin Conley
Stephen Angerer
Owen Lambert

Project Idea

“What color is grass?”



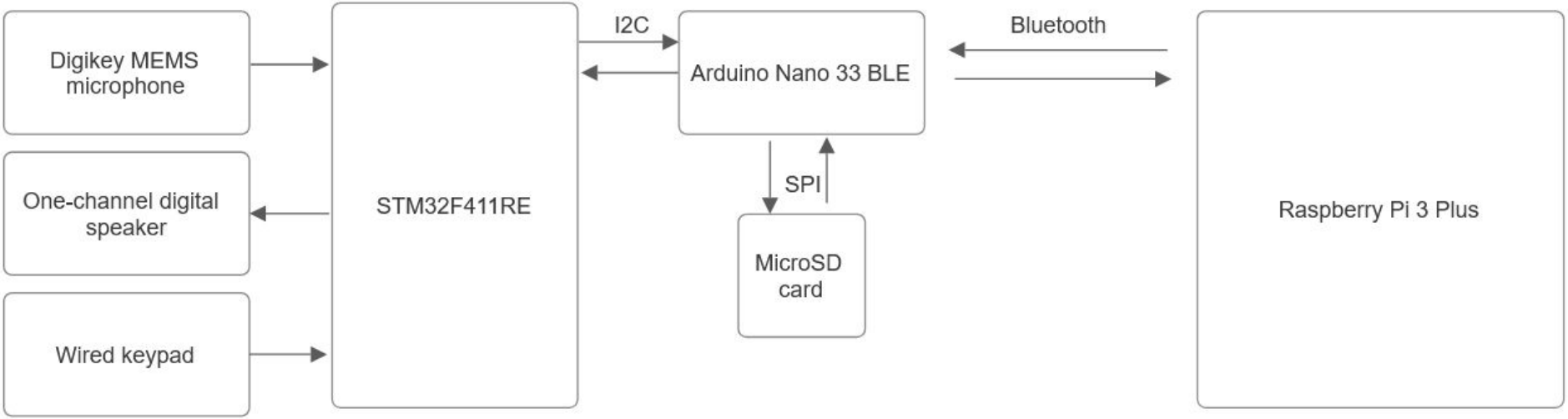
Communicates with hub device



Plays sound → “Grass is green.”

Embedded LLM creates answer

Block Diagram



Device Control

- Buttons
  - Security
    - Sequence input prevents unauthorized use
    - Correct sequence = ready to start
  - Reset and Interrupt
    - Sets processor to sleep, disabling peripherals
    - Press again to wake up back to initial state
  - Microphone Input
    - Hold sound button down to get input
    - Releasing saves audio file

Audio Recording

- Timer-driven one-channel ADC using DMA, non-circular mode. 12 bit samples @ 44 KHz

- Uses conversion half/complete callbacks to save samples

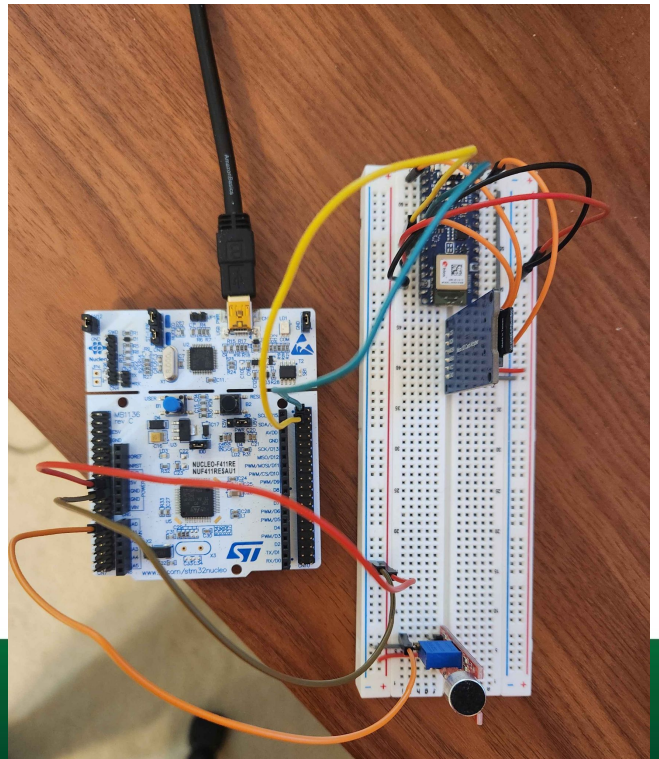
- Create WAV file on SD card: write file header, then data blocks of size (½*buffer length)

Data Storage

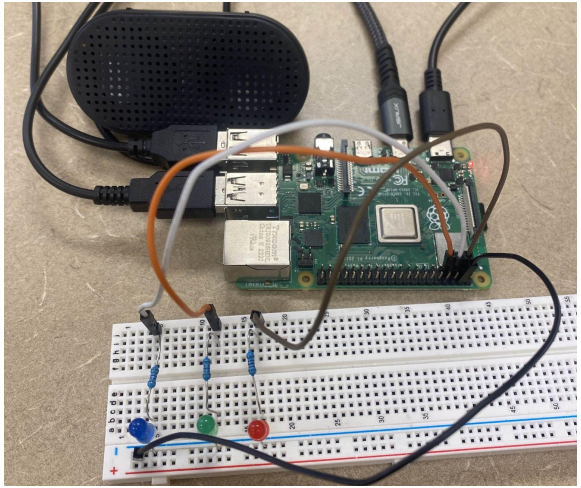
- Audio samples are stored on a FAT32 SD card driven by an Arduino Nano

- Samples are sent by 100KHz I2C (Master: STM, Slave: Nano)

- When the buffer is filled, write to the SD card using SPI



Wireless



- Arduino uses BLE to broadcast a service
- Raspberry Pi listens to the Bluetooth network and connects to Arduino
- Arduino then sends chunks of data from the .wav file on its SD card to the Raspberry Pi
- Raspberry Pi then listens to the data and saves it locally on it’s drive

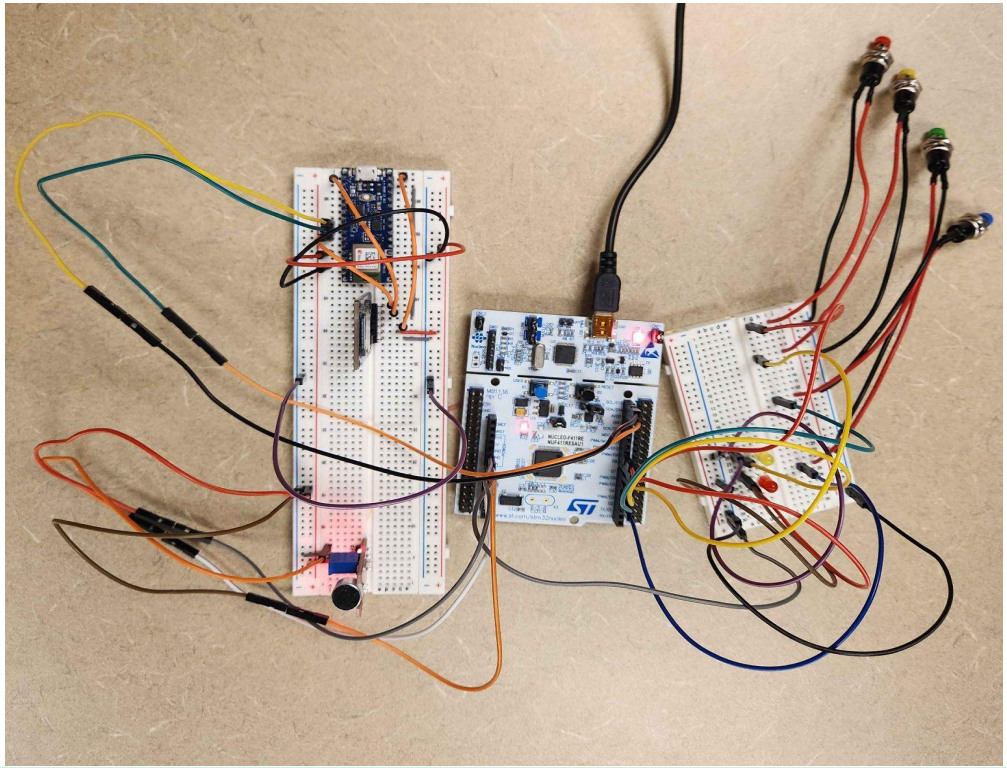
AI Generation

- Google Speech Recognition AI translates voice to text

- Locally ran LLM generates a response to the question

- Raspberry Pi then outputs the result using text to speech

Prototype



Remaining Work

- Additional protocols for robustness against signal drop

- Fine-tuning language model

- Record final demo

- Switch between onboard Pi speaker and secondary speaker