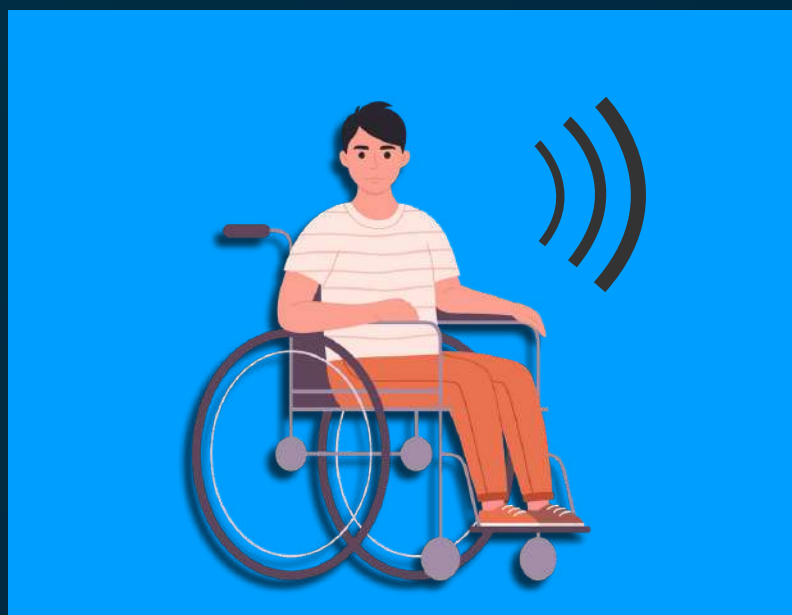


# PROYECTO VOICELOGIC ROBOLOGIC

Sofía Estrada Hernández - A01666608  
Emma Sofia García Montealegre - A01659535  
Christian Damar Marín Ramírez - A01659334  
Carolina Pérez Valencia - A01665909

# PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA



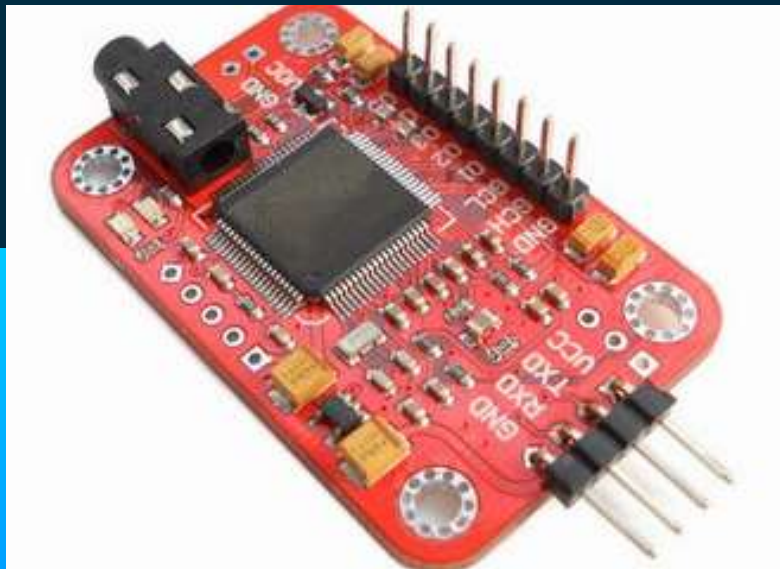
- El acceso a los hogares aún presenta barreras para personas con discapacidad visual o dificultades motrices.
- Sistema de seguridad activado por reconocimiento de voz.

- Dispositivo más dinámico, atractivo, fácil de usar y accesible.

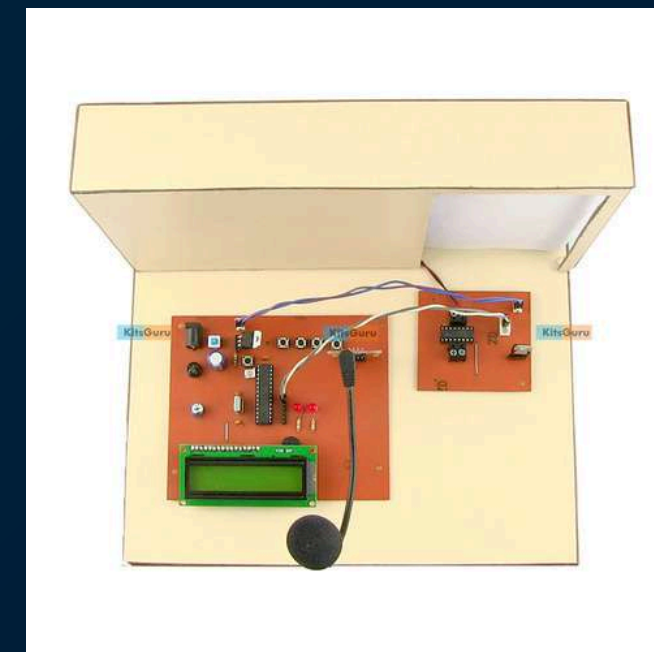


# ANTECEDENTES

## Proyectos Similares



- Candado controlado por voz implementado en el microcontrolador de 16 bits Sungyang SPCE061A
- Sistema de candado controlado por contraseña de voz





# ANTECEDENTES



## Ventajas

- Facilitar acceso
- Mayor seguridad
- Conexión a diversos sistemas inteligentes

## Desventajas

- Costo elevado comparado con cerraduras por pin, huella digital o tarjeta de proximidad.
- Configuración más complicada
- Poca disponibilidad



Candado de grabación de audio y reproducción MP3 (8 GB de memoria)

**93,75 EUR**



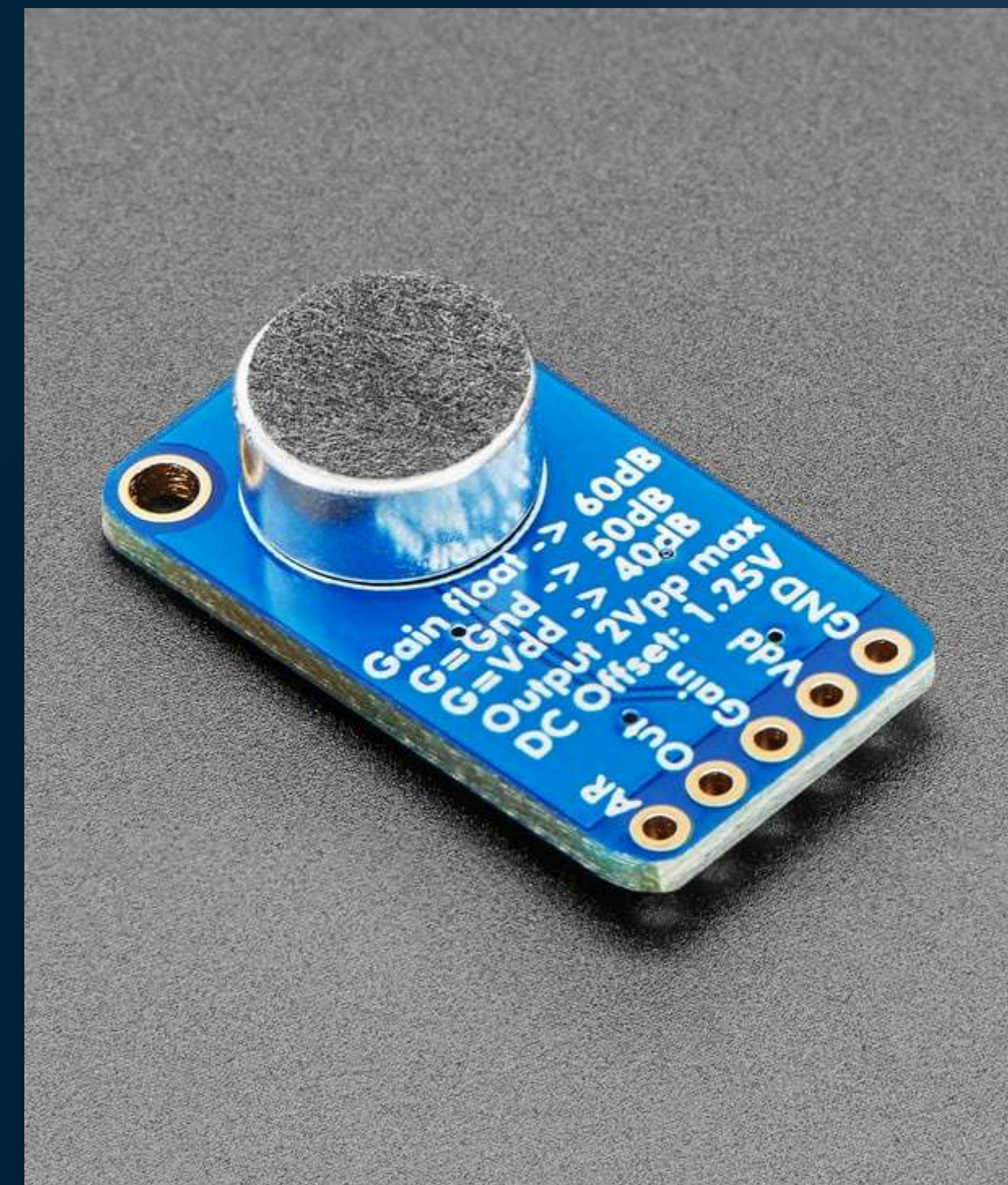
# OBJETIVOS

**Diseñar e implementar un sistema de reconocimiento de voz basado en el microcontrolador ATmega16 con el fin de implementarlo en sistemas de seguridad accesibles para personas con dificultades de movilidad y visibilidad, asegurando que el costo total del prototipo no exceda los \$700.00.**



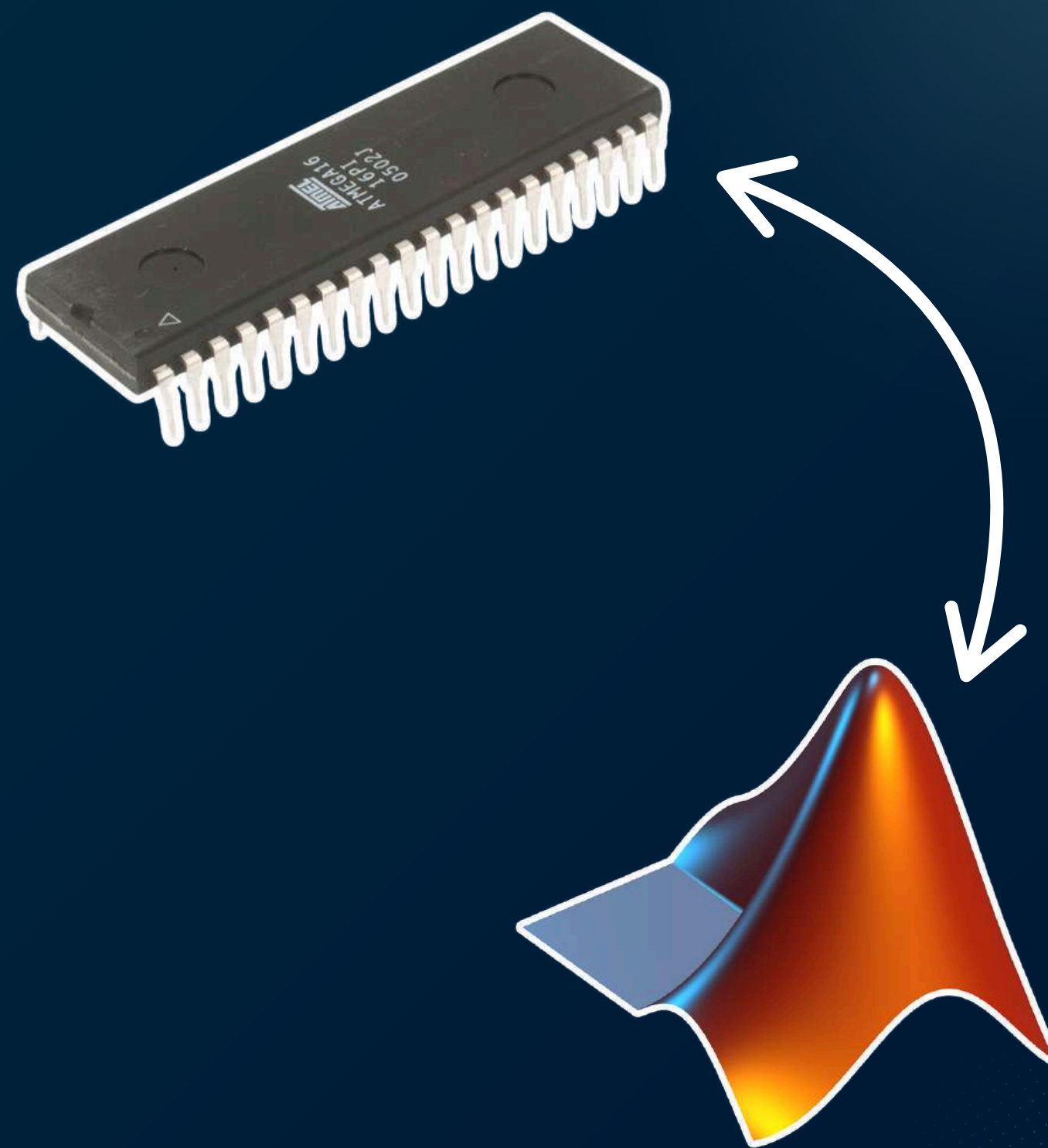
# OBJETIVO SECUNDARIO #1

Incluir un módulo de entrada de audio, utilizando un micrófono integrado en el sistema, y lograr una captura y envío de voz de manera simultanea.



## OBJETIVO SECUNDARIO #2

Generar un algoritmo que con 70 audios por usuario, en total 280 audios, genere un modelo predictivo, y detecte similitudes entre el audio entrante con aquellos en la base de datos.





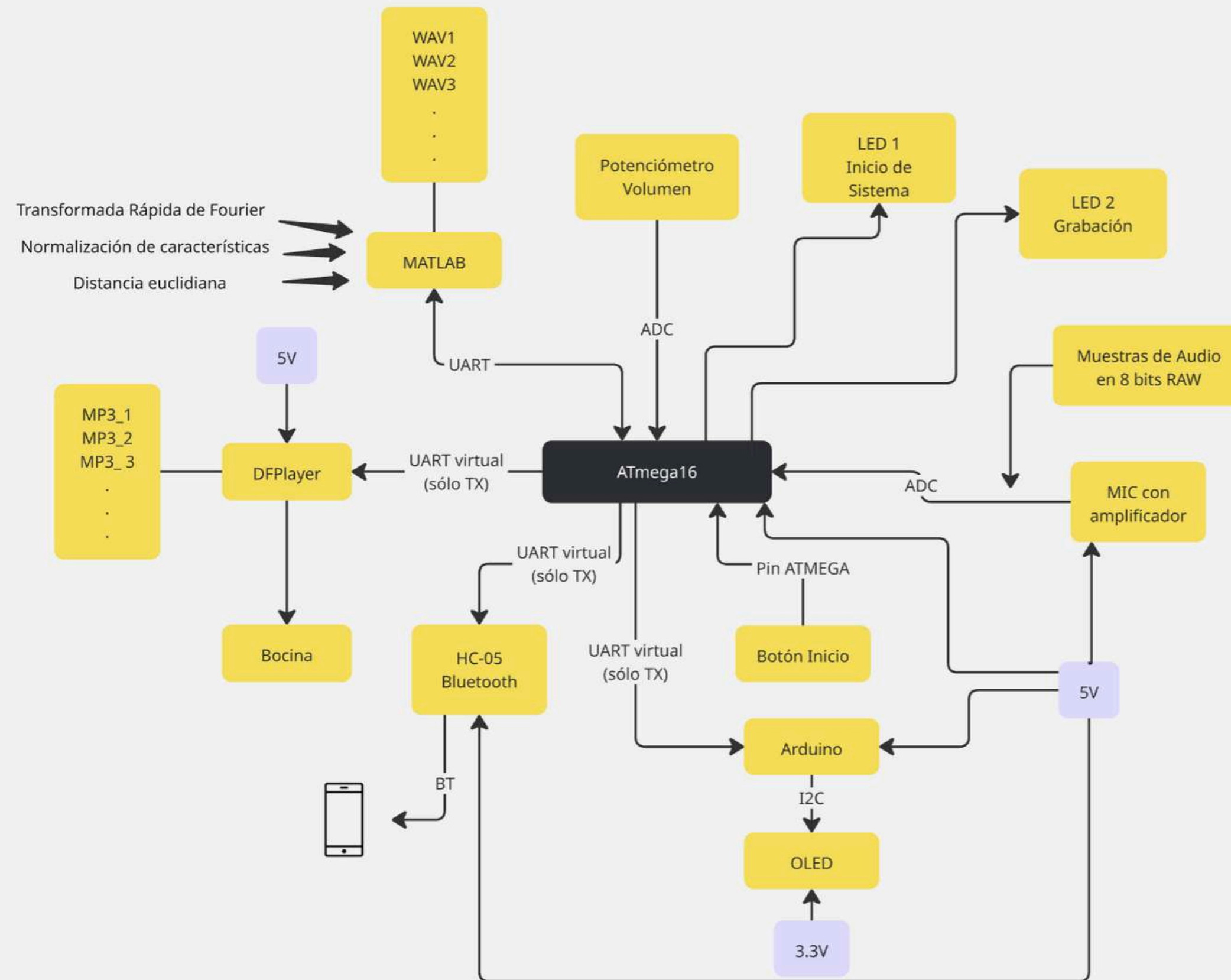
## OBJETIVO SECUNDARIO #3

Desarrollar un sistema de reproducción de audio que permita la salida del mismo por medio de la bocina integrada dentro del sistema y por una aplicación móvil. Con el fin de acercarnos al usuario y tener otra forma de verificar con usuarios con dificultad visual.

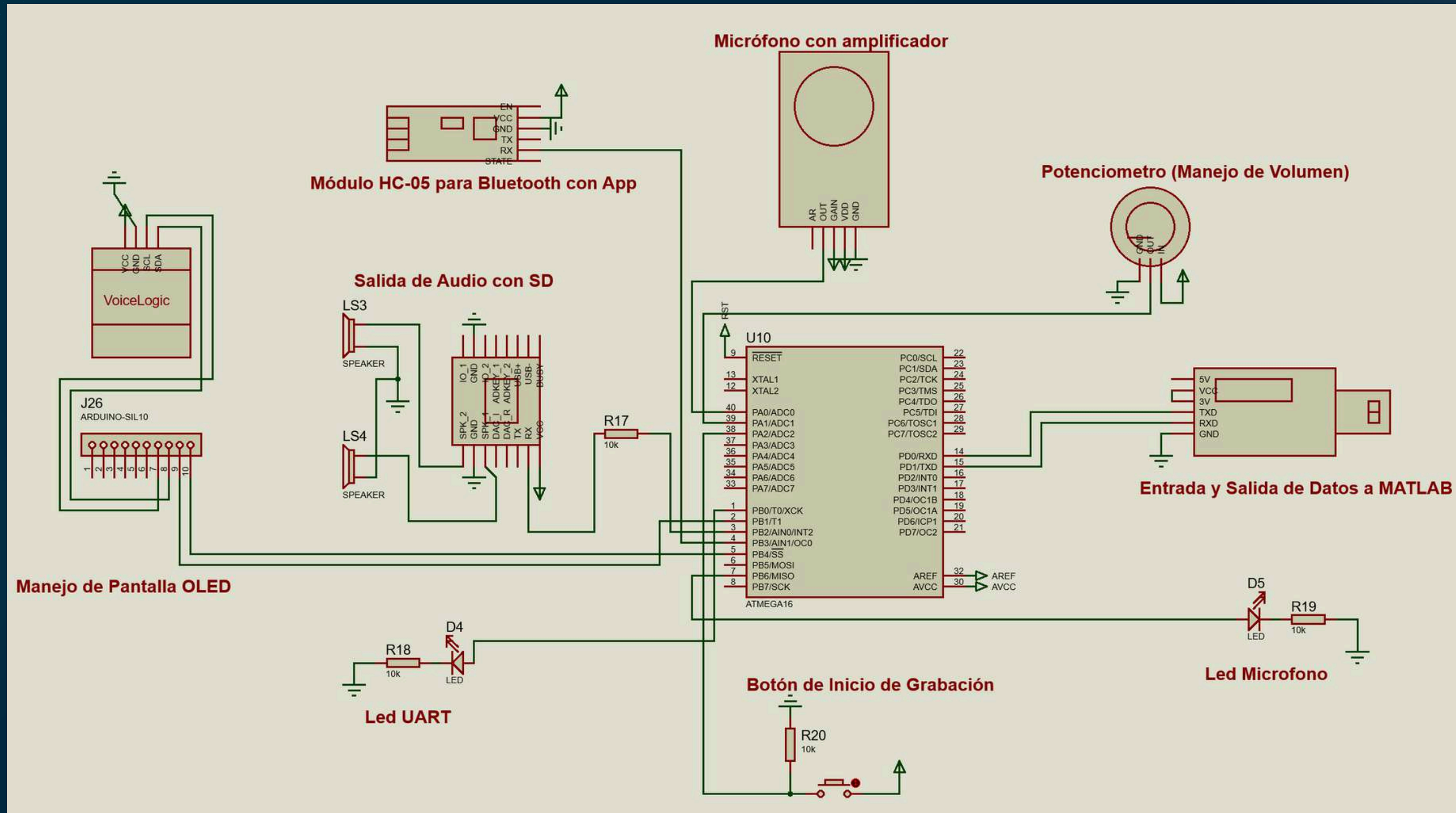




# DIAGRAMA GENERAL



# ESQUEMÁTICO

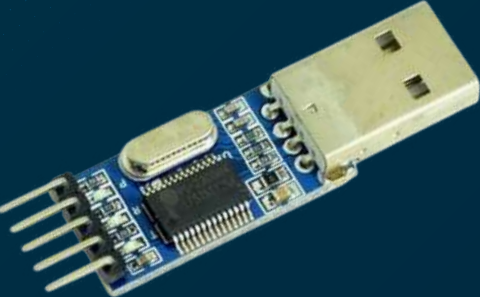
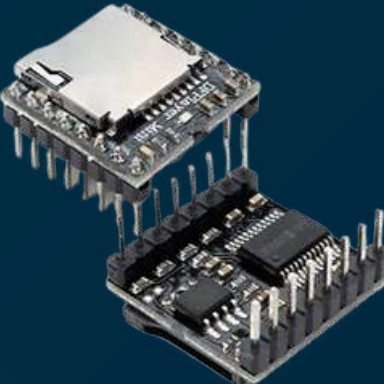
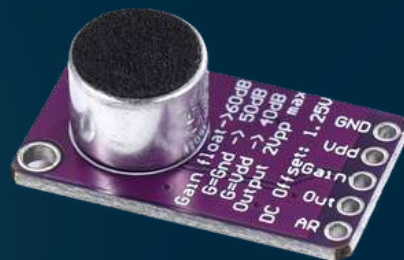






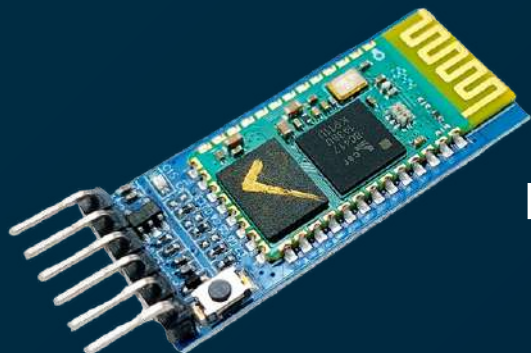


# ESQUEMÁTICO

Pin físico	Nombre lógico	Dirección	Conectado a...	Función
PA0	ADC0	Entrada	Micrófono con amplificador	ADC – Señal de voz
PA1	ADC1	Entrada	Potenciómetro	ADC – Volumen
PB5	PB2	Salida	DFPlayer	UART virtual – TX a DFPlayer
Arduino	SDA / SCL	Salida	Arduino (Pantalla OLED)	UART virtual – TX a Arduino
PB3	RX	Salida	HC-05 Bluetooth	UART virtual – TX a Bluetooth
PB1	T1	Salida	Arduino	Señal de grabación activa
PB0	T0	Salida	LED UART	Indicador de transmisión
PD1	TXD	Salida	Módulo UART USB (MATLAB)	UART – TX hacia MATLAB
PD0	RXD	Entrada	Módulo UART USB (MATLAB)	UART – RX desde MATLAB

# SENSORES Y ACTUADORES

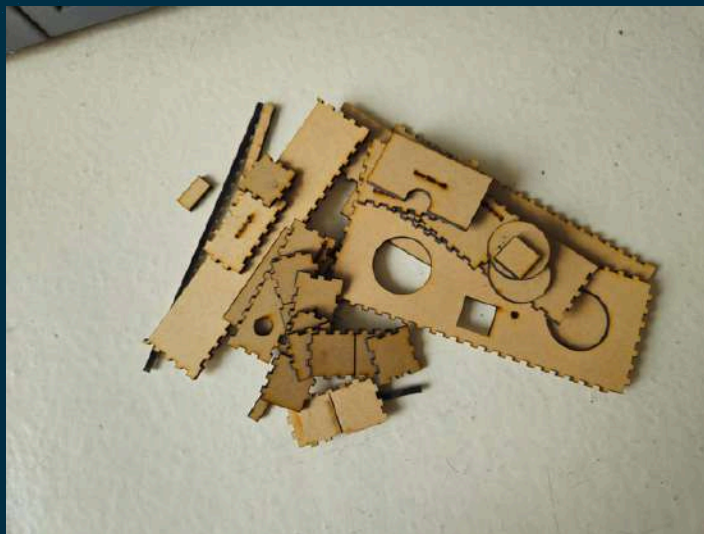
Sensores y Actuadores	Precio	¿Por qué se eligió?
 <p>Zx-022</p>	\$159.00	Se utiliza para el mandado de los datos de voz por segmentos de muestras a MATLAB.
 <p>DFPlayer</p>	\$ 60.00	Obtiene del ATMEGA16 los identificadores de cada usuario y así poder reproducir el mensaje correspondiente al reconocido o no reconocido.
 <p>MAX9814</p>	\$ 79.00	Se utiliza para la captura de audio del usuario, este audio ya amplificado se manda al ATMEGA16, que regulará el mandado de muestras en lotes para su procesamiento.
 <p>Pantalla OLED</p>	\$ 50.00	En ella se muestra la pantalla de inicio y los indicadores visuales al usuario para indicar que está grabando y cuál fue registrado.



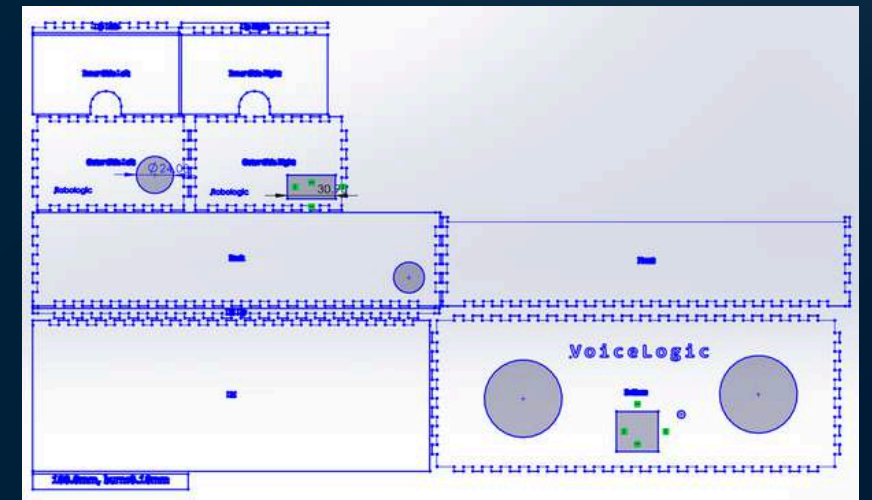
Sensores y Actuadores	Precio	¿Por qué se eligió?
 <p>Arduino</p>	\$200.00	Obtiene un comando de control de el ATMEGA16 para indicar el cambio entre las animaciones de las pantallas.
 <p>Bluetooth</p>	\$ 60.00	Se utiliza para mandar los mismos identificadores que llegan a DFPlayer por Bluetooth y además poder visualizar el usuario y su canción identificadora en la aplicación móvil.
 <p>Potenciómetro</p>	\$ 12.00	Se utiliza para poder regular el audio por ADC.
 <p>Bocina</p>	2 de \$ 30.00	Se utiliza como una de las salidas de audio del sistema.

# SENSORES Y ACTUADORES

# MATERIALES Y CONSIDERACIONES DEL PROTOTIPO FÍSICO



- **Exterior el sistema:**
  - Se diseñó una caja por medio de corte láser de madera



- **Alimentación:**
  - Eliminador 29 W a 5 V



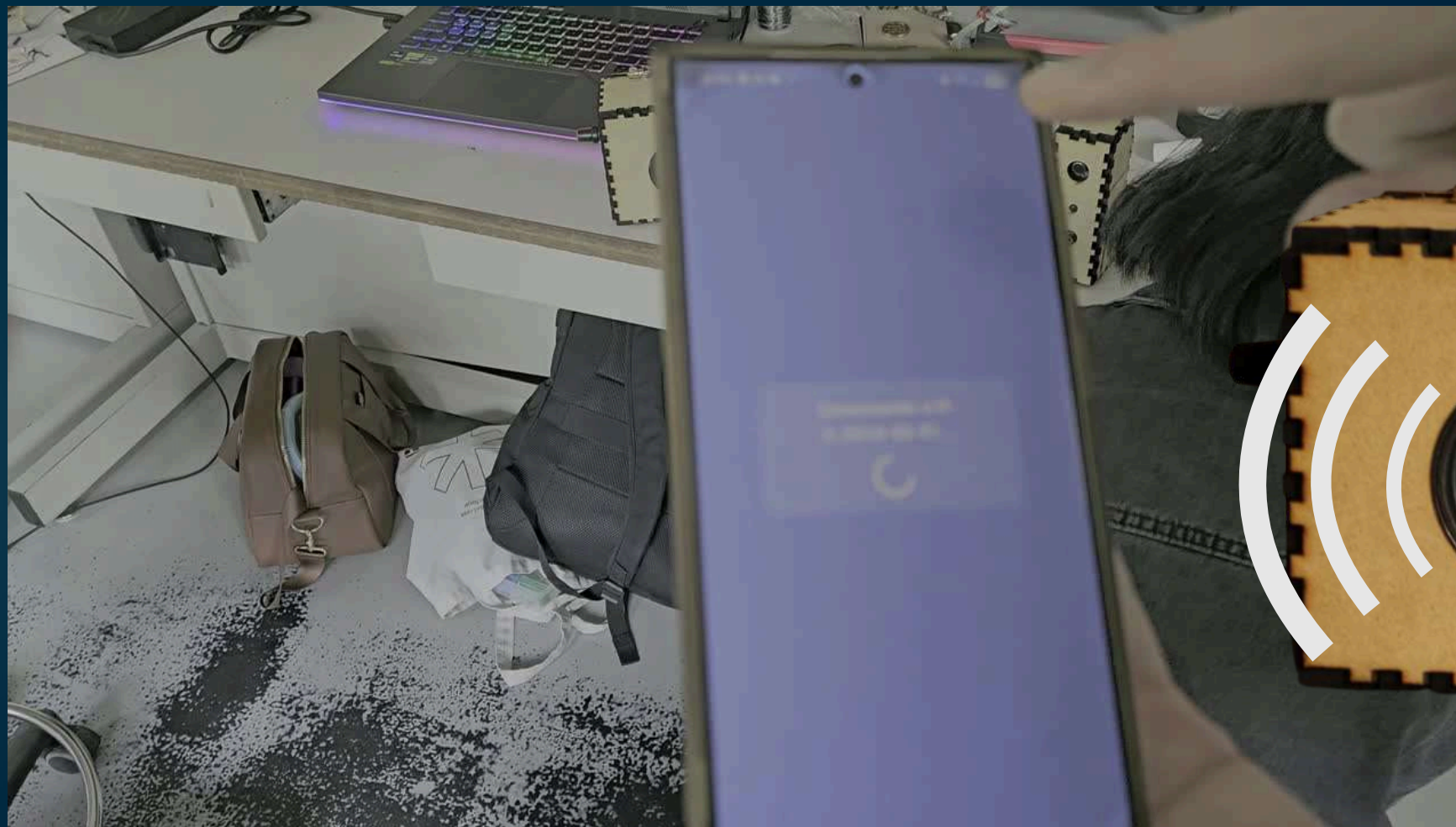
- **Hardware externo:**
  - Laptop para correr MatLab y celular Android para descargar nuestra aplicación.





# RESULTADOS

El sistema resultante se observa a continuación:



# MODO DE USO

1. Conectar sistema a la alimentación y a laptop con MatLab



2. Abrir app móvil



3. Seleccionar módulo Bluetooth en lista de dispositivos



voicelogic





# IMODO DE USO

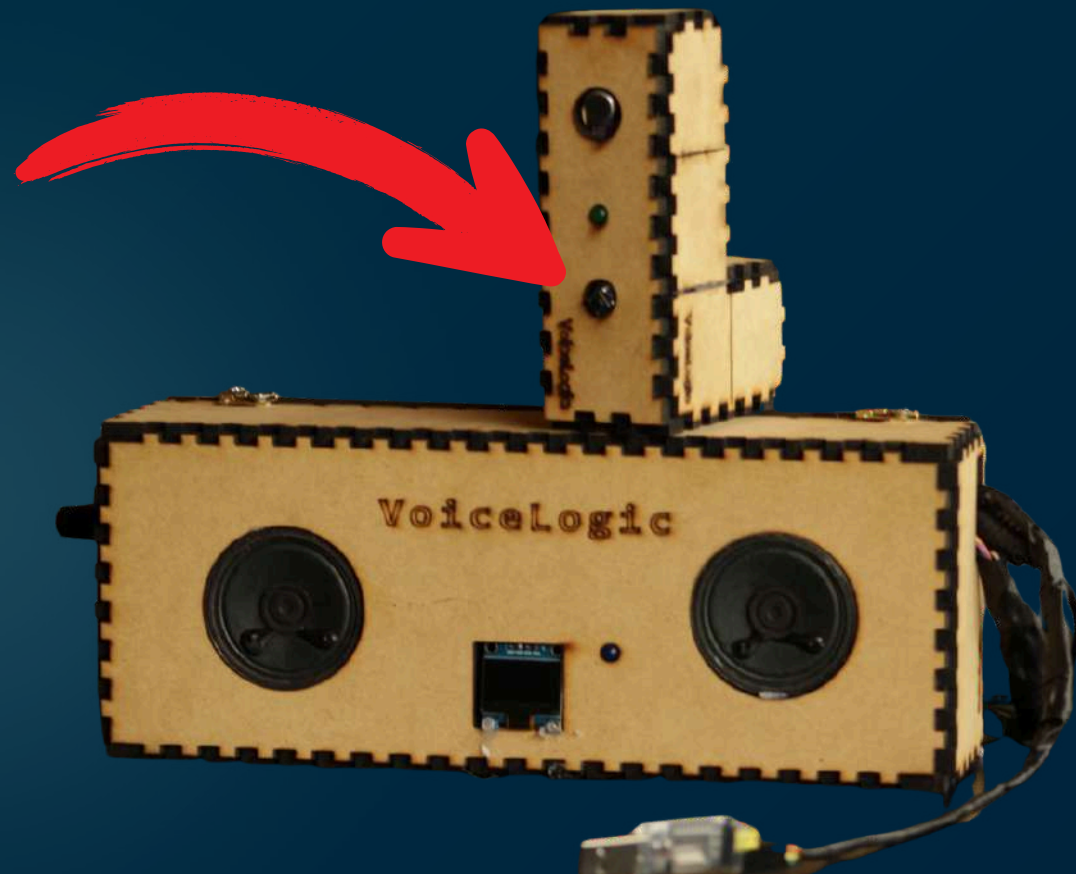
4. Presionar botón debajo del micrófono.



5. Mantener presionado y hablar hasta que se hayan recolectado 8000 muestras en MatLab. (aprox 30 seg)

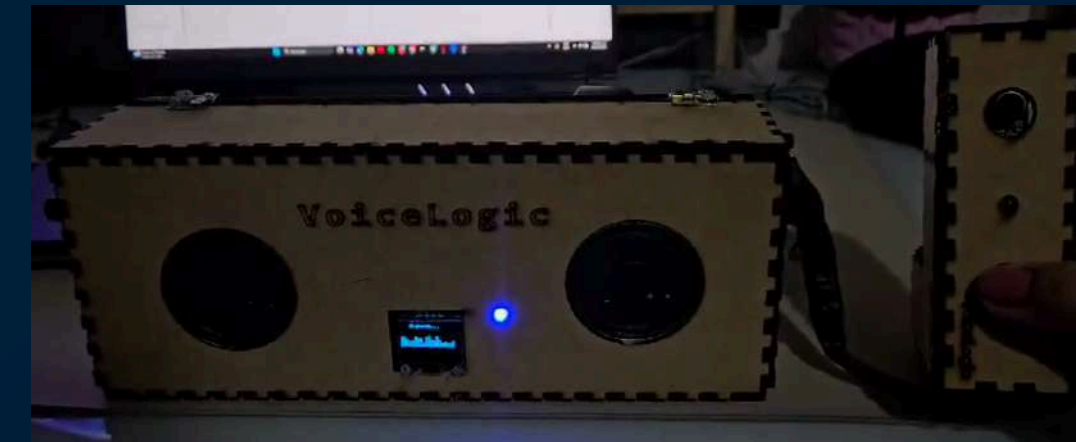
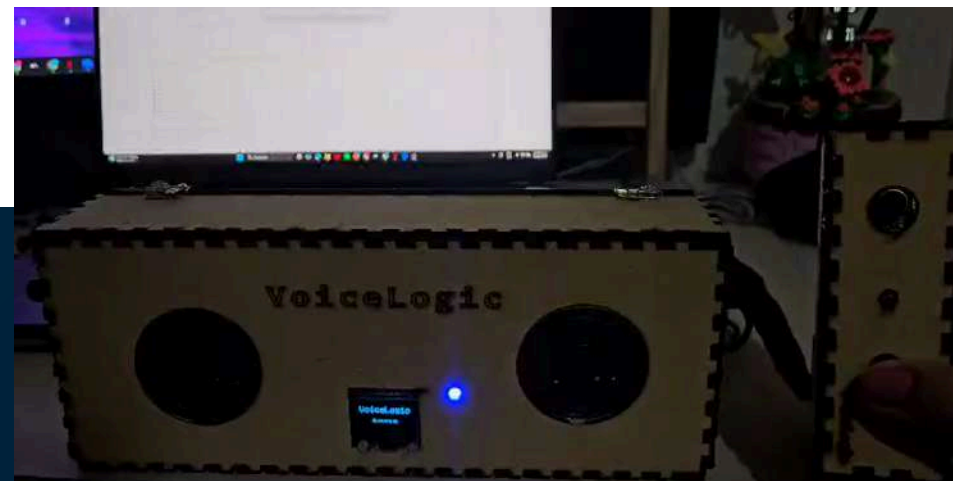


6. Acceso otorgado / denegado



```

Esperando botón...
Capturando audio...
Muestras leídas: 1000
Muestras leídas: 2000
Muestras leídas: 3000
Muestras leídas: 4000
Muestras leídas: 5000
Muestras
Muestras
Muestras
    
```

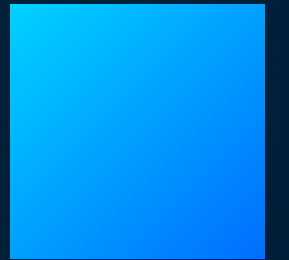


```

X Voz no reconocida.
Índice 5 enviado al microcontrolador
Esperando botón...
    
```

# PROCESAMIENTO DE AUDIO

---



## Extracción de características: .

- Energía, Varianza, Media absoluta, Frecuencia dominante, Curtosis, Entropía espectral, ZCR

## Clasificación (ECOC)

- Se usa el algoritmo ECOC (Error-Correcting Output Codes), para un modelo multiclase con la función fitcecoc, para la clasificación de voces de múltiples usuarios a partir de las características extraídas del audio.

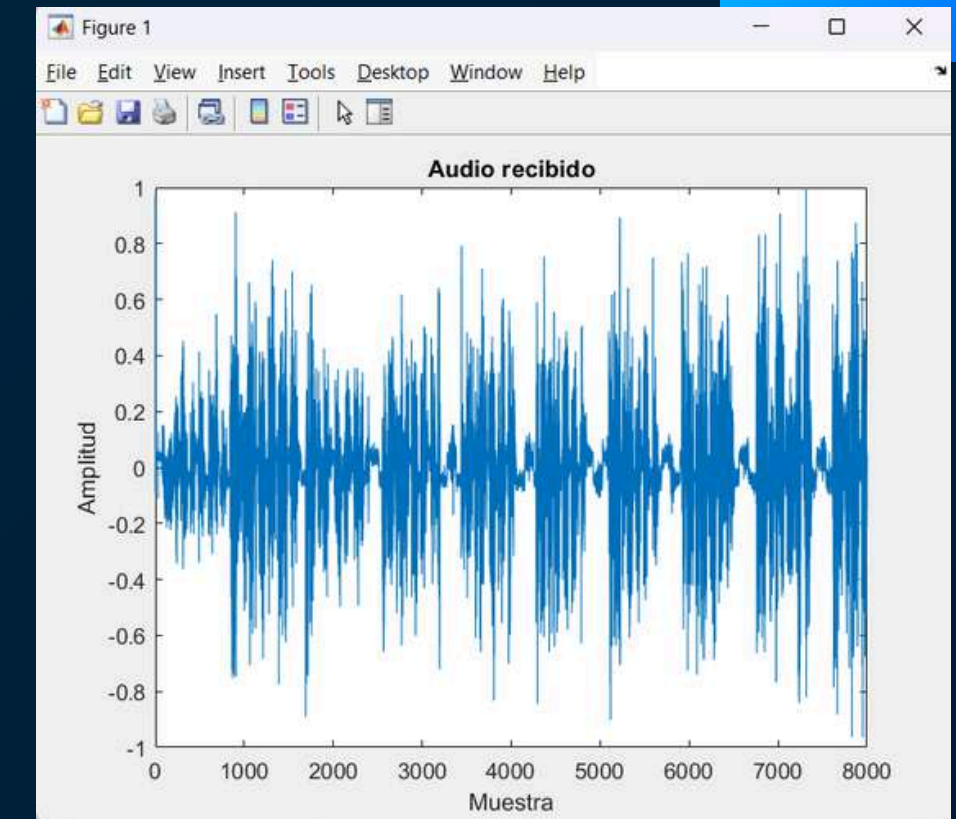
## Comparación Euclidiana

- Se normaliza el audio recibido y se compara con los vectores de usuarios registrados, calculando la distancia euclidiana. Se identifica el usuario más cercano si la distancia promedio está por debajo de un umbral definido manualmente.



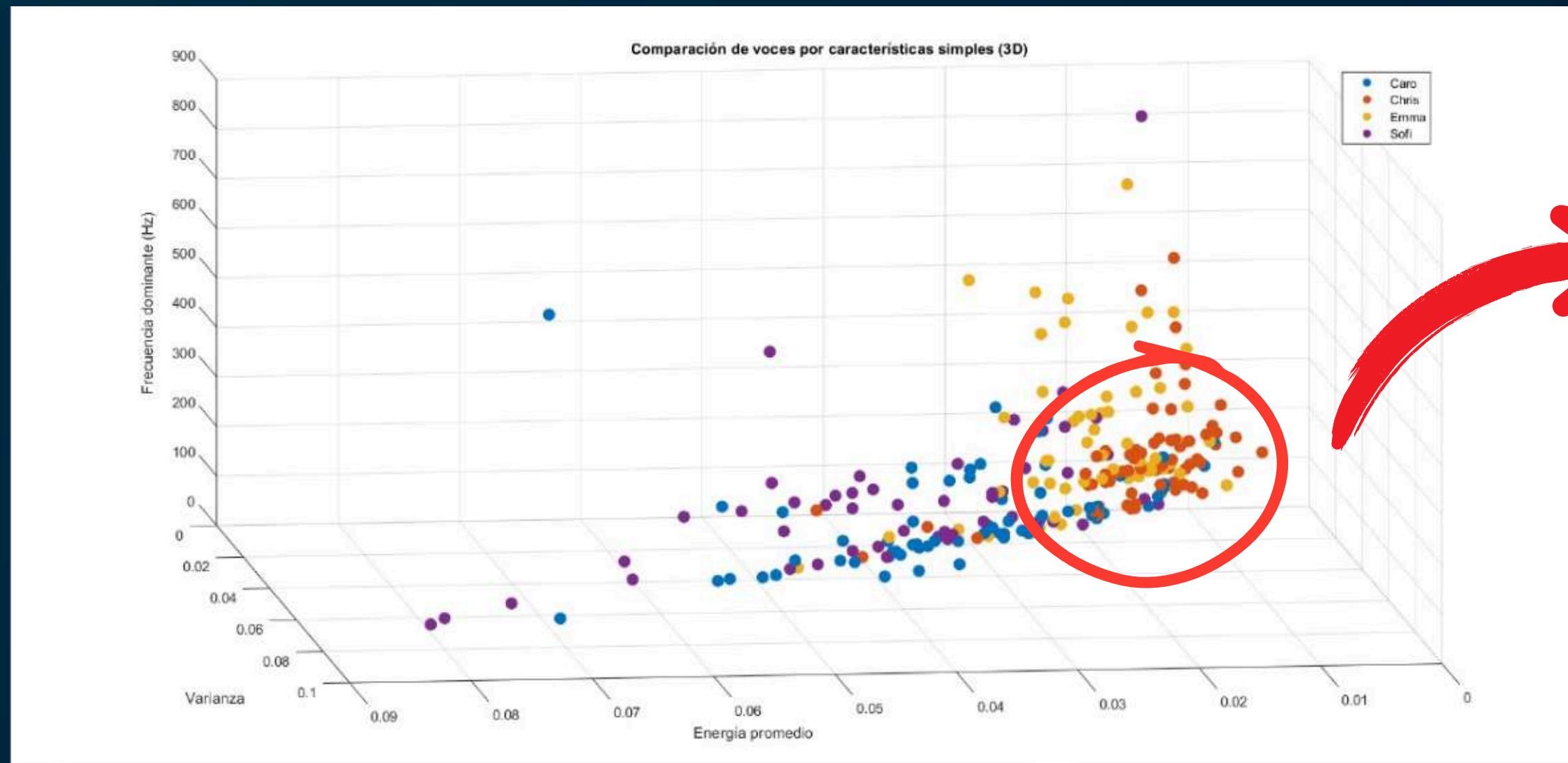
# RESULTADOS

- Se logró crear el sistema de reconocimineto de voz con una precisión de 70% en un promedio de 70 pruebas de cada individuo del equipo.
- El muestreo tomó un tiempo promedio de 20 segundos (8000 muestras equivalente a 5 segundos de audio).
- No se logró el costo de prototipaje de \$700.00, el prototipo costó aproximadamente \$900.00.



# ANÁLISIS DE RESULTADOS

- El margen de error se vio principalmente en voces con tono y volumen similares.



Mayor  
similitud

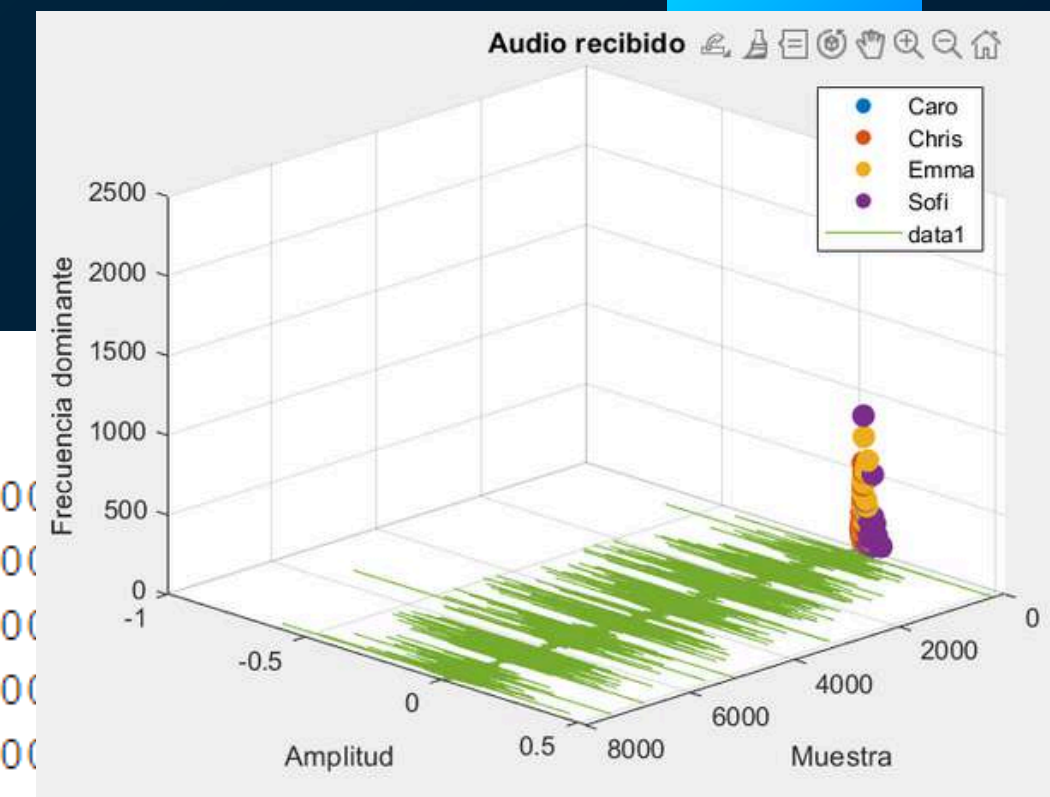


# ANÁLISIS DE RESULTADOS

El tiempo requerido en la serialización y transmisión de los datos de audio se vio reflejado en el muestreo lento.

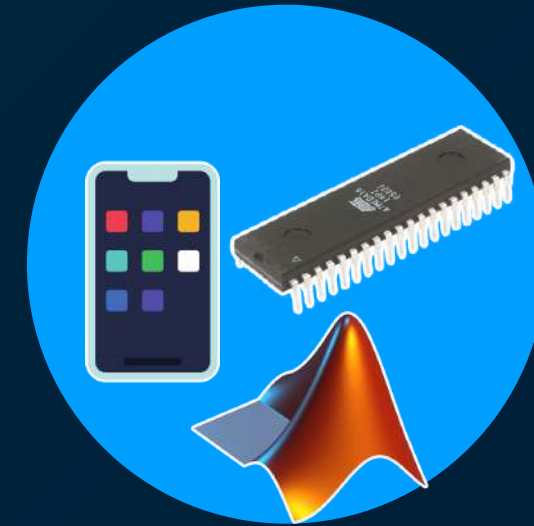
```

Esperando botón...
Capturando audio...
Muestras leídas: 1000
Muestras leídas: 2000
Muestras leídas: 3000
Muestras leídas: 4000
Muestras leídas: 5000
Muestras leídas: 6000
Muestras leídas: 7000
Muestras leídas: 8000
Guardado: grabacion_223058.wav
Distancia promedio a Caro: 9.8369
Distancia promedio a Chris: 9.1891
Distancia promedio a Emma: 8.2221
Distancia promedio a Sofi: 10.0808
X Voz no reconocida.
Índice 5 enviado al microcontrolador.
Esperando botón...
    
```



# CONCLUSIONES

- Se logró el funcionamiento completo, logrando comunicación entre ATmega16, MatLab y la app móvil
- Se logró reconocer correctamente la voz del usuario en la mayoría de los casos.
- El sistema presenta áreas de oportunidad en términos de velocidad y portabilidad.





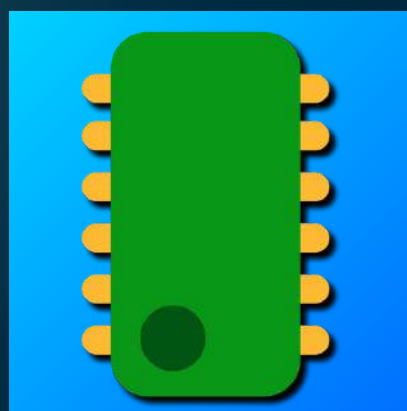
# TRABAJO A FUTURO



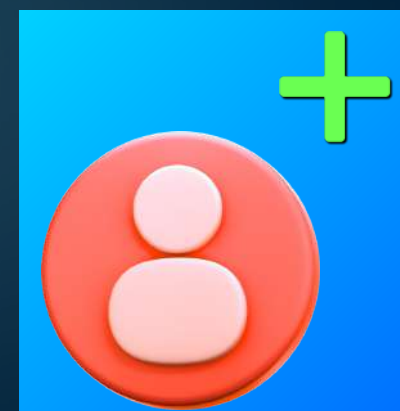
- Micrófono con mayor calidad de grabación



- Acortar tiempo de muestreo

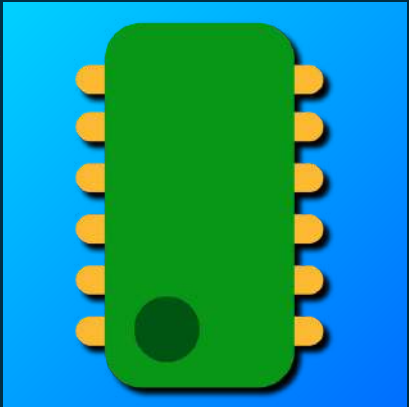


- Reemplazar laptop con microcontrolador y volverlo completamente embebido. Además de aplicar context switching.



- Ingresar nuevos usuarios desde aplicación

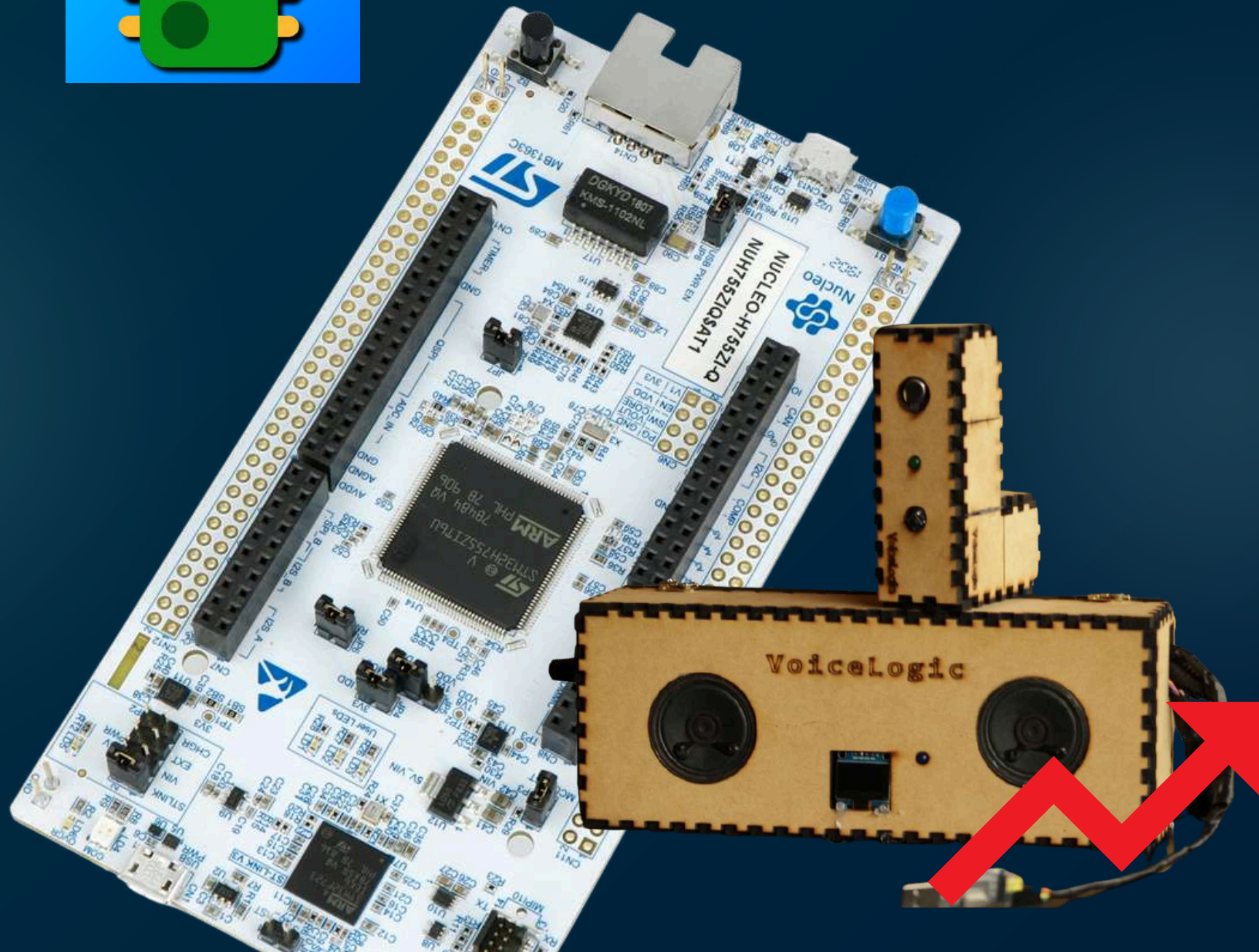
# TRABAJO A FUTURO



- Usar STM32H745 NUCLEO para lograr un sistema embebido al 100%. No depender de externos.

## Mejoras:

- Capacidad de cargar modelos directamente:
  - Normalización
  - FFT con CMSIS DSP
  - Extracción de características (energía, varianza, ZCR, etc.)
  - Clasificación local (con SVM, KNN o TinyML)
- Reproducción y visualización más fluida (Múltiples UART nativos)
- Mejores capturas de audio. ( 12 o 16 bits)
- Almacenar buffers de audio grandes. (RAM 1MB)





# REFERENCIAS

---

- Lock Keychain Audio Recorder and Mp3 Player with 8 GB Integrated Memory. (s. f.). SpyShopEurope.com.  
<https://www.spyshopeurope.com/en/lock-keychain-audio-recorder-and-mp3-player-with-8-gb-integrated-memory/875/177>
- DIY Advanced Voice Password Based Door lock System @KitsGuru. (2021, 13 noviembre). Recuperado de  
<https://kitsguru.com/products/advanced-voice-password-based-door-lock-system>
- Voice electronic door lock system based on 16-bit microcontroller. (s. f.). Recuperado de  
<https://en.eeworld.com.cn/news/mcu/eic13394.html>

