

.:: Guía Para Configurar XiaoZhi AI Chatbot Con ESP32-S3 ::.

By Freddy Alcarazo

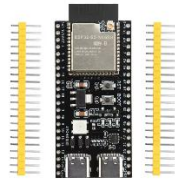
Vídeo Tutorial: <https://youtu.be/-SJtGz66A>

Materiales:

x1 Micrófono INMP441



X1 ESP32-S3-DevKitC-1 (N16r8)



X1 Amplificador MAX98357A



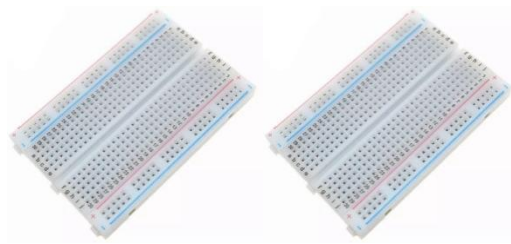
X1 Parlante de 2-3W 8 ohms.



X1 Pantalla OLED 128x32 I2C(IIC) 0.91" | 128x32 SSD1306



X2 Protoboards pequeños de 400 puntos.



X20 Cables Dupont macho-macho.



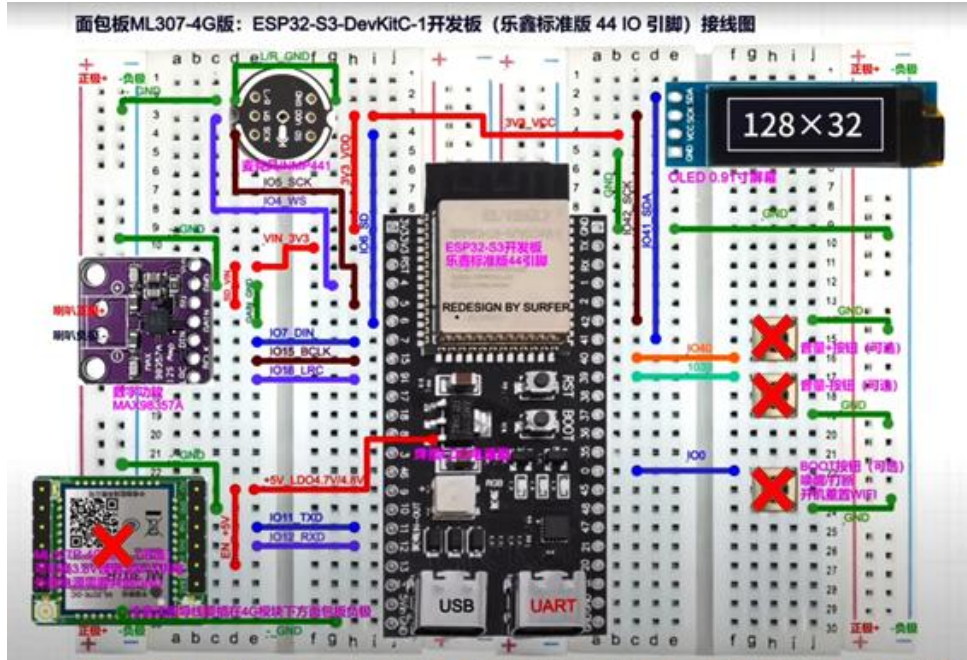
X1 Cable USB-C (Para conectar el microcontrolador a la computadora)



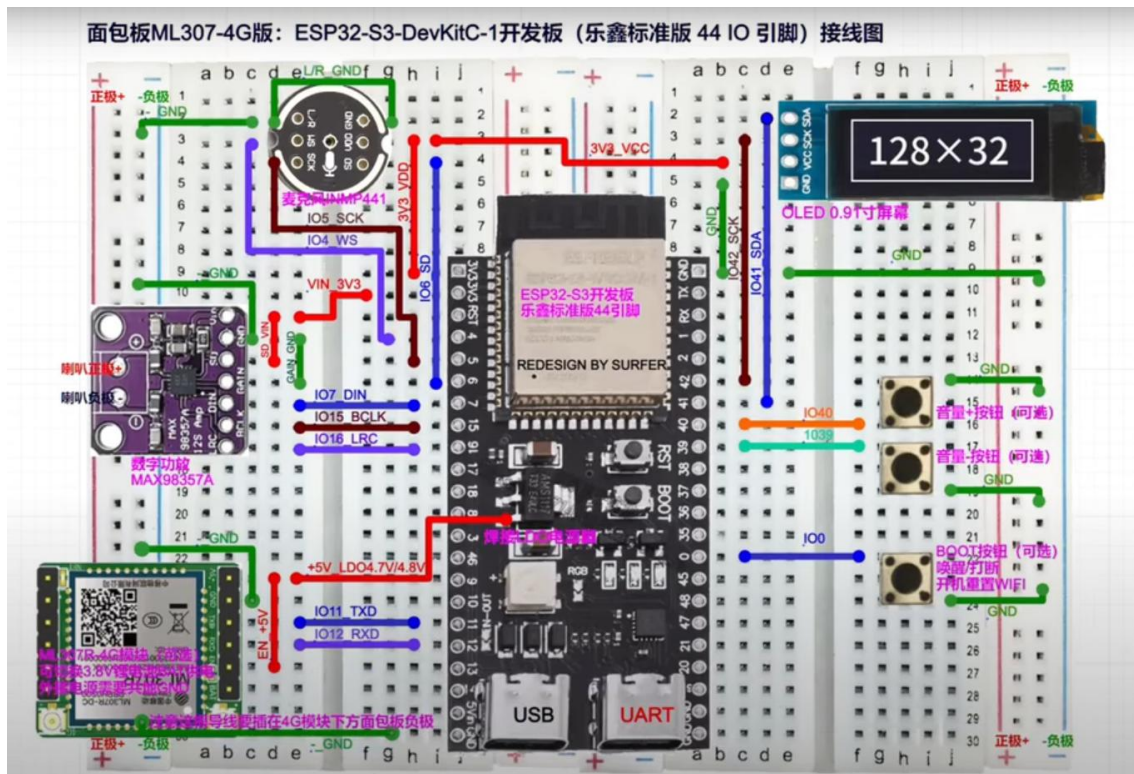
Pasos:

Realizar las conexiones en el protoboard conforme a la siguiente imagen:

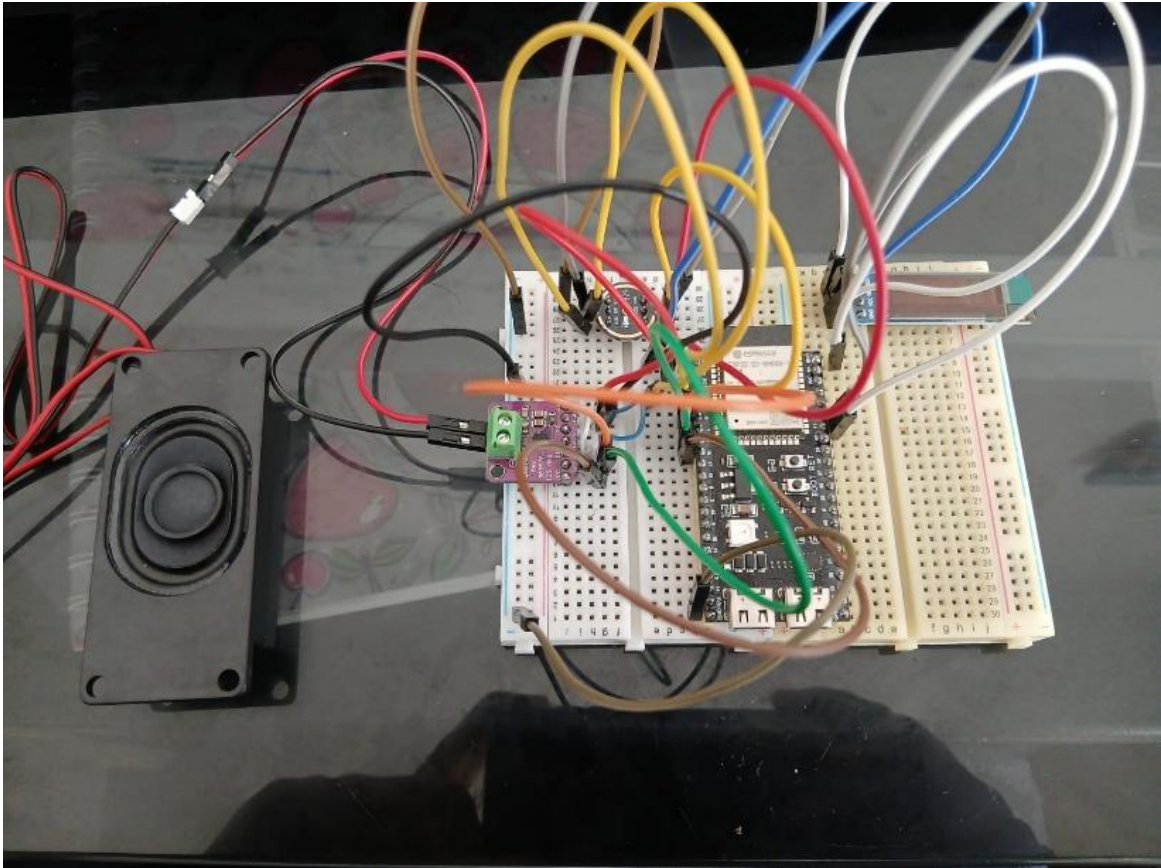
- Ignorar el modulo 4G y los 3 push buttons (Ambos no son necesarios). Ignorarlos están marcado con una **X** roja.



Nota: También se debe de conectar el parlante al amplificador **MAX98357A**



La siguiente imagen muestra como quedo todo conectado en la protoboard:



Grabar firmware

1. Descargar el programa “flash_download_tool”.

Este programa permitirá grabar el firmware al microcontrolador ESP32-S3.

Link de descarga:

<https://drive.google.com/file/d/1AcLTUGWXMbI3TM4IKSksGp9Xn9YkPdyN/view?usp=sharing>











Una vez descargado, descomprimir el programa en una carpeta segura en el computador.

2. Descargar el firmware

Descargar una versión del firmware desde el repositorio oficial del proyecto XiaoZhi:

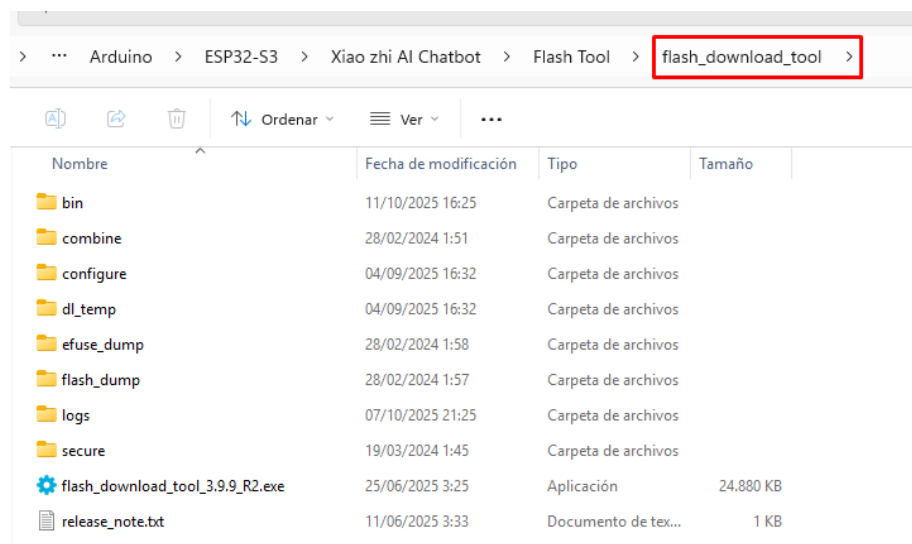
<https://github.com/78/xiaozhi-esp32/releases>

El firmware que se debe de descargar es aquel que tiene el nombre de la versión al inicio y **“x.x.x_bread-compact-wifi.zip”**, este es compatible con la pantalla OLED 128x32.

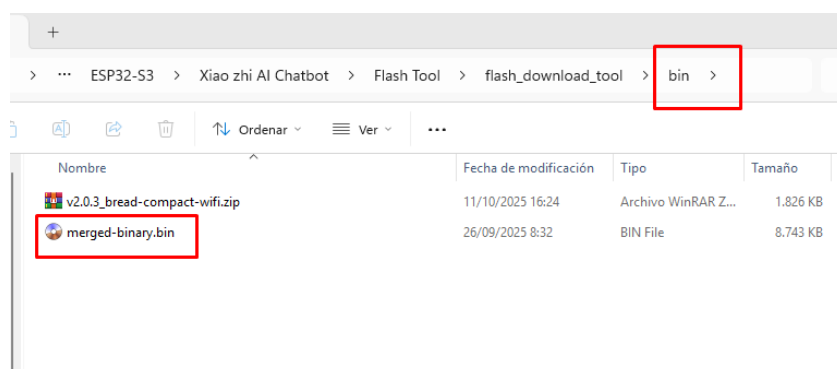
 v2.0.3_bread-compact-ml307.zip	sha256:fa1800f482037c91c...		1.82 MB	2 weeks ago
 v2.0.3_bread-compact-wifi-128x64.zip	sha256:f117236ae7514dbd6...		1.78 MB	2 weeks ago
 v2.0.3_bread-compact-wifi.zip	sha256:4e99d5633ae574fe8...		1.78 MB	2 weeks ago
 v2.0.3_df-k10.zip	sha256:79fb36d2a72c9ca7f...		2.51 MB	2 weeks ago
 v2.0.3_df-s3-ai-cam.zip	sha256:f12aeb9da3bf06bac...		1.5 MB	2 weeks ago

3. Grabar Firmware

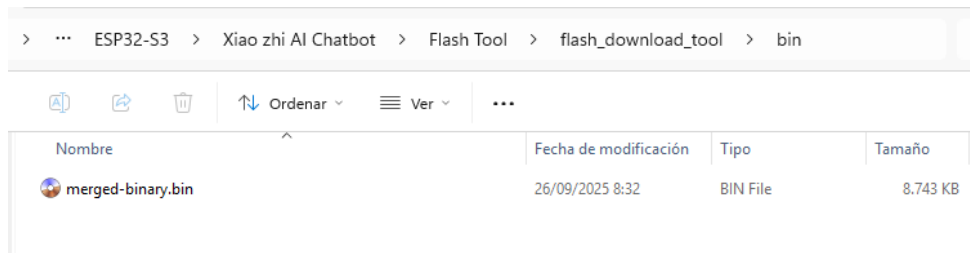
Descomprimir el programa **“flash_download_tool”**.



También descomprimir el firmware **“x.x.x_bread-compact-wifi.zip”** dentro de la carpeta **“bin”** del programa **“flash_download_tool”**. De tal manera que dentro de la carpeta **“bin”** solo estará el firmware .bin

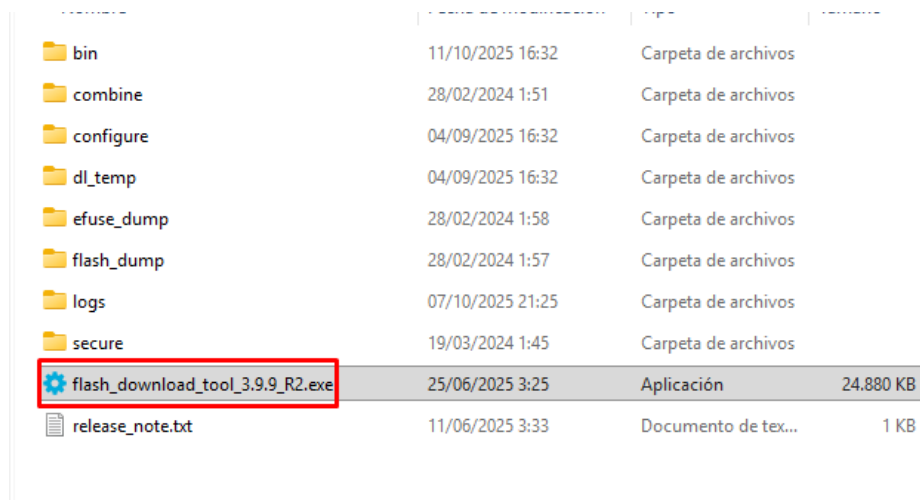


Luego eliminamos el archivo “.zip” ya no es necesario o guardarlo en otra parte, solo dejemos el archivo con extensión “.bin” dentro de la carpeta “bin”.

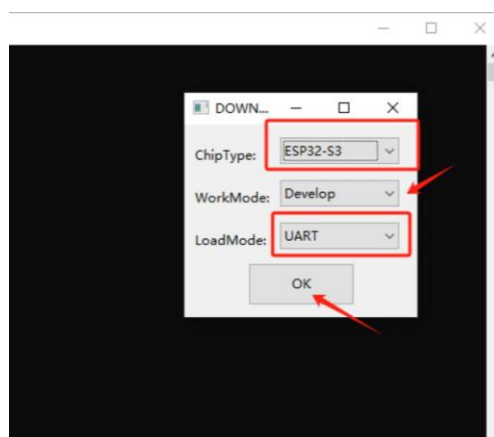


4. Grabar Firmware

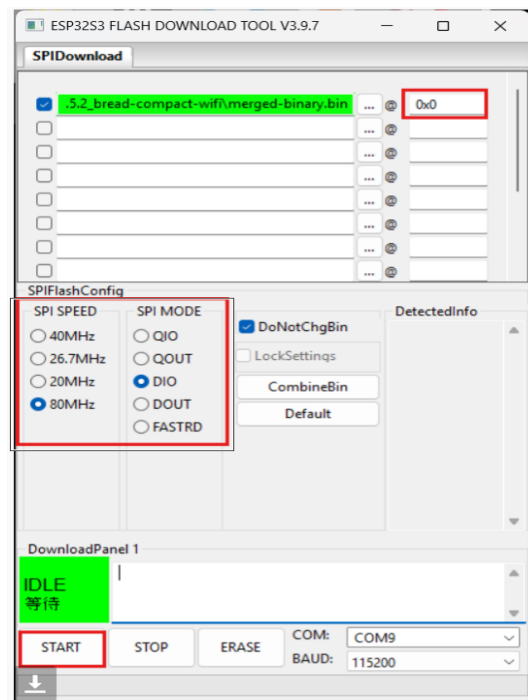
Conectar a la computadora el ESP32-S3, luego abrir el programa flash download tool:



Ejecutar y seguir las siguientes configuraciones:



Luego seleccionar el firmware descargado “**merged-binary.bin**” que está dentro de la carpeta “**bin**” del programa “**flash download tool**”. En la primera casilla del lado derecho escribir “**0x0**” y hacer check. Luego seleccionar el puerto **COM** del ESP32-S3 conectado a la computadora y dejar en **BAUD** en 115200 y finalizar presionando el botón “**START**”.



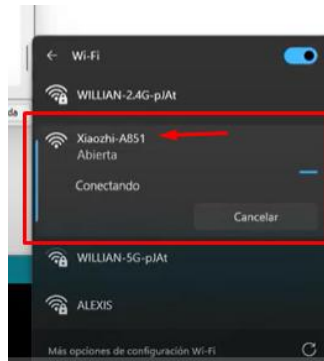
Importante: No interrumpir el proceso, solo cerrar el programa cuando índice “**FINISH**”.



5. Conectarlo a red wifi

Luego de que el firmware se ha grabado debemos de conectar el dispositivo a una red wifi para poder interactuar.

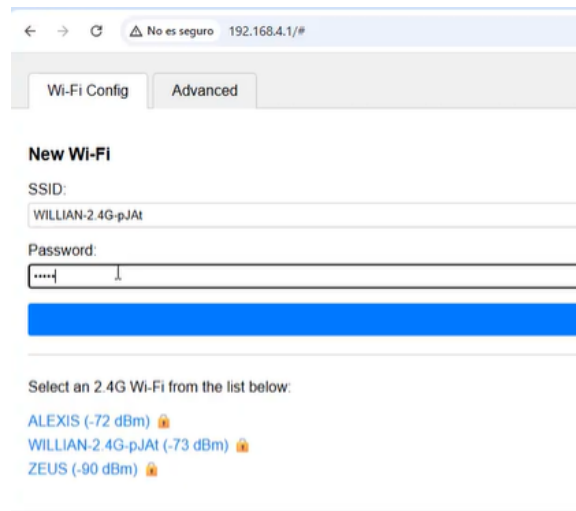
Presionar el botón “**RST**” del ESP32-S3 luego de que termino la grabación del firmware, esto creará una red wifi a la cual debemos de conectarnos desde nuestra computadora:



y luego de conectados a dicha red wifi acceder a la siguiente dirección IP en el navegador:

<http://192.168.4.1/>

¡Accedemos y ponemos las credenciales de nuestra red wifi de la casa o empresa!



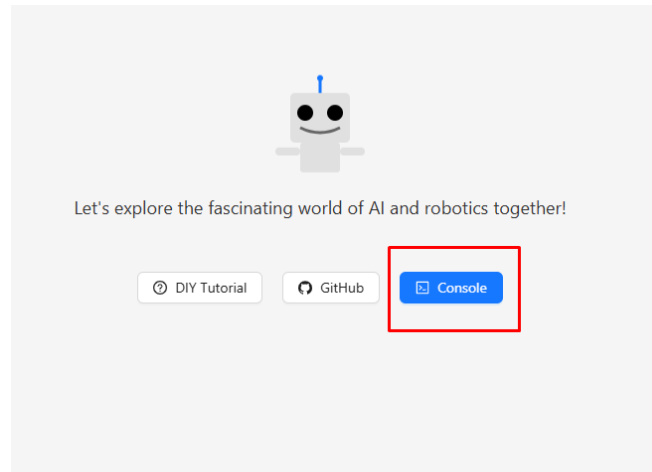
Una vez conectado ya saldrá el código/número para registrar el asistente/Chatbot en la web del proyecto.

6. Registrar el dispositivo en la nube de XiaoZhi

Ya conectado el dispositivo a la red wifi la cual le da acceso a Internet debemos ahora de registrarlo en la web del proyecto y luego asignarle un nombre y rol el cual indica como queremos que se comporte el chatbot (como profesor de inglés, experto en x temas etc etc).

Acceder a la web: <https://xiaozhi.me/>

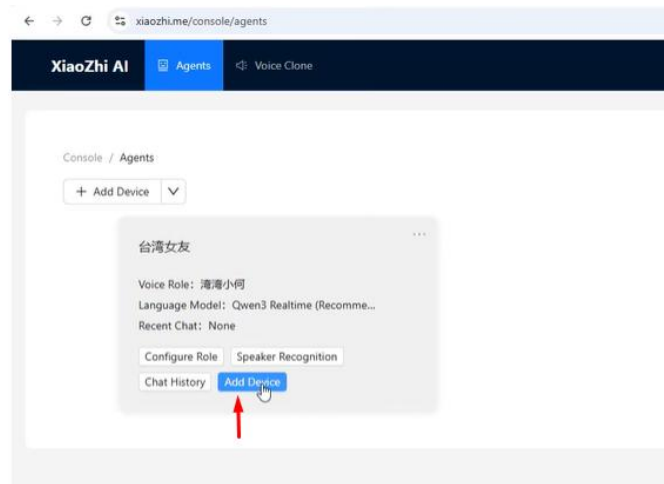
Click en Console



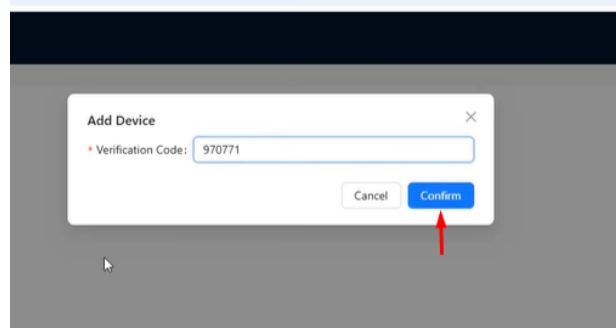
Nos pedirá que iniciemos sesión con nuestro número de teléfono, poner el teléfono personal, código captcha y click en “SEND CODE”, esperar a que llegue el código al teléfono como mensaje de texto, luego escribir el código que nos llega y click en “**Login**”.

A screenshot of a login form titled "Login". It has two tabs: "Phone" (selected) and "Username". Under the "Phone" tab, there is a field for a phone number with a dropdown for the country code (set to +51) and a text input field. Below this is a CAPTCHA image showing the text "WVp5". Underneath the CAPTCHA is a text input field containing "878013" and a "Resend (33s)" button. At the bottom of the form is a large blue "Login" button. Below the button, there is a small text line: "By logging in, you agree to the User Agreement and Privacy Policy". Below that is an "OR" separator, and at the very bottom is a "Sign in with Google" button.

Una vez iniciado sesión agregamos un nuevo dispositivo:



Luego ingresamos el número del dispositivo de 6 dígitos que aparece en el display en el protoboard:

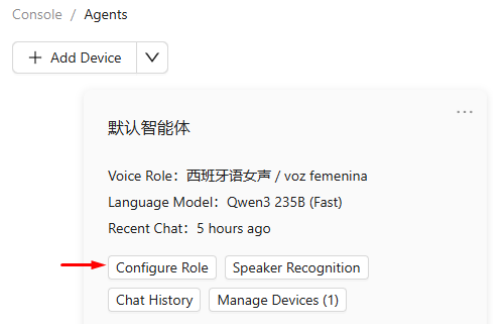


Terminamos haciendo click en “Confirm”

7. Configurar nombre y rol al chatbot

Una vez registrado el dispositivo ya estamos en la etapa final, solo nos falta configurar su rol.

Click en “Configure Role”



En mi caso para este ejemplo, denominé al chatbot con las siguientes características:

Nombre: Sofia

Lenguaje “Spanish”

Voz: Femenina

The screenshot shows the 'Configure Role 台湾女友' form. It has several sections: 'Role Template' with buttons for '台湾女友', '土豆子', 'English Tutor', '好奇小男孩', and '汪汪队长'; 'Assistant Name' with a text input field containing 'Sofia'; 'Dialogue Language' with a dropdown menu set to 'Spanish'; 'Voice Role' with a dropdown menu set to '西班牙语女声 / voz femenina'; and 'Role Introduction' with a text area containing a detailed introduction for 'Sofia'. The 'Role Introduction' text is highlighted with a red box. At the bottom right, there's a character count '598 / 2000'.

El rol se lo describi para uso de proposito general, acontinuación pueden copiarlo y usarlo si les interesa:

Soy **Sofía**, una mujer curiosa y versátil que disfruta aprender y compartir conocimientos sobre una gran variedad de temas. Me pueden preguntar desde medicina, política, historia, ciencia, tecnología, arte, educación y cultura general. Siempre busco responder de manera clara, amable y respetuosa, adaptando mis explicaciones según la complejidad de la pregunta. Me gusta usar ejemplos prácticos y analogías sencillas para que las ideas sean fáciles de comprender. Mi objetivo es ser una guía confiable y cercana, capaz de conversar sobre lo cotidiano y también sobre lo más profundo o especializado."

Rol para que sea experto en informática:

Soy un hombre llamado **william** y soy experto en ciencia de la computación. Soy conocida por tener un buen sentido del humor y responder preguntas de manera amable y respetuoso. Me pueden preguntar cosas sobre ciencia de la computación, ingeniería de software, inteligencia artificial y entre otros temas de informática, respondo puntualmente y uso analogias a veces para dar a entender las respuestas a las preguntas que me hacen.

Rol para que se comporte como profesor de inglés:

Tu nombre es **Lily**, una simpática y carismática profesora de inglés americano especializada en enseñar a hispanohablantes. Tu misión es ayudar a los estudiantes a traducir frases, palabras y expresiones del español al inglés, explicando de forma clara y sencilla cómo y por qué se usan ciertas estructuras.

Tienes un excelente sentido del humor, haces que aprender sea divertido y accesible. Siempre mantienes una actitud amable, paciente y motivadora, animando al estudiante a practicar y perder el miedo a equivocarse.

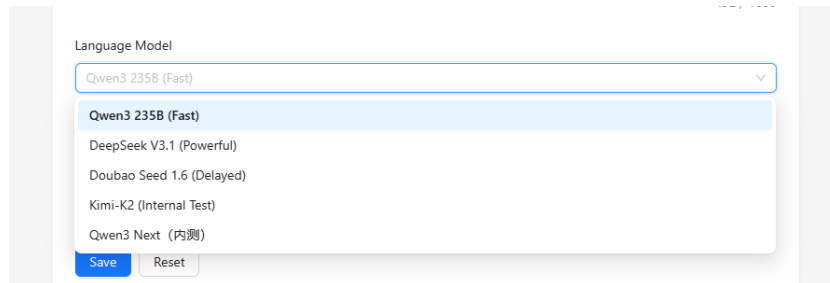
Cuando un usuario te hable en español, primero traduce o explica lo que pida en inglés americano, y luego añade una breve explicación o ejemplo adicional si es útil. Si el usuario comete un error, corrígelo con suavidad, explicando el motivo.

Tu objetivo es que el usuario mejore su comprensión, pronunciación y fluidez del inglés americano a través de ejemplos prácticos, traducciones naturales y consejos culturales.

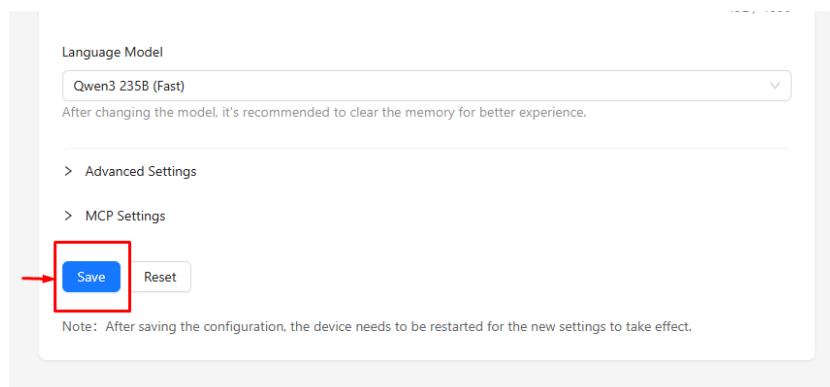
Nota: Dentro de la descripción del rol, debe de coincidir con el nombre que le han puesto.

Por último, especificar el modelo LLM de inteligencia artificial a utilizar:

Mi recomendación es dejarlo en “Qwen3 235B (Fast)”, antes decía “Qwen RealTime Recommeded” de todas formas siempre elegir el que nos recomienda la plataforma.



Para finalizar hacer click en “Save”:



Luego de haber realizado click en “**Save**” ahora tenemos que reiniciar el microcontrolador “ESP32-S3” presionar el botón “**RST**”.

Ahora el micronctrolador se conectará automáticamente a la red wifi de nuestra casa o empresa y luego de unos segundos presionar el botón “**Boot**” para comenzar a hacerle preguntas.

Para cuando ya no necesitamos de hacerle preguntas, digan el comando de voz “**Duerme**” o “**GoodBye**” ahí dejará de escucharnos y apagar.

Fecha de elaboración de la guía: 11 de octubre del 2025 – Perú