

## ¿Me Oyes?

El reto inicia con un código QR. Al escanearlo, se redirige a un enlace de Vocaroo, donde se encuentra un archivo de audio. A simple escucha, el contenido no es comprensible, lo que sugiere que el mensaje está en otro idioma.



Al descargar el audio y reproducirlo, se identifica que el idioma es japonés, por lo que no es posible obtener la información directamente solo escuchándolo. Esto indica que es necesario transcribir o traducir el audio.

Para ello, se utiliza Whisper (ggml / whisper.cpp), una herramienta de reconocimiento de voz basada en IA.

<https://ggml.ai/whisper.cpp/>

whisper.cpp : WASM example

File Edit View Window Help

- Firefox cannot load files larger than 256 MB - use Chrome instead

More examples: [bench](#) | [stream](#) | [command](#) | [wchess](#) |

Model fetched: base.en

Input: ☒ File ☐ Microphone

Audio file:  Vocaroo 15...E6URW.mp3

Language:  Threads:  16

```
js: audio loaded, size: 500352
loadRemote: storage quota: 16881006476 bytes
loadRemote: storage usage: 6028 bytes
loadRemote: created IndexedDB whisper.ggerganov.com version 1
loadRemote: "https://huggingface.co/ggerganov/whisper.cpp/resolve/main/ggml-base.en.bin" is not in the IndexedDB
fetchRemote: downloading with fetch()...
fetchRemote: fetching 0% ...
fetchRemote: fetching 10% ...
fetchRemote: fetching 20% ...
fetchRemote: fetching 30% ...
fetchRemote: fetching 40% ...
fetchRemote: fetching 50% ...
fetchRemote: fetching 60% ...
fetchRemote: fetching 70% ...
fetchRemote: fetching 80% ...
fetchRemote: fetching 90% ...
fetchRemote: fetching 100% ...
loadRemote: "https://huggingface.co/ggerganov/whisper.cpp/resolve/main/ggml-base.en.bin" stored in the IndexedDB
storeFS: stored model: whisper.bin size: 147964211
```

| Build time: Wed Dec 17 13:20:22 2025 | Commit hash: [6c22e792](#) | Commit subject: talk-llama : sync llama.cpp | [Source Code](#) |

Luego, se carga el archivo de audio en Whisper, se selecciona un modelo ligero (por ejemplo, tiny), se define el idioma como Japanese y se ejecuta la opción Translate.

whisper.cpp : WASM example

File Edit View Window Help

- Firefox cannot load files larger than 256 MB - use Chrome instead

More examples: [bench](#) | [stream](#) | [command](#) | [wchess](#) |

---

Model loaded: tiny

Input: ☒ File ☐ Microphone

Audio file:  Vocaroo 15...E6URW.mp3

---

Language:  Threads:  8

```
system info: n_threads = 8 / 20 | whisper_cer : compile = 0 | ctranslate2 = 0 | cpu : WASM SIMD = 1 | fallback = 1 |
operator(): processing 500352 samples, 31.3 sec, 8 threads, 1 processors, lang = ja, task = translate ...

[00:00:00.000 --> 00:00:02.440] This is the final 3.
[00:00:02.440 --> 00:00:06.440] This is the first time I have a hard time.
[00:00:06.440 --> 00:00:08.440] Please remember this.
[00:00:08.440 --> 00:00:11.440] Frag Start F
[00:00:11.440 --> 00:00:13.080] I S
[00:00:13.080 --> 00:00:14.680] Hidari Kaku
[00:00:14.680 --> 00:00:18.920] L I M P I A T E
[00:00:18.920 --> 00:00:20.480] Andarbar
[00:00:20.480 --> 00:00:21.920] L A
[00:00:21.920 --> 00:00:23.480] Andarbar
[00:00:23.480 --> 00:00:26.480] O R E J A
[00:00:26.480 --> 00:00:27.960] Mini Kaku
[00:00:27.960 --> 00:00:29.800] Frag O R I
[00:00:29.800 --> 00:00:31.280] Thank you.

whisper print timings: load time = 108.48 ms
whisper print timings: fallbacks = 0 p / 0 h
whisper print timings: mel time = 85.31 ms
```

| Build time: Wed Dec 17 13:20:22 2025 | Commit hash: [6c22e792](#) | Commit subject: talk-llama : sync llama.cpp | [Source Code](#) |

En la transcripción aparece claramente una secuencia que sigue el formato de flag del CTF, la cual es dictada directamente en el audio.

FIS{LIMPIATE\_LA\_OREJA}