



MI4ALL: Manual de uso de la API

La API de MI4ALL pone a disposición del usuario las funciones necesarias para realizar las tareas de speech-to-text, machine translation y text-to-speech. Para ello se debe realizar una petición POST a la siguiente URL:

lexytrad.es/dev/mi4all_test/api/

A continuación, se muestran, para cada función, los datos que se deben enviar a la API, así como la respuesta (en formato JSON) que se recibirá en cada caso.

Función Speech-to-Text

Recibe un audio (en formato base64) y un idioma, y devuelve un texto con la transcripción.

```
datos = {  
    action: "stt",  
    audio: <String>,  
    lang: <String>,  
    tgt_lan: <String>,  
    stt_to_use: <String>  
}
```

action	Debe ser el string “stt”.
audio	String que representa un audio en Base64
lang	Código de idioma (es, en, ar).
tgt_lan	Código de idioma (es, en, ar). Solo se utiliza en el caso de modelos end-to-end.
stt_to_use	Modelo Speech-to-text que se va a utilizar. Puede tomar los siguientes valores: - <i>deepspeech</i> : Reconocimiento en inglés. Es el valor por defecto. - <i>nemo</i> : Reconoce los idiomas ES y EN. - <i>huggingface-asr</i> : Reconoce los idiomas ES, EN y AR. - <i>huggingface-s2t</i> : Este es un modelo end-to-end. Reconoce audio en inglés y lo puede traducir a los idiomas ES y AR.

```
respuesta = {  
    status: <String>,  
    transcript: <String>,  
    error: <String>  
}
```

status	Puede ser “success” (si todo ha ido bien) o “error” (en caso contrario).
transcript	Transcripción del audio (traducción de la transcripción para modelos end-to-end).
error	Mensaje de error. Solo aparece en el caso en el que status sea “error”.

Función Machine Translation

Recibe un texto y un par de lenguas (origen y destino), y devuelve la traducción del texto.

```
datos = {  
    action: "mt",  
    text: <String>,  
    src_lan: <String>,  
    tgt_lan: <String>,  
    trans_to_use: <String>  
}
```

action	Debe ser el string “mt”.
text	Texto que queremos traducir.
src_lan	Código del idioma origen (es, en, ar).
tgt_lan	Código del idioma destino (es, en, ar).
trans_to_use	Modelo de Machine Translation que se va a utilizar. Puede tomar los siguientes valores: - <i>apertium</i> : Multidioma. Es el valor por defecto. - <i>nemo</i> : Realiza traducciones ES-EN y EN-ES. - <i>huggingface-mt</i> : Realiza traducciones ES-AR, AR-ES, EN-AR y AR-EN.

```
respuesta = {  
    status: <String>,  
    translation: <String>,  
    error: <String>  
}
```

status	Puede ser “success” (si todo ha ido bien) o “error” (en caso contrario).
translation	Traducción del texto.
error	Mensaje de error. Solo aparece en el caso en el que status sea “error”.

Función Text-to-Speech

Recibe un texto y devuelve un audio (en formato base64) con la síntesis de voz.

```
datos = {  
    action: "stt",  
    text: <String>,  
    tts_to_use: <String>  
}
```

action	Debe ser el string “tts”.
text	Texto del que queremos realizar la síntesis de voz.
tts_to_use	Modelo de Text-to-Speech que se utilizará. Puede tomar los siguientes valores: - <i>mozillaTTS</i> : Genera audio en español. Es el valor por defecto. - <i>nemo</i> : Genera audio en inglés.

```
respuesta = {  
    status: <String>,  
    audio: <String>,  
    error: <String>  
}
```

status	Puede ser “success” (si todo ha ido bien) o “error” (en caso contrario).
audio	String que representa un audio en Base64.
error	Mensaje de error. Solo aparece en el caso en el que status sea “error”.