

# StressTech API

*Procesamiento de audios de voz humana para cuantificación y monitorización de estados emocionales.*

La API de StressTech es una herramienta avanzada de análisis de audio que permite evaluar el bienestar emocional y la personalidad de los usuarios mediante el procesamiento de voz y texto. Ofrece análisis detallados de estrés, emociones y rasgos de personalidad, y realiza transcripciones y evaluaciones de sentimiento casi en tiempo real. Su tecnología incluye modelos especializados que clasifican tonos emocionales y detectan entidades y temas conversacionales, lo que facilita la creación de experiencias personalizadas en aplicaciones de salud mental y bienestar. Con un coste de procesamiento económico y una latencia rápida, StressTech proporciona un apoyo útil y adaptable para el monitoreo de la salud emocional en aplicaciones dinámicas y centradas en el usuario.

<b>Introducción</b>	<b>1</b>
<b>Beneficios</b>	<b>1</b>
<b>Análisis de Estrés y Bienestar Emocional</b>	<b>1</b>
<b>Perfiles de Personalidad</b>	<b>2</b>
<b>Análisis Avanzado de Voz</b>	<b>2</b>
<b>Transcripción de Voz y Análisis Lingüístico</b>	<b>2</b>
<b>Interacción en Tiempo Real y Retroalimentación Inmediata</b>	<b>2</b>
<b>Optimización y Uso en Aplicaciones de Salud Mental</b>	<b>3</b>
<b>Detalles técnicos</b>	<b>3</b>
Modelos de procesamiento de la señal	4
Modelos de procesamiento de voz	4
Modelos de procesamiento del lenguaje	5
Modelos de creación de sonidos binaurales (audios relajantes)	6
<b>Output API</b>	<b>7</b>
<b>Preguntas sobre el Cliente y el Producto</b>	<b>10</b>
<b>Preguntas sobre Implementación y Logística</b>	<b>11</b>
<b>Preguntas sobre la Tecnología y los Detalles Técnicos</b>	<b>11</b>
<b>Preguntas sobre el Futuro del Producto</b>	<b>12</b>
<b>Unidades económicas</b>	<b>12</b>

## Introducción

La API de StressTech ofrece un completo sistema de análisis de audio para evaluar aspectos de bienestar emocional y personalidad, brindando herramientas avanzadas para el procesamiento de voz y texto. A través de esta API, es posible transformar grabaciones de audio en información detallada sobre estrés, emociones, personalidad, y otros aspectos del estado mental, permitiendo crear experiencias y recomendaciones personalizadas.

## Beneficios

### Análisis de Estrés y Bienestar Emocional

- **Detección de estrés:** La API mide el nivel de estrés en tres categorías (alto, medio y bajo), ayudando a identificar estados emocionales críticos.
- **Detección de depresión y autoeficacia:** Evalúa la presencia de depresión y autoeficacia en niveles alto, medio y bajo, permitiendo realizar un seguimiento de la salud mental.
- **Indicadores de impotencia y autocompasión:** Incluye medidas para detectar la sensación de impotencia y autocompasión, lo que ofrece información detallada sobre el bienestar emocional del usuario.

### Perfiles de Personalidad

- **Características de personalidad:** Evalúa rasgos como creatividad, autoestima, compasión y comunicación, proporcionando una visión personalizada de la estructura emocional y mental del usuario.
- **Datos útiles para personalización:** Estos datos permiten crear contenido y recomendaciones personalizadas basadas en la personalidad de cada usuario.

### Análisis Avanzado de Voz

- **Análisis de frecuencias:** Proporciona estadísticas sobre la frecuencia vocal (media, desviación, moda, entre otros), además de clasificaciones musicales (notas), útil para evaluar patrones en la voz.
- **Análisis de tono emocional y polaridad:** Permite clasificar el tono emocional del usuario en tiempo real, detectando emociones como tristeza, calma, enojo, etc.
- **Información sobre entidades y temas:** La API reconoce entidades como nombres de personas y temas de conversación, así como tópicos relacionados con el audio grabado, lo cual es valioso para analizar el contexto de las interacciones.

### Transcripción de Voz y Análisis Lingüístico

- **Transcripción automática de audio:** Convierte el audio en texto, detectando el idioma, probabilidad de silencio y otras métricas como la entropía, que es útil para el análisis semántico.
- **Análisis de sentimientos:** La API evalúa la polaridad y la subjetividad del lenguaje utilizado, útil para detectar si el discurso tiene una connotación positiva, negativa o neutra.
- **Detección de emociones en texto:** Proporciona una lista de emociones detectadas en el discurso textual, facilitando la identificación de estados de ánimo específicos.

### Interacción en Tiempo Real y Retroalimentación Inmediata

- **Latencia de respuesta rápida:** Con una latencia de 60 segundos, permite que la aplicación genere análisis casi en tiempo real, ideal para integrarse en experiencias de usuario dinámicas.
- **Generación de contenido de apoyo:** Con la información recopilada, es posible crear audios relajantes o contenidos de apoyo adaptados a las necesidades emocionales identificadas.

### Optimización y Uso en Aplicaciones de Salud Mental

- **Seguimiento y apoyo en salud mental:** Al detectar estrés, depresión y emociones, la API es ideal para aplicaciones de salud mental que busquen monitorear y apoyar el bienestar emocional.
- **Posibilidad de realizar intervenciones personalizadas:** Los datos recopilados ayudan a que las aplicaciones respondan con intervenciones efectivas en función de los estados emocionales detectados.

## Detalles técnicos

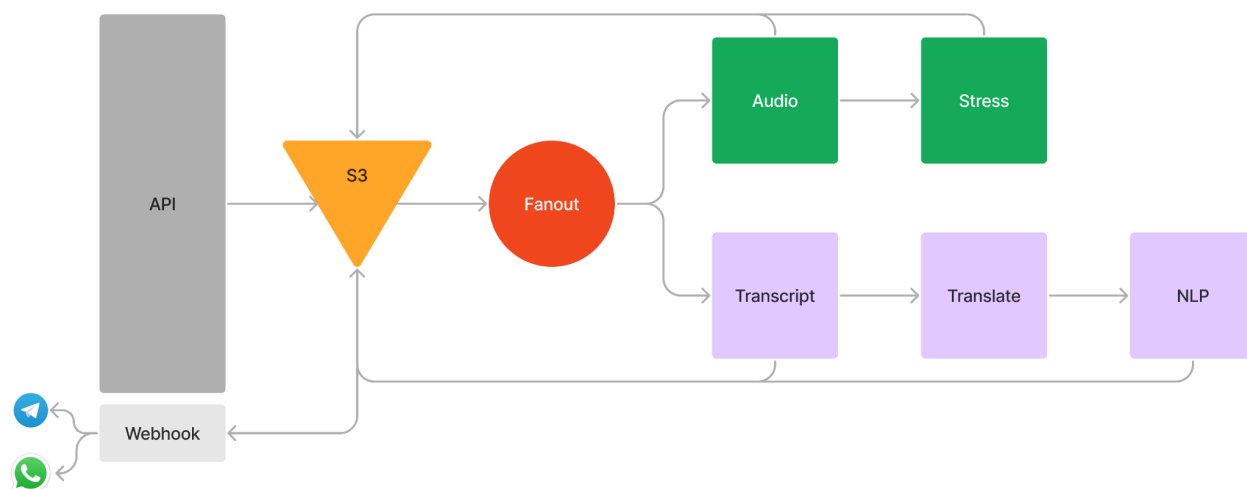


Figura 1. Arquitectura StressTech API.

### Input: Archivo de Audio:

- **Formato:** Soporta los formatos .ogg, .wav y .mp3, que son comunes en la grabación y almacenamiento de audio.
- **Entrada:** Es el punto de partida del proceso, el archivo que se desea analizar o procesar.

**API:** Es la interfaz que actúa como la puerta de entrada al sistema.

### Output: Módulos de Procesamiento:

- **Análisis de Audio:** Limpia y examina las características del audio, como el tono, frecuencias, el timbre, el ruido de fondo, etc.
- **Clasificación de Audio:** Categoriza el audio para la predicción de los niveles de estrés, depresión y emociones.
- **Transcripción:** Convierte el audio en texto escrito.
- **Traducción:** Traduce el texto transcrito a otro idioma.
- **Procesamiento del Lenguaje Natural (NLP):** Realiza un análisis más profundo del texto transcrito, identificando entidades, temáticas, emociones, sentimientos, intenciones, etc.

### Webhook:

- **Notificación:** Es un mecanismo para notificar a otros sistemas o aplicaciones cuando se completa el procesamiento o cuando ocurre un evento específico.
- **Integración:** Permite conectar este sistema con otros servicios o aplicaciones externas.

En el siguiente enlace se puede acceder al portal de desarrolladores de nuestra API que permite enviar un audio y analizarlo. [Portal de desarrollador de StressTech API.](#)

## Modelos de procesamiento de la señal

- **Limpieza y Análisis de Audio.** Procesos de limpieza de audio, análisis de frecuencia, tono y creación de espectrogramas para obtener información precisa de la señal de voz.
- **Análisis de frecuencias por nota musical.** Permitiendo asociarlo a centros emocionales y características de la personalidad.

## Modelos de procesamiento de voz

Los siguientes modelos tienen como input el audio y/o espectrograma de los audios.

### Voice emotional stress transformer (Audio classification).

- **Objetivo:** Predice niveles de estrés, autoeficacia e impotencia basados en la voz.
- **Datos de entrenamiento:** Propios a partir de 850 audios generados por 140 individuos y cuestionario Perceived Stress Questionnaire (PSQ).
- **Desempeño:** F1-Score de 0.95, ROC AUC de 0.95 y precisión de 0.87.
- **Estado:** En producción
- [VEST model documentation](#)

### Voice depression transformer (Audio classification).

- **Objetivo:** Predice niveles de depresión en el usuario.
- **Datos de entrenamiento:** EATD Corpus
- **Desempeño:** F1-Score de 0.95, ROC AUC de 0.96 y precisión de 0.94.
- **Estado:** En producción + necesita reentrenamiento

### Voice emotional transformer (Multi audio classification).

- **Objetivo:** Predecir los niveles emocionales (ira, tristeza, disgusto, miedo, sorpresa, felicidad, calma) que está sintiendo un individuo en su sistema nervioso a partir de un audio de su voz.
- **Datos de entrenamiento:**
  - Crowd-sourced Emotional Multimodal Actors Dataset (Crema-D)
  - Ryerson Audio-Visual Database of Emotional Speech and Song (Ravdess)
  - Surrey Audio-Visual Expressed Emotion (Savee)
  - Toronto emotional speech set (Tess)
- **Desempeño:** Precisión de 0.64 en multi-clasificación.
- **Estado:** En producción + necesita reentrenamiento

### Speech age gender transformer (Audio regression and classification) - TBD

- **Objetivo:** Estima la edad y el género del usuario a partir de la voz.
- **Datos de entrenamiento:** <https://arxiv.org/pdf/2306.16962>
- **Desempeño:** MAE (error medio absoluto) entre 7.1 años y 10.8 años para la edad, y al menos un 0.91 de precisión para el género (femenino, masculino, infantil).
- **Estado:** Poner en producción.

**Whisper (Audio translation).**

- **Objetivo:** Reconocimiento, traducción e identificación de idioma. Para más información, consulte [Whisper en OpenAI](#).
- **Estado:** En producción.

**Modelos de procesamiento del lenguaje**

Los siguientes modelos como input el texto generado por Whisper.

**Reconocimiento de Entidades (NER).**

- **Objetivo:** Detecta entidades como Personas, lugares, eventos temporales, eventos vitales, enfermedades, síntomas, actividades, interacciones sociales, dinero y personajes ficticios.
- **Datos de entrenamiento:** GLiNER
- **Desempeño:** F1-Score de 0.83.
- **Estado:** En producción.

**Reconocimiento de temáticas (Multilabel text classification).**

- **Objetivo:** Detecta temas como Arte y cultura, negocios, cultura pop, rutina diaria, familia, moda, cine y TV, salud, alimentación, juegos, aprendizaje y educación, música, relaciones sociales, ciencia y tecnología, deportes, viajes, vida juvenil.
- **Datos de entrenamiento:** Modelo RoBERTa-large ajustado para clasificación de temáticas en el texto.
- **Desempeño:** F1-Score de 0.83.
- **Estado:** En producción + revisar.

**Reconocimiento de emociones (Multilabel text classification).**

- **Objetivo:** Detecta emociones en el texto tales como Admiración, diversión, enfado, molestia, aprobación, cuidado, confusión, curiosidad, deseo, decepción, desaprobación, disgusto, vergüenza, excitación, miedo, gratitud, dolor, alegría, amor, nerviosismo, optimismo, orgullo, realización, alivio, remordimiento, tristeza y sorpresa.
- **Datos de entrenamiento:** Modelo RoBERTa-large ajustado para clasificación de temáticas en el texto.
- **Desempeño:** F1-Score de 0.78.
- **Estado:** En producción + revisar.

**Reconocimiento de tiempo verbal utilizado (Pos-tagging).**

- **Objetivo:** Identifica el tiempo verbal dominante: pasado, presente o futuro.
- **Datos de entrenamiento:** N/A.
- **Desempeño:** N/A.
- **Estado:** En producción.

**Reconocimiento de sentimiento (Sentiment analysis).**

- **Objetivo:** Analiza la polaridad y subjetividad en el texto.

- **Datos de entrenamiento:** N/A.
- **Desempeño:** F1-Score de 0.97.
- **Estado:** En producción.

## Modelos de creación de sonidos binaurales (audios relajantes)

A partir de las frecuencias de la voz se crean audios personalizados de meditación para crear coherencia del sistema nervioso.

- **Objetivo:** Crear audio binaural y con sonido blanco de fondo para inducir ondas theta asociadas con la relajación profunda y la creatividad. También se relacionan con estados meditativos ligeros, ideales para la calma.
- **Estado:** En producción.
- **Beneficios:**
  - **Relajación del sistema nervioso:** Al entrar en un estado de calma, el sistema nervioso parasimpático se activa, promoviendo la relajación y reduciendo los niveles de cortisol, la hormona del estrés.
  - **Mejora de la concentración y claridad mental:** Las ondas Alpha y Theta pueden mejorar la concentración y reducir la sensación de ansiedad.
  - **Facilita la entrada en un estado meditativo:** Los binaural beats pueden ser una herramienta útil para personas que encuentran difícil la meditación, al inducir un estado meditativo con menos esfuerzo consciente.
  - **Promueve el bienestar emocional:** Se ha observado que la exposición a frecuencias calmantes reduce la percepción de estrés y mejora el estado de ánimo.

## Documentación API

<https://stresstech-api-main-8287389.d2.zuplo.dev/docs/routes/index>

## Output API

### Metadatos Generales

- **created\_at:** Marca de tiempo en formato Unix que indica el momento en que se creó el análisis de audio.
- **id:** Identificador único del análisis que combina un prefijo de token y el identificador único del archivo de audio, asegurando trazabilidad.
- **aid:** Identificador único del archivo de audio en el sistema, utilizado para vincular el audio a su proceso de análisis.
- **token:** Cadena de autenticación o identificador de sesión para el análisis, necesaria para acceder a recursos específicos.

## Almacenamiento en S3

- **s3**: Información sobre la ubicación del archivo de audio en el almacenamiento en la nube.
  - **bucket**: Nombre del bucket en Amazon S3 donde se almacena el archivo.
  - **key**: Ruta completa del archivo dentro del bucket, incluyendo el nombre y formato del archivo.

## Enlaces y Formato

- **result\_url**: URL relativa que permite obtener los resultados del análisis de audio.
- **format**: Especifica el formato del archivo de audio analizado, en este caso, **.wav**.

## Estado de Proceso

- **status**: Indicador del estado de cada proceso del análisis, donde cada clave representa una etapa.
  - **AUDIO\_STORED**: **True** si el audio fue almacenado correctamente.
  - **AUDIO\_ANALYSED**: **True** si el audio fue analizado.
  - **AUDIO\_TRANSCRIBED**: **True** si el audio fue transcrito.
  - **STRESS\_ANALYSED**: **True** si se realizó el análisis de estrés.
  - **TEXT\_ANALYSED**: **True** si se completó el análisis del texto.
  - **SOUND\_CREATED**: **True** si se generó un sonido de salida en base al análisis.

## Variables Externas

- **external\_vars**: Contenedor para variables o datos adicionales externos, actualmente vacío.

## Datos de Análisis

### Rasgos de Personalidad

- **traits**: Evaluación de varios rasgos de personalidad.
  - Valores entre 0 y 1 indican el nivel estimado de cada rasgo:
    - **survival**: Supervivencia.
    - **creativity**: Creatividad.
    - **self\_esteem**: Autoestima.
    - **compassion**: Compasión.
    - **communication**: Habilidades de comunicación.
    - **imagination**: Imaginación.
    - **awareness**: Conciencia.

### Estadísticas de Frecuencia



- **frequencies**: Análisis de frecuencia de la voz.
  - **mean**: Media de las frecuencias.
  - **sd**: Desviación estándar.
  - **median**: Mediana de las frecuencias.
  - **mode**: Moda.
  - **Q25 y Q75**: Percentiles 25 y 75.
  - **IQR**: Rango intercuartílico.
  - **skewness**: Medida de asimetría.
  - **kurtosis**: Curtosis, medida de forma de la distribución.
  - **mean\_note**: Nota musical promedio.
  - **median\_note** y **mode\_note**: Nota mediana y nota moda.
  - **Q25\_note** y **Q75\_note**: Nota en los percentiles 25 y 75.
  - **rmse**: Error cuadrático medio de la señal.

#### Duración

- **duration**: Duración total del archivo de audio en segundos.

#### Análisis de Estrés y Depresión

- **stress**: Probabilidades de estrés en tres niveles.
  - **low**: Probabilidad de estrés bajo.
  - **medium**: Probabilidad de estrés medio.
  - **high**: Probabilidad de estrés alto.
- **depression**: Probabilidades de depresión en tres niveles.
  - **low**: Probabilidad de depresión baja.
  - **medium**: Probabilidad de depresión media.
  - **high**: Probabilidad de depresión alta.

#### Autoeficacia e Impotencia

- **self\_efficacy**: Probabilidades de autoeficacia en tres niveles.
  - **low, medium, high**: Niveles de autoeficacia.
- **helplessness**: Probabilidades de impotencia.
  - **low, medium, high**: Niveles de impotencia.

#### Emociones Detectadas en Voz

- **vemotions**: Probabilidades de emociones en el tono de voz.
  - Cada clave representa una emoción (**sad, disgust, neutral**, etc.) con su respectiva probabilidad.

## Análisis de Sentimiento y Emociones

- **sentiment**: Análisis de sentimiento en el texto.
  - **polarity**: Polaridad, indicador de si el sentimiento es positivo o negativo.
  - **subjectivity**: Nivel de subjetividad del texto.
- **emotions**: Lista de emociones detectadas en el texto con sus probabilidades.

## Reconocimiento de Entidades

- **entities**: Entidades identificadas en el texto transcrito.
  - Cada elemento contiene:
    - **entity**: Nombre de la entidad.
    - **label**: Tipo de entidad (por ejemplo, **person**, **activity**).

## Reconocimiento de Temas

- **topics**: Lista de temas identificados en el texto, con sus puntuaciones de relevancia.
  - Cada elemento incluye:
    - **label**: Tema detectado (por ejemplo, **business\_&\_entrepreneurs**).
    - **score**: Puntuación que indica la relevancia de dicho tema en el texto.

## Transcripción de Audio

- **transcription**: Información sobre el texto transcrito.
  - **text**: Texto transcrito.
  - **language**: Idioma detectado en el audio.
  - **no\_speech\_prob**: Probabilidad de silencio en el audio.
  - **entropy**: Entropía del texto, medida de diversidad de información.
  - **tense**: Presencia de tiempos verbales en el texto.
    - **past**, **present**, **future**: Probabilidad de que el texto esté en cada tiempo verbal.

## Latencia de Respuesta

- **latency**: Tiempo estimado en segundos para completar el análisis, incluyendo procesamiento y almacenamiento.

## Preguntas sobre el Cliente y el Producto

### ¿Qué es la API de StressTech?

La API de StressTech es una herramienta avanzada de análisis de audio diseñada para evaluar el bienestar emocional y la personalidad de un usuario. A través del procesamiento de

grabaciones de voz, convierte la información sonora en datos sobre estrés, emociones, personalidad y otros aspectos psicológicos, facilitando experiencias personalizadas y recomendaciones para el bienestar emocional.

### **¿Por qué debería elegir este producto?**

La API de StressTech permite detectar y monitorear estados emocionales complejos como el estrés, la depresión y la autoeficacia. Es útil para aplicaciones de salud mental y bienestar, ofreciendo un análisis detallado y preciso que permite responder a las necesidades emocionales de los usuarios de manera personalizada.

### **¿Qué beneficios ofrece frente a otros productos del mercado?**

A diferencia de otros productos, StressTech ofrece una evaluación integral en tiempo real de diferentes aspectos emocionales y de personalidad, incluyendo análisis de tono emocional, perfil de personalidad, transcripción automática de voz, y retroalimentación casi inmediata. Además, la API tiene una latencia baja, menor a 60 segundos, ideal para aplicaciones de salud y bienestar en tiempo real.

### **¿Cómo me ayuda este producto a resolver problemas de estrés y bienestar emocional?**

La API de StressTech ayuda a resolver problemas de bienestar emocional al detectar niveles de estrés, depresión y autoeficacia. También evalúa emociones como tristeza, calma y enojo, facilitando la creación de contenido y recomendaciones que pueden ayudar al usuario a gestionar su salud emocional.

## **Preguntas sobre Implementación y Logística**

### **¿Cuándo y dónde estará disponible el producto?**

La API de StressTech ya se encuentra en producción y disponible para integrarse en aplicaciones de salud mental y plataformas de bienestar en línea. El acceso puede gestionarse a través del portal de desarrolladores de StressTech.

### **¿Cuál será el precio y las opciones de compra?**

Para obtener información precisa sobre precios y opciones de suscripción, es necesario revisar la documentación en el portal de desarrolladores de StressTech o contactar al equipo de ventas para configuraciones de uso personalizadas.

### **¿Se necesitará algún equipo o software adicional?**

Para utilizar la API, se requiere un archivo de audio en formato .ogg, .wav o .mp3. No es necesario equipo adicional, pero se recomienda contar con una conexión estable a internet para enviar y recibir los datos de procesamiento.

## **¿Cómo es el proceso de instalación/configuración?**

La integración se realiza mediante API, permitiendo que el sistema de StressTech procese archivos de audio y devuelva datos de análisis. El portal de desarrolladores ofrece documentación para conectar la API a otros sistemas mediante webhooks.

## **Preguntas sobre la Tecnología y los Detalles Técnicos**

### **¿Cómo funciona la tecnología principal detrás de este producto?**

La API de StressTech utiliza tecnologías avanzadas de procesamiento de audio y lenguaje natural (NLP), junto con modelos de machine learning para clasificar y analizar tonos emocionales y de personalidad a partir de la voz y texto. Procesa los datos en tiempo real y devuelve indicadores de salud emocional y personalidad.

### **¿Cuáles son los requisitos de hardware/software?**

Para la interacción con la API, se necesita acceso a internet y un archivo de audio en un formato compatible (.ogg, .wav o .mp3). Los desarrolladores pueden integrarla en aplicaciones de software compatibles.

### **¿Qué tan seguro es el producto para el usuario?**

StressTech asegura la protección de datos mediante identificadores únicos y almacenamiento en la nube (Amazon S3), con opciones de autenticación de sesión. Cada análisis es identificado de manera única para garantizar privacidad y trazabilidad.

### **¿Cómo se actualiza el producto?**

Las actualizaciones de software y los modelos de aprendizaje se aplican periódicamente. Los modelos como "Voice depression transformer" y "Voice emotional transformer" están configurados para reentrenamiento, mejorando su precisión y adaptación a nuevos datos.

## **Preguntas sobre el Futuro del Producto**

### **¿Cuál es la visión a largo plazo para este producto?**

StressTech planea evolucionar integrando más características y funciones para ampliar su alcance en el monitoreo emocional y la personalización en el bienestar, optimizando continuamente sus modelos y agregando compatibilidad con otros tipos de análisis de voz y texto.

### **¿Qué mejoras futuras están planeadas?**

Entre las mejoras previstas se incluye el perfeccionamiento de los modelos para mejorar la precisión de predicciones emocionales y la adición de análisis de características como edad y

género. También se contempla la posibilidad de integración con sistemas externos para una mayor accesibilidad.

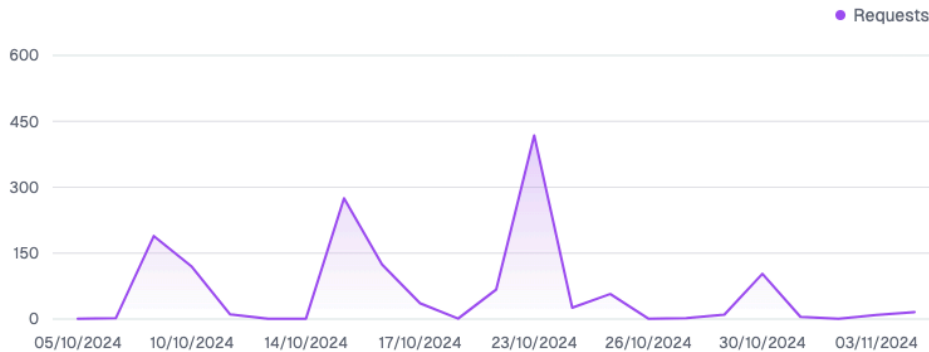
### ¿El producto tendrá compatibilidad con otros sistemas/plataformas?

Sí, StressTech permite integraciones mediante webhooks y opciones de API, lo que facilita su conexión con otros sistemas y plataformas. Esto incluye la posibilidad de interconectar aplicaciones de salud y bienestar para crear un ecosistema de monitoreo emocional centralizado.

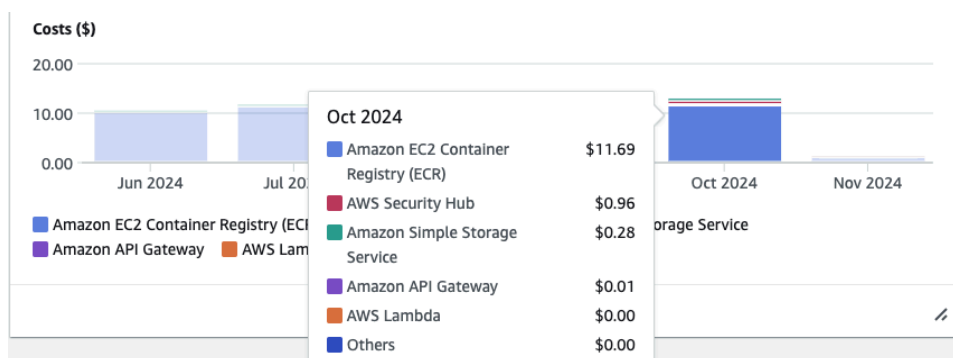
## Unidades económicas

Número de llamadas a la API desde el 1 de octubre al 1 de noviembre ~1500

### Requests



Gasto en AWS desde el 1 de octubre al 1 de noviembre ~ 15\$ - 13,75€



Coste de procesamiento por audio ~ 0,01€