

ElevenLabs

Plataforma de inteligencia artificial líder para departamentos de marketing, productoras audiovisuales y desarrolladores que necesitan generación de voz sintética de alta fidelidad. Permite crear locuciones con carga emocional humana, clonación de voz profesional y doblaje automático en múltiples idiomas. Es la solución ideal para escalar la producción de audio en videojuegos, e-learning y medios de comunicación, eliminando la dependencia constante de locutores presenciales y estudios.

[Visitar Sitio Oficial](#) [Preguntar a ChatGPT](#) [Preguntar a Claude](#) [Preguntar a Grok](#)

Contenido del Dossier

- [Información de la Herramienta](#)
- [Consejos de Implantación](#)
- [Preguntas Frecuentes](#)
- [Contratos y Condiciones](#)

INFORMACIÓN DE LA HERRAMIENTA

Qué y para quién es

ElevenLabs es una plataforma de inteligencia artificial líder en la generación de voz sintética, clonación de voz y doblaje automatizado. Utiliza modelos de aprendizaje profundo para crear locuciones con una carga emocional y entonación humanas, superando las limitaciones de los sistemas de texto a voz (TTS) tradicionales.

En el ámbito profesional, está diseñado para departamentos de marketing, productoras audiovisuales, desarrolladores de software y creadores de contenido que buscan una producción de audio de alta calidad, escalable y en múltiples idiomas sin depender de locutores presenciales para cada iteración.

Principal ventaja profesional

La capacidad de generar voces con una naturalidad extrema y la opción de **Professional Voice Cloning (PVC)**, que permite crear una réplica digital de una voz específica con una fidelidad indistinguible de la original, permitiendo la personalización de marca a escala global.

Para quién no es

No es la herramienta ideal para empresas que requieren locuciones puramente robóticas o funcionales sin matices (como avisos de megafonía simples), ni para aquellas con presupuestos extremadamente reducidos que no puedan cubrir los costes por caracteres si el volumen de producción es masivo y constante. Tampoco es apta para usuarios que busquen anonimato total en el uso de voces ajenas, debido a sus estrictas políticas de ética y verificación.

Funcionalidades clave

- **Text-to-Speech (TTS):** Conversión de texto a audio con control de estabilidad, claridad y exageración de estilo.
- **Speech-to-Speech:** Transforma una grabación de voz propia en la voz de otro personaje manteniendo la emoción y el ritmo.
- **Doblaje Automático:** Traducción y sustitución de voz en vídeos manteniendo la voz original del hablante en el nuevo idioma.
- **Professional Voice Cloning:** Creación de clones digitales de alta fidelidad (requiere muestras de audio extensas y verificación).
- **Voice Design:** Generación de voces sintéticas únicas ajustando parámetros como edad, género y acento.
- **Audio Native:** Widget para sitios web que convierte artículos escritos en audio de forma automática.

Precios

- **Versión gratuita:** 10.000 caracteres al mes. No permite uso comercial y requiere atribución. Incluye 3 slots de voz personalizados (no clonación instantánea).
- **Rango de precios:** 5\$ a +1.300\$ al mes (con descuentos por pago anual).
- **Starter (5\$/mes):** 30.000 caracteres, licencia comercial básica y clonación de voz instantánea.
- **Creator (22\$/mes):** 100.000 caracteres, acceso a Professional Voice Cloning y mayor calidad de audio (192kbps).
- **Pro / Scale / Business:** Desde 99\$ hasta 1.320\$/mes para volúmenes de entre 500.000 y 11.000.000 de caracteres, con soporte de mayor calidad (44.1kHz), menor latencia y gestión de equipos.
- **Enterprise:** Precios personalizados con SSO, cumplimiento HIPAA/BAA y soporte dedicado.

Perfil del usuario

- **Agencias de Marketing:** Creación de anuncios personalizados y contenido para redes sociales.
- **Desarrolladores de Videojuegos:** Locución de NPCs y prototipado rápido de diálogos.
- **Sectores de e-Learning:** Producción de cursos multilingües con voces consistentes.
- **Medios de Comunicación:** Automatización de versiones en audio de noticias escritas.
- **Empresas Tecnológicas:** Integración de voces naturales en asistentes virtuales y aplicaciones mediante API.

Nivel técnico requerido

- **Uso:** bajo. La interfaz web es intuitiva y funciona mediante editores de texto.
- **Configuración API:** medio/alto. Requiere conocimientos de desarrollo para integrar las claves API en aplicaciones externas o flujos de trabajo automatizados.
- **Clonación Profesional:** medio. Requiere seguir procesos específicos de grabación y limpieza de audio

para obtener resultados óptimos.

Ejemplos de uso profesional

- **Localización de producto:** Traducir y doblar una demo de software a 29 idiomas en minutos manteniendo una voz corporativa uniforme.
- **Audiobooks:** Generar narraciones de libros completos con voces que mantienen la coherencia narrativa y emocional.
- **Personalización de ventas:** Crear mensajes de audio personalizados con el nombre del cliente utilizando una voz de marca preestablecida.

Uso y distribución

- **Versión web:** Plataforma completa accesible desde cualquier navegador.
- **API:** Documentación técnica exhaustiva para integración en productos propios.
- **Versión móvil:** App disponible para consumo y generación rápida de contenido.
- **Reader App:** Aplicación específica para escuchar cualquier texto o documento con sus voces.

Integraciones

- **Facilidad de integración:** Full code mediante API REST avanzada.
- **API propia:** API potente que permite TTS en streaming con latencia ultrabaja.
- **Integraciones nativas:** Conexión con plataformas de creación de contenido y automatizadores como Make o Zapier.

Notas finales

Información legal, licencias y contratos

- Los derechos comerciales pertenecen al usuario en todos los planes de pago.
- En el plan gratuito, la propiedad intelectual del modelo es de ElevenLabs y el uso comercial está prohibido.
- La plataforma incluye herramientas de marca de agua de audio (AI Speech Classifier) para verificar si un audio fue generado con su tecnología, mitigando riesgos de seguridad.

Para más información:

- [Sitio web oficial](#)
- [Precios](#)
- [Documentación técnica \(API\)](#)
- [Políticas de seguridad y licencias](#)

CONSEJOS DE IMPLANTACIÓN

Aplicación profesional

- **Empresas de escala global:** Ideal para corporaciones que gestionan atención al cliente en múltiples idiomas (ej. Revolut, Klarna) o productoras que requieren doblaje rápido.
- **Startups tecnológicas:** Integración de "agentes de voz" en productos SaaS para automatizar entrevistas de usuarios o soporte técnico.
- **Presupuesto:** Variable; desde implementaciones de bajo coste para prototipado (\$22/mes) hasta soluciones Enterprise de alto volumen que superan los \$1,300/mes.
- **Puntos clave:** Reducción de hasta 8 veces en tiempos de resolución de incidencias y ahorro de meses en desarrollo de infraestructura de audio propia.

Madurez digital requerida

- **Usuarios:** Capacidad para diseñar prompts complejos y gestionar flujos de conversación lógica (si se usan ElevenAgents).
- **Empresa:** Necesidad de una infraestructura de datos organizada (CRM/ERP) si se desea que la IA realice acciones (reembolsos, agendar citas) y no solo hable.

Plan orientativo de implantación

Pasos necesarios y estimaciones

- **Tiempo de despliegue:** Desde 1 semana para una capa de voz lista en producción (usando ElevenAgents API) hasta 3 meses para sistemas complejos integrados con telefonía.
- **Fase 1: Evaluación (1-2 semanas):** Identificación de casos de uso de bajo riesgo pero alto impacto (ej. preguntas frecuentes o traslados de llamadas).
- **Fase 2: Configuración y Clonación (1-2 semanas):** Creación de la voz de marca (Professional Voice Cloning) y ajuste de estabilidad y claridad del modelo.
- **Fase 3: Prueba de Concepto (2-4 semanas):** Lanzamiento de un piloto controlado. Selección de LLM (ej. Gemini o GPT) para equilibrar latencia y razonamiento.
- **Fase 4: Escalado:** Integración con sistemas de telefonía empresarial o flujos de trabajo automatizados (Make, Zapier, n8n).

Necesidades de formación del equipo

- Entrenamiento en diseño de "Prompts de Sistema" para evitar que el agente sea redundante o robótico.
- Capacitación en ética de IA y cumplimiento de privacidad (GDPR/RGPD) para el manejo de grabaciones de voz.

Perfiles necesarios

- **Perfiles técnicos:** Desarrolladores Backend (integración API REST) y especialistas en IA/NLP.
- **Personal externo recomendado:** Consultores en experiencia de usuario de voz (VUI) y especialistas en seguridad de identidad (prevención de deepfakes).

Retorno de la inversión (ROI)

- **Tiempos:** Recuperación de la inversión visible en los primeros 3-6 meses mediante ahorro en costes de locución profesional y reducción de carga en soporte humano.
- **KPIs clave:** Tasa de éxito en llamadas (objetivo >90%), reducción del tiempo medio de gestión (AHT) y tasa de resolución en el primer contacto (FCR).

Otros

- **Seguridad:** ElevenLabs incluye AI Speech Classifier, una herramienta esencial para que las empresas verifiquen la autenticidad de sus propios contenidos y eviten fraudes por suplantación.
- **Latencia:** Clave en aplicaciones en tiempo real; el uso de modelos "Turbo" es mandatorio para evitar pausas incómodas en la conversación humana.

PREGUNTAS FRECUENTES

¿Qué es ElevenLabs y en qué se diferencia del TTS tradicional?

ElevenLabs es una plataforma de inteligencia artificial especializada en síntesis de voz que utiliza modelos de aprendizaje profundo para generar locuciones con entonación, ritmo y carga emocional humana. A diferencia de los sistemas de texto a voz (TTS) tradicionales, que suelen sonar metálicos o monótonos, esta tecnología permite una expresividad avanzada y la capacidad de clonar voces con una fidelidad indistinguible de la original.

¿Tiene una versión gratuita y qué limitaciones profesionales presenta?

Sí, existe una versión gratuita que ofrece 10.000 caracteres mensuales y permite probar las funciones básicas. Sin embargo, para uso profesional presenta tres limitaciones críticas: no permite el uso comercial de los audios generados, exige la atribución obligatoria a la plataforma y no incluye acceso a la clonación de voz instantánea ni profesional.

¿Cuáles son los costes y cómo escalan según el uso profesional?

Los planes varían desde el nivel Starter (5\$/mes por 30.000 caracteres) hasta el nivel Business (1.320\$/mes por 11 millones de caracteres). Para profesionales independientes, el plan Creator (22\$/mes) es el estándar recomendado, ya que desbloquea el Professional Voice Cloning y una calidad de audio superior a 192kbps. Para grandes organizaciones, existen planes Enterprise con soporte personalizado y cumplimiento normativo específico.

¿Cómo funciona el Professional Voice Cloning (PVC) y qué requisitos tiene?

El PVC permite crear una réplica digital exacta de una voz específica. A diferencia de la clonación instantánea, requiere que el usuario cargue muestras de audio extensas y de alta nitidez (mínimo 30 minutos recomendados) para entrenar un modelo dedicado. Además, por motivos de seguridad, exige un proceso de verificación donde el locutor original debe leer un texto generado dinámicamente para confirmar su consentimiento.

¿Es posible integrar ElevenLabs en aplicaciones externas de forma automatizada?

Sí, la plataforma ofrece una API REST avanzada diseñada para desarrolladores, que permite integrar la generación de voz en streaming con latencia ultrabaja en videojuegos, aplicaciones móviles o asistentes virtuales. También es compatible con herramientas de automatización como Make o Zapier para flujos de trabajo sin código.

¿A quién pertenecen los derechos de propiedad intelectual de los audios generados?

Los derechos de explotación comercial pertenecen íntegramente al usuario siempre que se utilicen planes de pago. En el plan gratuito, ElevenLabs retiene la propiedad intelectual del modelo generado y prohíbe cualquier uso con fines de lucro.

¿Cómo aborda la plataforma la ética, la privacidad y el uso de deepfakes?

ElevenLabs implementa políticas estrictas para evitar el uso malicioso de su tecnología. Incorpora el 'AI Speech Classifier', una herramienta que permite verificar si un archivo de audio ha sido creado con su red neuronal. Además, prohíbe la clonación de voces de terceros sin su consentimiento explícito y utiliza sistemas de marca de agua de audio para la trazabilidad de los contenidos.

¿Cumple con normativas de seguridad y privacidad empresarial?

La plataforma ofrece cumplimiento con estándares de seguridad avanzados en sus niveles empresariales, incluyendo opciones para acuerdos HIPAA y BAA para sectores que manejan datos sensibles (como salud). También permite la gestión de accesos mediante Single Sign-On (SSO) en el plan Enterprise para garantizar que solo personal autorizado utilice la infraestructura de voz de la compañía.

¿Qué idiomas soporta y cómo funciona el doblaje automático?

Soporta más de 29 idiomas mediante su modelo Multilingual v2. La función de doblaje automático es capaz de traducir el contenido de un vídeo y sustituir la pista de audio original por una versión en otro idioma, manteniendo la identidad vocal del hablante inicial, lo que resulta ideal para la localización de contenidos audiovisuales y marketing internacional.

CONTRATOS Y CONDICIONES

Principales recomendaciones

- Es obligatorio el uso de planes de pago (Starter o superior) para cualquier fin profesional, ya que el plan gratuito prohíbe el uso comercial.
- Se debe obtener el consentimiento explícito y por escrito antes de realizar una clonación de voz (Professional Voice Cloning) de cualquier empleado o colaborador, detallando el propósito y tiempo de uso.
- En caso de utilizar la plataforma para procesar datos de terceros o locuciones con datos personales, la empresa debe firmar un Anexo de Procesamiento de Datos (DPA) con ElevenLabs.
- Es necesario informar a los usuarios finales de que están interactuando con contenido generado por IA para cumplir con las obligaciones de transparencia.

Ley de Inteligencia Artificial (AI Act)

- ElevenLabs se clasifica como un sistema de IA de propósito general con capacidades de generación de contenido (Deepfakes).
- La empresa usuaria tiene la obligación legal de etiquetar de forma clara y legible que el contenido de audio ha sido generado artificialmente o manipulado.
- El sistema de "Professional Voice Cloning" requiere medidas de autenticación reforzadas para evitar la suplantación de identidad, responsabilidad que recae parcialmente en el cliente profesional.

Privacidad y protección de datos

- Responsabilidades: La empresa española actúa como Responsable del Tratamiento y ElevenLabs como Encargado del Tratamiento.
- Ubicación de los datos: Los datos se procesan principalmente en servidores ubicados en Estados Unidos.
- Transferencia internacional: Se basa en Cláusulas Contractuales Tipo (SCC). ElevenLabs (Eleven Labs Inc.) está sujeta a la jurisdicción de EE. UU., lo que implica un impacto medio/alto en el análisis de riesgos para empresas de la UE bajo el marco del RGPD.
- Derechos ARCO: El usuario debe gestionar las solicitudes de acceso o eliminación de muestras de voz de sus empleados directamente a través del panel de ElevenLabs o mediante soporte técnico.

Propiedad intelectual

- Propiedad de datos: El usuario conserva los derechos sobre los archivos de voz originales (muestras) subidos para el entrenamiento de clones.
- Propiedad del resultado: En planes de pago, el usuario ostenta la propiedad comercial y los derechos de explotación sobre los archivos de audio generados.
- En el plan gratuito, la licencia es de atribución no comercial, manteniendo ElevenLabs derechos de uso limitados sobre el contenido para mejora de servicios.

Usos y prohibiciones

- Usos prohibidos: Crear contenido para fraude, suplantación de identidad sin consentimiento, acoso, desinformación política o contenido que infrinja derechos de terceros. No se permite el uso de voces de figuras públicas sin autorización legal.
- Usos admitidos: Narración de libros, doblaje de videos, creación de contenido publicitario, videojuegos y asistencia de accesibilidad.

Seguridad y certificaciones

- Seguridad: Implementación de encriptación de datos en reposo (AES-256) y en tránsito (TLS 1.2+).
- Certificaciones: ElevenLabs cuenta con certificación SOC2 Tipo II, lo que garantiza controles internos rigurosos sobre la seguridad de la información.
- Dispone de herramientas de "AI Speech Classifier" para identificar fragmentos de audio generados por su propia plataforma y verificar autenticidad.

Otros

- La plataforma aplica un sistema de moderación automatizado que puede bloquear la generación de texto si detecta contenido sensible o prohibido por sus políticas de seguridad (Safety).
- Para el uso en sectores regulados como salud (HIPAA), es obligatorio contratar el plan Enterprise y firmar los acuerdos específicos de cumplimiento.

Fuentes consultada:

- [Términos de servicio de ElevenLabs](#)
- [Política de privacidad](#)
- [Centro de seguridad y cumplimiento](#)
- [Documentación sobre licencias comerciales](#)
- [Compromiso con el AI Act y Ética](#)

Para más información y herramientas:

Explora look4.tools para descubrir las mejores soluciones tecnológicas del mercado.

[Inicio](#) [Todas las herramientas](#) [Categorías](#)

Este documento ofrece recomendaciones generadas mediante análisis humano y sistemas de IA automatizados. La información tiene carácter meramente informativo y no constituye asesoramiento legal, profesional ni garantía de resultados. Las marcas, logotipos y nombres comerciales pertenecen a sus respectivos propietarios y se utilizan únicamente con fines identificativos.