

Reconocimiento

para la interacción vocal



Verbio ASR (Automatic Speech Recognition – Reconocimiento automático del habla) **convierte de forma automática una locución de habla natural en un texto con la máxima fiabilidad y en varios idiomas.**

Permite una interacción muy natural puesto que reconoce cualquier palabra y, también, a cualquier persona sin entrenamiento previo. Puede usarse para reconocer palabras aisladas o comandos y para el reconocimiento en habla continua.

Vaya más allá. Permita que sus aplicaciones dialoguen

Con el motor de reconocimiento de voz Verbio ASR conseguirá interactividad hombre-máquina a través de la voz, en aplicaciones interactivas entre usuarios y sistemas informáticos.








Una tecnología de nueva generación que permite una interacción natural basada en la interfaz vocal.

PRINCIPALES VENTAJAS

- 1 Motor independiente del locutor o usuario.** Reconocimiento de cualquier persona sin entreno previo.
- 2 Robustez** frente al ruido y en especial para telefonía móvil. Arquitectura abierta para crear y disponer de modelos especiales para entornos acústicos concretos.
- 3 Bajas tasas de error** gracias al sistema con interpretación semántica del resultado, que también permite parametrizar el desarrollo de las aplicaciones o diálogos de voz, ampliando las posibilidades del desarrollo de diálogos naturales. Ajustable para dispositivos y entornos acústicos concretos.

APLICACIONES

Especialmente indicado para trabajar en entornos telefónicos, aunque tiene aplicabilidad en otros entornos multimedia.

-  **Telefonía y CTI:** Call center. IVR's. Mensajería unificada. Operadoras automáticas. Portales de voz.
-  **Multimedia:** Realización de prototipos, CD's de información genérica y cambiante, formación interactiva, ebooks, quioscos de información.
-  **Internet:** Mediante las tecnologías asociadas a la VoIP, las aplicaciones con reconocimiento a través de la Red abren un ilimitado mundo de aplicaciones para navegar, interactuar o simplemente para rellenar formularios.
-  **Domótica:** Comandos de control de sistemas domóticos.
-  **Discapacitación:** Aplicaciones de ayuda a distintos niveles de discapacidad, tanto motriz como visual.
-  **Industrial:** Automatización industrial de procesos mediante la voz (logística, maquinaria, etc.)
-  **Terminales portátiles (teléfonos móviles, PDA, etc.):** Control del flujo de información y comandos en móviles a través de manos libres.

CARACTERÍSTICAS

Diversidad de idiomas soportados. Español, catalán, euskera, gallego, inglés, francés*, portugués europeo, brasileño y árabe*. Variantes fonéticas del español en Latinoamérica para: México, Argentina, Uruguay, Paraguay, Chile, Colombia, Puerto Rico, Cuba, República Dominicana, Panamá, Venezuela, Perú, Ecuador, Bolivia, Costa Rica, Honduras, Guatemala, El Salvador y Nicaragua.

Sistema multilingüe. Permite el uso simultáneo en varios idiomas.

Detección de interrupción vocal gracias a sistemas avanzados de barge-in que detectan si un usuario desea interrumpir al sistema mediante su comando o locución.

Gramáticas específicas básicas de reconocimiento ya incorporadas (built-in). Cadenas de dígitos, deletreos, sí/no, fechas, horas, DNI, tarjetas de crédito o números telefónicos.

Integración directa con el sistema de verificación de locutor, Verbio Speaker ID.

Simplicidad de integración. Para lograrlo dispone de múltiples SDK de programación, incluyendo un SDK específico para técnicas de Word Spotting.

Especificación de los vocabularios de interés. Gracias a un listado de palabras, gramáticas (ABNF) o modelos de lenguaje estadísticos (SLM).

ESPECIFICACIONES TÉCNICAS

Arquitecturas	<ul style="list-style-type: none"> • Monopuesto [Desktop] • Cliente-Servidor
Requerimientos de memoria	<ul style="list-style-type: none"> • Motor de reconocimiento (Vox Server): 10 MB • Configuración de reconocimiento monolingüe: 16 MB • Configuración de reconocimiento bilingüe: 29 MB • El consumo de memoria depende, principalmente, del vocabulario o gramática activo, por lo que la estimación de consumos debe hacerse en función de la gramática más compleja de la aplicación
Tasa de muestreo	<ul style="list-style-type: none"> • 8 KHz (ley A o ley Mu)
CPU	<ul style="list-style-type: none"> • Requisitos mínimos recomendables Pentium 4 - 3 Ghz 512 MB RAM • La carga computacional depende, principalmente, del vocabulario o gramática activo, por lo que la estimación de consumos debe hacerse en función de la gramática más compleja de la aplicación
Plataforma	<ul style="list-style-type: none"> • Windows NT, 2000, XP, 2003, Vista • Linux • Pocket PC
Interfaces	<ul style="list-style-type: none"> • MRCP v1 • Verbio API (SDK) • VoiceXML
Multi-idioma / voces	<ul style="list-style-type: none"> • Castellano, catalán, gallego, euskera, portugués, brasileño, mexicano, argentino, colombiano, chileno, venezolano, inglés US*, francés* y árabe*
SDK	<ul style="list-style-type: none"> • Diferentes SDK según entorno (Dialogic, Eicon, CTADE, AvayaIR, genérico, etc.)
Otras características especiales	<ul style="list-style-type: none"> • Gramáticas List-of-words, W3C SRGS (ABNF y XML), modelos lingüísticos (SLM) • Reconocimiento independiente del locutor • Optimizado para entornos telefónicos (fijo y móvil) y ruidosos • Posibilidad de desarrollo de nuevos idiomas "on-demand" • Herramientas de "tuning" • Configuraciones de reconocimiento multilingües • Múltiples hipótesis de reconocimiento • Índices de confianza por palabra y por frase • Gramáticas básicas (built-in) incorporadas • Retorno de resultados en formato NLSML • Integración con SpeakerID (verificación del locutor) • WordSpotting para la indexación y localización de palabras clave en ficheros de audio

* Consultar

