SESTEK

Conversational Automation Company

TEXT-TO-SPEECH

DATASHEET



SESTEK
Text-to-Speech

Premium TTS Voice For a Cent
As Natural As Human Speech

Text-to-Speech (TTS) is a powerful speech synthesis technology that can vocalize written text into audible speech with a human-like voice. The technology helps businesses to deliver high-quality self-service applications to customers while improving the experience.

DIFFERENTIATORS

1. Owned Technology

SESTEK controls the model, the roadmap, and the pricing. No dependency on a third party changing an API or deprecating a voice you've already deployed with.

2. Proven Quality

Voice quality is validated through blind MOS (Mean Opinion Score) testing against the market, not asserted without evidence.

3. SSML support

We provide more human-like experiences with pauses, numbers, date, time formatting and more.

4. Multi-language

Currently, we give brands voices in **more than 20 different languages.**

5. Fast Adaptation

We can develop and expand brand voices generated from limited voice samples.

6. Deploy Where Your Data Lives

On-premises for full data control and lower latency, or cloud for zero infrastructure overhead, same voice quality either way.

BENEFITS



NATURAL EXPERIENCE

Built on LLM-based neural architecture, TTS delivers natural-sounding, neutral voices that closely resemble human speech, minimizing the mechanical, robotic quality of synthetic speech.



LOWER COSTS

Fully automated voice output reduces the need for human voice actors on every announcement, and updates go live instantly, with no re-recording.



FASTER INTEGRATION

APIs and SDKs add voice capability to existing software without significant infrastructure changes, so developers ship voice-enabled experiences in days, not months.

© 2025 SESTEK. All rights reserved.

SESTEK
Text-to-Speech

1. Custom Voice Branding

Our Text-to-Speech technology includes a Custom Voice Creation feature that enables businesses to create unique, branded voices that set them apart from competitors. Rather than relying on a generic voice profile, a voice is developed to reflect your brand identity and audience expectations, and it is exclusive to your organization.

The feature uses machine learning and deep neural networks to analyze and replicate the vocal characteristics of a designated voice talent. The resulting voice can then generate high-quality, natural-sounding speech for applications such as e-learning, virtual assistants, and audiobooks. Creating a custom voice involves recording and processing large amounts of speech data, then fine-tuning the model to produce the desired vocal characteristics.



2. Language Support

SESTEK Text-to-Speech technology supports multiple languages, enabling businesses to communicate effectively with a diverse customer base. We are continually exploring new opportunities to expand our language support, so businesses can communicate effectively with customers around the world. The following languages are supported by SESTEK Text-to-Speech:

Arabic, Azerbaijani, Brazilian Portuguese, English, Filipino, French, French (Canadian), German, Hindi, Indonesian, Italian, Japanese, Kurdish, Malay, Mandarin Chinese, Pashto, Persian, Russian, Spanish, Swahili, Thai, Turkish, Ukrainian, and Urdu.

Accent-Specific Voices: SESTEK TTS goes beyond language support by offering accent-specific voices that reflect the rich diversity of speech patterns around the world. For example:

Arabic - voices are available for MSA, Najdi, Kuwaiti, and Emirati accents, each with distinct pronunciation and intonation
English - separate voice sets are available for American (en-US) and British (en-GB) accents

This capability enhances the naturalness and relatability of synthesized speech, enabling more personalized and engaging interactions.

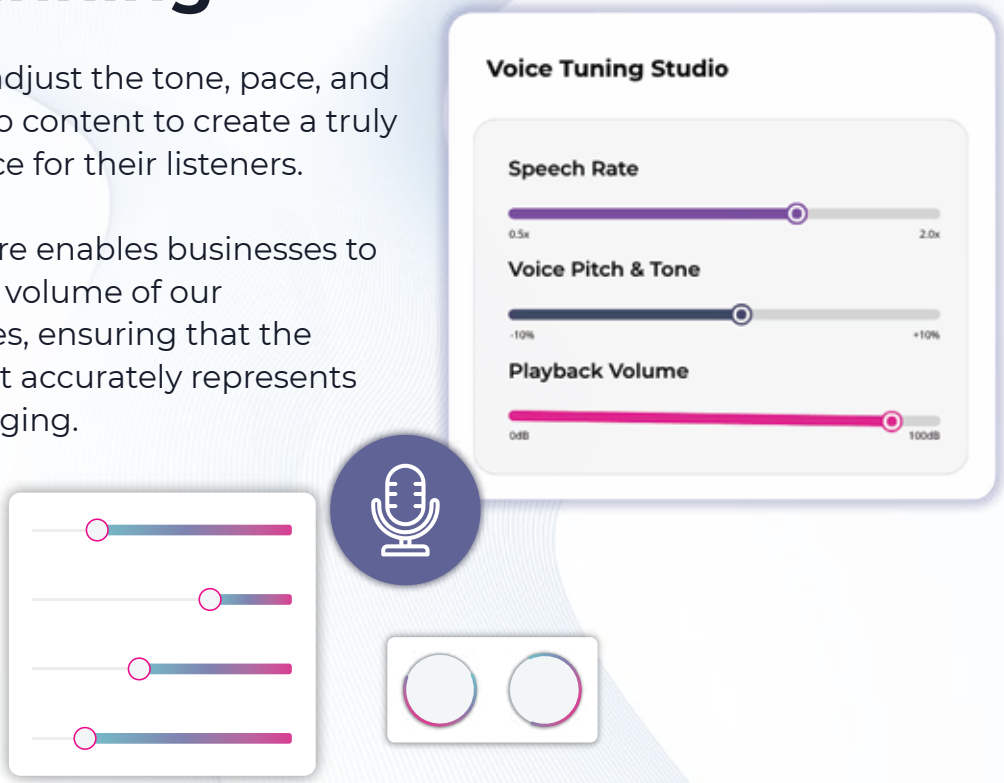
© 2025 SESTEK. All rights reserved.

SESTEK
Text-to-Speech

3. Voice Tunning

Businesses can easily adjust the tone, pace, and emphasis of their audio content to create a truly personalized experience for their listeners.

Our voice tuning feature enables businesses to change the speed and volume of our natural-sounding voices, ensuring that the resulting audio content accurately represents their brand and messaging.



4. SSML Tag Support

We offer our customers the ability to use Speech Synthesis Markup Language (SSML) tags with our Text-to-Speech technology. SSML tags allow businesses to customize the pronunciation, intonation, and emphasis of their audio content, creating a more engaging and personalized experience for the listener. Our support for the most commonly used SSML tags, including break, say as, and even insert audio files, all with just a few lines of code.



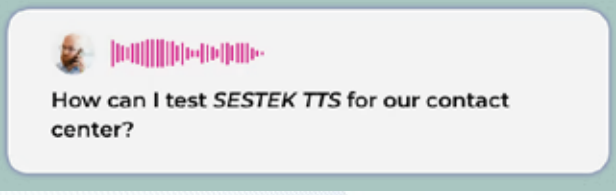
© 2025 SESTEK. All rights reserved.

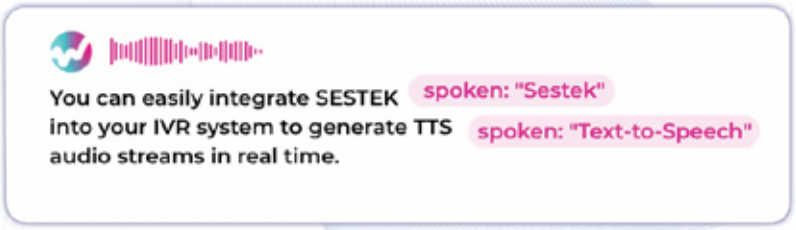
SESTEK
Text-to-Speech

5. Custom Abbreviations



It enables businesses to define their own abbreviations and pronunciations for specific words, ensuring that the resulting audio content accurately represents their brand and messaging. This level of control enables businesses to create highly customized and engaging audio content that resonates with their audience. By leveraging our custom abbreviations feature, businesses can ensure that the resulting audio content is free from mispronunciations and other errors that can undermine the credibility of their message.





6. Flexible Deployment

On-premise Deployment

With the on-premise deployment of SESTEK Text-to-Speech, businesses can have complete control over their data, ensuring maximum security and privacy. The software can be installed on your own hardware, allowing for faster data processing and reduced latency. Additionally, SESTEK Text-to-Speech offers easy integration with existing software and systems, making it a seamless addition to your business's infrastructure.

Cloud

SESTEK Text-to-Speech is also available as SaaS, delivering all the same capabilities without the burden of implementing and running it within your own team and environment. Key advantages include:

Scalability: Increase or decrease computing resources on demand
Reduced maintenance: No hardware or technical infrastructure to manage
Accessibility: Available from any device with an internet connection

© 2025 SESTEK. All rights reserved.

SESTEK
Text-to-Speech

7. Integration Types

REST (Representational State Transfer)

SESTEK Text-to-Speech offers a API that enables businesses to easily integrate our technology into their applications and workflows. With the REST integration, businesses can send HTTP requests to our API to convert text to speech and receive the resulting audio content in a variety of formats. The REST integration is easy to use and offers businesses complete control over the resulting audio content. The REST integration is a popular choice for businesses that require flexibility and ease of use in their integrations.

MRCP (Media Resource Control Protocol)

Knovvu Text-to-Speech also offers MRCP integration, which enables businesses to use our technology through a standardized protocol that supports both local and remote media servers. It is the recommended integration method for IVR systems. The MRCP integration offers businesses high-level control over the speech synthesis process and enables them to use our technology across a wide range of platforms and devices.



© 2025 SESTEK. All rights reserved.

SESTEK
Text-to-Speech

Bank Gives Voice to its ATMs for Visually Impaired Customers

Using Text-to-Speech technology, one of the most reputable banks in Turkey, Yapı Kredi introduced voice-enabled ATMs to serve visually impaired customers.

1st

First bank to introduce voice-enabled ATMs

↑100%

Increased accessibility for visually impaired customers

Contributed to brand reputation as innovative and customer-focused



© 2025 SESTEK. All rights reserved.

SESTEK

Conversational Automation Company

SESTEK is a conversational automation company helping organizations with conversational solutions to be data-driven, increase efficiency and deliver better experiences for their customers. Sestek's AI-powered solutions are built on text-to-speech, speech recognition, natural language processing and voice biometrics technologies.

SESTEK is a part of **UNIFONIC**

sales@sestek.com

www.sestek.com

/sestek

© 2025 SESTEK. All rights reserved.