SESTEK

Conversational Automation Company

SPEECH RECOGNITION

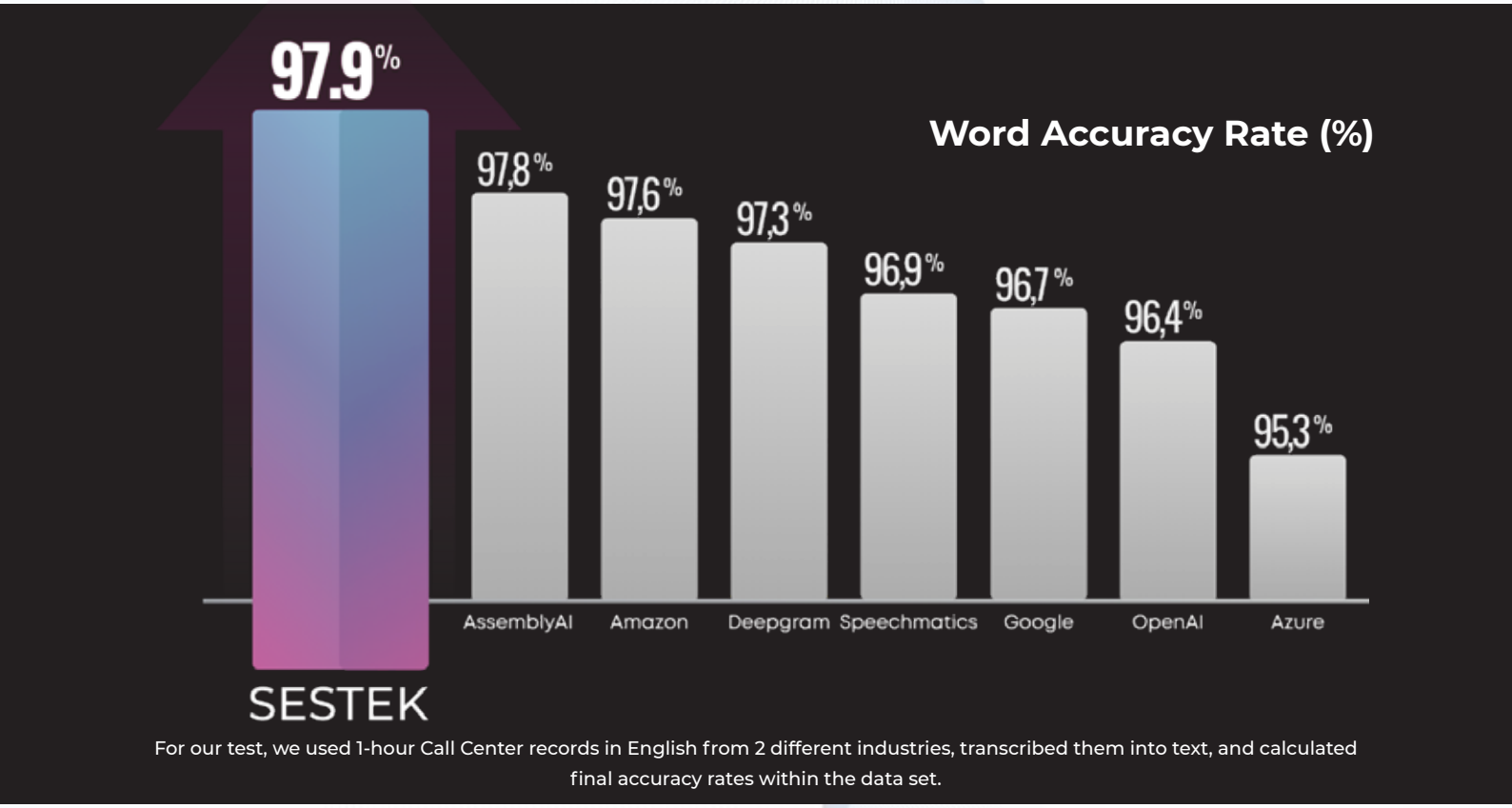
DATASHEET



SESTEK
Speech Recognition

Market-Leading Accuracy Rate of ~98%

Our fully in-house developed SR technology can capture and interpret human speech with ~98% accuracy. Speech Recognition learns your product names, terms, and accents, so every product in the Agentic CX Suite understands your customers this well.



For our test, we used 1-hour Call Center records in English from 2 different industries, transcribed them into text, and calculated final accuracy rates within the data set.

BENEFITS

LEARNS YOURVOCABULARY

Fix common terms directly in the interface, or let our R&D team retrain the model on your own recordings for deeper, lasting accuracy.

BEST OF BOTH

Proprietary models cover the languages your business runs on with full fine-tuning control. Best-in-class external engines extend that to 100+ languages, all through the same platform.

PREDICTABLE COST

One platform, one price, whether it's SESTEK's own engine or an integrated external model, giving you a cost structure you can plan around with confidence.

© 2023 SESTEK. All rights reserved.

KNOVVU
Speech Recognition

1. Features

Our SR technology features help businesses improve the customer experience, streamline their operations and make decisions more informed.

Voice Activity Detection

Detect when a person is speaking to optimize speech recognition technology and improve accuracy.

Emotion Detection

Understand the emotional context of a conversation to respond appropriately and improve customer experience.

Punctuation

Improve the readability and accuracy of transcriptions for better communication and decision-making.

Language Identification

Automatically detect the language spoken by a speaker in an audio stream for effective communication with a diverse customer base.

Inverse Text Normalization

Accurately transcribe colloquial or non-standard speech for improved accuracy and customer satisfaction.

Speaker Diarization

Identify individual speakers in a conversation for improved data analysis and operational efficiency.

Profanity Filtering

Ensure that transcriptions are professional and appropriate for improved brand image and customer satisfaction.

Sentiment Analysis

Understand the emotional context of a conversation to respond appropriately and improve customer experience.

© 2023 SESTEK. All rights reserved.

knovvu
Speech Recognition

2. Language Support

SESTEK SR works where your customers are. Proprietary models deliver deep accuracy for your core languages, and best-in-class external engines extend coverage to **100+ languages, all through one API.**

72%
WITHOUT CUSTOMIZATION
No call center data

real customer data
>>>>>>

85%
WITH CUSTOMIZATION
Trained on call center audio

Built on real calls
A low-resource language with no pior call center data reached 85%+ after training o real calls.

Based on real customer deploment.

Monolingual Models

Afrikaans · Albanian · Amharic · Arabic · Armenian · Assamese · Azerbaijani · Bashkir · Basque · Belarusian · Bengali · Bosnian · Breton · Bulgarian · Burmese · Cantonese · Catalan · Croatian · Czech · Danish · Dutch · English · Estonian · Faroese · Persian · Filipino · Finnish · Flemish · French · Galician · Georgian · German · Greek · Gujarati · Haitian Creole · Hawaiian · Hebrew · Hindi · Hungarian · Icelandic · Indonesian · Italian · Japanese · Javanese · Kannada · Kashmiri · Kazakh · Khmer · Korean · Kurdish · Kyrgyz · Lao · Latin · Latvian · Lingala · Lithuanian · Luxembourgish · Macedonian · Malagasy · Malay · Malayalam · Maltese · Mandarin Chinese · Maori · Marathi · Mongolian · Nepali · Norwegian · Occitan · Odia · Pashto · Polish · Portuguese · Punjabi · Romanian · Russian · Sanskrit · Serbian · Shona · Sindhi · Sinhala · Slovak · Slovenian · Somali · Spanish · Sundanese · Swahili · Swedish · Tagalog · Tamil · Tatar · Telugu · Thai · Tibetan · Turkish · Turkmen · Ukrainian · Uighur · Urdu · Uzbek · Vietnamese · Welsh · Yiddish · Yoruba

Multilingual Models

English+Turkish · Azerbaijani+Russian · Dutch+French · Arabic+English · English+Spanish · English+Malay · Latvian+Russian

· Arabic+English+Turkish · English+French+Turkish · Arabic+English+French+Urdu

· Mandarin+Tamil+Malay+English

· English+French+Portuguese+Spanish

· English+French+Portuguese+Spanish+Dutch+German+Italian

· English+Arabic+French+Spanish

· English+Turkish+Arabic+Russian+French+Spanish

· Arabic+Danish+Dutch+English+Finnish+French+German+Hindi+Italian+Latvian+Mandarin +Norwegian+Portuguese+Russian+Spanish+Swedish+Tagalog+Turkish+Urdu

knovvu
Speech Recognition

3. Language Models

We offer two types of language models: Hybrid models and End-to-End (E2E) models. Our E2E models are our new generation models, providing numerous advantages over traditional hybrid models.

1. Hybrid Models

Hybrid models use a combination of acoustic and language models to transcribe speech to text. While hybrid models have been used in the industry for many years, they require manual feature engineering and are often limited in their ability to handle diverse accents and speaking styles. Hybrid models can also be more resource-intensive and may require significant time and technical expertise to deploy.

2. End-to-End Models

End-to-End (E2E) models are a new generation of language models that can transcribe speech to text with high accuracy and speed. E2E models utilize deep learning techniques and advanced algorithms to eliminate the need for manual feature engineering, resulting in faster and more accurate transcription. Additionally, E2E models can handle a wide variety of accents and speaking styles, making them ideal for businesses that operate in diverse markets.

Advantages of Next Generation E2E Models

At Knovvu, we believe that E2E models provide numerous advantages over traditional hybrid models.

1

E2E models can transcribe speech to text faster and more accurately than hybrid models, resulting in improved efficiency and customer satisfaction.



2

Additionally, E2E models are highly scalable, enabling businesses to handle large volumes of transcription requests with ease.



3

Finally, E2E models can provide businesses with improved accuracy and reliability, resulting in fewer errors and improved decision-making.



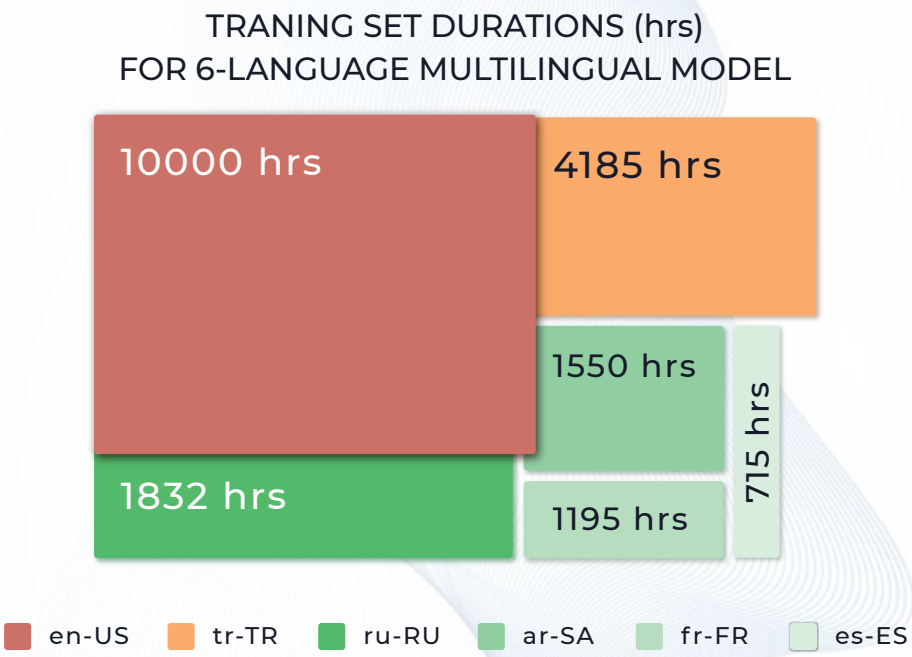
© 2023 SESTEK. All rights reserved.

knovvu
Speech Recognition

3. Multi-Language Models

With newly developed multi-language approach, single model can identify and transcribe multiple languages.
6-language, 4-language and bilingual models can be offered depending on business needs.

- **4-language multilingual model:** English, Arabic, French, Spanish
- **6-language multilingual model:** English, Turkish, Arabic, Russian, French, Spanish
- **Bilingual models:** Arabic-English, English-French, French-Dutch, English-Spanish





© 2023 SESTEK. All rights reserved.

knovvu
Speech Recognition

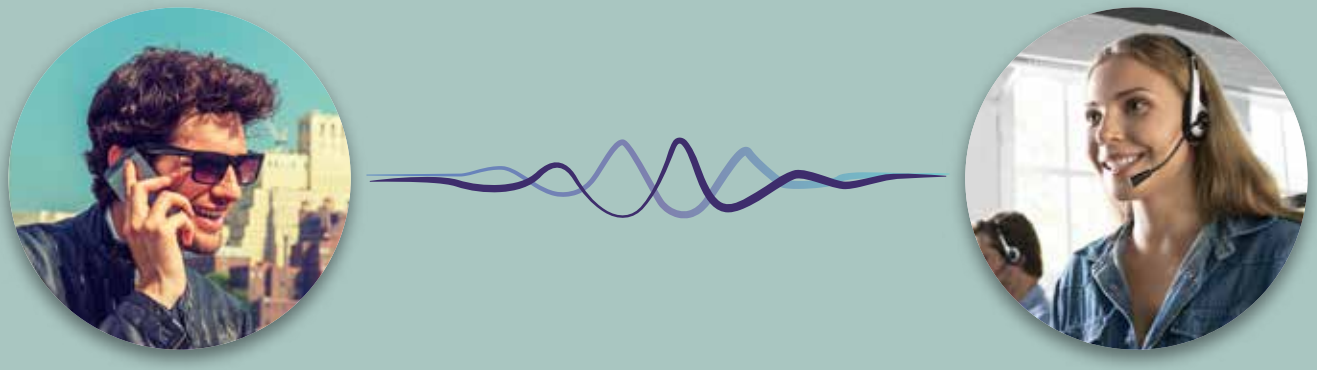
4. Transcription Types

Batch Transcription

You can transcribe pre-recorded media files whenever it's convenient for you. Save time and resources by automating the transcription process. Improve the accuracy of transcriptions by providing ample time for the software to analyze and transcribe speech.

Real-time Transcription

You can transcribe in real-time and collect actionable info in industries such as healthcare, finance, and customer service, where timely and accurate communication is critical. With Knovvu Speech Recognition's real-time transcription, businesses can improve their response times, reduce errors, and streamline their operations.



00:00:00

Hello. I am Jen from ABC Company. How may I help you today?

00:00:05

Hello. I'm planning a trip to Europe next week and I want to learn about your data roaming plans.

00:00:13

Absolutely. We have a campaign for our international traveling customers, and I can offer our Premium Roma plan. You only pay $10 on top of your current plan.

00:00:28

Thanks for the offer. Please upgrade my plan to Premium perfect.

00:00:32

Welcome to our premium club.

© 2023 SESTEK. All rights reserved.

knovvu
Speech Recognition

5. Easy Deployment

On-premise Deployment

With on-premise deployment of Knovvu Speech Recognition, 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, Knovvu Speech Recognition offers easy integration with existing software and systems, making it a seamless addition to your business's infrastructure.

Cloud

This feature offers scalability, enabling businesses to freely adjust their computer resources as needed. All the advantages of the Knovvu Speech Recognition are provided by our SaaS, but without the difficulties of implementing it within your own team and environment. It offers scalability, enabling businesses to increase or decrease their computing resources as needed. This deployment option eliminates the need for businesses to maintain your own hardware, reducing maintenance costs and minimizing the need for technical expertise. With Knovvu Speech Recognition's cloud-based solution, businesses can easily access the software from anywhere, on any device with an internet connection.



© 2023 SESTEK. All rights reserved.

knovvu
Speech Recognition

6. Integration Types

REST (Representational State Transfer)

The fastest integration method for your application. Businesses can easily integrate speech recognition into their existing applications and systems. REST integration is highly versatile and can be used in a variety of industries, including healthcare, finance, and customer service.



MRCP (Media Resource Control Protocol)

It is the recommended integration method for IVR systems. MRCP SR Server is a standalone service, this service can be integrated to most IVR systems (e.g., Avaya, Genesys, Cisco) through simple configuration. MRCP integration provides businesses with a more efficient and cost-effective way of handling large volumes of calls.

WebSocket

It provides full-duplex communication channels over a single TCP connection. It allows for partial recognition results to be displayed in the client app until the full sentence is uttered, so it is more suitable for mobile applications especially. Businesses can easily integrate speech recognition into their web applications such as chatbots and virtual assistants. This integration is highly responsive and enables businesses to provide real-time feedback to users.

© 2023 SESTEK. All rights reserved.

knovvu
Speech Recognition

Unilever boosts sales after listening to customers

> "For this project, we decided to work with SESTEK, known as the best company in their line of work. Especially, their expertise in voice technologies attracted our attention."
>
> — Unilever

THE CUSTOMER

Unilever is a global consumer goods company, offering beauty and personal care products, food and refreshments, and home care products.

THE CHALLENGE

Unilever aimed to collect customer insights and identify their preferences for its Domestos bleach product. However, they had to analyze a vast amount of customer calls at its call center.

THE SOLUTION

Using Sestek's Speech Recognition technology allowed Unilever Turkey to analyze large volumes of customer calls quickly and accurately, enabling Unilever to make data-driven decisions that improve customer satisfaction and drive sales.



120K
of calls converted into text and analyzed

6
New product variants introduced to the market

↑48%
Increase in sales growth

© 2023 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

© 2023 SESTEK. All rights reserved.