

yobe

Product Brief

VIENNA

VOICE INTELLIGENCE EXTRACTION FOR NEURAL NETWORK ANALYTICS



OVERVIEW

VIENNA is a cutting-edge voice extraction solution engineered to capture and transcribe targeted speech in everyday noisy environments with exceptional accuracy. By overcoming the challenge of extracting clean identified speech, VIENNA unlocks limitless possibilities for various voice applications. Using Yobe's unique blend of voice biometrics, intelligent voice activity detection, and large language models (LLMs), VIENNA ensures accurate voice for a customer's analytics. Voice-captured data allows for the extraction of valuable insights from natural, unstructured conversations, providing a richer context compared to text data alone. This integration enhances decision-making processes, improves customer experiences, and drives innovation by uncovering deeper insights from voice interactions. Furthermore, the automation of voice data analysis can lead to significant time and cost savings for organizations, allowing them to allocate resources more efficiently.

Accessible via a simple API call, the solution is input device-agnostic and fully scalable.

Target Audience:

Corporate Sector, Consumer Service Centers, Industrial Automation, Hospitality & Tourism, Gaming & Virtual Reality, Automotive, Retail & E-commerce, Government & Public Sector, DoD & Surveillance, Healthcare, Virtual Assistant Developers, Sports & Fitness, Energy & Utilities, Entertainment & Media, Transportation & Logistics, and Education.

KEY FEATURES

INTELLIGENT VOICE ACTIVITY DETECTOR:

- Uses advanced biometrics to activate the input channel only when the target speaker is talking.
- Significantly reduces unwanted Automatic Speech Recognition (ASR) data charges by filtering out irrelevant noise and non-target speech.

VOICE BIOMETRICS TEMPLATE MATCHING:

- Personalizes the user experience through voice template matching.
- Enables efficient voice data tagging for easy identification and retrieval.

ENHANCED SPEECH-TO-TEXT CAPABILITIES:

- Delivers robust transcription in environments with cross-talk and other noise sources.
- Ensures high accuracy, making it suitable for a variety of demanding applications.

EMOTION DETECTION:

- Detects emotions such as anger or frustration within voice signals.
- Provides platform alerts, enabling timely interventions based on emotional cues.

DATA ANALYTICS USING LLMs:

- Analyzes unstructured voice data using Large Language Models (LLMs).
- Offers actionable insights, helping businesses make informed decisions based on voice interactions.

PRODUCT BENEFITS

UNMATCHED ACCURACY:

Superior voice extraction and transcription capabilities in noisy environments.

COST EFFICIENCY:

Reduces ASR data charges with intelligent voice activity detection.

PERSONALIZATION:

Enhances user experience with voice biometrics and template matching.

EMOTIONAL INTELLIGENCE:

Alerts based on detected emotions for better customer engagement and response.

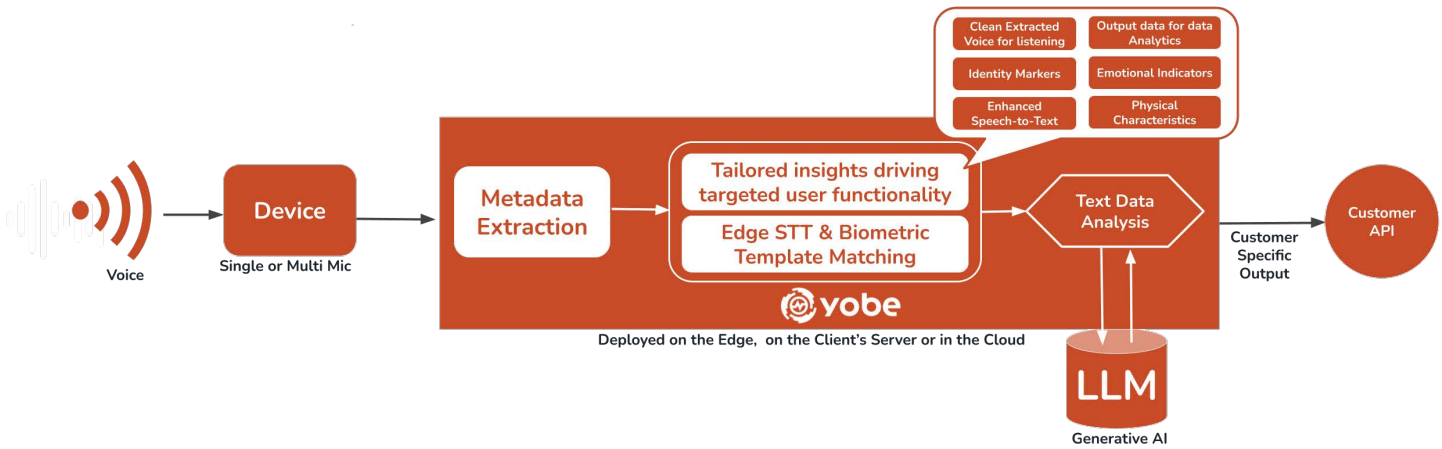
ACTIONABLE INSIGHTS:

Transforms unstructured voice data into valuable business intelligence.

Use Case Examples	Benefit 1	Benefit 2	Benefit 3
MANUFACTURING & MAINTENANCE	Hands-Free Operation: Enables workers to perform tasks without needing to use their hands for data entry, increasing efficiency.	Condition Monitoring: Collects real-time data on equipment conditions through voice inputs, allowing for predictive maintenance and reducing downtime.	On-the-Job Training: Provides real-time, voice-activated training and guidance, helping new employees learn procedures quickly.
TRAVEL & HOSPITALITY	Language Translation: Ensures travelers can communicate effectively with staff, reducing misunderstandings.	Guest Experience: Analyzing guest interactions to enhance service quality and personalize guest experiences.	Operational Savings: Reduces operational costs by automating routine tasks and improving efficiency through voice interfaces.
CUSTOMER SERVICE & SUPPORT	Sentiment Analysis: Real-time analysis of customer sentiment during calls to identify dissatisfaction and potential churn risks.	Issue Trending: Identifying common issues and trends from customer interactions to improve products and services.	Agent Performance: Monitoring and analyzing agent interactions to provide feedback and training for improved customer service.
CORPORATE COMMUNICATIONS	Automated Note-Taking: Automatically captures meeting notes and key points, reducing the need for manual note-taking.	Meeting Summaries: Generates concise summaries of meetings and discussions, ensuring all participants are aligned and have access to the same information.	Knowledge Sharing: Captures and analyzes voice interactions to share best practices and organizational knowledge.
HEALTHCARE	Patient Feedback: Collecting patient feedback through voice interactions to improve care quality and patient experience.	Audit Trails: Creates detailed, voice-logged audit trails for various processes, improving transparency and accountability.	Expert Consultations: Facilitates easy access to remote experts through voice interfaces, enabling instant consultation and problem-solving.
RETAIL	Customer Preferences: Gathering insights on customer preferences and buying behavior through voice interactions.	Effortless Shopping: Enables hands-free shopping, allowing customers to search for products, check prices, and make purchases using voice commands.	Upselling Opportunities: Identifies upselling opportunities by analyzing customer interactions and preferences.

Assessability & Scalability	
API Intergration	The solution can be accessed via an easy-to-use API call.
Device Agnostic	Ensures compatibility with any input device, making it fully scalable and adaptable to various platforms.

Metric	Results
Processing Latency	less than 500 ms
Rate of Environmental Adaptation	Every 128ms
ASR/STT Accuracy Enhancement	From 70% to 90%+ in -SNRs
SNR Improvement	15 dB+



Microphone Capture Technology Stack

Yobe's solution comes in single and multi mic configurations (interleaved) . The format can be either 8 or 16-bit PCM data or normalized data.

Yobe can provide real-time audio capture support for (TBD).

Frontend: JavaScript (React or Angular)
 Backend: Python, Node.js
 Machine Learning: TensorFlow, PyTorch for voice recognition, speech-to-text, and emotion detection models
 Analytics: Integration with LLMs and BI tools like Tableau or Power BI

Next Steps

VIENNA sets a new standard in voice data processing by combining advanced voice extraction, biometrics, and powerful analytics. It is an essential tool for organizations looking to enhance security, improve transcription accuracy, and gain valuable insights from voice data in noisy environments.

Discover how Yobe can revolutionize your voice interaction experience. Integrate our solution today and unlock the potential of accurate speech-to-text transcription in any environment. Visit www.yobeinc.com or contact us at contact.us@yobeinc.com



Enabling the evolution of voice

www.yobeinc.com

contact.us@yobeinc.com