

Deepgram

The **Voice AI** Platform for Business Use Cases

APIs for speech-to-text, text-to-speech, and full speech-to-speech voice agents. Over 200,000+ developers use Deepgram to build voice AI products and features.

Voice Agent API

A single, unified voice-to-voice API for building real-time, context-aware voice agents that deliver natural, responsive conversations. It provides the simplicity developers want and the infrastructure control enterprises need.



"We believe that integrating AI voice agents from Deepgram will be one of the most impactful initiatives for our business operations over the next five years, driving unparalleled efficiency and elevating the quality of our service." – **Doug Cook**, CTO, Jack in the Box

Real-time Performance

Build voice agents that respond instantly, speak naturally, and handle mid-turn interruptions.

Enterprise Scale

Deploy at scale with unmatched compute efficiency

Maximum Controllability

Use your own LLMs or external APIs with full control over orchestration, deployment, and behavior.

Learn more about our
Voice Agent API



Flux

New model

Conversational speech recognition for real-time voice agents with built-in turn detection, natural interruption handling, and ultra-low latency.

Nova-3

High-performance speech-to-text for production transcription with top accuracy, multilingual support, and noise robustness.

Industry-Tuned

Specialized speech-to-text models optimized for industry-specific vocabulary and structure for domains like healthcare, legal, and finance.

Custom

Custom speech-to-text models trained on proprietary or novel datasets for maximum accuracy in edge-case scenarios.

Speech to Text API

Build your apps with world-class speech to text. Powers everything from transcription and analytics to real-time, human-like voice agents:

Accuracy >90%
word recognition rate (1 - WER)

Speed 40x
faster inference time

Flexible Deployment:
On-premises, VPC, Cloud

2x Lower cost
than cloud providers

Learn more about our
Speech to Text API



Text to Speech API

Deepgram's Aura-2 meets the rigorous, real-time, and scalable demands of enterprise environments.

Natural

voices designed for business settings rather than entertainment use cases

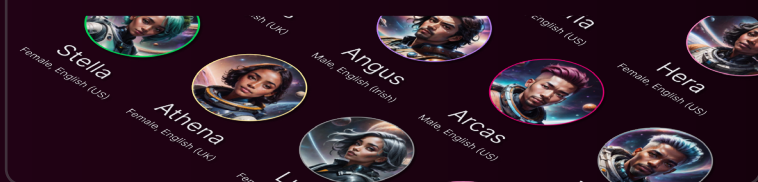
Accurate

pronunciation of names, URLs, dates, and industry-specific terms (e.g. healthcare)

Scalable

high-throughput architecture to handle concurrent interactions at low latency

Choose from a diverse set of voices fine-tuned for conversational use cases.



Learn more about
Aura-2

Aura-2

New model

Aura-2 delivers enterprise-grade TTS with professional, context-aware naturalness, precise pronunciation, and ultra-low latency.

- Deliver authentic, **expressive voices** with optimized pacing, tone, and contextual awareness to enhance engagement and build trust.
- Ensure **precise, consistent pronunciation** of numerals, proper nouns, and industry-specific terminology (e.g., healthcare, finance, legal) for clear and reliable communication.
- Achieve **ultra-low latency conversational experiences** while ensuring affordable, high-throughput enterprise scaling.

Features:

- **Summarization:** Capture conversation essence to improve workflows.
- **Sentiment Analysis:** Detect sentiment to enhance experiences and analyze trends.
- **Intent Recognition:** Track speaker intent to enable smarter, satisfaction-driven actions.
- **Topic Detection:** Pinpoint key topics to optimize workflows, and post-call analysis.

Audio Intelligence

Unlike costly, slow language models, our Audio Intelligence models are lightweight, purpose-driven, and fine-tuned on domain and task-specific conversational data sets. The result: precise accuracy on specialized topics, fast speed, and low costs for high-performance applications.

Learn more about our
Audio Intelligence API



About Deepgram

Deepgram is the leading voice AI platform for business use cases, offering speech-to-text, text-to-speech, or full speech-to-speech capabilities—all powered by our enterprise-grade runtime. Over 200,000 developers build with Deepgram's voice-native foundational models - accessed through cloud APIs or as self-hosted / on-premises APIs – due to our unmatched accuracy, latency and pricing. With 50,000+ years of audio processed and over 1 trillion words transcribed, no one understands voice better than Deepgram. Learn more at deepgram.com

Deepgram

Voice AI Platform for Business Use Cases