

Deepgram

The **Voice AI** Platform for Developers

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.

Speech to Text API

Build your apps with world-class speech to text. Everything developers need to build with confidence and ship faster.
Fast, accurate and cost effective:

Accuracy 30%
lower word error rate (WER)¹

Cost 3-5x
lower price

Speed 40x
faster inference time

Deploy Anywhere:
On-premises, VPC, Cloud



Learn more about our
Speech to Text API

Nova-2

- Multi-language support
- Optimized for multiple use cases
- Starting at \$0.0036/min (with Growth plan)

Nova-2 Medical

- 43% higher accuracy compared to alternatives
- Secure and HIPAA Compliant
- Accurately transcribe healthcare terms (symptoms, medications, etc.)

Custom

- Enterprise-ready high-throughput models
- Unlock maximum model efficiency with your data
- Designed to adapt to your specific use cases

Aura

Deepgram Aura is a natural-sounding, high-throughput text-to-speech model for real-time voicebots and conversational AI applications.

- Industry-leading low latency with batch processing and real-time TTS.
- Diverse, fine-tuned male and female voices for conversational use cases.
- Enterprise Scale: Faster, more compute-efficient, and perfect for large-scale conversational AI.

Text to Speech API

Bring your applications to life with Aura's natural, responsive voices for conversational use cases from AI agents to IVRs.

Quality

Human-like tone, rhythm, and emotion

Speed

Less than 250 ms latency

Scale

Optimized for high-throughput apps

Learn more about our
Text to Speech API



Voice Agent API

A unified voice-to-voice API that enables natural-sounding conversations between humans and machines. With one powerful API, create LLM-powered AI agents that listen, think, and speak with the same intelligence and emotive quality that a person can.



"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 real-time agents that speak naturally, with human-like voices and seamless responses.

Enterprise Scale

Our optimized API delivers unmatched compute efficiency for large-scale AI voice agents.

Maximum Controllability

Use your LLM or choose open/closed-source options with flexible, secure deployments.

Learn more about our
Voice Agent API



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 designed for developers building voice-first products using speech-to-text, text-to-speech, or speech-to-speech APIs. Over 200,000 developers build with Deepgram's voice-native foundational models - accessed through APIs or as self-managed software - due to our unmatched accuracy, latency and pricing. Contact us to learn more at deepgram.com/contact-us.

Deepgram

Essential Building Blocks for Voice AI