

# Deepgram

## The Most Powerful Speech-to-Text API

Convert speech to text with unmatched accuracy, ultra-low latency, and enterprise scalability. 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

Cost

**2x** lower price

Multilingual

**30+** languages and dialects to choose from

## Speech Recognition Models

Deepgram models power everything from real-time conversations to domain-specific transcription, with options for speed, accuracy, and full customization.

### 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

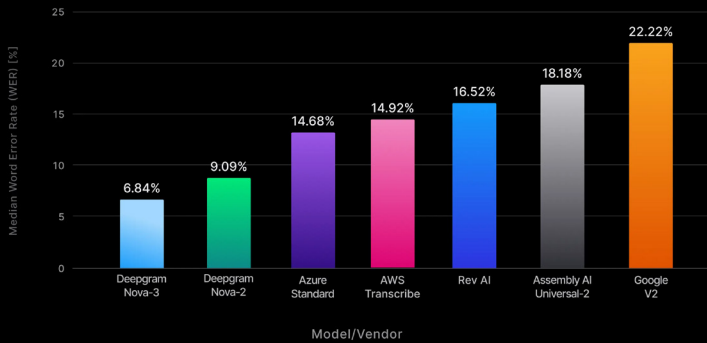
Custom speech-to-text models trained on proprietary or novel datasets for maximum accuracy in edge-case scenarios.



## The Most Accurate STT Model for Pre-recorded Audio

- Median WER of 5.26%, a 47.4% improvement over the next-best competitor
- 70% relative WER improvement over Whisper (Large)

Word Error Rate (WER) - Streaming



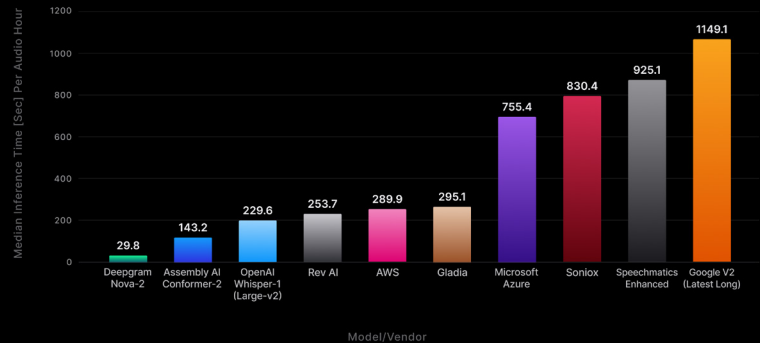
## The Most Accurate STT Model for Streaming Audio

- Median WER of 6.84%
- 54.2% improvement over the next-best competitor

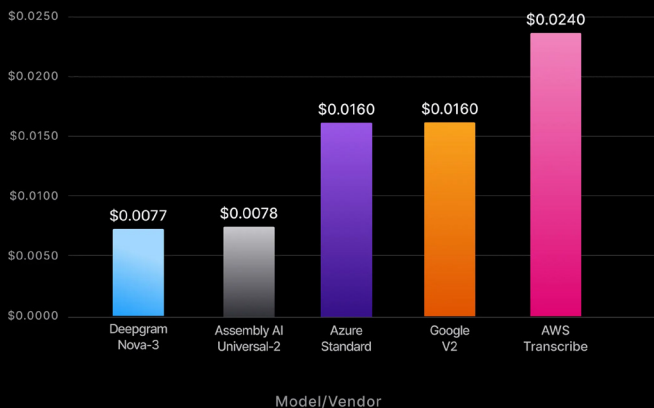
## Hands-Down the Fastest Models<sup>2</sup>

- Median inference time of 29.8 seconds per audio hour
- 5 to 40 times faster than comparable models offering diarization

2. Nova-3 demonstrates comparable latency to Nova-2.



Benchmark - Pricing (Streaming)



## The Most Affordable STT Model in the Market

- Deepgram Nova-3 starts at just \$0.0077 per minute of streaming audio
- Over 2x more affordable than cloud providers

### 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. Contact us to learn more at [deepgram.com](https://deepgram.com)

**Deepgram**

The Voice AI Platform for Business Use Cases