# Open-source AI voice model Dia captures raw human emotions including screaming and laughter



A new open-source text-to-speech (TTS) model called Dia is pushing the boundaries of AI-generated voices by incorporating a wide range of emotional expressions, including realistic screaming, laughter, coughing, and throat-clearing. Developed by a small team at Nari Labs, Dia aims to capture the emotional depth of human speech, expanding beyond the usual friendly and relaxed tones common in most AI voices.

Unlike typical AI voices that prioritise a smooth, consistent tone designed to sound helpful and cheerful, Dia seeks to replicate the complexities of human emotion during speech. This includes the difficult-to-mimic vocalisations of intense emotions such as yelling and screaming, which require different speech modes rather than simply speaking louder. Dia’s design incorporates an understanding of nonverbal sounds as integral parts of communication rather than incidental noises, delivering nuanced timing, pitch changes, and breath control that contribute to a more authentic listening experience.

One user even demonstrated Dia's capabilities by recreating a famous moment from the Leroy Jenkins World of Warcraft sketch, highlighting its potential for expressive performance in digital entertainment contexts.

The innovation is notable because, while commercial AI voice systems like those from OpenAI, ElevenLabs, Google, and Sesame have advanced in expressing emotion and reactiveness, they typically remain within a spectrum of pleasant and positive demeanors. OpenAI, for example, offers an Advanced Voice Mode capable of different emotional tones, and ElevenLabs adjusts speech based on punctuation and capitalization, but neither captures the raw intensity of, say, a genuine surprised yelp or hearty wheeze of laughter as Dia does. Sesame's voice models similarly excel at sounding reactive and conversational but generally avoid venturing into more dramatic emotional territory.

Dia's creators, two undergraduate students—one currently serving in the military—and operating without external funding, successfully developed the model with the intention of rivaling established AI voice technologies like NotebookLM Podcast, ElevenLabs Studio, and Sesame CSM. Their approach emphasises the importance of emotional realism in AI voices, marking an important step forward in the growing field of emotionally intelligent AI.

The emergence of Dia highlights a broader trend within artificial intelligence towards equipping digital assistants and virtual characters with the ability to not only say the right words but to convey them with appropriate emotional inflection. This capability could prove valuable in various applications, such as customer support bots that sound genuinely apologetic, educational tools that sound encouraging, and in-game characters whose emotional reactions enhance player immersion.

However, the enhanced emotional expressiveness also raises questions about the potential for AI voices to be more persuasive, and possibly manipulative, as convincingly emotive speech becomes an AI tool. The power to emulate human emotions convincingly in voice synthesis could have wide-ranging implications across media, entertainment, customer service, and beyond.

Despite these considerations, Dia’s development opens new creative possibilities. For instance, it could add dramatic impact to storytelling by not just reading a ghost story but effectively performing it—complete with spine-chilling screams and authentic laughter.

The Tech Radar report underscores Dia as a remarkable achievement from a small, resource-limited team, signalling a significant advancement in how AI voices might evolve to interact with people on a deeper emotional level in the coming years.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://www.segmind.com/models/dia> - This page describes Dia as a 1.6 billion parameter advanced open-weights text-to-speech model by Nari Labs that incorporates natural speech, emotions, and nonverbal cues such as laughter, coughing, and throat-clearing, supporting the claim that Dia pushes the boundaries of AI-generated voices by including a wide range of emotional expressions.
2. <https://fal.ai/models/fal-ai/dia-tts> - Fal.ai details Dia’s capability to directly generate realistic dialogue from transcripts with audio conditioning that enables emotion and tone control, confirming Dia's design to replicate complex human emotions and nonverbal vocalizations for authentic listening experiences.
3. <https://www.ainvest.com/news/nari-labs-dia-1-6b-outperforms-competitors-emotional-speech-2504/> - This article highlights that Dia-1.6B outperforms competitors like ElevenLabs and Sesame in emotional speech, including realistic laughing, coughing, and screaming, and discusses the model's nuanced handling of emotionally intense vocalizations, corroborating the article's points on Dia's advancement beyond typical AI voice systems.
4. <https://github.com/nari-labs/dia> - The official GitHub repository for Dia shows the open-source nature of the model, its ability to generate highly realistic dialogue with emotional and tone controls via audio conditioning, supporting the claim about the model being developed by a small team at Nari Labs with a community-driven approach.
5. <https://www.youtube.com/watch?v=fyA7NT1xTdQ> - This YouTube video demonstrates Dia’s expressive capabilities including emotional speech and nonverbal sounds like laughter and gasps, and compares it with premium commercial AI voice tools, substantiating the example of a user showcasing Dia’s creative and expressive performance potential.
6. <https://openai.com/blog/advanced-voice-modes> - OpenAI's blog on Advanced Voice Modes explains how commercial voice AIs like OpenAI's models provide emotional tones but mostly within pleasant and positive ranges, supporting the claim that while established systems offer emotional nuance, they do not capture the raw intensity of emotional vocalizations as Dia does.
7. <https://www.techradar.com/computing/artificial-intelligence/the-ai-that-cried-aaaaaahhh> - Please view link - unable to able to access data