

Meta unveils Voicebox, an AI model for audio processing for creators

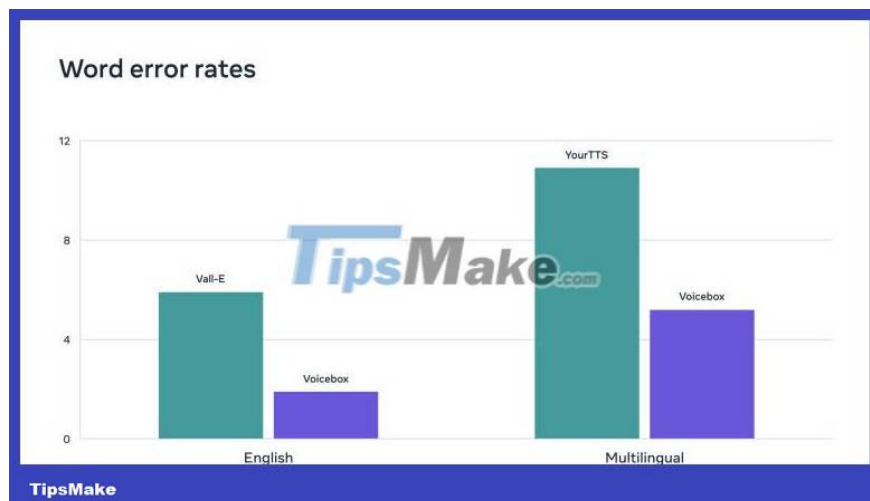
After countless rumors, Meta has finally officially announced its latest generation AI model called Voicebox.

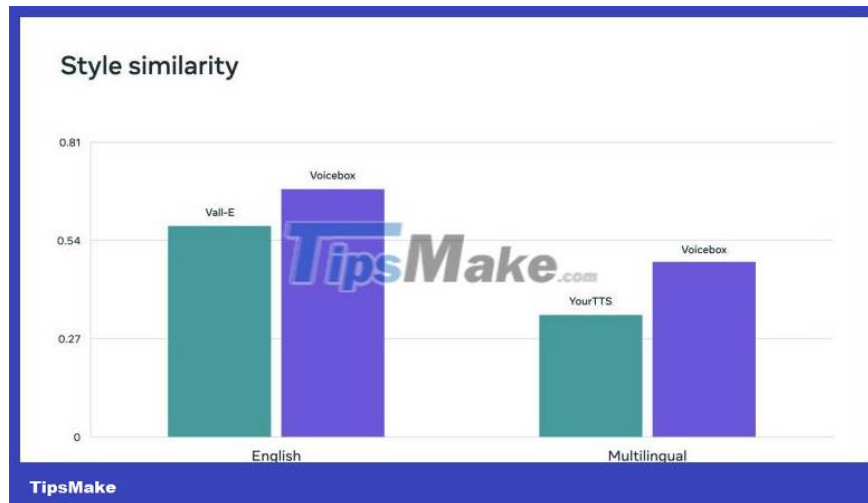
After countless rumors, Meta has finally officially announced its latest generation AI model called Voicebox. This model is designed to help content creators efficiently handle speech generation-related tasks such as audio editing, sampling, and stylization, even if it hasn't been specifically trained through in-context data.

Meta confidently asserts that this new AI model will 'benefit many people around the world', not just in the field of content creation. For example, helping visually impaired people hear text messages by voice, as well as allowing people to speak foreign languages in their own voice.

Voicebox itself is also touted as being able to both create high-quality audio clips and edit pre-recorded audio clips to remove unwanted interruptions, such as car horns, while preserving the same multilingual content and style, (generating speech in six different languages). Future developments that have been planned by Meta for the model include providing natural voices for visual assistants or characters in games in the metaverse.

Meta also compared Voicebox to other audio-enabled AI models currently on the market, namely key competitors like Vall-E and YourTTS. Overall, Meta's model is advanced and outperforms the competition when it comes to Word error rates and Style similarities.





Voicebox is built on Flow Matching. This is Meta's latest non-regressive generalization model, which can handle highly indeterminate mapping between text and speech. This allows Voicebox to learn from different types of speech data without having to carefully label them, giving them access to a wider variety of training data and at scale. To date, Voicebox has trained over 50,000 hours of speech recordings and transcripts from audiobooks in English, French, Spanish, German, Polish and Portuguese.

While this technology could usher in a new era of AI in audio processing, Meta acknowledges that it can bring the potential for abuse and unintended harm. In the research paper that Meta shared about Voicebox will include details on how the company built a highly effective classifier that can distinguish between authentic voices and voices generated by Voicebox.

Meta will not make the AI ??Voicebox program available to the public for use, nor will the source code be released, at least for the time being.

You finished reading the article "**Meta unveils Voicebox, an AI model for audio processing for creators**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.