

Introducing SAM Audio: Segment Sound Using Text, Visual, and Time Prompts

December 16, 2025

Takeaways

- **SAM Audio** is the first unified AI model that can segment sound from complex audio mixtures using text, visual, and time span prompts.
- This technology has the potential to transform audio and video editing and drive innovation in areas like music, podcasting, television, film, scientific research, accessibility, and more.
- You can download SAM Audio or explore its capabilities in the Segment Anything Playground today.



Video Source: [View Header Video](#)

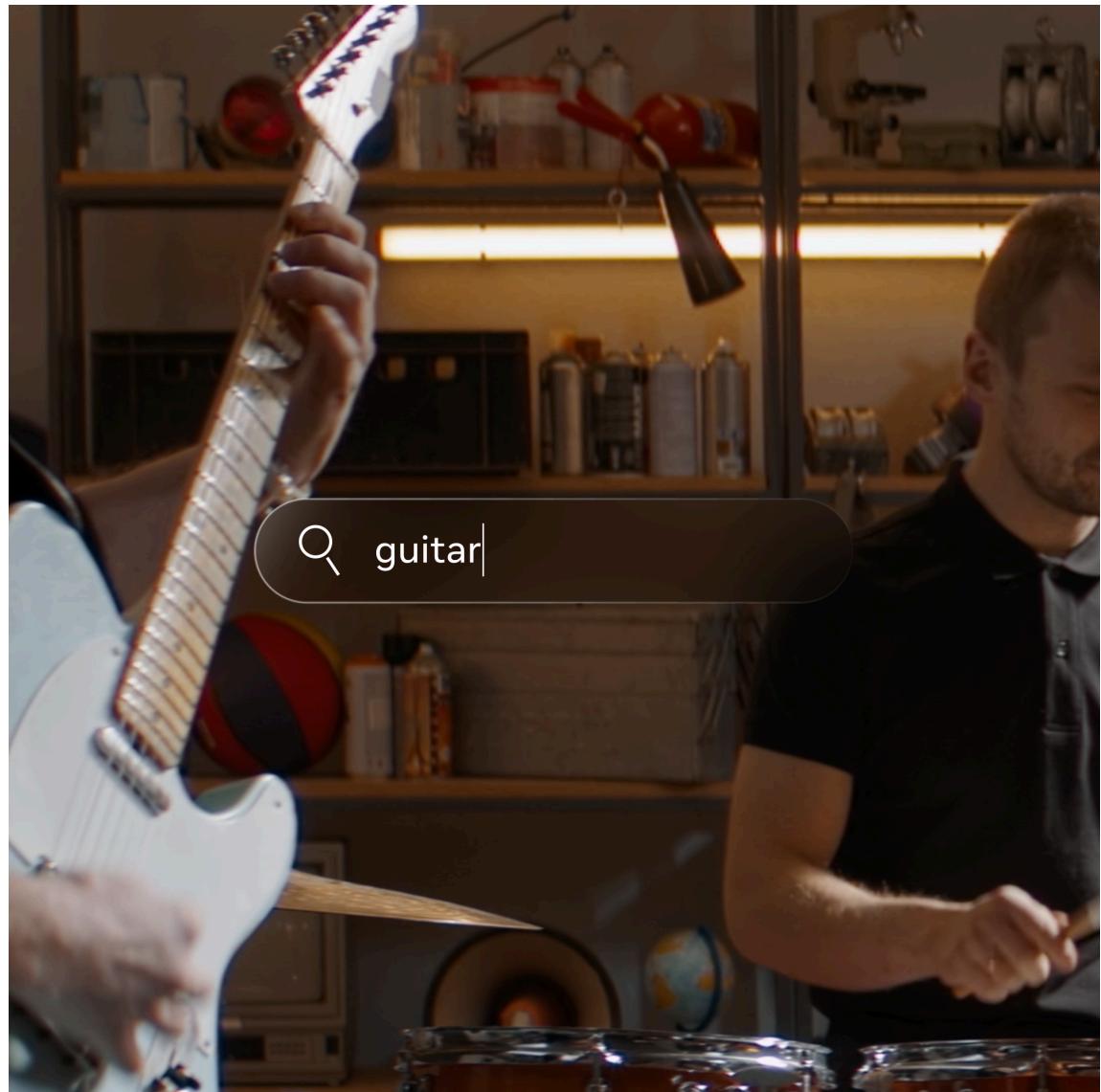
Today, we're introducing **SAM Audio**, a state-of-the-art AI model that enables you to segment sound. Imagine recording a video of your favorite band and isolating the guitar or vocals with a single click, using text prompts to filter traffic noise from a video filmed outside, or removing the sound of a dog barking from your entire podcast recording. SAM Audio, the latest addition to our Segment Anything collection, transforms audio processing by making it easy to isolate any sound from complex audio mixtures using text, visual, and time span prompts.

This intuitive approach mirrors how people naturally engage with sound, making professional-grade audio separation more accessible and easier than ever before. SAM Audio has the potential to transform audio and video editing and drive innovation in areas like music, podcasting, television, film, scientific research, accessibility, and more.

Until now, audio segmentation and editing has been a fragmented space, with a variety of tools designed for single-purpose use cases. As a unified model, SAM Audio is the first to support use cases that match how people naturally think about audio, and achieves cutting-edge performance across diverse, real-world scenarios. SAM Audio supports three kinds of prompts:

Text prompting

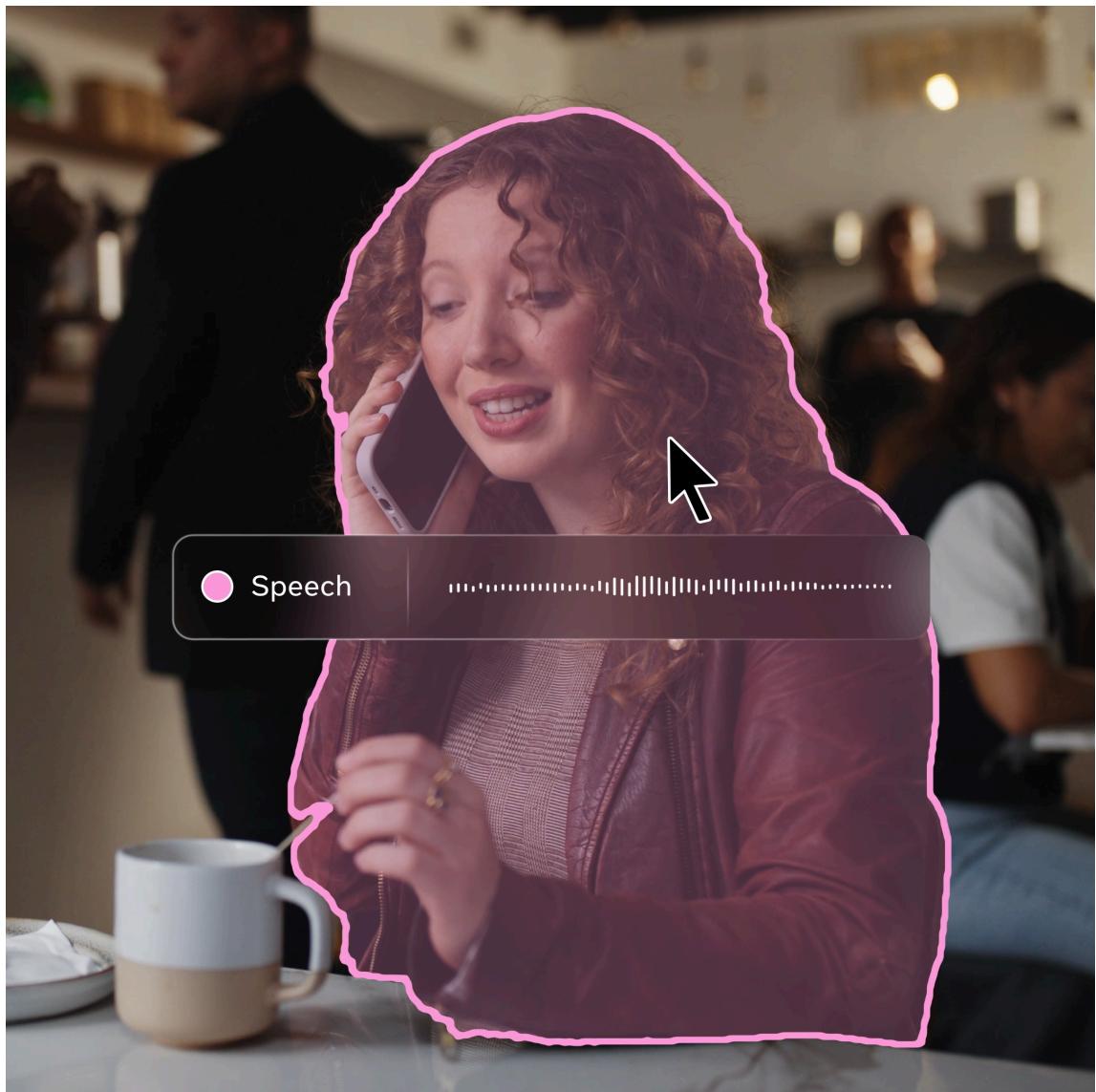
Type “dog barking” or “singing voice” to extract specific sounds.



Video Source: [View Text Prompt Demo](#)

Visual prompting

Click on the person or object in the video that's making a sound to isolate their audio.



| **Video Source:** [View Visual Prompt Demo](#)

Span prompting

An industry first, this method lets you mark time segments where target audio occurs.



Video Source: [View Span Prompt Demo](#)

These prompting methods can be used alone or in any combination, giving you precise and intuitive control over how audio is separated. We see so many potential use cases, including sound isolation, noise filtering, and more to help people bring their creative visions to life, and we're already using SAM Audio to help build the next generation of creative media tools.

You can try SAM Audio in the Segment Anything Playground, our new platform that enables anyone to try our latest models. Starting today, people can select from our collection of audio and video assets or upload their own to explore the capabilities of SAM Audio. The model is also available for download.

We're excited to bring audio to the Segment Anything collection of models and we believe SAM Audio is the all-around best audio separation model available. Learn more about SAM Audio and try it on the Segment Anything Playground today.