Voxtral TTS: Voice Synthesis and Cloning

Course syllabus

Version 2026-11

Overview

Create custom voices and generate speech with Voxtral TTS: zero-shot cloning, real-time streaming, and optimized audio formats.

Learning objectives

- Create a custom voice via the Voices API from a quality audio sample
- Generate speech and clone a voice on the fly from a reference excerpt
- Stream generated audio in real time to reduce perceived latency
- Control pronunciation and text segmentation to ensure reliable vocal output
- Assemble a complete voice pipeline respecting ethical voice cloning rules

Audience and prerequisites

Developers, engineers, and integrators looking to leverage Mistral AI via API in their applications.

Ability to read Python and JSON, command line basics. A Mistral AI Studio API key to try the examples.

Duration and format

Estimated duration : 1 h 30 min. 100% online self-paced course, accessible 24/7 from a browser. Progress saved, resume anytime. Estimated duration: 1 h 30 min.

Content

- Introduction to TTS (3 lessons)
- Creating Custom Voices (4 lessons)
- Generating Speech (4 lessons)
- Formats and Optimization (3 lessons)

Assessment

Practice exercises offered throughout the lessons. Each lesson ends with its Key takeaways, supporting self-assessment. A final quiz of 5 corrected questions, at the end of the last lesson, validates learning.

Accessibility

self-paced course accessible to people with disabilities: text content compatible with screen readers, keyboard navigation, and compliant contrasts. For any adaptation needs, email hello@corsen.ai.

Price and recognition

Free

Certificate of completion issued by Corsen AI.

Access time

Immediate after account creation and email confirmation.

Corsen AI — Paris, France — SIRET 10297040700015
academy@corsen.ai — https://academy.corsen.ai