OpenAI Audio operations

# OpenAI Audio operations#

Use this operation to generate an audio, or transcribe or translate a recording in OpenAI. Refer to OpenAI for more information on the OpenAI node itself.

## Generate Audio#

Use this operation to create audio from a text prompt.

Enter these parameters:

• Credential to connect with: Create or select an existing OpenAI credential.

• Resource: Select Audio.

• Operation: Select Generate Audio.

• Model: Select the model you want to use to generate the audio. Refer to TTS | OpenAI for more information.  
TTS-1: Use this to optimize for speed.  
TTS-1-HD: Use this to optimize for quality.

• TTS-1: Use this to optimize for speed.

• TTS-1-HD: Use this to optimize for quality.

• Text Input: Enter the text to generate the audio for. The maximum length is 4096 characters.

• Voice: Select a voice to use when generating the audio. Listen to the previews of the voices in Text to speech guide | OpenAI.

### Options#

• Response Format: Select the format for the audio response. Choose from MP3 (default), OPUS, AAC, FLAC, WAV, and PCM.

• Audio Speed: Enter the speed for the generated audio from a value from 0.25 to 4.0. Defaults to 1.

0.25

4.0

1

• Put Output in Field: Defaults to data. Enter the name of the output field to put the binary file data in.

data

Refer to Create speech | OpenAI documentation for more information.

## Transcribe a Recording#

Use this operation to transcribe audio into text. OpenAI API limits the size of the audio file to 25 MB. OpenAI will use the whisper-1 model by default.

whisper-1

Enter these parameters:

• Credential to connect with: Create or select an existing OpenAI credential.

• Resource: Select Audio.

• Operation: Select Transcribe a Recording.

• Input Data Field Name: Defaults to data. Enter the name of the binary property that contains the audio file in one of these formats: .flac, .mp3, .mp4, .mpeg, .mpga, .m4a, .ogg, .wav, or .webm.

data

.flac

.mp3

.mp4

.mpeg

.mpga

.m4a

.ogg

.wav

.webm

### Options#

• Language of the Audio File: Enter the language of the input audio in ISO-639-1. Use this option to improve accuracy and latency.

• Output Randomness (Temperature): Defaults to 1.0. Adjust the randomness of the response. The range is between 0.0 (deterministic) and 1.0 (maximum randomness). We recommend altering this or Output Randomness (Top P) but not both. Start with a medium temperature (around 0.7) and adjust based on the outputs you observe. If the responses are too repetitive or rigid, increase the temperature. If they’re too chaotic or off-track, decrease it.

1.0

0.0

1.0

Refer to Create transcription | OpenAI documentation for more information.

## Translate a Recording#

Use this operation to translate audio into English. OpenAI API limits the size of the audio file to 25 MB. OpenAI will use the whisper-1 model by default.

whisper-1

Enter these parameters:

• Credential to connect with: Create or select an existing OpenAI credential.

• Resource: Select Audio.

• Operation: Select Translate a Recording.

• Input Data Field Name: Defaults to data. Enter the name of the binary property that contains the audio file in one of these formats: .flac, .mp3, .mp4, .mpeg, .mpga, .m4a, .ogg, .wav, or .webm.

data

.flac

.mp3

.mp4

.mpeg

.mpga

.m4a

.ogg

.wav

.webm

### Options#

• Output Randomness (Temperature): Defaults to 1.0. Adjust the randomness of the response. The range is between 0.0 (deterministic) and 1.0 (maximum randomness). We recommend altering this or Output Randomness (Top P) but not both. Start with a medium temperature (around 0.7) and adjust based on the outputs you observe. If the responses are too repetitive or rigid, increase the temperature. If they’re too chaotic or off-track, decrease it.

1.0

0.0

1.0

Refer to Create transcription | OpenAI documentation for more information.

## Common issues#

For common errors or issues and suggested resolution steps, refer to Common Issues.