

Q1: Hi, who are you?

A1: Hi, I'm AudioGPT, a powerful audio dialogue assistant tool that can help with a wide range of tasks and provide valuable insights and information on a wide range of topics.

Q2: What can you do?

A2:I can help with a wide range of tasks related to audio, such as generating audio from user input text, generating audio from an image, synthesizing speech from user input text, generating singing voice from user input text, generating text from an audio, audio inpainting, detecting sound events from an audio, synthesizing binaural audio from a mono audio input, and so on.

Q3: I am tired today, can you generate a music for me?

A3: The singing voice is generated and saved as  audio/1b870c1d.wav. Enjoy!

Q4: I try to dub for movie. Can you help me generate a background sound that includes birds singing and people speaking?

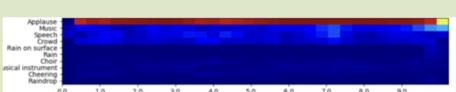
A4: The audio is generated and saved as
♫ [audio/c2dfd256.wav](#). Enjoy!

O5: Please write an caption based on the audio.

A5: The caption for  audio/c2dfd256.wav is: a woman is speaking with a man, while some birds are singing.

Q6: What events in this audio:
 [audio/932dfe0.wav](#)

A6: The audio includes applause and speech events. The predicted probability as following:



Q7: Synthesize a speech with text ‘My name is Audio GPT, I am happy to help you!’

A7: Synthesized speech saved as
 [audio/71a3503f.wav](#).

Q8: Please transfer the synthesized audio into binaural audio.

A8: Binaural audio saved as  [audio/a08211d3.wav](#). Is there anything else I can help you with?

Q9: Speak using the voice of this audio. The text is "here we go". [audio/a806f68f.wav](#)

A9: Received! Here is the audio file you request:  [audio/9df31081.wav](#)

Q10: Generate an audio based on the picture.



A10: Received! Here is the audio file you
request: Here is the audio you requested:
 [audio/10d1ac8b.wav](#)

O11: I want to inpaint the audio.

A11: Please enter the mask region:

Q12: The mask region as follow:



A12: Here are the predict audio and the mel spectrum.  [audio/d109610d.wav](#)
 [image/9181abac.png](#)