VoiceGPT - AiKodex

Documentation

NOTE:

Please allow Unity up to 4 hours to generate an Invoice Number from the time of your purchase. Once it is available, you can enter this number in the invoice field on top to access the voice generator.

VoiceGPT is a highly-realistic Text To Voice AI solution for Fantasy Voices. This Unity Editor's Extension tool can create voices from text, clone voices, trim, combine and equalize audio files. Choose from 60+ voices. This service does not entail subscriptions or repetitive payments. This documentation will help you get started with using the extension and provide information on how to use the asset to its best capacity.

Examples

Accents

Now with accent support, you can create different accent combinations easily.

To create an accent, type the text in the main language you want the voice to be in. Select a different language to create an accent onto the voice.

Text - *Now, retired, I sit in my small dacha, sipping hot tea, memories of comrades and distant battles warming my heart. My babushka's borscht, a taste of home, brings comfort in the quiet days. Life was tough.*

Language: Russian - ***Play***

Text: *Jaa! It's all so different now..... Used to bike through tulip fields, and now, dodging zombies! I hate this!*

Language: Dutch - *__Play__*

Pause

There are a few ways to add pauses to the generated text. One trick that seems to provide the most control and predictability is a simple dash (-) or the em-dash (—).

Ellipsis (…) also works to add a pause between words but usually also adds some "hesitation" or "nervousness" to the voice that might not always fit.

So - I think - this is something I want - __Play__

Or

Eah... that seems a bit weird - __Play__

The system is always being improved upon, and we are currently working on adding features such as the ability to add pauses and change the speed of the generated voice overs.

Emotion

If you want the AI to express a specific emotion, the best approach is to write in a style similar to that of a book. To find good prompts to use, you can flip through some books and identify words and phrases that convey the desired emotion.

For instance, you can use dialogue tags to express emotions, such as he said, confused, or he shouted angrily. These types of prompts will help the AI understand the desired emotional tone and try to generate a voiceover that accurately reflects it. With this approach, you can create highly customized voice overs that are perfect for a variety of applications.

Please note that you need to put a backslash (\) before quotes else you may receive a 400 error.
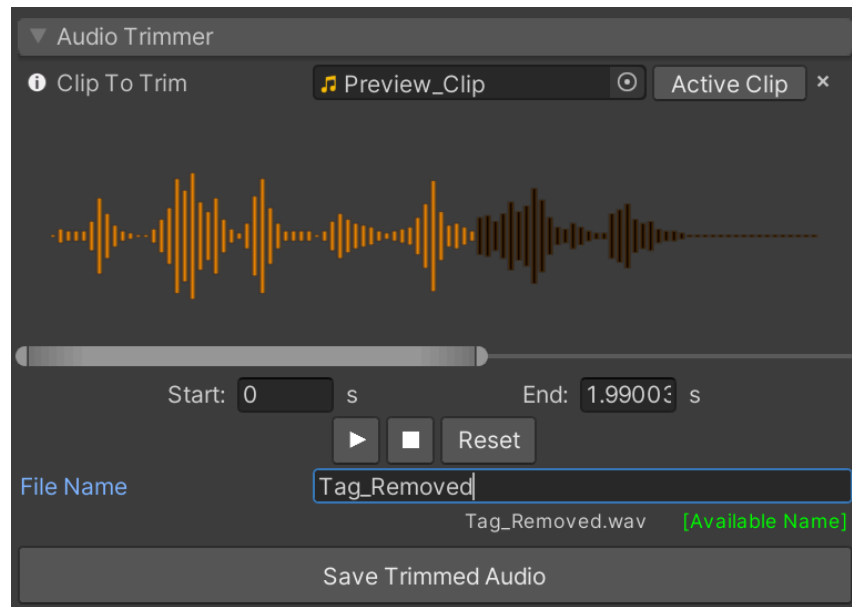
\"*Are you sure about that?*\" he said, confused. *__Play__*

You will also have to somehow remove the prompt as the AI will read exactly what you give it.

Tag Removal

Since the AI says exactly what the input is, you need to trim the part where the AI says

adds context at the end. To do this, please use the Audio Trimmer.

The AI can also sometimes infer the intended emotion from the text's context, even without the use of tags.

\"*Do you really think so?*\" **Play**

This is not always perfect since you are relying on the AI discretion to understand if something is sarcastic, funny, or just plain from the context of the text.

Pacing

To control the pacing of the speaker, you can use the same approach as in emotion, where you write in a style similar to that of a book. While it's not a perfect solution, it can help improve the pacing and ensure that the AI generates a voiceover at the right speed. With this technique, you can create high-quality voice overs that are both customized and easy to listen to.

\"*I wish you were right, I truly do, but you're not*\" she said **calmly**. **Play**
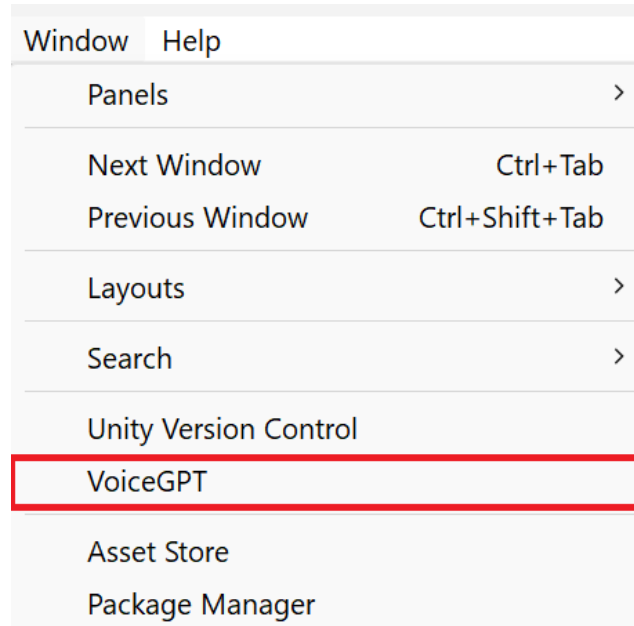
Use the audio trimmer to cut out the part "she said calmly" tag.

Dependencies

This asset requires the external package **Editor Coroutines 1.0.0** which can be found in Window > Package Manager > **Editor Coroutines**.
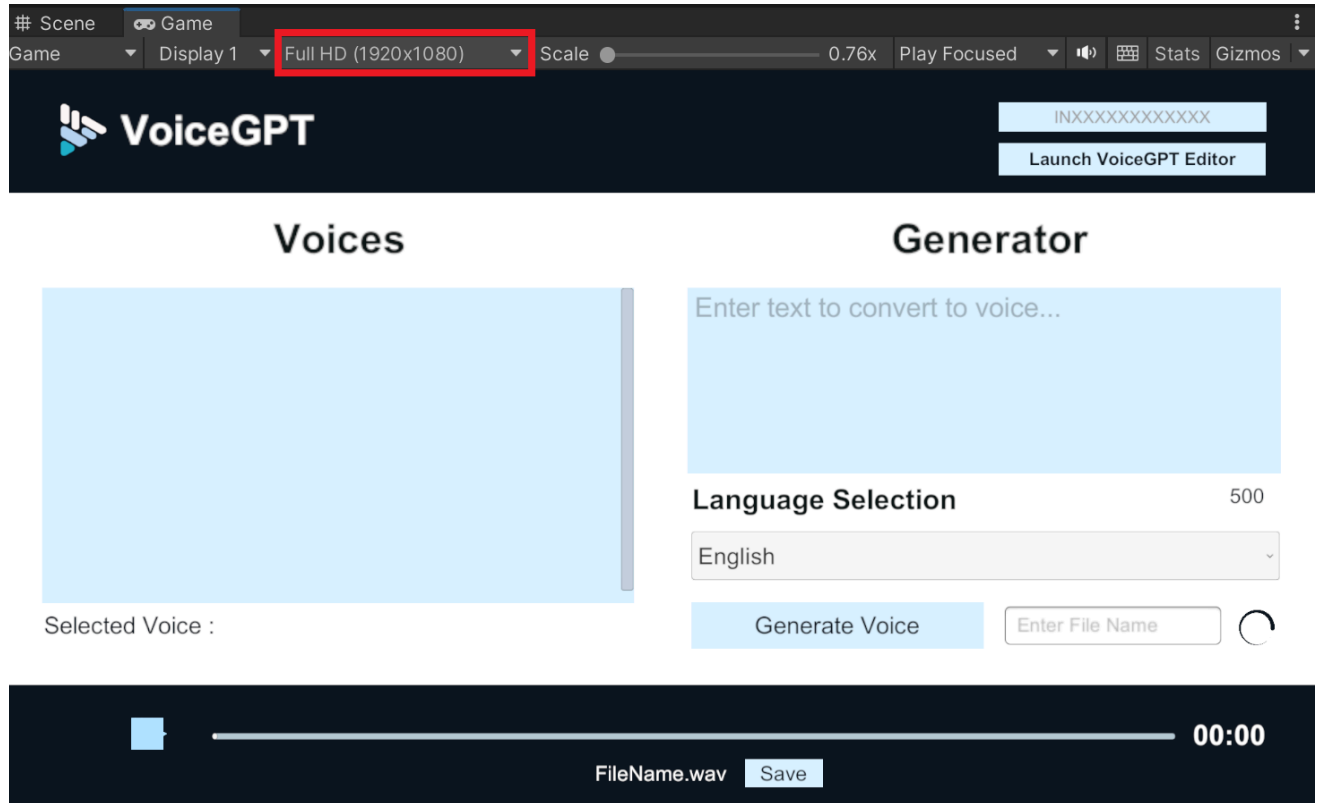
Usage

To use the Editor's Extension please go to Window > VoiceGPT
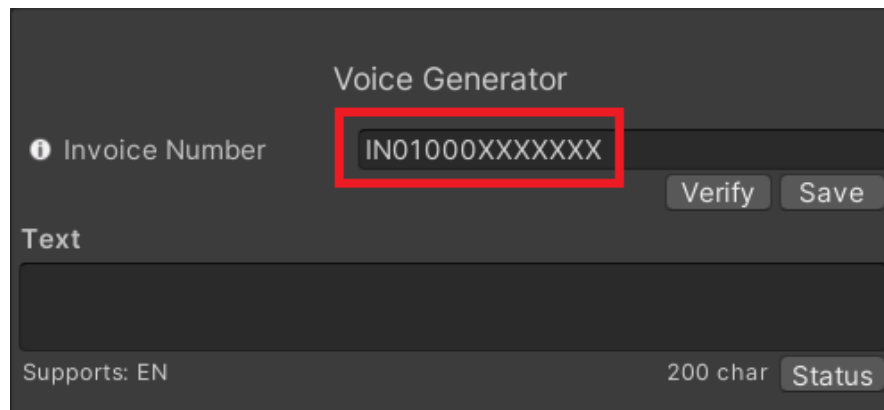
Features

Scene

Please set the aspect ratio to 16:9 to view all the contents of the scene. Press the Intro button to load the intro file on to the preview player. The detailed functionality of the asset is explained further in the documentation including the working of this scene and other audio operations.

Invoice Number

We use this to assign you the number of characters every month. You can find the invoice number here in the My Orders section of the Unity Asset Store.
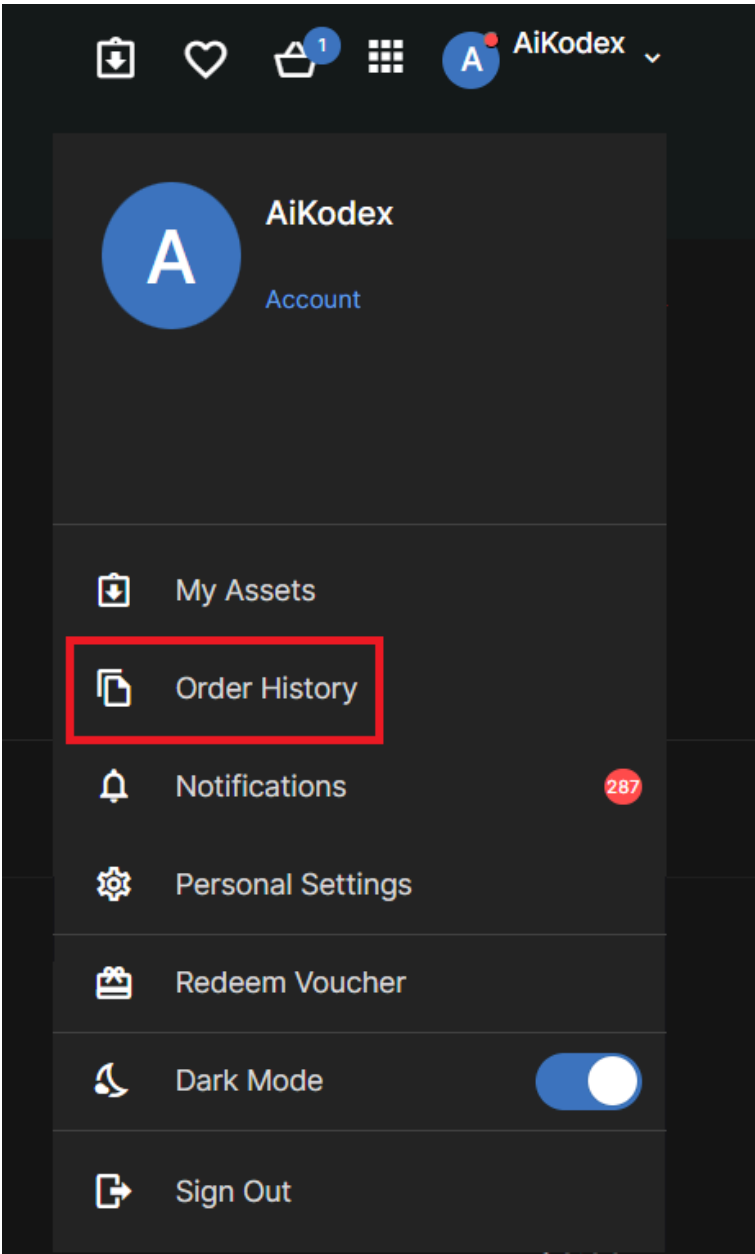
Order History

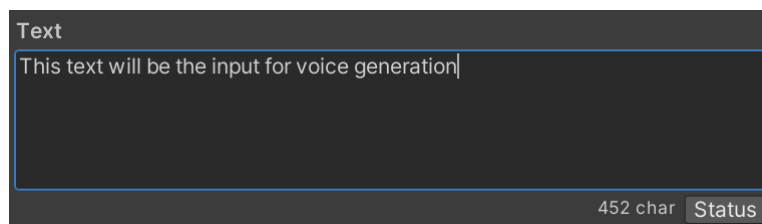| Order Date ⌄	Payment Type	Order Id	Invoice Number	Order Total	Status	
		2474	IN01000		COMPLETED	View Order Details
		2474	IN01000		COMPLETED	View Order Details
		2474	IN01000		COMPLETED	View Order Details
		2474	IN01000		COMPLETED	View Order Details

Once you have entered the number, you can then click on **verify** to check if the number you have entered is correct. If it is, you can save the number so when you launch the window again, you do not need to fill this field again. Once you have generated voices with the plugin, you can check the status of generation by clicking on the Status button. It should tell you the number of characters you have used.

For the scene, you can enter the invoice number on the top right section.

Editor Window

Voice Generator

Text: This field allows you to input text for narration. Enter the text you want to convert to a voice in the "Text" field. For now, the text is limited to a character count of 500. This is due to longer processing times for more tokens.

Model: Model specifies text-to-speech (TTS) model file to be used for the generation of the voice. VoiceGPT offers 1 model currently which is the VoiceGPT_X. We will be adding more versions to this for better and more up-to-date performance.

Voice: Selects the Voice name to use for the given model. Choose from a variety of different voices and find the best fit for your character.
Voices are synthetically created and do not resemble any public figure. Any resemblance to actual persons is purely coincidental.

Language: Select the language you want to generate the audio in. As of now, English is the most performant with the best quality.

Segment Semantics: Splits texts into sentences. Retains context if switched to false but comes at the cost of lowering the number of characters that can be processed at once.

Voice Cloning
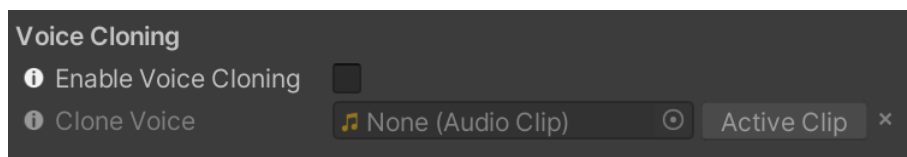
Enable Voice Cloning: Enables the voice cloning field and disables the voice field. Also, if the voice cloning field is enabled and no file is present in the Clone Voice field, then the Generate button will be grayed out.

Clone Voice: Input a .wav file (.mp3 and other compressed formats are not supported). The audio file should be at least 3 seconds long. Preferably in the range of 6-15 seconds. The file should not be more than 20MB. The speech should be in the target language of audio to be generated.
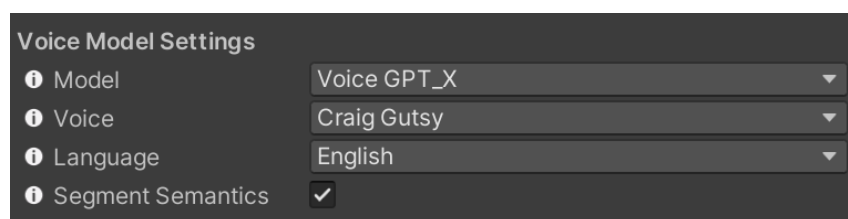

File Name: Is automatically assigned based on the selected voice. Additionally, increments the take field after voice processing. Available names are checked if the name field is modified by the user.

[Overwrite Name]: This file name already exists. Clicking on generate will overwrite and replace the current file. Proceed with precaution.

[Available Name]: This file name is available to use

Voices Folder: Changes the directory of generation. We recommend the directory be kept as the default.

Click on the Generate Voice button to send the inputs for processing.
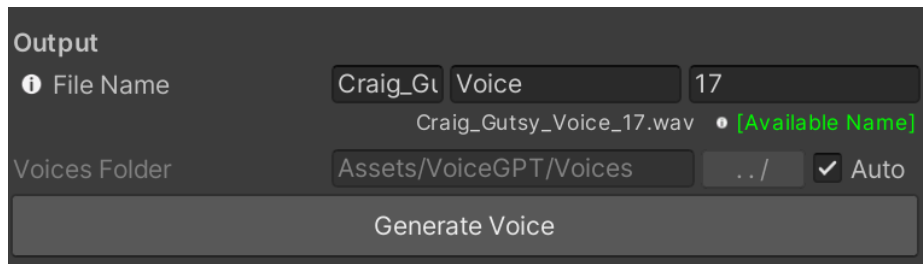
Preview

The preview section is a convenient feature that allows you to preview sound files directly within the interface, eliminating the need to switch between different windows or applications. By utilizing this functionality, you can quickly assess the contents of audio files without interrupting your workflow.

To access the preview section, simply single-click on a file within your project. Keep the plugin window active beside the scene. Hover your cursor over this panel, and you will notice that the preview section becomes enabled.

Once the preview section is activated, you can utilize the playhead, which is a visual representation of the current playback position, to navigate through different sections of the audio file. By scrubbing the playhead back and forth along the timeline, you can easily preview specific segments of the audio, allowing you to pinpoint certain parts or listen to the file in its entirety.

This functionality is particularly useful when you need to review and evaluate the contents of multiple voice recordings. Instead of opening each file individually in separate applications or audio players, the preview section enables you to conveniently listen to the audio files directly within the project interface.
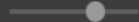
By offering a seamless and efficient way to preview sound files, this feature helps streamline your workflow and enhance productivity. It provides a centralized platform where you can not only organize your project's audio assets but also evaluate and make informed decisions about the content of each file. Whether you're a sound designer, audio editor, or working on any project involving audio files in Unity the preview section offers a valuable tool to facilitate your creative process and ensure the optimal outcome of your work.

ℹ Preview

Preview_Example_1, 44100Hz, 1.43s

▶ ‖ ■

Playhead ——●—— [1] s

Audio Utility

Audio Trimmer: Select an audio file you wish to trim. Once selected, use the slider to cut portions of the audio. When satisfied, save the audio by entering a valid name for the audio file. Click on the Active Selection button to select the clip active in the project. To remove the selection, simply click on the x button on the right side of the clip selection field.
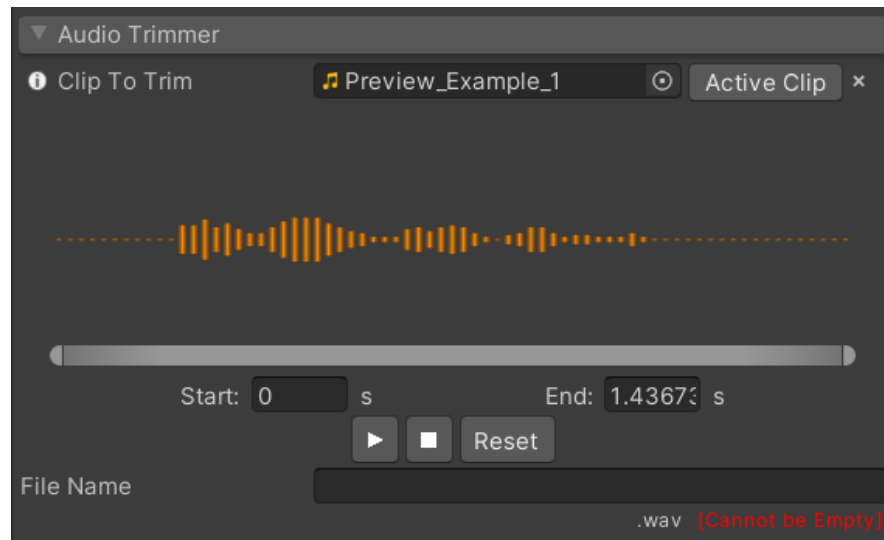
Audio Joiner: Select two or more audio files you wish to combine. Select the audio files from the project and click on "Set Selected" to auto populate the queue with the selected files. Please note that you cannot manually assign clips using the editor, you may only use the Set Selected Button to assign clips in this version of the asset. You can rearrange the audio clips in the hierarchy by dragging the clips. Once satisfied with the arrangement of the clips, enter a suitable name and save the file. You can clear the queue using the x button on the right of the Set Selected Button.

The Set Selected Button will be disabled until two files from the project are selected.
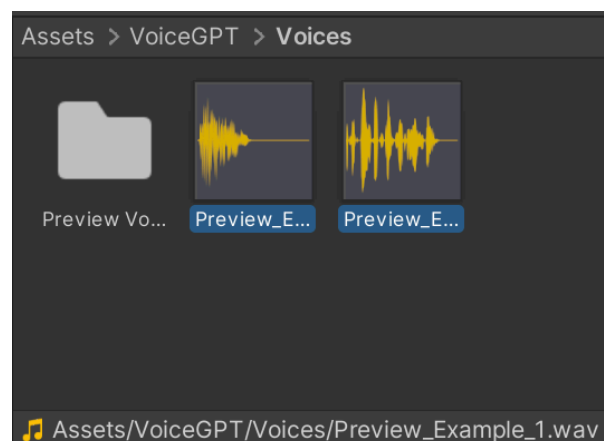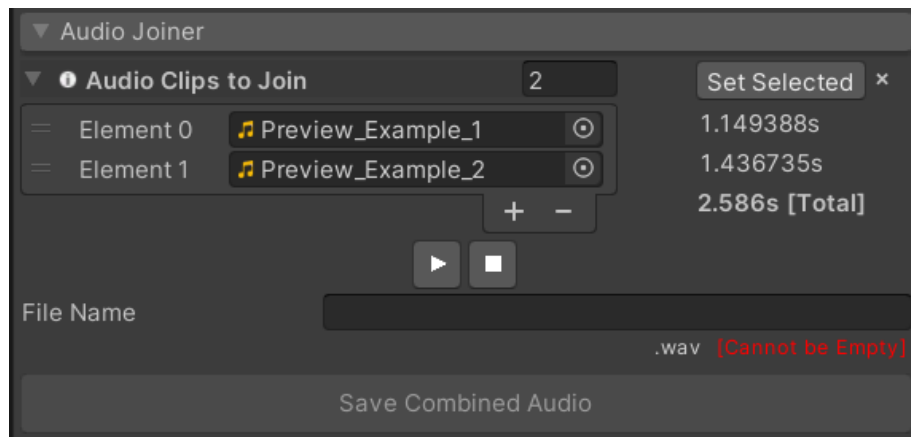
Audio Equalizer: Select an audio file you wish to equalize. You can adjust the sliders to make the voice loud, low, bassy or shrill. Once satisfied with the changes, enter a suitable name and save the file. You can reset the settings for the equalizer using the reset button.
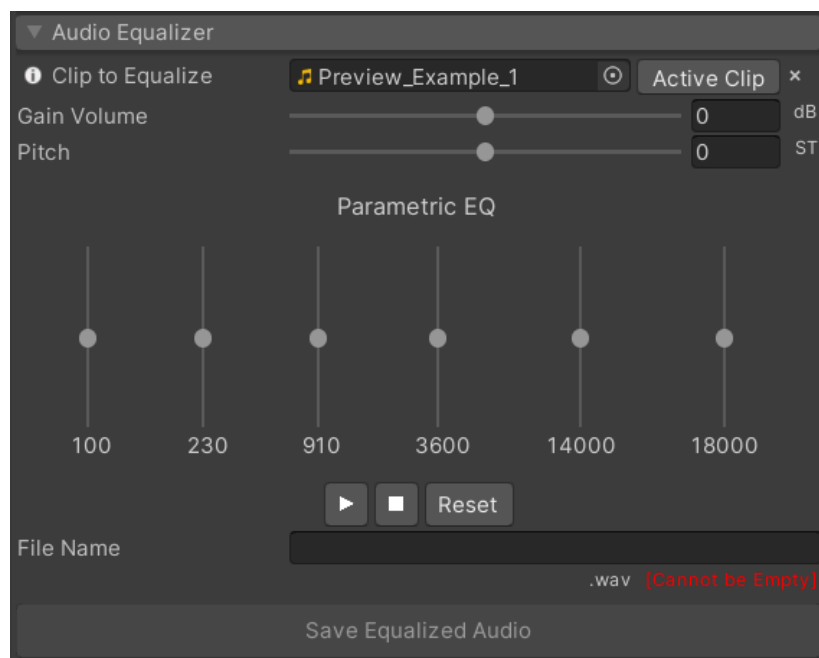
If you have any questions or difficulties in generation of voice from text, trimming the audio clips, combining them or equalizing them, please reach out to info@aikodex.com

Saving Files

Files are saved as the name and directory provided in the asset. It is important to note that the name of the file cannot be left empty, as it serves as a unique identifier for each asset. If you attempt to save a file without entering a name, a red tag labeled "[Cannot be Empty]" will appear, indicating that a name must be provided.

Furthermore, it is crucial to exercise caution when working with files that have identical names. In such cases, there is a possibility that you may accidentally overwrite an existing file. Overwriting a file means replacing its content with the new file being saved.

To assist you in avoiding accidental overwrites, the plugin window includes a warning mechanism. If you attempt to save a file with a name that already exists within the designated directory, a red tag labeled "[Overwrite]" will appear as a warning. This visual indicator serves as a reminder that saving the file with the current name will result in the existing file being replaced. This warning helps you make informed decisions and take necessary precautions before proceeding with the save operation.

On the other hand, if the file name you have chosen is available and unique within the specified directory, you will see a green tag labeled "[Available name]". This green tag serves as a confirmation that the chosen name can be used without any risk of overwriting existing files.

By providing clear visual cues and warnings, the system aims to prevent accidental file overwrites and ensure the integrity of your audio assets. This helps maintain organization and avoid potential data loss or confusion that may arise from unintentional overwrites.
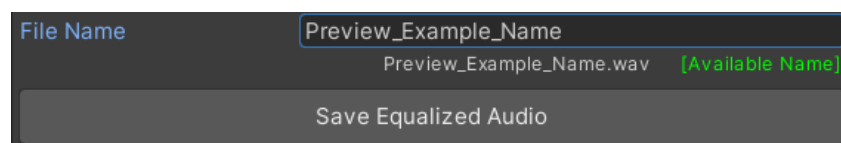
Troubleshooting:

If you see the network errors:

500: Internal Server error / Cannot connect to destination host

Possible causes:

There may be times that the server is down. In this unlikely event, please check the forum for announcements by us pertaining to server maintenance or contact info@aikodex.com if this issue persists for over a day. Please check your internet connection and try again in a few hours.

429: Too many requests

Possible Causes

Occurs when the requests are sent too often. Verification is never required after the first time. Status of characters can be checked less frequently as well. The request limit is set on per second, per minute and per day maximums. For verification and status, please re-check in a minute. Voice generation limits are set on a per second and per day basis. This is to prevent automation and the misuse of the service.

400: Bad Request

Possible Causes

Occurs when the information passed is not recognized either due to syntax error or other reasons. Using special characters like ("") that have a specific meaning in coding can cause this error. Please send us an email with the Unity version included outlining your issue in as much detail. A screenshot or video of the problem will help us serve you better.

Request Timeouts (408 Request Timeout response status code)

Possible Causes

The voice model could have run into a generative error. When this happens, the audio

file becomes extremely large and is not able to be passed through the API built. Please try again with a fewer number of words using sentence chunking.

Privacy, legal terms and misuse of service

At AiKodex, we believe that protecting the privacy of our users is of utmost importance. We provide a secure and private environment for users to utilize our Voice Generation services, without compromising on their privacy.

Our privacy policy is designed to ensure that we do not store any data or personal information that is shared between users and our service. We do not store input text, generated voices, IP addresses or any other data, and our remote servers are configured to automatically delete any data that is left on the system.

We understand that privacy is a fundamental right, and we are committed to upholding this right for our users. We will continue to invest in the latest technologies and security measures to ensure that our users can enjoy a safe and private experience with VoiceGPT.

Welcome to VoiceGPT by AiKodex ("we," "us," or "our"). By accessing or using our services, including but not limited to our asset, applications, and any other related services (collectively, the "Services"), you agree to comply with and be bound by these Terms of Service ("Terms"). If you do not agree to these Terms, please do not use our Services.

As long as you have the necessary intellectual property rights in the Content you Generate with our Services, you may use this Generated Content for commercial purposes.

a. Content: Refers to audio clips generated by VoiceGPT

b. Generated Content: Content created, or produced using our Services.

c. Intellectual Property Rights: Refers to patents, copyrights, trademarks, trade secrets, and any other intellectual property rights recognized by applicable law.

We may revise and update these Terms from time to time. Any changes will be effective upon posting the revised Terms on the Asset Store. Your continued use of our Services after such changes constitutes your acceptance of the updated Terms.

Voice Cloning Guidelines

These guidelines will help you to understand how you can use this voice technology safely, follow best practice, and respect our terms of service. What constitutes a lawful use of voice cloning may vary depending on who created the content you are reusing, where you are located and where you publish your materials. These guidelines describe what constitutes safe uses of voice cloning in most countries, with a focus on countries in the European Union and the United Kingdom. This guide does not constitute legal advice.

What do we mean by voice cloning?

"Voice cloning" is the digital imitation of a person's voice with high fidelity results.

When you use this voice cloning technology, you upload a recorded voice, and our model creates its digital representation. You can then use this digital representation to generate new speech. Our cloning technology produces voices which sound similar but are not technically identical to the materials you have uploaded.

Voice cloning : safe use and best practice explained

You can voice clone in three scenarios, with exceptions described below:

Voice cloning will be safe if you use your own voice and content to generate a clone. You can use this clone for any purpose.

Voice cloning will be safe if you use someone else's voice or content to generate a clone with their permission. You can use this clone for any purpose you received permission for.

Voice cloning is only sometimes safe if you use someone else's voice and materials to generate a clone without their permission. You can use this cloning technology if you don't impact the person's privacy or economic interests.

A very important rule applies to all uses of voice cloning technology: you cannot clone a voice for abusive purposes such as fraud, discrimination, hate speech or for any form of online abuse without infringing the law.

Communicate your own rights and permissions in the voice clone. Your voice clone, that you have created with your voice or someone else's provided it is not an accurate representation of their voice may be original enough to be protected by its own intellectual property rights. If this is the case, you may be the owner of the rights in the voice clone. You may also own the rights in other source materials you have created, like the text you have used to prompt new speech.

Prohibited Activities and Misuse

You shall not use these services, data or content provided by VoiceGPT in any manner that is illegal, unethical, or inconsistent with the intended use of the service. You may not abuse the server by overloading requests on it or using the service from outside Unity. You may not try to automate this service in any manner or form. This may lead to termination of service. Please refrain from using profanity. Do not use this asset to promote hate or discrimination. We aim to create a welcoming and inclusive environment for all users. Thank you for your understanding and cooperation.

Happy Dubbing!

- Offered By AiKodex