

Audio Chunk

Audio Chunk User's Guide

Version 2.2

By Dr. Eric O. Flores

December 2024



Public domain under CC0: No rights reserved.

Table of Contents

Public Domain Dedication.....	3
Introduction.....	4
Chapter 1: Getting Started.....	5
Installation:.....	5
Launching Audio Chunk:.....	5
The Main Window:.....	5
Chapter 2: Understanding the Interface.....	6
Chapter 3: Step-by-Step Guide.....	8
Load an Audio File:.....	8
Select an Output Directory:.....	8
Set Chunk Duration:.....	8
Choose Output Format:.....	8
Set Base Name (Optional):.....	8
Start Splitting:.....	8
Monitoring Progress:.....	9
Completion:.....	9
Chapter 4: Best Practices for AI Voice Training.....	9
Chapter 5: Troubleshooting.....	9
Chapter 6: Conclusion.....	10
Disclaimer.....	11
No Copyright.....	11
Other Information.....	11

Public Domain Dedication

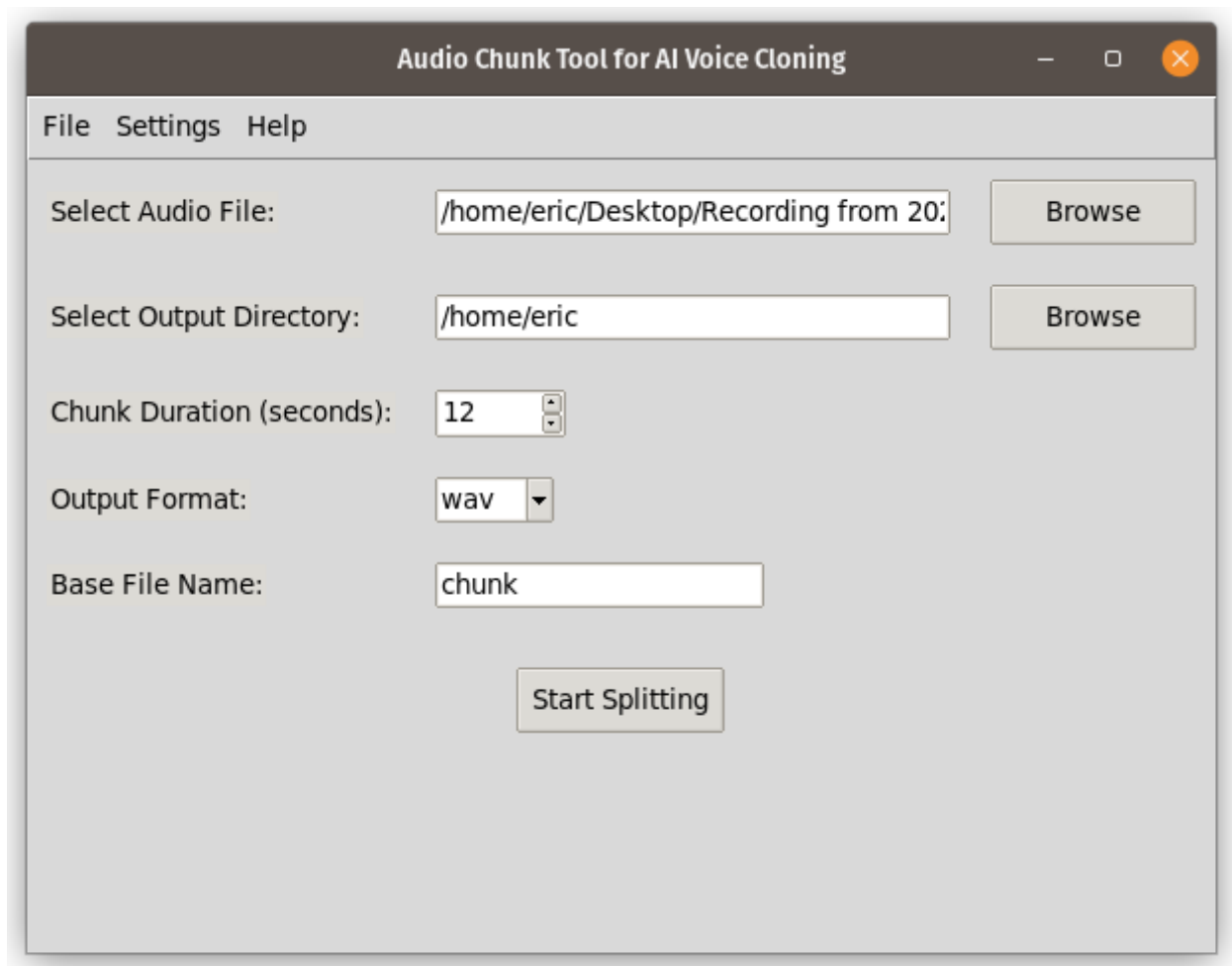
To the extent possible under law, the author(s) have dedicated all copyright and related and neighboring rights to this software and documentation to the public domain worldwide. This software and documentation are distributed without any warranty.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission. See the Creative Commons Public Domain Dedication for more details:

<https://creativecommons.org/publicdomain/zero/1.0/>

Introduction

Welcome to Audio Chunk, a user-friendly tool designed to split audio files into smaller, equally timed segments. This process is crucial for preparing audio data for use in AI voice training models, ensuring that the data is consistently formatted for optimal training results. This guide will walk you through each step, ensuring you can effectively use Audio Chunk for your audio processing needs.



Chapter 1: Getting Started

Installation:

- Audio Chunk is a Python application and requires Python 3.x and the tkinter library (which is often included with Python installations). You will need to have librosa and soundfile installed. You can install it using: `pip install librosa soundfile`.

Launching Audio Chunk:

- Locate the Audio Chunk.py script and run it using Python. You can typically do this from your terminal or command prompt by navigating to the directory and typing `python Audio Chunk.py` and pressing enter.

The Main Window:

- Once launched, Audio Chunk presents a simple graphical user interface (GUI) with a menu bar at the top and several input fields and buttons.

Chapter 2: Understanding the Interface

The Audio Chunk interface is designed to be intuitive and easy to use. Here's a breakdown of the main elements:

- **Menu Bar:** The menu bar at the top provides access to different functions.
 - **File:**
 - **Load a file:** Opens a file selection dialog to choose your input audio file.
 - **Quit:** Closes the application.
 - **Settings:**
 - **Set Output Directory:** Opens a dialog to select the directory where split audio files will be saved.
 - **Set Chunk Duration:** Focuses on the spinbox where the chunk duration is set.
 - **Set Output Format:** Focuses on the combobox where the output format is selected.
 - **Start Split:** Start the audio splitting using user input from settings, or the GUI input fields.
 - **Help:**
 - **About:** Displays information about Audio Chunk, including the version and details.
- **Input Fields and Buttons:**
 - **Select Audio File:**
 - **Text Field:** Displays the path of the selected audio file.
 - **Browse Button:** Opens a file selection dialog to select an audio file (.mp3, .wav, .flac, .ogg, .m4a formats are supported).
 - **Select Output Directory:**
 - **Text Field:** Displays the path of the chosen output directory.
 - **Browse Button:** Opens a directory selection dialog to select where the split audio files will be saved.
 - **Chunk Duration (seconds):**
 - **Spinbox:** A numeric input field to specify the length (in seconds) of each audio chunk. Default is set to 10 seconds. This is the most crucial setting for creating equally timed audio frames.
 - **Output Format:**

- **Combobox:** A dropdown list to select the output file format for the split audio chunks (.wav, .flac, .ogg, .aiff formats are supported). Default is set to .wav.
- **Base File Name:**
 - **Text Field:** A text input to specify the base name for the output files. The default name is 'chunk'.
- **Start Splitting:** A button that initiates the splitting process.

Chapter 3: Step-by-Step Guide

Here's a step-by-step guide on how to use Audio Chunk:

Load an Audio File:

- Click on "Load a File" under the "File" Menu or Click the "Browse" button in the "Select Audio File" section.
- Navigate to the directory containing the audio file you want to process.
- Select your file and click "Open". The file path should now appear in the text field next to the browse button.

Select an Output Directory:

- Click on "Set Output Directory" under the "Settings" Menu, or the "Browse" button in the "Select Output Directory" section.
- Navigate to the directory where you would like to save the output files.
- Select the directory and click "Open". The directory path should now appear in the text field next to the browse button.

Set Chunk Duration:

- In the "Chunk Duration (seconds)" field, use the spinbox to set the desired length of each audio chunk in seconds. This is crucial for creating consistent audio frames for AI training. It's recommend to keep it between 5-10 seconds. You can either use the arrow keys to increase or decrease the number, or you can directly enter it into the spinbox. You can alternatively click on "Set Chunk Duration" under the "Settings" Menu to bring focus to the spinbox.

Choose Output Format:

- Use the dropdown list in the "Output Format" field to choose the format of the output audio files. .wav is often recommended for its compatibility, while lossless formats such as .flac may be preferable for training models. You can alternatively click on "Set Output Format" under the "Settings" Menu to bring focus to the dropdown.

Set Base Name (Optional):

- In the "Base File Name" field, enter a desired base name for the output files, if the user does not, the default name chunk will be used. The output files will have their index number appended to this name (e.g. chunk_001.wav, chunk_002.wav...).

Start Splitting:

- Once all settings are entered click on the "Start Splitting" Button, or you can click on "Start Split" under the "Settings" Menu.

- A confirmation dialog will appear asking you to confirm the approximate number of chunks to be generated. Click "Yes" to proceed.
- Audio Chunk will proceed to load the file and split the file into chunks and create output files inside a subfolder called `audio_chunks` in your selected output directory.

Monitoring Progress:

- While the process is running, you'll see messages on the terminal/command prompt showing each created file.

Completion:

- Once completed, you'll be prompted with a message indicating the successful completion of the splitting process and a message box showing the amount of files created.

Chapter 4: Best Practices for AI Voice Training

- **Consistent Chunking:** Using consistent durations for all your training data is critical for the performance of your AI models.
- **Output Format:** .wav is recommended for AI voice training model input due to broad compatibility; however, .flac offers better audio fidelity, but also a larger file size.
- **Data Augmentation:** Consider additional data augmentation techniques to improve model performance. This can involve manipulating audio speed, adding noise, etc.
- **Labeling:** The training data, in combination with audio files, needs to have a labeling system so that an AI model knows who is speaking.
- **Hardware Considerations:** High-quality audio training models require significant hardware.

Chapter 5: Troubleshooting

- **File Format Issues:** Ensure that your input file is in a supported format (.mp3, .wav, .flac, .ogg, .m4a).
- **Output Directory Errors:** If a file can not be written, please ensure that you have the permissions to create directories and files. Also, ensure that there is enough storage space.
- **Unexpected Errors:** Please save the error messages in the terminal/command prompt, and also any error messages and review these.

Chapter 6: Conclusion

Audio Chunk is a practical and efficient tool for preparing your audio data for AI voice cloning models. With a user-friendly interface, you can easily split your audio files into equally timed chunks, ensuring that your AI training data is well-formatted. This user's guide should provide you with all the required knowledge to begin using this software.

If you have any further questions, or have identified any issues, please let us know so we can improve this for other future users.

Disclaimer

This software is intended for research purposes only and to prepare training data for AI models. Please ensure you use this software in an ethical and responsible way.

No Copyright

The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission. See Other Information below.

Other Information

In no way are the patent or trademark rights of any person affected by CC0, nor are the rights that other persons may have in the work or in how the work is used, such as publicity or privacy rights.

Unless expressly stated otherwise, the person who associated a work with this deed makes no warranties about the work, and disclaims liability for all uses of the work, to the fullest extent permitted by applicable law.

When using or citing the work, you should not imply endorsement by the author or the affirmer.