



# CROSSFADE

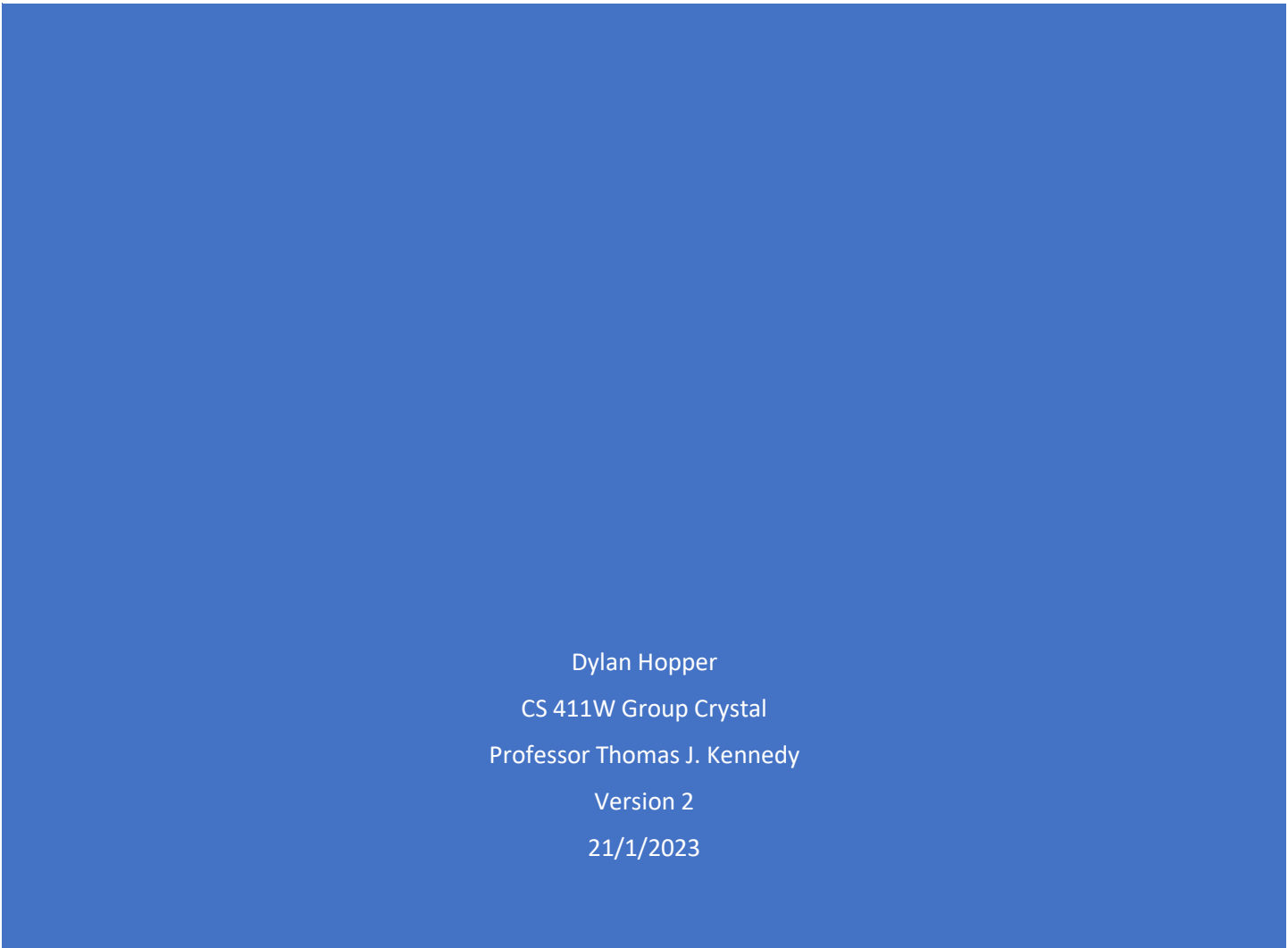
Dylan Hopper

CS 411W Group Crystal

Professor Thomas J. Kennedy

Version 2

21/1/2023



## Table of Contents.

1. Introduction.....	2
1.1 Purpose.....	2
1.2 Scope.....	2
1.3 Definitions, Acronyms, and Abbreviations .....	3
1.4 References .....	4
2. Overall Description .....	6
2.1 Purpose.....	6
2.2 Product Functions .....	6
2.3 User Characteristics .....	6
2.4 Constraints.....	7
2.5 Assumptions and Dependencies .....	7

## **1. Introduction**

This is the SRS documentation for the CrossFade application, a tool used for transcription and composing music. This document is explained below.

### **1.1 Purpose**

The purpose of the SRS documentation is to provide users with a description of what CrossFade does and how it will be expected to perform. The SRS document details the functional requirements of the software, including user interfaces, and algorithms used. It also provides the expected performance requirements, and non-functional requirements such as security, maintainability, and reliability. This information has been provided to the user so that they have a clearer image of what to expect from CrossFade.

### **1.2 Scope**

CrossFade intends to provide professional & beginner musicians, composers, and music instructors with the ability to easily transcribe and edit music in the form of audio, MIDI, or MusicXML. The transcribed audio can be manually edited, and uses artificial intelligence to recognize errors in sheet music

### **1.3 Definitions, Acronyms, and Abbreviations**

Convolutional Neural Network (CNN) - Deep learning algorithm which can differentiate one image from another by assigning weights and biases to different aspects of the images. It is used in audio to differentiate different frequencies in a visual format.`

Deep Learning - Subfield of machine learning which uses neural networks to solve complex problems. Learning comes directly from the data, instead of being hand-engineered by humans.

Keyboard - An electronic piano used to produce sound and MIDI information.

Monophony - A phrase of music in which only a single voice is played at a time.

Musical Instrument Digital Interface (MIDI) - A communications protocol used to connect physical and virtual music devices and instruments. MIDI files store note information which can be used to trigger instruments and devices.

MusicXML - A markup language format used to interchange and distribute digital sheet music.

Polyphony - A phrase of music in which more than a single voice is played at a time

## 1.4 References

- Automatic Music Transcription and Ethnomusicology: A user study. (n.d.). Retrieved from <https://diva-portal.org/smash/get/diva2:1474663/FULLTEXT01.pdf>
- H. Takeda, T. Nishimoto and S. Sagayama, "Rhythm and Tempo Analysis Toward Automatic Music Transcription," 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, Honolulu, HI, USA, 2007, pp. IV-1317-IV-1320, doi: 10.1109/ICASSP.2007.367320.
- Huang, Z., Jia, X., & Guo, Y. (2019, June 29). State-of-the-art model for music object recognition with Deep Learning. MDPI. Retrieved February 8, 2023, from <https://www.mdpi.com/2076-3417/9/13/2645>
- Jovanovic, J. (2015, February 2). How does Shazam work? music recognition algorithms, fingerprinting, and processing: Toptal®. Toptal Engineering Blog. Retrieved February 8, 2023, from <https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-recognition>
- Scarlatos, L. L. (n.d.). Continuous media. Audio. Retrieved March 1, 2023, from <https://www3.cs.stonybrook.edu/~lori/classes/GUI/sound.htm>
- Solanki, A., & Pandey, S. (2019, January 30). Music instrument recognition using deep convolutional neural networks - International Journal of Information Technology.

SpringerLink. Retrieved February 8, 2023, from

<https://link.springer.com/article/10.1007/s41870-019-00285-y>

Zhang, X. (2022, March 11). Aided recognition and training of music features based on

the internet of things and Artificial Intelligence. Computational intelligence and

neuroscience. Retrieved February 8, 2023, from

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8933112/>

## **2. Overall Description**

This section details the CrossFade application itself, including its purpose as a tool, and the fundamental parts that make it up.

### **2.1 Purpose**

CrossFade will allow users (musicians, composers, & teachers) to transcribe audio & MIDI to MusicXML format. It will also allow the user to import their own XML files. Once a file has been opened in CrossFade, the user will be allowed to manually edit the song, as well as have our advanced AI automatically check for – and correct – any detected errors. The file can then be saved and exported as an XML or MIDI file.

### **2.2 Product Functions**

- Live audio transcription.
- File import support to transcription (.mp3, .midi, .xml).
- Real time sheet music editing.
- Transcription error detection.
- Automatic error correction.
- Audio playback.
- File export support (.midi, .xml)

### **2.3 User Characteristics**

CrossFade is aimed at both beginner musicians – as a learning tool to help them visualize music and audio, and for experienced users in the music field, including advanced musicians, composers, and music educators.

## **2.4 Constraints**

N/A

## **2.5 Assumptions and Dependencies**

CrossFade has various libraries it depends on. For our artificial intelligence, we are relying on TensorFlow. To interface Python and MusicXML we are relying on the pymusicxml & Music21 libraries.