

# LAB 1 - CROSSFADE PRODUCT DESCRIPTION

## Lab 1 - Crossfade Product Description

Hyacinth Abad, Mohamed Abdullahi, Arnaud Blay, Dylan Hopper, Selai Naim,

Virginia Vano Ruano, Joseph Wassell,

Old Dominion University

CS411W

CrossFade

11/17/2023

Version 1

## Table of Contents

1 Introduction.....	
2 Product Description.....	
2.1 Key Product Features and Capabilities.....	
2.2 Major Components (Hardware/Software).....	
3 Identification of Case Study.....	
4 Product Prototype Description.....	
4.1 Prototype Architecture (Hardware/Software).....	
4.2 Prototype Features and Capabilities .....	
4.3 Prototype Development Challenges.....	
5 Glossary.....	
6 References.....	

## Figures and Tables

Figure 1 Major functional component diagram.....	
Figure 2 Phase 1 prototype major functional component diagram.....	
Table 1 Feature comparison between full product and prototype.....	

## **1 Introduction**

There is a large barrier involved with entering music or the music industry. Many individuals struggle with reading music, music theory, musical techniques, or a technical requirement associated with music. As technology advances, to produce and share music, individuals involved with the field now must understand how to use certain commonplace file types for storing musical information such as MIDI and MusicXML along with sheet music. With tools to produce MIDI and MusicXML files becoming more common this technical requirement is daunting to aspiring musicians. Because of the many difficulties in music, roughly 50% of music students decide to discontinue lessons by the time the students turn 17 years old (Ruth, 2021). To lessen the burden involved with the music industry, as well as promote the development of creative minds from many backgrounds, a tool will need to be developed that allows aspiring musicians to focus solely on music when entering the field. This tool should help reduce the rate at which music students decide to stop pursuing music and allow music students to enter musical fields with less immediate difficulties. The desktop application CrossFade is a software solution that will allow musicians to focus on music. Instead of requiring specialized tools or equipment, CrossFade will create a MIDI and MusicXML file for a musician when passing an audio file or playing music through a microphone. CrossFade allows for the transcription of shareable music for individuals exploring the music industry, and CrossFade aims to alleviate the technical requirements that have become required with the creation of music in recent years.

## **2 CrossFade Product Description**

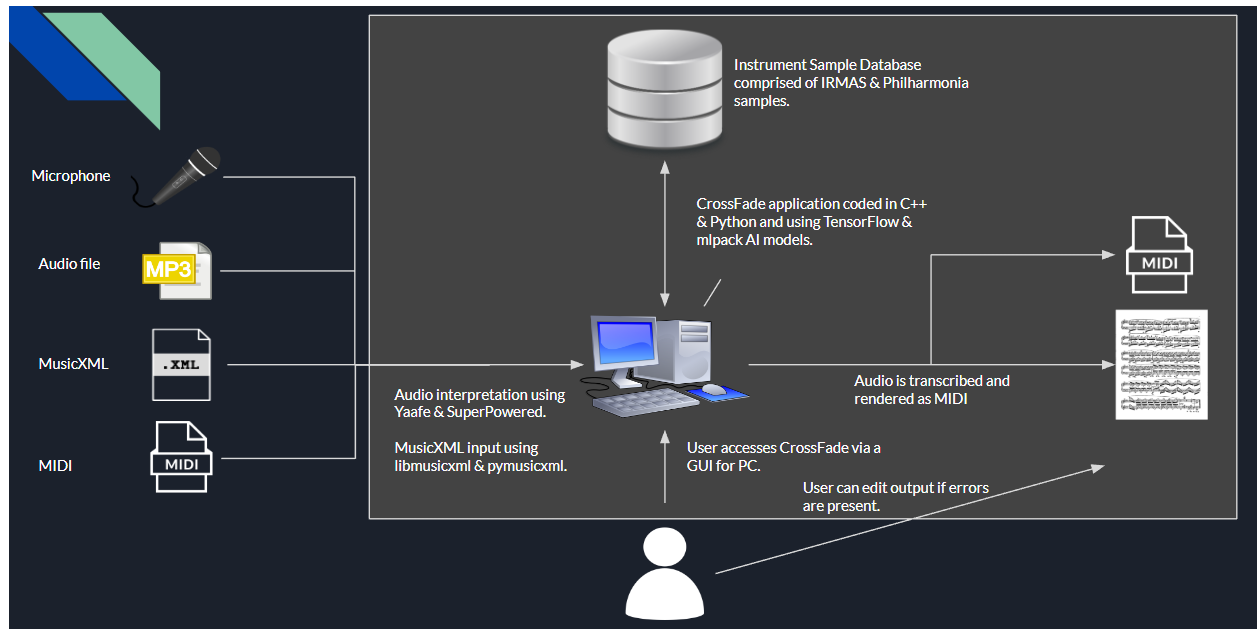
CrossFade is a desktop application that allows for the transcription of shareable music. CrossFade takes an audio file, a MIDI file, or audio played through a microphone as input, and creates a MIDI file and a MusicXML file. CrossFade allows individuals to focus solely on learning about music when entering the music industry, by creating shareable music files, and preventing the need to learn about music files and how to create sheet music.

### **2.1. Key Product Features and Capabilities**

CrossFade is an automatic transcription software with the capability to create written music through audio information. Unlike other transcription software CrossFade is able to correct musical transcriptions using a database of songs and artificial intelligence to determine likely errors and correct them. Crossfade also has the ability to recognize multiple different instruments at the same time. To prevent the user from being unsatisfied with the output, or the output containing errors, the user is able to adjust the output manually. CrossFade will highlight potential problem sections to allow for easier edits to the music.

### **2.2. Major Components (Hardware/Software)**

CrossFade requires a computer, and if the live audio feature is used CrossFade requires a microphone and musical instruments. CrossFade allows the user to interact with the application through a GUI on the user's PC. CrossFade takes in an audio file, a MIDI file, live audio through a microphone, and a MusicXML file. CrossFade then makes comparisons to a database based on the type of input passed and produces a MIDI file and a MusicXML file that the user can edit if there are errors present.



### 3 Identification of Case Study

CrossFade is intended for starting musicians with little to no experience. CrossFade can also be used by more experienced musicians and music teachers who may want to have students create music files using the software. Due to the error transcription correction feature, CrossFade is able to be utilized by more experienced musicians and producers.

### 4 CrossFade Product Prototype Description

The CrossFade prototype will be mainly focused on transcription correction. The prototype will have no features associated with the transcription functionality. This removes many of the capabilities that would likely be useful to beginners but offers a correction to many of the existing software in the hope of making automatic transcriptions more natural and understandable. The exact feature comparisons are listed in the tables below.

Transcription		
<u>Feature</u>	<u>Real World Product</u>	<u>Prototype</u>
Live Audio Transcription	✓	
MIDI Input/File Transcription	✓	
MIDI Transcription Correction	✓	✓
Multitrack Transcription	✓	
MusicXML Compatibility	✓	✓
MIDI File Export	✓	✓

Note Recognition		
<u>Feature</u>	<u>Real World Product</u>	<u>Prototype</u>
Monophonic Note Recognition	✓	
Polyphonic Note Recognition	✓	
Instrument Recognition	✓	
Instrument Distinction	✓	

Transcription Error Correction		
<u>Feature</u>	<u>Real World Product</u>	<u>Prototype</u>
Transcription Error Correction	✓	✓
Highlight Possible Errors	✓	✓
Take User Feedback	✓	✓
Offer Possible Solutions	✓	✓
Compare Single Note from Original to Written	✓	✓
Compare a Segment from Original to Written	✓	✓

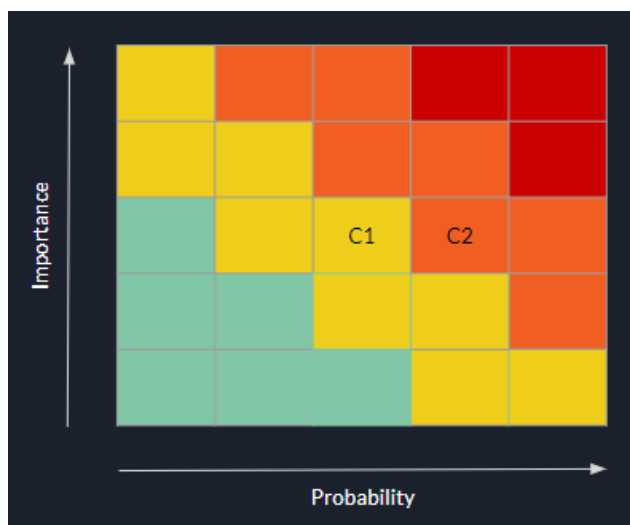
#### 4.1. Prototype Architecture (Hardware/Software)

The prototype will be able to accept MusicXML and will be able to provide automatic transcription correction features. The prototype will still need to make use of a database for comparisons, but some of the software components for parsing and interpreting audio are no longer required (Yaafe and SuperPowered).

## 4.2. Prototype Features and Capabilities

The prototype demonstrates the most useful features of the application are able to be accomplished. While many of the features of CrossFade are helpful for beginners they are also offered by other software. Automatic transcription is known to have many errors involved with the most prevalent type of error being rhythmic errors. By solving this issue the prototype shows that all features aside from the instrument recognition feature are possible. Instrument recognition has been determined to be a difficult feature with many risks involved, but there are software that accomplishes this feature in part. Creating the error transcription correction features shows that an accurate and natural transcription can be made with minimal knowledge lowering the amount of technical knowledge that a beginner would need to learn to enter the music industry. The prototype aims to be able to correct common errors in automatic transcription and create a shareable music file in MIDI and MusicXML format.

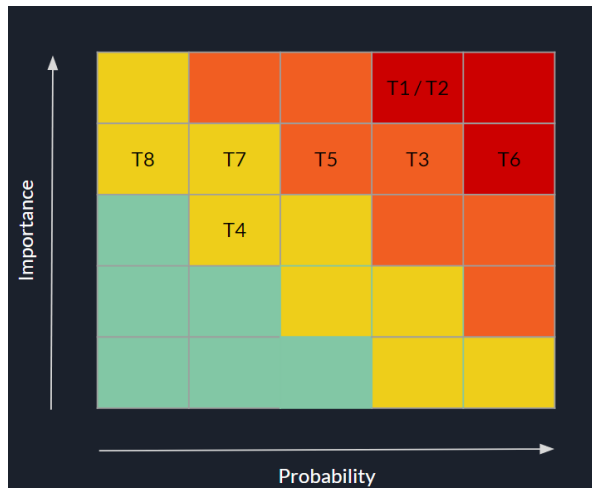
## 4.3. Prototype Development Challenges



The above figure shows the customer risks involved with the application.

- C-1 Difficulty using the application.

- A tutorial will be added to allow for customers to understand how to use the application.
- C-2 Unfamiliarity with the terms.
  - A list of definitions and links to explanations will be provided to help educate users on any unfamiliar terms that are commonly used.



The above figure shows the technical risks involved with the application. Some of these risks will not be associated with the prototype and have been separated in the list below.

Prototype risks:

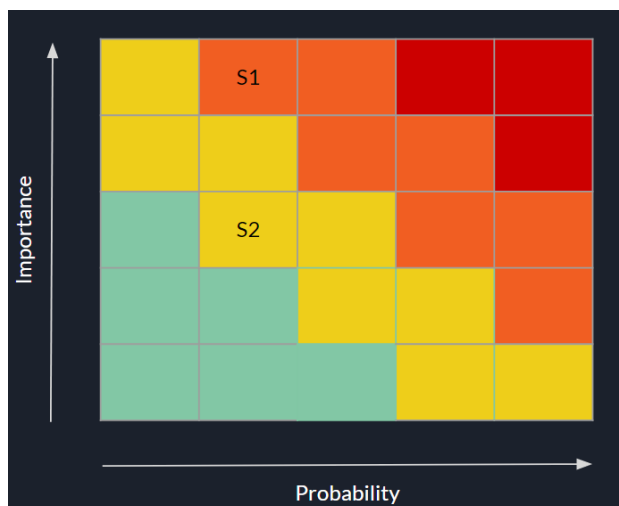
- T-4 Limited system performance due to musical diversity.
  - To mitigate this risk we will train the A.I. algorithms with multiple different types of music.
- T-5 Library Availability.
  - To mitigate this risk we will pull sample songs from many different sources.
- T-7 Applying incorrect correction rules when correcting music notation.
  - To mitigate this risk we will have many different sample songs to pull data from.



- T-8 Bias in the correction process.
  - To mitigate this risk we will attempt to limit bias as much as possible while training the A.I. algorithms.

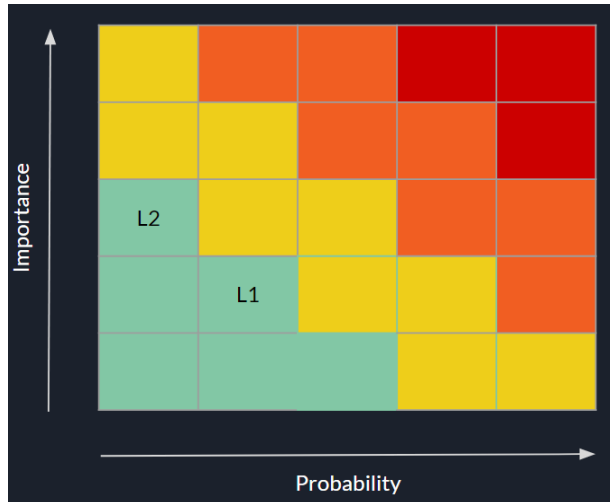
#### Real-World Product Risks:

- T-1 Multi-pitch detection Feature Difficulty.
- T-2 Note Tracking Feature Difficulty.
- T-3 Overestimating system performance.
- T-6 User Inputted Recognizability.



The above figure shows the security risks involved in the creation of the prototype

- S-1 Intellectual property.
  - Ensuring data privacy and security to prevent data leaks should mitigate the possibility of illegally obtaining a user's intellectual property.
- S-2 Application vulnerability.
  - Following best practices for application security will mitigate the ability of bad actors from exploiting vulnerabilities.



#### Legal Risks:

- L-1 Unlawful product use.
  - A watermark will be added to music to help prevent illegal distribution of sheet music.
- L-2 Copyright Law.
  - To prevent violation of copyright law A.I. algorithms will be trained using non-copyrighted works.

## 5 Glossary

**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.

## 6 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>
- Ruth, N., & Müllensiefen, D.. (2021). Survival of musical activities. When do young people stop making music?. PLOS ONE, 16(11), e0259105. <https://doi.org/10.1371/journal.pone.0259105>
- Scarlato, 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/>