June 13, 2021

Dr. Craig Scratchley
School of Engineering Science
Simon Fraser University
Burnaby, British Columbia
V5A 1S6

Re: ENSC 405W Requirement Specification for the **Sheet Music Transcriber** by **HappyJam**

Dear Dr. Scratchley,

As per the ENSC 405W Capstone A course Instructions, please find attached to this letter the functional requirements for the SMT (Sheet Music Transcriber) by HappyJam. The SMT takes in an audio sample of music played on any instrument or a small group of instruments and converts the music to a score ready to be played or further edited by a musician.

The document below covers the outline and requirement for our proposed product. Requirements are divided into sections and releases for a clear description of different subsystems and for the timescale these systems will be developed across. These sections include the algorithmic engine, the UI, documentation, economic concerns, and sustainability and safety requirements.

HappyJam is a multinational, diverse, and multidisciplinary team of passionate senior engineering students: Computer Engineers Matthew Marinets, Polina Bychkova, Haoran Hu, and Avital Vetshchaizer; System Engineer Jaskirat Arora; and Electronics Engineer Akaash Parajulee.

Thank you very much for your time and consideration. We truly appreciate your concern and time investment. Please let us know if you have any questions or concerns. You could contact our Chief Communications Officer Polina Bychkova anytime at pbychkov@sfu.ca.

Regards,

Matthew Marinets
Chief Executive Officer
HappyJam

SFU

School of Engineering Science

Requirements Specification for The Music Transcriber

ENSC 405: HappyJam (Company 1)

Matthew Marinets (301311930)
Akaash Parajulee (301308798)
Polina Bychkova (301269789)
Avital Vetshchaizer (301301019)
Jaskirat Arora (301319559)
Haoran Hu (301350932)

Abstract

The SMT (Sheet Music Transcriber) created by HappyJam aims to be an assistant for musicians of all levels. Our tool will be able to detect notes played within an audio sample and determine their properties, such as duration, volume, pitch, and timbre. It may then display the result as sheet music or a piano roll for a musician to begin practicing with. It may also export this result as MIDI or a MusicXML file for a musician to further edit, clean up, or re-use the result. This simplifies the process of transcription so that even a beginner may begin working with it. When applied in real-time situations, the SMT can even assist learners and teachers, giving them another view into what they're playing.

The requirements for our tool from a hardware, software, algorithmic, and UI perspective will be outlined in this document. Of particular note will be the algorithmic section of which the crux of our project lies. Furthermore, considerations from a documentative, economic, sustainability, safety, and engineering standard perspective will be discussed.

Table of contents

Version History

Date	Section	Change	Old	New
2021-06-17	2.7	Modify	H means the feature is… Capstone professors.	H means the feature was …for our project.
2021-06-17	4	add	-	R4.3-B
2021-06-17	4	add	-	R4.12-V1
2021-06-17	4	add	-	R4.13-B
2021-06-17	4	add	-	R4.14-B
2021-06-17	5, 6, 7	modify	Requirement names ending in -H	Now coloured red to indicate we're not doing them
2021-06-17	4	add	-	R4.13-V1
2021-06-17	4	add	-	R4.14-B
2021-06-17	4	add	-	R4.15-B
2021-06-17	4	add	-	R4.16-V1
2021-06-17	4	add	-	R4.17-B
2021-06-17	4	add	-	R4.18-B
2021-06-17	5	modify	R5.1-A	R5.1-B
2021-06-17	5	modify	R5.1.1-A	R5.1.1-B
2021-06-17	5	modify	R5.1.2-A	R5.1.2-B
2021-06-17	5	modify	R5.1.3-A	R5.1.3-B
2021-06-17	12	add	-	Spectrogram definition

1 Introduction

In the world of music, violins, concertos, electric guitars, and treble clefs, there is one thing that is common among them: sheet music. In the same way that recitation of a book aloud will impart a story upon the listener, musicians read from standardized notation called “sheet music” and transmit their story through their instrument to the listener. Performing sheet music from a score is generally easy and can be done in real time with some training. Listening to music and transcribing it is a comparatively much harder problem, taking longer and requiring even more specialized practice. This can stand as a barrier between musicians and their wish to create their own pieces of music or recreate what they hear. Furthermore, those who are just starting out in playing an instrument can have difficulty in properly identifying the difference between what they are playing and the notes on a page. If there were a way for beginners to visualize how their playing differs from that of the notes they should be playing, it would significantly accelerate their learning capabilities.

We at HappyJam have designed the SMT (Sheet Music Transcriber) to make this process easier. The SMT is an automated assistant which listens to an audio sample, picks out the notes and their properties, and compiles them into a format that can be read or further cleaned up by a musician. With this tool, the transcription of music into sheet music format can be transformed from a process that requires a large amount of specialized practice and long arduous hours to one that requires little to no effort. In doing so, the barrier between a musician and their creativity can be lowered significantly.

2 Product Overview

The SMT is built on three main components — the algorithms that detect notes and musical properties of an audio sample; the user interface that displays this data clearly to the user; and the user documentation that gives users better understanding and control of the product.

2.1 Algorithms

The SMT will utilize both whole-sample and time-progressive algorithms to transcribe the melody. When played, the audio will be converted into a spectrogram, which will then be processed through several signal processing stages to filter and analyze the audio. With further processing from neural networks, the algorithms determine a list of notes and their properties, a tempo measurement, and a set of timbre profiles that define how different instruments in the audio sound.

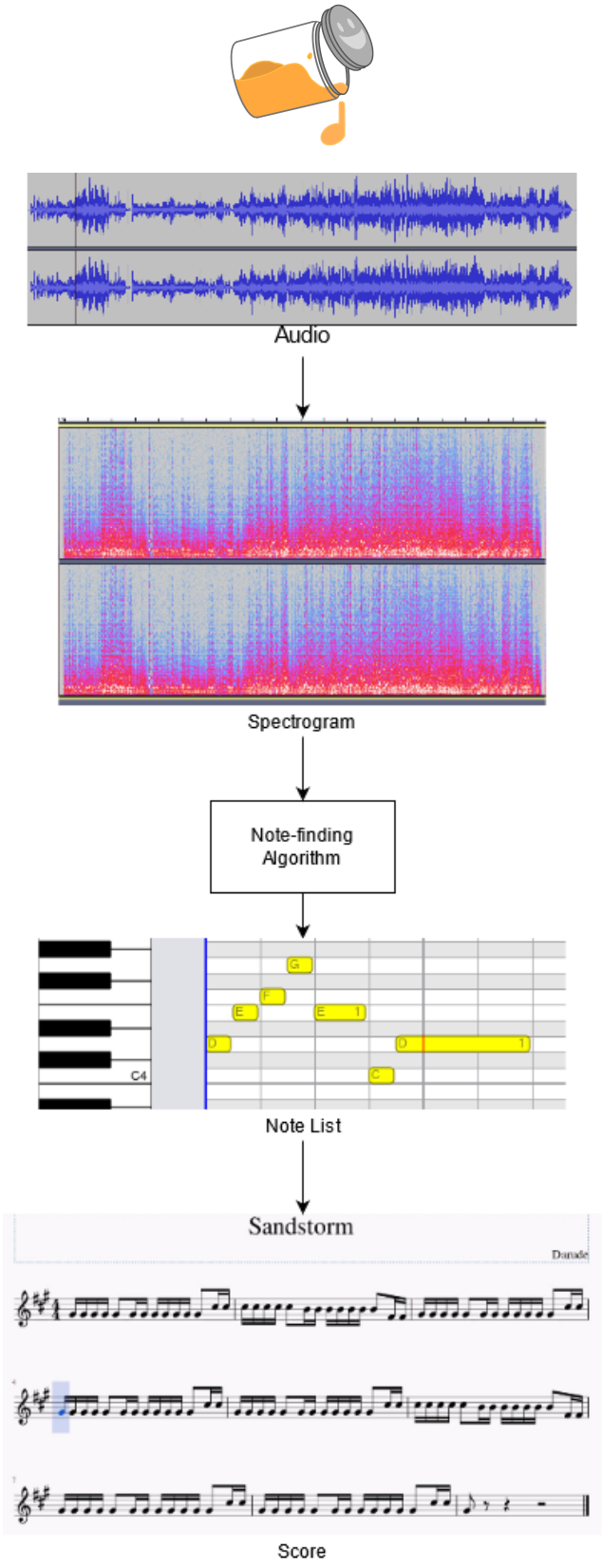

Figure 2.1.1: Flow diagram of different data steps under algorithmic processing

2.2 User Interface

The user interface will provide at least two views — piano roll and sheet music — created from the data processed by the algorithmic component. Each view offers specific functionalities that focus on different tasks. The user will be able to import recordings and export MIDI and sheet music through the interface. The user interface will also allow the user to change display settings and specify parameters controlling the note-finding algorithm.

2.3 Hardware

The primary hardware component of the sheet music transcriber will consist of a microphone recommendation in the user manual. HappyJam will be testing various microphones with different pick-up patterns mounted in a specific manner for best audio input to the software and user interaction.

2.4 User Documentation

The user documentation includes the functionality of the product, the general workflow of the product, and any safety or accessibility considerations. The document should help users familiarize themselves with the product quickly, and use the product safely.

2.5 Project timeline

Development of the project will happen in distinct phases.

- Alpha accounts for early prototyping and proof-of-concept, done by August 1 2021.
- The beta release will be a usable product ready for user testing, completed by October 31, 2021.
- Version 1 will be the first complete user-ready release, completed at the end of capstone in December 2021.
- Prior to a meeting with Craig Scratchley and Andrew Rawecz on June 14, 2021, there had been plans for a hardware device to act as a platform for this device. After the meeting, this component was removed from the scope of the project.

2.6 Requirements Format

The requirements will be identified with a character string in the format of

R[Section].[Requirement].[Elaboration Numbers]-[Phase]

Where

- [Section] is the section number of this document containing the requirement

- [Requirement] is the requirement number within that section
- [Elaboration Numbers] is an optional section; containing a string of period-separated numbers, used to number elaborations or specific sub-requirements relating to the main requirement.
- [Phase] refers to which phase of development the Requirement should be completed by
    - A means the feature is required for alpha
    - B means the feature is required for beta
    - V1 means the feature is required for the first full release
    - H means the feature was part of the project but after a recent meeting with Capstone professors it is not a required component for our project.

3 Algorithmic Requirements

The algorithmic system is responsible for finding notes and their qualities from an audio sample.

R3.1-A	The SMT must be able to analyze an audio sample to record a list of notes and their properties
R3.1.1-A	The start time of a detected note must be recorded
R3.1.2-A	The pitch of a detected note must be recorded
R3.1.3-B	The duration of a detected note be recorded
R3.1.4-B	The relative volume of a detected note should be recorded
R3.1.5-B	The timbre of a detected note should be recorded
R3.1.5.1-B	The algorithm must have an input specifying an upper bound on the number of distinct timbres to record
R3.1.5.2-B	The SMT should determine a list of timbres that best accounts for the power distribution across the times and frequencies in the audio sample
R3.2-B	The Sheet Music Transcriber must determine the tempo of music in an audio sample
R3.2.1-V1	The Sheet Music Transcriber must detect tempo changes in the music
R3.2.2-V1	The Sheet Music Transcriber should detect the time-signature of the music

R3.3-A	The Sheet Music Transcriber should locate peaks in power at particular frequencies and times in the audio sample's spectrogram

4 UI Requirements

The UI is the main method users have for viewing and interacting with the results of the algorithmic analysis.

R4.1-A	The software must be able to open a window on a desktop PC running Windows
R4.1.1-A	The user must be able to stop the software and close the window by pressing the window's X button.
R4.2-A	The software must have a submenu to select an audio file to import
R4.2.1-A	The import feature must take uncompressed audio files in .wav format
R4.2.2-B	The import feature must be able to take compressed audio formats, including .mp3 and .ogg
R4.3-B	The software must be able to record from an external microphone
R4.4-A	The user interface must display pitch, timing, and duration of notes
R4.5-A	The user interface must have multiple formats for displaying notes and audio that a user may choose between
R4.5.1-A	One format must display an audio waveform of the audio sample
R4.5.2-B	One format must display a piano roll of detected notes
R4.5.3-B	One format must display a score of detected notes in sheet music notation
R4.5.3.1-B	The user must be able to change the key signature of displayed sheet music
R4.5.3.2-B	The user must be able to change the base note duration for the displayed sheet music between half notes, quarter notes, eighth notes, or sixteenth notes
R4.5.4-B	One format may display a spectrogram of audio input
R4.5.4.1-B	The user must be able to place markers with text annotations on specific

	points of the spectrogram
R4.5.5-V1	One format may display a timbre analysis view for detected timbres
R4.5.6-A	The user must be able to zoom in and out of the audio view formats
R4.6-B	The user interface must have a panel of analysis parameters that a user may configure
R4.6.1-B	The user must be able to configure thresholds on what the software considers a note
R4.6.2-V1	The user must be able to configure the maximum number of timbres to detect in an audio sample
R4.7-V1	The user should be able to edit existing notes in the piano roll view
R4.8-B	The user must be able to export generated MIDI to a file
R4.8.1-B	The user must have an input method to write track information to an exported MIDI file
R4.9-V1	The user must be able to export sheet music in MusicXML format
R4.9.1-V1	The user must be able to specify score presentation such as title and performer
R4.9.2-V1	The user must be able to export only certain parts, filtered by timbre
R4.10-V1	The user must be able to label identified timbres
R4.10.1-V1	Timbre labels should be reflected in the UI and in the part names of exported scores
R4.11-A	The UI must display in dark mode
R4.12-V1	The UI must not crash
R4.13-V1	The software may be able to recognise instrument(s) being analysed
R4.14-B	The user must be able to force stop the analysis if it is taking too long
R4.15-B	The software must be able to display an error message corresponding to the error it encounters
R4.16-V1	The user should be able to modify the note list, including adding notes

	or changing note properties
R4.17-B	The user must have the option to save their current project for later
R4.17.1-B	The save file must include the currently loaded audio data
R4.17.2-B	The save file must include the current notes and their properties
R4.17.3-B	The save file must include the generated timbre profile
R4.17.4-B	The save file must include the sheet music view options
R4.18-B	The user must have the ability to load a previously saved project

5 Hardware Requirements

The Sheet Music Transcriber hardware is a small device that can fit on a music stand to record audio for analysis.

R5.1-B	The user should be able to record the music being played through the microphone
R5.1.1-B	The microphone must have low current consumption (< 500 uA)
R5.1.2-B	The microphone must have high signal-to-noise Ratio (SNR) (over 60 dBA)
R5.1.3-B	The microphone must be able to operate in human audible range (20 Hz-20kHz)
R5.2-H	The device must be able to playback the recognized note through speakers
R5.2.1-H	The speaker must be small in size to fit inside the device (upto 60mm diameter)
R5.2.2-H	The speaker must have low power consumption (< 1 W)
R5.2.3-H	The speaker must have low input impedance (< 20 ohms)
R5.3-H	The device should have a sizable screen (at least 4”)
R5.3.1-H	The screen resolution should be at least 480x272 pixels
R5.3.2-H	The screen life hours should be decent

R5.4-H	The microcontroller/processor being used should be able to compute the algorithm efficiently and quickly (less than 0.1 sec for individual notes)
R5.4.1-H	The microcontroller should not be bulky (less than 100g)
R5.4.2-H	All the microcontroller pins must be programmable
R5.4.3-H	The microcontroller must have a power jack
R5.4.4-H	The microcontroller must be compatible with battery or AC-DC adapter
R5.4.1-H	The system should have adjustable gain for optimal recording
R5.5-H	The battery should be able to power the device for at least 3-4 hours
R5.5.1-H	The battery must be Li-ion battery
R5.5.2-H	The battery capacity must be at least 2500 mAh
R5.5.3-H	The battery life cycle should be over 1000 charge cycles
R5.6-H	All components used must be able to operate within 3.3-5V supply voltage
R5.7-H	The user interface components should be able to withstand regular use
R5.8-H	The component housing should be portable (at max 18x24cm)

6 Documentation Requirements

A digital or physical manual (if a hardware component is involved) helps a user to understand the functionality of the product, the general workflow of the product, the safety measurements and maintenance information of the product.

R6.1-B	The manual must specify the difference between the views of the software and the purpose of each view
R6.2-B	The manual must demonstrate how to change between views
R6.3-B	The manual must clarify the analysis parameters the user may configure
R6.4-B	The manual must provide an example of the general workflow
R6.5-B	The manual must demonstrate how to import music

R6.6-V1	The manual must demonstrate how to export sheet music
R6.7-B	The manual must demonstrate how to export MIDI files
R6.8-B	The manual must list all the discrete components of the hardware with pictures
R6.9-H	The manual must describe the function of each hardware button
R6.10-H	The manual must show how to set up the device for recording
R6.11-H	The manual must specify the power source requirement
R6.10-H	The manual must show how to boot up/turn off the hardware device
R6.11-H	The manual must inform user how to store the device properly
R6.12-H	The manual must include disposal and recycling instructions for the hardware

7 Economic Requirements

R7.1-H	The microphone must be less than $10 [4]
R7.2-H	The speaker must be less than $3 [5]
R7.3-H	The battery must be less than $2 [3]
R7.4-H	The screen must be less than $20 [6]
R7.5-H	The microcontroller/processor should be less than $15 [7]
R7.6-H	The housing should be less than $20
R7.8-H	The overall cost of hardware device should be less than $80
R7.9-H	The retail price for the hardware should be $150
R7.10-V1	Desktop software will be open-source

8 Sustainability & Safety

HappyJam is committed to maintaining the safety and sustainability of the community it is a part of. We intend our products to be maintainable and have a planned lifecycle from its production to disposal and recycling. As most of our product is largely software-based, its environmental impact is dependent on the power consumption of

our software as well as the cleanliness of the energy used to power the computer the software runs on. The hardware components therefore represent a larger proportion of potential environmental impact, and as such we are committed to ensuring its recyclability, maintainability, and re-usability.

8.1 Software

In Canada there are no specific regulations related to consumer software safety outside of conventional laws already in the penal code. As such we are committed to maintaining an internal safety policy related to our software

R8.1-A	Use of microphone must be constrained to user activity
R8.2-A	Use of speakers must be constrained to user activity
R8.3-V1	Active power consumption must be reasonable
R8.4-V1	Idle power consumption must be reasonable
R8.5-B	Software will not create any security vulnerabilities in a device

8.2 Hardware

Any required hardware would either be owned by the user already or be made of conventional materials. Most components will be RoHS compliant. The battery of the device would be a rechargeable Lithium Ion battery. A rechargeable battery will cut down on waste and may be recycled through existing battery recycling programs.

R8.7-H	Speaker	Electronics Recycler	RoHS compliant
R8.8-H	Microphone	Electronics Recycler	RoHS compliant
R8.9-H	Microcontroller	Electronics Recycler	RoHS compliant
R8.10-H	Screen	Electronics Recycler	RoHS compliant
R8.11-H	Lithium Ion Battery	Battery Recycling Program	RoHS compliant
R8.12-H	Housing	Curbside/Plastic Recycler	Recycled Plastic

9 Engineering Standards

As the Sheet Music Transcriber is largely stand-alone, it does not need to conform to many standards. Its reliance on standards is mainly in the input and output layers — exporting note values and scores, and importing audio.

9.1 MIDI 1.0 Standard

The Sheet Music Transcriber's MIDI export must comply with Standard MIDI Files 1.0 [1].

9.2 MusicXML 4.0

The Sheet Music Transcriber's score export must comply with the MusicXML Specification [2].

9.3 RoHS (Restriction of Hazardous Substances) Compliance

All the components used should be RoHS compliant including battery [3], microphone [4], speaker [5], screen [6], and the microcontroller [7]

9.4 Canada Consumer Product Safety Act

The device must be compliant with the Consumer Product Safety Act in order to be safe for users in Canada.

10 Alpha Acceptance Test Plan

Requirements labeled for completion in the alpha release will have an acceptance criterion listed

10.1 Software Acceptance Test Plan

R3.1-A	The SMT must be able to analyze an audio sample to record a list of notes and their properties
Criterion:	The program can output text representing the note
R3.1.1-A	The time of a detected note must be recorded
Criterion:	The output note list includes the offset into the audio sample, in either seconds or samples, at which every note begins
R3.1.2-A	The pitch of a detected note must be recorded

Criterion:	The output note list includes the pitch name, including the octave number, of every note in the list
R4.1-A	The software must be able to open a window on a desktop PC running Windows
Criterion:	Running the executable opens a window.
R4.1.1-A	The user must be able to stop the software and close the window by pressing the window's X button.
Criterion:	After opening the window, press the X button and ensure the window closes. Check Task Manager to ensure the process is no longer running.
R4.2-A	The software must have a submenu to select an audio file to import
Criterion:	Clicking the audio import option in the submenu opens a file explorer window.
R4.2.1-A	The import feature must take uncompressed audio files in .wav format
Criterion:	Opening the audio import file explorer and selecting a .wav file correctly opens and displays the audio data.
R4.4.5-A	The user must be able to zoom in and out of the audio view formats
Criterion:	After importing an audio sample, turning the scroll wheel stretches or squashes the audio waveform display.
R4.10-A	The UI must display in dark mode
Criterion:	On opening the UI window, the user's eyes are soothed with dark-grey colours.

10.2 Hardware Acceptance Test Plan

R5.1-A	The user should be able to record the music being played through the microphone
R5.1.3-A	The microphone must be able to operate in human audible range (20 Hz-20kHz)
Criterion:	Audio recorded should be noise free
R5.2-H	The device must be able to playback the recognized note through

	speakers
Criterion:	Playback should be clear and noise free
R5.3.2-H	The screen life hours should be decent
Criterion	Screen on/off life cycle should be tested
R5.4-H	The microcontroller/processor being used should be able to compute the algorithm efficiently and quickly (less than 0.1 sec for individual notes)
Criterion:	Testing for heat generation: temperature of the microcontroller should not exceed 50 degrees C
R5.5-H	The battery should be able to power the device for at least 3-4 hours
R5.5.2-H	The battery capacity must be at least 2500 mAh
R5.5.3-H	The battery life cycle should be over 1000 charge cycles
Criterion:	Battery life cycle should be decent when in use
Criterion:	Working hours of device on one full charge should be tested
R5.7-H	The user interface components should be able to withstand regular use
Criterion:	Operating temperature of the device should be between -4 to 60 degrees C

11 Conclusion

With the HappyJam's SMT (Sheet Music Transcriber), the work of transcribing sheet music from audio will be much easier and faster. Learners, teachers, performers, and composers alike will have the barrier between them and their musical creativity significantly lowered. Learners and teachers will be able to analyze what they have played in more formats, and thus accelerate their learning. Composers and performers will have an easier time creating and viewing music and analyzing their performances. Transcribers will have a powerful tool for working quickly, with a strong starting point for creating scores.

We at HappyJam want to see this tool become reality. In the above document we have enumerated the specifications we have determined for the product that will make this dream a reality. The milestones for each of the Alpha, Beta, and Version 1 phases are as follows: August 2021, October 2021, and December 2021. The Sheet Music Transcriber will be founded on a clear UI and intelligent note-finding algorithms, with affordable hardware recommendations to back it up.

As a nature-friendly company, HappyJam also sees the importance of sustainability and recycling. We aim to be an example to future generations on the importance of providing a sustainable and safe product.

12 Glossary

Chord

A group of notes played at the same time.

Note

The smallest unit of musical analysis. A note is a pulse of sound with a few properties:

- Timing: when the note is played
- Duration: How long the note lasts
- Pitch: the fundamental frequency of the sound
- Volume: the amplitude of the sound
- Timbre: waveform of the sound; timbre varies with instrument

Pitch

Pitch is the name musicians give to the fundamental frequency of a note. Pitches have names that loop through [A, A#, B, C, C#, D, D#, E, F, F#, G, G#], sometimes notated with a number following the pitch symbol to indicate which octave it falls in. Pitch names are spaced logarithmically, so A4 (440 Hz) is twice as high as A3 an octave below (220 Hz).

Score

A score is a document written in sheet music that describes how to play a song.

Sheet Music

Sheet music is the primary notation used in Western music to describe how to play a particular part or song.

Spectrogram

A two-dimensional image acquired from an audio sample representing the relative power carried by across time and frequency axes.

Timbre

Timbre is the quality of a note that differs between instruments, manifesting physically as a different waveform. As the wave is periodic, timbre may be described as the relative power of a note's harmonics — that is, the relative power present in frequencies around integer multiples of the fundamental frequency.

Tempo

The speed by which a section of the music is played, expressed in beats per minute (bpm).

13 References

[1] Standard MIDI Files 1.0. Los Angeles, CA: The MIDI Manufacturers Association, 1996.

[2] “MusicXML 4.0,” w3c.github.io, Jun. 03, 2021. https://w3c.github.io/musicxml/version-history/40/ (accessed Jun. 14, 2021).

[3] Best Price Rechargeable Li Ion Battery 3.7v 2600mah 2800mah 3000mah 18650 Li Ion Battery Pack - Buy 18560 Li Ion Battery,Li Ion Battery,Li Ion Battery Pack Product on Alibaba.com https://www.alibaba.com/product-detail/Best-price-rechargeable-li-ion-battery_50043020744.html?spm=a2700.galleryofferlist.topad_creative.d_title.6ca7c634FjVrhY

[4] SparkFun Analog MEMS Microphone Breakout - ICS-40180 - BOB-18011 - SparkFun Electronics https://www.sparkfun.com/products/18011

[5] Speaker - 0.5W (8 Ohm) - COM-09151 - SparkFun Electronics https://www.sparkfun.com/products/9151

[6] 4" Inch Lcd 480x800 With Touch Panel Tft Display Screen - Buy 4" Inch Lcd,4" Inch Lcd 480x800,4" Inch Lcd 480x800 Tft Display Screen Product on Alibaba.com https://www.alibaba.com/product-detail/4-inch-LCD-480x800-with-touch_60235653025.html?spm=a2700.galleryofferlist.normal_offer.d_title.5e0c2e2bsv0JYp

[7] Mega R3 Neutral Without Logo Mega R3 For Arduinos - Buy Mega R3,Mega,Neutral Without Logo Product on Alibaba.com https://www.alibaba.com/product-detail/Arduinos-Mega-R3-Neutral-Without-LOGO_60864016075.html?spm=a2700.galleryofferlist.normal_offer.d_title.463d5463rQCHsq&s=p