



Project Report

Name Recorder Automation

April 2021
Version 0.1

ISA-PM-2021-01-13-MTech1-2M-NameRecorderAutomation

Team Member: Vincent Ng, Zhou Zhe



1. Project Summary

The Name Recorder Automation is an intelligent software agent product that is able to automate record, normalize, add and remove silent padding of sound of student names read by voice talent.

Republic Polytechnic (RP) deploys a name reader system for graduation ceremonies every year that makes use of pre-recorded student names that will be played as the students go up on stage to receive their diploma certificates.

The reason for using pre-recorded audio is because when a person is reading hundreds of names live, it is very stressful on the reader to pronounce the name right each time. Mispronouncing the student's name is embarrassing for both the student and the school.

Thus, every year before the graduation ceremonies, thousands of student names are pre-recorded and verified by RP staff.

The recording process involves a voice talent who will read and record the names, and an audio technician who post-edits the recordings. This is a very repetitive and error-prone process, with each name taking on average 6 seconds to record, and 60 seconds to post-process. Multiplied by the number of students graduating each year, around 90 man-hours are spent by RP staff on this process. This makes it an excellent candidate for process automation.

This automation system is able to help RP save manpower costs and also improve the quality and consistency of the name recordings. The benefits include:

1. Less human errors in post-processing
 2. Consistency in both volume and silence padding
 3. Manpower savings of 82 man-hours/year
-



2. Product Plan & Market Research

Product Plan

Through interviews with the RP staff tasked to perform the recordings, we understood pipeline of the recording process, which are the following:

1. Voice talent is prompted to pronounce the student's name
2. Audio technician performs the recording
3. Audio technician then does post-processing which includes
 - Removal of background room noise
 - Adjust pre and post recording silence to 0.5 second
 - Normalize recordings to optimal playback volume
 - Save recording as <student ID>.wav file


Our complete automation system will automate all of these tasks.

The Name Recorder Automation (NRA) system will be released and marketized in two phases:

1. Firstly, the NRA product will be developed, tested and utilized for graduation ceremonies in Republic Polytechnic, which will be fine-tuned and improved based on the feedback from RP staff or students.
2. Once the NRA product has been optimized and stabilized that will be applied for most of ceremonies in different schools such as primary school, secondary school etc. The NRA product will be improved after getting more and more feedback from different customers. Additionally, we will add new features into NRA product in order to fulfill various requirements based on market demands.

Market Research

According to statistic data from Singapore government's one-stop portal (data.gov.sg), there are 326 government or gov-aided or other public supported



schools from primary school to university in Singapore (approx. 380,000 students in 2019). There will be huge impact and cost-save for name reading in each graduation ceremony.

The RNA product can be also generalized and utilized in different event where there is requirement to auto reading name.

3. System Design

The Name Recorder Automation system orchestrates the full audio reading and processing using friendly GUI to automate the repetitive error-prone tasks:

1. Record name read by voice talent
2. Remove background room noise (if any)
3. Normalize audio to optimal playback volume
4. Adjust pre- and post- recording silence to 0.5 seconds
5. Save audio file as <Student ID>.wav

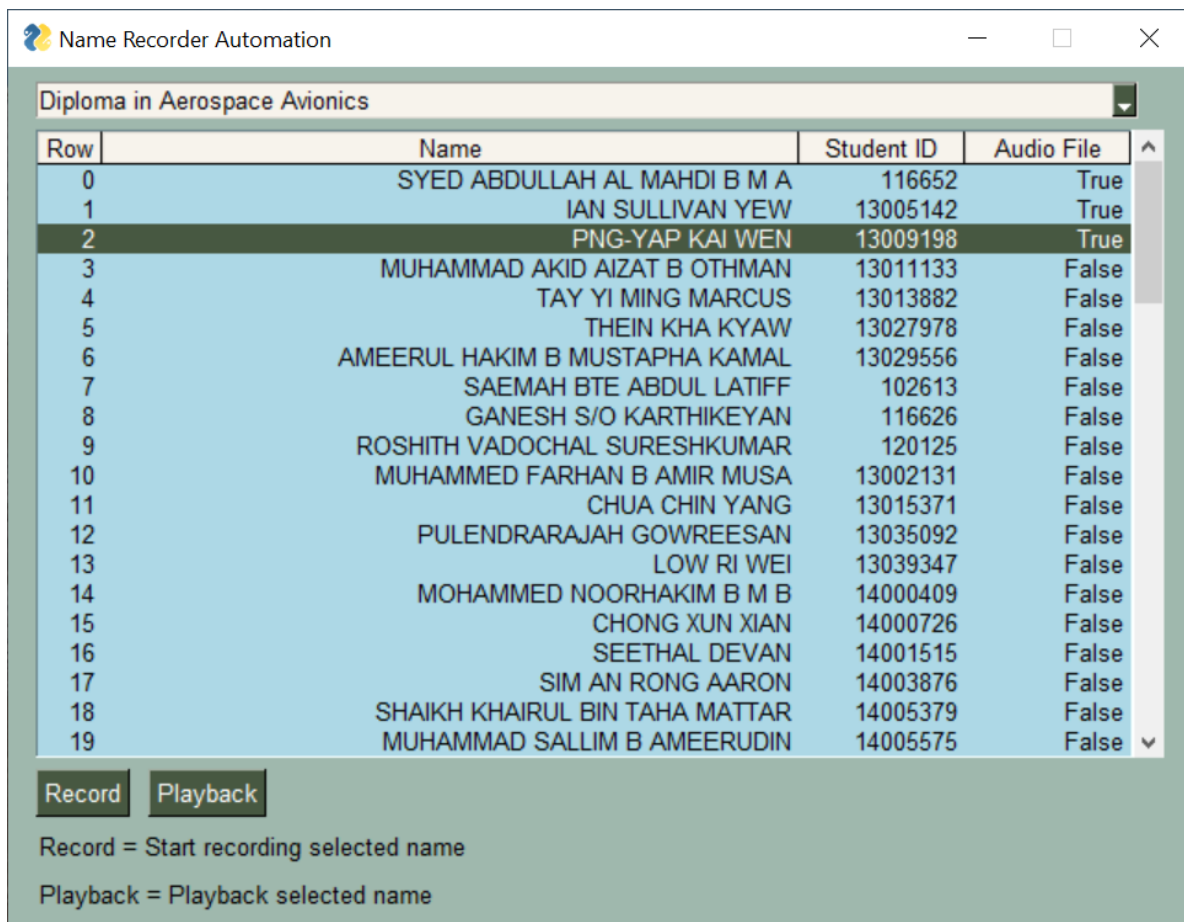
The Name Recorder Automation system comprises of five software agents to perform different tasks, they are followings,

1. Prompter agent
2. Noise removal agent
3. Normalization agent
4. Silence detection and truncation agent
5. Output agent

Prompter Agent

The prompter agent is a visualized and user-friendly GUI that shows all students from different majors and allow voice talent to filter diplomas using dropdown list and pronounce any selected student name. there is no need for cue from audio technician and just click 'Record' button to start recording and click 'Playback' button to review the recording quality. Any selected name in list that voice talent can re-record whenever it's needed.

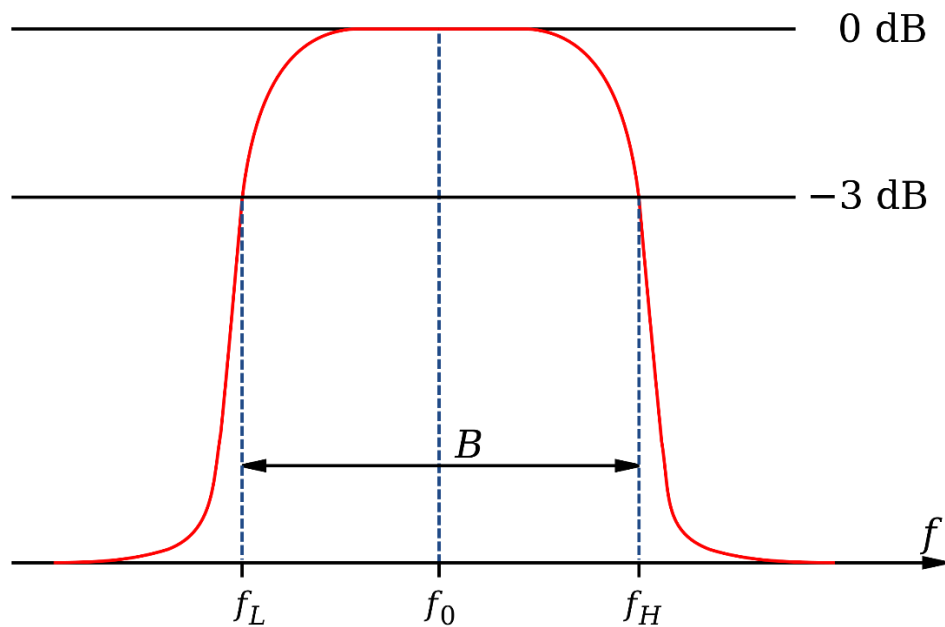
The below figure shows the functions that voice talent can use:



Noise Removal Agent

The Noise Removal Agent is able to detect the background room noise generated during recording period, which performs band-pass filtering on recording. As human voice normally resides the vocal range between 50Hz and 8KHz. Those

non-essential background room noises are removed by band-pass filtering techniques.



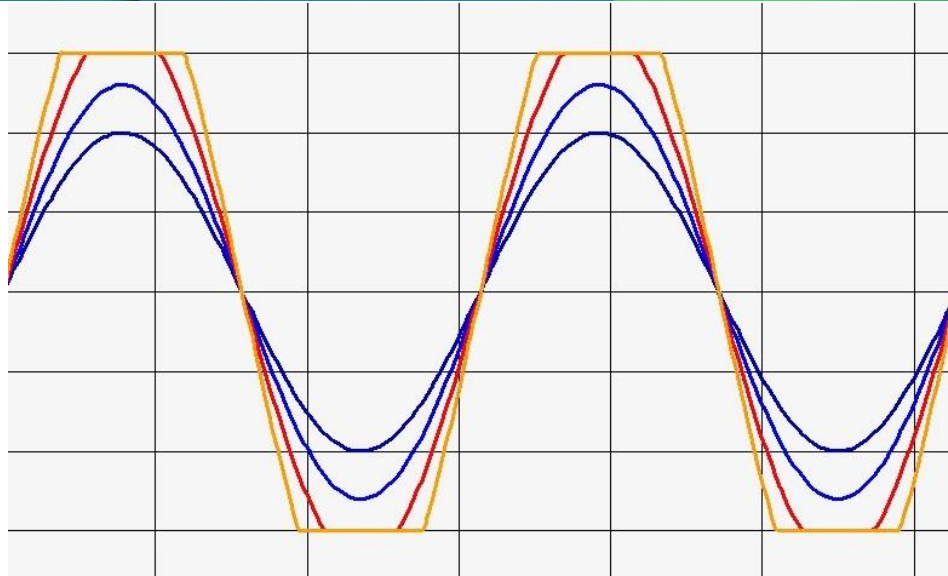
Normalization Agent

The Normalization Agent is useful when the same voice talent may pronounce differently at different times of period.

1. Recording that are too loud are softened
2. Recordings that are too soft are amplified

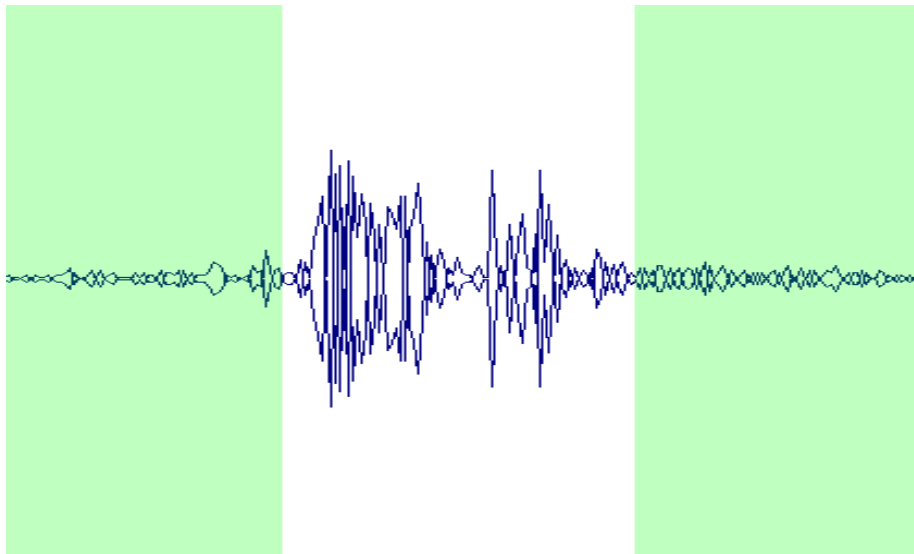
The audio can play consistently in playback volume after audio file is processed by the normalization agent.

As shown the below figure, the voice is normalized within consistency range.



Silence Detection and Truncation Agent

As requirement in RP graduation ceremony, 0.5 second silent must be added pre and post for each student name in order to ensure playback will not start or stop abruptly. any silent voice must be detected and removed before adding required silent voice.



Output Agent

The output agent is used to save audio files with alignment where the student's name is selected in Name Recorder Automation GUI. The following benefits when output agent is applied:

- Recordings are automatically named as <Student ID>.wav in Audio folder
- GUI will indicate names that have/have not been recorded
- Allows for recording done over multiple sessions without losing track of progress

The screenshot displays the 'Name Recorder Automation' application window. The main window has a title bar and a menu bar. Below the menu bar is a dropdown menu set to 'Diploma in Aerospace Avionics'. The main area contains a table with four columns: 'Row', 'Name', 'Student ID', and 'Audio File'. The 'Audio File' column is highlighted with a red box. Below the table are 'Record' and 'Playback' buttons. At the bottom, there are instructions: 'Record = Start recording selected name' and 'Playback = Playback selected name'.

Row	Name	Student ID	Audio File
0	SYED ABDULLAH AL MAHDI B M A	116652	True
1	IAN SULLIVAN YEW	13005142	True
2	PNG-YAP KAI WEN	13009198	True
3	MUHAMMAD AKID AIZAT B OTHMAN	13011133	False
4	TAY YI MING MARCUS	13013882	False
5	THEIN KHA KYAW	13027978	False
6	AMEERUL HAKIM B MUSTAPHA KAMAL	13029556	False
7	SAEMAH BTE ABDUL LATIFF	102613	False
8	GANESH S/O KARTHIKEYAN	116626	False
9	ROSHITH VADOCHAL SURESHKUMAR	120125	False
10	MUHAMMED FARHAN B AMIR MUSA	13002131	False
11	CHUA CHIN YANG	13015371	False
12	PULENDRARAJAH GOWREESAN	13035092	False
13	LOW RI WEI	13039347	False
14	MOHAMMED NOORHAKIM B M B	14000409	False
15	CHONG XUN XIAN	14000726	False
16	SEETHAL DEVAN	14001515	False
17	SIM AN RONG AARON	14003876	False
18	SHAIKH KHAIRUL BIN TAHA MATTAR	14005379	False
19	MUHAMMAD SALLIM B AMEERUDIN	14005575	False

Record Playback

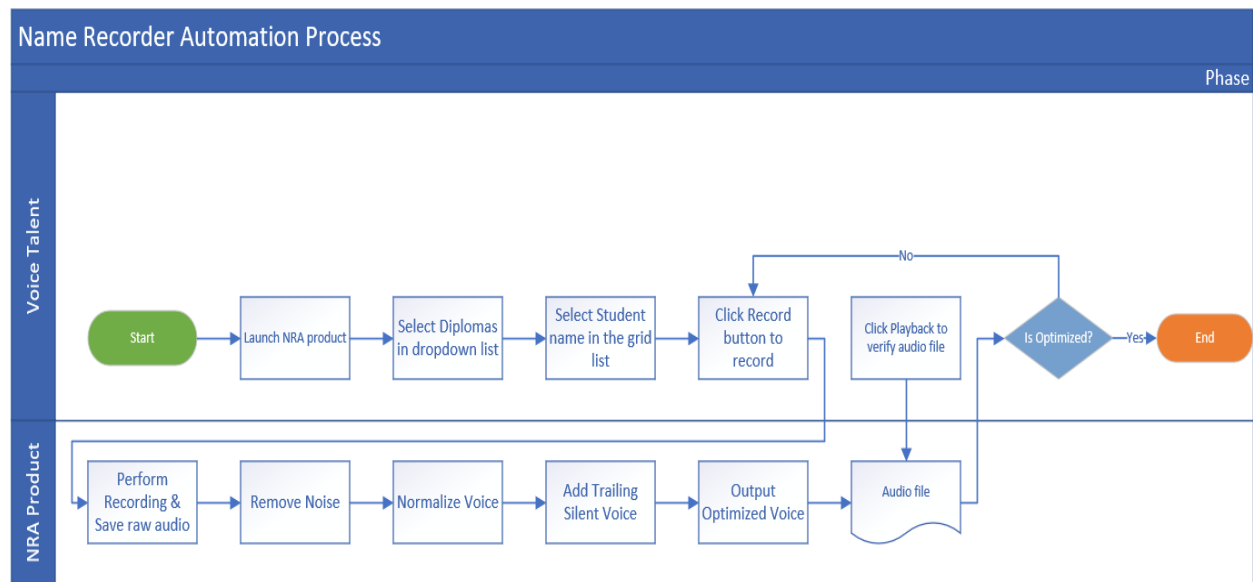
Record = Start recording selected name
Playback = Playback selected name

The file explorer shows a list of audio files in a folder. The files are named with student IDs followed by '.wav'. The columns are 'Name', 'Size', and 'Date modified'.

Name	Size	Date modified
102613.wav	560 KB	3/31/2021 6:12 PM
116618.wav	982 KB	4/6/2021 9:48 AM
116652.wav	517 KB	3/31/2021 5:52 PM
122945.wav	982 KB	4/6/2021 11:10 A...
13000183.wav	979 KB	4/6/2021 9:52 AM
13001075.wav	981 KB	4/6/2021 9:53 AM
13005142.wav	981 KB	4/6/2021 10:56 A...
13009198.wav	517 KB	3/31/2021 5:52 PM
13011133.wav	517 KB	3/31/2021 5:59 PM
13013882.wav	305 KB	3/31/2021 6:05 PM
13027978.wav	466 KB	3/31/2021 6:06 PM
13029556.wav	696 KB	3/31/2021 6:12 PM
14000730.wav	1,034 KB	4/6/2021 11:13 A...
14000776.wav	1,034 KB	4/6/2021 11:13 A...

Process Flowchart

The below process flowchart illustrates the whole process flow where voice talent pronounces name recording and NRA system processes and saves optimized audio files into audio folder.





4. System Implementation

The Name Recorder Automation product is the application-based system that is fully designed and implemented using python programming language and leverages on several python open-source libraries such as pysimplegui, sounddevice and pydub etc.

The library pysimplegui is used to create GUI windows using different widgets easily and quickly.

The library sounddevice is used to perform sound recording and generating to physical files.

The library pydub is used to perform sound silence detection and truncation as well as add pre and post silent sound.

5. Appendix
