Deepcut: A Robot that Raps
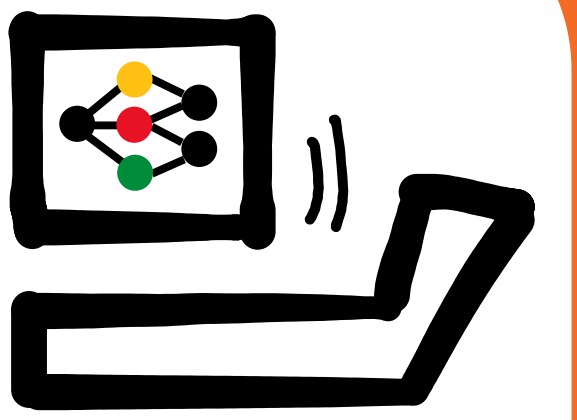
Jarrod Luckenbaugh (EE) **Heriberto Felipe (EE)** **Shah Misbahul Huda (EE)**

Hamza Aamer Bhatti (CE) **Luz Martinez-Lucas (EE)** **Delia Victoria Trejo (CE)**

UTDALLAS

Senior Design II **Faculty Advisor: Dr. Neal Skinner** **Team Number: 1145**
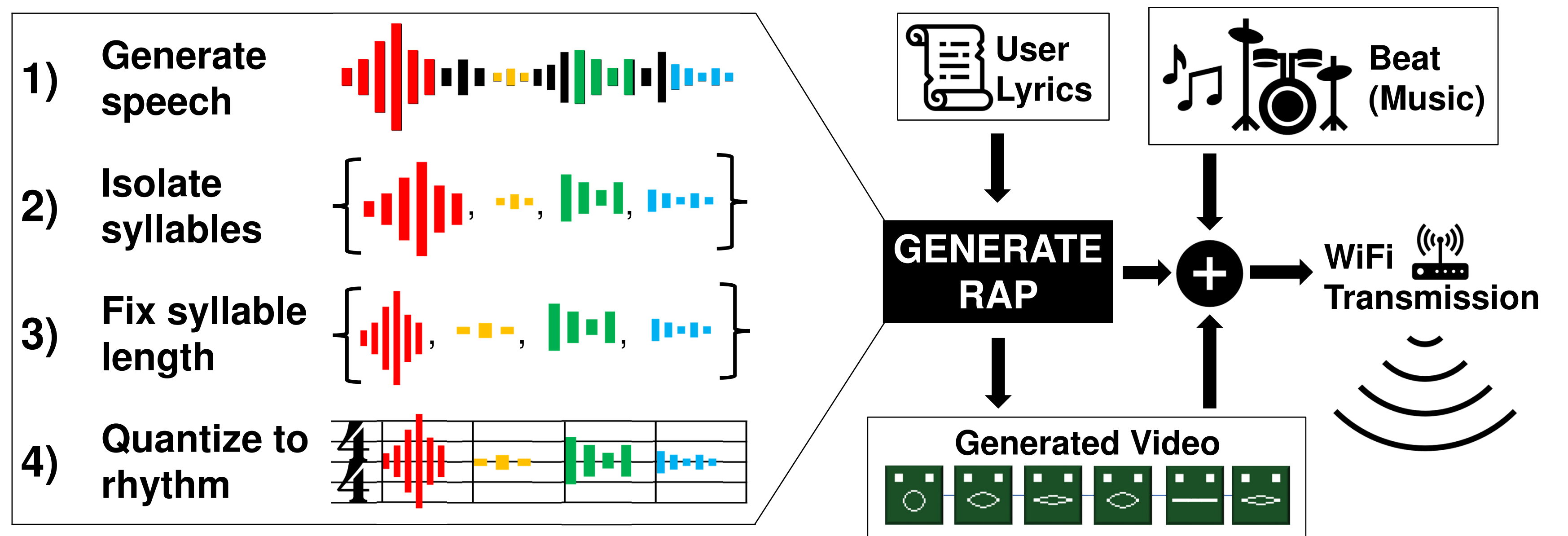
GOALS

- **Our team will create a system that is capable of the synthesis and performance of music with a rapped vocal style**
- Three main subsystems:
  - Front End Interface – UI and look of the robot
  - Vocal Synthesizer – creates the rapped vocals
  - Performance Module – performs the music on the robot
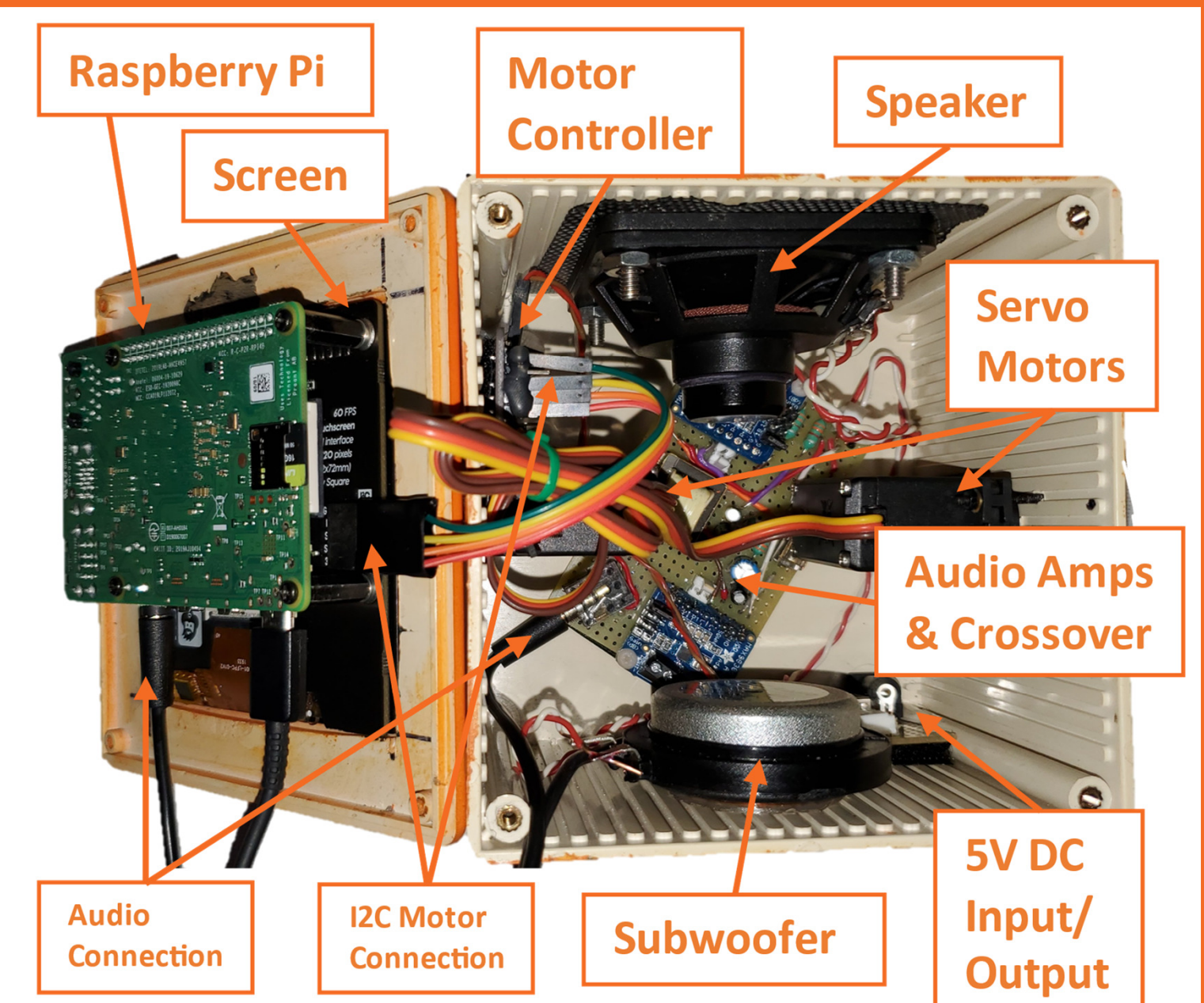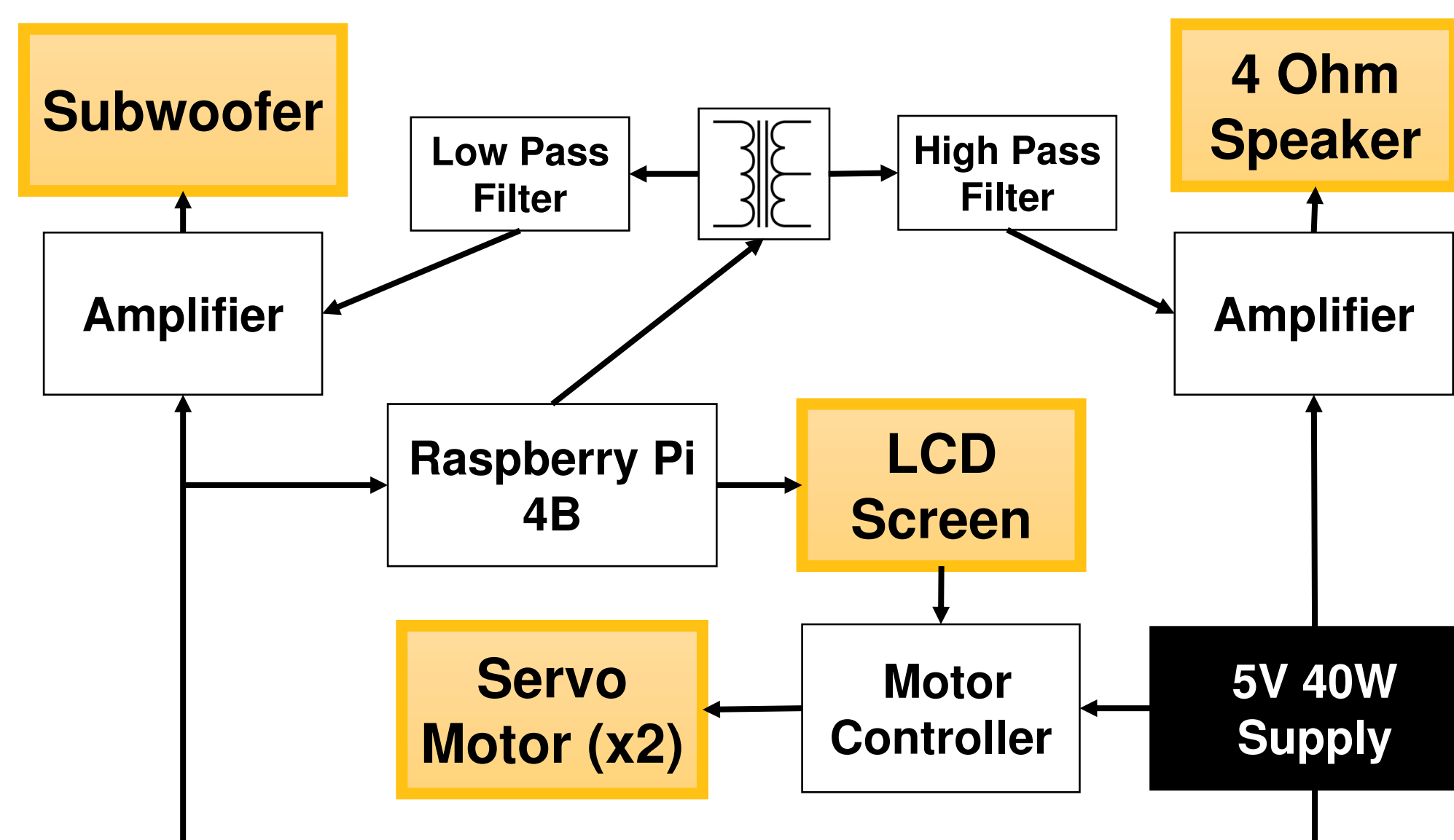
USER INTERFACE
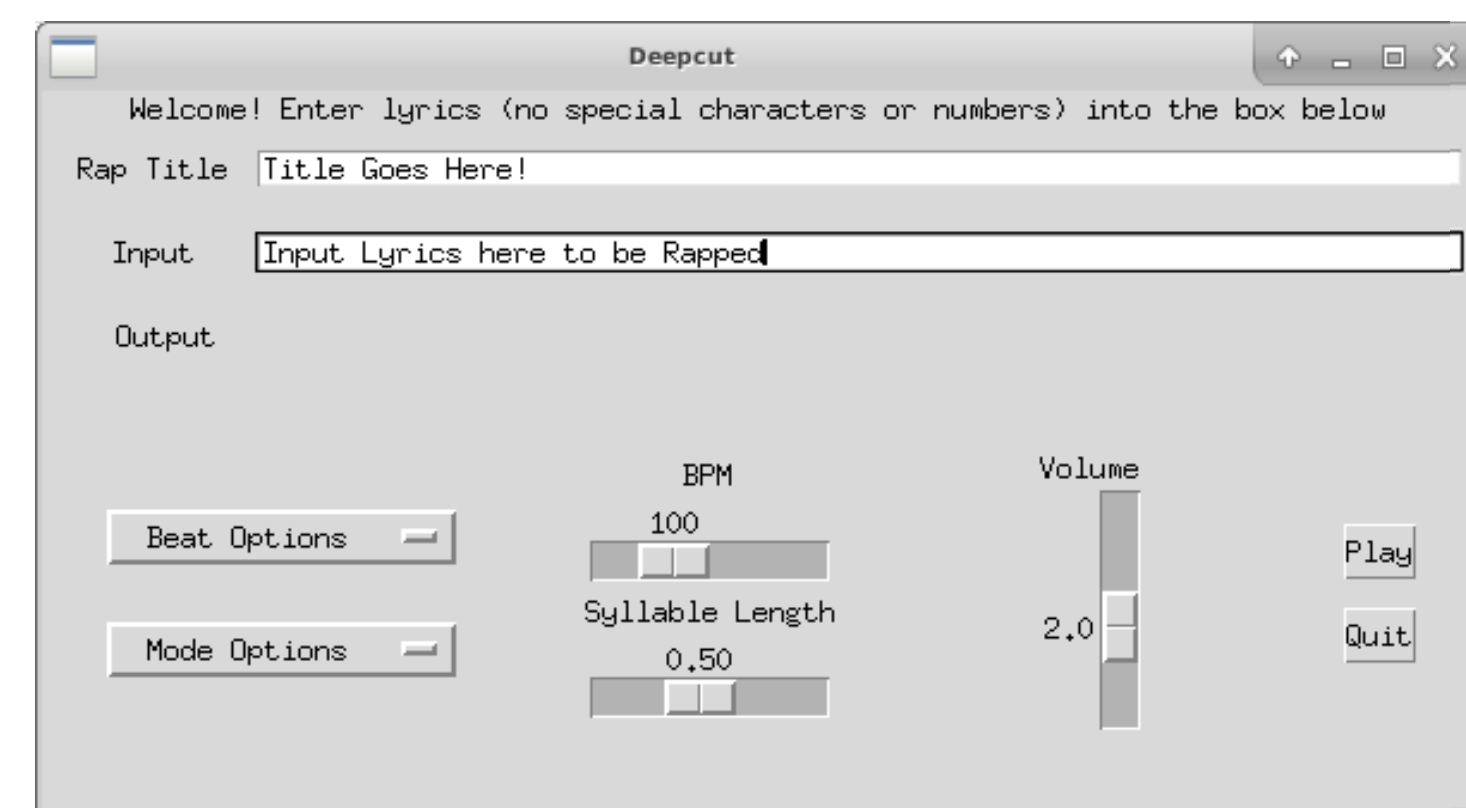
VOCAL SYNTHESIZER

- End to end rap synthesis system
- Four step speech-to-rap conversion using NVIDIA's Flowtron test-to-speech and the Montreal Forced Aligner
- User Input
  - Title – file naming
  - Input text – rap lyrics/transcript
  - Beat option – choice of background music
  - Mode option – choice of rhythm
  - BPM – beats per minute
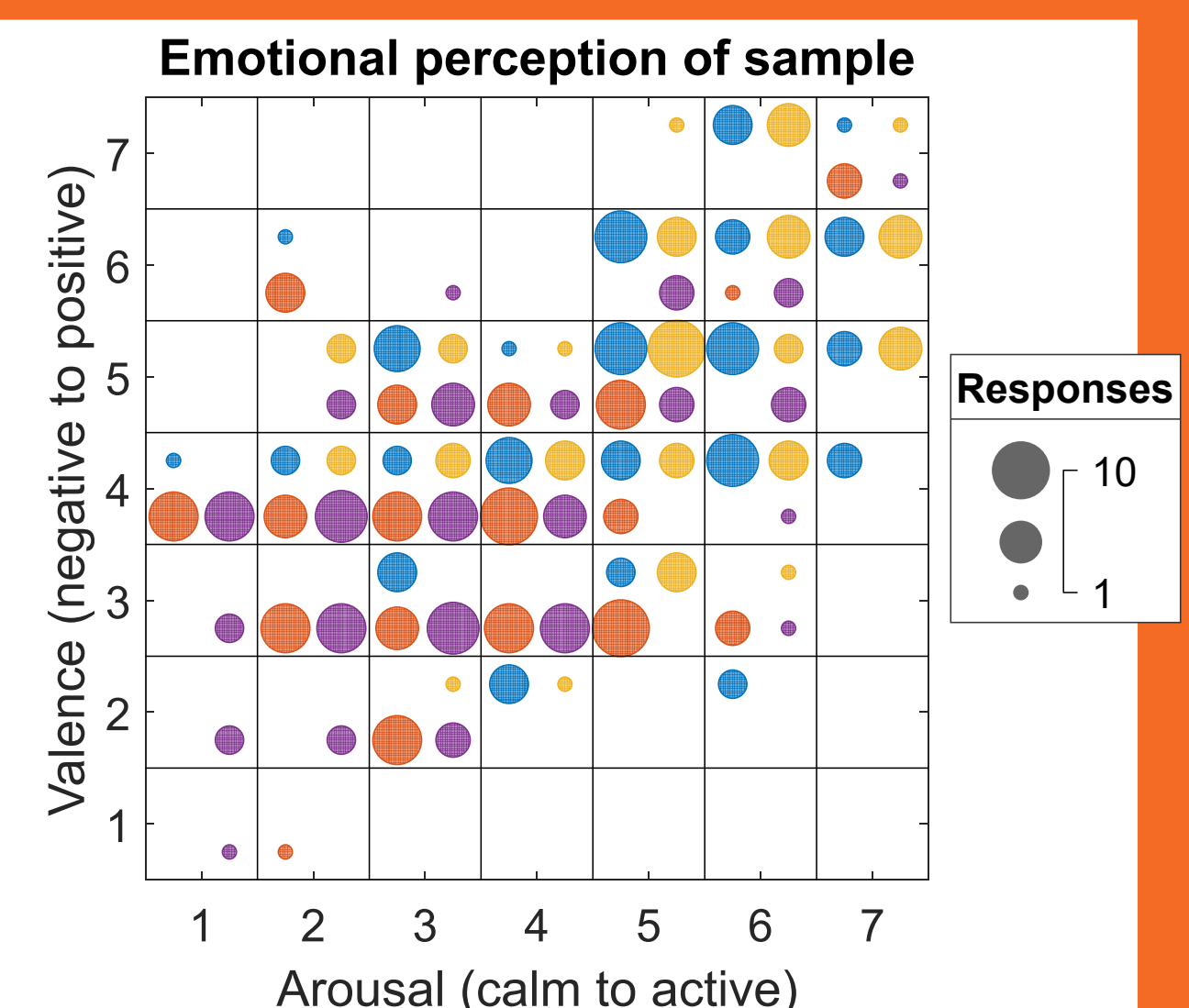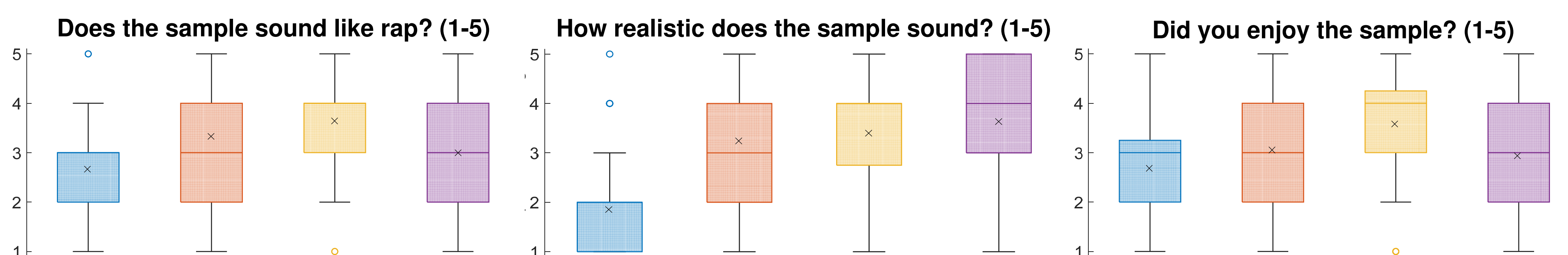 - Syllable length (with respect to rhythm choice)

1) Generate speech
2) Isolate syllables
3) Fix syllable length
4) Quantize to rhythm

User Lyrics → GENERATE RAP ← Beat (Music) → + → WiFi Transmission

Generated Video

PERFORMANCE MODULE

- **Robot Components:**
  - AC/DC Power Adapter: 40W 5V
  - Micro-Computer: Raspberry Pi 4B
  - LCD Screen: Pimoroni Ltd PIM470
  - Servo Motors: TowerPro MG996R
  - Motor Controller: Adafruit PCA9685
  - Speaker: Visaton SL 70 NVD – 4 OHM
  - Subwoofer: PUI Audio ASX05404-HD-R
  - Amplifiers: Adafruit MAX98306
  - 4.12" ABS Plastic Chassis

Subwoofer — Amplifier — Low Pass Filter — High Pass Filter — 4 Ohm Speaker — Amplifier — Raspberry Pi 4B — LCD Screen — Servo Motor (x2) — Motor Controller — 5V 40W Supply

Raspberry Pi · Motor Controller · Speaker · Screen · Servo Motors · Audio Amps & Crossover · Audio Connection · I2C Motor Connection · Subwoofer · 5V DC Input/Output

SURVEY RESULTS

- Text-to-speech + Speech-to-rap
- Natural speech + Speech-to-rap
- Text-to-speech
- Natural speech

Does the sample sound like rap? (1-5)

How realistic does the sample sound? (1-5)

Did you enjoy the sample? (1-5)

Emotional perception of sample — Valence (negative to positive) — Arousal (calm to active) — Responses 10 / 1

ETHICAL STATEMENT

- Our project is not an artist replacement
  - The software will be an open-source tool for artists to experiment with and enhance their work
  - Any audio generated by our system will be owned by the creator of the lyrics inputted
- We do not intend to create a caricature of rap culture or rap performers, which is why our robot's defining characteristics will not be human-like

CONCLUSION

- Application:
  - The software can be used by others in the music industry to test out ideas for new lyrics and tempos without excessive effort
  - The robot could be an interesting toy for kids
- Future Work:
  - Further training the Flowtron text-to-speech model on rap
  - Smoothing the syllables for better output
  - Increase the freedom of the user to choose a rhythm