Project Vision

E96 Final Presentation

Presented and Designed By:

Nolan Gunsolley

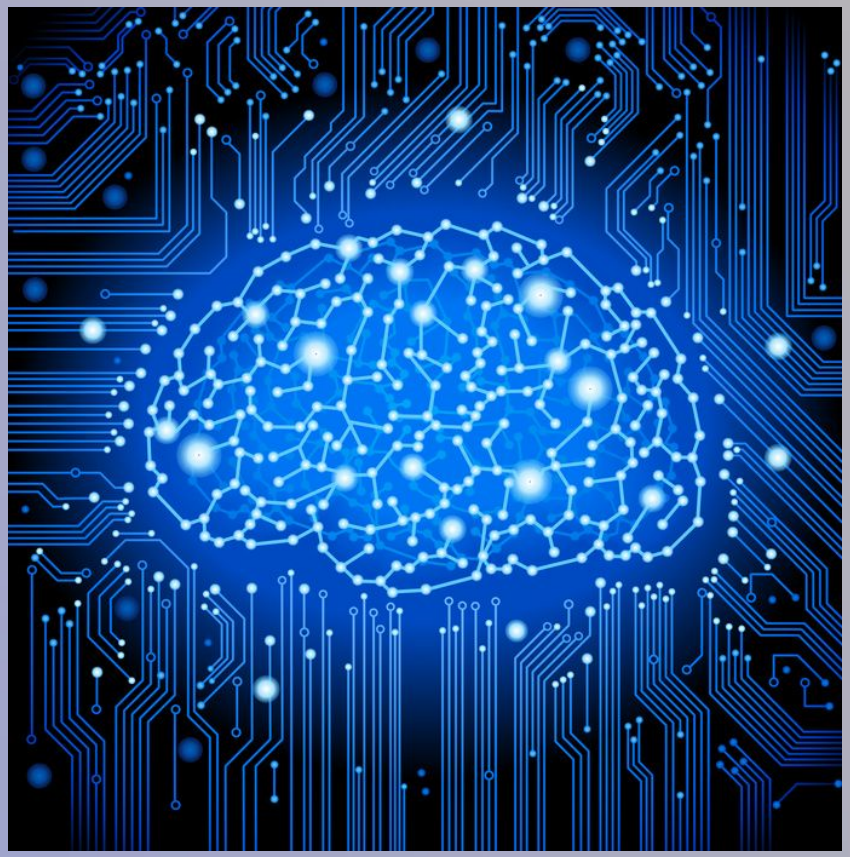
Jun Kai Ong

Kevin Tolby

Hongyi Zhang

Product Development Mission

- Utilizing image recognition technology for social causes.
- Prototype to showcase the possibilities of machine learning integrated into everyday lives.



Product Development Mission

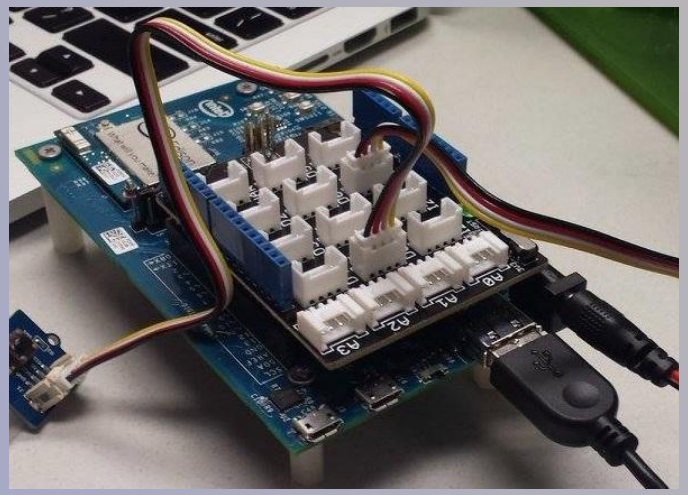
- Applications
  - Assistance for the blind/disabled.
  - Education aid for persons with developmental disabilities. (e.g. ASD)
  - Emotion detection - context sensitive robots.

Design Overview: Technical Challenges

- Finding a method of visualization that represents the human eye, both in object recognition and perception.
- Connecting to the internet for a near-infinite supply of objects for object comparison.
- Finding a form of communication between the user and the Edison.
- Translating received text files into playable sound bytes.

Design Overview: Technical Challenges

- Working with hardware and software compatible with IoT/Edison
- Interactions between different programming language scripts, GPIO, and bash scripts



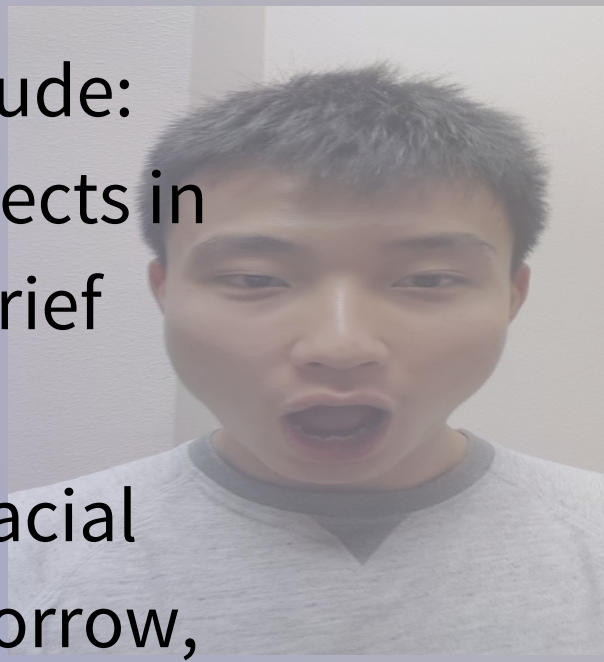
```
root@edison:~# opkg install espeak
Installing espeak (1.48.04-r0) on root.
Downloading http://repo.opkg.net/edison/repo/core2-32/espeak_1.48.04-r0_core2-32.ipk.
Installing libportaudio2 (v19+svnr1387-r0) on root.
Downloading http://repo.opkg.net/edison/repo/core2-32/libportaudio2_v19+svnr1387-r0_core
2-32.ipk.
Installing libjack (0.121.0-r0) on root.
Downloading http://repo.opkg.net/edison/repo/core2-32/libjack_0.121.0-r0_core2-32.ipk.
```

Design Overview: Technical Approach

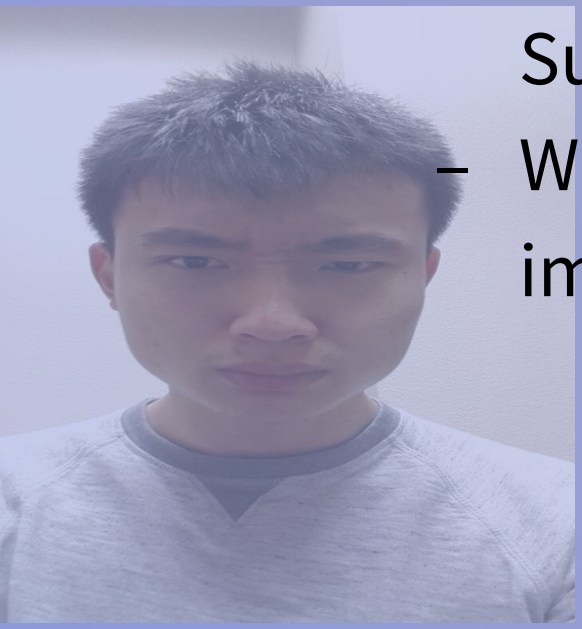
- The project is centered around an API (Application Programming Interface) developed by Google called the Google Vision API.

- The API receives an image from the user and returns a JSON file with a detailed description of the image.

Google Cloud Platform



- The image descriptions include:
  - The categories of the objects in the image along with a brief description.
  - For a human, his or her facial expression (e.g. Anger, Sorrow, Surprise, Joy)
  - Whether the person in the image is wearing a hat



Design Overview: Technical Approach

- A button is connected to the Intel Edison as a trigger.
- When the button is pressed, the Intel Edison runs a script that snaps a photo using the USB camera attached and sends the photo to the Google API.
- The API then returns a JSON file.
- Once the JSON file is received, the JSON file is passed to another API called the IBM Watson Text-to-Speech API.
- A .wav file is returned.
- The .wav file is then played via a bluetooth speaker connected to the Intel Edison.

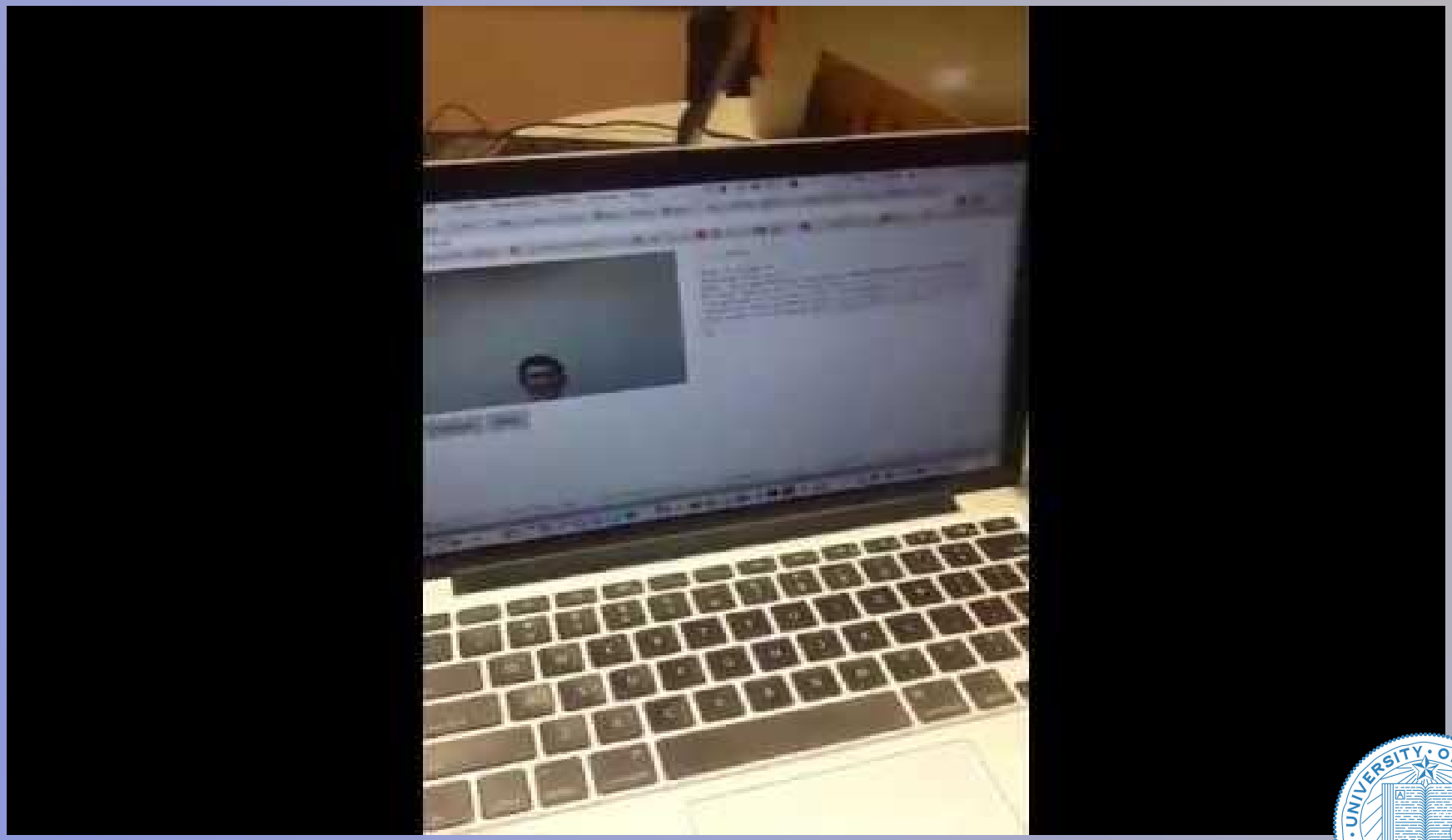
Live Demonstration

Snap Process Clear

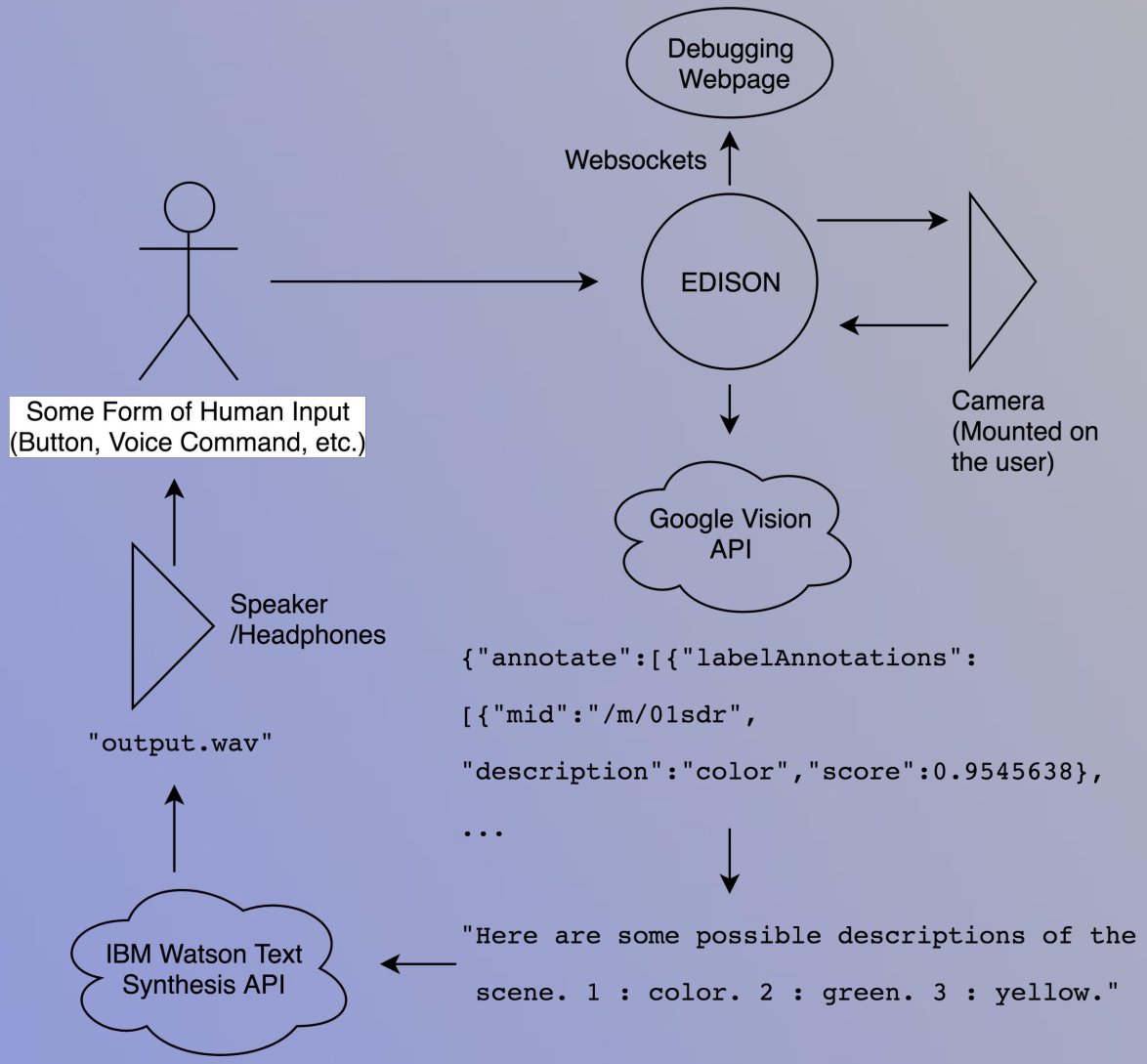
```
----- OUTPUT ------

Snap is triggered!
Received Image Data!
DATA: /9j//gAQTGF2YzU3LjI0LjEwMgD/2wBDAAgGBgcGBwgICAgICAkJCQoKCgkJCQkK
Received Annotations, synthesizing to sound...
{"annotate":[{"faceAnnotations":[{"boundingPoly":{"vertices":[{"x":681
Converting text to speech:There is one person in front of you. 1 of th
Done converting to sound, now writing to disk and playing...
Played speech!
```

Video Demonstration



IoT System Implementation



Team Responsibilities and Contributions

Hongyi Zhang:

- Vision API Script and implementation with hardware
- Creation of debugging website for use with Vision API
- JavaScript Programming

Jun Kai:

- IBM Watson Text-to-Speech API Script
- Bluetooth auto-connect feature development
- C program that plays audio output from IBM Watson API
- The photos

Team Responsibilities and Contributions

Kevin Tolby:

- Iterative design on TTS
  - Method to feed output from Vision API to TTS API
- Integration of GPIO interrupts into code
- Development of startup-script to send Edison IP address info to cell phone

Nolan Gunsolley:

- Establishment of Bluetooth communication between Edison and external speaker
- Integration of the early-stage espeak text-to-speech program into javascript code format
- Bluetooth and text-to-speech debugging

Future IoT Missions

- A **wearable system** with auto-distance detection using **ultrasound sensors** to allow vision-impaired users to identify nearby objects.
- Incorporation into educational curriculum to assist in object and emotional recognition for students with developmental disabilities.
- Artificially-intelligent robotics with object recognition capabilities.
- A more extensive interconnected library of images and emotional expressions.

Additional Optional Information

- Why we chose JavaScript
  - No memory management!
  - Rapid prototyping
  - Native language of the internet
    - Ease of communicating over HTTP / to APIs / Websockets

socket.io

Questions?

Code available at
https://github.com/hongee/engr96-vision-proj