IN4MATX 133: User Interface Software

Lecture:
Alternative interaction

Today's goals
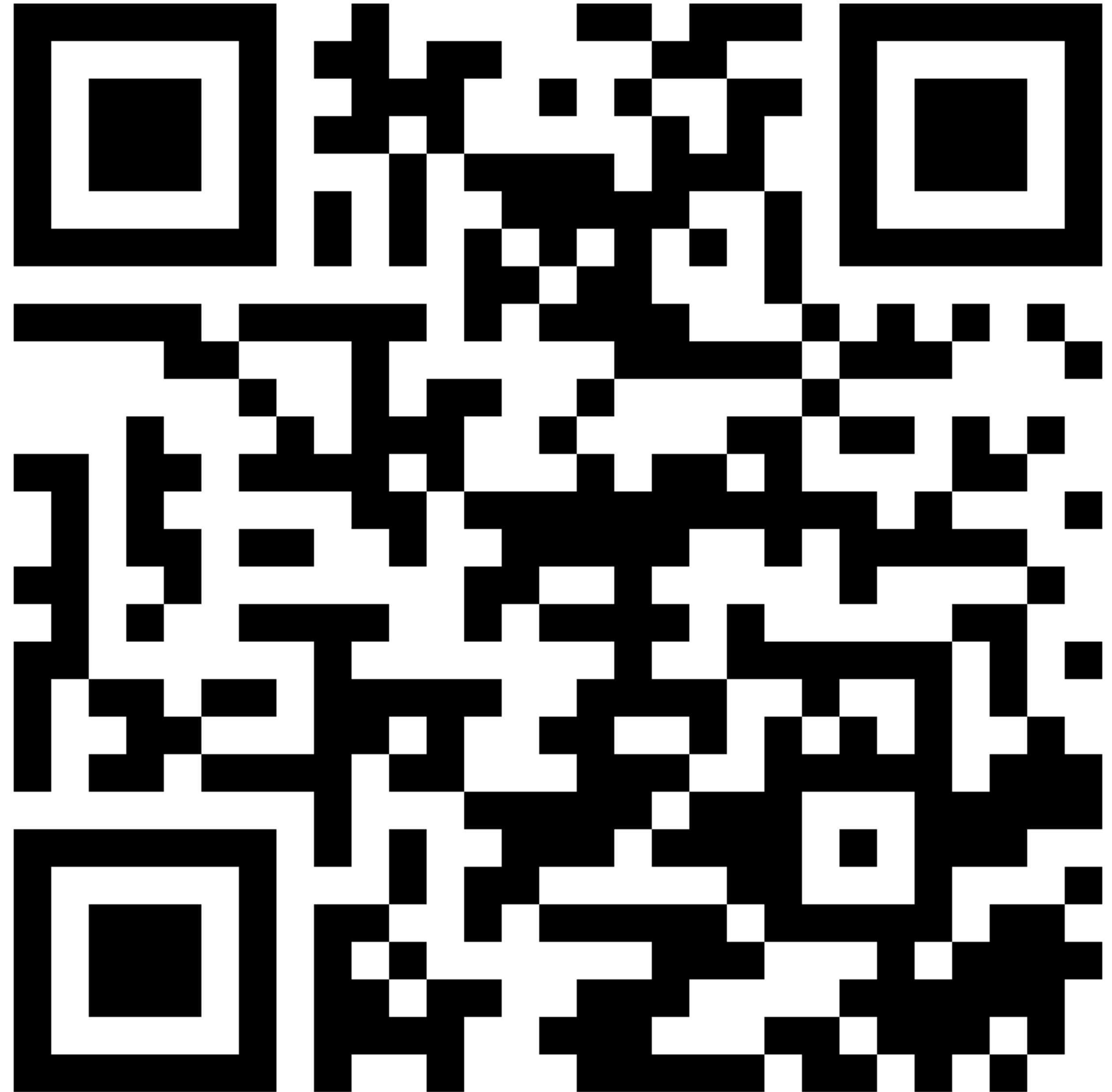
By the end of today, you should be able to…

- Describe how machine learning works

- Describe how HandTrack.js and other alternative interactions make use machine learning

- Understand multimodal UI's and how they can lead to a better user experience

# Socrative Quiz!

Enter your UCI Email when prompted for name!!!
e.g.,

xxxxx@uci.edu

https://api.socrative.com/rc/52QwBu

Machine Learning

- The application of an [algorithm] on a [set of data] to make [predictions for computational actions].

- Algorithms

  - Supervised, unsupervised, and reinforcement learning algorithms

- Set of data

  - Sample data, or training data, is assembled into a model.

  - Model is input to algorithm

- Predictions for computational actions

  - Based on input model, algorithm predicts what the next action should be

  - E.g., avoid person, recognize object

https://www.ironhack.com/en/data-analytics/what-is-machine-learning
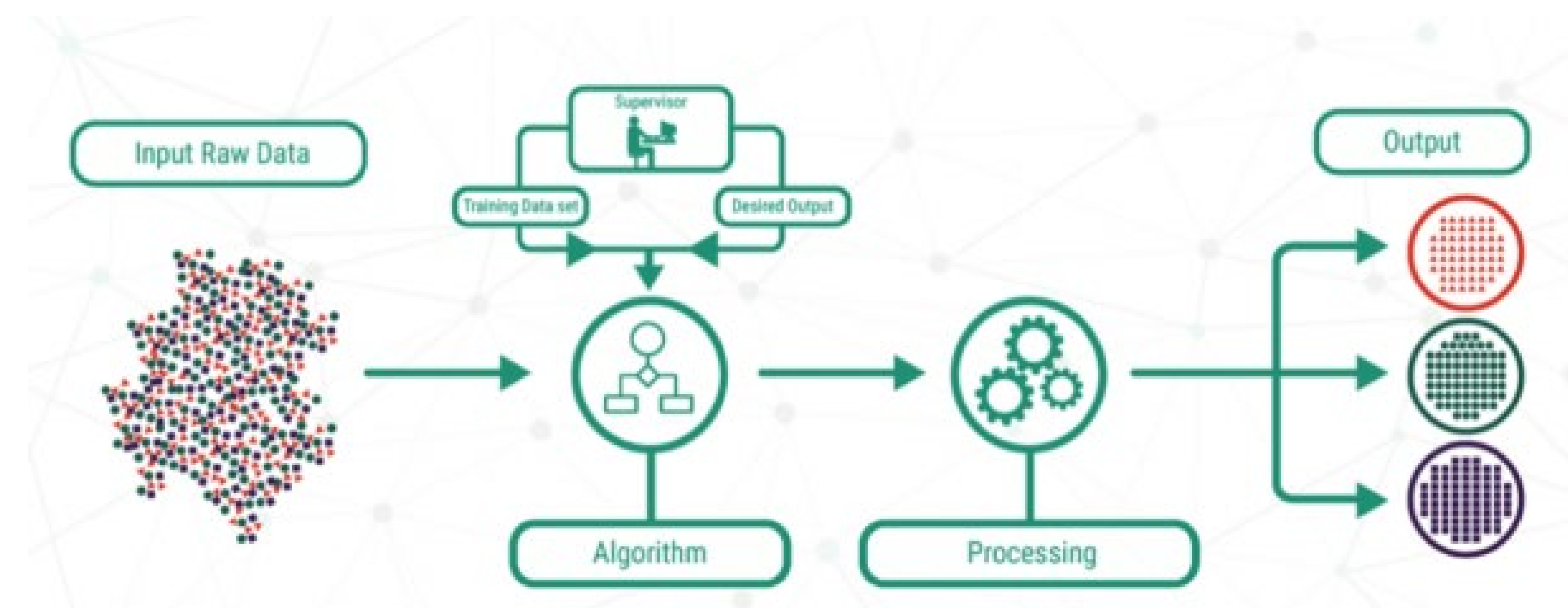
Machine Learning - Algorithms

- Supervised

  - Algorithms that learn from models with desired input and output

  - Ex. A data set of animals.

- Unsupervised

  - Algorithms that learn from models with desire input only

  - Output is predicted

  - Ex. Other movies you might like to watch

- Reinforcement

  - Algorithms that learn do not assume knowledge from a model, but rather predict output based on dynamic input.

  - Ex. Game A.I., self-driving vehicles

Machine Learning - Models

- Models are built from existing data

  - Dataset for model must be large (larger the better)

  - Data must be annotated (e.g., giraffe labeled as giraffe)

- Many datasets are kept private (costly to produce), but there are plenty also available for free

  - UCI provides many! (https://archive.ics.uci.edu/ml/datasets.php)

  - Also, Kaggle (https://www.kaggle.com/)

  - Likely others

https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research

Machine Learning - Tools

- Many exist, but TensorFlow has become the primary entry point

- A platform for building and deploying machine learning

  - Can build new models

  - Apply existing models

  - Apply existing algorithms

  - Support for numerous languages, but Python is most used

- TensorFlow.js adds ML capability to the browser

HandTrack.js

- Handtrack.js

  - Built on TensorFlow.js

  - Our version uses a model derived from the EgoHands dataset

    - Contains 15,000 examples of hands

    - Assembled from videos captured with Google Glass

http://vision.soic.indiana.edu/projects/egohands/

HandTrack.js

- Model is built with TensorFlow Object Detection API

- Trained model is then converted to a TensorFlow.js webmodel using TensorFlow.js tooling

- The handtrack.js code then loads the webmodel to make predictions based on inputs collected (static image or video)

https://towardsdatascience.com/handtrackjs-677c29c1d585
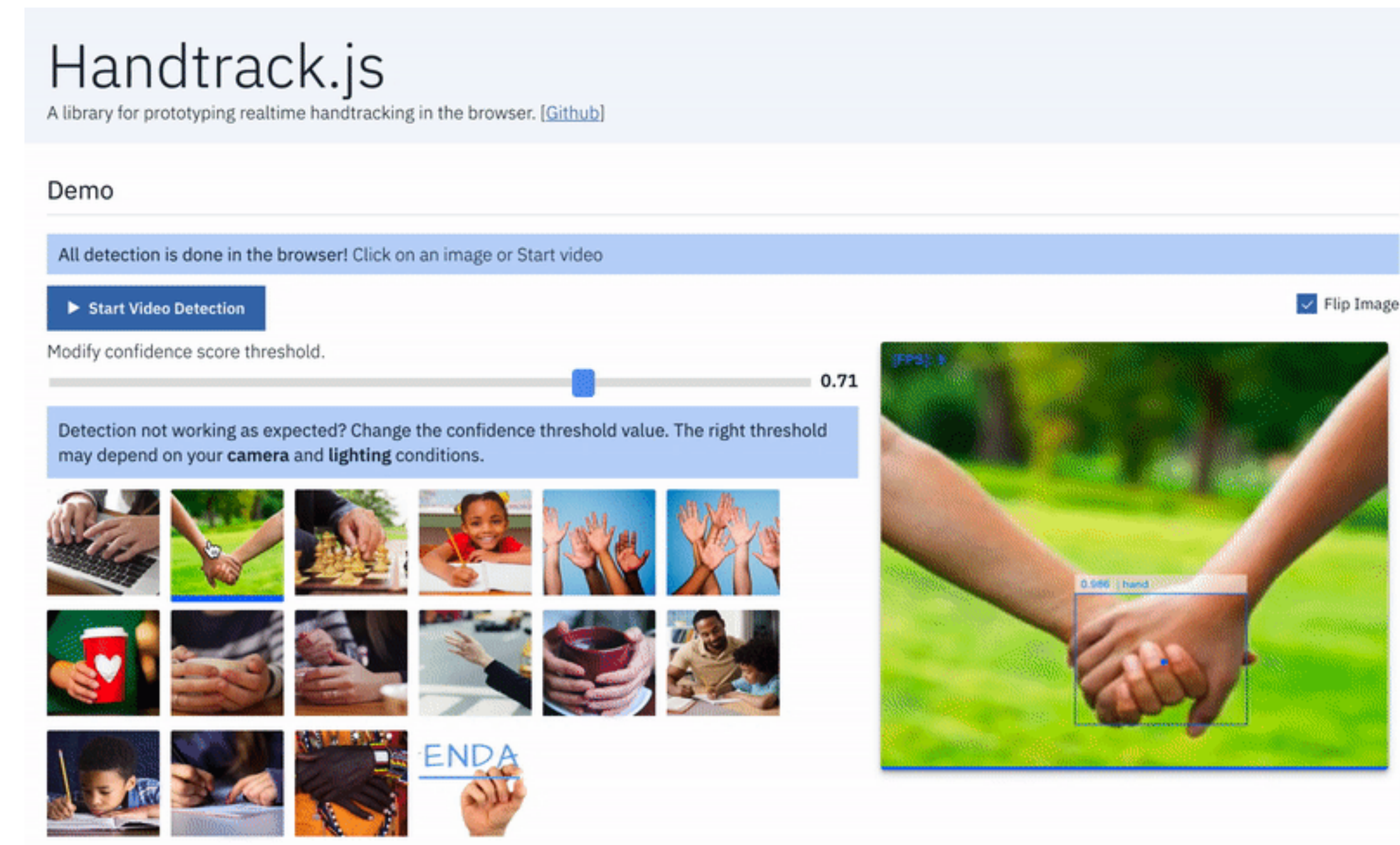
HandTrack.js

- Basic functionality of the library

  - Handtrack.js is a library that provides core functions and helper functions to process images

  - Load(params): accepts a params object and performs the initialization of the library. Returns a model object, required to make a prediction

  - Model object provides a "detect" function that will take an image as input and predict hand pose against the pre-built model

https://towardsdatascience.com/handtrackjs-677c29c1d585

HandTrack.js

- Basic functionality of the library

  - Params:

    - imageScaleFactor – Set to reduce the size of your image, thus improving prediction time (possibly reducing accuracy). Default is 1:1.

    - maxNumBoxes – The number of boxes to detect. If you only need one 1 hand, then set it to 2 (1 for face)

    - iouThreshold – "intersection of union," how sensitive do you want your overlapping objects to be?

    - scoreThreshold – The higher the score, the lower the amount of detections you will get. Find a balance!

```
private modelParams = {
    flipHorizontal: true,
    imageScaleFactor: 0.7,
    maxNumBoxes: 20,
    iouThreshold: 0.5,
    scoreThreshold: 0.6,
};
```

https://towardsdatascience.com/handtrackjs-677c29c1d585

HandTrack.js

- Basic functionality of the library

  - Prediction results:

    - Bounding box (bbox): The position and size of the detected object

    - Class: The predicted object (e.g., hand, face)

    - Label: The predicted gesture (e.g., open, closed, point, pinch)

    - Score: The confidence of the prediction. Higher score == greater likelihood of accurate prediction.

https://towardsdatascience.com/handtrackjs-677c29c1d585

[{
  bbox: [x, y, width, height],
  class: "hand",
  label: "open",
  score: 0.8380282521247864
}, {
  bbox: [x, y, width, height],
  class: "hand",
  label: "closed",
  score: 0.74644153267145157

}]

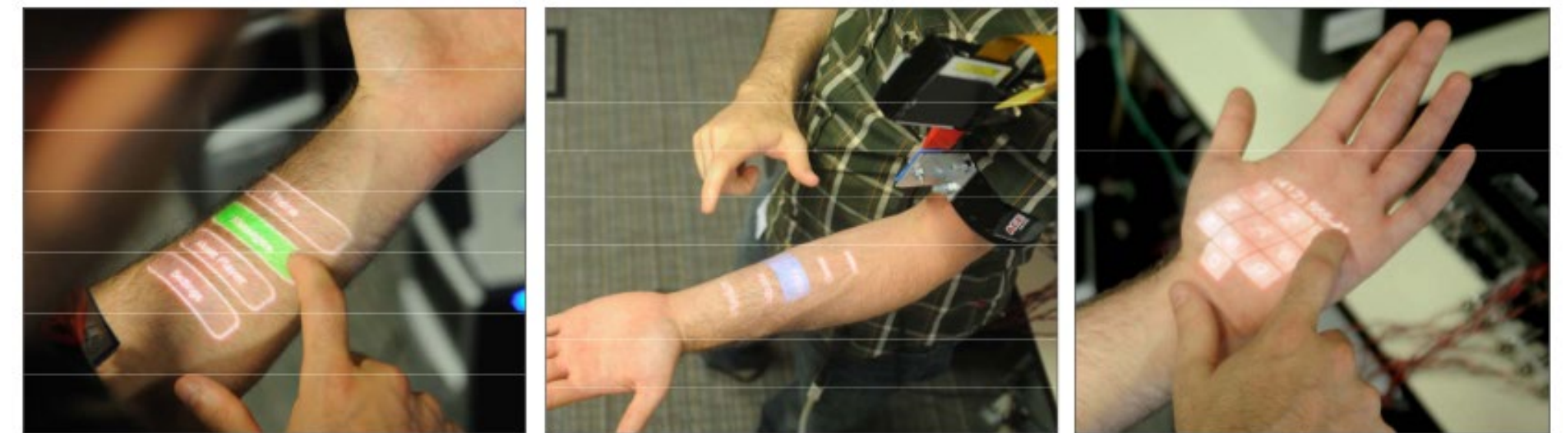# Machine Learning in UI

- How does all this apply to our work with user interfaces?

  - The ability to predict opens new directions for UI's

    - Gestures (hand and full body)

    - On-body interaction

    - Eye-gaze

    - Speech

Machine Learning in UI

- On-body Interaction

  - Armband detects muscle deflection

  - Deflections can be modeled to predict on body location

  - Projected overlay coordinates with predictions to identify input and perform output

Figure 5. Prototype armband.

Harrison, C., Tan, D., & Morris, D. (2010, April). Skinput: appropriating the body as an input surface. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 453-462).

Machine Learning in UI

- Eye Gaze

  - Standard web camera, tracks eye movement

  - Eye movement combined with cursor position, predicts location of gaze

  - Slightly lower resolution, but able to reproduce effects of expensive commercial eye tracking devices

Papoutsaki, A., Laskey, J., & Huang, J. (2017, March). Searchgazer: Webcam eye tracking for remote studies of web search. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval* (pp. 17-26).

Speech

- Speech

  - It's complicated!

  - Reserved for organizations large enough to create very large datasets

  - Automatic speech recognition is hard

  - Text-to-speech generation is hard

  - Open source speech engines exist, but they far fall below the quality of services like Alexa and Siri

https://medium.com/react-native-institute/amazon-alexa-tutorial-build-your-own-skill-a83cf71f8aed

Speech

- Common Voice

  - The Mozilla Foundation is working on an open dataset

  - Ask US to help build the dataset by recording preset speech passages

  - But…datasets are released back to public!

https://commonvoice.mozilla.org/en

Multiple Modalities

- What's wrong with just the keyboard and mouse or a touchscreen?

  - In many cases these three input devices are just fine

  - Accessibility ensures that software is designed with a "common language" that all systems can interpret

Multiple Modalities

- What's wrong with just the keyboard and mouse or a touchscreen?

  - However…

  - Accessible systems do not ensure efficiency, ergonomics, or enjoyment

  - We can do better!

Multiple Modalities

- What's wrong with just the keyboard and mouse or a touchscreen?

  - Offset non-critical tasks with alternatives in UI's

  - Examples?

Multiple Modalities

- What's wrong with just the keyboard and mouse or a touchscreen?

  - Offset non-critical tasks with alternatives in UI's

  - Examples?

    - Gestures or speech to control – freeing hands to continue work or other tasks

    - Text-to-speech to read the screen

    - Eye-gaze magnification

    - Modal input to increase functionality of existing interactions
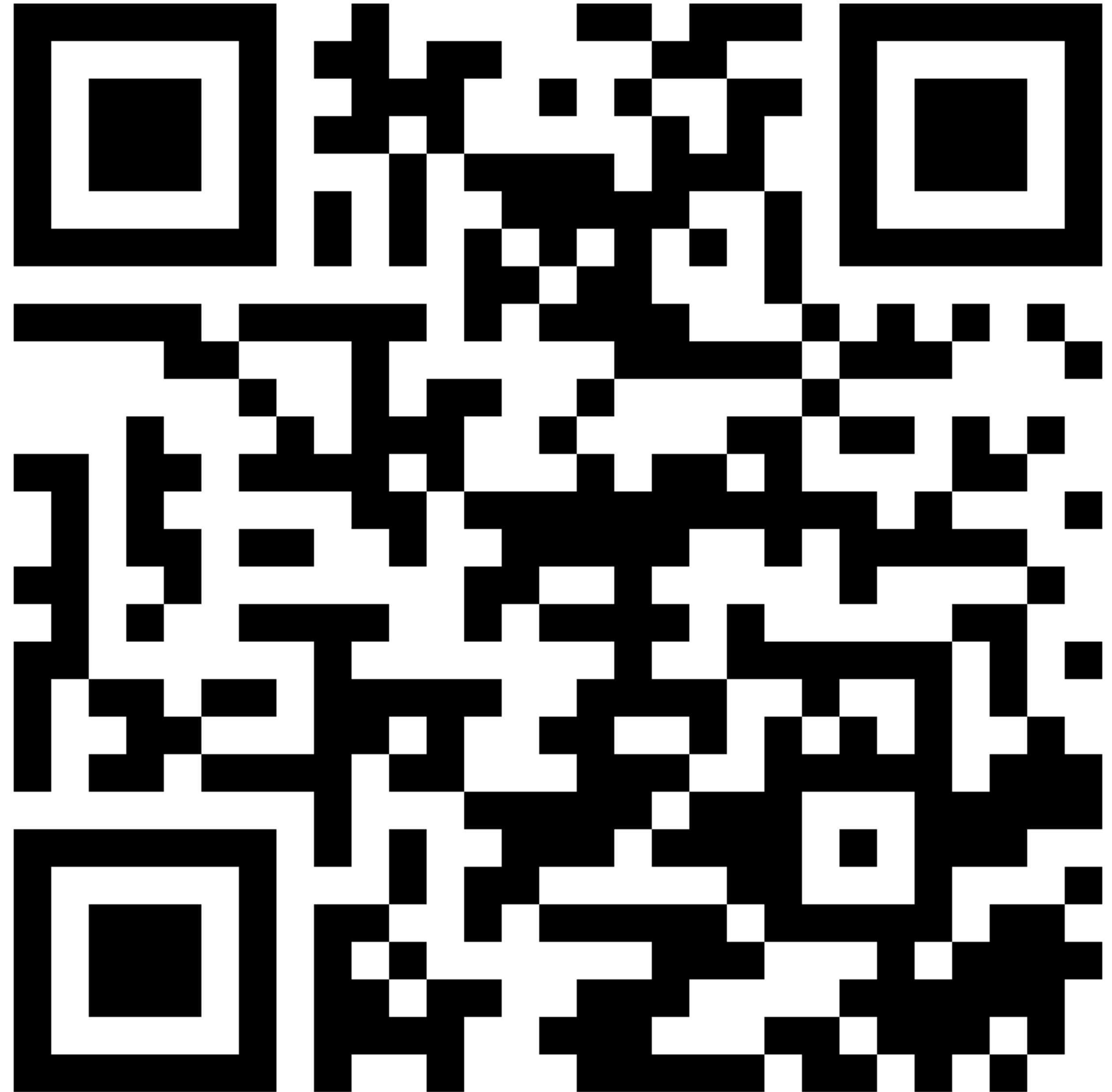
Multiple Modalities

- Bridging alternative interactions with touch, keyboard, and mouse, we broaden the ways in which we can control user interfaces

- Alternative interactions can increase usability for people with physical disability, both temporary or permanent

Socrative Quiz!

Enter your UCI Email when prompted for name!!!
e.g.,

xxxxx@uci.edu

https://api.socrative.com/rc/52QwBu

Multiple Modalities

A5 Demos?

Today's goals

By the end of today, you should be able to…

- Describe how machine learning works

- Describe how HandTrack.js and other alternative interactions make use machine learning

- Understand multimodal UI's and how they can lead to a better user experience