

# Modular Framework for Data Acquisition and Annotation (MFDAA)

## Authors:

- Pedro Durval n° 103173
- João Afonso n° 103037
- Eduardo Fernandes n° 102648
- Guilherme Claro n° 98432
- Tiago Mostardinha n° 103944

## Supervisors:

- Samuel Silva
- Fábio Barros

# Agenda

- Context and Challenges
- Related Work
- Goals
- Methodology
- Requirements
- Concept and Architecture
- Iterative development
- Live Demo
- Critical Discussion

# Context

- Technology has been an **exponentially growing** area for well over a decade, creating a increasing pool of **stakeholders** that look to integrate **user interactive features** into their systems.
- Whether it is postures, gestures or facial cues, the **complexity** of the methods used to recognize these interactions has **been increasing**.





# Context

- **Machine Learning** is a tool widely used across systems due to its versatility of applications.
- This versatility comes from the ability to use data from a certain phenomenon to model it.
- As such, it plays an important role in developing novel interaction methods, e.g., to **support the recognition of non-verbal cues**.

# Challenges

- Lack of readily available datasets that fit customers needs.
- Creation of datasets is a **resource and time consuming** task as it includes:
  - Systematic data acquisition;
  - Annotation / Labelling.



# Related Work

## GOOGLE ML KIT POSE DETECTION API

Is a lightweight versatile solution for app developers to detect the pose of a subject's body in real time from a continuous video or static image (Movement Recognition using Machine Learning)

## HAGRID

One of the largest data sets for HGR (Hand Gesture Recognition) systems, this dataset contains 552,992 full HD RGB images, divided into 18 classes of gestures

## QUIC SENSE

Git hub repository that is an inference engine to serve powerful neural networks for action recognition, with a low computational footprint.

# Main Goal

Create a modular platform that allows users to acquire,  
annotate and visualize data as well as their features

# Methodology



- **User-Centered Development**
  - Designed to **fit the view of the stakeholders** and improved through time;
  - Constructed with the **established goals** in mind.

- **Weekly Sprints**
  - **Iterative** methodology;
  - The **weekly discussions** helped the team stay on pace and determine what the next steps should be;
  - The **backlog** was also always updated with new tasks from each sprint (**JIRA SOFTWARE**).

# Target Users & Scenarios

## Researchers

- *Motivation*

Looking to create a **collection** of videos with **specific labels** to use in a study.

- *Solution*

Use MFDAA's **automative capabilities** to record several videos with corresponding annotations in an **efficient** manner.

## Developers

- *Motivation*

**Find a dataset** fitted to train an AI in the recognition of a specific **non-verbal cue**.

- *Solution*

With our framework users can **browse** various public projects to find one that fits their needs. In addition, these can also be **imported** and **expanded** by third-parties.

# Functional Requirements

---

- Data Acquisition: The platform should be able to capture video data inside the framework and import from various sources.
- Annotation and Labeling: should provide to users an intuitive approach for labeling and annotating videos.
- Allow collaboration between users by allowing projects to be public and expandable with export and import features.
- Feature Extraction platform must include tools or integration for extracting relevant features from the videos (ML Integration).



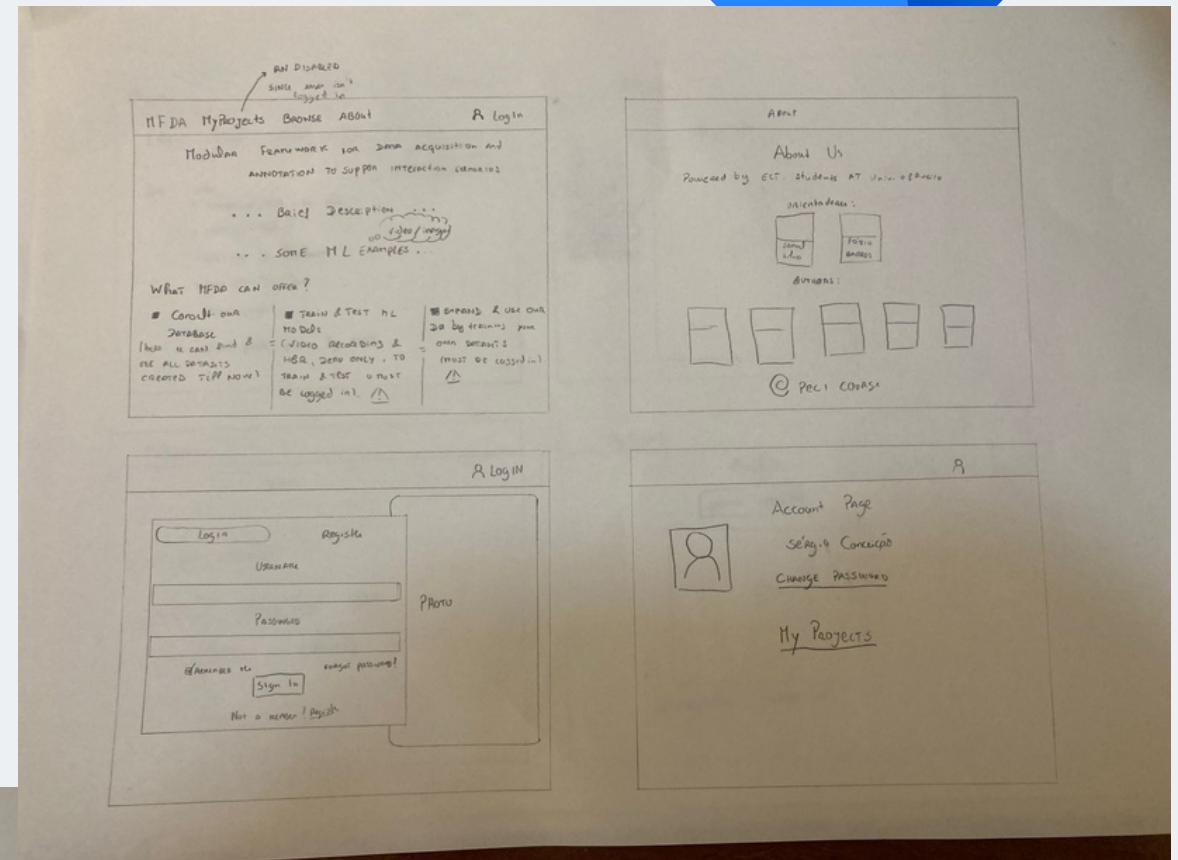
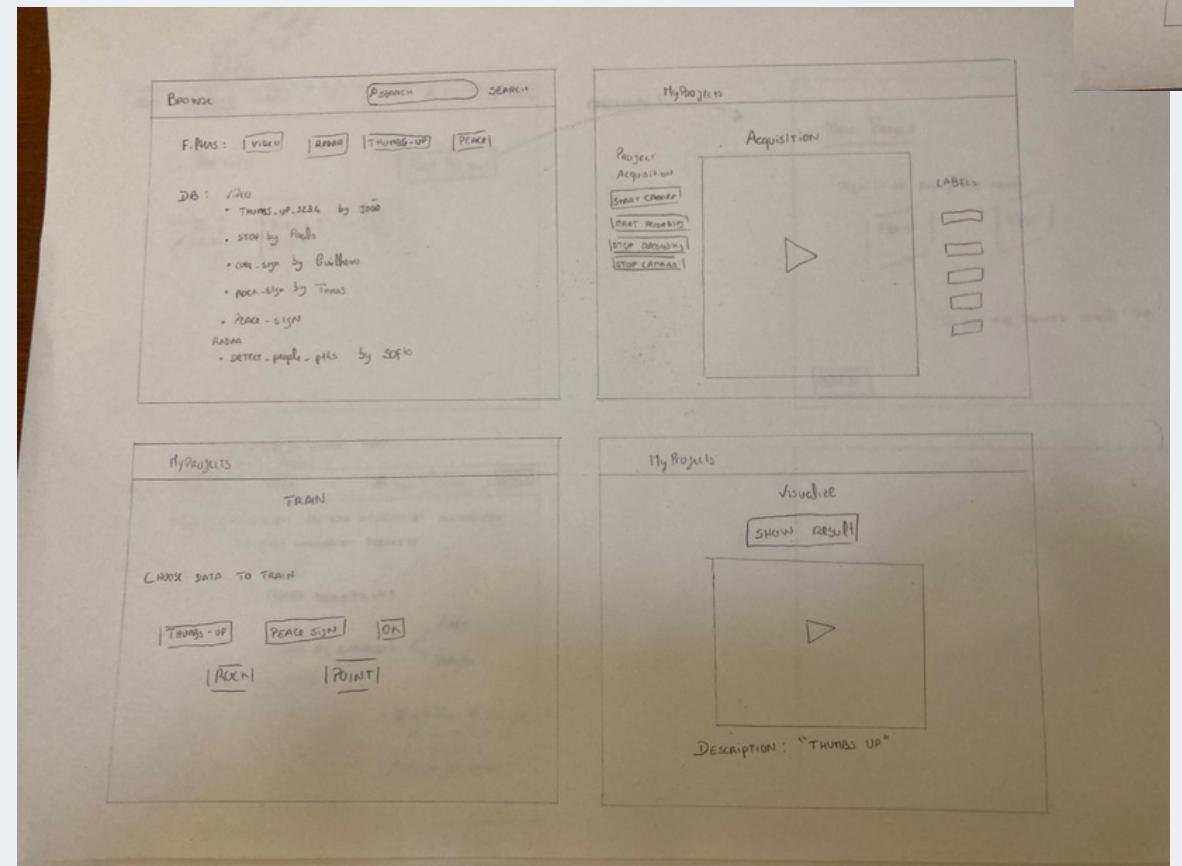
# Non-functional Requirements

---

- Performance and Scalability, framework should be efficient and scalable to handle large video datasets.
- User-Friendly Interface.
- Reliability, platform should perform accurately and consistently without errors or failures.
- Accessibility and Compatibility with different browsers and systems

# Prototype

- In the early stages, we developed a **Low-Fidelity prototype**
- User **Test Prototype**



# Concept and Architecture

- **Concept**

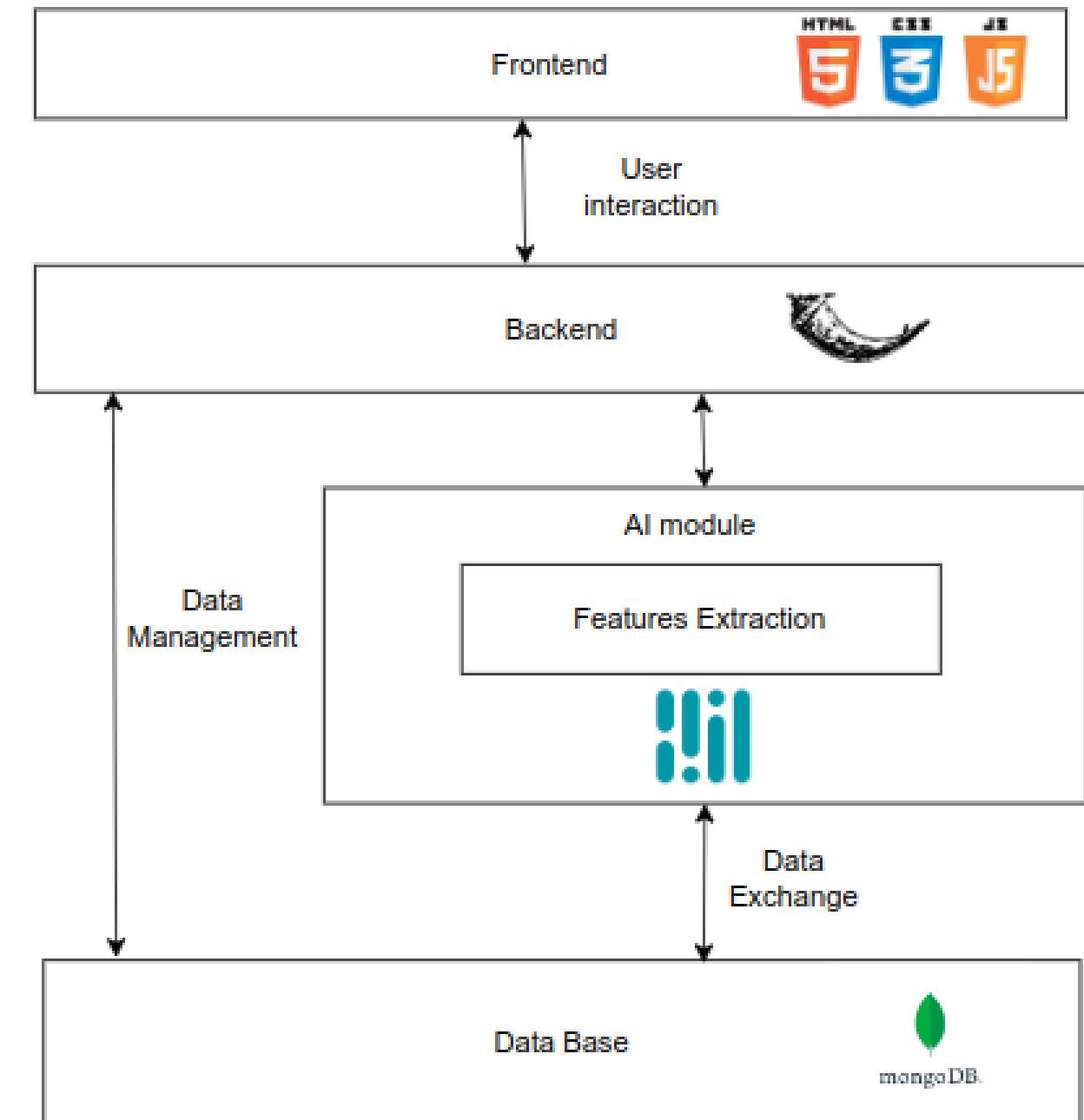
- While developing, we faced **problems**, and we needed the **right tools** to solve them.

- **Front-end Technologies:**

- **HTML**
- **CSS**
- **JavaScript**

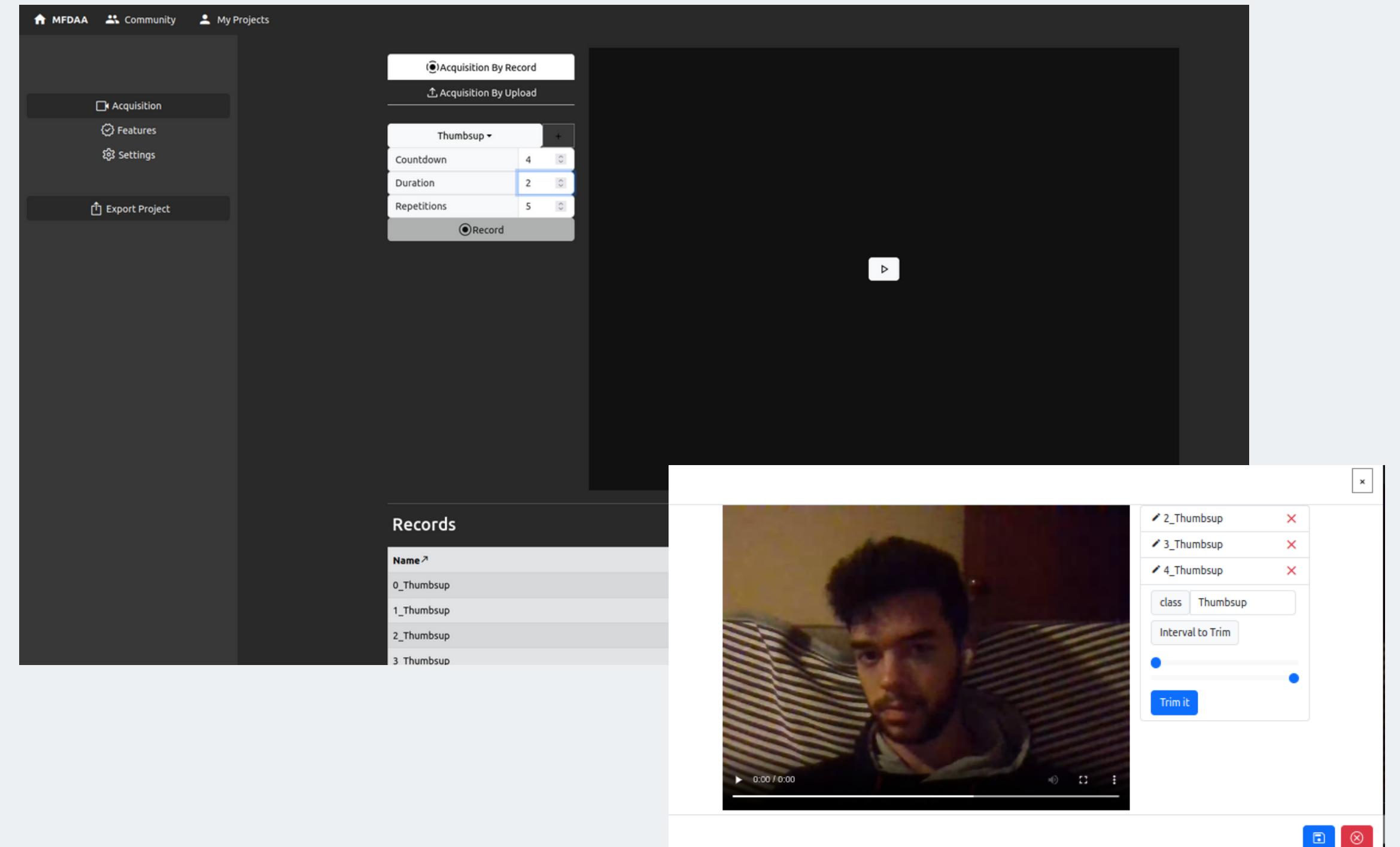
- **Back-end Technologies:**

- **Python**
- **Flask**
- **MongoDB**
- **MediaPipe** from TensorFlow framework



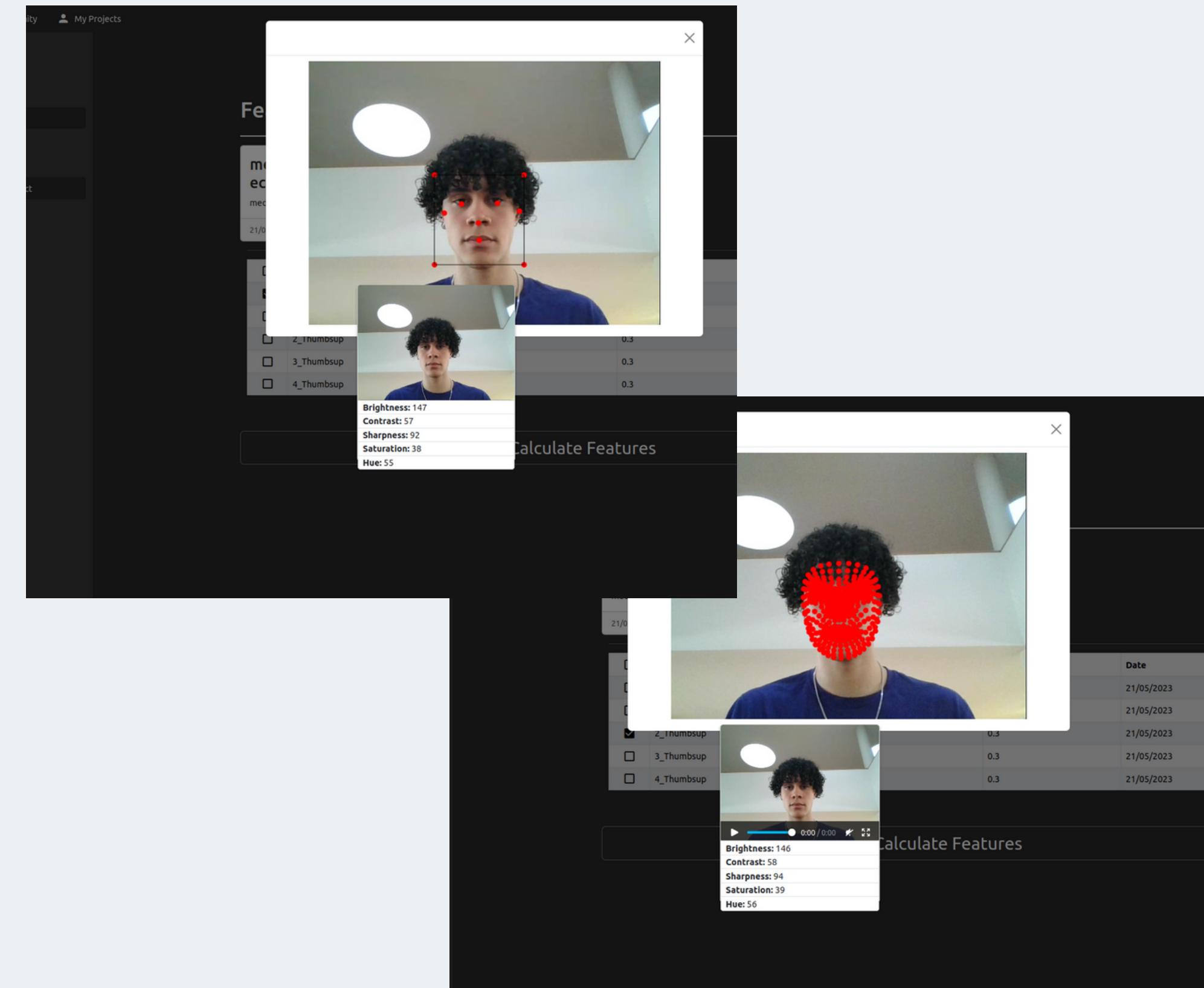
# Development - Acquisition

- Variety of options
- Correct mistakes easily
- Automated Acquisition
- Data Visualization



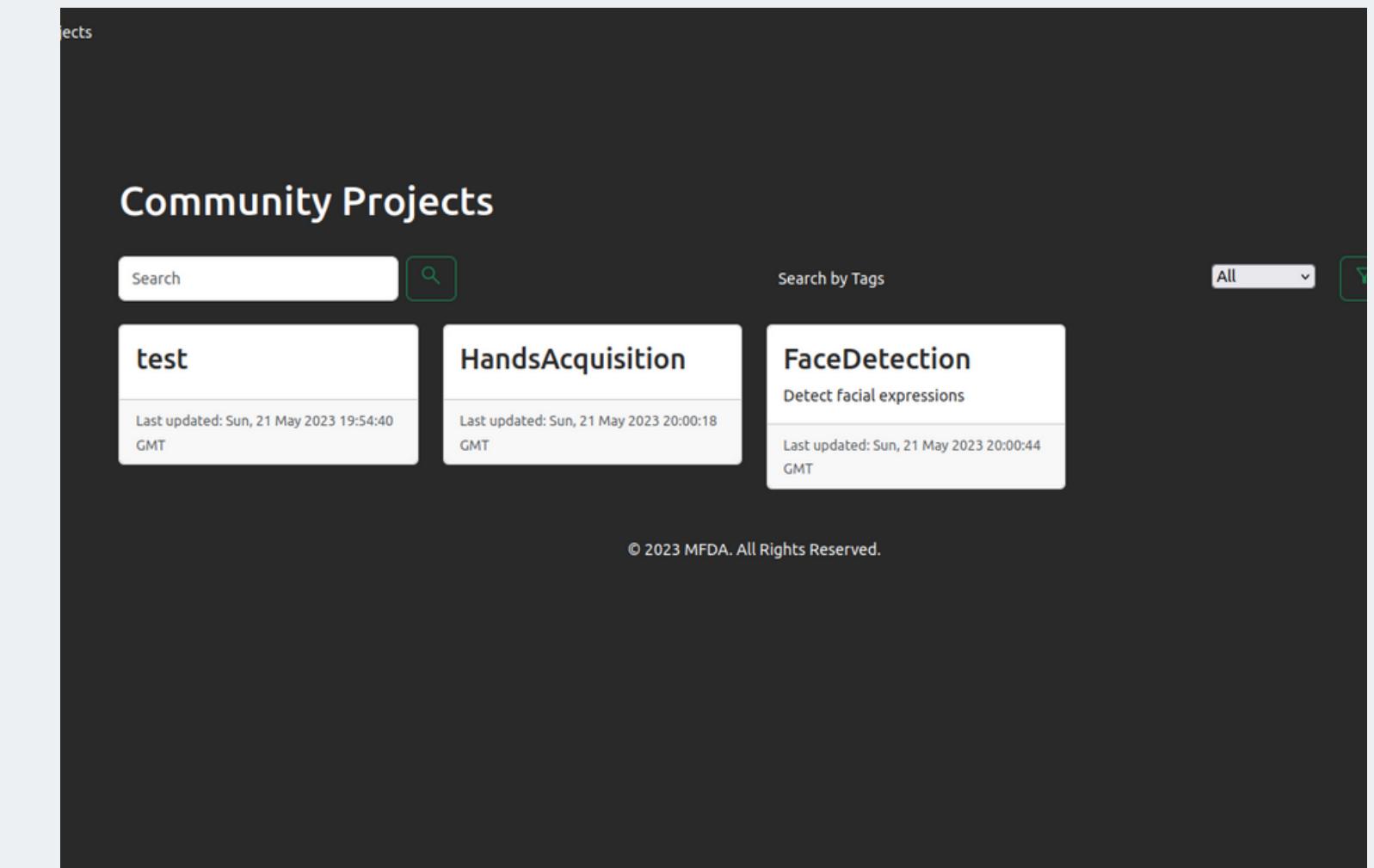
# Development - Feature Extraction

- Process Acquisitions
- Video Processing
- Important Points Extraction
  - MEDIAPIPE



# Development - Community Interaction

- Big DataBase of Public Projects
- Community working together
- Search Community projects by keywords and filters



# Development - Export Project

- Export Project in a readable way
- Use the data gathered to train AI Models



The screenshot shows a dark-themed user interface for managing a project. At the top, there is a large, empty text input field. Below it, the interface is organized into four main sections:

- Description**: A field containing a small icon and the text "Description".
- Category**: A field containing a small icon and the text "Video" with a dropdown arrow.
- Features**: A field containing a small icon and the text "No feature calculated for now".
- Characteristics**: A field containing a small icon and the text "Contrast: 58 Brightness: 147 Sharpness: 95 Saturation: 39 Hue: 56".

At the bottom of the interface, there are three buttons: "Extend Project" (white background), "Download Project" (blue background with a download icon), and "Delete Project" (red background with a trash icon).



# Development Validation

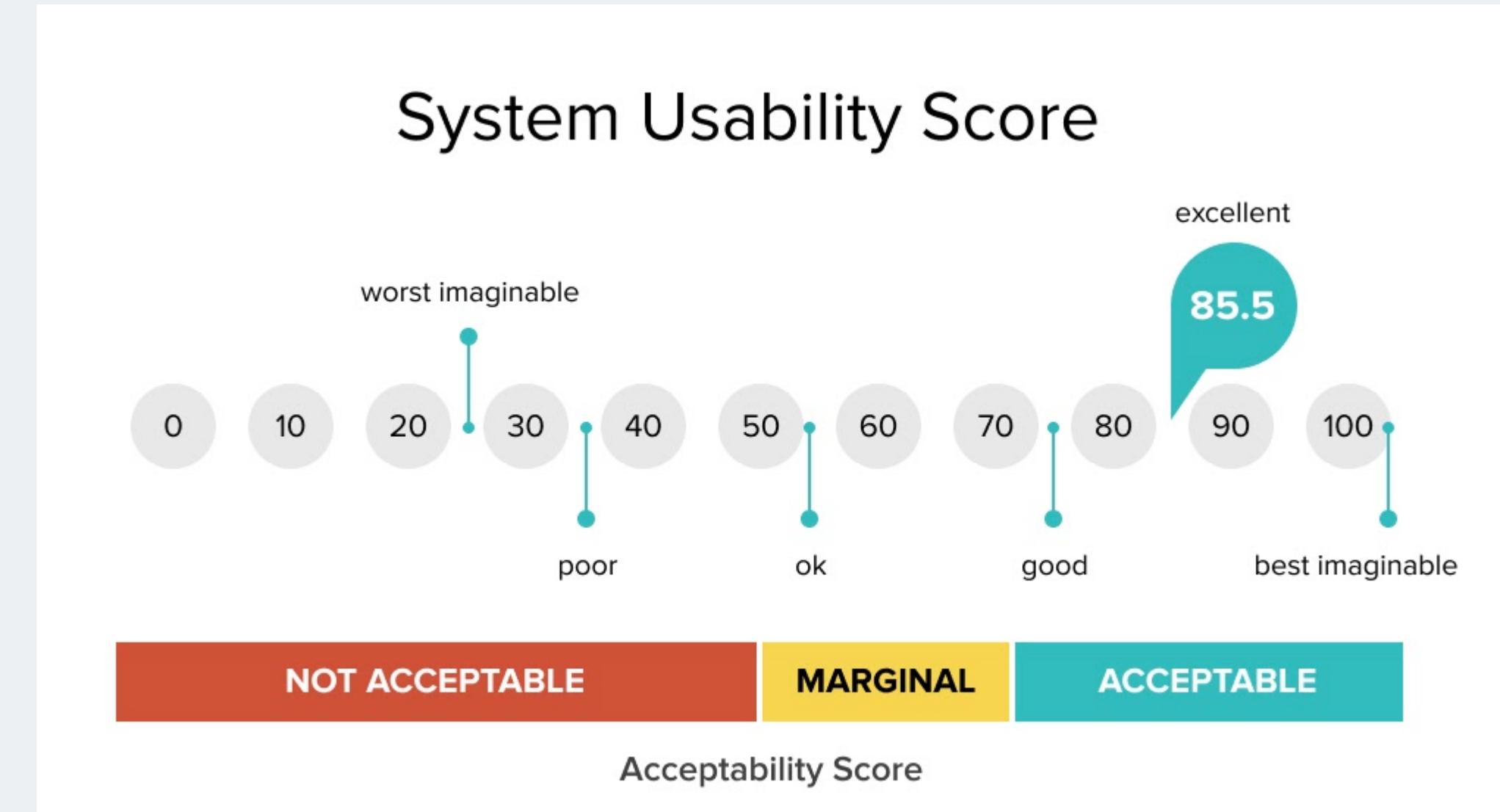
- **USER TESTS**

- **Tasks** were created for people to test out our interface;
- **Target users**
- The outcome was useful, **we were able to make the changes** needed in our interface to make it more intuitive.

	Tester 1	Tester 2	Tester 3	Tester 4
I think that I would like to use this system frequently.	Totally agree	Somewhat agree	Neutral	Somewhat agree
I found the system unnecessarily complex.	Somewhat disagree	Somewhat disagree	Somewhat disagree	Totally disagree
I thought the system was easy to use	Totally agree	Somewhat agree	Neutral	Neutral
I think that I would need the support of a technical person to be able to use this system	Somewhat disagree	Disagree	Disagree	Somewhat disagree
I found the various functions in this system were well integrated	Neutral	Somewhat agree	Somewhat agree	Somewhat agree
I thought there was too much inconsistency in this system	Disagree	Netral	Disagree	Disagree
I found the system very cumbersome to use	Somewhat disagree	Somewhat disagree	Somewhat disagree	Somewhat disagree
I would imagine that most people would learn to use this system very quickly	Agree	Agree	Somewhat Agree	Somewhat Agree
I felt very confident using the system	Agree	Agree	Neutral	Somewhat Agree
I needed to learn a lot of things before I could get going with this system	Disagree	Disagree	Disagree	Disagree

# System Usability Scale (SUS)

- **Industry standard** to validate interface
- Average **SUS score**: 68



The background features a large, abstract graphic composed of overlapping blue and white shapes. It includes a prominent white 'X' shape and several overlapping circles and triangles in varying shades of blue, creating a dynamic, layered effect.

# Live Demo

# Attained points

- Provide a **dynamic interface**;
- **Record and label** their content;
- Help users **visualize** their trained **data**;
- Give a wide **range of options** to **edit and manage** each project;
- **Automated acquisition and annotation** processes;
- **Accurate recognition** of face and hand gesture cues;
- **Import projects** to the platform;
- **Download their projects** from the platform.



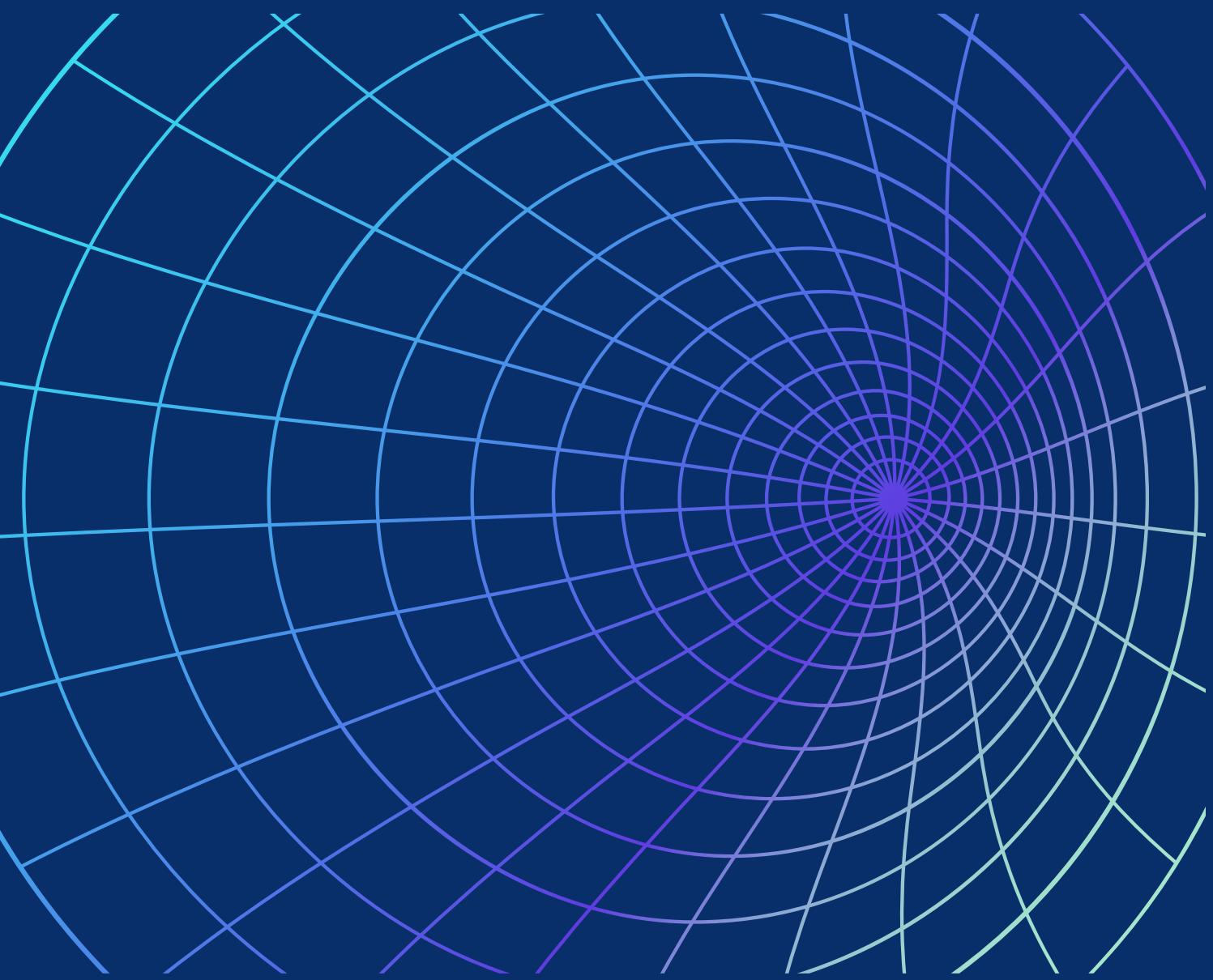
# Weaknesses

- Interface design;
- Lack of automation options;
- Versatility between annotation and acquisition processes.



# Future Work

- Allow for the processing of other types of data, such as: audio, radar, and more.
- Improve efficiency, accuracy and variety of our data analysis classes
- The ecosystem of our framework can be expanded to mobile to favor acquisition



# Any questions left?



MFDAA   Community   My Projects

Modular Framework for Data Acquisition and Annotation to Support Interaction Scenarios

Unleash insights hidden in Non-Verbal Communications!

Explore

A screenshot of a website titled "MFDAA". The top navigation bar includes links for "Community" and "My Projects". The main content area features the text "Modular Framework for Data Acquisition and Annotation to Support Interaction Scenarios" and "Unleash insights hidden in Non-Verbal Communications!". Below this is a button labeled "Explore" with a dropdown arrow icon. The background of the page is dark.