MICHIGAN STATE UNIVERSITY

Project Plan Presentation

Image Analysis Tool for Biphasic Solutions

The Capstone Experience

Team AbbVie

Hank Murdock
Colton Leslie
Noel Vazquez
Chirag Solanki
Alex Chirillo
Joesph Cook

Department of Computer Science and Engineering
Michigan State University

Fall 2024

From Students…
…to Professionals

Project Sponsor Overview

- Biopharmaceutical company
- North Chicago, IL
- Immunology, oncology, and neuroscience
- Humira, Imbruvica, Botox, Lexapro

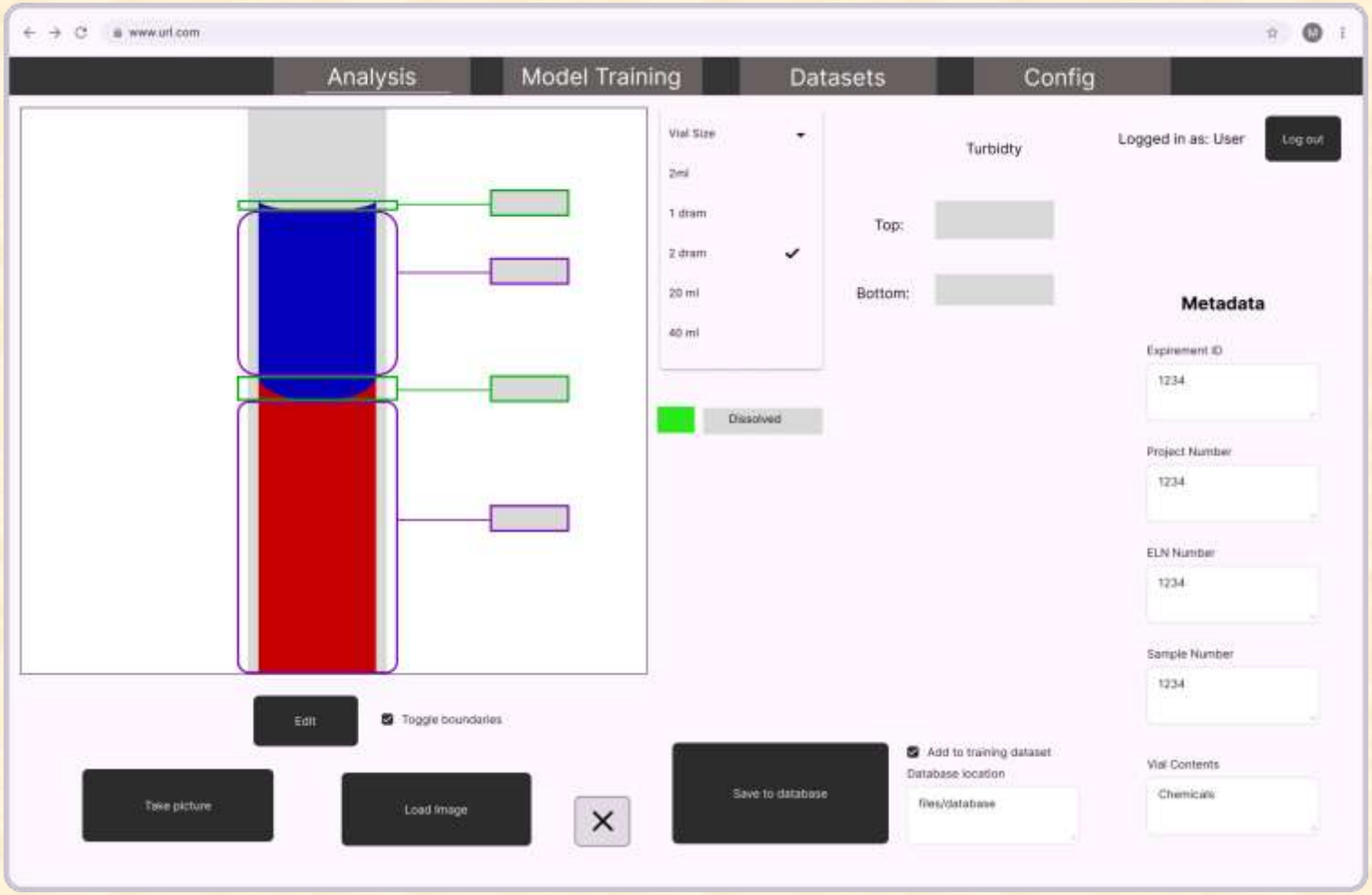
Project Functional Specifications

- AbbVie relies on manual, time-consuming methods to analyze biphasic solutions which introduces inefficiencies and limits scalability in high-throughput experiments.
- This project provides an automated image analysis tool, streamlining the process of measuring vial properties such as phase boundaries, turbidity, and emulsification, with a machine learning model that can be retrained as needed to maintain accuracy.

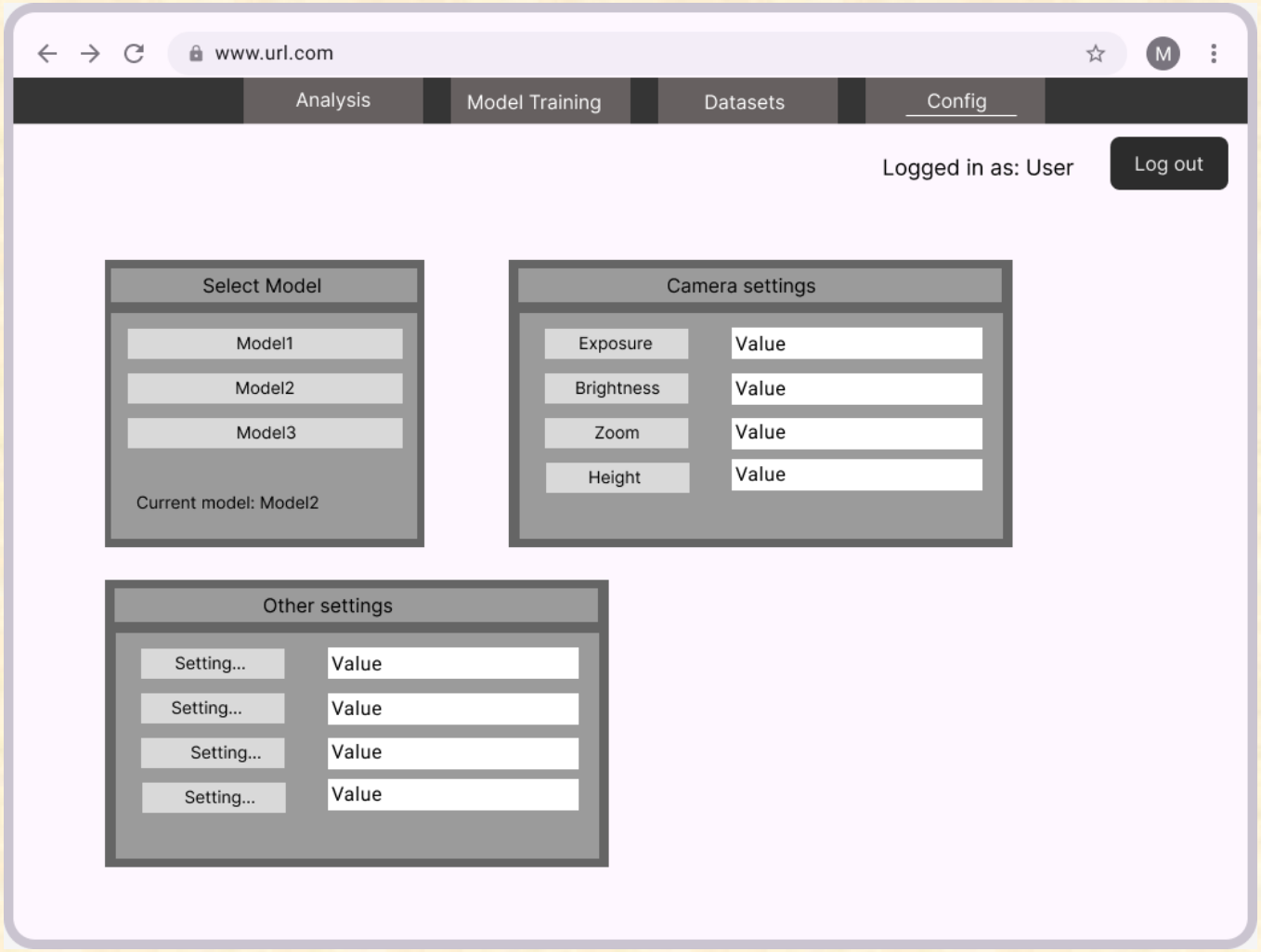
Project Design Specifications

- Users will access the Image Analysis Tool through their web browser
- Login with AbbVie username and password
- Analysis page provides lab results for biphasic solutions and stores them in a database
- Ability for users to retrain models when they deem fit
- Dataset Viewer page allows for easy manipulation of datasets

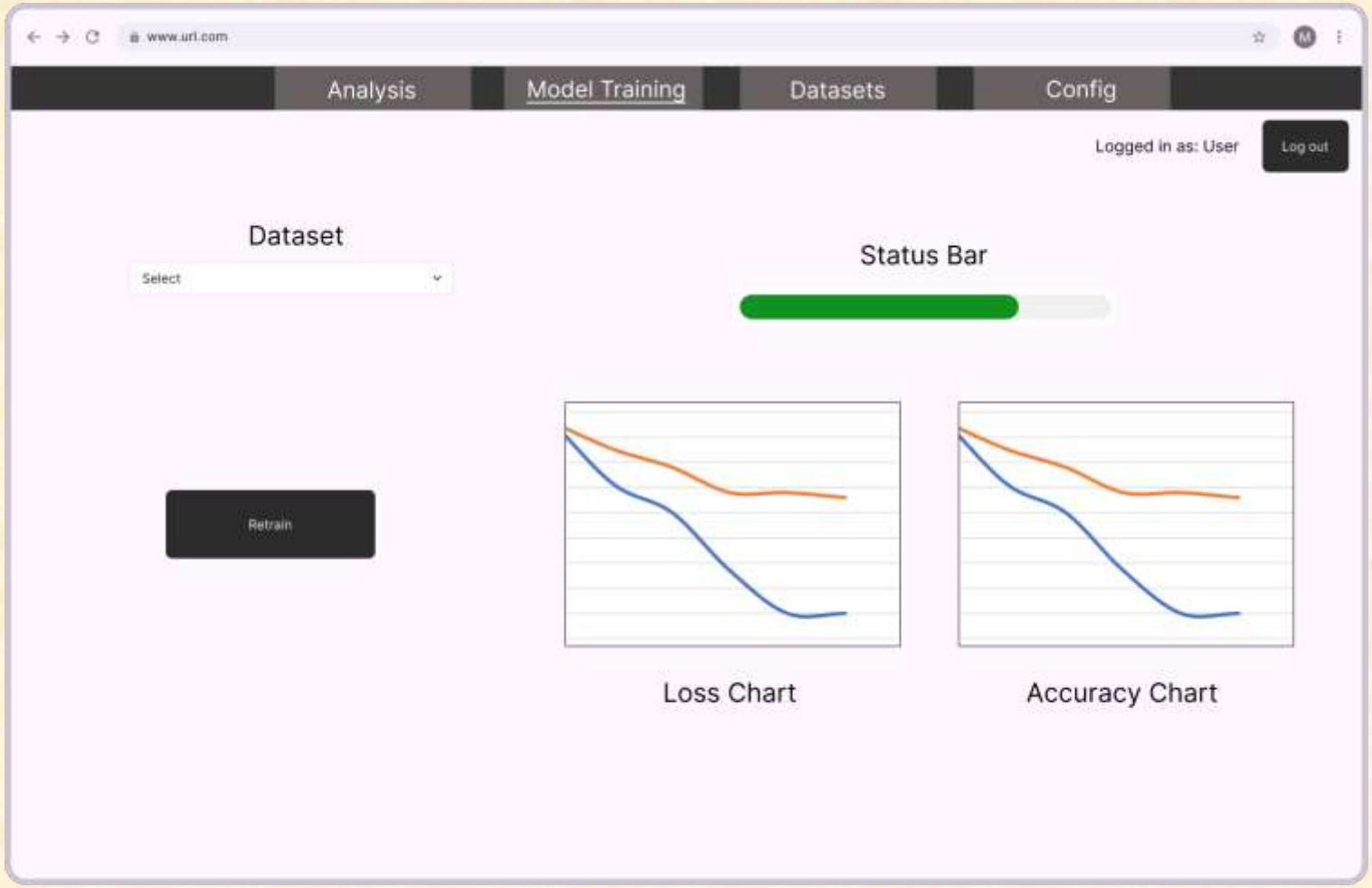
Screen Mockup: Analysis Page



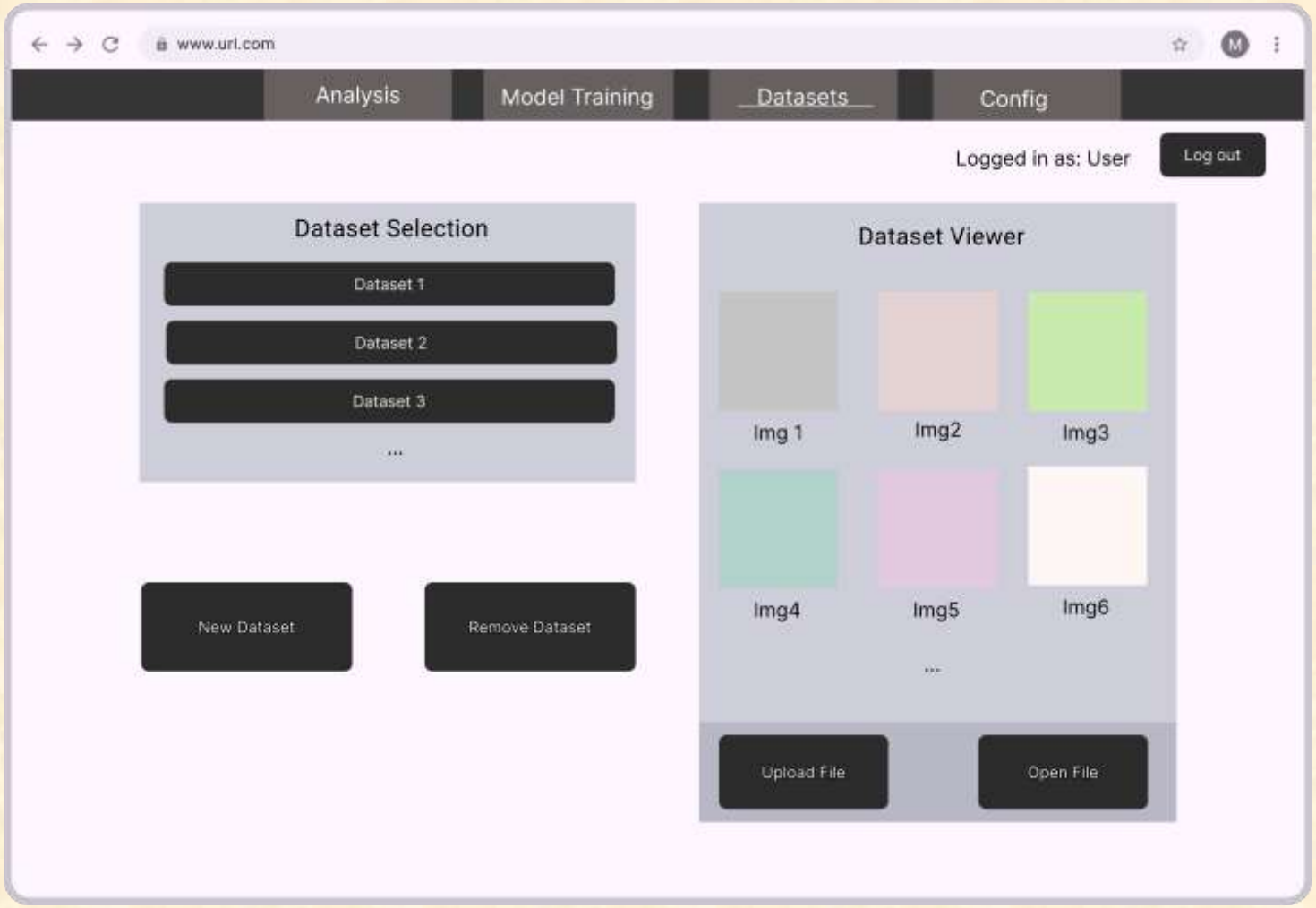
Screen Mockup: Config Page



Screen Mockup: Model Retraining Page



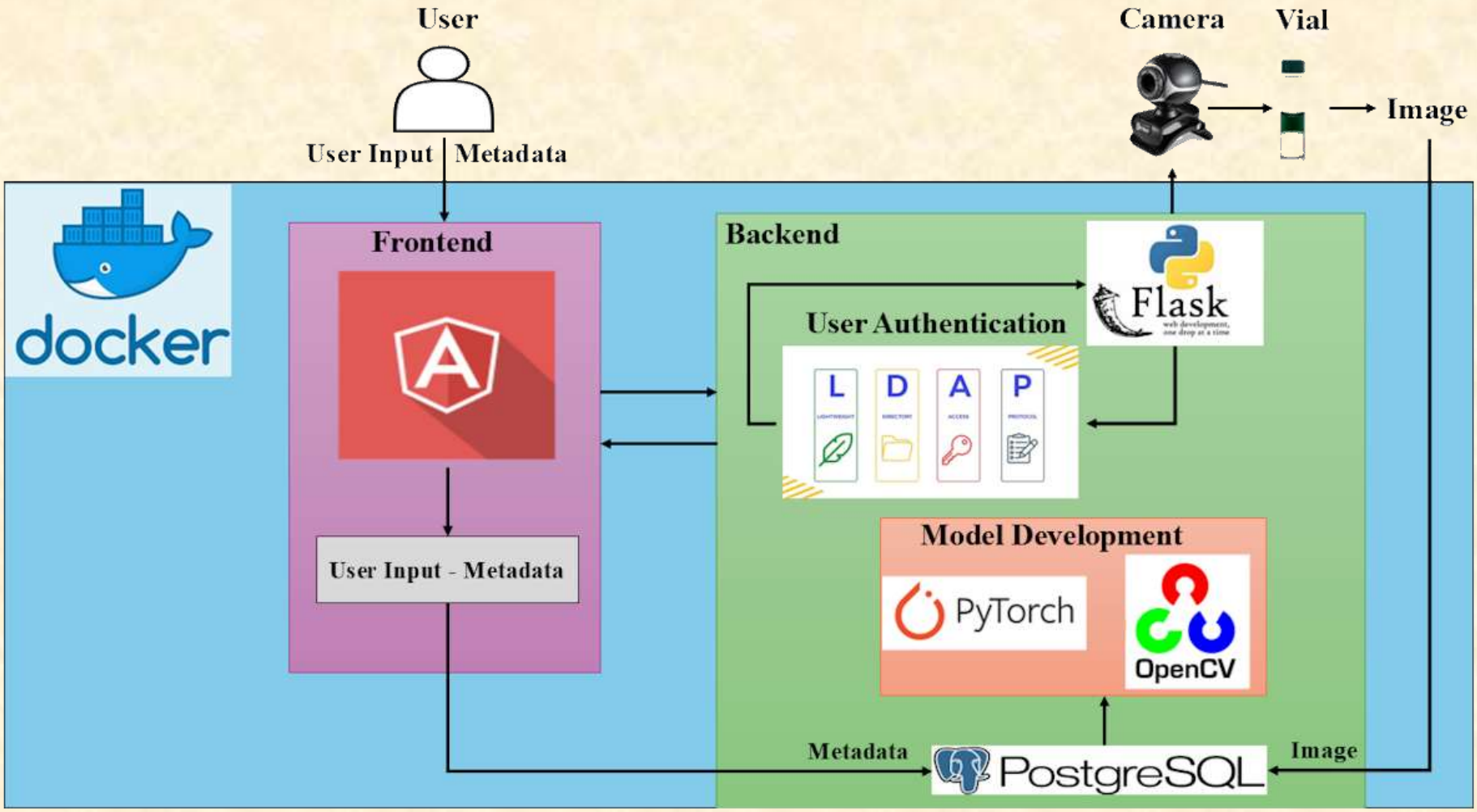
Screen Mockup: Dataset Viewer Page



Project Technical Specifications

- **Automated Image Processing**
  - The system uses Flask (Python) with a PyTorch-powered Mask-RCNN model for detecting vial boundaries, phase separation, turbidity, and emulsification through image segmentation.
- **Containerized Architecture**
  - Docker containers ensure consistency across development and production environments, with the backend using Flask and the frontend built using Angular for user interaction and metadata management.
- **Database/Data Management**
  - PostgreSQL, running in a Docker container, stores all the image data and metadata, while LDAP manages user authentication for secure access to the system
- **Model Retraining**
  - The system supports model retraining, allowing for scientists to classify images and maintain accuracy over time, adapting to new data patterns.

Project System Architecture



Project System Components

- Hardware Platforms
  - Camera
  - Light
- Software Platforms / Technologies
  - PyTorch
  - Flask
  - OpenCV
  - Docker
  - Angular
  - LDAP server
  - PostgreSQL

Project Risks

- Risk 1: **Non-Trivial Image Processing Task**
  - We don't know what algorithms will work, and many of the biphasic solutions are transparent, and we are not experts in the field of computer vision.
  - **Mitigation:** We have a solid first round of algorithms to test, and hopefully those algorithms will show that they are close enough to expected results and we can then refine them. If they do not work, then we can talk with MSU professors to get back on track.
- Risk 2: **Dataset Collection**
  - No prior datasets → Need to create our own dataset. Dataset may not be close enough to true images, could lead to an inability to generalize.
  - **Mitigation:** Using everyday chemicals from a hardware store, we can create true biphasic solutions. However, we don't know whether the solutions we can create will be representative enough of true datasets. We do have some images obtained with real chemicals from AbbVie that we can test against.
- Risk 3: **Embedded Model Retraining**
  - Model retraining takes a long time, we don't want the system to be unusable to the user for an extended period.
  - **Mitigation:** Write the current model to a path in our code base. Then, start an asynchronous background task that will train the new model. Once training is complete, write the newly trained model to the path.
- Risk 4: **Camera Resolution**
  - The model will have to find very small menisci; The camera AbbVie provided may not be have a high enough resolution to allow the model to find them.
  - **Mitigation:** Fix the distance, focus, and exposure to capture optimal images for the model and have as many pixels on features as possible.

Questions?