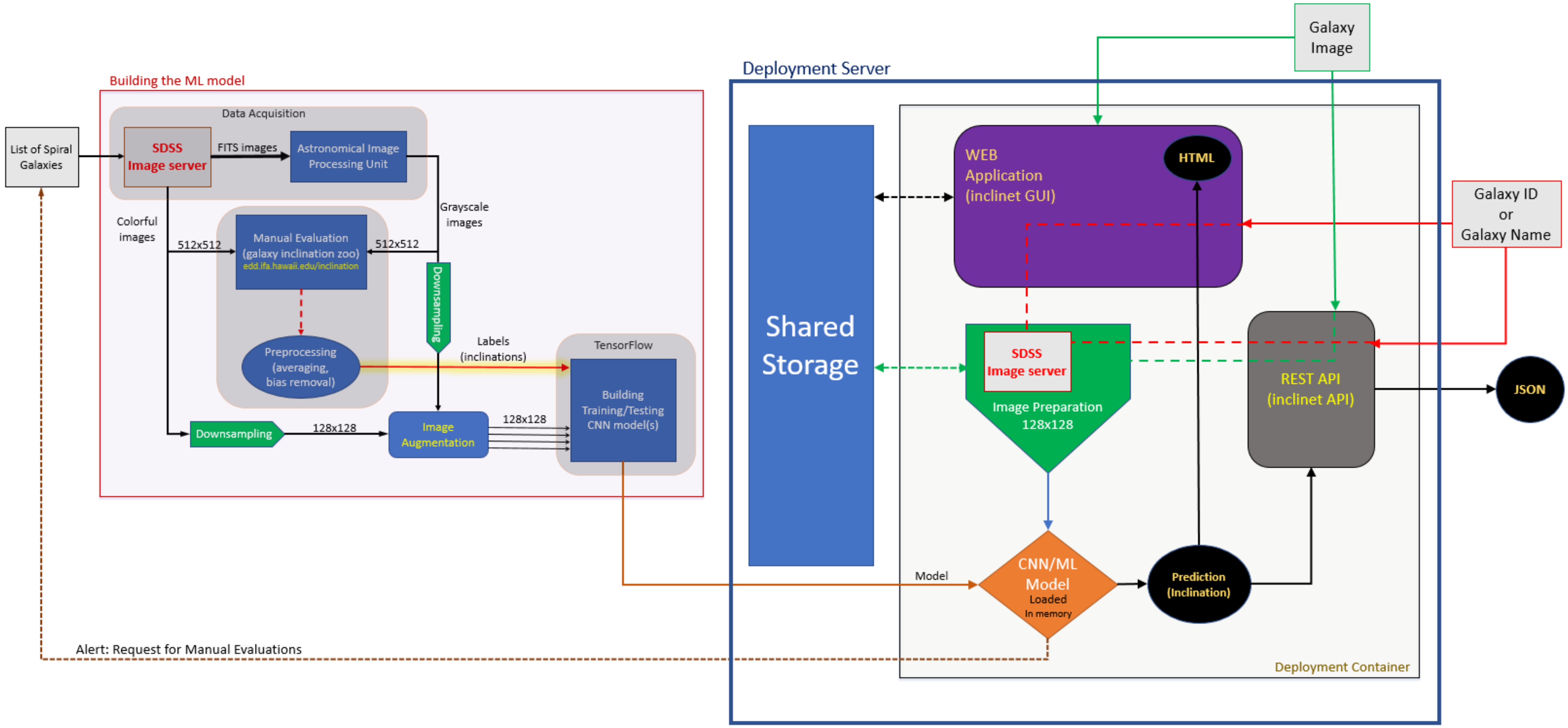

Main Components:

1. Model building Unit: This module consists of 4 main components:
    a. Data acquisition: Images are downloaded from the SDSS images web server (for this project we chose images taken from a specific telescope. For other telescopes, data acquisition and preprocessing might be different)
    b. Label generation: Galaxy images are evaluated by human users using an online GUI
    c. Image resizing and data augmentation: images are downsampled to the resolution of 128x128 to facilitate the training process, images are augmented to avoid overfitting

   d. ML model training/testing: CNN model(s) are trained and tested. Model(s) are then stored on disk and shipped to the deployment unit. Multiple models might be generated to compare their results at the deployment stage. If the results are far from each other, there would be an alarm going off asking for manual investigations to resolve the discrepancies.

2. Deployment Unit: the deployment has a monolithic architecture. The deployment container is executed on the deployment server and has access to the shared storage of the server through mounting the required folders.
   a. Shared Storage, stores the image products and log files of the online operation. Different components of the deployment container need the proper access to the storage to read and write data
   b. Deployment Unit: This is the heart of the service. The main component is developed using the Python package **Flask**. This unit hosts
      - Online web service, which includes a **web application**. Online users can communicate with the service through the online dashboard
      - API service, where users can send requests and images through a **REST API**
      - Models are preloaded into the deployment container. The CNN models reside in the memory as soon as the application is deployed, and there is no need to load them over at each API call. This accelerates the response time of the provided services
      - There is an image processing unit that prepares images to comply with the input format of the model network

Main Problem:

- Data is presented in the form of images. Non-square images are padded to have square dimensions, and then are resized to the appropriate shape
- Labels are the spatial inclinations of spiral galaxies from face-on

Can this problem be solved manually?

- Galaxy Inclination Zoo (GIZ) enables human users to cross compare galaxies with a set of standard galaxies with known inclinations and determine the inclinations of the target galaxy

- In ~60% of cases where images are clear with no ambiguity/anomaly, the axial ratio of the ellipse that encloses the galaxy can be used to determine the inclination

How Regularly the Model needs to be retrained?

- This model has been trained using the SDSS data and therefore it performs very well on the similar data set. If data is generated using other telescopes or have very different color schema, then the model training needs to be revisited. Note that SDSS has not covered the entire sky, and mainly surveyed the northern sky (due to the geographic limitations of the telescope)
- In case of planning for the evaluation of a huge batch of new galaxies with very different types of data, the model needs to be tested in advance. If the performance is not satisfactory, for a small subset of images (~1,000 galaxies) the manual evaluation needs to be performed and retrain the model to adapt the new data.
- The retrained model would be shipped to the deployment container in **h5** or **pickled** format.

In the case of retraining, where the data is stored?

- Normally data should be preprocessed to be compatible with our training pipeline. Each telescope and instrument has its own characteristics. The best practice is to generate 512x512 postage stamp images of galaxies. Data is accessed through the telescope's APIs, or the FTP/HTTP.

How to improve the model?

- The best way is to create many synthetic galaxy images, with known inclinations. The results of galaxy simulations can be visualized at various spatial inclinations under controlled situations. Later, the resolution can be tuned to different levels and the foreground, background objects can be superimposed on the image. Additional noise and ambiguities can be added to images. All of these factors allow the network to gain enough expertise on different examples of galaxy images.

How to make more complicated models (room for possible improvements):

- We can present all images taken at different wavebands in separate channels, instead of parsing them single entities. All passbands can be fed into the CNN at once.
- Images can be used in raw format. The dynamical range of the astronomical images are way beyond the 0-255 range. Instead of downscaling the dynamical range to produce the visualizable images, we can use the full dynamical range of the observations

Service Cost:

- This service is hosted on a preowned server located at the University of Hawaii data center. The initial cost of purchasing the physical server was ~$5,000.
- The annual charge per year for hosting the server in the data center is $125