# README

This application was developed in Google Colab and will work best there. If you want to run the notebooks as is, without having to change the code to point to local storage, please send us a message so we can add you to the Google Drive directory.

# Data Processing

1. Begin in the “Get\_Flickr\_Pics/Store\_in\_MongoDB” notebook located in Google Colab.
   1. Get the ip address of the Colab instance and save to MongoDB Atlas
   2. Run the notebook to get 500 image datasets, validate the descriptions and upload the subset with valid English descriptions to MongoDB Atlas
   3. Repeat to get the desired number of training photos
2. Go to the “Download\_Images\_To\_Google\_Drive” notebook located in Google Colab.
   1. Get the ip address of the Colab instance and save to MongoDB

# Model Training

After the training data has been downloaded to a local disk, do the following:

1. Run the 03\_model\_training\_prep.ipynb notebook to create the files to be used to train the model
2. Run the 04\_model\_training.ipynb notebook to train the model - this will take about 2 hours
3. Run the 05\_predict.ipynb notebook to make predictions