Sprint 4 Requirements Artifacts

Team 13

Group Members: Landon Pyko, Aiden Patel, Kai Achen, John Newman, Andrew Mcferrin

1) Break up functions to allow exclusions from processing
    - Separate scraping function from ML models
    - Output scraped data to CSV
    - Allow CSV modification to remove images
    - Pass in CSV to processing function to compute
    - Return computed outputs

2) Front end support to exclude images from scraped site
    - Process updated back end functions by calling scrape and process separately
    - Output scraped images to user first
    - Allow user to select and deselect images to process
    - Take in selection and update CSV
    - Pass in updated CSV to process function
    - Display returned output

3) Store past computations in database
    - Compute some hash with image and text tuples
    - Store hash alongside computer alt-text
    - Retrieve alt-text when same image and text combinations are scraped from a site to reduce process time
    - Store full sets of data as well so users can pull up entire sites that have been computed in the past

4) Create demo site to show precomputed results
    - Run the program on multiple different sites in advance
    - Save their output
    - Create site that allows users to select demo sites to view output
    - Shows user proof of concept without spending time computing