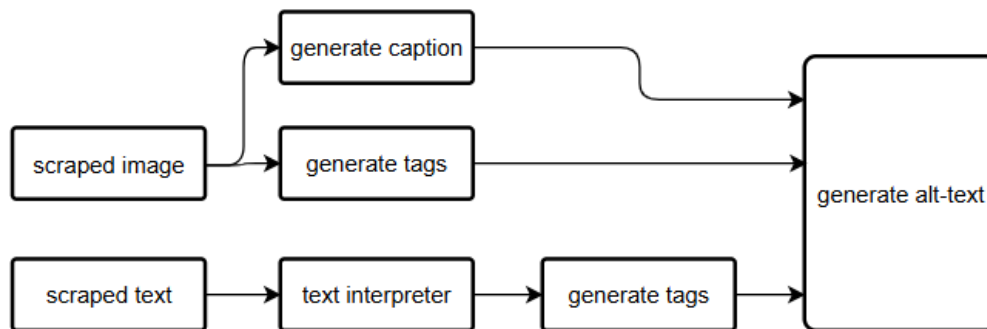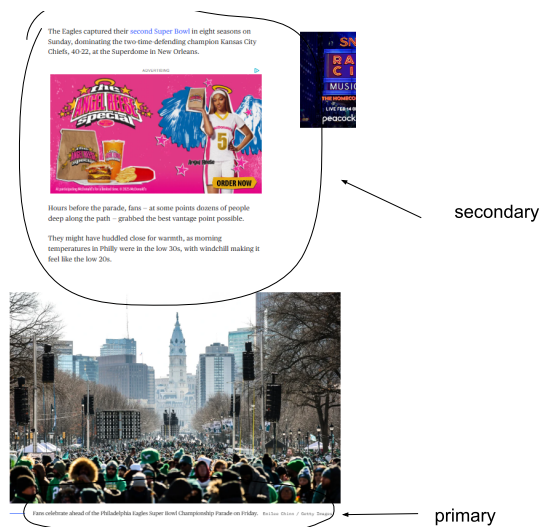# Sprint 2 Requirements Artifacts
Team 13
Group Members: Landon Pyko, Aiden Patel, Kai Achen, John Newman, Andrew Mcferrin

1. Image->alt-text intermediate step

2. Create text web scraper
   - Scrape important text from websites
     - text relating to images, such as captions or in-text references to provide primary context
     - Overall text of webpage to gain secondary context

3. Create image web scraper
   - Scrape images from websites
   - Download images into structure/format the model will be able to read
   - Extract metadata from each image to provide additional context

4. Interpret text scraped data
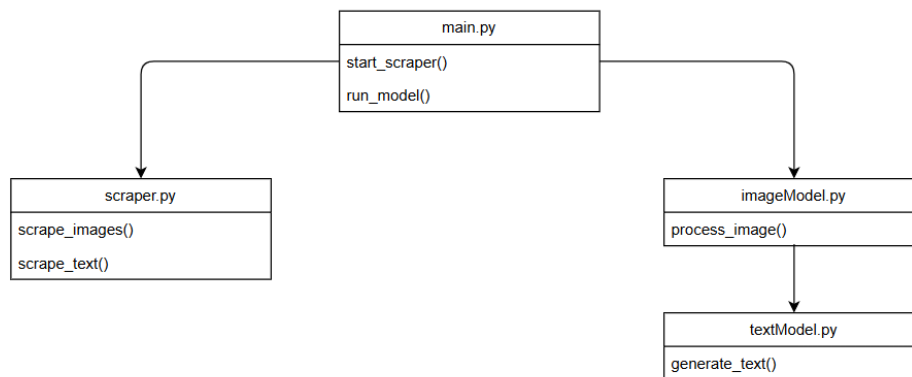   - Extract relevant tags and words from data
     - Named entities (people, companies, places, etc)
   - Determine relevant relationships between tags

5. Interpret image scraped data
   - Handle if the image needs to be converted to some specific type
   - Format the image
   - Pass the image to the model to extract tags

6. Create function interface for scraper
   Have a main.py to run each part and pass data from scraper to model

7. Test model on scraped data
   - Ensure models works for all possible input of scraped data
   - Test edge cases (website with only images, website with no images, etc.)
   - Ensure accuracy of provided alt-text- note inaccurate text