CS418/518 Project – Milestone 2 – Project A (Figure annotator)

In this milestone, you will build the search function in your search engine using Elasticsearch and other functions using MySQL.

Base requirements

1. The website should index the metadata fields of all compound figures by Elasticsearch viewable for both the admin and users;
2. Next to the search box, there should be a button called “Search”. After this button is pressed, the search engine will return a page containing compound figures whose captions match the keywords. This should result in a search result page showing the thumbnails of figures and the number of results.
3. Next to the search box, there should be another button called “Search Annotation Tasks”.
4. When clicking the “Search Annotation Tasks” button without putting anything in the search box, the search engine will return a page containing compound figures that are assigned to the current user (see below).
5. The search result page should show the compound figure file names, the thumbnails of the compound figures. The number of results should be displayed at the top.
6. The search results are paginated.
7. The admin should see a view to assign annotation tasks to one or more users. The schematic view looks like below

User	Assigned group (a drop-down menu)
Jwu	1

8. The view displays the progress of a user’s annotation tasks. For example, out of X number of compound figures assigned to user John, Y number of figures have been finished. The schematic view looks like below. Note that one user can be assigned with multiple groups and the same group can be assigned to multiple users.

User	Assigned Group	Assign Datetime	#Compound figures	#Finished
Jwu	1	2022-10-11 7:31:00	50	25
Jwu	2	2022-10-01 6:30:00	50	50
Jhe	2	2022-10-11 8:00:00	50	10

9. On the backend, when the admin assigns a group of compound figures to a certain user, all the compound figure names are copied from the main index to a new table called *annotation* in Elasticsearch. The schema of this table looks like below:

{

“compoundfigure_file”: “USD0543554-20070529-D00001”,

“assignments”:[{“assign_id”:1, “user_id”:”jwu”, ”datetime”:”2022-10-11 07:31:00”,

”annotations”:{“seg_correct”:”yes”,”n_subfigure”:2,“subfigures”:[

{“subfigure_filie”:”USD0543554-20070529-D00001_1”, ”object_correct”:”yes”, “object”:”shoe”, ”aspect_correct”:”no”, “aspect”:”front view”},

{“subfigure_id”:”USD0543554-20070529-D00001_2”, ”object_correct”:”no”, ”object”:”shoe”, ”aspect”:”cross section”}] // end subfigure

} // end annotation

] //end assignments

} // end root

In the example above, a compound figure file “USD0543554-20070529-D00001” is assigned to user “jwu” for annotation. This compound figure was segmented into two subfigures. The annotator (a user that use the web portal to annotate figures) needs to determine whether the segmentation is correct or not. The information will be written into the “annotations” object. The annotator will also need to determine whether object names and aspects are correctly extracted from the original figure caption. The information will be written into the “subfigures” object. When the assignment was initially made, each compound figure file needs to be assigned with the designated user and the annotations object contains empty fields, except the subfigure_file field, until the annotator fills in.

Submission

All source codes and should be committed to GitHub by noon, Thursday, November 10, 2022. The report should be submitted to Canvas by November 15, 2022.