Project: Figure Annotation (Project A)

Student name: Priyanka Gummagatta Anil Kumar

Student UIN: 01125294

1. **Overview**

The goal of the Project is to build a web-based Figure annotation interface. This project's objective is to create a user interface for figure annotations that enables users to search figures using a variety of criteria. Additionally, users can indicate figures that have been assigned, assigned to the user, and assigned but unannotated. The "Admin Role" and "User Role" are the two distinct roles for this application. The features deployed in this project up to this Milestone include user login for users, role-based authentication, two-factor authentication using Twilio, and user registration based on email id. User group relationships, elastic search, search annotated tasks, password encryption, password change, forgotten password, landing page, home page with a search button, admin approval of users and assigning them to various groups are all features.

I have used HTML, CSS with Bootstrap, Javascript and PHP to build the project and MySQL as the Database and Elastic Search for data indexing.

2. **Milestone Accomplishments**

List ALL specifications of the CURRENT MILESTONE and specify which specifications are fulfilled or not.

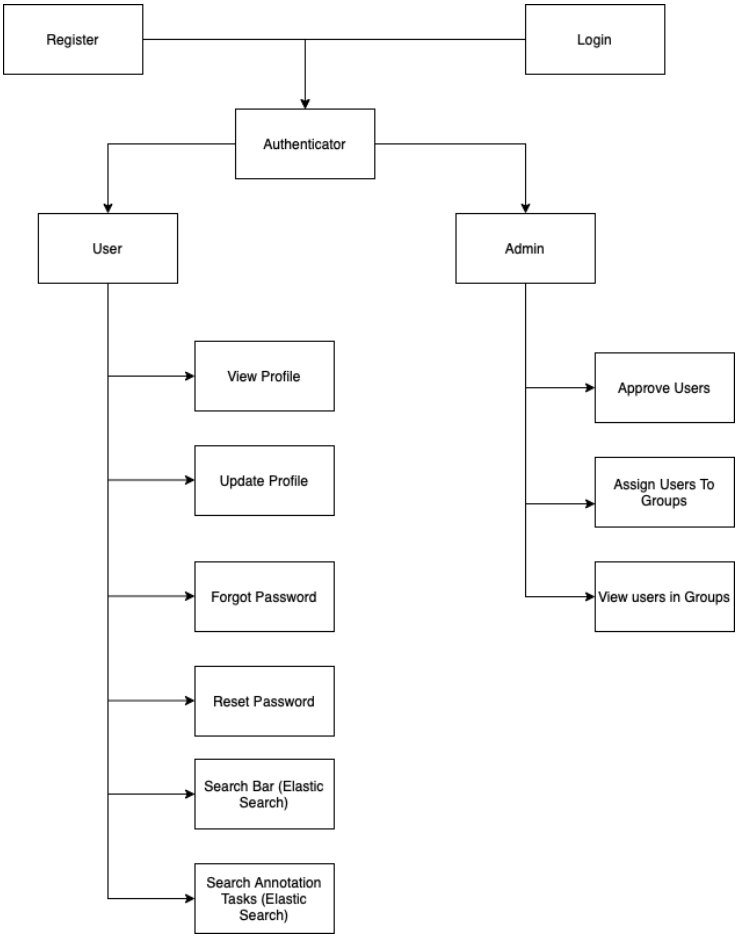
Table 1: Status of milestone specifications.

Fulfilled	Feature#	Specification
Yes	1	The website should index the metadata fields of all compound figures by Elasticsearch viewable for both the admin and users
Yes	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
Yes	3	Next to the search box, there should be another button called “Search Annotation Tasks”
Yes	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

Yes	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
Yes	6	The search results are paginated.
Yes	7	The admin should see a view to assign annotation tasks to one or more users
Yes	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
Yes	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

3. Architecture

The project has Been developed in HTML and CSS framework using PHP and MySQL as the database. The project has the landing page which will have the login option for the user. Once the user signs in based on the credentials he will either be directed to the admin page or the user page



4. Data

The dataset consists of several figures and their sub-figures from numerous patents. 300 artifacts, 359 patents, 831 compound figures, and 1998 single figures make up the collection. One folder had the photographs for the segmented figures, and the other contained the images for the compound figures. The figure_segmented_nipseval_test2007 provides the metadata for all compound figures and subfigures. This table was initially imported into the MySQL database. These fields in the table are:

Field	**Type**	**Key**	**Example**
id	INT	Primay key	1
patentID	VARCHAR(255)	not null	763
patentdate	DATE	not null	2022/11/11

figid	VARCHAR(255)	not null	3
caption	VARCHAR(1000)	not null	FIG. 2 is a perspective view thereof;
object	VARCHAR(1000)	not null	Portions of a connector plug and receptacle pair
aspect	VARCHAR(1000)	not null	perspective view
figure file	VARCHAR(255)	not null	USD0544845-20070619-D00001.png
subfigure file	VARCHAR(255)	not null	USD0544845-20070619-D00001_2.png
object title	VARCHAR(255)	not null	Portions of a connector plug and receptacle pair
groupID	VARCHAR(255)	not null	1

There are unique subfigure IDs in every row, but there may be repetition of compound finger files and patent IDs since every patent may contain multiple subfigues. The caption, object, aspect, and object title are all details on the subfigures figures. These figures are further categorised into groups and are identified by group IDs. Every group consists of 100 compound figures. To fetch the images to the Front-End, I have used Front-End folder structure and used the relative path name to fetch the figures.

A new table - “UserGroups” is created which has a many-to-many relationship to store the relationship between users and groups to implement the user group feature.

Field	**Type**	**Key**	**Example**
id	INT	Primary	1
email	VARCHAR(255)	Foreign (Reference USERS), not null	pgumm001@odu.edu
groupID	VARCHAR(255)	Foreign(Reference figure_segmented_nipseval_test2007 table), not null	10
created_at	DATETIME	null	2022-11-02 00:08:45

5. Elasticsearch Design

We must first establish a database before we can save the metadata in Elasticsearch. To start elastic Database in the local machine, I first navigated to the folder where Elasticsearch was located and ran the command bin/elasticseacrch. This starts the Database on the url - http://localhost:9200. I used Kibana's interface to use the Elasticsearch Database. To run this, I navigated to the folder where Kibana was downloaded and ran bin/kibana. This started the kibana interface on http://localhost:5601. To insert the data into Elasticsearch Database, I used Kibana's interface and import the CSV data, exported from MySQL Database. Every document entry into the Database looks as below:

{

"patentID": "USD0544845-20070619",

"patentdate": "2007-06-19",

"figid": "1",

"caption": "FIG. 1 is a front elevation view of a portion of a connector plug, showing our new design;",

"object": "Portions of a connector plug and receptacle pair",

"aspect": "front elevation view",

"figure_file": "USD0544845-20070619-D00001.png",

"subfigure_file": "USD0544845-20070619-D00001_1.png",

"object_title": "Portions of a connector plug and receptacle pair",

"groupID": "1",

"figure_url":"https://firebasestorage.googleapis.com/v0/b/web-project-eb94d.appspot.com/o/compoundimages%2FUSD0544845-20070619-D00001.png?alt=media&token=324ddb56-cde9-4bf9-88ce-abc52f800afa"

}

The below query can be used to retrieve all data from the Elastic search DB

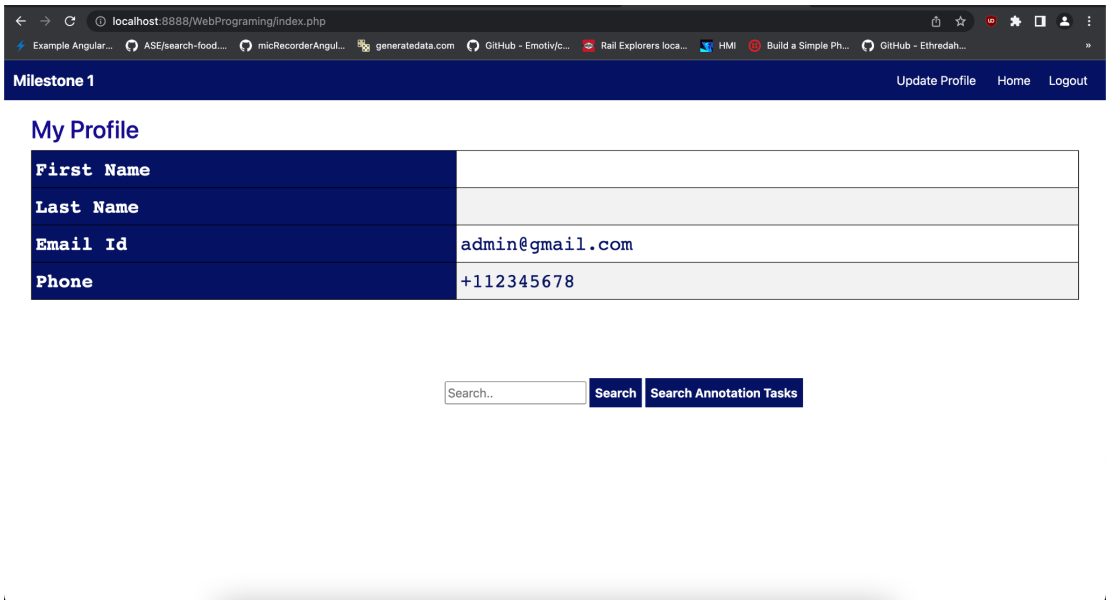
```
GET annotation/_search
{
  "size" :2000,
  "query": {
    "match_all": {}
  }
}
```

6. Implementation

User Home Page with Search and Search Annotation

The User Home page has two buttons:- “Search And Search Annotation Tasks” and a search box where the user can enter a word/search phrase. The Elasticsearch function is called using the button next to the search bar to return all the pages that contain the keyword in the captioncolumn. When the "Search Annotated Tasks" button is pressed without any search criteria set, a list of all the compound figures that are members of the group the user belongs to is displayed.

Code file: index.php



Search Result Page:

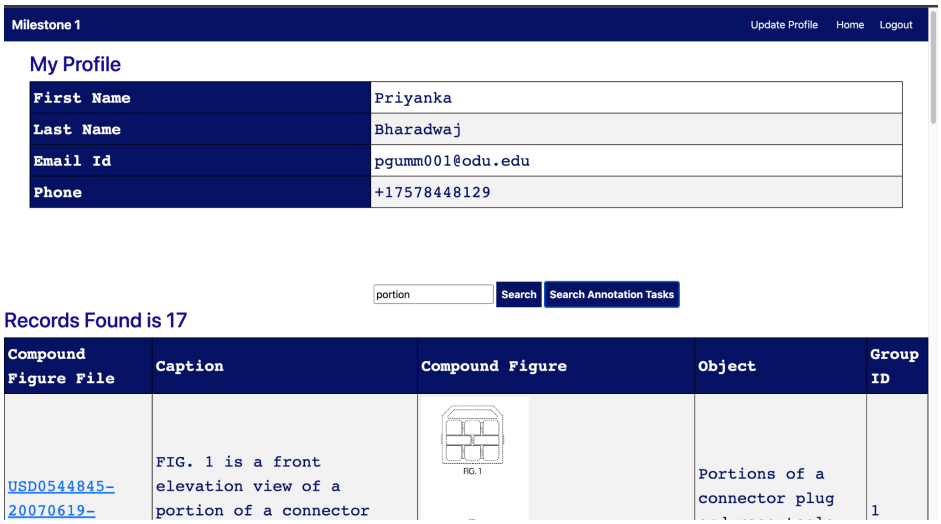
On this page, you may see a list of compound figures together with their file names, images, and captions. The top of the search result shows the total number of results that were fetched. This is page paginates the results and displays 10 records per page.

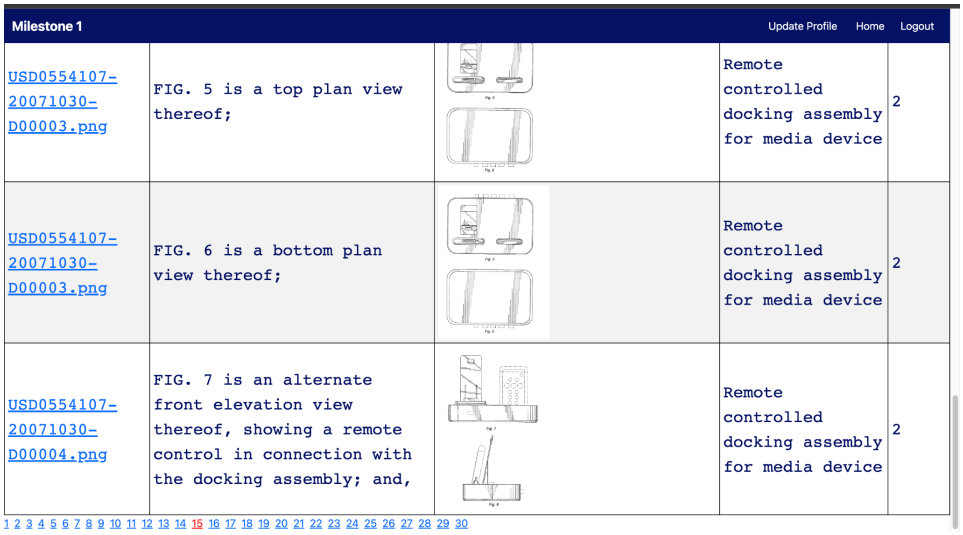
The Elastic search query used to get the results is :

```
GET annotation/_search
{
 "query": {
  "bool": {
   "must": [],
   "filter": [
    {
     "bool": {
      "should": [
       {
        "match": {
         "caption": "shelf"
```

```
          }
         }
        ],
        "minimum_should_match": 1
       }
      }]
   }
  }
}
```

Code file: searchCaption.php



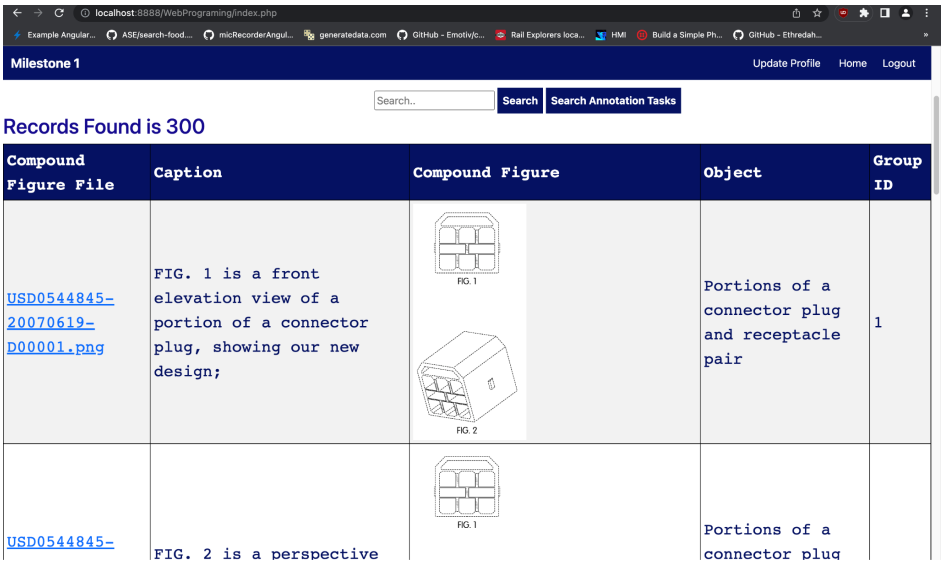


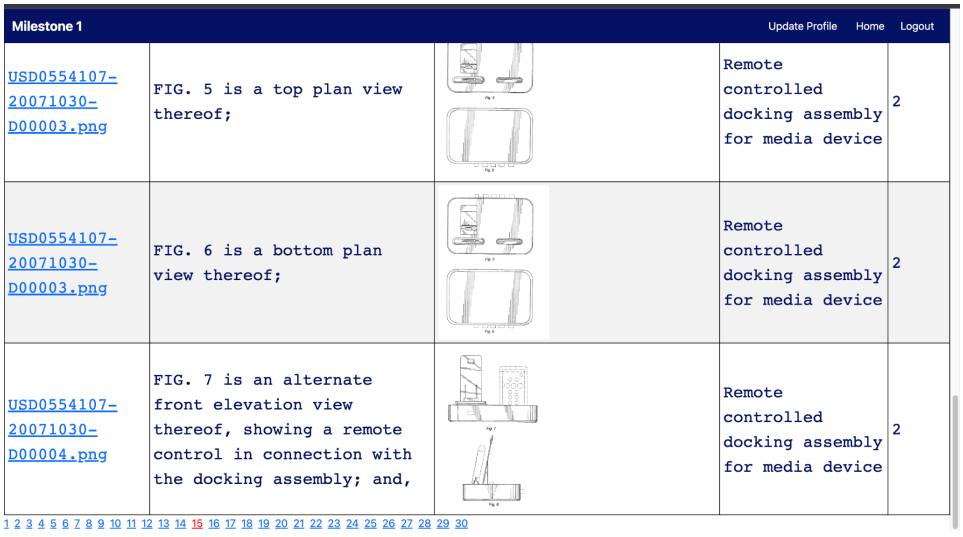
Search Annotated Tasks:

If the search bar is left blank and the "Search Annotated Tasks" button is pressed, all compound figures in the groups the user is a part of are displayed. This is page paginates the results and displays 10 records per page.

Code file:searcAnnotatedTasks.php

```
GET annotation/_search
{
  "size":1998,
 "query": {
   "query_string": {
     "query": "(caption:'' OR object:'') AND (groupID:1)"
   }
 }
}
```

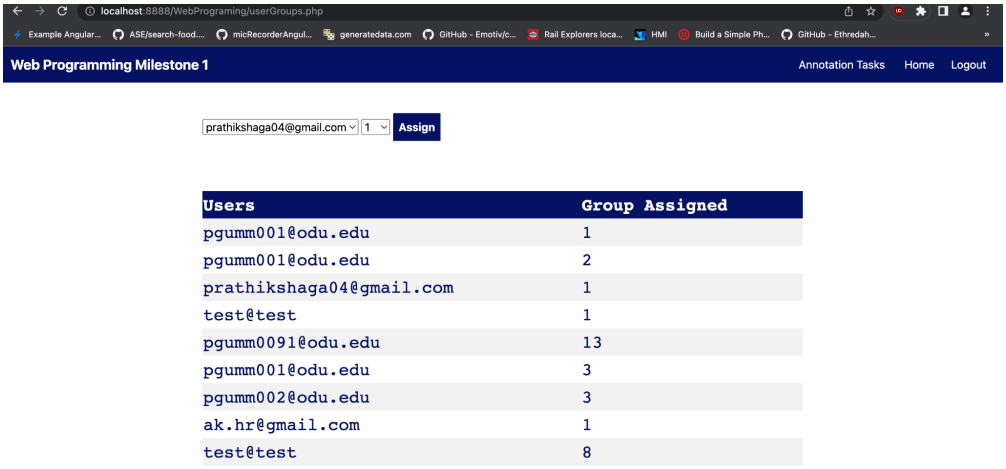




Add User to a Group:

Any user can be assigned to any group by the admin. Users may belong to more than one group, and each group may have more than one user.Admin can choose from a list of users in the user drop-down, and groups can be selected from a list of groups in the group drop-down. All compound figure names are copied from the main index to a new table in Elasticsearch called annotation when a user is assigned to a group.

Code file: assignUser.php

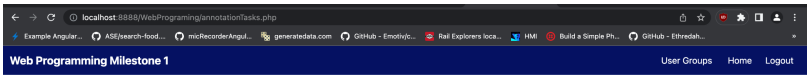


Users	Group Assigned
pgumm001@odu.edu	1
pgumm001@odu.edu	2
prathikshaga04@gmail.com	1
test@test	1
pgumm0091@odu.edu	13
pgumm001@odu.edu	3
pgumm002@odu.edu	3
ak.hr@gmail.com	1
test@test	8

List all User assigned to Groups:

The list of all user group relationships and the dates that each user was allocated to a certain group is available to the administrator. Additionally, it shows how many of the compound figures assigned to the user in that particular group have been annotated or completed. Instead of the total number of figures allocated to the user, this number refers to the figures in that particular group that the user is assigned to.

Code file:assignUser.php



Users	Group Assigned	Date Assigned	#Compound Figures	#Finished
pgumm001@odu.edu	1		100	0
pgumm001@odu.edu	2		100	0
prathikshaga04@gmail.com	1		100	0
test@test	1		100	0
pgumm0091@odu.edu	13		100	0
pgumm001@odu.edu	3	2022-11-19 00:00:00	100	0
pgumm002@odu.edu	3	2022-11-19 07:01:32	100	0
ak.hr@gmail.com	1	2022-11-19 07:04:11	100	0
test@test	8	2022-11-19 07:04:21	100	0