Cloud Computing and Big Data – Summer 2022

Homework Assignment 3

Assignment:

Build the database for the application.

Outline:

1. Load the data in Elasticsearch using the data from the ES.CSV file provided.
    a. Create an index named ‘posts’ in Elasticsearch and load the CSV data into it
    b. Use the field ‘id’ in CSV file as the document id in the Elasticsearch index.
2. Load the data in DynamoDB using the data from the DynamoDB.CSV file provided.
    a. Create a Table named ‘posts’ in DynamoDB and load the CSV data into it.
    b. Use the field ‘id’ in CSV file as the Partition key in the DynamoDB table.
3. Upload and index questions:
    a. Modify the Lambda function LF0 from assignment 2, to implement upload functionality for the questions provided by the user.
    b. Load the data into DynamoDB under the ‘posts’ table using the POST method in the API and LF0.
4. Search for the answers:
    a. Create a lambda function (LF1).
    b. Given a search query ‘q’ return 3 posts with the same tag. To implement this feature, search the ‘posts’ index in Elasticsearch using the tag provided as input by the user and return the post id. Use these posts id’s to query the DynamoDB table, get the post body from the DynamoDB, and return it as the result to the user.
    c. If no match is found return a message “no answers found for this category”.

Acceptance criteria:

1. Given a question you should be able to use the API and Lambda function created to load the question into DynamoDB.
2. Your website should be able to search and display 3 answers related to the tag provided by the user and handle all the cases as mentioned above.

Submission Instructions:

1. This assignment is to be done individually.
2. Please upload your files (Lambda function code) on GitHub and create a release. No changes to the repository after the assignment deadline.
3. Submit a document with your name, S3 URL to your website and GitHub link on Brightspace.