Steps to setup project

1) DOWNLOAD CODE FROM GITHUB REPOSITORY

Open VSCode or any other Editor

- a) Download code from Git Repository
 - i) gh repo clone nittysharma/metdata-enrichment
 - ii) cd metdata-enrichment

2) INSTALL REQUIRED NODE PACKAGES

```
npm install aws-amplify
npm install -g @aws-amplify/cli
npm i —save @fortawesome/fontawesome-svg-core
npm i —save @fortawesome/free-solid-svg-icons
npm i —save @fortawesome/free-regular-svg-icons
npm i —save @fortawesome/react-fontawesome@latest
npm install aws-sdk
```

3) SET-UP AMPLIFY

- a) Run below commands in code editor to setup amplify
 - i) amplify configure

Please refer to https://docs.amplify.aws/cli/start/install/ to know how to initialize in detail if you are new to amplify It will also ask you to create a user, so create an IAM user with access to all the related services mentioned below.

Download the credentials and save it.

Update Configuration parameters

Navigate to src/config.json in your editor and update below properties

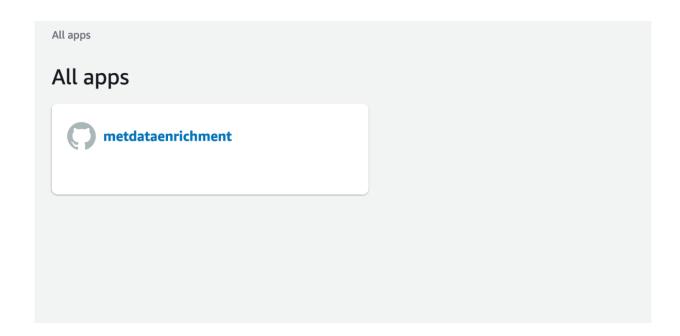
```
AWSRegion - Region where this application will be deployed

TableName - Name of the Dynamo DB table that will be created by CDK temporates.

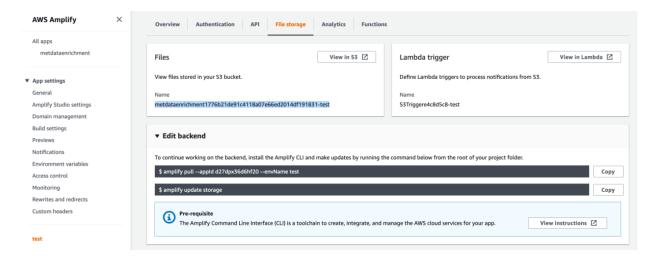
e.g. assetMetadata is default table name set
```

Leave OpenSearchUrl for now, we will update once we have deployed the entire workflow.

- ii) amplify init This will initialize the amplify settings, set-up the profile using access key and secret key
- iii) amplify status Verify the Amplify status
- iv) amplify publish Publish the application, after successful publish it will
- b) Go to AWS Amplify service in AWS console and you will see that application is published over here



Once Deployed go to File Storage menu and copy bucket name, you will be using it later



Go to Lambda trigger created as part of amplify and remove that manually from S3 Bucket as we will be creating the lambda trigger from CDK.

5) CDK TEMPLATE

The following

- a) Following software installations are required. CLI needs to be configured to access you AWS account. Instructions can be found here.
 - Open JDK 11 or higher
 - Maven 3.9 or higher

- CDK v2.7
- AWS CLI 2.11
- b) Check out the project source code from the Git repository.
 gh repo clone deeptichugh28/metadata-enrichment-java-2
 Structure looks like shown here.



Open cdk.json file and update the s3BucketName key with the name of S3 bucket created as part of the Amplify. Also update tableName withe value assetMetaData. Update the apiGatewayRegion to the region you are running from.

```
"@aws-cdk/aws-redshift:columnId": true,
"s3BucketName": "rtananthan-test-cdk-image-upload",
"tableName": "assetMetadata",
"apiGatewayRegion": "us-east-1"
```

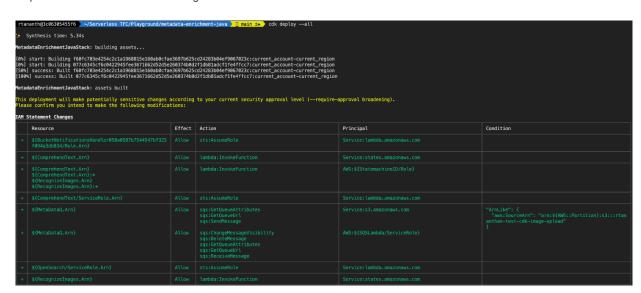
- c) In the terminal.command window of the operating system navigate to the root folder of the project(metadata-enrichment-java)
- d) Bootstrap CDK environment with the command

```
cdk bootstrap aws://<YOUR_AWS_ACCOUNT_ID>/<REGION>

~/Serverless TFC/Playground/metadata-enrichment-java > [] main ±+ > cdk bootstrap aws:// //ap-southeast-2
```

e) To deploy the application execute the following command

Output will look similar to the following

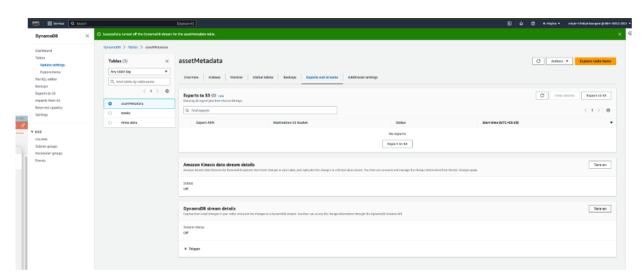


	Policy Changes				
Reso	urce	Managed Policy ARN			
	cketNotificationsHandler050a0587b7544547bf325f094a3db834/Role}	arn:\${AWS::Partition}:iam::aws:policy/service-role/AWSLambdaBasicExecutionRole			
	mprehendText/ServiceRole} mprehendText/ServiceRole}	$ann: \$ \{AWS:: Partition\}: iam:: aws: policy/service-role/AWSLambdaBasicExecutionRole arn: \$ \{AWS:: Partition\}: iam:: aws: policy/AmazonDynamoDBFullAccess$			
		arn:\${AWS::Partition}:iam::aws:policy/service-role/AWSLambdaBasicExecutionRole			
+ \${Re + \${Re + \${Re	cognizeImages/ServiceRole} cognizeImages/ServiceRole} cognizeImages/ServiceRole} cognizeImages/ServiceRole} cognizeImages/ServiceRole}	<pre>arn:\${AWS::Partition}::ams::policy/service-role/AWSLambdaBasicExecutionRole arn:\${AWS::Partition}::ams::policy/AmazonS3FullAccess arn:\${AWS::Partition}::am::aws::policy/AmazonRekognitionFullAccess arn:\${AWS::Partition}::am::aws::policy/AmazonBynamoDBFullAccess arn:\${AWS::Partition}::am::aws::policy/ComprehendFullAccess</pre>			
	SLambda/ServiceRole}	arn:\${AWS::Partition}:iam::aws:policy/service-role/AWSLambdaBasicExecutionRole			
(NOTE: There may be security-related changes not in this list. See https://github.com/aws/aws-cdk/issues/1299)					
Do you wish to deploy these changes (y/n)? y MetadataEnrichmentJavaStack: deploying [1/1] [0%] start: Publishing f60fc703e4254c2c1a1968815e160ab0cfae3697b625cd24283b04ef9067023c:current_account-current_region [0%] start: Publishing 077c6345cf6c0422945fee3671662d52d5e260374b0d2f1db81adcf1fe4ffcc7:current_account-current_region [50%] success: Published 077c6345cf6c0422945fee3671662d52d5e260374b0d2f1db81adcf1fe4ffcc7:current_account-current_region [100%] success: Published 670fc703e4254c2c1a1968815e160ab0cfae3697b625cd24283b04ef9067023c:current_account-current_region MetadataEnrichmentJavaStack: creating CloudFormation changeset					
Metad	MetadataEnrichmentJavaStack				
+ Deploy	Deployment time: 189.29s				
stack ARN: irn:aws:cl	tack ARN: rn:aws:cloudformation:ap-southeast-2:094595487946:stack/MetadataEnrichmentJavaStack/20885710-ced7-11ed-a641-0652b6293ed6				
	Total time: 194.63s				

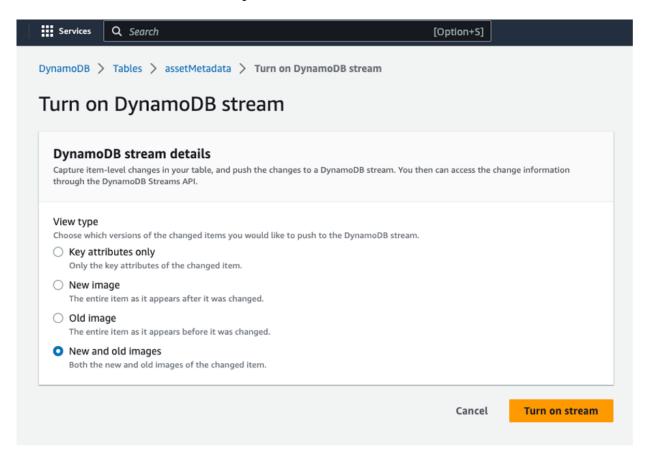
f) To remove the application and associated resources execute the following command.

cdk destroy --all

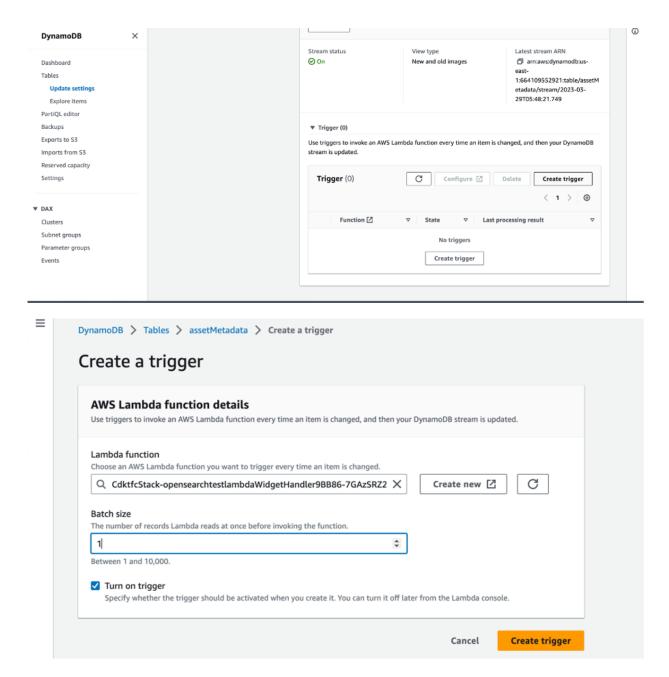
Step1: Please make sure Dynamo DB Stream is enabled for table "assetMetaData" you have created. Go to DynamoDB—Tables—click assetMetaData—Export and Streams—scroll down by default DynamoDb Stream and if it is disabled click on "Turn On" button on right



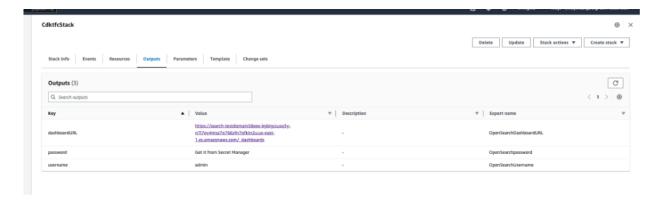
On Next screen Select both "New and Old Images" and click "Turn on Stream"



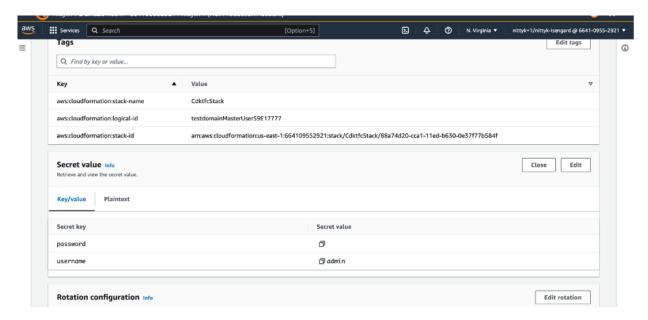
Add lambda trigger which we have just created.



Step2: Configure Open Search Go to Cloud-formation template Click on "" go to output session, You will see OpenSearchDashboard URL. Copy this URL and Open it on another browser tab. Enter username as "admin"



For password go to AWS Secret manager



Use this password to login in Open Search Dashboard.



Please login to OpenSearch Dashboards

If you have forgotten your username or password, please ask your system administrator

2	admin		
ı	••••••		
<u>Log In</u>			

After Login skip add sample option, On left side menu go to Dev Tools, Execute below Script

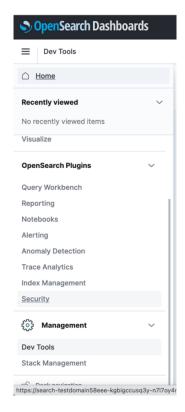
```
PUT /news
"settings": {
"index": {
"number_of_shards": 2,
"number_of_replicas": 1
"mappings": {
"properties": {
"assetKey": {
"type": "text"
"CelebrityData": {
"type": "text"
"Objects": {
"type": "text"
"Sentiment": {
"type": "text"
"Texts": {
"type": "text"
```

```
},
"aliases": {
"news1": {}
```

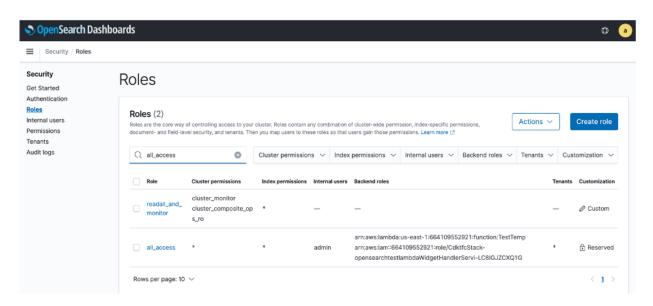
```
OpenSearch Dashboards
                                                                                             (C) (a)
 ■ Dev Tools
 Console
History Settings Help
                                             ⊅ଷ୍| 1 |
```

This will create index with name of "news".

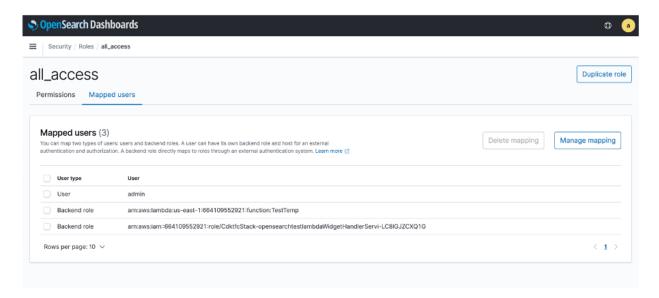
Now we need to add backend role of lambda to openSearch. Go to left side menu--->Security



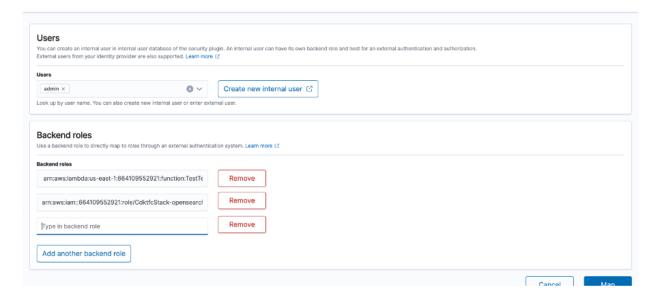
Click on Roles—search for "all_access" from search bar. Click on all_access



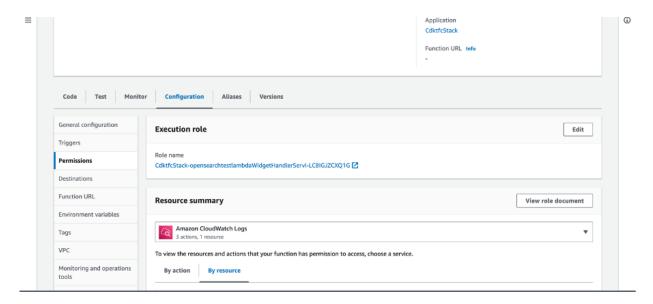
On Next screen click on Mapped User Tab: Here we need to add Role of our current lambda which is in between Click on Manage Mapping



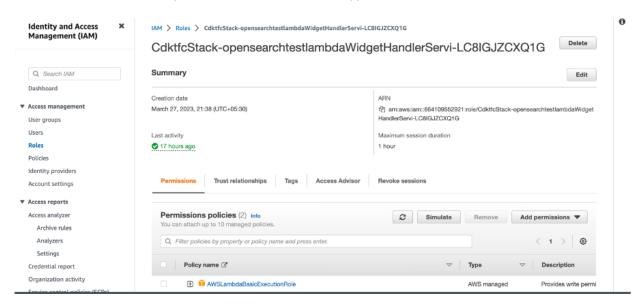
We have two sections User and backend role. We need to add new backend role.



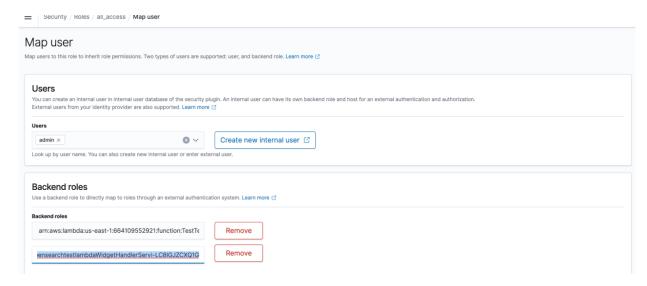
Go To aws console again to get Lambda role ARN,



click on this roleName It will take you to IAM screen where we can copy ROLE ARN.



Paste this role in Opensearch Backend Role and save it.



For every add/update/delete in dynamodb table we will get DynamoDb Stream that will trigger lambda and lambda will feed open search. we can check same by running.



7) EXPOSING AS API GATEWAY

After we have successfully setup the Opensearch service, we will expose this via API gateway. There is already one Lambda function "opensearch-lambda" created by CDK template. This function will query OpenSearch Service and return results.

CREATE AND CONFIGURE THE API

TO CREATE YOUR API USING THE API GATEWAY CONSOLE

- 1. Navigate to the API Gateway console at https://console.aws.amazon.com/apigateway/home
- On the left navigation pane, choose APIs.
- Click on Create API
- Locate REST API (not private) and choose Build.
- On the following page, locate the Create new API section and make sure New API is selected.

- Configure the following fields:
- API name: opensearch-api
- Description: API for searching assets using Metadata using an Amazon OpenSearch Service
- Endpoint Type: Regional
- Choose Create API.
- Choose Actions and Create Method.
- Select GET in the dropdown and click the checkmark to confirm.
- Configure the following settings, then choose Save:
 - o Integration type Lambda Function
 - o Use Lambda proxy integration Yes
 - Lambda Region <REGION>
 - o Lambda Function opensearch-lambda
 - Use Default Timeout Yes

CONFIGURE THE METHOD REQUEST

CHOOSE METHOD REQUEST AND CONFIGURE THE FOLLOWING SETTINGS:

	Setting	Value
1	Authorization	NONE
2	Request Validator	Validate query string parameters and headers
3	API Key Required	FALSE

Under URL Query String Parameters, choose Add query string and configure the following parameter:

	Setting	Value
1	Name	q
2	Required	Yes

DEPLOY THE API AND CONFIGURE A STAGE

THE API GATEWAY CONSOLE LETS YOU DEPLOY AN API BY CREATING A DEPLOYMENT AND ASSOCIATING IT WITH A NEW OR EXISTING STAGE.

- 1. Choose Actions and Deploy API.
- 2. For Deployment stage choose New Stage and name the stage opens earch-api-test.
- 3. Choose Deploy.
- 4. Configure the following settings in the stage editor, then choose Save Changes:

	Setting	Value
1	Enable throttling	Yes
2	Rate	1000
3	Burst	500

These settings configure an API that has only one method: a GET request to the endpoint root (https://some-id.execute-api.<Region>.amazonaws.com/search-es-api-test). The request requires a single parameter (q), the query string to search for. When called, the method passes the request to Lambda, which runs the opensearch-lambdafunction

Also, Enable CORS to access it from your frontend. You can enable CORS by following the steps mentioned here https://docs.aws.amazon.com/apigateway/latest/developerguide/how-to-cors-console.html

You can copy the created API gateway endpoint to search the metadata after you have uploaded images through UI/S3 bucket.

8) RUN APPLICATION AND UPLOAD THE IMAGES

You can upload the images and test this.

Metadata Management



Now you can successfully search the metadata through the URL copied in step 7 and it will show you the output in JSON. This can be integrated with the UI as well.

JSON Raw Data Headers Save Copy Collapse All Expand All 🗑 Filter JSON timed_out: false ▼ _shards: total: 2 skipped: failed: ▼ hits: value: relation: max_score: 0.28907335 ▼ hits: "news"
"_doc"
"public/GettyImages.jpg" _index: _type: _id: _score: 0.28907335 "{\"Score\": 0.999950476554871, \"Type\": \"MME\", \"BeginOffset\": 134, \"EndOffset\": 149}}"

"[{\"CelebrityMame\": \"Priyanka Chopra\", \"Confidence\": 9.6.1427307128966}]"

"[{\"TokenIdv\": 1, \"Text\": \"public\", \"BeginOffset\": 0, \"RedOffset\": 0, \"PartOfSpeech\": \"\"ap\": \"Ap\": \"Ap\": \"Ap\": \"BeginOffset\": 1, \"Text\": \"public\", \"BeginOffset\": 0, \"BeginOffset\" ₩ PII: ▼ SyntaxElements: