

MinIO Data Search (Backend)

Guide to run the code:

1. Clone from -> <https://github.com/Shao96/Celeron>
2. Make sure you have 'docker' and 'docker-compose' commands installed.
3. On the main branch from the ./Celeron directory run ->

docker-compose up -d

This will start the MinIO, RabbitMQ and ElasticSearch servers

4. Start the Event Receiver

- a. Under the directory -> ./EventReceiver, create a file called .env with the following content:

```
RM_QUEUE=my-queue
RM_HOST=localhost
RM_PORT=5672
RM_USERNAME=guest
RM_PASSWORD=guest
ELASTICSEARCH_HOST=localhost:9200
```

- b. Follow the steps from step 2 in the readme.md under the ./EventReceiver

5. Upload the datasets to MinIO and index them on elasticsearch.

- a. Under the directory -> ./FileUploader, create a file called .env with the following content:

```
MINIO_ENDPOINT=localhost:9000
MINIO_ACCESS_KEY=minio
MINIO_SECRET_KEY=miniostorage
RM_QUEUE=my-queue
RM_HOST=localhost
RM_PORT=5672
RM_USERNAME=guest
RM_PASSWORD=guest
ELASTICSEARCH_HOST=localhost:9200
```

- b. Follow the steps from step 2 in the ReadMe.md under the ./FileUploader
- c. Sample commands to upload and index the datasets:

```
python3 file_uploader.py --dir_path Covid19-dataset
```

```
python3 file_uploader.py --dir_path OralDB
```

6. **Start the File Searcher server.**

- a. Under the directory -> ./FileSearcher, create a file called .env with the following content:

ELASTICSEARCH_HOST = localhost:9200

SERVER_IP=localhost

SERVER_PORT=3000

- b. Follow the steps from step 2 in the ReadMe.md under the ./FileSearcher
c. Sample commands to

7. Go to localhost:3000/api-docs to access the swagger page of the api.

8. Sample requests on the Swagger page:

Sample request 1:

Search over Dataset names

GET

/datasets

Search for datasets

Parameters

Cancel

Name	Description
datasetName string (query)	Name of the dataset to search for
fromDate string(\$date) (query)	Starting date range for dataset creation
toDate string(\$date) (query)	Ending date range for dataset creation
minSize integer (query)	Minimum size of the dataset in bytes
maxSize integer (query)	Maximum size of the dataset in bytes
fileType string (query)	Comma-separated list of file types to filter by (e.g. jpg, csv, png)

Execute

Sample response 1 :

Responses

Curl

```
curl -X 'GET' \
  http://localhost:3000/datasets?datasetName=oral&fromDate=2023-03-28&toDate=2023-03-31&fileType=jpg' \
  -H 'accept: application/json'
```

Request URL

http://localhost:3000/datasets?datasetName=oral&fromDate=2023-03-28&toDate=2023-03-31&fileType=jpg

Server response

Code

Details

200

Response body

```
{
  "index": "datasets",
  "id": "PgFMNYcBYXnrXf2jSm0J",
  "score": 1,
  "source": {
    "name": "OralDB",
    "bucket": "oraldB",
    "date": "2023-03-31",
    "size": 16041666,
    "filetype": "jpg"
  }
}
```

Download

Response headers

```
access-control-allow-origin: *
content-length: 162
content-type: application/json; charset=utf-8
date: Fri, 31 Mar 2023 04:27:52 GMT
etag: W/"a2-Xm/YoFlUmzNeo8LVcSLp08oe+fe"
x-powered-by: Express
```

Responses

Code

Description

Links

Sample request 2 :

Search inside Datasets

GET

/search/{datasetName}

Search for data in dataset.

Parameters

Cancel

Name	Description
datasetName required (path)	OralDB
query (query)	Query string to search with nuance
eq_filters (query)	Key value pairs separated with '&'. For eg: color=blue&shape=round eq_filters
fields (query)	Comma separated list of fields on which you want to search. fields
pagination (query)	Page to return (expects the page number) 1
pageSize (query)	Number of items returned per page. 10

Execute

Clear

Sample response 2:

responses

Curl

curl -X 'GET' \n'http://localhost:3000/search/OralDB?query=nuance&pagination=1&pageSize=10' \n-H 'accept: */*'

Request URL

http://localhost:3000/search/OralDB?query=nuance&pagination=1&pageSize=10

Server response

Code

Details

200

Response body

```
{
  "total": {
    "value": 145,
    "relation": "eq"
  },
  "pagination": "1",
  "pageSize": "10",
  "results": [
    {
      "Spectral camera": " Nuance EX (CRI, Inc., USA).",
      "Illumination": " Thorlabs OSL2 halogen light source (100% power) + Thorlabs Ring Illuminator.",
      "Reference sample used": " Matt diffuse gray sample (\\\"Matt Diff Grey\\\", Ceram Research, Ltd., UK).",
      "Blood pressure and pulse": " -/-, -.",
      "Blood oxygen saturation": " -.%",
      "name": "ff6bc35d_top_inpainted_GDBICP_similarity.jpg",
      "bucket": "oraldb",
      "path": "OralDB/images/ff6bc35d_top_inpainted_GDBICP_similarity.jpg"
    },
    {
      "Spectral camera": " Nuance EX (CRI, Inc., USA).",
      "Illumination": " Thorlabs OSL2 halogen light source (100% power) + Thorlabs Ring Illuminator.",
      "Objective lens": " Navitar 5mm f/2.8.",
      "Aperture": " f/8.",
      "Reference sample used": " Matt diffuse gray sample (\\\"Matt Diff Grey\\\", Ceram Research, Ltd., UK).",
      "Blood pressure and pulse": " -/-, -.",
      "Blood oxygen saturation": " -.%",
      "name": "616b9dea_wideangle_GDBICP_similarity_2.jpg",
      "bucket": "oraldb",
      "path": "OralDB/images/616b9dea_wideangle_GDBICP_similarity_2.jpg"
    }
  ]
}
```

Response headers

```
access-control-allow-origin: *
connection: keep-alive
content-length: 5150
content-type: application/json; charset=utf-8
date: Fri, 31 Mar 2023 04:29:18 GMT
etag: W/"141e-x6V1z8q7Hu86d21D2PjeG2cX0X8"
keep-alive: timeout=5
x-powered-by: Express
```

Responses