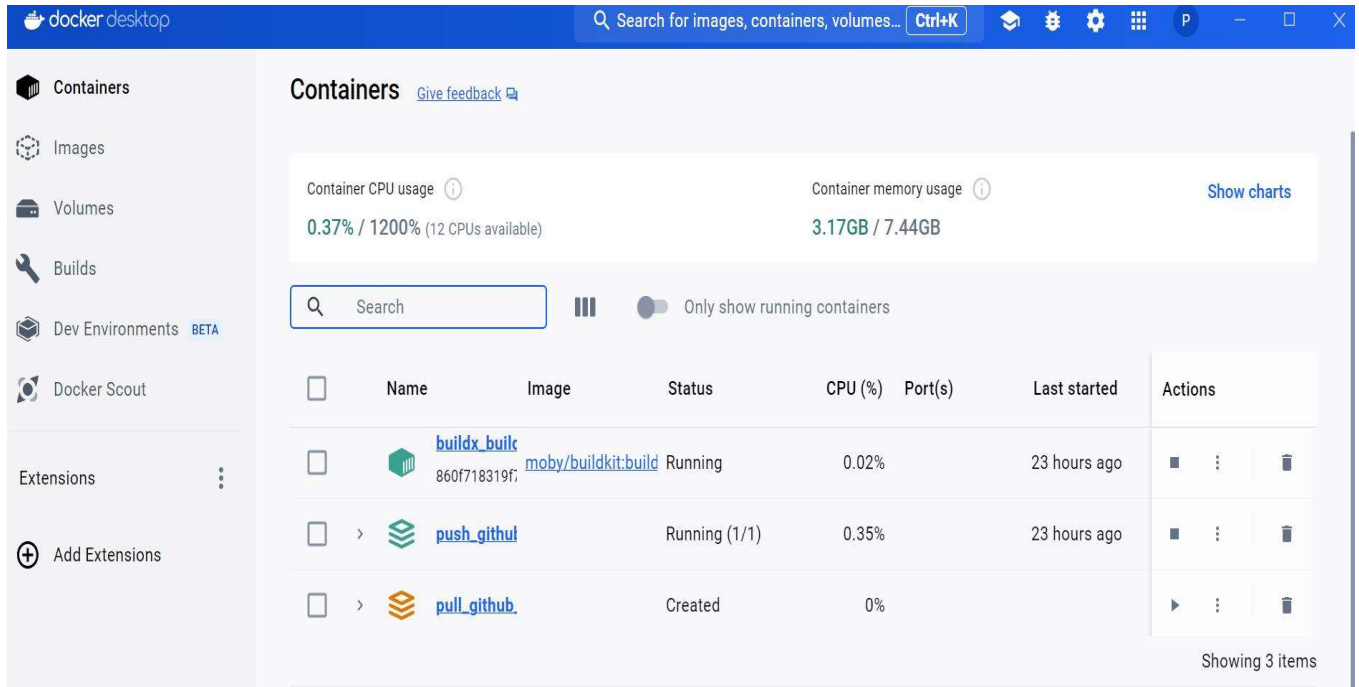


ASSIGNMENT – 5 – PART 2

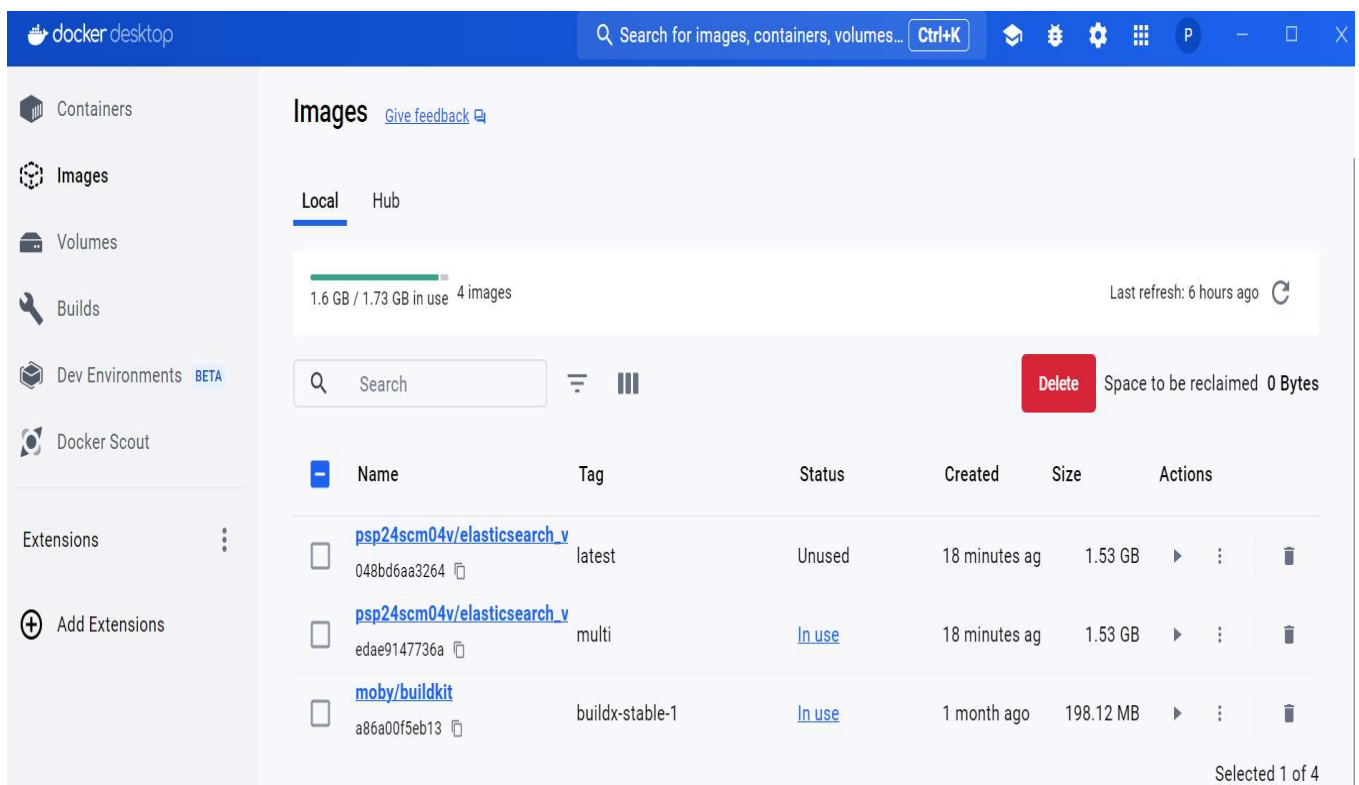
PAVITRA SAI VEGIRAJU
A20525304

Docker Website and Desktop App Screenshot:



The screenshot shows the Docker Desktop interface with the 'Containers' tab selected. The left sidebar contains navigation options: Containers, Images, Volumes, Builds, Dev Environments (BETA), Docker Scout, Extensions, and Add Extensions. The main area displays container statistics: CPU usage at 0.37% / 1200% (12 CPUs available) and memory usage at 3.17GB / 7.44GB. A search bar and a toggle for 'Only show running containers' are present. A table lists three containers: 'buildx_build' (Running), 'push_github' (Running 1/1), and 'pull_github' (Created). Each container row includes a checkbox, name, image, status, CPU usage, ports, last started time, and actions (stop, restart, delete). The bottom right indicates 'Showing 3 items'.

	Name	Image	Status	CPU (%)	Port(s)	Last started	Actions
<input type="checkbox"/>	buildx_build	moby/buildkit:build	Running	0.02%		23 hours ago	Stop, Restart, Delete
<input type="checkbox"/>	> push_github		Running (1/1)	0.35%		23 hours ago	Stop, Restart, Delete
<input type="checkbox"/>	> pull_github		Created	0%			Start, Restart, Delete



The screenshot shows the Docker Desktop interface with the 'Images' tab selected. The left sidebar is the same as the previous screenshot. The main area displays image statistics: 1.6 GB / 1.73 GB in use, 4 images, and a last refresh time of 6 hours ago. A search bar and a 'Delete' button are present. A table lists three images: 'psp24scm04v/elasticsearch_v' (Unused), 'psp24scm04v/elasticsearch_v' (In use), and 'moby/buildkit' (In use). Each image row includes a checkbox, name, tag, status, created time, size, and actions (refresh, restart, delete). The bottom right indicates 'Selected 1 of 4'.

	Name	Tag	Status	Created	Size	Actions
<input type="checkbox"/>	psp24scm04v/elasticsearch_v	latest	Unused	18 minutes ag	1.53 GB	Refresh, Restart, Delete
<input type="checkbox"/>	psp24scm04v/elasticsearch_v	multi	In use	18 minutes ag	1.53 GB	Refresh, Restart, Delete
<input type="checkbox"/>	moby/buildkit	buildx-stable-1	In use	1 month ago	198.12 MB	Refresh, Restart, Delete

Push_Github_Issues:

```
[9]: #Sample Issue
pprint(issues[2])

{'_body': '          ****\r\n'
         '\r\n'
         'Originally posted by @sarbazvatanatan in '
         'https://github.com/openai/openai-cookbook/issues/1103#issuecomment-2063022189_\r\n'
         ',
         '_closedAt': '2024-04-18T05:22:50Z',
         '_createdAt': '2024-04-18T05:21:58Z',
         '_issueNumber': '1159',
         '_repo': 'openai-cookbook',
         '_state': 'closed',
         '_title': '****',
         '_type': 'issue'}
```

```
[ ]:
```

```
[10]: #Number of Issues in the given timeframe
pprint(len(issues))

1994
```

```
[14]: ## Embedding creation using openAI of Github Issues.
```

```
from openai import OpenAI
from tqdm import tqdm
import time

# Initialize OpenAI client with API key
client = OpenAI(api_key="sk-proj-qvAcF7Ph208m9ZGhebNUT381bkFJIAbrvFPh70bYnHv572p0")

Issue_embeddings = []

# Batch size for processing data
batch_size = 500

# Initialize data structure for storing text
data = [
    [], # Titles
]
count=0;
# Embed and insert in batches
for i in tqdm(range(0, len(df_Issues))):
    title = str(df_Issues.iloc[i]['_title']).replace("\n", "") or ''
    body = str(df_Issues.iloc[i]['_body']).replace("\n", "") or ''

    # Merge 'repository name', 'title' and 'body' of the Github Issue
    combined_text = f"Repository:{owner}/{repo} Issue Title:{title} Issue Body:{body}"
    data[0].append(combined_text)
    if len(data[0]) % batch_size == 0:
        print("Embedding batch...")

        embeddings_batch = embed(data[0])
        Issue_embeddings.extend(embeddings_batch)
        data = [[]]
        print("Waiting for 1 minute before the next batch...")

        time.sleep(60)

# Embed the remaining data if any
if len(data[0]) != 0:
    embeddings_rem = embed(data[0])
    print(len(embeddings_rem))
    Issue_embeddings.extend(embeddings_rem)
```

```
0%|          | 0/1994 [00:00<?, ?it/s]
Embedding batch...
Waiting for 1 minute before the next batch...
25%|██████    | 500/1994 [01:03<03:10, 7.85it/s]
Embedding batch...
Waiting for 1 minute before the next batch...
50%|██████████| 1000/1994 [02:05<02:04, 7.96it/s]
Embedding batch...
Waiting for 1 minute before the next batch...
100%|██████████| 1994/1994 [03:07<00:00, 10.62it/s]
494
```

```
[17]: # Configure Elasticsearch connection
from elasticsearch import Elasticsearch, helpers
es = Elasticsearch(['http://localhost:9200'])
es.ping() #connection testing
```

```
[17]: True
```

```
[ ]:
```

```
[18]: #Index Mapping for githubissues
```

```
index_mapping = {
    "properties": {
        "Github_Issue_vector": {
            "type": "dense_vector",
            "dims": 1536,
            "index": "true",
            "similarity": "cosine"
        },
        "_type": {"type": "text"},
        "_repo": {"type": "text"},
        "_issueNumber": {"type": "long"},
        "_title": {"type": "text"},
        "_createdAt": {"type": "date"},
        "_closedAt": {"type": "date"},
        "_state": {"type": "text"},
        "_body": {"type": "text"}
    }
}

if es.indices.exists(index="github_issues"):
    es.indices.delete(index="github_issues")

es.indices.create(index="github_issues", body={"mappings": index_mapping})

[18]: ObjectApiResponse({'acknowledged': True, 'shards_acknowledged': True, 'index': 'github_issues'})
```

Pull_GitHub_Issues:

```
[3]: # To search documents

# Import Elasticsearch and helpers from elasticsearch

from elasticsearch import Elasticsearch, helpers
import json

es = Elasticsearch(['http://localhost:9200'])
es.ping() # connection testing
res = es.search(index="github_issues", body={"query": {"match_all": {}}})

# Extract the total hit count
print("Got %d Hits:" % res['hits']['total']['value'])

# Extract hits from the response
hits = res.get('hits', {})

# Pretty print the hits
print(json.dumps(hits, indent=2))
```

```
Got 1994 Hits:
{
  "total": {
    "value": 1994,
    "relation": "eq"
  },
  "max_score": 1.0,
  "hits": [
    {
      "_index": "github_issues",
      "_id": "mIne_o48VFULauTgpX8C",
      "_score": 1.0,
      "_source": {
        "_type": "issue",
        "_repo": "openai-cookbook",
        "issueNumber": "1161",
        "title": "Update registry.yaml",
        "createdAt": "2024-04-19T18:32:11Z",
```

```
[4]: hits['hits'][5]['_source'] # Choose an Issue from Issues in repository
```

```
[4]: {'_type': 'issue',
      '_repo': 'openai-cookbook',
      'issueNumber': '1156',
      'title': 'Codespace stunning space umbrella jjg7w544957fq7vq',
      'createdAt': '2024-04-18T04:23:21Z',
      'closedAt': '2024-12-31T00:36:30Z',
      'state': 'open',
      'body': '## Summary\r\n\r\nBriefly describe the changes and the goal of this PR. Make sure the PR title summarizes the changes effectively.\r\n\r\n## Motivation\r\n\r\nWhy are these changes necessary? How do they improve the cookbook?\r\n\r\n---\r\n\r\n## For new content\r\n\r\nWhen contributing new content, read through our [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md), and mark the following action items as completed:\r\n\r\n- [ ] I have added a new entry in [registry.yaml](https://github.com/openai/openai-cookbook/blob/main/registry.yaml) (and, optionally, in [authors.yaml](https://github.com/openai/openai-cookbook/blob/main/authors.yaml)) so that my content renders on the cookbook website.\r\n- [x] I have conducted a self-review of my content based on the [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md#rubric):\r\n- [ ] Relevance: This content is related to building with OpenAI technologies and is useful to others.\r\n- [ ] Uniqueness: I have searched for related examples in the OpenAI Cookbook, and verified that my content offers new insights or unique information compared to existing documentation.\r\n- [ ] Spelling and Grammar: I have checked for spelling or grammatical mistakes.\r\n- [ ] Clarity: I have done a final read-through and verified that my submission is well-organized and easy to understand.\r\n- [ ] Correctness: The information I include is correct and all of my code executes successfully.\r\n- [ ] Completeness: I have explained everything fully, including all necessary references and citations.\r\n\r\nWe will rate each of these areas on a scale from 1 to 4, and will only accept contributions that score 3 or higher on all areas. Refer to our [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md) for more details.\r\n\r\n'}

```

```
[6]: # We get Issues that are similar to the Issue chosen and sorted based on score
# First result is the most relevant and has the score 1.0

for response in responses:
    print(response)
    print()
```

```
registry.yaml](https://github.com/openai/openai-cookbook/blob/main/registry.yaml) (and, optionally, in [authors.yaml](https://github.com/openai/openai-cookbook/blob/main/authors.yaml)) so that my content renders on the cookbook website.\r\n- [x] I have conducted a self-review of my content based on the [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md#rubric):\r\n- [ ] Relevance: This content is related to building with OpenAI technologies and is useful to others.\r\n- [ ] Uniqueness: I have searched for related examples in the OpenAI Cookbook, and verified that my content offers new insights or unique information compared to existing documentation.\r\n- [ ] Spelling and Grammar: I have checked for spelling or grammatical mistakes.\r\n- [ ] Clarity: I have done a final read-through and verified that my submission is well-organized and easy to understand.\r\n- [ ] Correctness: The information I include is correct and all of my code executes successfully.\r\n- [ ] Completeness: I have explained everything fully, including all necessary references and citations.\r\n\r\nWe will rate each of these areas on a scale from 1 to 4, and will only accept contributions that score 3 or higher on all areas. Refer to our [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md) for more details.\r\n\r\n'}

```

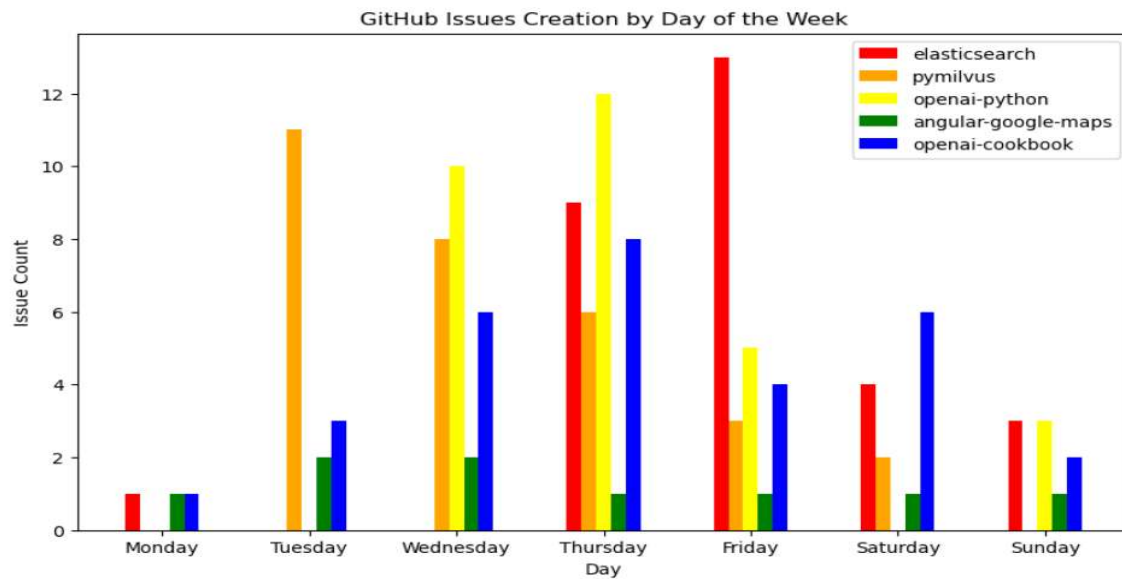
```
{ '_score': 0.96706605, '_type': 'issue', '_repo': 'openai-cookbook', 'issueNumber': '1155', 'title': 'Initial commit', 'createdAt': '2024-04-17T14:37:36Z', 'closedAt': '2024-04-18T16:18:18Z', 'body': '## Summary\r\n\r\nBriefly describe the changes and the goal of this PR. Make sure the PR title summarizes the changes effectively.\r\n\r\n## Motivation\r\n\r\nWhy are these changes necessary? How do they improve the cookbook?\r\n\r\n---\r\n\r\n## For new content\r\n\r\nWhen contributing new content, read through our [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md), and mark the following action items as completed:\r\n\r\n- [ ] I have added a new entry in [registry.yaml](https://github.com/openai/openai-cookbook/blob/main/registry.yaml) (and, optionally, in [authors.yaml](https://github.com/openai/openai-cookbook/blob/main/authors.yaml)) so that my content renders on the cookbook website.\r\n- [x] I have conducted a self-review of my content based on the [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md#rubric):\r\n- [ ] Relevance: This content is related to building with OpenAI technologies and is useful to others.\r\n- [ ] Uniqueness: I have searched for related examples in the OpenAI Cookbook, and verified that my content offers new insights or unique information compared to existing documentation.\r\n- [ ] Spelling and Grammar: I have checked for spelling or grammatical mistakes.\r\n- [ ] Clarity: I have done a final read-through and verified that my submission is well-organized and easy to understand.\r\n- [ ] Correctness: The information I include is correct and all of my code executes successfully.\r\n- [ ] Completeness: I have explained everything fully, including all necessary references and citations.\r\n\r\nWe will rate each of these areas on a scale from 1 to 4, and will only accept contributions that score 3 or higher on all areas. Refer to our [contribution guidelines](https://github.com/openai/openai-cookbook/blob/main/CONTRIBUTING.md) for more details.\r\n\r\n'}

```


Analytics Script:

Requirement – 1 Output:

```
ax.set_xlabel("Day")
ax.set_ylabel("Issue Count")
plt.show()
display_weekly_issue_distribution(weekly_data)
```



Requirement – 2 Output:

```
Repository: openai-cookbook
Issue 1: Codespace stunning space umbrella jfg7w544957fq7vq - Score: 5.1836243
Issue 2: Initial commit - Score: 5.1586986
Issue 3: H w - Score: 5.1584526
Issue 4: Evals conclusion - Score: 5.1428755
Issue 5: Eval Registry - Score: 5.1417246
Issue 6: Update registry.yaml - Score: 4.1836243
Issue 7: [FEATURE] - Score: 4.1836243
Issue 8: **** - Score: 4.1836243
Issue 9: [[C519aebd-8614-4b45-8afa-e8e6fac2ef07](https://github.com/openai/openai-cookbook/assets/166670647/707f943a-73a1-49d3-83e8-a1191e255d8a) - Score: 4.1836243
Issue 10: !!20240417-164047](https://github.com/openai/openai-cookbook/assets/166670647/2db64bdf-698d-4a63-9a15-cccd69977877) - Score: 4.1836243

Repository: elasticsearch
Issue 1: Codespace stunning space umbrella jfg7w544957fq7vq - Score: 1.0
Issue 2: Initial commit - Score: 0.96786605
Issue 3: H w - Score: 0.96682817
Issue 4: Evals conclusion - Score: 0.95845085
Issue 5: Eval Registry - Score: 0.95810014
Issue 6: Fix monitoring-es-mb.json - Score: 0.08074688
Issue 7: Document behaviour of 'SubscribableListener' on rejection - Score: 0.08074688
Issue 8: Muting DownsamplerActionIF testRollupNonIndex - Score: 0.08074688
Issue 9: [C] AzureBlobStoreRepositoryTests testRequestStats failing - Score: 0.08074688
Issue 10: Muting some DivTests failing tests - Score: 0.08074688

Repository: openai-python
Issue 1: chore(types): include discriminators in unions - Score: 4.5225544
Issue 2: chore(internal): improve deserialisation of discriminated unions - Score: 4.5225544
Issue 3: feat(cli): add support for fine_tuning.jobs - Score: 4.5225544
Issue 4: chore: export NOT_GIVEN sentinel value - Score: 4.5225544
Issue 5: chore(internal): support parsing Annotated types - Score: 4.5225544
Issue 6: chore(internal): bump pyright - Score: 4.5225544
Issue 7: Incompatible: with_streaming_response "stream_to_file" "response_format='aac'" - Score: 4.5225544
Issue 8: chore(internal): add core support for deserializing into number response - Score: 4.5225544
Issue 9: fix(streaming): improve error messages - Score: 4.5225544
Issue 10: Messages with non-latin languages texts returns server errors (500 status code) - Score: 4.5225544

Repository: pymilvus
Issue 1: Codespace stunning space umbrella jfg7w544957fq7vq - Score: 1.0
Issue 2: Initial commit - Score: 0.96786605
Issue 3: H w - Score: 0.96682817
Issue 4: Evals conclusion - Score: 0.95845085
Issue 5: Eval Registry - Score: 0.95810014

Repository: angular-google-maps
Issue 1: Codespace stunning space umbrella jfg7w544957fq7vq - Score: 1.0
Issue 2: Initial commit - Score: 0.96786605
Issue 3: H w - Score: 0.96682817
Issue 4: Evals conclusion - Score: 0.95845085
Issue 5: Eval Registry - Score: 0.95810014
```

Issue URL: <https://api.github.com/repos/openai/openai-cookbook/issues/1161>

Similar Issue URL: <https://api.github.com/repos/openai/openai-cookbook/issues/1140>, Similarity Score: 0.19410538199503427

Similar Issue URL: <https://api.github.com/repos/openai/openai-cookbook/issues/1131>, Similarity Score: 0.1328772404372487

Similar Issue URL: <https://api.github.com/repos/openai/openai-cookbook/issues/1147>, Similarity Score: 0.12383576830835173

Similar Issue URL: <https://api.github.com/repos/openai/openai-cookbook/issues/1137>, Similarity Score: 0.12091448989540268

Similar Issue URL: <https://api.github.com/repos/openai/openai-cookbook/issues/1154>, Similarity Score: 0.11774072238252697

Analytics Script:

```
Issue URL: https://api.github.com/repos/elastic/elasticsearch/issues/107666
Similar Issue URL: https://api.github.com/repos/elastic/elasticsearch/issues/107661, Similarity Score: 0.13183719996084844
Similar Issue URL: https://api.github.com/repos/elastic/elasticsearch/issues/107662, Similarity Score: 0.11312517057906255
Similar Issue URL: https://api.github.com/repos/elastic/elasticsearch/issues/107645, Similarity Score: 0.07735961618775425
Similar Issue URL: https://api.github.com/repos/elastic/elasticsearch/issues/107646, Similarity Score: 0.06989010310364346
Similar Issue URL: https://api.github.com/repos/elastic/elasticsearch/issues/107664, Similarity Score: 0.06892534656363378

Issue URL: https://api.github.com/repos/openai/openai-python/issues/1348
Similar Issue URL: https://api.github.com/repos/openai/openai-python/issues/1342, Similarity Score: 0.6183330369057484
Similar Issue URL: https://api.github.com/repos/openai/openai-python/issues/1335, Similarity Score: 0.23364063972074936
Similar Issue URL: https://api.github.com/repos/openai/openai-python/issues/1339, Similarity Score: 0.2011644028411152
Similar Issue URL: https://api.github.com/repos/openai/openai-python/issues/1349, Similarity Score: 0.1760216764266032
Similar Issue URL: https://api.github.com/repos/openai/openai-python/issues/1325, Similarity Score: 0.1660989129261239

Issue URL: https://api.github.com/repos/milvus-io/pymilvus/issues/2040
Similar Issue URL: https://api.github.com/repos/milvus-io/pymilvus/issues/2024, Similarity Score: 0.2328811991921522
Similar Issue URL: https://api.github.com/repos/milvus-io/pymilvus/issues/2041, Similarity Score: 0.17026325416464042
Similar Issue URL: https://api.github.com/repos/milvus-io/pymilvus/issues/2033, Similarity Score: 0.14959849300484573
Similar Issue URL: https://api.github.com/repos/milvus-io/pymilvus/issues/2044, Similarity Score: 0.12536005050174517
Similar Issue URL: https://api.github.com/repos/milvus-io/pymilvus/issues/2032, Similarity Score: 0.11402518217225054

Issue URL: https://api.github.com/repos/sebholstein/angular-google-maps/issues/2034
Similar Issue URL: https://api.github.com/repos/sebholstein/angular-google-maps/issues/2047, Similarity Score: 0.2682738041880502
Similar Issue URL: https://api.github.com/repos/sebholstein/angular-google-maps/issues/2044, Similarity Score: 0.22716729870668176
Similar Issue URL: https://api.github.com/repos/sebholstein/angular-google-maps/issues/1946, Similarity Score: 0.21283289964784025
Similar Issue URL: https://api.github.com/repos/sebholstein/angular-google-maps/issues/2046, Similarity Score: 0.19660872807692284
Similar Issue URL: https://api.github.com/repos/sebholstein/angular-google-maps/issues/1890, Similarity Score: 0.10127174084703525
```

LOCAL HOST 9200 WEBSITE:

```
{
  "_source": {
    "_type": "issue",
    "_repo": "openai-cookbook",
    "issueNumber": "1161",
    "title": "Update registry.yaml",
    "createdAt": "2024-04-19T18:32:11Z",
    "closedAt": "2024-04-19T19:05:01Z",
    "state": "closed",
    "body": "Forgot to update registry.yaml with modified filename.\r\n",
    "GitHub_Issue_vector": [
      0.0046083759516477585,
      0.006052379962056875,
      0.004484012257307768,
      -0.016844408586621284,
      -0.009555298835039139,
      0.022371696308255196,
      -0.025549886748194695,
      -0.0058243791572749615,
      -0.0026479163207113743,

```

[16]: