Elasticsearch

Exercícios

Escreva um script para importar os dados tags.csv de ml-latest-small em um novo índice "tags"

Nossa Solução

```python
import csv
from collections import deque
import elasticsearch
from elasticsearch import helpers

def readMovies():
    csvfile = open('ml-latest-small/movies.csv', 'r')

    reader = csv.DictReader( csvfile )

    titleLookup = {}

    for movie in reader:
            titleLookup[movie['movieId']] = movie['title']

    return titleLookup

def readTags():
    csvfile = open('ml-latest-small/tags.csv', 'r')

    titleLookup = readMovies()

    reader = csv.DictReader( csvfile )
    for line in reader:
        tag = {}
        tag['user_id'] = int(line['userId'])
        tag['movie_id'] = int(line['movieId'])
        tag['title'] = titleLookup[line['movieId']]
        tag['tag'] = line['tag']
        tag['timestamp'] = int(line['timestamp'])
        yield tag


es = elasticsearch.Elasticsearch()

es.indices.delete(index="tags",ignore=404)
deque(helpers.parallel_bulk(es,readTags(),index="tags"), maxlen=0)
es.indices.refresh()
```