

Um script mais simples

Bibliotecas clientes gratuitas estão disponíveis na maioria das linguagens

- [Java](#) tem um cliente mantido pela elastic.co
- [Python](#) tem um package elasticsearch
- [elasticsearch-ruby](#)
- Várias opções para [scala](#)
- Módulo elasticsearch.pm para [perl](#)

- [Você não precisa lidar com JSON](#)

```
es = elasticsearch.Elasticsearch()
```

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

Script Completo

```
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 readRatings():
    csvfile = open('ml-latest-small/ratings.csv', 'r')

    titleLookup = readMovies()

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

es = elasticsearch.Elasticsearch()

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



