

# Connecting to a NoSQL Server with Python

## Elastic Search

In [1]:

```
from elasticsearch import Elasticsearch
from elasticsearch_dsl import Search
import certifi
import json
import csv
import pandas as pd
from tqdm import tqdm
```

In [2]:

```
##### Parameters #####
index_name = 'index_name'
type_name = 'type_name'
csv_file = "base_ids.csv"
json_file = "results.json"
#####
```

In [ ]:

```
# Connecting to Elastic Search
try:
    es = Elasticsearch(
        ['server.us-east-1.aws.found.io'],
        http_auth=('user_login', 'user_pass'),
        port=port, #9243
        use_ssl=True,
        verify_certs=True,
        ca_certs=certifi.where(),
    )
    print ("\nConnected to ElasticSearch"), es.info()

except Exception as ex:
    print ("\nError to connect to ElasticSearch:"), ex
```

In [ ]:

```
# Get ids that was already downloaded
df = pd.read_csv(csv_file)
base = list(df.Ids.unique())
print("\n")
print("File '%s' imported" % csv_file)
print("Number of rows: "+str(len(base)))

print("\n")
```

In [ ]:

```
# Get and compare the existing Ids at Elastic Search

print("Starting to get Ids from server")
s = Search(using=es, index=index_name, doc_type=type_name)
s = s.source([])
ids_total = [h.meta.id for h in s.scan()]

print("\n")
print("ElasticSearch has %s Ids" % len(ids_total))

ids=list(set(ids_total)-set(base))
print('\n')
print("Total of new rows: "+str(len(ids)))

# You can Limit the n ids
#n=50000
#ids=ids[0:n]
#print(len(ids))
#res= es.get(index=index_name,doc_type=type_name,id=i))
print("\n")
print("Starting to get the new rows")
```

In [ ]:

```
# Capture information for each new Id
json_antigo = json.load(open(json_file))

data=[]
for i in tqdm(ids):

    # print(i)
    # print("\n")
    data.append(es.get(index=index_name,doc_type=type_name,id=i))

print("\n")
print("Retrieved %s logs" % len(data))

print("\n")
#print(data[0])
```

In [ ]:

```
# Update Local JSON file
with open(json_file, 'w') as file:
    json.dump(data+json_antigo, file)
print("Ficheiro '%s' atualizado com sucesso!" % json_file)
print("\n")
```

In [ ]:

```
# Update base_ids with the new ids

base_ids = open(csv_file, "a")

for i in ids:
    base_ids.write(str(i)+"\n")
base_ids.close()
print("File was '%s' successfully updated!" % csv_file)
print("\n")
```