

lab2MF

October 16, 2020

```
[1]: #Lab2
     #Marek Fudaliński

[2]: from elasticsearch import Elasticsearch

[3]: es = Elasticsearch(HOST="http://localhost", PORT=9200)
     es = Elasticsearch()
     index = "my-index-000001"

[4]: es.indices.delete(index=index)

[4]: {'acknowledged': True}

[5]: #Defin analyzer
     es.indices.create(
         index = index,
         body = {
             "settings": {
                 "analysis": {
                     "analyzer": {
                         "polish": {
                             "type": "custom",
                             "tokenizer": "standard",
                             "filter": [
                                 "lowercase",
                                 "synonym",
                                 "morfologik_stem"
                             ]
                         }
                     }
                 },
                 "filter": {
                     "synonym": {
                         "type": "synonym",
                         "tokenizer": "standard",
                         "synonyms": [
                             "kpk, kodeks postępowania karnego",
                             "kpc, kodeks postępowania cywilnego",
```

```

        "kk, kodeks karny",
        "kc, kodeks cywilny"
    ]
    }
}
},
"mappings": {
    "dynamic": "strict",
    "properties": {
        "bill": {
            "type": "text",
            "analyzer": "polish",
        },
        "file_name": {"type": "text"}
    }
}
)

```

```
[5]: {'acknowledged': True, 'shards_acknowledged': True, 'index': 'my-index-000001'}
```

```
[6]: from os import listdir
from os.path import isfile, join
billsPath = './bills/'
onlyBills = [f for f in listdir(billsPath) if isfile(join(billsPath, f))]
print(len(onlyBills))
```

1180

```
[7]: for i in range(len(onlyBills)):
    f = open(billsPath+onlyBills[i], "r", encoding='utf-8')
    text = f.read()
    body = {"bill": text, "file_name": onlyBills[i]}
    f.close()
    es.index(index=index, id=i, body=body)
```

```
[8]: # 6. Determine the number of legislative acts containing the word ustawa (in
    ↪ any form)
body = {
    "query": {
        "match": {
            "bill": "ustawa"
        }
    }
}
res = es.count(index=index, body=body)
```

```
res
```

```
[8]: {'count': 1179,  
      '_shards': {'total': 1, 'successful': 1, 'skipped': 0, 'failed': 0}}
```

```
[9]: # 7 Determine the number of legislative acts containing the words kodeks  
      ↳postępowania cywilnego in the specified order, but in an any inflection form.  
body = {  
    "query": {  
        "match_phrase": {  
            "bill": {  
                "query": "kodeks postępowania cywilnego",  
                "slop": 0  
            }  
        }  
    }  
}  
res = es.count(index=index, body=body)  
res
```

```
[9]: {'count': 100,  
      '_shards': {'total': 1, 'successful': 1, 'skipped': 0, 'failed': 0}}
```

```
[10]: # 8 Determine the number of legislative acts containing the words wchodzi w  
       ↳życie (in any form) allowing for up to 2 additional words in the searched  
       ↳phrase.  
body = {  
    "query": {  
        "match_phrase": {  
            "bill": {  
                "query": "wchodzi w życie",  
                "slop": 2  
            }  
        }  
    }  
}  
res = es.count(index=index, body=body)  
res
```

```
[10]: {'count': 1175,  
      '_shards': {'total': 1, 'successful': 1, 'skipped': 0, 'failed': 0}}
```

```
[11]: #9 Determine the 10 documents that are the most relevant for the phrase  
      ↳konstytucja.  
body = {  
    "size":10,  
    "_source": ['file_name'],
```

```

        "query": {
            "match": {
                "bill": "konstytucja"
            },
        },
    }
}

res = es.search(index=index, body=body)
res
for i in range(len(res['hits']['hits'])):
    print(str(i + 1) + ". " + res['hits']['hits'][i]['_source']['file_name'] + "
↳: " + str(res['hits']['hits'][i]['_score']))

```

```

1. 1997_629.txt : 6.87158
2. 2000_443.txt : 6.667763
3. 1997_604.txt : 6.637804
4. 1996_350.txt : 6.631998
5. 1997_642.txt : 6.2563715
6. 2001_23.txt : 6.0651627
7. 1996_199.txt : 5.935479
8. 1999_688.txt : 5.8563156
9. 1997_681.txt : 5.473405
10. 2001_1082.txt : 5.473405

```

[12]: #10 Print the excerpts containing the word konstytucja (up to three excerpts
↳per document) from the previous task.

```

body = {
    "query": {
        "match": {
            "bill": {
                "query": "konstytucja"
            }
        },
    },
    "size": 10,
    "_source": [ "file_name" ],
    "highlight": {
        "number_of_fragments": 3,
        "fragment_size": 100,
        "fields": {
            "bill": {}
        }
    }
}

res = es.search(index=index, body=body)
res['hits']['hits']
for i in range(len(res['hits']['hits'])):

```

```

print(str(i + 1) + ') ' + res['hits']['hits'][i]['_source']['file_name'])
for j in range(len(res['hits']['hits'][i]['highlight']['bill'])):
    print("\t" + str(j+1) + ". " +
↳res['hits']['hits'][i]['highlight']['bill'][j])

```

1) 1997_629.txt

1. o zmianie ustawy konstytucyjnej o trybie przygotowania i uchwalenia Konstytucji Rzeczypospolitej
2. W ustawie konstytucyjnej z dnia 23 kwietnia 1992 r. o trybie przygotowania i uchwalenia Konstytucji
3. Do zgłoszenia projektu Konstytucji załącza się wykaz obywateli popierających zgłoszenie

2) 2000_443.txt

1. umowy międzynarodowej i nie wypełnia przesłanek określonych w art. 89 ust. 1 lub art. 90 Konstytucji
2. międzynarodowej lub załącznika nie wypełnia przesłanek określonych w art. 89 ust. 1 lub art. 90 Konstytucji
3. co do zasadności wyboru trybu ratyfikacji umowy międzynarodowej, o którym mowa w art. 89 ust. 2 Konstytucji

3) 1997_604.txt

1. Jeżeli Trybunał Konstytucyjny wyda orzeczenie o sprzeczności celów partii politycznej z Konstytucją
2. Jeżeli Trybunał Konstytucyjny wyda orzeczenie o sprzeczności z Konstytucją celów lub działalności
3. Ciężar udowodnienia niezgodności z Konstytucją spoczywa na wnioskodawcy, który w tym

4) 1996_350.txt

1. Za naruszenie Konstytucji lub ustawy, w związku z zajmowanym stanowiskiem lub w zakresie
2. W zakresie określonym w art. 107 Konstytucji odpowiedzialność przed Trybunałem Stanu ponoszą
3. Członkowie Rady Ministrów ponoszą odpowiedzialność przed Trybunałem Stanu za naruszenie Konstytucji

5) 1997_642.txt

1. wnioskami o:
 - 1) stwierdzenie zgodności ustaw i umów międzynarodowych z Konstytucją
 2. stwierdzenie zgodności przepisów prawa wydawanych przez centralne organy państwowe, z Konstytucją
 3. ratyfikowanymi umowami międzynarodowymi i ustawami,

- 4) stwierdzenie zgodności z Konstytucją
- 6) 2001_23.txt
1. W Dzienniku Ustaw Rzeczypospolitej Polskiej, zwanym dalej "Dziennikiem Ustaw", ogłasza się:
- 1) Konstytucję
2. akty prawne dotyczące:
- 1) stanu wojny i zawarcia pokoju,
- 2) referendum zatwierdzającego zmianę Konstytucji
3. ministra, któremu Sejm wyraził wotum nieufności,
- h) powoływania lub odwoływania na określone w Konstytucji
- 7) 1996_199.txt
1. 2c i art. 9-11 ustawy konstytucyjnej z dnia 23 kwietnia 1992 r. o trybie przygotowania i uchwalenia Konstytucji
2. Prezydent Rzeczypospolitej Polskiej zarządza poddanie Konstytucji pod referendum w trybie określonym
3. Przyjęcie w referendum Konstytucji następuje wówczas, gdy opowiedziała się za nią większość biorących
- 8) 1999_688.txt
1. postępowania w sprawie wykonywania inicjatywy ustawodawczej przez obywateli, o której mowa w art. 118 ust. 2 Konstytucji
2. Projekt ustawy nie może dotyczyć spraw, dla których Konstytucja Rzeczypospolitej Polskiej zastrzega wyłączną
3. Projekt ustawy wniesiony do Marszałka Sejmu powinien odpowiadać wymogom zawartym w Konstytucji i Regulaminie
- 9) 1997_681.txt
1. Rzecznik Praw Dziecka, zwany dalej Rzecznikiem, stoi na straży praw dziecka określonych w Konstytucji
2. uroczyście, że przy wykonywaniu powierzonych mi obowiązków Rzecznika Praw Dziecka dochowam wierności Konstytucji
- 10) 2001_1082.txt
1. Krajowa Rada Sądownictwa, zwana dalej "Radą", realizuje cele określone w Konstytucji Rzeczypospolitej
2.) podejmuje uchwały w sprawach wystąpienia do Trybunału Konstytucyjnego o zbadanie zgodności z Konstytucją

[]: