



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА КОМПЬЮТЕРНЫЕ СИСТЕМЫ И СЕТИ (ИУ6)

НАПРАВЛЕНИЕ ПОДГОТОВКИ 09.04.01 Информатика и вычислительная техника

МАГИСТЕРСКАЯ ПРОГРАММА 09.04.01/07 Интеллектуальные системы анализа,
обработки и интерпретации больших данных

О Т Ч Е Т

по лабораторной работе № 5

Название: Работа с документной NoSQL БД Elasticsearch и реализация
процесса data science

Дисциплина: Технология параллельных систем баз данных

Студент

ИУ6-12М
(Группа)

(Подпись, дата)

С.В. Астахов
(И.О. Фамилия)

Преподаватель

(Подпись, дата)

(И.О. Фамилия)

Москва, 2023

Введение

1. Цель работы: студент должен изучить модель представления данных и способы работы с документной NoSQL базой данных Elasticsearch. Получить навыки инсталляции, индексации и поиска в Elasticsearch, реализовать этапы процесса Data Science, протестировать работу с большим объёмом данных.

Ход выполнения

2. Установка и запуск ES.

Для упрощения установки воспользуемся докер-контейнером с ES, пробросив порт 9200 в сеть основной машины (рисунок 1).


Name	Image	Status	CPU (%)	Port(s)
 es4 b4e97354d53b	bitnami/elasticsearch:latest	Running	0.98%	9200:9200 Show all ports (2)

Рисунок 1 — контейнер с ES

Проверим работу ES с помощью GET-запроса к url “http://localhost:9200/?pretty” (рисунок 2).

```
{
  "name" : "b4e97354d53b",
  "cluster_name" : "elasticsearch",
  "cluster_uuid" : "z9602ra5Q_WZa0JkKqFBHQ",
  "version" : {
    "number" : "8.10.4",
    "build_flavor" : "default",
    "build_type" : "tar",
    "build_hash" : "b4a62ac808e886ff032700c391f45f1408b2538c",
    "build_date" : "2023-10-11T22:04:35.506990650Z",
    "build_snapshot" : false,
    "lucene_version" : "9.7.0",
    "minimum_wire_compatibility_version" : "7.17.0",
    "minimum_index_compatibility_version" : "7.0.0"
  },
  "tagline" : "You Know, for Search"
}
```

Рисунок 2 — информация о кластере ES

Так же установим пакет для работы с ES в python (рисунок 3).

```
PS C:\Users\sergey.astakhov> pip install Elasticsearch
WARNING: Ignoring invalid distribution -atplotlib (c:\users\sergey.astakhov\pyenv\pyenv-win\versions\3.9.0\lib\site-packages)
Requirement already satisfied: Elasticsearch in c:\users\sergey.astakhov\pyenv\pyenv-win\versions\3.9.0\lib\site-packages (8.10.1)
```

Рисунок 3 — установленный python-пакет

3. Реализация процесса Data Science на примере диагностики болезней

Перенесем и проиндексируем данные на сервере ES с помощью программы, код которой приведен в листинге 1.

Листинг 1 — программа для индексации данных

```
import json
from elasticsearch import Elasticsearch

# подключение к серверу
client = Elasticsearch([{"host": "127.0.0.1", "scheme": "http", "port": 9200}])

# создание индекса
indexName = "medical"
client.indices.create(index=indexName)

# описание схемы
diseaseMapping = {
    "properties": {
        "name": {
            "type": "text",
            "fielddata": True
        },
        "title": {
            "type": "text",
            "fielddata": True
        },
        "fulltext": {
            "type": "text",
            "fielddata": True
        }
    }
}

# отправка схемы
client.indices.put_mapping(index=indexName,
                           # doc_type="diseases",
                           # include_type_name="true",
                           body=diseaseMapping)

# считывание данных из файла
with open('data.json', 'r') as f:
    dataStore = json.load(f)

# отправка данных в ES
for data in dataStore:
    try:
        client.index(
            index=data["index"],
            # doc_type=data["doc_type"],
```

```
        id=data["id"],
        body=data["body"]
    )
except Exception as e:
    print(e)
```

Напишем свой первый запрос в ES (листинг 2, рисунок 4). Будем запрашивать первые 3 подходящих записи, ища запрашиваемые слова во всех полях документа с различными весами. Результат будет отсортирован по весу (“score”), который зависит от того, в каких полях располагались искомые слова.

Листинг 2 — запрос для поиска по ключевым словам

```
indexName = "medical"
docType = "diseases"
searchFrom = 0
searchSize = 3

# 1
searchBody = {
    "_source": ["name"], # возвращать только названия болезней
    "query": {
        "simple_query_string": {
            "query": '+fatigue+fever+joint pain',    # список симптомов
            "fields": ["fulltext", "title^5", "name^10"] # веса
        }
    }
}
result = client.search(index=indexName, body=searchBody, from_=searchFrom, size=searchSize)
```

```
result['hits']

{'total': {'value': 33, 'relation': 'eq'},
 'max_score': 116.35371,
 'hits': [{'_index': 'medical',
            '_id': 'Chronic fatigue syndrome',
            '_score': 116.35371,
            '_source': {'name': 'Chronic fatigue syndrome'}},
          {'_index': 'medical',
            '_id': 'Chronic fatigue immune dysfunction syndrome',
            '_score': 99.84779,
            '_source': {'name': 'Chronic fatigue immune dysfunction syndrome'}},
          {'_index': 'medical',
            '_id': 'Encephalomyelitis, myalgic',
            '_score': 39.252914,
            '_source': {'name': 'Encephalomyelitis, myalgic'}}]}
```

Рисунок 4 — результат выполнения запроса

Первые 3 найденных заболевания: “Синдром хронической усталости”, “Синдром хронической усталости и иммунной дисфункции”, “Энцефаломиелит миалгический”.

Запрос 2 — как изменятся результаты поиска болезней, если в "query" добавить новый симптом rash (сыпь)? Параметры запроса и ответ на него приведены в листинге 3 и на рисунке 5 соответственно.

Листинг 3 — запрос для поиска по ключевым словам

```
indexName = "medical"
docType = "diseases"
searchFrom = 0
searchSize = 3

# 3
searchBody = {
  "_source": ["name"], # возвращать только названия болезней
  "query": {
    "simple_query_string": {
      "query": '+fatigue+fever'+joint pain'+rash',      # список симптомов
      "fields": ["fulltext", "title^5", "name^10"]      # веса
    }
  }
}
result = client.search(index=indexName, body=searchBody, from_=searchFrom, size=searchSize)
```

```
result['hits']

{'total': {'value': 16, 'relation': 'eq'},
 'max_score': 13.86566,
 'hits': [{ '_index': 'medical',
  '_id': 'Chikungunya',
  '_score': 13.86566,
  '_source': {'name': 'Chikungunya'}},
 {'_index': 'medical',
  '_id': 'Cutaneous lupus erythematosus',
  '_score': 13.500596,
  '_source': {'name': 'Cutaneous lupus erythematosus'}},
 {'_index': 'medical',
  '_id': 'Lupus erythematosus',
  '_score': 13.500596,
  '_source': {'name': 'Lupus erythematosus'}}]}
```

Рисунок 5 — результат выполнения запроса

Как видно, уменьшилось число результатов (value) и изменились их веса (score). Первые 3 соответствующие запросу болезни: “Чикунгунья”, “Кожная красная волчанка” и “Красная волчанка”.

Запрос 3 — Как изменятся результаты поиска болезней, если в "query" добавить ещё один симптом sensitivity to sunlight (чувствительность к солнечному свету)?

Параметры запроса и ответ на него приведены в листинге 4 и на рисунке 6 соответственно.

Листинг 4 — запрос для поиска по ключевым словам

```
# 3
searchBody = {
  "_source": ["name"],
  "query": {
    "simple_query_string": {
      "query": '+fatigue+fever+joint pain+rash+sensitivity to sunlight',
      "fields": ["fulltext", "title^5", "name^10"]
    }
  }
}
result = client.search(index=indexName, body=searchBody, from_=searchFrom, size=searchSize)
```

```
result['hits']

{'total': {'value': 2, 'relation': 'eq'},
 'max_score': 20.476803,
 'hits': [{ '_index': 'medical',
             '_id': 'Cutaneous lupus erythematosus',
             '_score': 20.476803,
             '_source': {'name': 'Cutaneous lupus erythematosus'}},
          { '_index': 'medical',
             '_id': 'Lupus erythematosus',
             '_score': 20.476803,
             '_source': {'name': 'Lupus erythematosus'}}]}
```

Рисунок 6 — результат выполнения запроса

Как видно, уменьшилось число результатов (value) и изменились их веса (score). Запросу соответствуют всего 2 болезни: “Кожная красная волчанка”, “Красная волчанка”.

Запрос 4 позволяет выявить симптомы болезней, в имени которых встречается слово `diabetes` (диабет). Параметры запроса и ответ на него приведены в листинге 5 и на рисунках 7-8 соответственно.

Листинг 5 — запрос с агрегацией

```
#4
searchBody={
  "_source": ["name"],
  "query": {
    "query_string": {
      "query": "name: diabetes"
    }
  },
  "aggregations": {
    "DiseaseKeywords": {
      "significant_terms": { "field": "fulltext", "size": 10 } # получить частые слова из описания
    }
  }
}
result=client.search(index=indexName, body=searchBody, from_ = searchFrom, size=searchSize)
```

```
result['hits']
```

```
{'total': {'value': 31, 'relation': 'eq'},
 'max_score': 5.9901114,
 'hits': [{ '_index': 'medical',
             '_id': 'Diabetes insipidus',
             '_score': 5.9901114,
             '_source': {'name': 'Diabetes insipidus'} },
           { '_index': 'medical',
             '_id': 'Diabetes mellitus',
             '_score': 5.9901114,
             '_source': {'name': 'Diabetes mellitus'} },
           { '_index': 'medical',
             '_id': 'Phosphate diabetes',
             '_score': 5.9901114,
             '_source': {'name': 'Phosphate diabetes'} } ] }
```

Рисунок 7 — список искомых болезней


```

: result['aggregations']['DiseaseKeywords']
: {
  'doc_count': 31,
  'bg_count': 5540,
  'buckets': [
    {'key': 'siphon',
     'doc_count': 10,
     'score': 52.08494938984013,
     'bg_count': 11},
    {'key': 'diabainein',
     'doc_count': 9,
     'score': 51.59313215400625,
     'bg_count': 9},
    {'key': 'bainein',
     'doc_count': 9,
     'score': 51.59313215400625,
     'bg_count': 9},
    {'key': 'ndi', 'doc_count': 8, 'score': 45.86056191467221, 'bg_count': 8},
    {'key': 'passer', 'doc_count': 9, 'score': 42.159776747706, 'bg_count': 11},
    {'key': 'sapere',
     'doc_count': 7,
     'score': 40.127991675338194,
     'bg_count': 7},
    {'key': 'tasty', 'doc_count': 7, 'score': 40.127991675338194, 'bg_count': 7},
    {'key': 'sapidus',
     'doc_count': 7,
     'score': 40.127991675338194,
     'bg_count': 7},
    {'key': 'avp', 'doc_count': 8, 'score': 36.636836628511965, 'bg_count': 10},
  ]
}

```

Рисунок 8 — список ключевых слов

Какие 5 симптомов диабета являются наиболее важными?

“siphon”, “diabainein”, “bainein”, “ndi”, “passer”.

Сколько документов (названий болезней, связанных с диабетом) было проанализировано?

result['hits']['value'] = 31.

Сколько времени выполнялся запрос?

result['took'] = 4884.

Запрос 5 — выявить симптомы болезней, в имени которых встречаются слова diabetes (диабет) и insipidus (несахарный). Параметры запроса и ответ на него приведены в листинге 6 и на рисунках 9-10 соответственно.

Листинг 6 — запрос с агрегацией

```
#5
searchBody={
  "_source": ["name"],
  "query": {
    "query_string": {
      "query": "name: diabetes+insipidus"
    }
  },
  "aggregations": {
    "DiseaseKeywords": {
      "significant_terms": { "field": "fulltext", "size": 10 }
    }
  }
}
result=client.search(index=indexName, body=searchBody, from_ = searchFrom, size=searchSize)
```

```
result['hits']

{'total': {'value': 31, 'relation': 'eq'},
 'max_score': 13.25312,
 'hits': [{'_index': 'medical',
            '_id': 'Diabetes insipidus',
            '_score': 13.25312,
            '_source': {'name': 'Diabetes insipidus'}},
          {'_index': 'medical',
            '_id': 'Central diabetes insipidus',
            '_score': 11.448059,
            '_source': {'name': 'Central diabetes insipidus'}},
          {'_index': 'medical',
            '_id': 'Nephrogenic diabetes insipidus',
            '_score': 11.448059,
            '_source': {'name': 'Nephrogenic diabetes insipidus'}}]}
```

Рисунок 9 — список искомых болезней

```
result['aggregations']['DiseaseKeywords']

{'doc_count': 31,
 'bg_count': 5540,
 'buckets': [{'key': 'siphon',
               'doc_count': 10,
               'score': 52.08494938984013,
               'bg_count': 11},
             {'key': 'diabainein',
               'doc_count': 9,
               'score': 51.59313215400625,
               'bg_count': 9},
             {'key': 'bainein',
               'doc_count': 9,
               'score': 51.59313215400625,
               'bg_count': 9},
             {'key': 'ndi', 'doc_count': 8, 'score': 45.86056191467221, 'bg_count': 8},
             {'key': 'passer', 'doc_count': 9, 'score': 42.159776747706, 'bg_count': 11},
             {'key': 'sapere',
               'doc_count': 7,
               'score': 40.127991675338194,
               'bg_count': 7},
             {'key': 'tasty', 'doc_count': 7, 'score': 40.127991675338194, 'bg_count': 7},
             {'key': 'sapidus',
               'doc_count': 7,
               'score': 40.127991675338194,
               'bg_count': 7},
             {'key': 'avp', 'doc_count': 8, 'score': 36.636836628511965, 'bg_count': 10},
```

Рисунок 10 — список ключевых слов

Запрос 6 — выполните нечёткий поиск (fuzzy) для имени болезни diabetse (в названии болезни случайно поменяли местами две последние буквы). Параметры запроса и ответ на него приведены в листинге 7 и на рисунках 11-12 соответственно.

Листинг 7 — запрос на нечеткий поиск

```
#6
searchBody={
  "_source": ["name"],
  "query": {
    "fuzzy": {
      "name": { "value": "diabetse" }
    }
  },
  "aggregations": {
    "DiseaseKeywords": {
      "significant_terms": { "field": "fulltext", "size": 10 }
    }
  }
}
result=client.search(index=indexName, body=searchBody, from_ = searchFrom, size=searchSize)
```

```
result['hits']
```

```
{'total': {'value': 35, 'relation': 'eq'},
 'max_score': 5.241348,
 'hits': [{ '_index': 'medical',
             '_id': 'Diabetes insipidus',
             '_score': 5.241348,
             '_source': {'name': 'Diabetes insipidus'} },
           { '_index': 'medical',
             '_id': 'Diabetes mellitus',
             '_score': 5.241348,
             '_source': {'name': 'Diabetes mellitus'} },
           { '_index': 'medical',
             '_id': 'Phosphate diabetes',
             '_score': 5.241348,
             '_source': {'name': 'Phosphate diabetes'} } ] }
```

Рисунок 11 — список искомых болезней

```
result['aggregations']['DiseaseKeywords']
```

```
{'doc_count': 35,  
  'bg_count': 5540,  
  'buckets': [{ 'key': 'siphon',  
    'doc_count': 10,  
    'score': 40.82745825602968,  
    'bg_count': 11},  
    { 'key': 'diabainein',  
      'doc_count': 9,  
      'score': 40.444897959183656,  
      'bg_count': 9},  
    { 'key': 'bainein',  
      'doc_count': 9,  
      'score': 40.444897959183656,  
      'bg_count': 9},  
    { 'key': 'ndi', 'doc_count': 8, 'score': 35.95102040816327, 'bg_count': 8},  
    { 'key': 'passer',  
      'doc_count': 9,  
      'score': 33.044526901669755,  
      'bg_count': 11},  
    { 'key': 'sapere',  
      'doc_count': 7,  
      'score': 31.457142857142866,  
      'bg_count': 7},  
    { 'key': 'tasty', 'doc_count': 7, 'score': 31.457142857142866, 'bg_count': 7},
```

Рисунок 12 — список ключевых слов

Как видно, список болезней и ключевых слов не изменились, следовательно, нечеткий поиск работает корректно.

Запрос 7 является примером запроса с фильтрацией (листинг 8, рисунок 13).

Листинг 8 — запрос с фильтрацией

```
#7  
filterQuery = {  
  "_source": ["name", "title", "fulltext"],  
  "query": {  
    "constant_score": {  
      "filter": {  
        "term": {  
          "title": "headache"  
        }  
      }  
    }  
  }  
}  
  
result=client.search(index="medical", body=filterQuery, size=1)
```

```
result['hits']
```

```
{'total': {'value': 2, 'relation': 'eq'},
 'max_score': 1.0,
 'hits': [{'_index': 'medical',
            '_id': 'Cluster headache',
            '_score': 1.0,
            '_source': {'name': 'Cluster headache',
                        'title': 'Cluster headache',
                        'fulltext': 'Cluster headache (CH) is a neurological disorder characterized by recurrent severe headaches on one side of the head, typically around the eye. There is often accompanying eye watering, nasal congestion, or swelling around the eye on the affected side. These symptoms typically last 15 minutes to 3 hours. Attacks often occur in clusters which typically last for weeks or months and occasionally more than a year. The cause is unknown. Risk factors include a history of exposure to tobacco smoke and a family history of the condition. Exposures which may trigger attacks include alcohol and nitroglycerin. They are a primary headache disorder of the trigeminal autonomic cephalalgias type. Diagnosis is based on symptoms. Recommended management includes lifestyle changes such as avoiding potential triggers. Treatments for acute attacks include oxygen or a fast-acting triptan. Measures recommended to decrease the frequency of attacks include steroid injections, civamide, or verapamil. Nerve stimulation or surgery may occasionally be used if other measures are'}}
```

Рисунок 13 — результат запроса

Запрос вернул 2 статьи о головной боли. Начало первой статьи:

“Кластерная головная боль (СН) - это неврологическое расстройство, характеризующееся повторяющимися сильными головными болями с одной стороны головы, обычно вокруг глаза. Часто наблюдается слезотечение, заложенность носа или припухлость вокруг глаза на пораженной стороне. Эти симптомы обычно длятся от 15 минут до 3 часов. Атаки часто происходят группами, которые обычно длятся недели или месяцы, а иногда и больше года. Причина неизвестна...”

Запрос 8 сортирует возвращаемые результаты (листинг 9, рисунок 14).

Листинг 9 — запрос с сортировкой

```
sortQuery = {
  "_source": ["name", "title"],
  "sort": [{"title": "desc"}]
}
result=client.search(index=indexName,body=sortQuery,from_=searchFrom, size=searchSize)
```

```
result['hits']
```

```
{'total': {'value': 5540, 'relation': 'eq'},
 'max_score': None,
 'hits': [{'_index': 'medical',
            '_id': 'Shared psychotic disorder',
            '_score': None,
            '_source': {'name': 'Shared psychotic disorder', 'title': 'Folie à deux'},
            'sort': ['à']},
          {'_index': 'medical',
            '_id': 'Zygomycosis',
            '_score': None,
            '_source': {'name': 'Zygomycosis', 'title': 'Zygomycosis'},
            'sort': ['zygomycosis']},
          {'_index': 'medical',
            '_id': 'Weissenbacher-Zweymüller syndrome',
            '_score': None,
            '_source': {'name': 'Weissenbacher-Zweymüller syndrome',
                        'title': 'Weissenbacher-Zweymüller syndrome'},
            'sort': ['zweymüller']}]}
```

Рисунок 14 — результаты запроса

В результате будут отсортированы все (value = 5540) записи.

Запрос 9 производит извлечение документа по идентификатору (листинг 10, рисунок 15).

Листинг 10 — запрос на извлечение документа по идентификатору

```
idQuery = {
  "query": {
    "terms": {
      "_id": ["Stomatitis", "Alkaptonuria"]
    }
  }
}
result=client.search(index=indexName, body=idQuery, size=1)
```

```
result['hits']

{'total': {'value': 1, 'relation': 'eq'},
 'max_score': 1.0,
 'hits': [{'_index': 'medical',
  '_id': 'Stomatitis',
  '_score': 1.0,
  '_source': {'name': 'Stomatitis',
  'title': 'Stomatitis',
  'fulltext': 'Stomatitis is inflammation of the mouth and lips. It refers to any inflammatory process affecting the mucous membranes of the mouth and lips, with or without oral ulceration. In its widest meaning, stomatitis can have a multitude of different causes and a ppearances. Common causes include infections, nutritional deficiencies, allergic reactions, radiotherapy, and many others. \nwhen inflammation of the gums and the mouth generally presents itself, sometimes the term gingivostomatitis is used, though this is also sometimes used as a synonym for herpetic gingivostomatitis. \nThe term is derived from the Greek stoma (στόμα), meaning "mouth", and the suffix -itis (-ίτις), meaning "inflammation". \n\n Causes == \n\n == Nutritional deficiency == \nMalnutrition (improper dietary intake) or m
```

Рисунок 19 — результат запроса

Начало текста статьи: “Стоматит - это воспаление полости рта и губ. Это относится к любому воспалительному процессу, поражающему слизистые оболочки полости рта и губ, с изъязвлением полости рта или без него. В самом широком смысле стоматит может иметь множество различных причин и проявлений. Распространенные причины включают инфекции, дефицит питательных веществ, аллергические реакции, лучевую терапию и многие другие. Когда обычно проявляется воспаление десен и полости рта, иногда используется термин "гингивостоматит", хотя иногда он также используется как синоним герпетического гингивостоматита. Термин происходит от греческого stoma (στόμα), означающего "рот", и суффикса -itis (-ίτις), означающего "воспаление"...”

Написать и выполнить запрос 10 bool query, позволяющий найти описания болезней, удовлетворяющих следующим условиям:

1. "must": в поле "title" встречается слово "Chronic"
2. "filter": в тексте документов ("fulltext") встречается слово "Americans"
3. "should": в поле "title" встречается слово "leukemia" или "syndrome"
4. "must_not": в поле "name" не должно встречаться сочетание "Epidemic encephalomyelitis"
5. Минимальное число низкочастотных терминов (для should), которые должны присутствовать, равно 1 ("minimum_should_match": 1).

Текст и результат такого запроса приведены в листинге 10 и на рисунке 16.

Листинг 10 — логический запрос

```
boolQuery = {
  "_source": ["name", "title", "fulltext"],
  "query": {
    "bool": {

      "must" : {
        "term" : { "title" : "chronic" }
      },

      "filter": {
        "term" : { "fulltext" : "americans" }
      },

      "should" : [
        { "term" : { "fulltext" : "leukemia" } },
        { "term" : { "fulltext" : "syndrome" } }
      ],

      "must_not" : {
        "term" : { "name" : "epidemic+encephalomyelitis" }
      },

      "minimum_should_match": 1
    }
  }
}

result=client.search(index=indexName, body=boolQuery, size=1)
```



```
result['hits']

{'total': {'value': 8, 'relation': 'eq'},
 'max_score': 9.468292,
 'hits': [{'_index': 'medical',
            '_id': 'Chronic myelogenous leukemia',
            '_score': 9.468292,
            '_source': {'name': 'Chronic myelogenous leukemia',
                        'title': 'Chronic myelogenous leukemia',
                        'fulltext': 'Chronic myelogenous leukemia (CML), also known as chronic myeloid leukemia, is a cancer of the white blood cells. It is a form of leukemia characterized by the increased and unregulated growth of myeloid cells in the bone marrow and the accumulation of these cells in the blood. CML is a clonal bone marrow stem cell disorder in which a proliferation of mature granulocytes (neutrophils, eosinophils and basophils) and their precursors is found. It is a type of myeloproliferative neoplasm associated with a characteristic chromosomal translocation called the Philadelphia chromosome.\nCML is largely treated with targeted drugs called tyrosine-kinase inhibitors (TKIs) which have led to dramatically improved long-term survival rates since 2001. These drugs have revolutionized treatment of this disease and allow most patients to have a good quality of life when compared to the former chemotherapy drugs. In Western countries, CML accounts for 15–25% of all adult leukemias and 14% of leukemias overall (including the pediatric population, where CML is less common).\n\n== Signs and symptoms ==\nThe way CML presents depends on the stage of the disease at diagnosis as it has been known to sk
```

Рисунок 16 — результаты запроса

Запрос вернул 8 заболеваний (value = 8). Начало первой статьи: “Хронический миелолейкоз (ХМЛ), также известный как хронический миелоидный лейкоз, является раком белых кровяных телец. Это форма лейкоза, характеризующаяся усиленным и нерегулируемым ростом миелоидных клеток в костном мозге и накоплением этих клеток в крови. ХМЛ - это клональное заболевание стволовых клеток костного мозга, при котором обнаруживается пролиферация зрелых гранулоцитов (нейтрофилов, эозинофилов и базофилов) и их предшественников...”.

Вывод: в ходе лабораторной работы были изучены модель представления данных и способы работы с документной NoSQL базой данных Elasticsearch, получены навыки инсталляции, индексации и поиска в Elasticsearch.