

The background features a series of concentric circles in light gray, some solid and some dashed. A large, solid blue oval is positioned in the center-right of the frame. A thick, dark gray curved line sweeps from the bottom left towards the blue oval.

Elasticsearch

Lukáš Špalek

17.5.2023

Agenda

- Theory
 - What is Elasticsearch
 - Examples
 - Elasticsearch vs Relation database
- Full-text search
 - Match query
 - Match_phrase
 - Multi_match
 - Bool query



The image features a blue speech bubble with a white outline and a small tail pointing downwards. Inside the bubble, the word "Theory" is written in a white, sans-serif font. The background is a light gray with a pattern of concentric circles, some solid and some dashed, creating a subtle geometric design.

Theory



What is Elasticsearch?

- NoSql database
- Search with relevance scoring
- Powerfull full-text search engine
- Use as analytics tool
- Very good scalability

When to use it?

- Search with relevance scoring
- Full text search
- Analyse application logs and system metrics
- Send events to Elasticsearch

Used by:

- Netflix
- Ebay
- Walmart

Elasticsearch / ELK stack

- Kibana - analytics and visualisation for elastic = dashboard
 - Possible to define change detection / ML
- Logstash
 - Logs are handled as events
- X-Pack
 - Security
 - Monitoring
 - Alerting
 - Reporting
- Beats

Elasticsearch vs. Relation database

Elasticsearch Relational DB	
Index	Table
Document	Row
Field	Column

Example

```
{
  "_index": "my_news",
  "_type": "_doc",
  "_id": "F_b-wocBJcXkNfyRH2-V",
  "_score": 12.764705,
  "_source": {
    "date": "2015-05-17",
    "short_description": "It's a
great green electrified and luxurious
car",
    "@timestamp": "2015-05-
17T00:00:00.000+02:00",

    "link": "https://www.huffingtonpost.com
/entry/test-driving-2015-
porsche_b_6957302.html",
    "category": "GREEN",
    "headline": "Test Driving Porsche
Panamera Plugin Hybrid Electric Car",
    "authors": "Seth Leitman,
ContributorThe Green Living Guy"
  }
}
```

id	date	short_description	category	headline	authors	...
F_b-wocBJcXkNfyRH2-V	2015-05-17	It's a great green electrified and luxurious car	GREEN	Test Driving Porsche Panamera Plugin Hybrid Electric Car	Seth Leitman, ContributorThe Green Living Guy	
...	

The background features a series of concentric circles in light gray, some solid and some dashed, creating a ripple effect. A large blue speech bubble is centered on the page, pointing downwards.

Full-text search

What we use in examples

- Elastic cloud: [Free Elastic Cloud Trial \(30 days\)](#)
- Dataset: [News Headlines Dataset](#)
- Queries:
 - Match query
 - Match_phrase
 - Multi_match
 - Bool query

Full-text search

- Match query
- Match_phrase
- Multi_match
- Bool query

Match query

- Order and proximity do not matter
- Search in single field

```
GET my_news/_search
{
  "query": {
    "match": {
      "short_description": {
        "query": "infant car seats",
      }
    }
  }
}
```

Match_phrase query

- Order and proximity matter
- Search in single field

```
GET my_news/_search
{
  "query": {
    "match_phrase": {
      "short_description": {
        "query": "car seats"
      }
    }
  }
}
```

Multi_match query

- Order and proximity do not matter
- Search in multiple fields
- "type": "phrase_prefix"
- Tuning relevance by boosting

```
GET my_news/_search
{
  "query": {
    "multi_match": {
      "query": "Roger Federer",
      "fields": [
        "headline^2",
        "short_description",
        "authors"
      ]
    }
  }
}
```

Bool query

- Search with combination of queries
- Has 4 parts:
 - Must
 - Must_not
 - Should
 - Filter

```
GET name_of_index/_search
```

```
{  
  "query": {  
    "bool": {  
      "must": [  
        {One or more queries can be specified here. A document MUST match all  
of these queries to be considered as a hit.}  
      ],  
      "must_not": [  
        {A document must NOT match any of the queries specified here. If it  
does, it is excluded from the search results.}  
      ],  
      "should": [  
        {A document does not have to match any queries specified here.  
However, it if it does match, this document is given a higher score.}  
      ],  
      "filter": [  
        {These filters(queries) place documents in either yes or no category.  
Ones that fall into the yes category are included in the hits. }  
      ]  
    }  
  }  
}
```


Combined query

- Find all articles
 - contains "Roger"
 - which are NOT in "Sport" category
 - On top we would like to see articles from "COLLEGE" category
 - Which were created before 2018-01-01

Summary

Usage of Elasticsearch

- full-text search
- log processing / search
- data analytics

Discussed various types of full text search queries

- Match query
- Match_phrase query
- Multi_match
- Bool query



Questions?

Links

- [Free Elastic Cloud Trial \(30 days\)](#)
- <https://github.com/LisaHJung/Part-2-Understanding-the-relevance-of-your-search-with-Elasticsearch-and-Kibana->
- <https://github.com/LisaHJung/Part-3-Running-full-text-queries-and-combined-queries-with-Elasticsearch-and-Kibana>
- https://www.youtube.com/playlist?list=PL_mJOmq4zsHbcdoeAwNWuhEWwDARMMBta
- <https://reflectoring.io/spring-boot-elasticsearch/>
- <https://www.knowi.com/blog/what-is-elastic-search/>