Advanced Elasticsearch

This is Heavy tech post, I would like to make different points which are very important when you are using elastic search for your enterprise projects.

## Disable Doc\_Values when you don’t need them

This column-store is known as doc values, and is enabled by default. Doc values are created at index-time: when a field is indexed, Elasticsearch adds the tokens to the inverted index for search.

Doc values are used in several places in Elasticsearch:

-Sorting on a field

-Aggregations on a field

If you don’t need to sort or aggregate on a field, or access the field value from a script, you HAVE to disable doc values in order to save disk space.

PUT demo\_index

{

"mappings": {

"\_doc": {

"properties": {

"username": {

"type": "keyword"

},

"session\_id ": {

"type": "keyword",

"doc\_values": false

}

}

}

}

}

## Synonymous

The concept of Synonymous is very interesting, handy and enriches the document in order to be retrieved with different combination of terms.

For example:

USA -> United States, U.S.A

Sr. -> Senior

Dev -> Developer, Software Developer

…………………………………………………………………………..

Now Let’s imagine we would like to search “USA” and get all documents contains either United States or USA, in order to solve this problem we have few options

1. At search time; we will need to convert a query for USA into a query for USA OR “United States”.
2. Using the synonym token filter in index time, we replace USA with the two terms usa and United States, then at search time we can search those terms.

Both solutions are valid but the first will drive us to have multiple responsibilities at application level and in my opinion it is not good practice.

With the second option we have centralized search engine will manage everything for us, converting the word and make it available for search ☺

Create Index [Code]

PUT /vergnaty\_index

{

"settings": {

"analysis": {

"filter": {

"special\_synonyms": {

"type": "synonym",

"synonyms": [

"USA,United States",

"Sr.,Senior"

]

}

},

"analyzer": {

"my\_synonyms": {

"tokenizer": "standard",

"filter": [“lowercase”, "special\_synonyms"]

}

}

}

}

}

GET /vergnaty\_index/\_analyze

{

"analyzer" : "my\_synonyms",

"text" : "Sr. Software developer in USA"

}

(USA)

(United States)

(basketball)

(dream)

(team)

<https://www.elastic.co/guide/en/elasticsearch/guide/master/symbol-synonyms.html>