Google Style Rank Based Search for Investment Funds

Mr. Fakruddin Mohammed

Motivation

Searching (text based) for information has become essential part in the era of digital world. We all know how the search should work but the reality is implementing a good search engine is extremely challenging task. Therefore, in this short tutorial I wanted to explore the search technologies.

Problem Statement

Nowadays every business have a Digital Front End providing informational services to the end users. When the user is looking for information, the ability to give the most relevant information is key ingredient to the success of the business who purely rely on digital data services. There are number of ways one can add text based search features to the application such as Relational Database queries or No-SQL based data queries or Remotely hosted CGI scripts. But the issues with these approaches is that it doesn't use contextual information from what the user is asking for, it doesn't give the results sorted by relevance and the free text based search implementation is extremely difficult task, often simplified or compromised to make implementation easier therefore, it doesn't retrieve the results accurately what the user is actually looking for. Therefore, in this short article the study explores the alternate ways to implement the search engine features using the investment funds data (mutual funds, exchange traded funds and hedge funds data). The reason for choosing the financial instruments data is due to the fact that a given investment product can be described in using thousands of attributes, therefore, it challenges the implementation. The data is web scraped from the publicly available websites and where data is not available, initialised with random data and normalised to the OpenFunds standards.
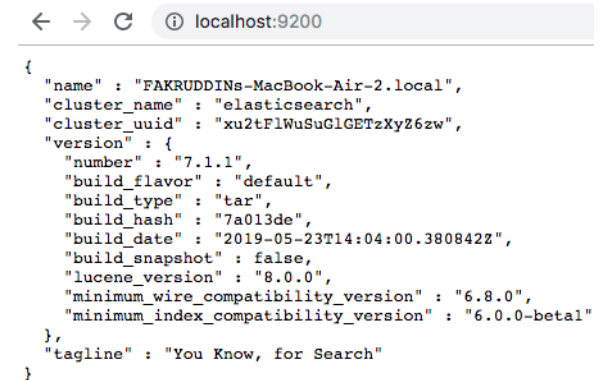
Architecture
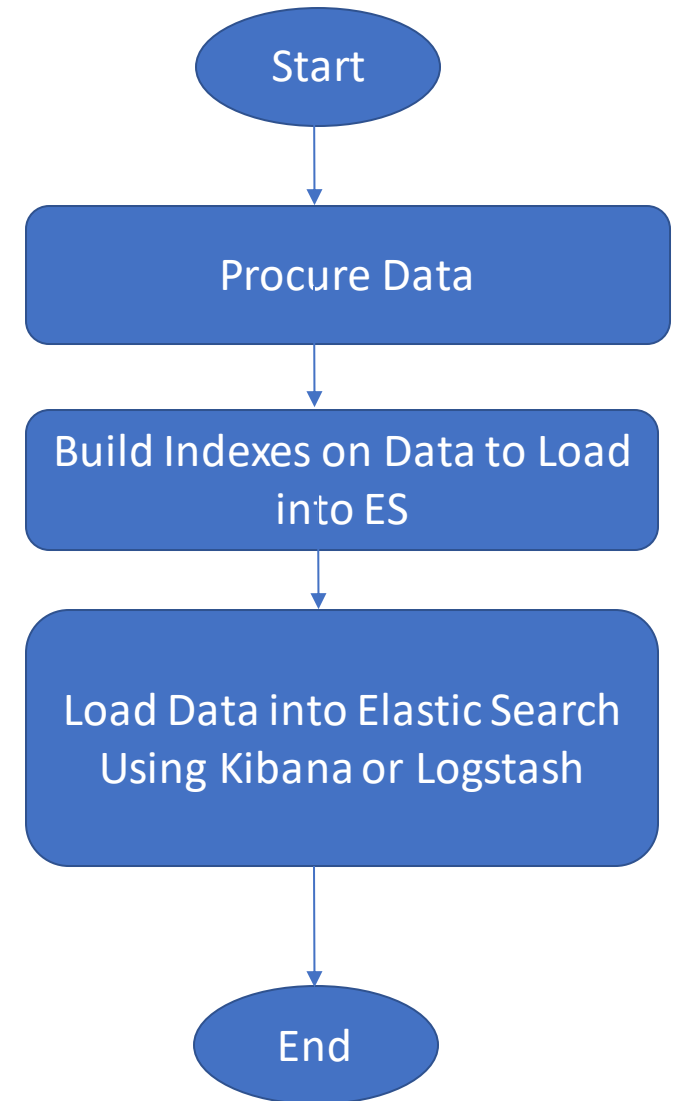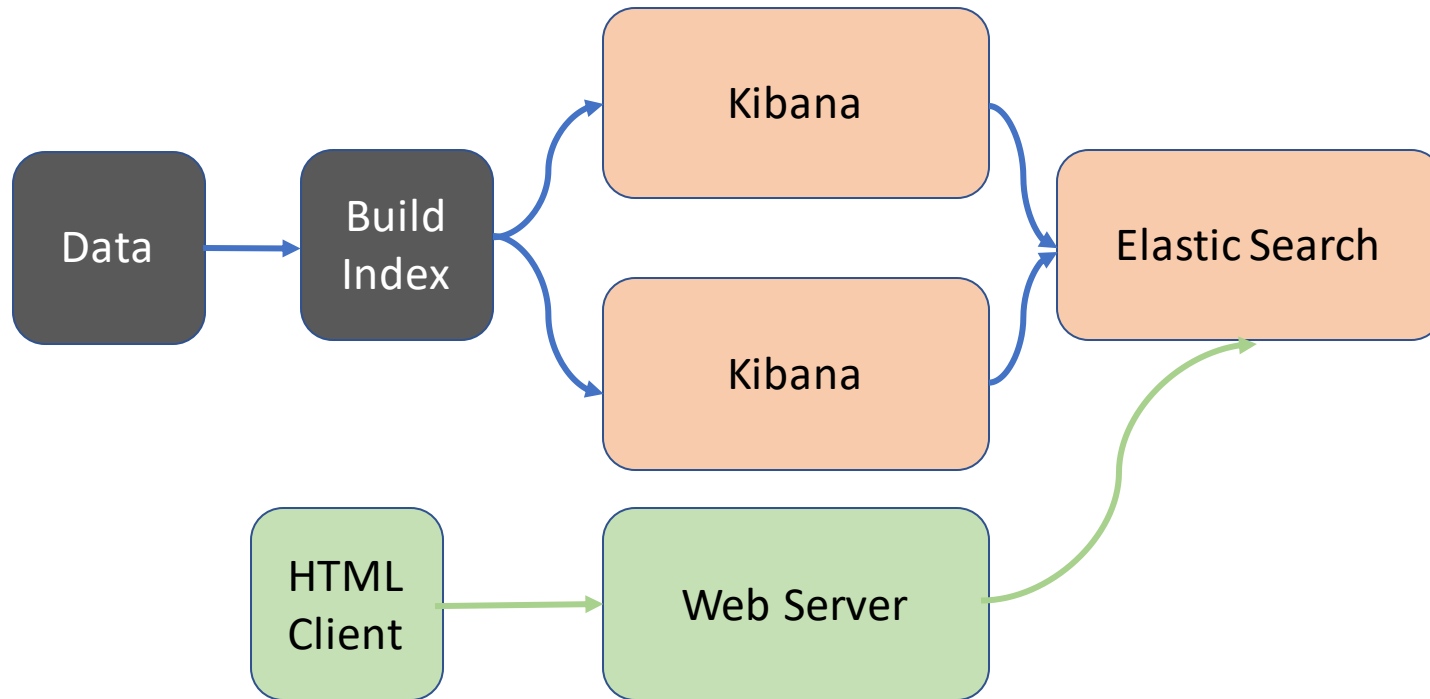
Software Installtion

ElasticSearch

- Download elasticsearch from the website: https://www.elastic.co/products/elasticsearch
- Extract the contents to a suitable folder
- Start the Elasticsearch by running the script the bin folder: elasticsearch.sh or elasticsearch.bat
- Check elasticsearch is running by going to the URL: http://localhost:9200/
- The output should look like the below

```
←  →  C      ⓘ localhost:9200

{
  "name" : "FAKRUDDINs-MacBook-Air-2.local",
  "cluster_name" : "elasticsearch",
  "cluster_uuid" : "xu2tFlWuSuGlGETzXyZ6zw",
  "version" : {
    "number" : "7.1.1",
    "build_flavor" : "default",
    "build_type" : "tar",
    "build_hash" : "7a013de",
    "build_date" : "2019-05-23T14:04:00.380842Z",
    "build_snapshot" : false,
    "lucene_version" : "8.0.0",
    "minimum_wire_compatibility_version" : "6.8.0",
    "minimum_index_compatibility_version" : "6.0.0-beta1"
  },
  "tagline" : "You Know, for Search"
}
```

Kibana

- Download elasticsearch from the website: https://www.elastic.co/downloads/kibana
- Extract the contents to a suitable folder
- Start the Kibana by running the script the bin folder: kibana.sh or kibana.bat
- Check Kibana is running by going to the URL: http://localhost:5601/

Data

Collate the Data

- Collate the data from the source systems

ISIN	Parent Company	Fund Currency	Fund Name
LUXXXXXXXXXX	Dummy Name - Bond Index Fund	EUR	Dummy Fund - Bond Index Fund EUR P

Convert the Above Data to Json

```
{
"ISIN": "LUXXXXXXXXXX",
"Parent Company": "Dummy Name - Bond Index Fund",
"Fund Currency": "EUR",
"Fund Name": "Dummy Fund - Bond Index Fund EUR P"
}
```

Load Data into Elasticsearch

Create Index

- Go to Kibana console (http://localhost:5601/app/kibana#/dev_tools/console?_g=())
- Create a new index called "dummy_product:

PUT dummy_product
- The output on the console should be

```
{
 "acknowledged" : true,
 "shards_acknowledged" : true,
 "index" : "dummy_product"
}
```

Load Data into Elasticsearch

Load Data

- Go to Kibana console (http://localhost:5601/app/kibana#/dev_tools/console?_g=())
- Load the first document into the elasticsearch

POST /dummy_product/default/_bulk
{"index": { "_id":1}}
{"ISIN": "LUXXXXXXXXXX","Parent Company":"Dummy Name - Bond Index Fund","Fund Currency": "EUR","Fund Name": "Dummy Fund - Bond Index Fund EUR P"}

- The output of the above command should be

This is manually added document index id

This is the document being loaded into the elasticsearch; on load the elasticsearch creates indexes for every word in the document

```
1   #! Deprecation: [types removal] Specifying
2 ▾ {
3       "took" : 805,
4       "errors" : false,
5 ▾     "items" : [
6 ▾       {
7 ▾         "index" : {
8             "_index" : "dummy_product",
9             "_type" : "default",
10            "_id" : "1",
11            "_version" : 1,
12            "result" : "created",
13 ▾          "_shards" : {
14              "total" : 2,
15              "successful" : 1,
16              "failed" : 0
17 ▾          },
18            "_seq_no" : 0,
19            "_primary_term" : 1,
20            "status" : 201
21 ▾        }
22 ▾      }
23 ▾    ]
24 ▾ }
```

Query Data from Elasticsearch

Query Data

- Go to Kibana console (http://localhost:5601/app/kibana#/dev_tools/console?_g=())
- Run the below query to check the document is loaded and indexes are built

```
GET dummy_product/default/_search
{
 "query":{
  "match_all":{}
 }
}
```

- The output of the above command
Is shown in the image
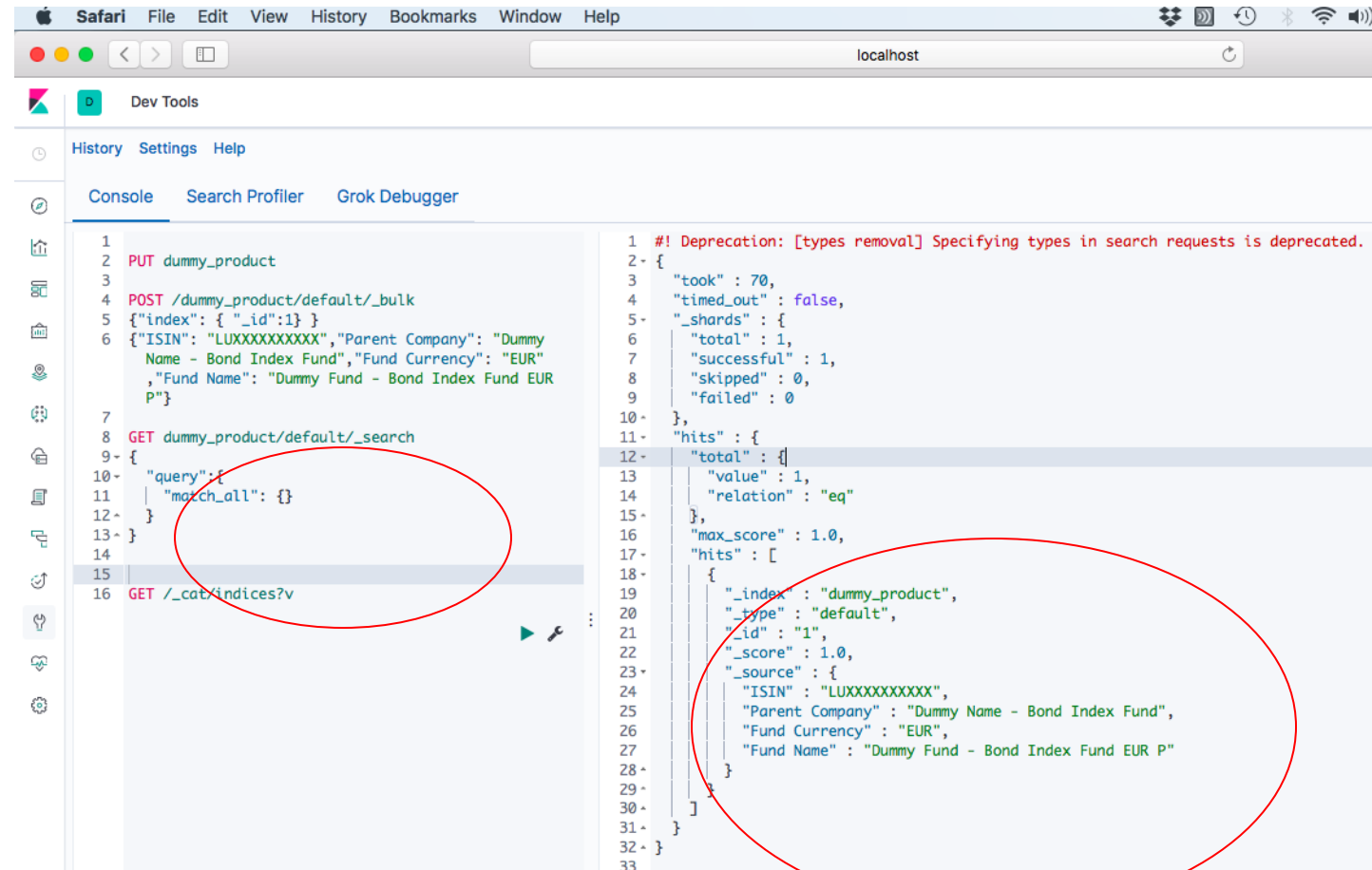
Bulk Load Data into Elasticsearch

Bulk Load All Documents

- For bulk loading documents we can use logstash or Kibana console.
- Using Kibana console, we can bulk load documents in one go by appending all documetns as shown below

```
POST /dummy_product/default/_bulk
{"index": { "_id":1}}
{"ISIN": "LUXXXXXXXXX","Parent Company": "Dummy Name - Bond Index Fund","Fund Currency": "EUR","Fund Name":
"Dummy Fund - Bond Index Fund EUR P"}
{"index": { "_id":2}}
{"ISIN": "LUXXXXXXXXX","Parent Company BBB": "Dummy Name - Bond Index Fund","Fund Currency": "GBP","Fund Name":
"Dummy Fund - Equity Index Fund GBP A"}
```

```
"items" : [
  {
    "index" : {
      "_index" : "dummy_product",
      "_type" : "default",
      "_id" : "1",
      "_version" : 2,
      "result" : "updated",
      "_shards" : {
        "total" : 2,
        "successful" : 1,
        "failed" : 0
      },
      "_seq_no" : 1,
      "_primary_term" : 1,
      "status" : 200
    }
  },
  {
    "index" : {
      "_index" : "dummy_product",
      "_type" : "default",
      "_id" : "2",
      "_version" : 1,
      "result" : "created",
      "_shards" : {
        "total" : 2,
        "successful" : 1,
        "failed" : 0
      },
      "_seq_no" : 2,
      "_primary_term" : 1,
      "status" : 201
    }
  }
]
```

HTML Client Demo

HTML Client Demo

Google Style Searching for Investment Funds

(With Thousands of Data Attributes)

Search for:
EUR EONIA Equity or Bond Funds registered in Switzerland with fee 2% and share class P

jpmorgan income growth fund EUR   Search

○ Parse   ○ Simple

Showing 186 results, took:4937ms

Rank	ISIN	Share Class	Matched Fields
8.391924	LU0697242484	undefined	Fund Promoter Name : JPMorgan Bank Luxembourg SA Fund Advisor Name : JPMorgan Asset Management UK Limited Fund Administrator Name : Morgan Bank Luxembourg S.A. Custodian Bank Name : Morgan Bank Luxembourg S.A. Full Share Class Name : JPMorgan Funds - JPM Global Multi Strategy Income A (div) - EUR
8.391924	LU0697242567	undefined	Fund Promoter Name : JPMorgan Bank Luxembourg SA Fund Advisor Name : JPMorgan Asset Management UK Limited Fund Administrator Name : Morgan Bank Luxembourg S.A. Custodian Bank Name : Morgan Bank Luxembourg S.A. Full Share Class Name : JPMorgan Funds - JPM Global Multi Strategy Income D (acc) - EUR

HTML Client Demo

Google Style Searching for Investment Funds

(With Thousands of Data Attributes)

Search for:
JPMorgan high yield with return 2% or more and with 35% in MSCI

jpmorgan index fund high yield 2% 35% MSC| | Search |

○ Parse ○ Simple

Showing 186 results, took:12471ms

Rank	ISIN	Share Class	Matched Fields
8.071962	LU0119174026	undefined	**PRIIP Category :** 2 **Benchmark :** BofA ML European High Yield Constrained **Full Share Class Name :** Aberdeen Global - Select Euro High Yield Bond Fund A MInc EUR
8.071962	LU0231473439	undefined	**PRIIP Category :** 2 **Benchmark :** BofA ML European High Yield Constrained **Full Share Class Name :** Aberdeen Global - Select Euro High Yield Bond Fund I Acc EUR
8.071962	LU0837982205	undefined	**PRIIP Category :** 2 **Benchmark :** BofA ML European High Yield Constrained **Full Share Class Name :** Aberdeen Global - Select Euro High Yield Bond Fund X MInc EUR
4.669957	LU0697242484	undefined	**Fund Promoter Name :** JPMorgan Bank Luxembourg SA **Fund Advisor Name :** JPMorgan Asset Management UK Limited **Benchmark :** 40% BLOOMBERG BARCLAYS US HIGH YIELD 2% ISSUER CAP INDEX (TOTAL RETURN GROSS) HEDGED TO EUR / 35% MSCI WORLD INDEX (TOTAL RETURN NET) HEDGED TO EUR / 25%,BLOOMBERG BARCLAYS GLOBAL CREDIT INDEX (TOTAL RETURN GROSS) HEDGED TO EUR **Fund Administrator Name :** Morgan Bank Luxembourg S.A. **Custodian Bank Name :** Morgan Bank Luxembourg S.A. **Full Share Class Name :** JPMorgan Funds - JPM Global Multi Strategy Income A (div) - EUR

HTML Client Demo

Search for:
EUR EONIA Equity or Bond Funds registered in Switzerland with fee 2% and share class P

Google Style Searching for Investment Funds
(With Thousands of Data Attributes)

EUR EONIA 0.02 equity bond P Switzerland   [Search]

○ Parse   ○ Simple

Showing 163 results, took:12892ms

Rank	ISIN	Share Class	Matched Fields
8.292718	LU0165130674	undefined	**Benchmark :** EUR EONIA **Abbreviated Share Class Name :** HSBC GLOBAL INVESTMENT FUNDS - EURO RESERVE AC **Full Share Class Name :** HSBC Global Investment Funds - HSBC GLOBAL INVESTMENT FUNDS - EURO RESERVE AC
8.292718	LU0165130591	undefined	**Benchmark :** EUR EONIA **Abbreviated Share Class Name :** HSBC GLOBAL INVESTMENT FUNDS - EURO RESERVE AD **Full Share Class Name :** HSBC Global Investment Funds - HSBC GLOBAL INVESTMENT FUNDS - EURO RESERVE AD
8.292718	LU0165131300	undefined	**Benchmark :** EUR EONIA **Abbreviated Share Class Name :** HSBC GLOBAL INVESTMENT FUNDS - EURO RESERVE IC **Full Share Class Name :** HSBC Global Investment Funds - HSBC GLOBAL INVESTMENT FUNDS - EURO RESERVE IC
7.775687	CH0002788591	undefined	**EFAMA Active EFC Classification :** Bond Aggregate MT EUR **Benchmark :** Barclays Euro Aggregate 500mio+ **Fund Administrator Name :** Northern Trust Switzerland AG **Abbreviated Share Class Name :** UBS (CH) Bond - EUR P **Custodian Bank Name :** UBS Switzerland AG **Subscription Fee Maximum :** 0.02 **EFAMA Main EFC Category :** Bond **Full Share Class Name :** UBS (CH) Bond Fund - EUR P

HTML Client Demo

Search for:

JPMorgan EUR Funds with Share Class A

OR

HSBC Funds registered anywhere except Luxembourg

Google Style Searching for Investment Funds

(With Thousands of Data Attributes)

(jpmorgan + EUR+A) | (HSBC + -LU*)   Search

○ Parse   ○ Simple

Showing 5 results, took:11073ms

Rank	ISIN	Share Class	Matched Fields
19.292982	FR0000438079	undefined	Fund Promoter Name : HSBC Global Asset Managmt FCP (France) Fund Advisor Name : HSBC Global Asset Management (France) Portfolio Managing Company Name : HSBC Global Asset Management (France) Fund Administrator Name : HSBC Global Asset Management (France) Abbreviated Share Class Name : HSBC Clic Euro 95 Full Share Class Name : HSBC Clic Euro 95
18.771591	FR0000438087	undefined	Fund Promoter Name : HSBC Global Asset Managmt FCP (France) Fund Advisor Name : HSBC Global Asset Management (France) Portfolio Managing Company Name : HSBC Global Asset Management (France) Fund Administrator Name : HSBC Global Asset Management (France) Abbreviated Share Class Name : HSBC Clic Euro 85 G Full Share Class Name : HSBC Clic Euro 85 G
18.771591	FR0000438061	undefined	Fund Promoter Name : HSBC Global Asset Managmt FCP (France) Fund Advisor Name : HSBC Global Asset Management (France) Portfolio Managing Company Name : HSBC Global Asset Management (France) Fund Administrator Name : HSBC Global Asset Management (France) Abbreviated Share Class Name : HSBC Clic Euro 90 G Full Share Class Name : HSBC Clic Euro 90 G
15.263008	LU0210533500	undefined	Fund Promoter Name : JPMorgan Bank Luxembourg SA Fund Advisor Name : JPMorgan Asset Management UK Limited Portfolio Managing Company Name : JPMorgan Asset Management UK Limited Full Share Class Name : JPMorgan Funds - JPM Global Convertibles (EUR) A (acc) - EUR
11.7577305	LU0697242484	undefined	Fund Promoter Name : JPMorgan Bank Luxembourg SA Fund Advisor Name : JPMorgan Asset Management UK Limited Full Share Class Name : JPMorgan Funds - JPM Global Multi Strategy Income A (div) - EUR