Maybe U2 should have used Elasticsearch

Jurgens du Toit

jrgns

Search is simple, right?

Search is simple, right?

SELECT * FROM `content`
WHERE `body` LIKE ‘%term%’

Search is simple, right?

SELECT * FROM `content`

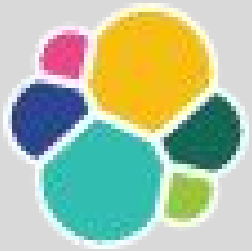
WHERE `body` LIKE '%term%'

Search

```
{
  "status" : 200,
  "name" : "Crimson Craig",
  "cluster_name" : "elasticsearch",
  "version" : {
    "number" : "1.7.0",
    "build_hash" : "929b9739cae115e73c346cb5f9a
    "build_timestamp" : "2015-07-16T14:31:07Z",
    "build_snapshot" : false,
    "lucene_version" : "4.10.4"
  },
  "tagline" : "You Know, for Search"
}
```

Search

What is it?



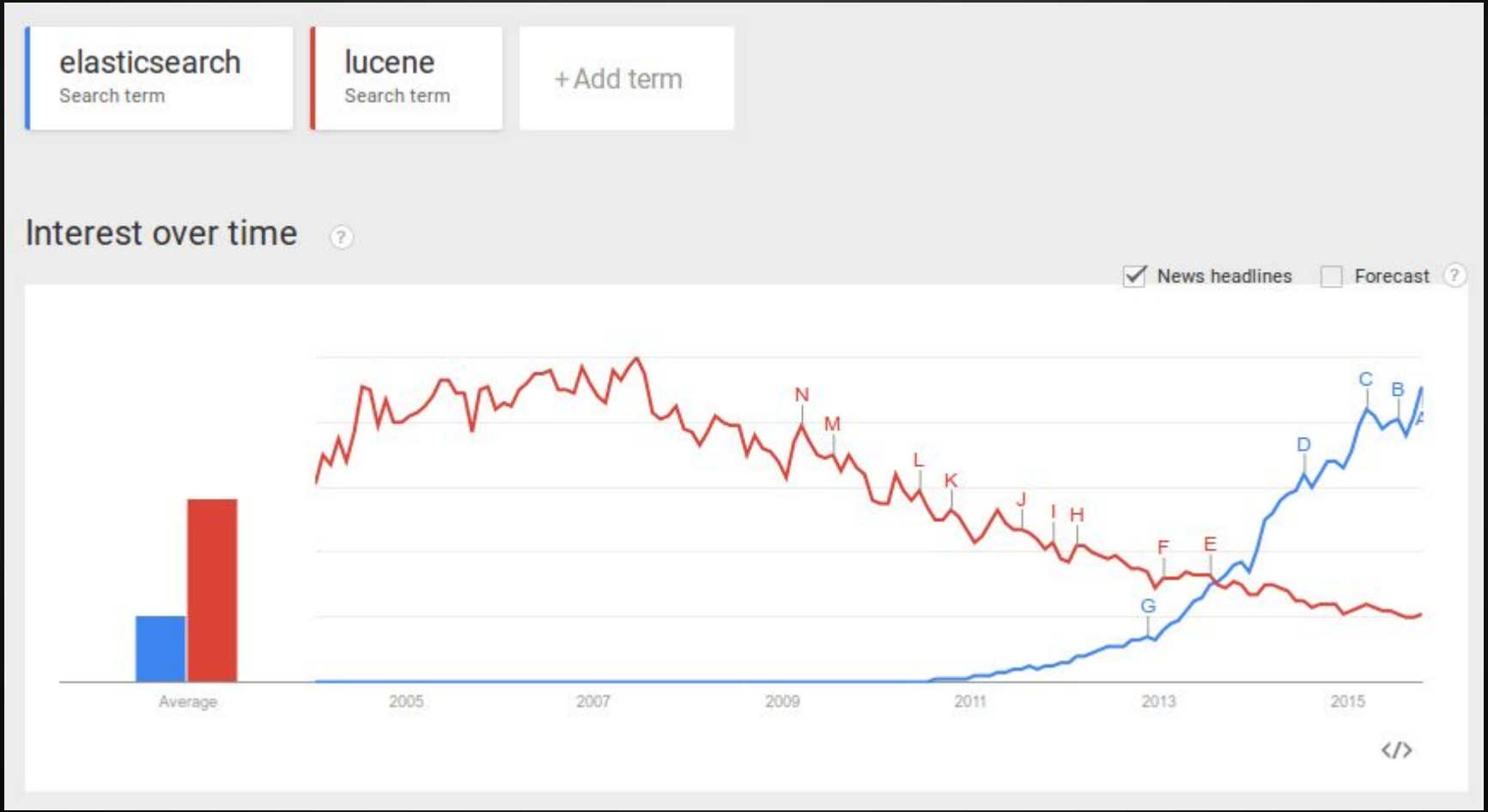
elastic

elasticsearch.

distributed restful search and analytics

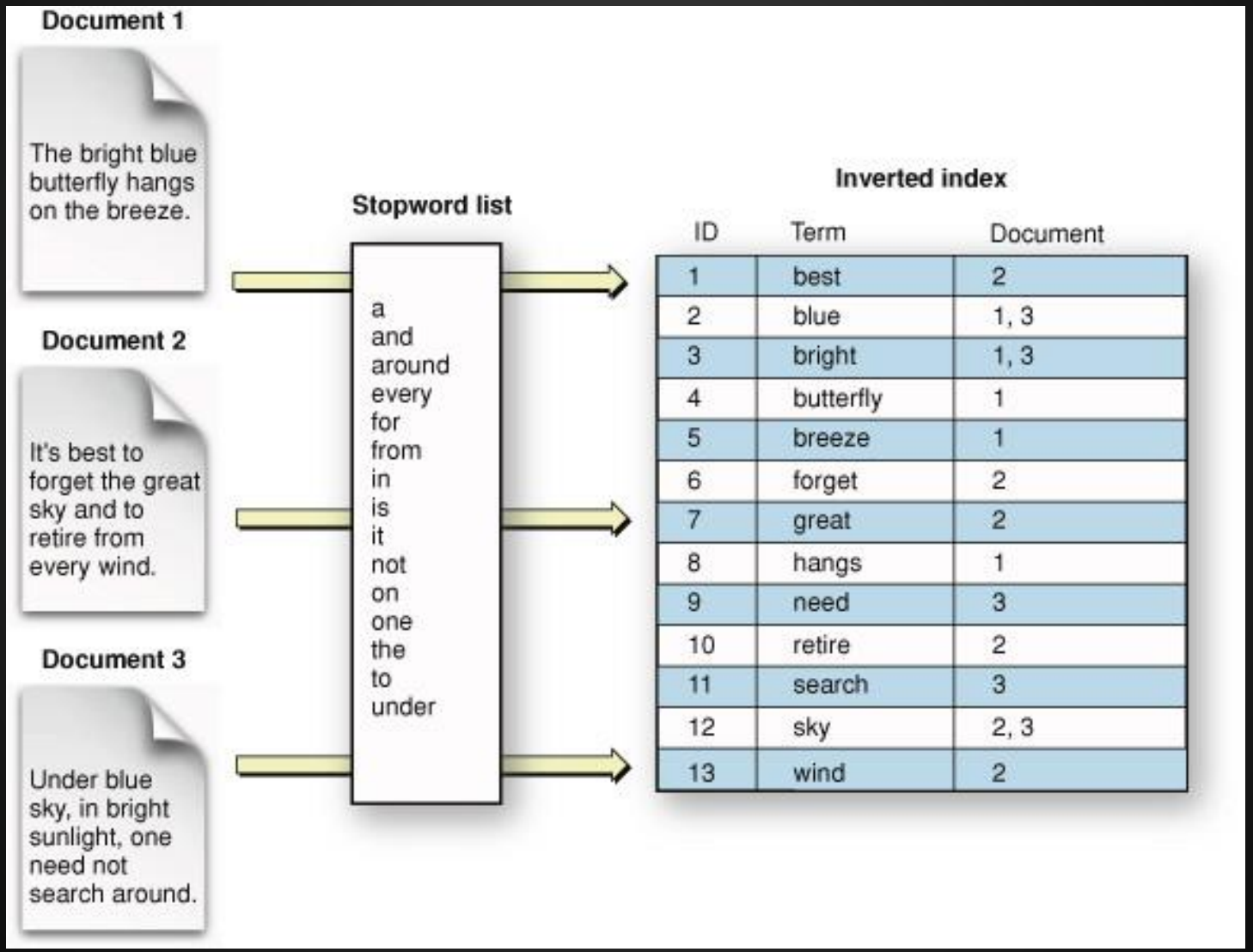
kibana

What is it?


elasticsearch
Search term
lucene
Search term
+ Add term
Interest over time
News headlines
Forecast
N
M
L
K
J
I
H
F
E
G
D
C
B
A
Average
2005
2007
2009
2011
2013
2015

Architecture

Indexing


Document 1
The bright blue butterfly hangs on the breeze.
Document 2
It's best to forget the great sky and to retire from every wind.
Document 3
Under blue sky, in bright sunlight, one need not search around.
Stopword list
a
and
around
every
for
from
in
is
it
not
on
one
the
to
under
Inverted index
ID
Term
Document
1 best 2
2 blue 1, 3
3 bright 1, 3
4 butterfly 1
5 breeze 1
6 forget 2
7 great 2
8 hangs 1
9 need 3
10 retire 2
11 search 3
12 sky 2, 3
13 wind 2

Use cases

Search

Time Series & Log Analysis

Application Database

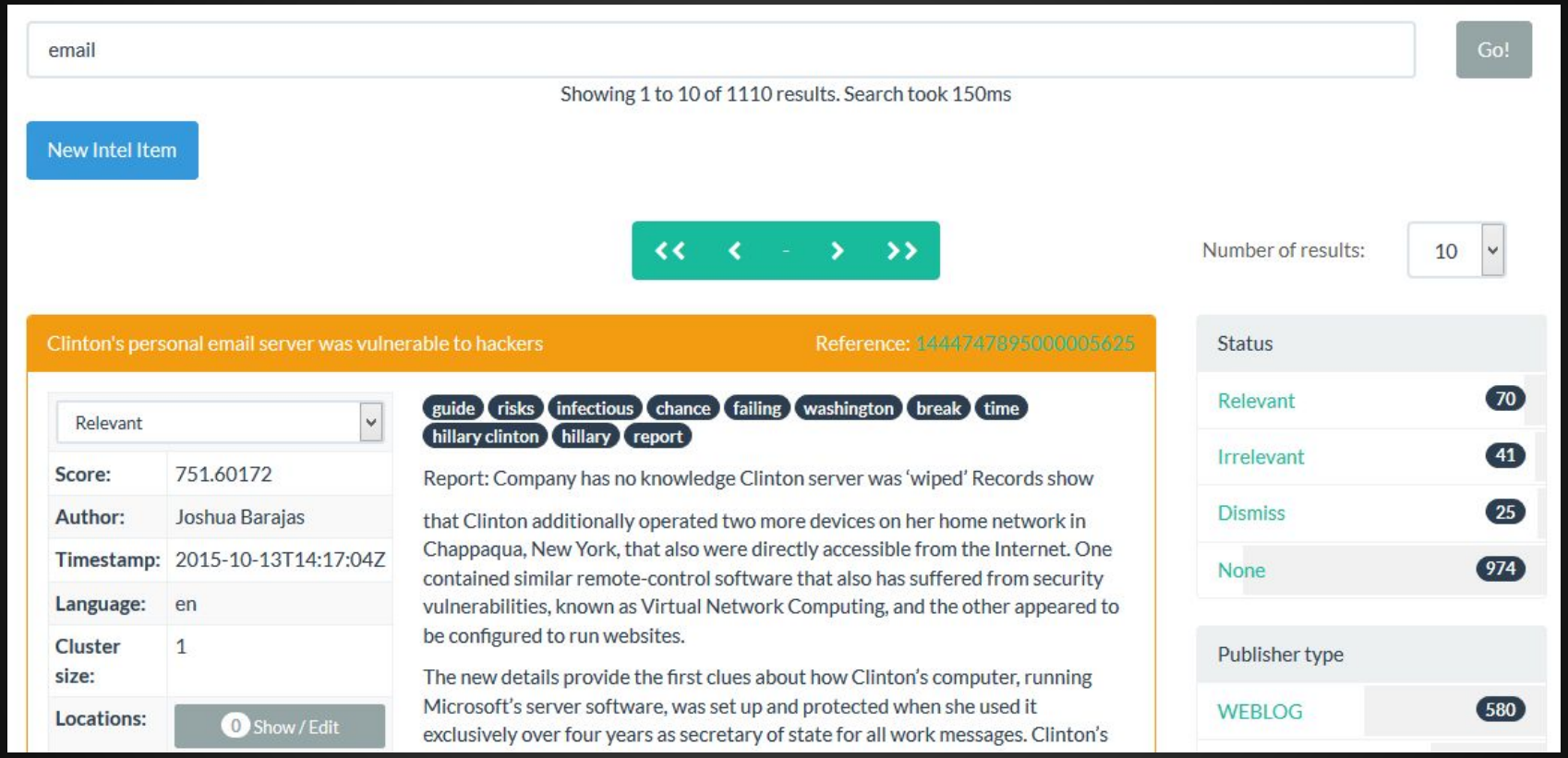
Use cases

Search

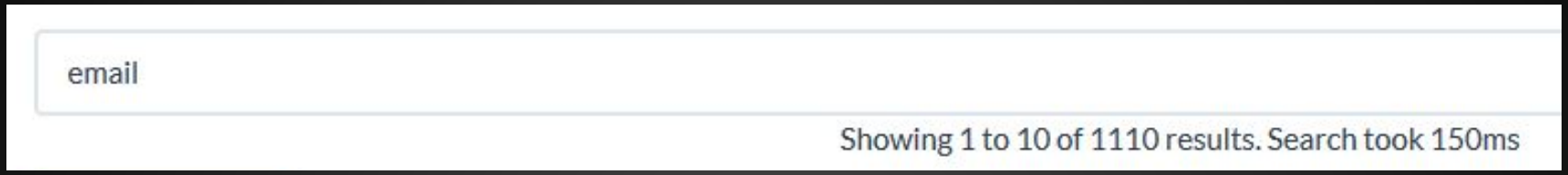
Time Series & Log Analysis

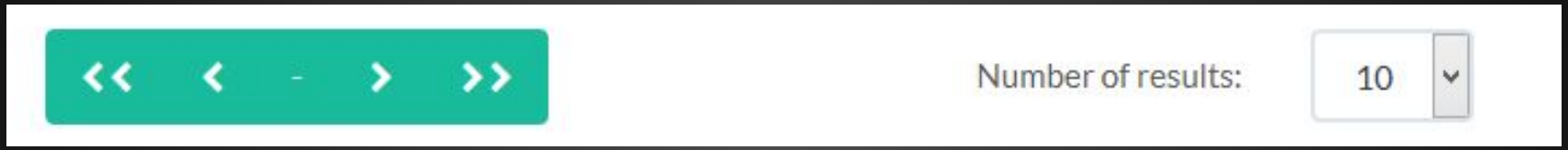
Application Database

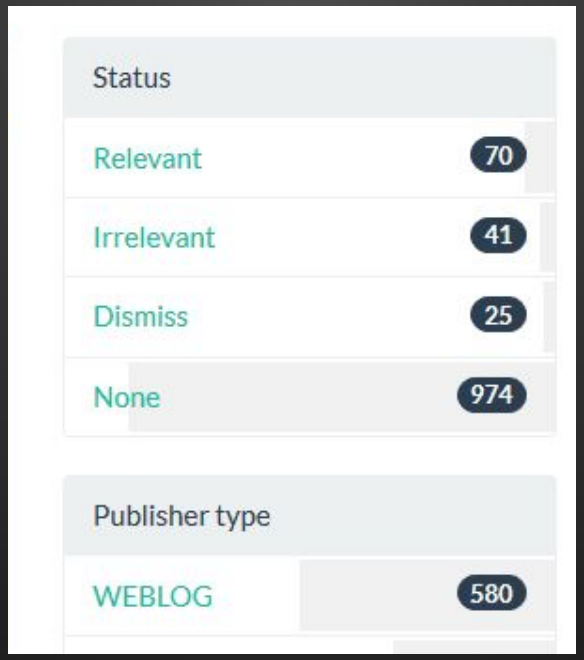
Search Features



Search Features







Search Features

```sql
-- Actual Search
SELECT * FROM `items` WHERE `content` LIKE '%email%'
-- Total number of items
SELECT COUNT(*) FROM `items` WHERE `content` LIKE '%email%'
-- Status aggregation - Search
SELECT `status`, COUNT(*) AS `count`
    FROM `items` WHERE `content` LIKE '%email%' GROUP BY `statu
-- Status aggregation - Global
SELECT `status`, COUNT(*) AS `count`
    FROM `items` GROUP BY `status`
```

Search Features

```
curl -s 'localhost:9200/_search?pretty' -d '
{
    "query": {
        "match": { "message": "email" }
    },
    "aggregations": {
        "status_values": {
            "terms": {
                "field": "status"
            }
        }
    }
}'
```

Search Features

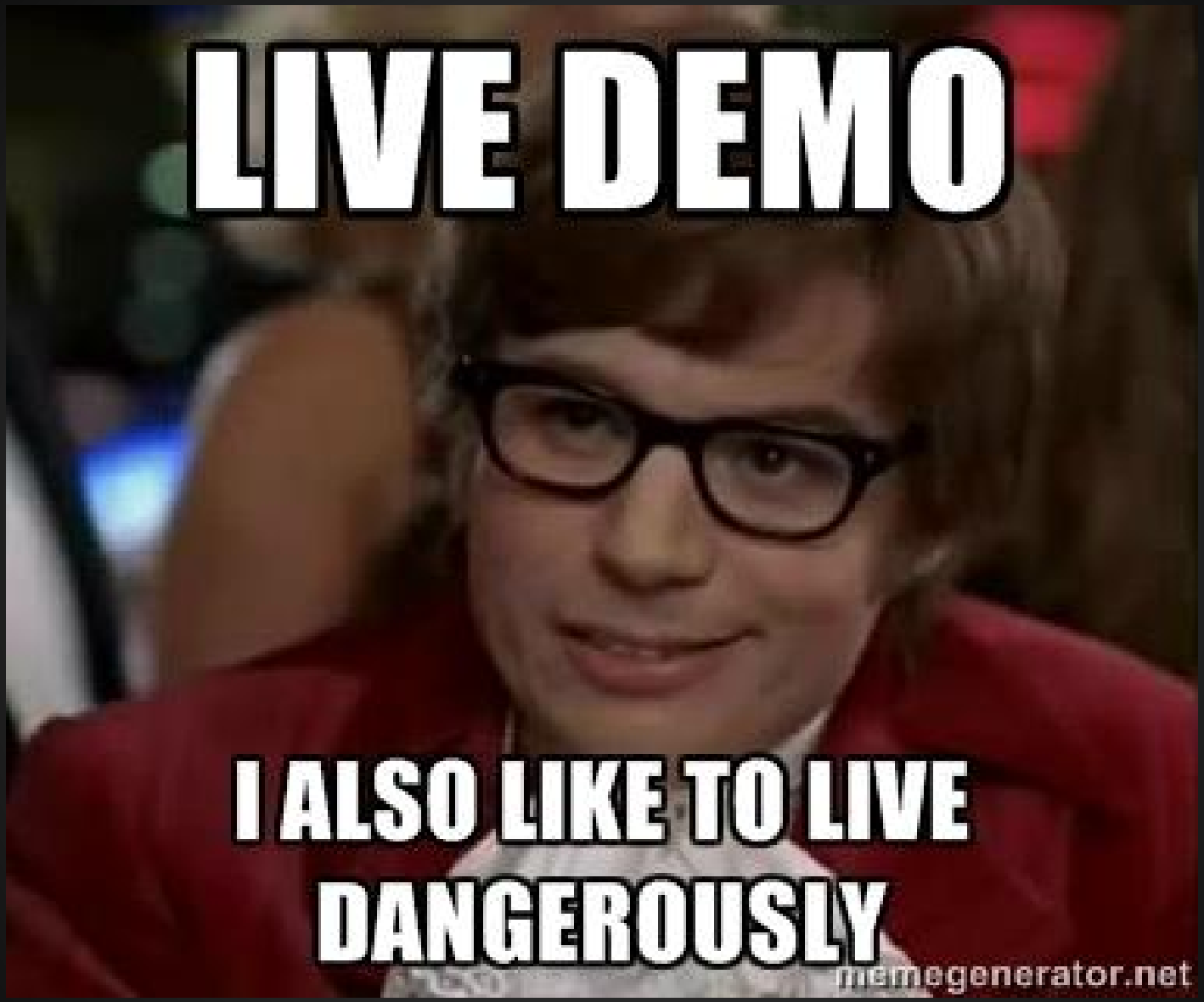
- Queries
- Filters
- Aggregations

Search Features

- Query String
- Match
- Match Phrase / Spans
- More Like This
- Boolean
- Fuzzy
- ...

Time Series Data / Logs

- A LOT of Data
- Difficult to manage
- Notorious to parse
- A wealth of information, if…


LIVE DEMO
I ALSO LIKE TO LIVE DANGEROUSLY
memegenerator.net

Resources

- **Reference Guide**
  https://www.elastic.co/guide/en/elasticsearch/reference/current/index.html
- **The Definitive Guide**
  https://www.elastic.co/guide/en/elasticsearch/guide/current/index.html
- **Resiliency Status**
  https://www.elastic.co/guide/en/elasticsearch/resiliency/current/index.html
- **Jo’burg Elastic Meetup**
  http://www.meetup.com/Elasticsearch-South-Africa/

Questions?