

Getting Started with Elasticsearch

Apr. 2022



Agenda



01

**What's
Elasticsearch?**



02

Use Cases



03

**Recent
Topics**

Takeaway



Takeaway

4

1. Elasticsearch (ES) is de facto standard for full-text search engines
2. There are some types of ES services, the main one is Elastic Stack
3. A similarity search system can be easily built in combination with Deep Learning



01

What's Elasticsearch?



- **Full-text search engine by Elastic co.**
 - De facto standard for full-text search engines
 - Search speed is faster than competitive services
- **Elasticsearch is adopted by many companies**
 - e.g.) Adobe, Uber, Shopify, Facebook, Netflix, Quora, ...
- **Flexible search queries with ML/DL**
 - kNN logic was added in Ver.7.3 later (Jul.2019)
 - Approx. kNN was added in Ver.8.0 later (Feb.2022)



Customers of Elastic Co.

7

Customers across various industries, geographies

Technology	Finance	Telco	Consumer	Healthcare	Public Sector	Automotive/ Transportation	Retail/ Ecommerce
							
							
							
							
							

[\[Q1-F22: Investor Presentation and Company Overview \(p.9\)\]](#)



Why Use Elasticsearch?

Why Use Elasticsearch?

9

(F-01) Very fast search speed

(F-02) Scalability

(F-03) Flexible search by simple query

(F-04) High affinity with ML/DL

(F-05) Faster development speed



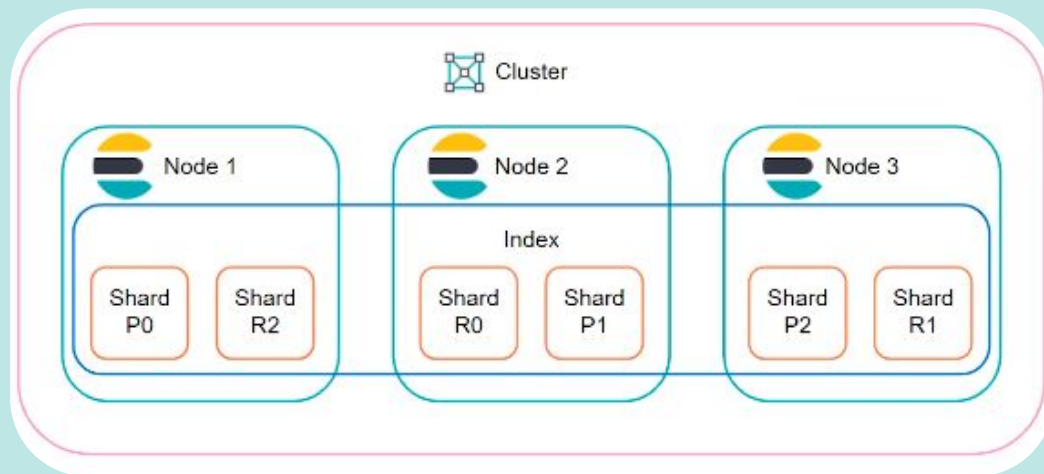
(F-01) Very fast search speed

10

- **Elasticsearch is built on top of Apache Lucene**
 - ES excels at full-text search
- **High latency on search**
 - Typically within one second
- **Well suited for time-sensitive use cases**
 - Security analytics, Infrastructure monitoring, ...



- **Elasticsearch is designed for cluster configuration**
 - Distribution / replica can be executed automatically
 - High affinity with Hadoop



(F-03) Flexible search by simple query

12

Exact match search and fuzzy search can be mixed

```
body = {  
  "from": skipN,  
  "size": topN,  
  "query": {  
    "bool": {  
      "filter": [  
        { 'term' : { 'xxx.keyword' : target_xxx } },  
        { 'term' : { 'yyy.keyword' : target_yyy } },  
      ],  
      "should" : [  
        { 'match' : { zzz : target_zzz } }  
      ]  
    }  
  },  
  "_source": {"includes": output_columns},  
  "sort" : [{"_score": "desc"}]  
}
```

Exact match
search

Fuzzy search



kNN Vector search can be available Ver.7.3 later (Jul.2019)

```
body = {
  'query': {
    'script_score': {
      'query': {
        'bool': {
          'filter': [
            {'term': {'xxx.keyword': target_xxx}},
            {'term': {'yyy.keyword': target_yyy}}, ],},
        'script': {
          'source': ('cosineSimilarity(params.query_vector, doc['image_vector']) + 1.0)*0.5,
          'params': {'query_vector': query_vector}
        }
      }
    },
    '_source': {'includes': output_columns },
    'sort': [{'_score': 'desc'}], 'from': skipN, 'size': topN
  }
}
```

**exact match
search**

vector search

**Normalize: 0.0~1.0
(Cosine: -1.0~1.0)**

(F-05) Faster development speed (1/2)

14

Minor updates are made every 1-3 months

Version 8.x

Released	Version
2022/03/09	8.1
2022/02/11	8.0

Version 7.x

Released	Version
2022/02/02	7.17
2021/12/08	7.16
2021/09/23	7.15
2021/08/04	7.14
2021/05/26	7.13
2021/03/24	7.12
2021/02/11	7.11
2020/11/12	7.10

Released	Version
2020/08/19	7.9
2020/06/19	7.8
2020/05/14	7.7
2020/02/12	7.6
2019/12/03	7.5
2019/10/01	7.4
2019/07/31	7.3
2019/06/26	7.2
2019/05/21	7.1
2019/04/11	7.0

Version 6.x





Released	Version
2019/05/21	6.8
2019/03/27	6.7
2019/01/29	6.6
2018/11/14	6.5
2018/08/23	6.4
2018/06/13	6.3
2018/02/07	6.2
2017/12/14	6.1
2017/11/15	6.0

*kNN logic can be add-on in ES Ver.7.3 (Jul.2019)

(F-05) Faster development speed (2/2)

15

Ref.) OSS Search Engines Comparison

No.		1	2	3	4	5
OSS Name		Elasticsearch	Apache Solr	Splunk	NGT (Neighborhood Graph and Tree for Indexing)	SPTAG (Space Partition Tree And Graph)
Logo						--
Developer		Elastic	Apache Software Foundation	Splunk Inc.	Yahoo	Microsoft
1st Released		2010y	2004y	2003y	2016y	2019y
Github	Latest Version	Ver.8.1.2 (2022/04)	Ver.8.11.1 (2022/04)	?	v.14.3 (2022/04)	1st Release
	#Contributors	1,729	275	--	11	26
	#Commits	63,686	35,803	--	139	126
Performance	Search Speed	◎	○	○	◎	◎
	Complex Search	○	△	○	△	△
	Scalable	◎	△	○	◎	◎
	Analytical Flexibility	◎	△	◎	--	--
	Affinity with ML	◎	△	△	◎	◎
Ranking	DB-Engines Ranking (As of Apr.2022)	7	20	13	--	--
	within Search Engine	1	3	2	--	--
	Popularity	◎	○	○	--	--



Service Type of Elasticsearch



Service Type of Elasticsearch (3-types)

17

(a) Elastic Stack  elastic stack

- OSS provided by Elastic co.
- Need a server to use it

(b) Elastic Cloud  Elastic Cloud

- Full-managed cloud service provided by Elastic co.

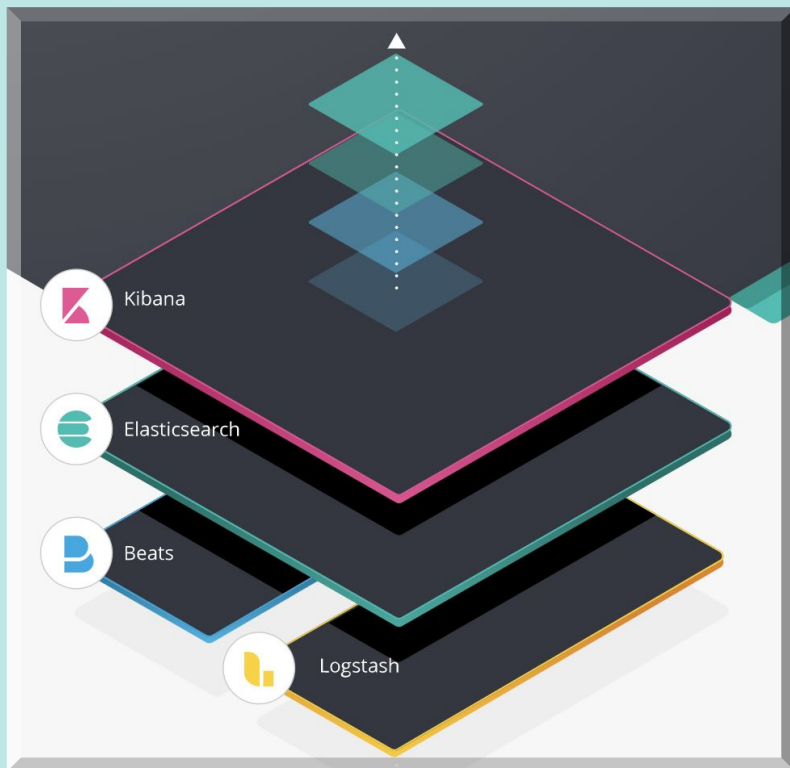
(c) Amazon OpenSearch Service  OpenSearch

- Elasticsearch provided by AWS
- Old Service Name : “Amazon Elasticsearch Service” (~Sep.2021)

☞ Full-Managed (b) Elastic Cloud is recommended, if price is ignored!!

(a) Elastic Stack (also known as the ELK Stack)

18



- **Elasticsearch:**
Search Engine
- **Logstash:**
ETL Pipeline
- **Kibana:**
Searching/visualizing tool
- **Beats:**
Lightweight data shippers

(a) Elastic Stack: Elasticsearch

19

- Elasticsearch is a distributed, RESTful search and analytics engine capable of addressing a growing number of use cases
 - **“Ask your data questions of all kinds”**



The heart of the free and open Elastic Stack

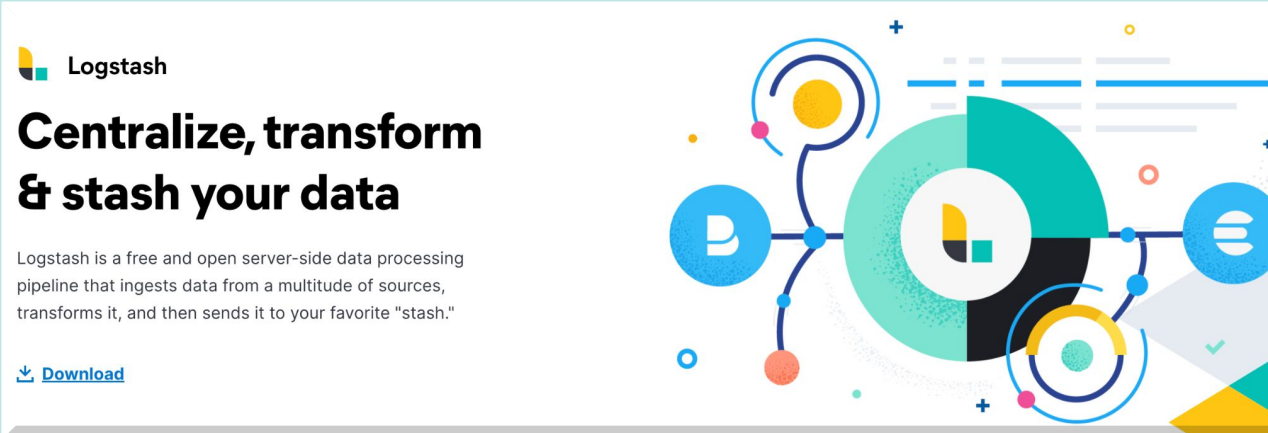
Elasticsearch is a distributed, RESTful search and analytics engine capable of addressing a growing number of use cases. As the heart of the Elastic Stack, it centrally stores your data for lightning fast search, fine-tuned relevancy, and powerful analytics that scale with ease.




(a) Elastic Stack: Logstash

20

- Logstash is a free and open server-side data processing pipeline that ingests data from a multitude of sources, transforms it, and then sends it to your favorite "stash."
 - **“Inputs, filters & outputs”**



 Logstash

Centralize, transform & stash your data

Logstash is a free and open server-side data processing pipeline that ingests data from a multitude of sources, transforms it, and then sends it to your favorite "stash."

[Download](#)

The image shows a promotional banner for Logstash. On the left, there is the Logstash logo (a square divided into four colored quadrants) followed by the text 'Logstash'. Below this is the headline 'Centralize, transform & stash your data' in a bold, sans-serif font. Underneath the headline is a short paragraph describing Logstash as a free and open server-side data processing pipeline. At the bottom left of the banner is a 'Download' link with a downward arrow icon. On the right side of the banner is a colorful, abstract graphic featuring a central circular element with a grid pattern, surrounded by various geometric shapes, lines, and icons, including a blue circle with a white 'D' and another blue circle with a white 'E'.



(a) Elastic Stack: Kibana

21

- Kibana is a free and open user interface that lets you visualize your Elasticsearch data and navigate the Elastic Stack
 - **“A picture's worth a thousand log lines”**



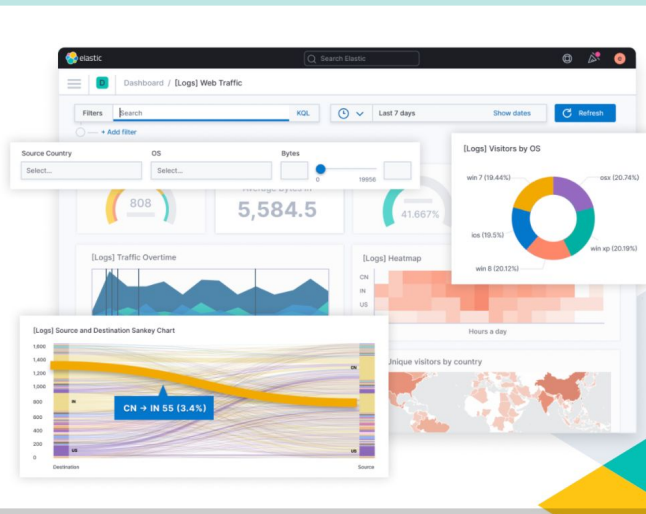
Kibana

Your window into the Elastic Stack

Kibana is a free and open user interface that lets you visualize your Elasticsearch data and navigate the Elastic Stack. Do anything from tracking query load to understanding the way requests flow through your apps.

Start free trial

[Download Kibana](#)



Ref.) Basic Terms in Elasticsearch

22

- Note that basic terms are different from RDB
 - The following terms correspond

RDB	Elasticsearch
database	index
table	mapping type
column	field
record	document



02

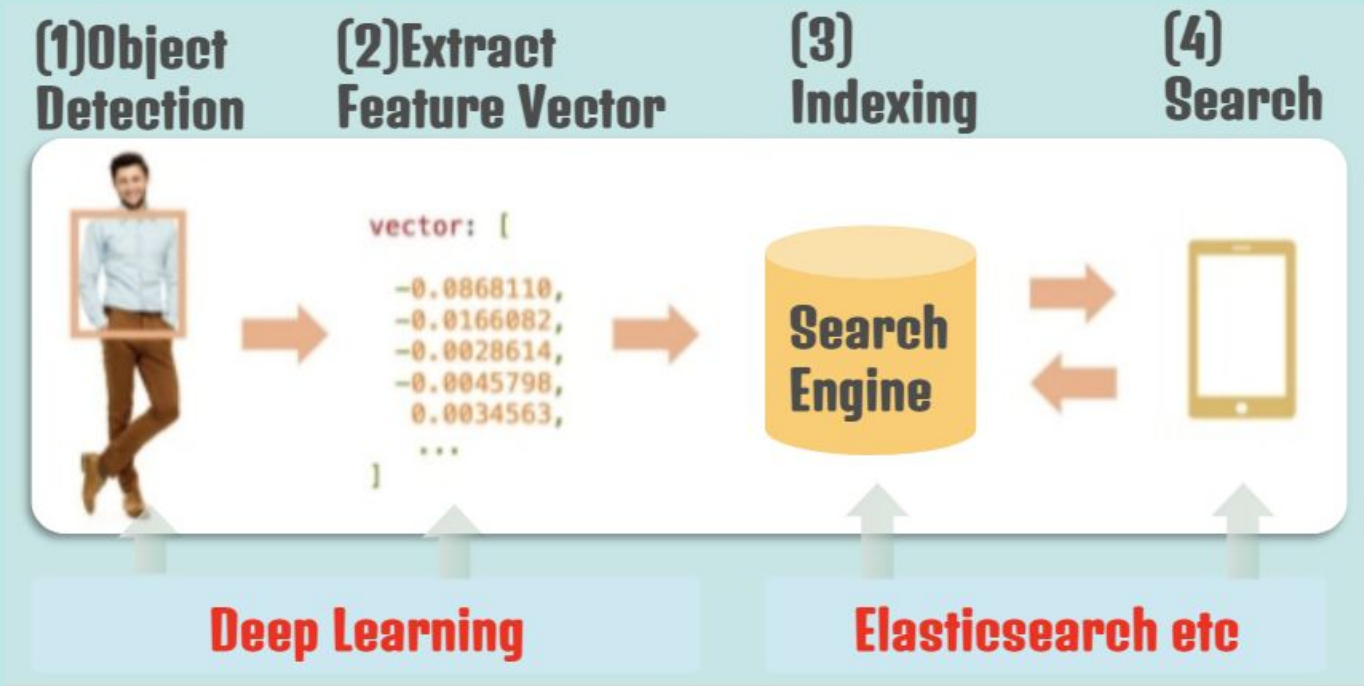
Use cases



Use-case(1): Image Similarity Search

24

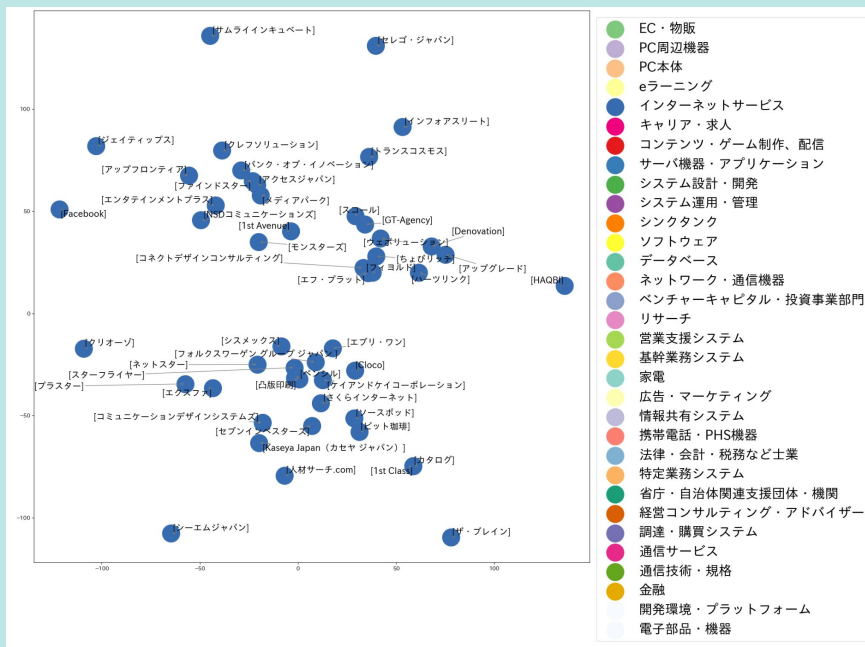
Easily create similar image searches with DL + ES



Similarities can be calculated by embedding tokenized texts

Ex.) Company Clustering

Similarity search based on the tokenized texts of company profiles



Use cosine similarity between “embedding vectors”

```
body = {
  'query': {
    'script_score': {
      'query': {
        'bool': {
          'filter': [
            { 'term': { 'xxx.keyword' : target_xxx } },
            { 'term': { 'yyy.keyword' : target_yyy } }, ],
          'script': {
            'source': ('cosineSimilarity(params.query_vector, doc["image_vector"]) + 1.0)*0.5,
            'params': {'query_vector': query_vector}
          }
        }
      }
    },
    '_source': {'includes': output_columns },
    'sort' : [{'_score':'desc'}], 'from': skipN, 'size': topN
  }
}
```

**exact match
search**

vector search

Normalize: 0.0~1.0
(Cosine: -1.0~1.0)

Embedding Vector (image/text → numerical vector)

27

Word Embedding Vector refers to
“assign each word to a unique vector”

Embedding example

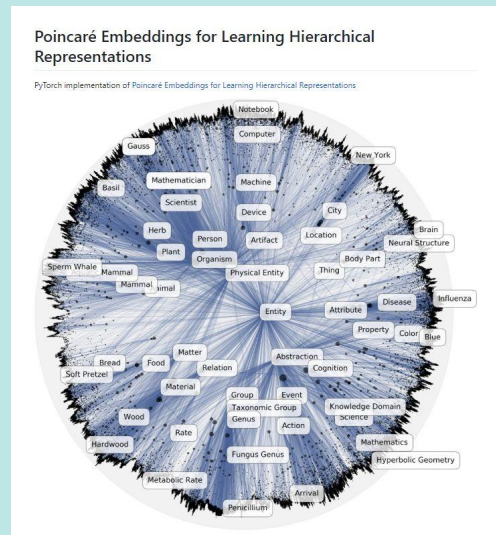
[apple, orange, banana]

"apple" = [1, 0, 0]

"banana" = [0, 0, 1]

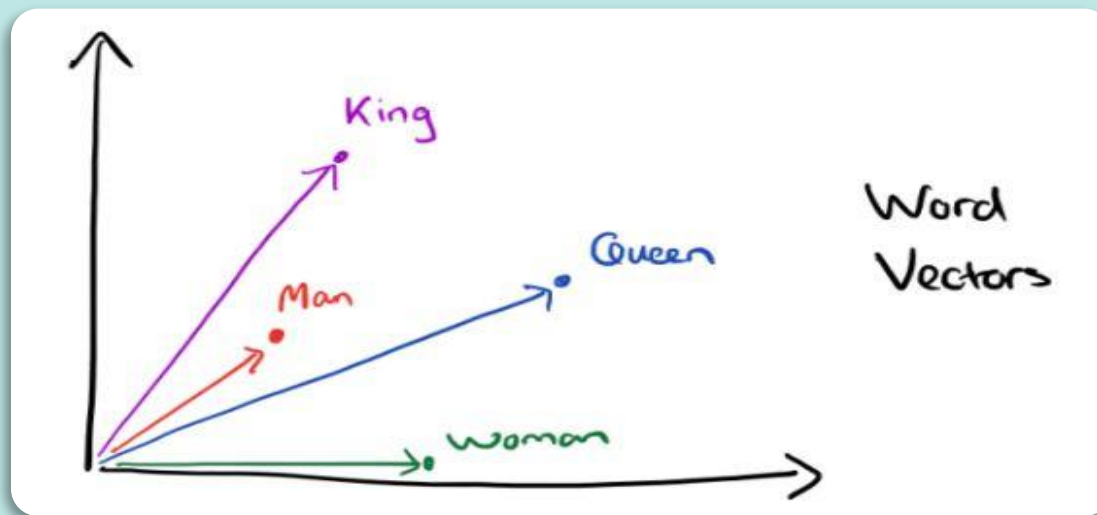
All words are represented by the same dim.
So, it can be easily input as valuables to ML

Word embedding to some space



[Facebook]: Poincaré Embeddings

King - Man + Woman = Queen!!



03

Recent Topics



(T-01) Elastic vs AWS

(T-02) New Features of Ver.8.0

(T-03) Breaking Changes in Major Updates

Topic-01 Elastic vs AWS



Topic-01: Elastic vs AWS (1/4)

32

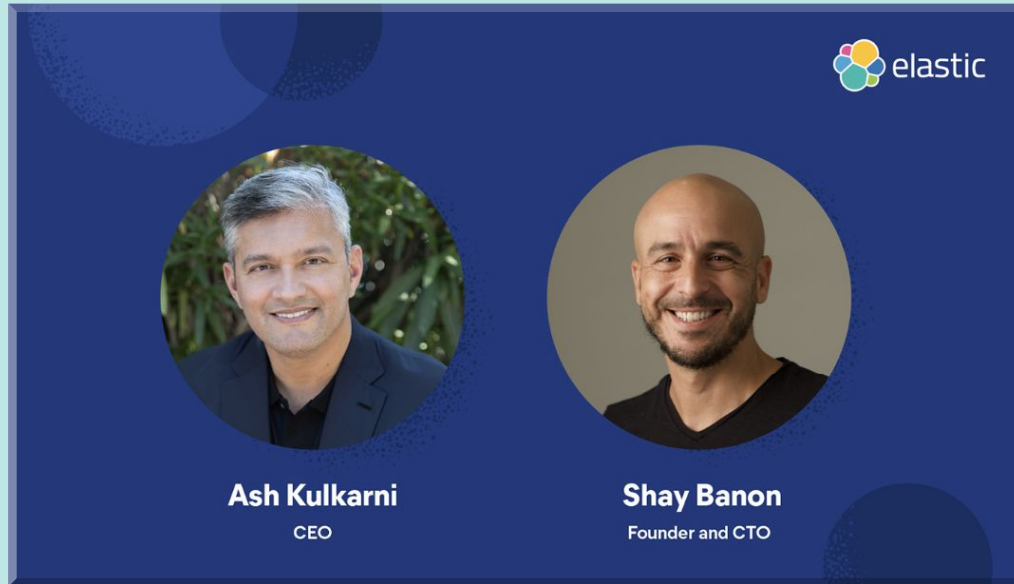


[\[Meet a CEO: Shay Banon from Elastic\]](#)

Topic-01: Elastic vs AWS (2/4)

33

Shay Banon: CTO(2012) -> CEO(2017) -> **CTO(2022)**

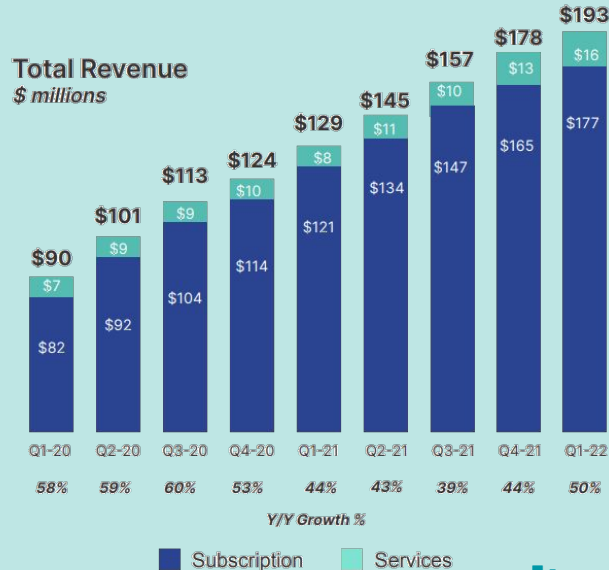


[Shay Banon to (re)assume the role of CTO,
Ash Kulkarni promoted to CEO (Jan.2022)]

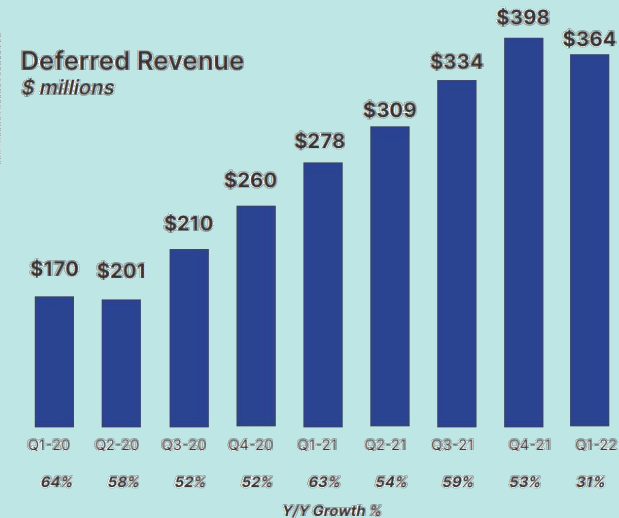
Ref.) Elastic's Quarterly Revenue

34

- Elastic's quarterly revenue is growing consistently
- FY2021 Q1-Q4 sales totaled \$ 608.5M, YoY + 42% growth



Sums may not add to totals due to rounding.



[\[Investor Presentation and Company Overview\]](#)



20 JANUARY 2021

NEWS

EN

ES

PT

KR

JP

DE

FR

CN

Amazon: NOT OK - why we had to change Elastic licensing

By [Shay Banon](#)

“

We have seen that this trademark issue drives confusion with users thinking Amazon Elasticsearch Service is actually a service provided jointly with Elastic, with our blessing and collaboration. This is just not true. NOT OK.

[\[Doubling down on open, Part II\]](#)

- Elastic changed to the dual license ([Jan.2021](#))
 - **Server Side Public License (SSPL)**
 - **Elastic License** (restrict commercial services)
- To prevent AWS from offering ES and Kibana as managed services on their own





Topic-02 New Features of Ver.8.0

The major update in Feb. 2022 (7.x → 8.x)

- **Approximate k-nearest neighbor (ANN) searches**
 - So far, only exact kNN, which is hard to scale
 - Large-scale vector-based searches can be performed faster
- **Native support for NLP models**
 - Starting with Elastic Stack 8.0, NLP using external models published on HuggingFace has become very easy to run



Approximate k-nearest neighbor (ANN) searches

- So far, only exact kNN, which is hard to scale
- Large-scale vector-based searches can be performed faster

Common use cases for kNN:

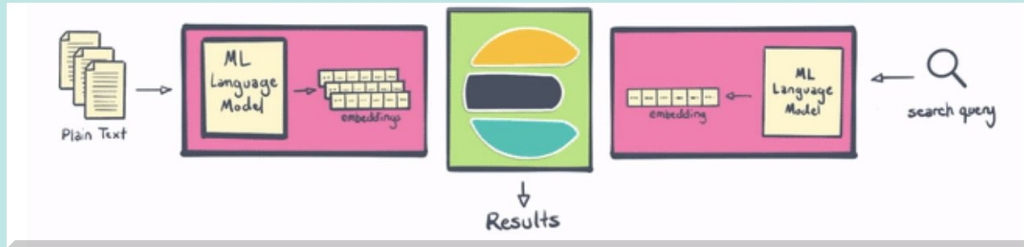
- Relevance ranking based on NLP algorithms
- Product recommendations
- Similarity search for images or videos

[\[k-nearest neighbor \(kNN\) search\]](#)

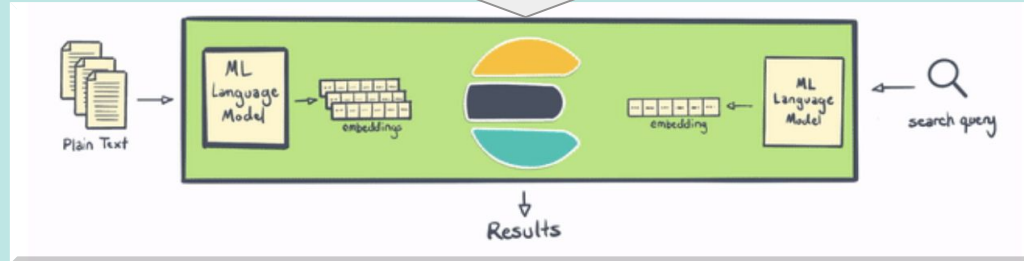


Power of NLP to move analysis to unexplored territory

7.x



8.x



[Elastic 8.0: A new era of speed, scale, relevance, and simplicity]

Easily try the pretrained NLP tasks in Ver.8.x

- (1) Select a trained model (HuggingFace NLP tasks)
- (2) Import the model and its tokenizer vocabulary

```
eland_import_hub_model
--url <clusterUrl> \
--hub-model-id elastic/[model identifier in the HuggingFace] \
--task-type [NLP_task]
```

```
=== NLP tasks ===
· fill_mask
· ner(NER)
· text_classification
· text_embedding
· zero_shot_classification
```

- (3) Deploy the model in your cluster
- (4) Inference (Try it out!!)

```
POST /_ml/trained_models/[NLP_task]/deployment/_infer
{
  "docs":{ "text_field" : "Sasha bought 300 shares of Acme Corp in 2022." }
}
```

Topic-03

Breaking Changes in Major Updates



ES contains many breaking changes in major update

- Note that unintended search results may be gotten if upgraded as it is (This is not limited to ES...)

[\[Migrating to 8.0\]](#)

Migrating to 8.0

This section discusses the changes that you need to be aware of when migrating your application to Elasticsearch 8.0.

Breaking changes

The following changes in Elasticsearch 8.0 might affect your applications and prevent them from operating normally. Before upgrading to 8.0, review these changes and take the described steps to mitigate the impact.



Ex.) kNN Comparison Ver.7.x and 8.x

[Ver.7.3~7.17]: Vanilla kNN

- Flexible search by mixing exact and vector searches
 - GET index/**_search**
 - A bit complicated due to script score query

[Ver.8.x]: Approximate kNN (ANN)

- ANN search was released as a **[Technical Preview] !!**
- Provided as a new API endpoint: **_knn_search**
 - GET index/**_knn_search**
- Easy to use, but not compatible with Query DSL

Bonus Slide



Is the motivation surprisingly simple?

46

Shay Banon created Elasticsearch while trying to index recipes for his wife, who was attending cooking school



[(Our story) It started with a recipe app]

Our story





References

❑ Elastic Co.:

- ❑ [\[\(Our story\) It started with a recipe app\]](#)
- ❑ [\[Q1-F22: Investor Presentation and Company Overview\]](#)

❑ Query DSL:

- ❑ [\[42 Elasticsearch Query Examples – Hands-on Tutorial \(Mar.2020\)\]](#)
- ❑ [\[Elasticsearch Queries: A Guide to Query DSL \(Aug.2021\)\]](#)
- ❑ [\[Elasticsearch 7.x Cheatsheet\]](#)

❑ Cluster Configuration:

- ❑ [\[Creating an Elasticsearch Cluster: Getting Started \(Jan.2020\)\]](#)
- ❑ [\[How to build an elastic search cluster for production? \(Mar.2021\)\]](#)

❑ **kNN Vector Search (Ver.7.3~; Jul.2019~):**

- ❑ [\[Text similarity search with vector fields \(Aug.2019\)\]](#)
- ❑ [\[Similarity Search and Similar Image Search in Elasticsearch \(Mar.2020\)\]](#)
- ❑ [\[Building a k-NN Similarity Search Engine using AWS \(Mar.2020\)\]](#)
- ❑ [\[Scalable Semantic Vector Search with Elasticsearch \(Jan.2021\)\]](#)
- ❑ [\[Speeding up BERT Search in Elasticsearch \(Mar.2021\)\]](#)

❑ **Ver.8.x (Feb.2022~):**

- ❑ [\[Introducing ANN search in Elasticsearch 8.0 \(Feb.2022\)\]](#)
- ❑ [\[Elastic 8.0: A new era of speed, scale, relevance, and simplicity \(Feb.2022\)\]](#)
- ❑ [\[Introduction to modern NLP with PyTorch in ES \(Feb.2022\)\]](#)
- ❑ [\[Elastic Gets New Vector Search and NLP Capabilities \(Feb.2022\)\]](#)

- ❑ **Beginner's Crash Course to Elastic Stack (YouTube):**
 - ❑ [\[Part 1: Intro to Elasticsearch and Kibana\]](#)
 - ❑ [\[Part 2: Relevance of a search\]](#)
 - ❑ [\[Part 3: Full text queries\]](#)
 - ❑ [\[Part 4: Aggregations\]](#)
 - ❑ [\[Part 5: Mapping\]](#)
 - ❑ [\[Part 6: Troubleshooting Errors\]](#)

End of Documents

