Getting Started with Elasticsearch

Apr.2022

elasticsearch

Agenda

01
What's Elasticsearch?
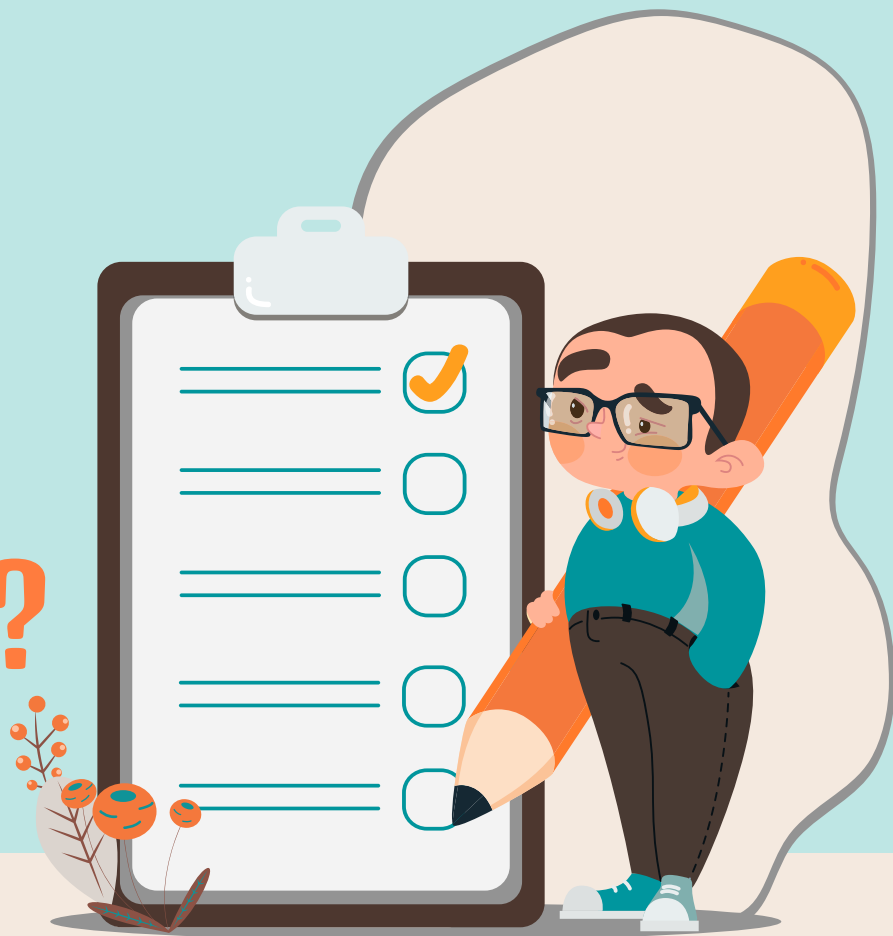
02
Use Cases

03
Recent Topics

Takeaway

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?

- **<u>Full-text search engine by Elastic co.</u>**
  - De facto standard for full-text search engines
  - Search speed is faster than competitive services

- **<u>Elasticsearch is adopted by many companies</u>**
  - e.g.) Wikimedia, Facebook, Netflix, Quora, …

- **<u>Abundant complex search queries</u>**
  - kNN logic can be add-on in ES Ver.7.3 later (Aug.2019)
  - Latest version is 8.1 (Apr.2022)

- **Very fast search speed**

- **Scalable**

- **Flexible search by simple query**

- **High affinity with ML/DL**

- **Faster development speed**

Elasticsearch has a faster development speed

	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	elasticsearch	Apache Solr	splunk>	NGT	--
	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)	?	v1.14.3 (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	◎	○	○	--	--

(a) <u>Elastic Stack</u>   elastic stack
- OSS provided by Elastic co.
- Need a server to use it

(b) <u>Elastic Cloud</u>   Elastic Cloud
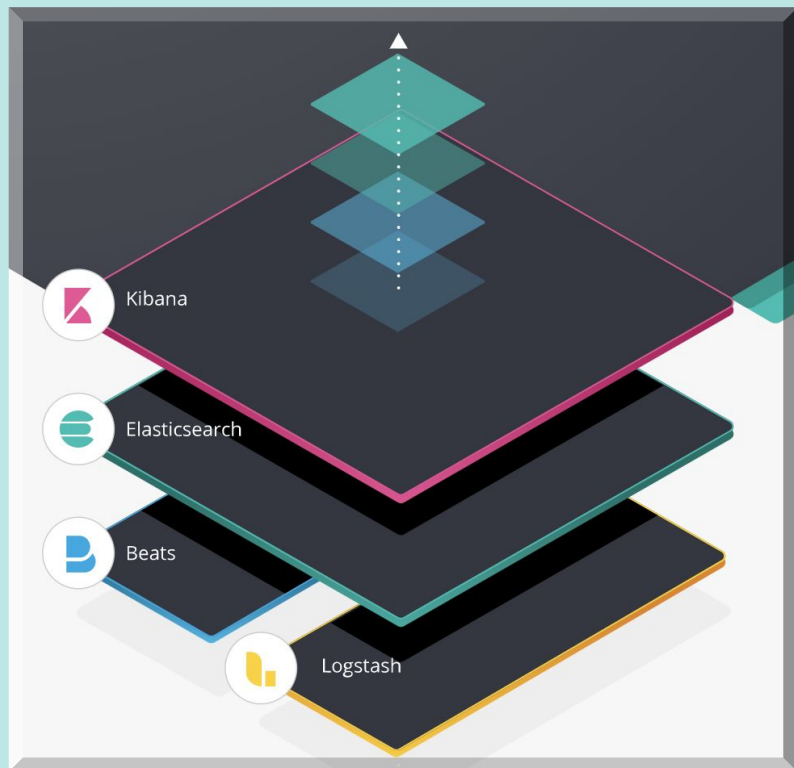- Full-managed cloud service provided by Elastic co.

(c) <u>Amazon OpenSearch Service</u>   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)

- **Elasticsearch:**
Search Engine

- **Logstash:**
ETL Pipeline

- **Kibana:**
Searching/visualizing tool

- **Beats:**
Lightweight data shippers

- **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"**

Elasticsearch

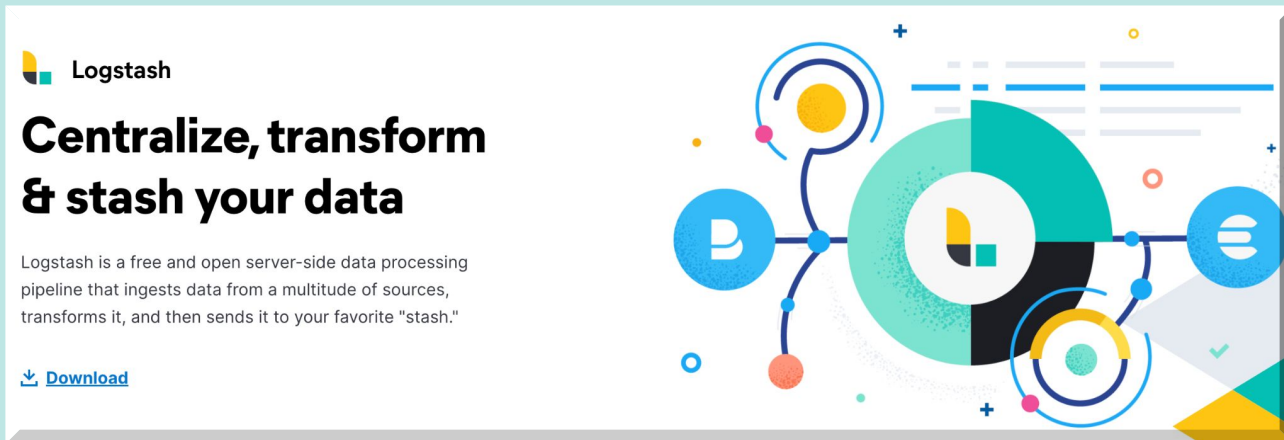# The heart of the free and open Elastic Stack

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.

- **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

(a) Elastic Stack: Kibana

- **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"**
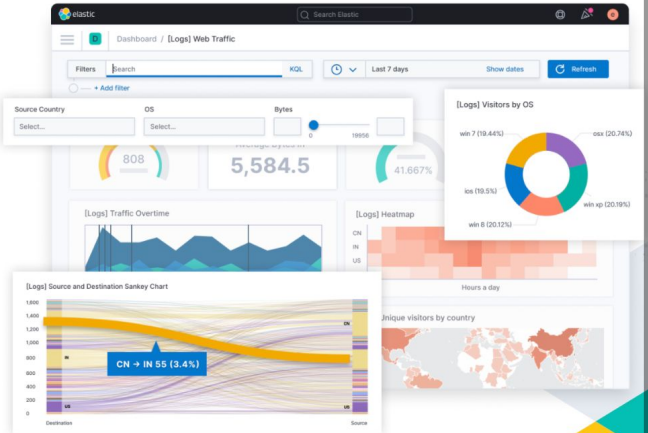
Kibana

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

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

RDB	Elasticsearch
database	index
table	mapping type
column	field
record	document

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

02

Use cases

Easily create similar image searches with DL + ES

(1)Object Detection　　**(2)Extract Feature Vector**　　**(3) Indexing**　　**(4) Search**

```
vector: [

    -0.0868110,
    -0.0166082,
    -0.0028614,
    -0.0045798,
     0.0034563,
     ...
]
```

Search Engine

Deep Learning　　　　　　**Elasticsearch etc**

Similarities can be calculated by embedding tokenized texts
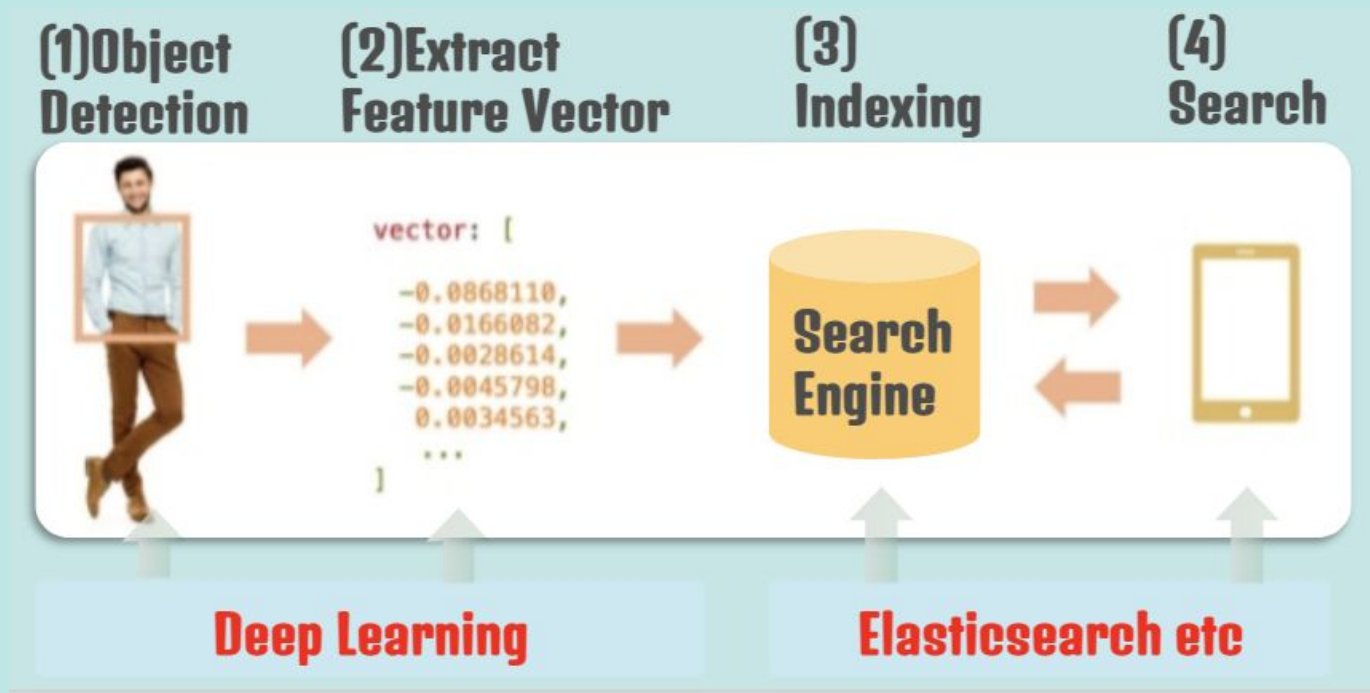
Ex.) Company Clustering

Similarity search based on the tokenized texts of company profiles
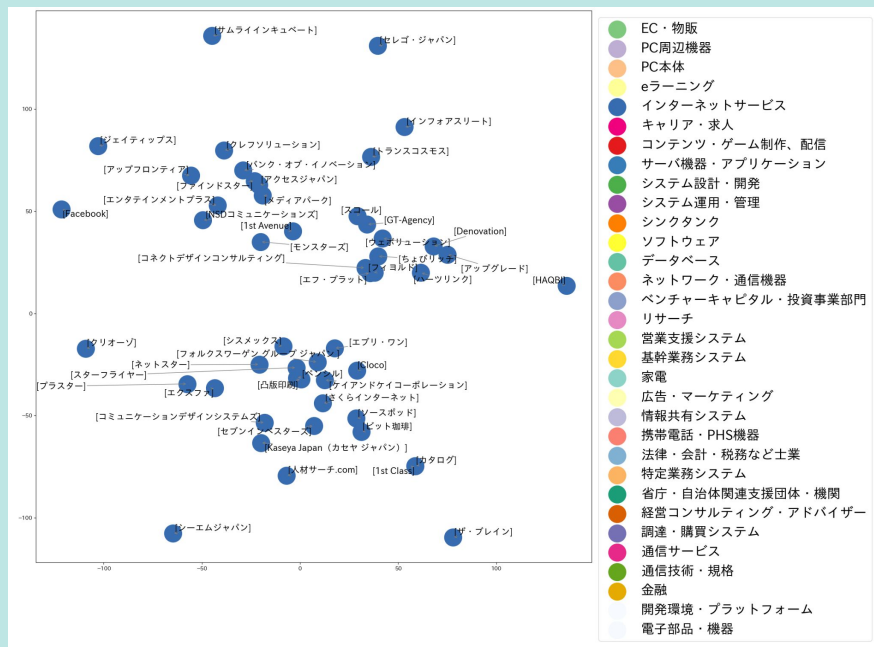
Use cosine similarity between "embedding vectors"

```
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

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

Word Embedding Vector refers to "assign each word to a unique vector"

Embedding example

[apple, orange, banana]
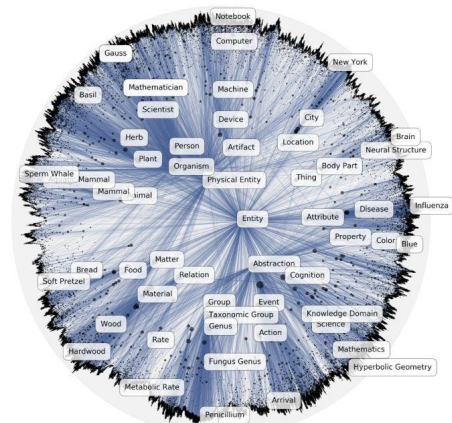
"apple"   = [1, 0, 0]
"banana" = [0, 0, 1]

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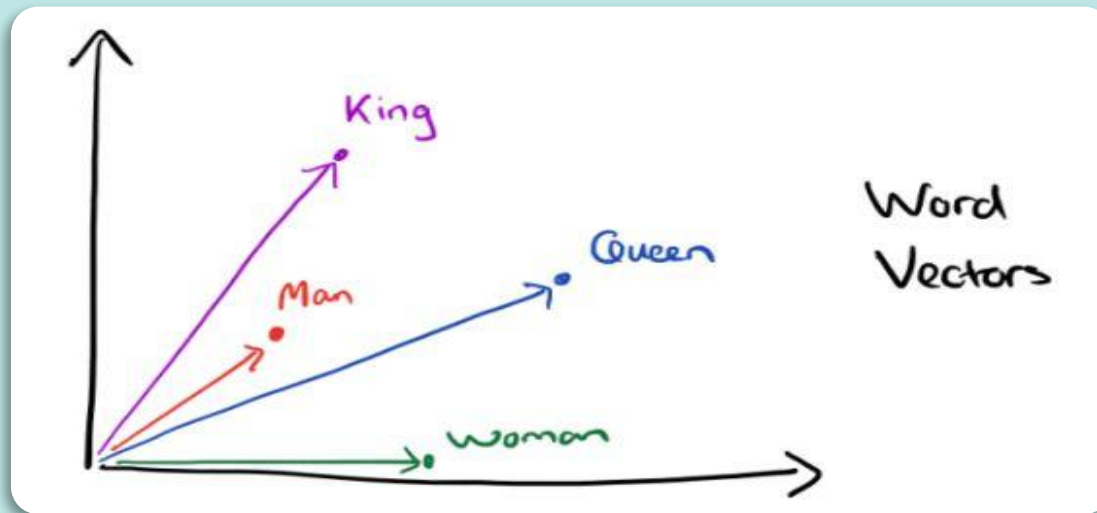
Word embedding to some space

Poincaré Embeddings for Learning Hierarchical Representations

PyTorch implementation of Poincaré Embeddings for Learning Hierarchical Representations

[Facebook]: Poincaré Embeddings

King – Man + Woman = Queen!!

King

Man

Queen

Woman

Word Vectors

03

Recent Topics

Topic-01
Elastic vs AWS

"Elasticsearch was created to put the power of data exploration in the hands of users."

Shay Banon, CEO Elasticsearch

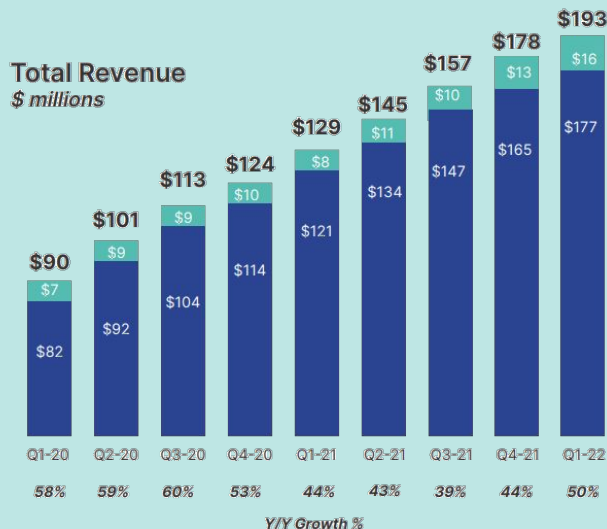[Meet a CEO: Shay Banon from Elastic]

Shay Banon: CTO(2012) -> CEO(2017) -> **CTO(2022)**
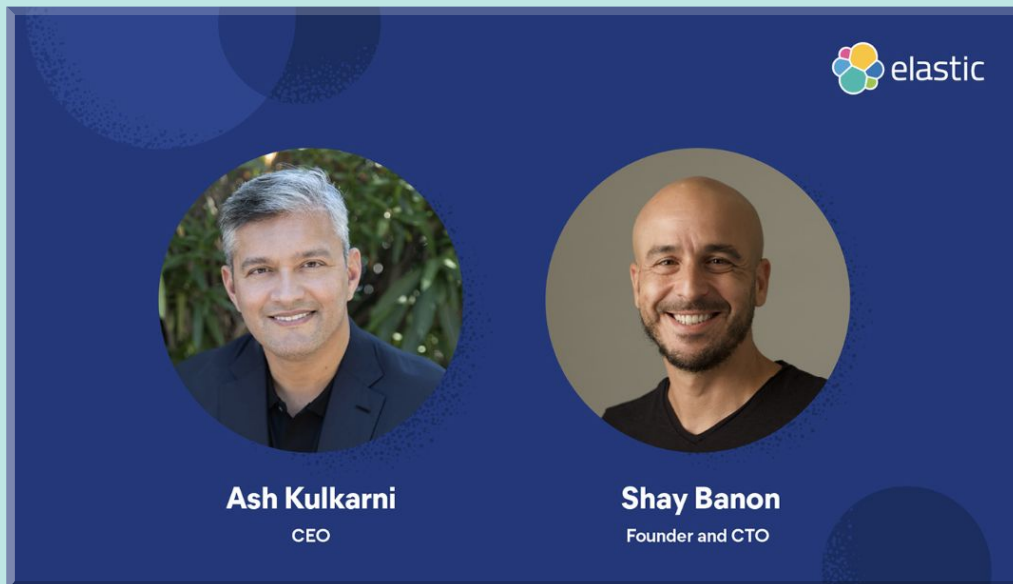
[Shay Banon to (re)assume the role of CTO,
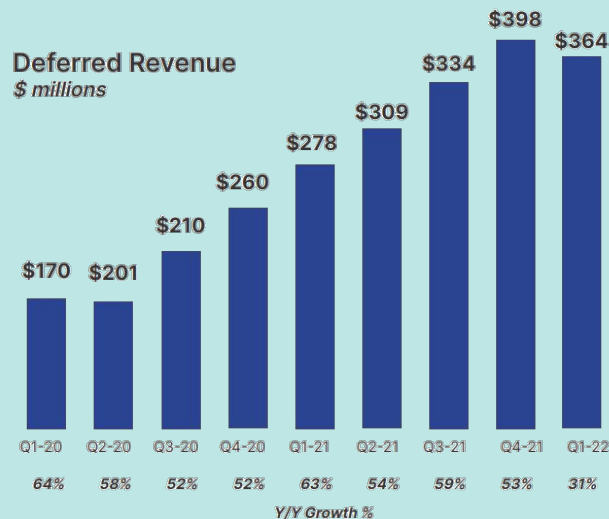Ash Kulkarni promoted to CEO (Jan.2022)]

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

Total Revenue
$ millions

	Q1-20	Q2-20	Q3-20	Q4-20	Q1-21	Q2-21	Q3-21	Q4-21	Q1-22
Total	$90	$101	$113	$124	$129	$145	$157	$178	$193
Services	$7	$9	$9	$10	$8	$11	$10	$13	$16
Subscription	$82	$92	$104	$114	$121	$134	$147	$165	$177
Y/Y Growth %	58%	59%	60%	53%	44%	43%	39%	44%	50%

■ Subscription   ■ Services

Sums may not add to totals due to rounding.

Deferred Revenue
$ millions

	Q1-20	Q2-20	Q3-20	Q4-20	Q1-21	Q2-21	Q3-21	Q4-21	Q1-22
Value	$170	$201	$210	$260	$278	$309	$334	$398	$364
Y/Y Growth %	64%	58%	52%	52%	63%	54%	59%	53%	31%

[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)

- **The purpose is to prevent AWS from offering Elasticsearch 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)

- <u>**Approximate k-nearest neighbor (A-kNN) searches**</u>
    - So far, only exact kNN, which is hard to scale
    - Large-scale vector-based searches can be performed faster

- <u>**Native support for NLP models**</u>
    - Starting with Elastic Stack 8.0, NLP using external models published on HuggingFace has become very easy to run

Approximate k-nearest neighbor (A-kNN) searches

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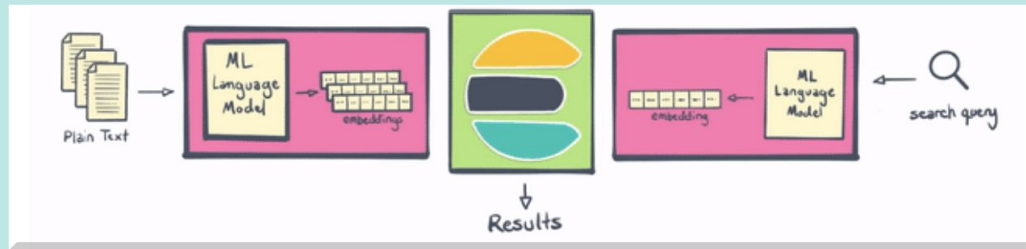- Large-scale vector-based searches can be performed faster

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

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

We can easily try the pretrained NLP tasks

(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 idendifier 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." }
}
```
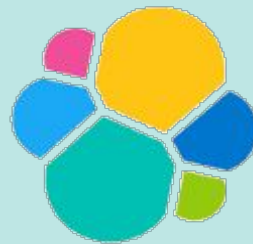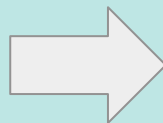
Bonus Slide

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

[Elastic CEO Shay Banon stepping down to become CTO,
with Ash Kulkarni as new CEO]

End of Documents