Elasticsearch Workshop for AI Engineer – ESRE for RAG use-cases

SSID: VTI – Workshop

Key: VTI@Gr3at2025

OpenAI API Key

sk-proj-b4bfD7\_654kfOa9TP3QvuFI9lRgw7-4MUFphJKG6f3pams90PbRqqZaMPq2Fherv7pVOlOcQLJT3BlbkFJKGImqyQMISGT9nmX5hlACEmdxRQL8hQI9vrdsx\_LaqwjZu948y0Qgx7kO75YWR6KqjtLjR7FwA

Pre-reqs:

* PyCharm (Python Editor)
* JDK (version 11 or later and properly configured JAVA\_HOME)
* Docker 28.0.1

## Create Deployment

1 . Akses Elastic Cloud pada [cloud.elastic.co](https://www.elastic.co/docs/explore-analyze/machine-learning/nlp/ml-nlp-model-ref) menggunakan email untuk membuat akun baru. Sesuaikan dengan kebutuhan anda. *Screenshot terlampir*

A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

A screenshot of a computer

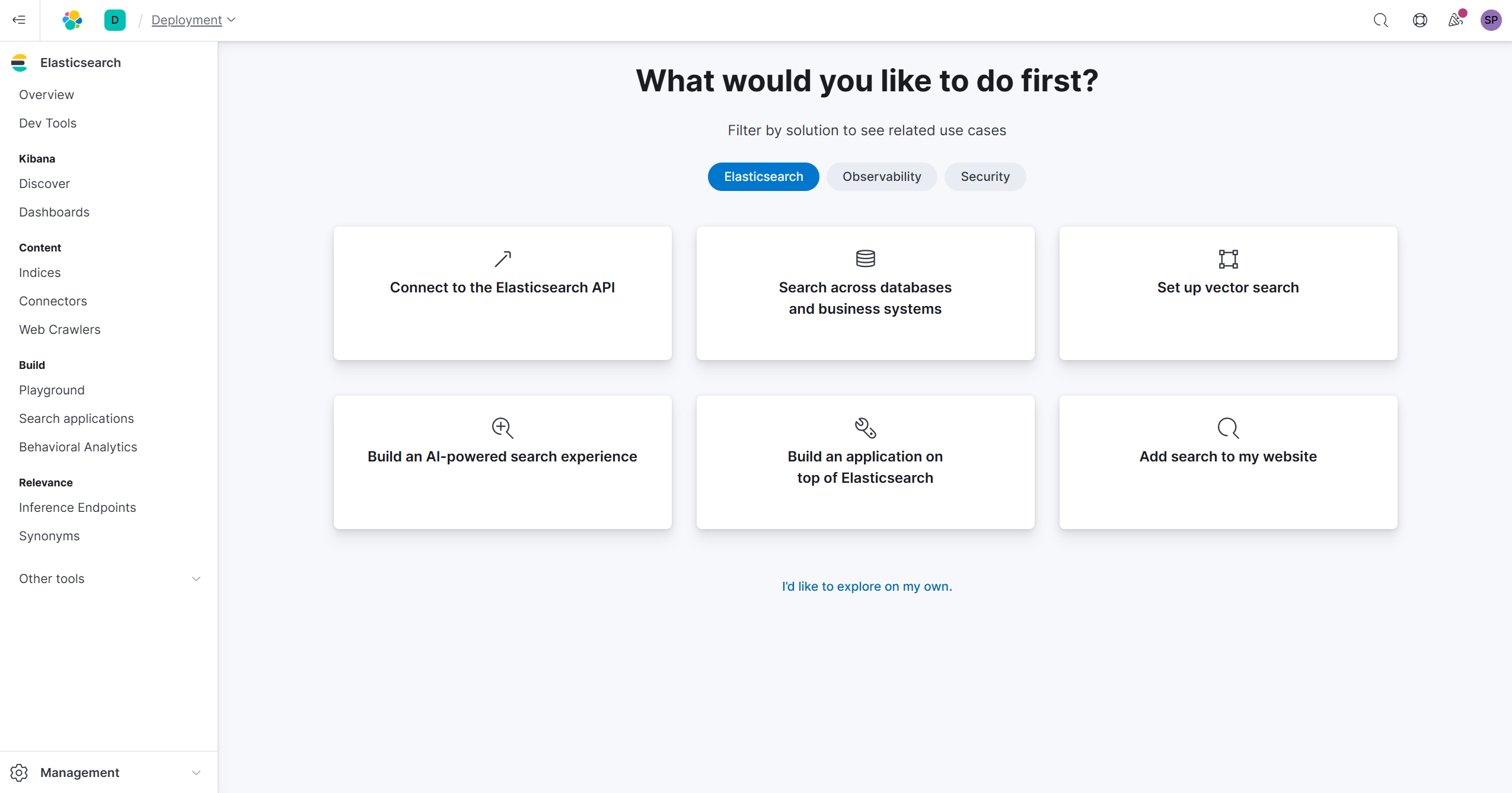
AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a video

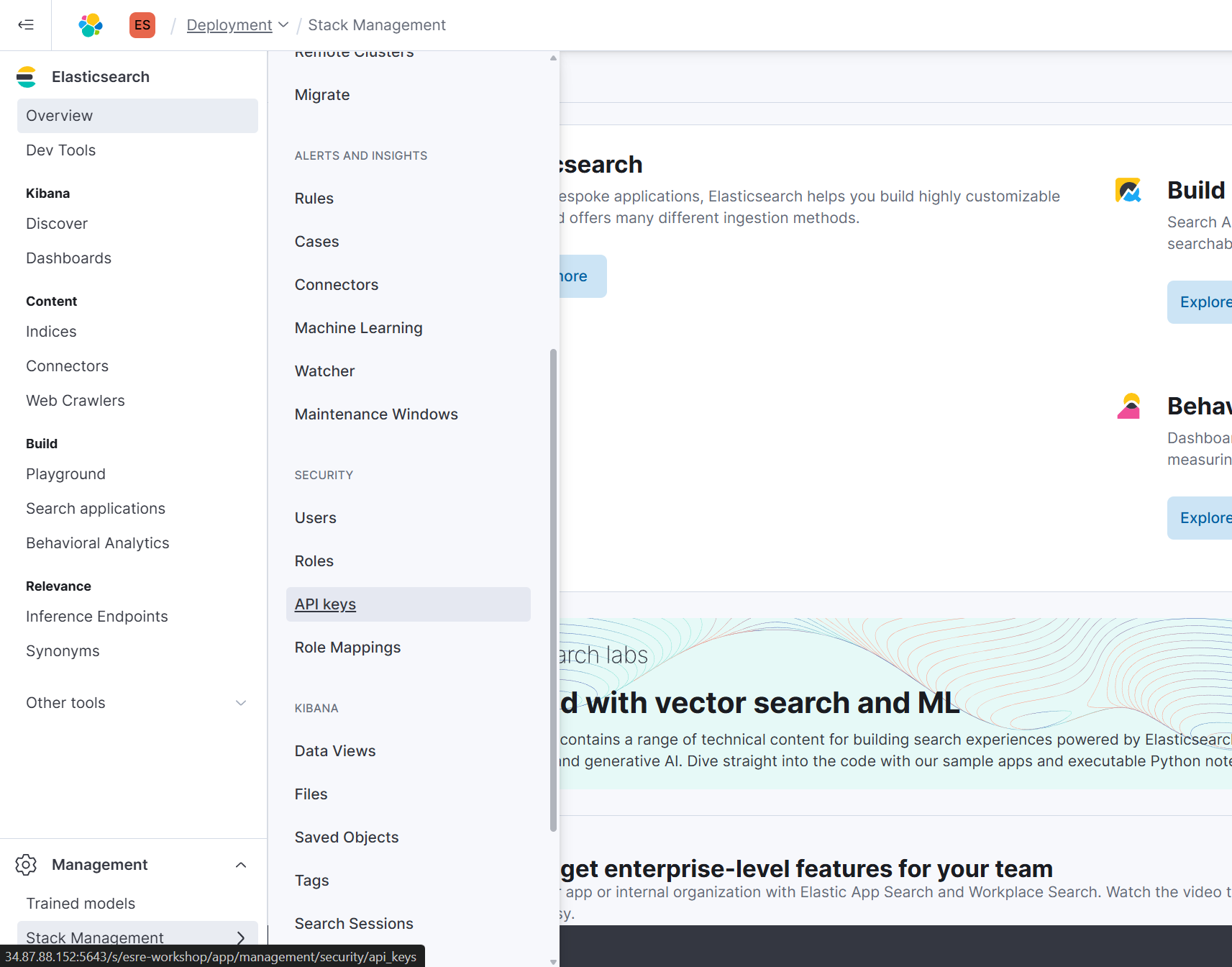
AI-generated content may be incorrect.



2. Create Elasticsearch API Key

API key akan digunakan untuk mengakses cluster Elastic tanpa menggunakan username dan password.

**Stack Management > Security > API keys**



Create a new API key

A screenshot of a application

AI-generated content may be incorrect.

Simpan API key dikarenakan value tidak dapat dilihat lagi setelah halaman ditutup

A screenshot of a computer

AI-generated content may be incorrect.

## File System Crawler (FSCrawler)

Download Link

<https://central.sonatype.com/repository/maven-snapshots/fr/pilato/elasticsearch/crawler/fscrawler-distribution/2.10-SNAPSHOT/fscrawler-distribution-2.10-20250519.193441-44.zip>

Setup job fscrawler baru menggunakan command berikut

1. .\fscrawler.bat –-setup <job\_name>

A screenshot of a computer screen

AI-generated content may be incorrect.

Konfigurasi job tersimpan pada user directory anda

A screenshot of a computer

AI-generated content may be incorrect.

Sesuaikan \_settings.yaml dengan kebutuhan

1. name: "general-rules"

2. fs:

3. url: "D:\\Libraries\\Documents\\SampleText\\general-rules"

4. update\_rate: "15m"

5. remove\_deleted: true

6. continue\_on\_error: true

7. ignore\_above: "10mb"

8. filename\_as\_id: true

9. lang\_detect: true

10. store\_source: false

11. index\_folders: false

12.

13. ocr:

14. enabled: "false"

15. language: "eng"

16. path: "C:\\Program Files\\Tesseract-OCR"

17. data\_path: "C:\\Program Files\\Tesseract-OCR\\tessdata"

18. output\_type: txt

19. pdf\_strategy: ocr\_and\_text

20.

21. elasticsearch:

22. nodes:

23. - url: "<ES\_URL>"

24. bulk\_size: 1000

25. flush\_interval: "5s"

26. byte\_size: "10mb"

27. api\_key: "<API\_KEY>"

28. ssl\_verification: "false"

Deploy model yang akan digunakan

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

Buat index mapping baru yang menggunakan inference id berdasarkan model yang sudah di-deploy pada step sebelumnya

1. PUT general-rules

2. {

3.   "mappings": {

4.     "properties": {

5.       "content": {

6.         "type": "text",

7.         "fields": {

8.           "semantic": {

9.             "type": "semantic\_text",

10.             "inference\_id": ".multilingual-e5-small-elasticsearch"

11.           }

12.         }

13.       }

14.     }

15.   }

16. }

Setelah mapping berhasil dibuat, jalankan fscrawler job yang sudah di-setup sebelumnya

1. .\fscrawler.bat <job name>

A screen shot of a computer

AI-generated content may be incorrect.

Setelah fscrawler sukses dijalankan, pencarian dokumen bisa dilakukan dengan query berikut

1. GET general-rules/\_search

2. {

3.   "query": {

4.     "semantic": {

5.       "field": "content.semantic",

6.       "query": "ketentuan mendapatkan bintang?"

7.     }

8.   },

9.   "highlight": {

10.     "fields": {

11.       "content.semantic": {

12.         "order": "score",

13.         "number\_of\_fragments": 1

14.       }

15.     }

16.   }

17. }

18.

## Import 3rd party model menggunakan Eland

Referensi model dan turunannya yang dapat digunakan terdapat pada <https://www.elastic.co/docs/explore-analyze/machine-learning/nlp/ml-nlp-model-ref>

Model tersebut dapat diunduh dari halaman web huggingface

1. docker run -it --rm docker.elastic.co/eland/eland eland\_import\_hub\_model --url https://xyperia:14041212@34.87.88.152:9243 --hub-model-id LazarusNLP/all-nusabert-base-v4 --task-type text\_embedding --insecure --start

A screen shot of a computer

AI-generated content may be incorrect.

A screen shot of a computer

AI-generated content may be incorrect.

Setelah model berhasil di-import, synchronizing perlu dilakukan terlebih dahulu agar model bisa muncul di halaman **Trained Models**

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.