OPENSEARCH
An analysis of the tool's key features

Group members

Ismael Liuzzi

Sofia Scattolini

Supervisor

Prof. Massimo Callisto de Donato

Contents

1. Project Description

OpenSearch is a community-driven, open source search and analytics platform that makes it easy to ingest, search, visualize, and analyze data.

1. Project Description

OpenSearch is a community-driven, open source search and analytics platform that makes it easy to ingest, search, visualize, and analyze data.

Full-Text Search

1. Project Description

OpenSearch is a community-driven, open source search and analytics platform that makes it easy to ingest, search, visualize, and analyze data.

Full-Text Search

Security

1. Project Description

OpenSearch is a community-driven, open source search and analytics platform that makes it easy to ingest, search, visualize, and analyze data.

Full-Text Search

Security

Analytics and machine learning

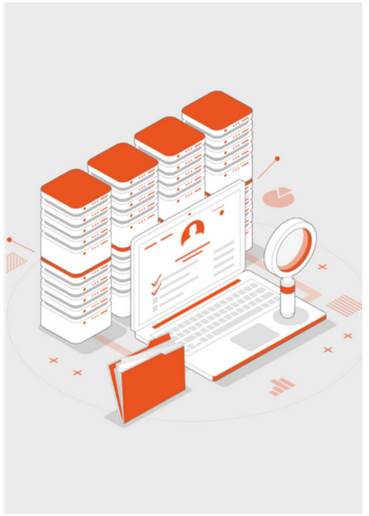
1. Project Description

OpenSearch is a community-driven, open source search and analytics platform that makes it easy to ingest, search, visualize, and analyze data.

Full-Text Search

Security

Analytics and machine learning



Observability

1. Project Description

Why OpenSearch?

💡 It provides an alternative to proprietary search engines like Elasticsearch while remaining fully open-source and extensible.

1. Project Description

Objectives:

1. Observability & Dashboard: Explore how OpenSearch collects logs, metrics, and traces for system monitoring.
2. OpenTelemetry Integration: Understand how OpenTelemetry helps standardize telemetry data.
3. AI & ML Capabilities: Look into OpenSearch's features for machine learning and AI.
4. Hands-on Implementation and Testing: Set up OpenSearch, configure observability pipelines, and test AI/ML models.

2. Setting Up OpenSearch

with Docker Compose

🛠️ Prerequisites

⚙️ Setup Instructions

1 → Obtain the docker-compose.yml file

2 → Edit .env or docker-compose.yml to set environment variables (e.g., admin, password)

3 → Start OpenSearch with docker-compose up -d

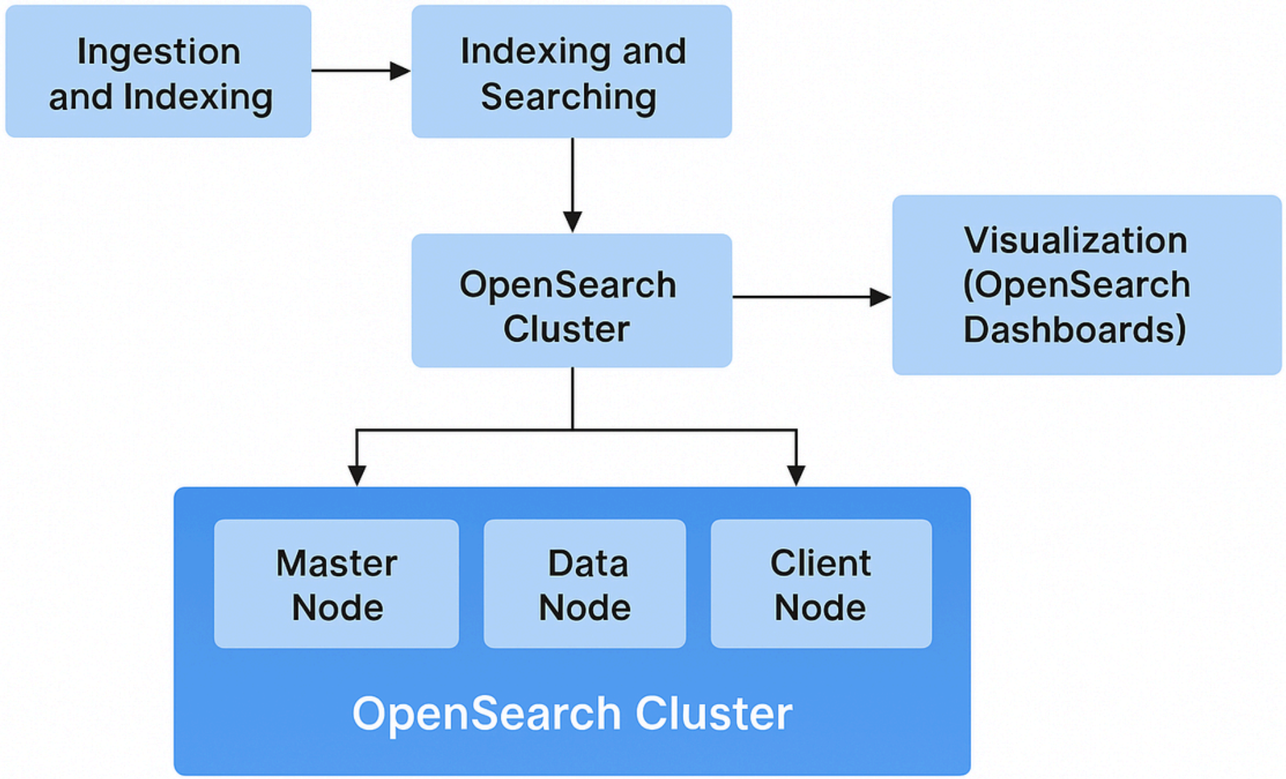
4 → Verify OpenSearch is running at https://localhost:9200

5 → Access OpenSearch Dashboards at https://localhost:5601 and log in (default: admin/admin)

6 Stop and remove the containers with docker-compose down

2. Setting Up OpenSearch

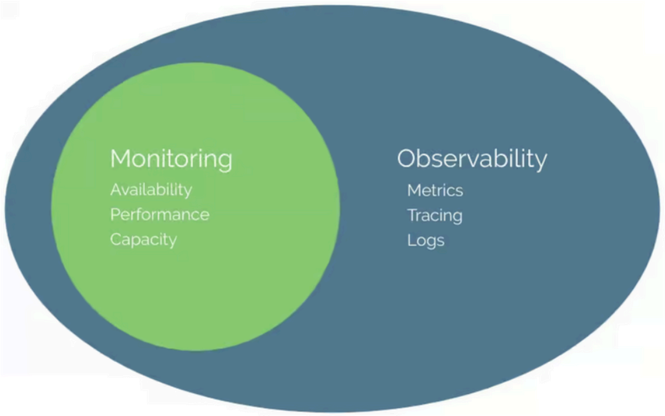
Understanding architecture



3. What is observability?

Observability is the ability of a system to make its internal state understandable, allowing developers to detect, diagnose and resolve issues affecting the performance, scalability or availability of software and infrastructure based on the data it generates.

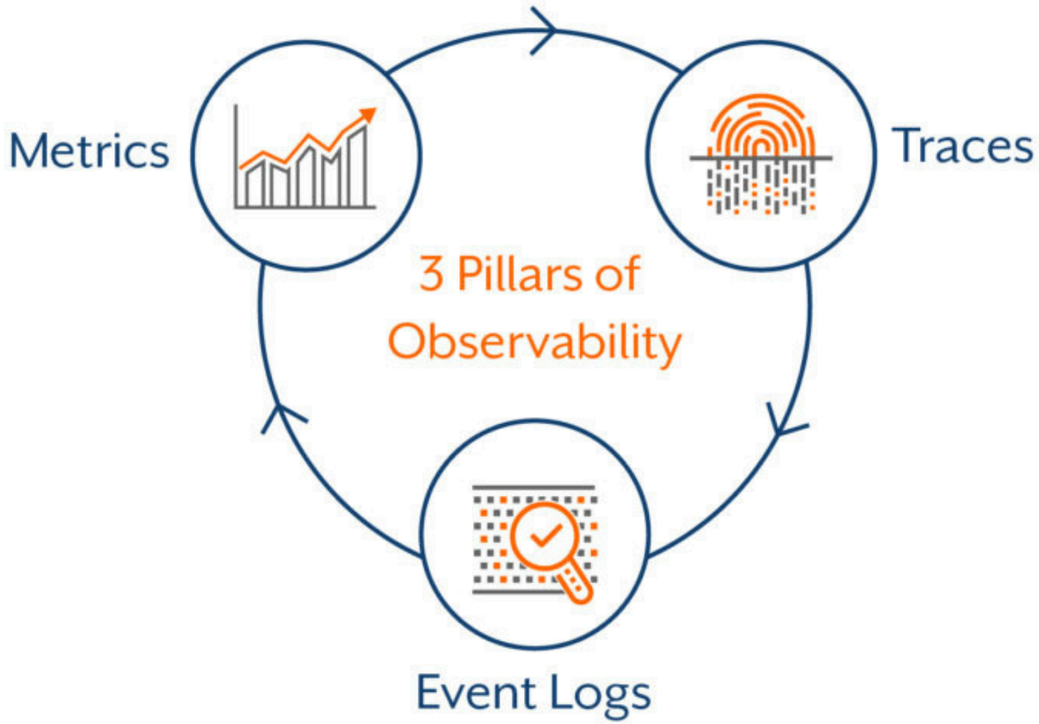
🔍 Monitoring tells us what is happening; observability tells us why.



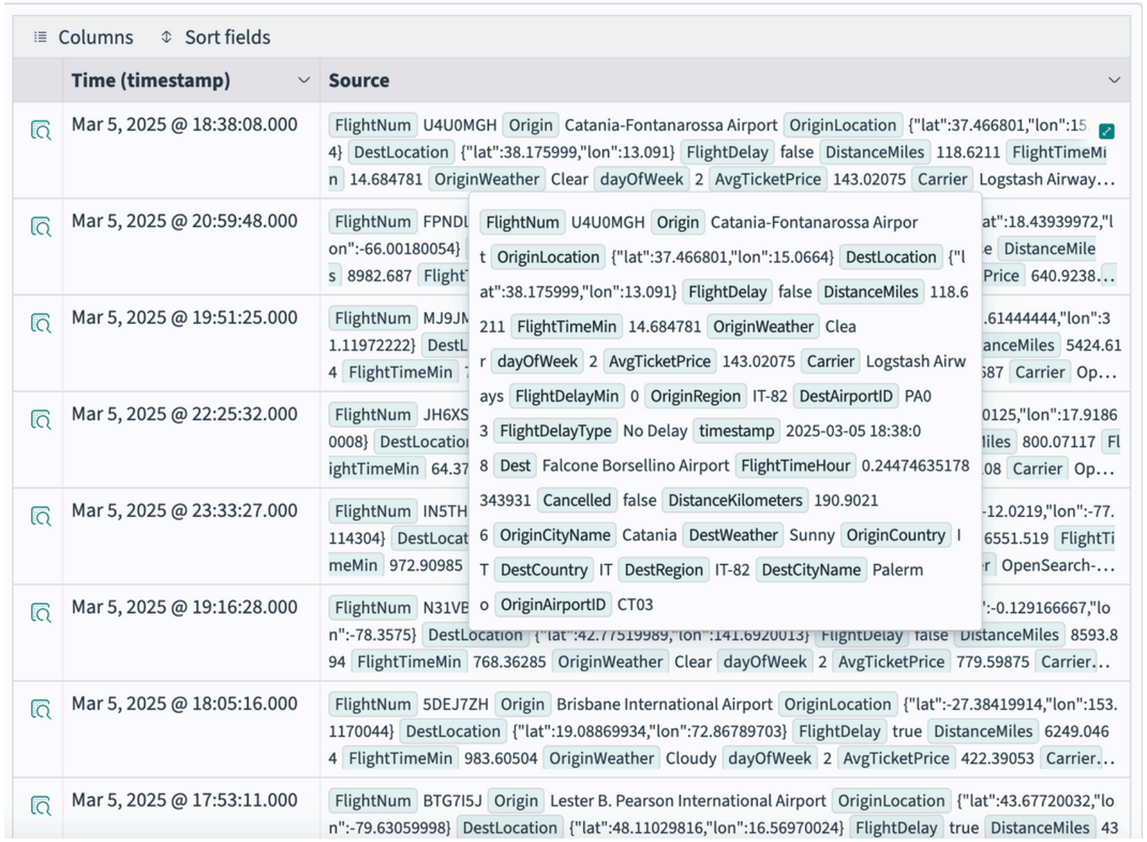
💡 *Why is it important?*

- Enables diagnosis of performance issues and failures.
- Provides real-time insights into system behavior.
- Essential for distributed architectures, microservices and IoT systems.

3. What is observability?

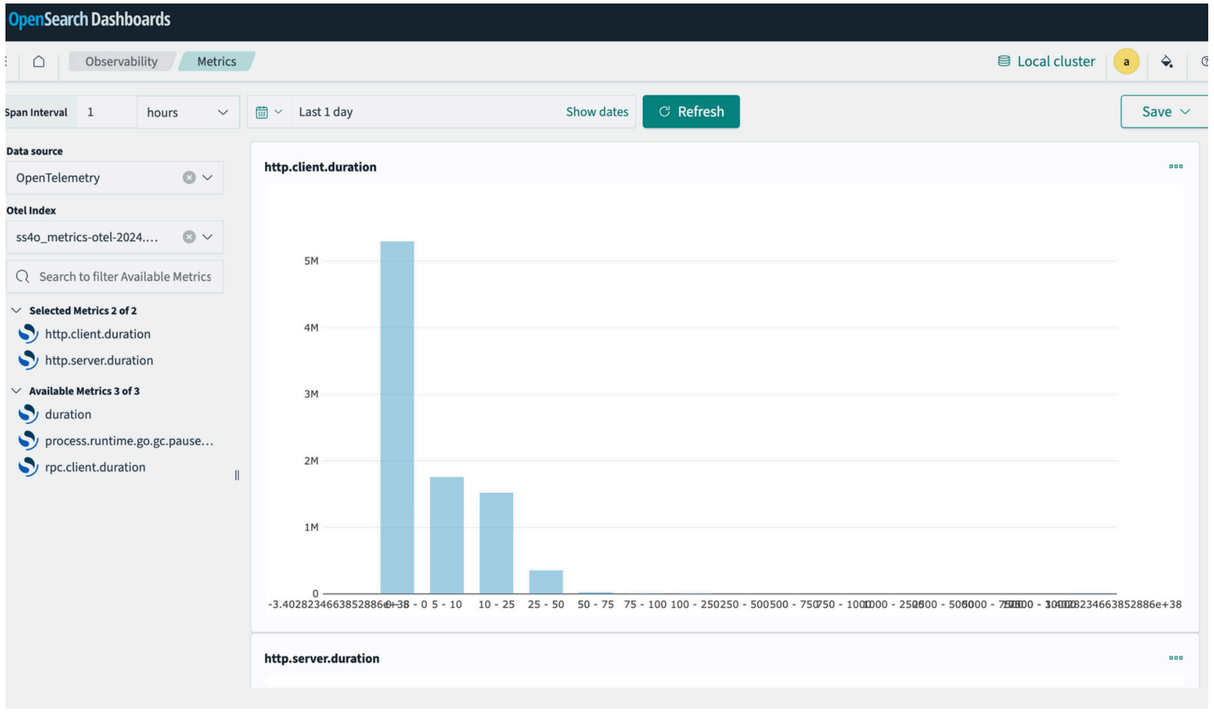


3. What is observability?



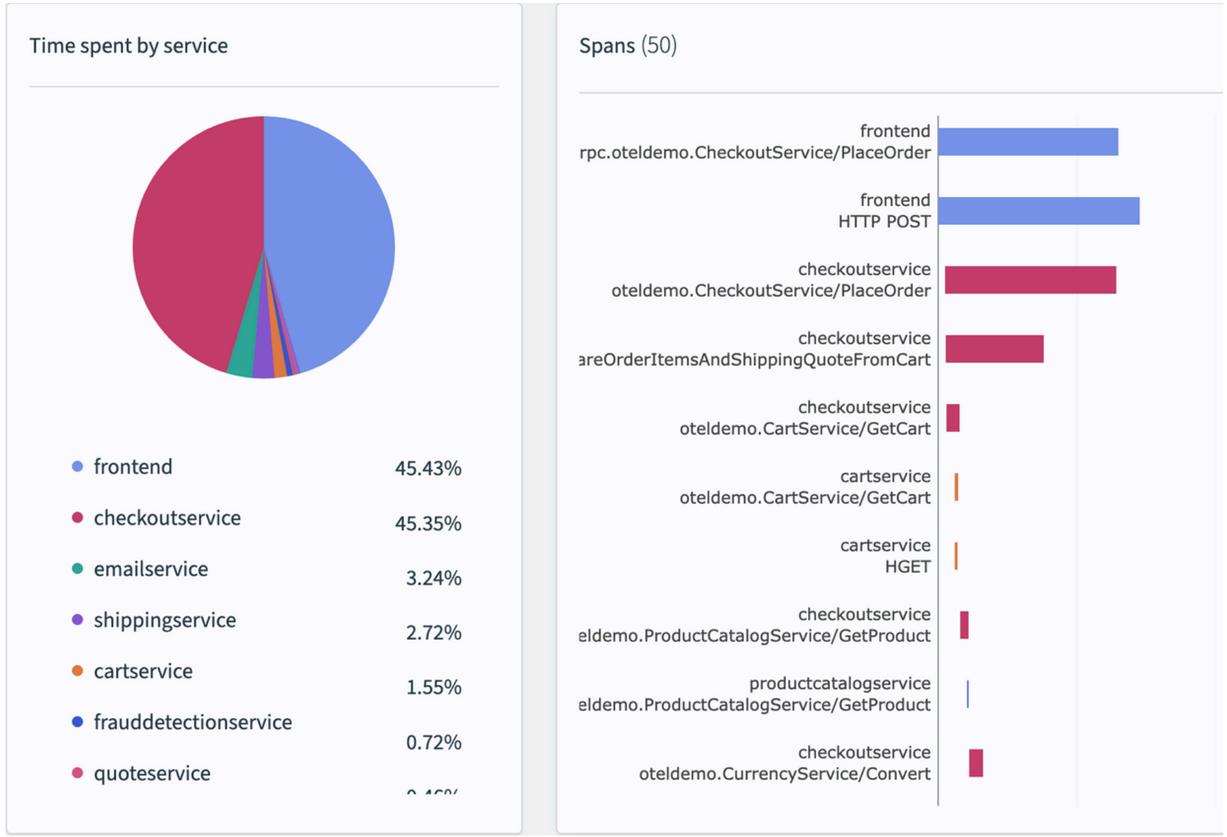
1. Logs: Text-based event records describing system actions (understand what happened and why).

3. What is observability?



1. Logs: Text-based event records describing system actions (understand what happened and why).
2. Metrics: Numerical data that measure system performance (e.g., CPU usage, response times).

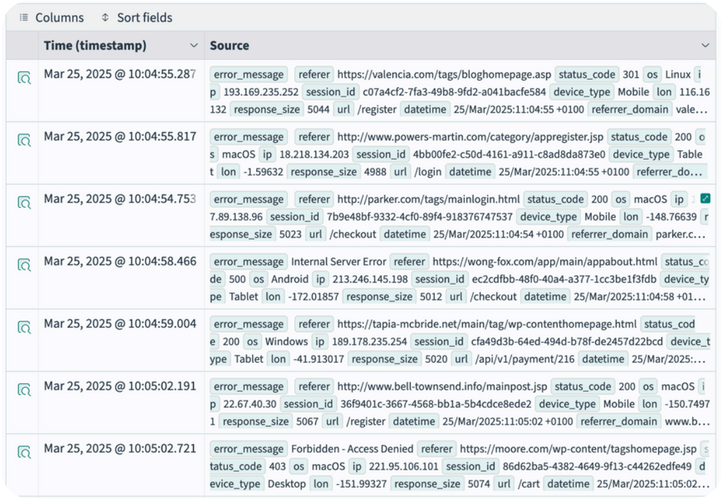
3. What is observability?



1. Logs: Text-based event records describing system actions (understand what happened and why).
2. Metrics: Numerical data that measure system performance (e.g., CPU usage, response times).
3. Traces: Track data flows of a request through a distributed system to see how everything functions together.

3. What is observability?

Technical implementation:
Fake Log Generator

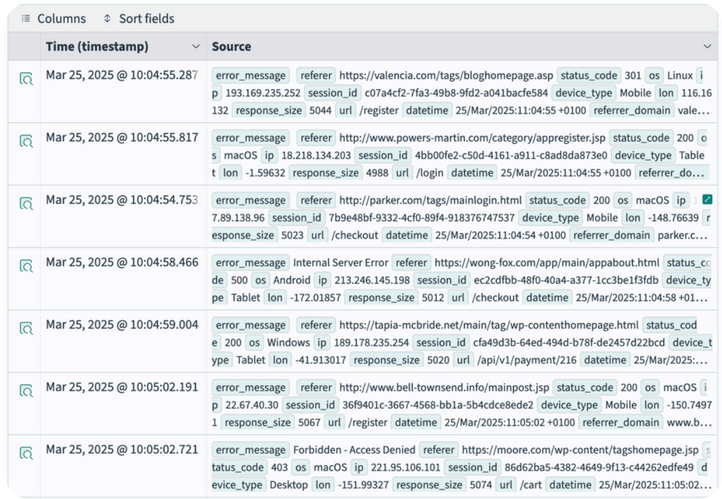

Columns Sort fields

Time (timestamp)	Source
Mar 25, 2025 @ 10:04:55.287	error_message referer https://valencia.com/tags/bloghomepage.asp status_code 301 os Linux ip 193.169.235.252 session_id c07a4cf2-7fa3-49b8-9fd2-a041bacfe584 device_type Mobile lon 116.16132 response_size 5044 url /register datetime 25/Mar/2025:11:04:55 +0100 referrer_domain vale…
Mar 25, 2025 @ 10:04:55.817	error_message referer http://www.powers-martin.com/category/appregister.jsp status_code 200 os macOS ip 18.218.134.203 session_id 4bb00fe2-c50d-4161-a911-c8ad8da873e0 device_type Tablet lon -1.59632 response_size 4988 url /login datetime 25/Mar/2025:11:04:55 +0100 referrer_do…
Mar 25, 2025 @ 10:04:54.753	error_message referer http://parker.com/tags/mainlogin.html status_code 200 os macOS ip 7.89.138.96 session_id 7b9e48bf-9332-4cf0-89f4-918376747537 device_type Mobile lon -148.76639 response_size 5023 url /checkout datetime 25/Mar/2025:11:04:54 +0100 referrer_domain parker.c…
Mar 25, 2025 @ 10:04:58.466	error_message Internal Server Error referer https://wong-fox.com/app/main/appabout.html status_code 500 os Android ip 213.246.145.198 session_id ec2cdfbb-48f0-40a4-a377-1cc3be1f3fdb device_type Tablet lon -172.01857 response_size 5012 url /checkout datetime 25/Mar/2025:11:04:58 +01…
Mar 25, 2025 @ 10:04:59.004	error_message referer https://tapia-mcbride.net/main/tag/wp-contenthomepage.html status_code 200 os Windows ip 189.178.235.254 session_id cfa49d3b-64ed-494d-b78f-de2457d22bcd device_type Tablet lon -41.913017 response_size 5020 url /api/v1/payment/216 datetime 25/Mar/2025:…
Mar 25, 2025 @ 10:05:02.191	error_message referer http://www.bell-townsend.info/mainpost.jsp status_code 200 os macOS ip 22.67.40.30 session_id 36f9401c-3667-4568-bb1a-5b4cdce8ede2 device_type Mobile lon -150.74971 response_size 5067 url /register datetime 25/Mar/2025:11:05:02 +0100 referrer_domain www.b…
Mar 25, 2025 @ 10:05:02.721	error_message Forbidden - Access Denied referer https://moore.com/wp-content/tagshomepage.jsp status_code 403 os macOS ip 221.95.106.101 session_id 86d62ba5-4382-4649-9f13-c44262edfe49 device_type Desktop lon -151.99327 response_size 5074 url /cart datetime 25/Mar/2025:11:05:02…

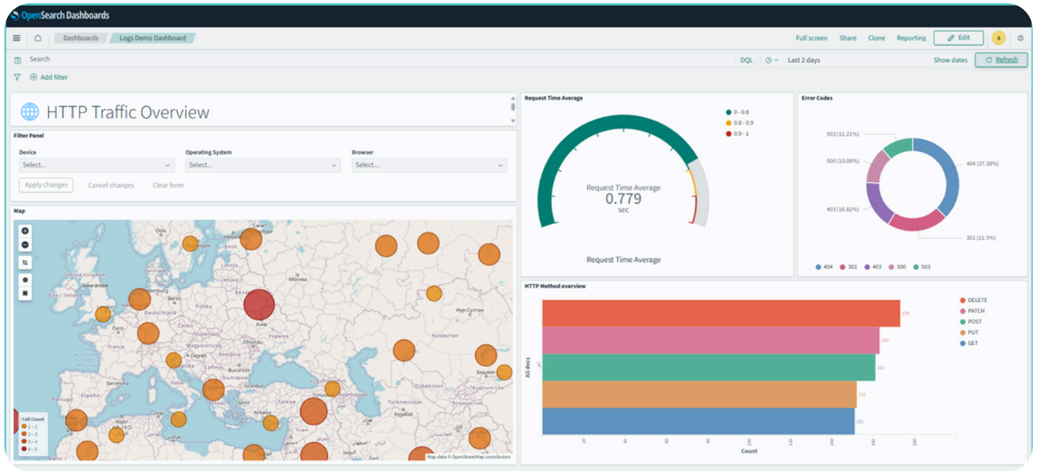
3. What is observability?

Technical implementation:
Fake Log Generator

“FakeLogger”
Python Script



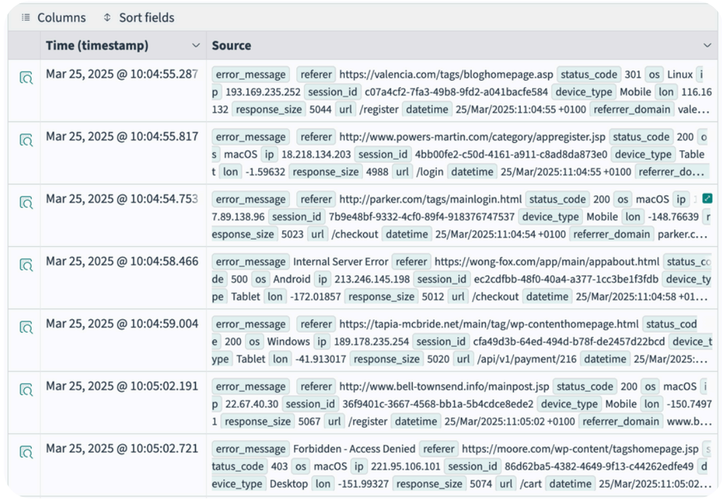
Dashboard



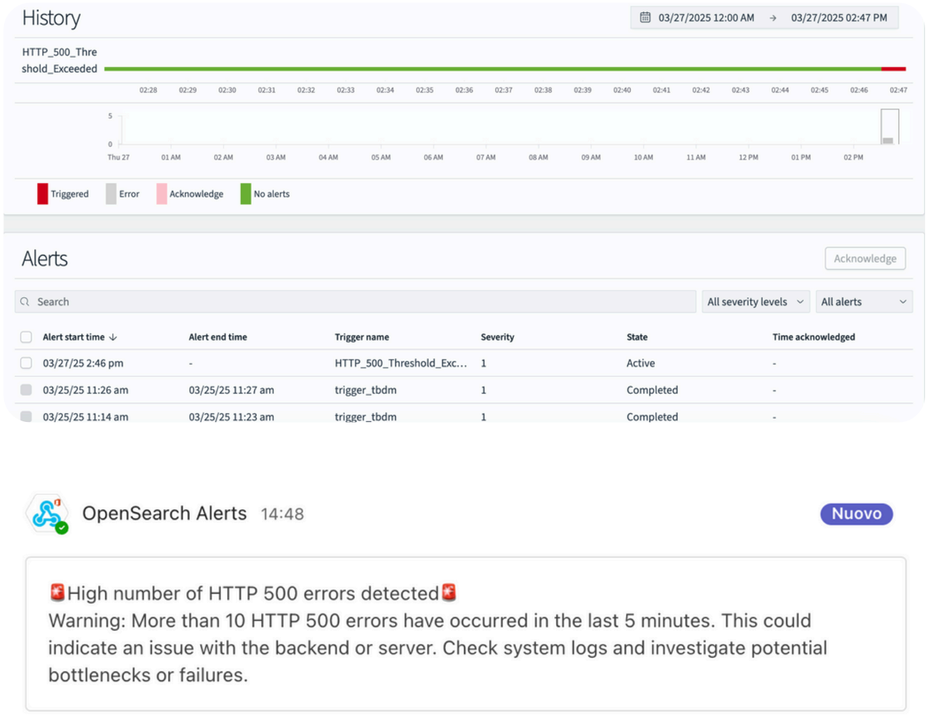
3. What is observability?

Technical implementation:
Fake Log Generator

“FakeLogger”
Python Script



Alerting & Monitoring



4. OpenTelemetry: A Standard for Observability

Observability relies on telemetry data (logs, metrics, traces) to provide a comprehensive view of a system's behavior.

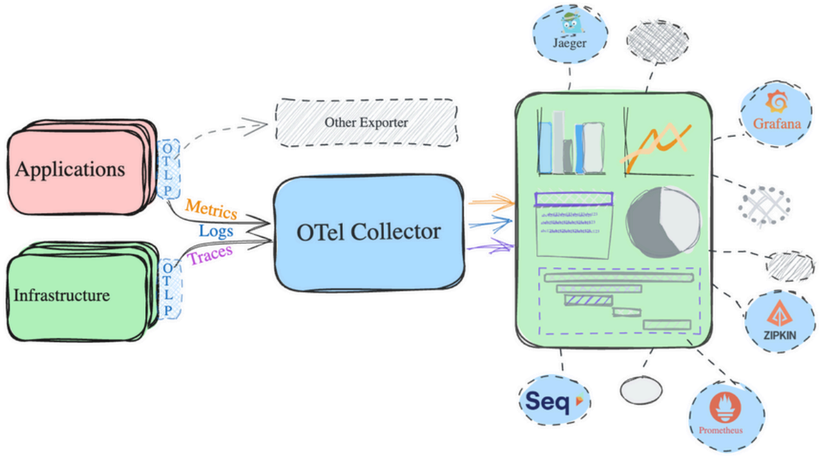
However, collecting and managing telemetry data from different services can be complex. This is where OpenTelemetry helps by providing a standardized framework for data collection.

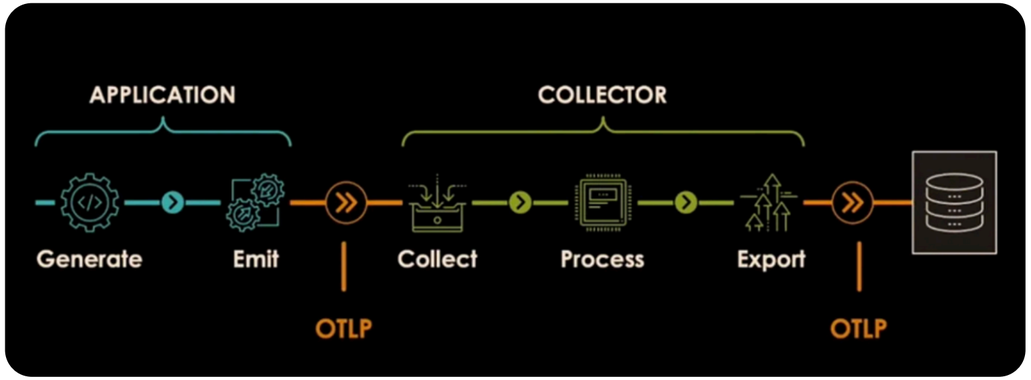
4. OpenTelemetry: A Standard for Observability

OpenTelemetry (OTel) is an open source observability framework that provides a unified standard for collecting and exporting telemetry data from different sources, simplifying integration with various back-end tools.

This means you can use the same tool to collect and analyze data from different sources.

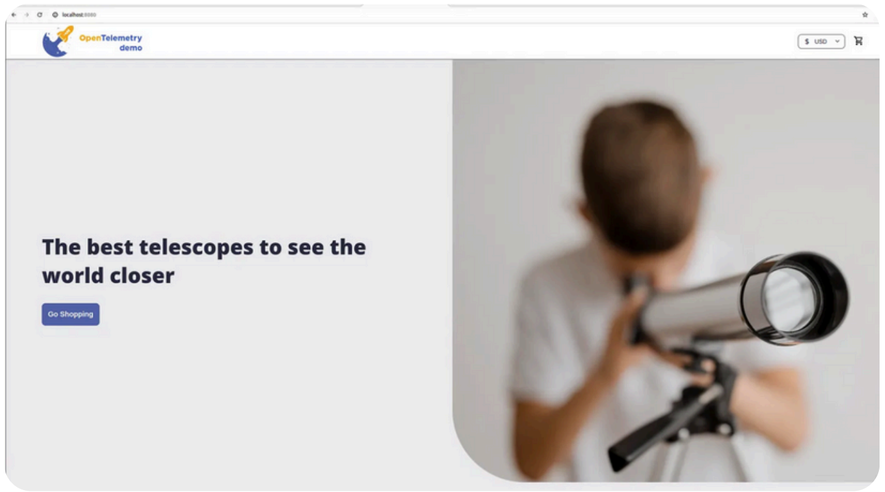
How does it works?





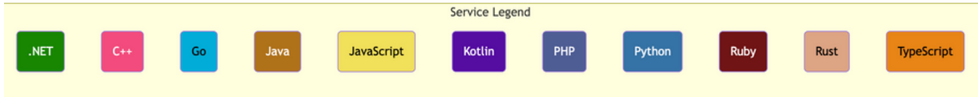
4. OpenTelemetry: A Standard for Observability

Technical implementation: opentelemetry-demo



https://github.com/open-telemetry/opentelemetry-demo

Core demo services



Dependent Services

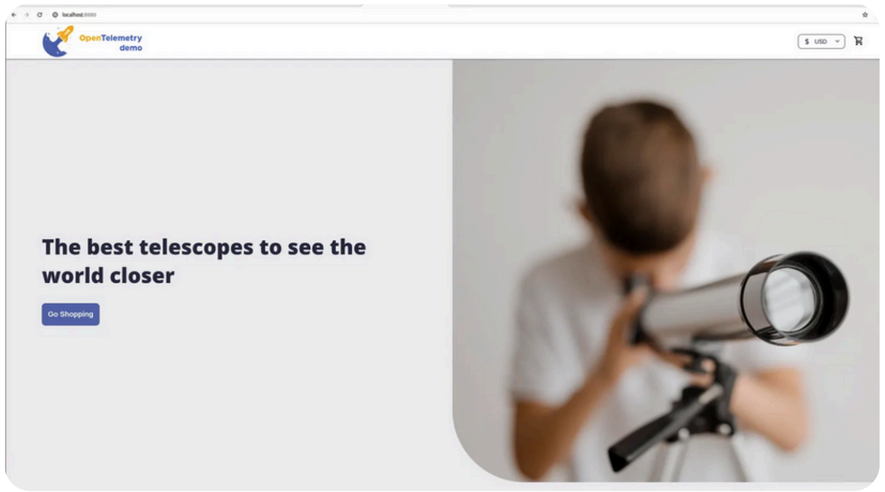
redis kafka Valkey

Telemetry Components

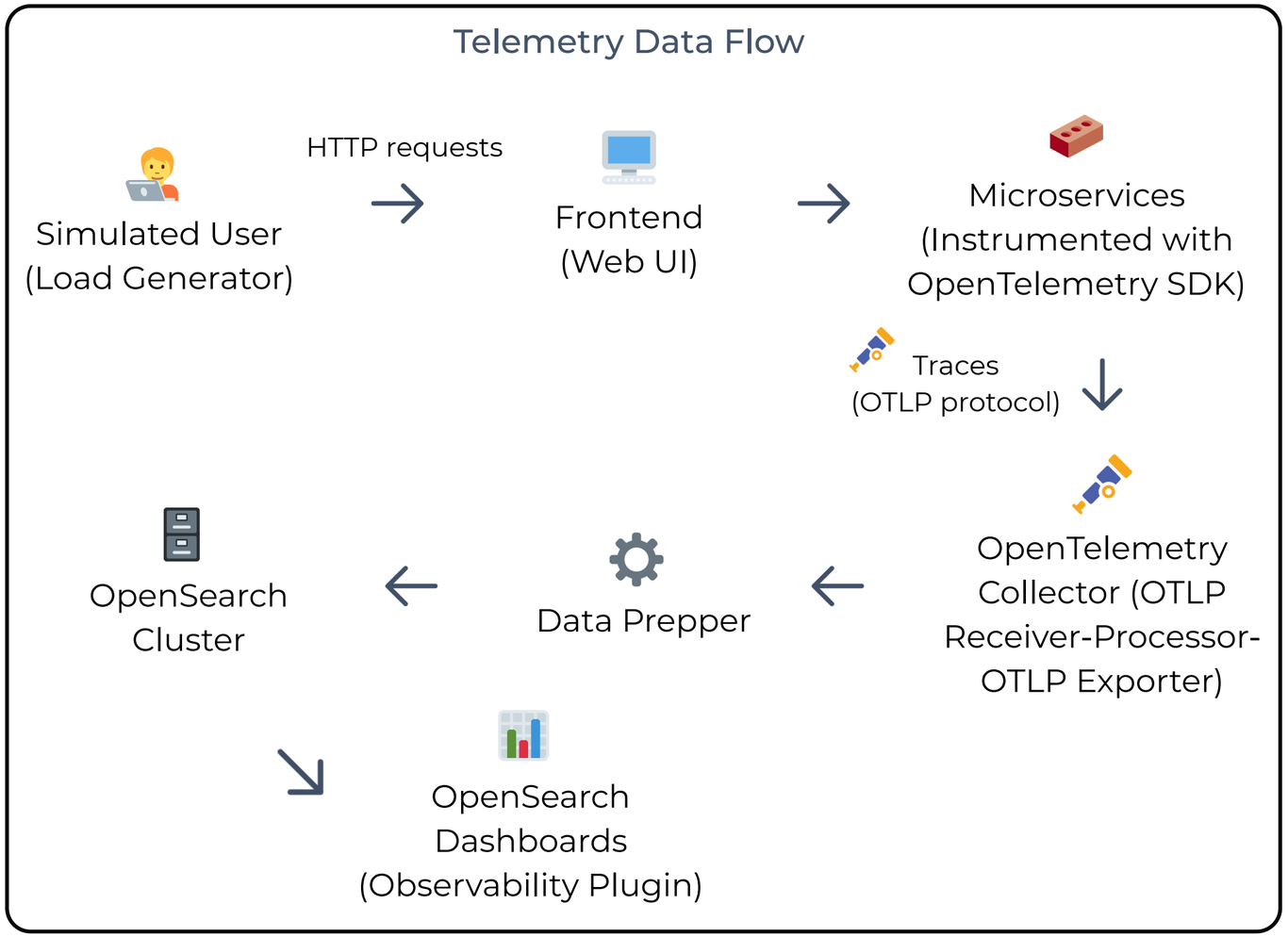
Utility services

4. OpenTelemetry: A Standard for Observability

Technical implementation: opentelemetry-demo

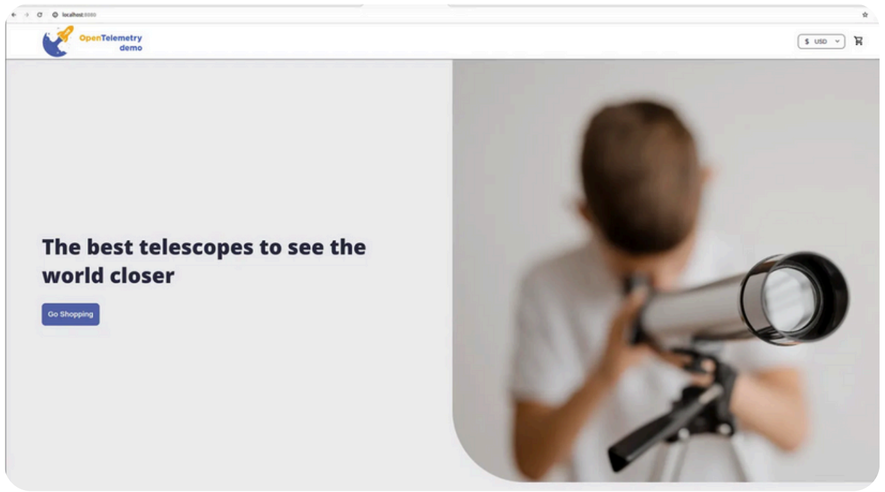


https://github.com/open-telemetry/opentelemetry-demo

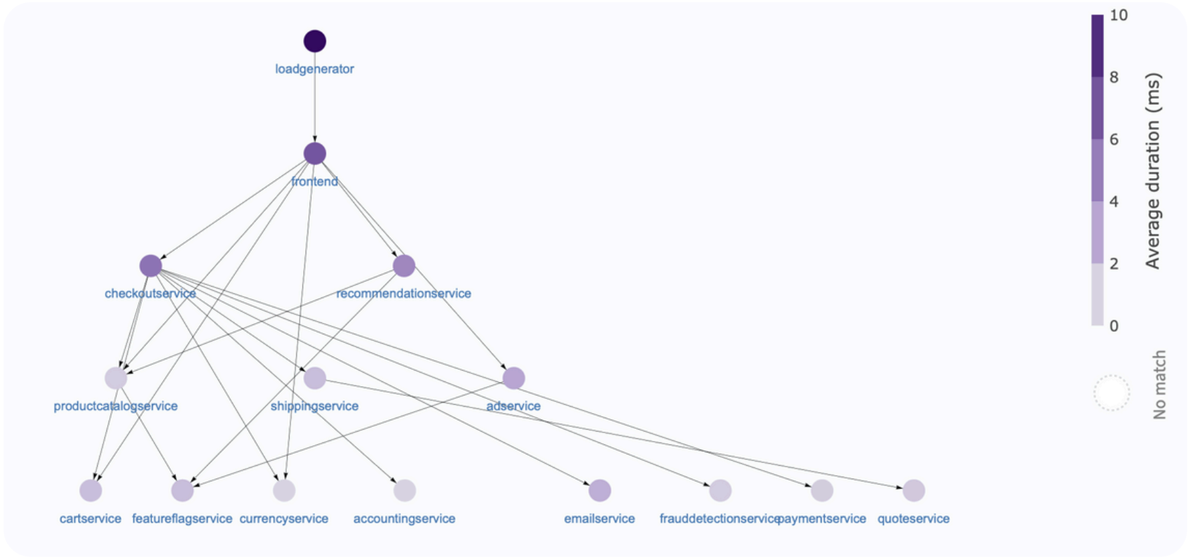


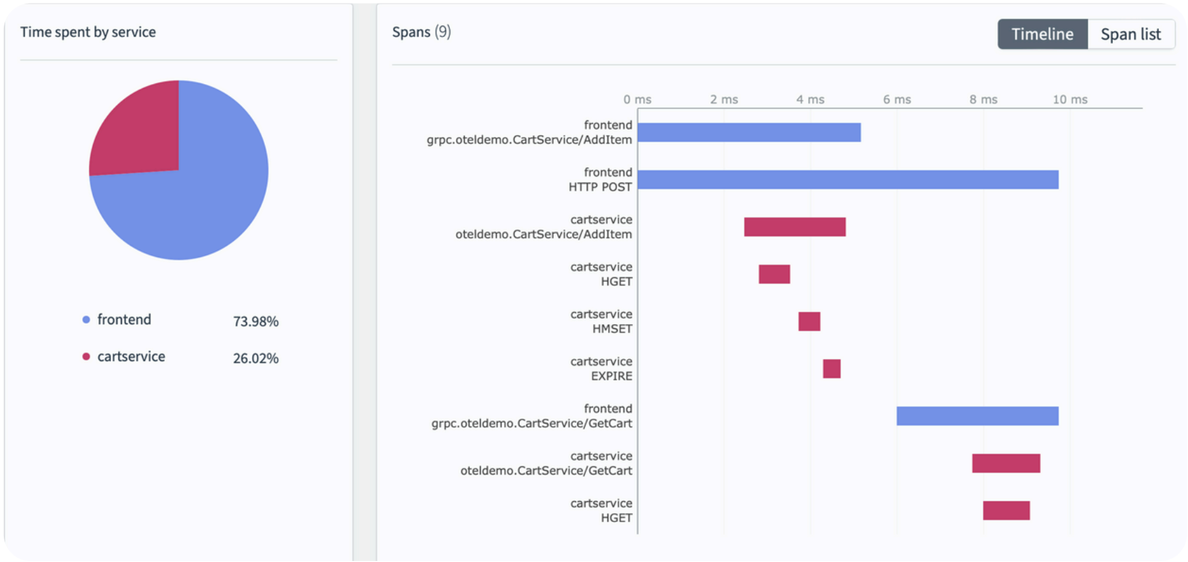
4. OpenTelemetry: A Standard for Observability

Technical implementation: opentelemetry-demo



https://github.com/open-telemetry/opentelemetry-demo





5. AI & ML

Anomaly Detection

- Detects anomalies in time-series data

ML Commons Plugin

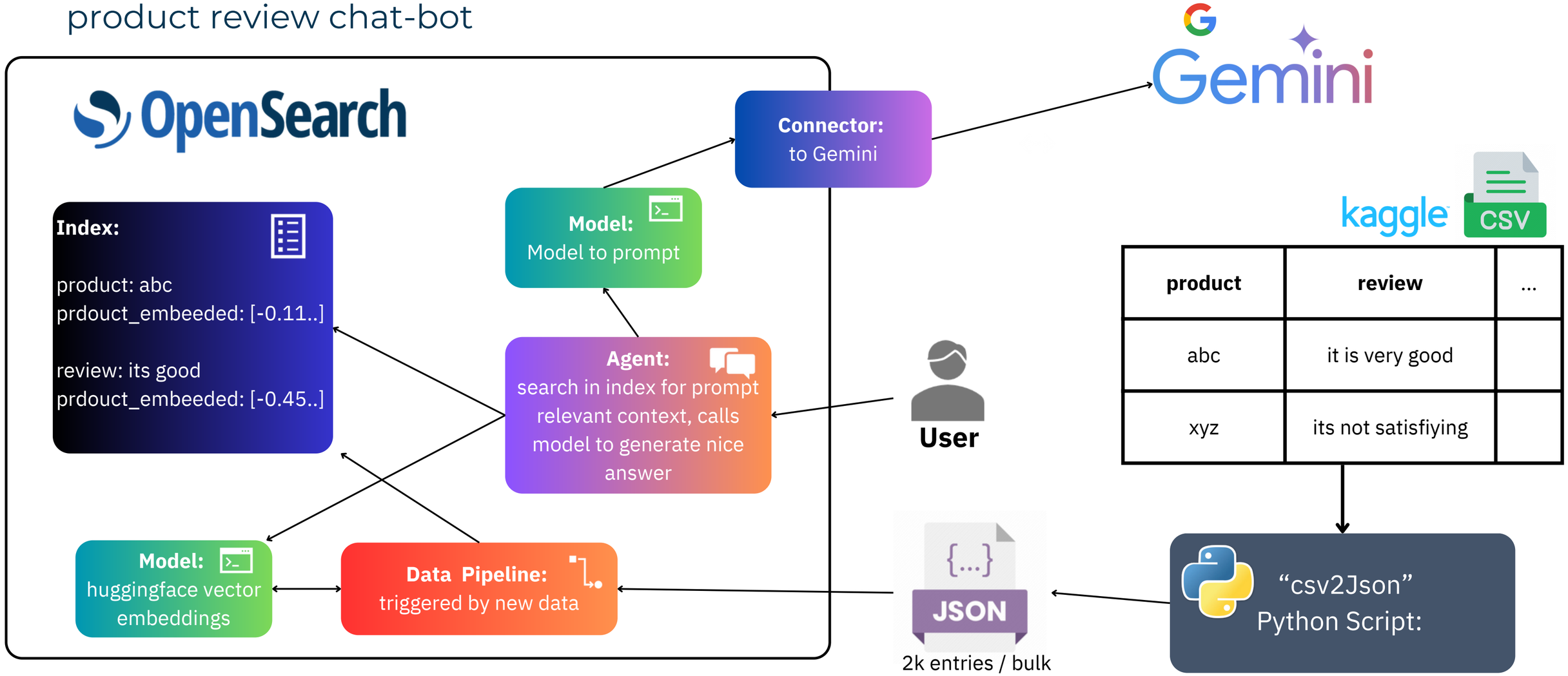
- Train, deploy, and run custom ML / AI models

Use Cases

  - Detect unusual system behavior
  - Predict trends (e.g., flight delays)
  - AI Chatbot (on indices)

5. AI & ML

Technical implementation: product review chat-bot



5. AI & ML

Technical implementation: product review chat-bot

Prompt by user

```
POST /_plugins/_ml/agents/7TWbuZUBcUgXhD6ry6DP/_execute
{
  "parameters": {
    "question": "Which product is the hail grail for some commenter?"
  }
}
```

Result by Model


```
"inference_results": [
  {
    "output": [
      {
        "name": "memory_id",
        "result": "qiylvpUBwdZjFirnUAEt"
      },
      {
        "name": "parent_message_id",
        "result": "qyylvpUBwdZjFirnUAFl"
      },
      {
        "name": "bedrock_claude_model",
        "result": """{"candidates":[{"content":{"parts":[{"text":"Based on the provided text, the Kandee Pop Glam Mermaid
          Dust Dry Shampoo is considered the \"best on the market\" by some commenters.  Therefore, this product appears to
          be the \"grail\" for at least some reviewers.\n"}],"role":"model"},"finishReason":"STOP","avgLogprobs":-0
          .029779279232025148}],"usageMetadata":{"promptTokenCount":321.0,"candidatesTokenCount":50.0,"totalTokenCount":371.0
          ,"promptTokensDetails":[{"modality":"TEXT","tokenCount":321.0}],"candidatesTokensDetails":[{"modality":"TEXT"
          ,"tokenCount":50.0}]},"modelVersion":"gemini-1.5-flash"}"""
      }
    ]
  }
```

6. Possible future improvements & applications

01 **Chatbot (AI agent) interface on Dashboard**

02 **Connect real application via data ingestion to OS**

03 **Real-time monitoring & prediction of Flight Operations**

Thanks for your attention!

https://github.com/Sophisss/OpenSearch