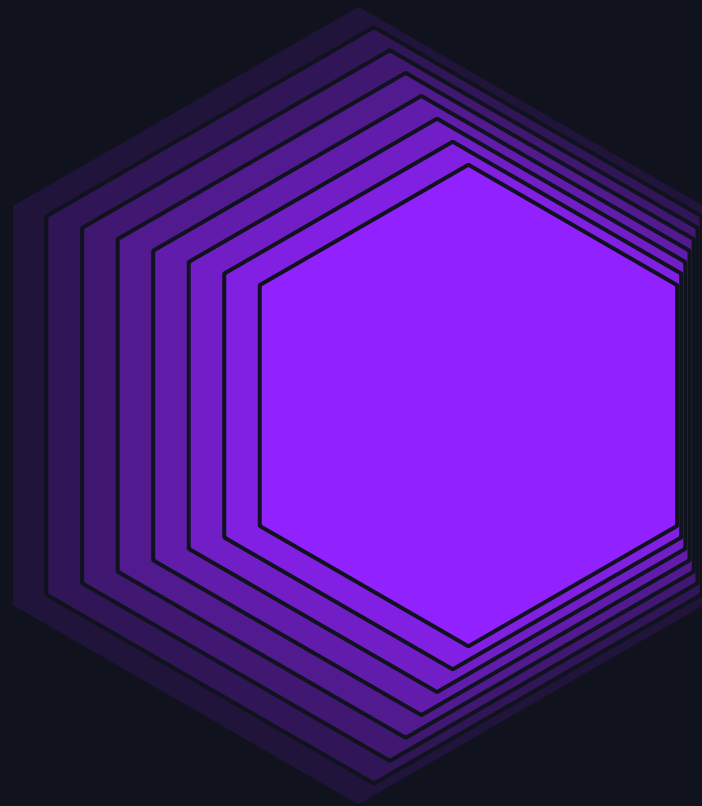


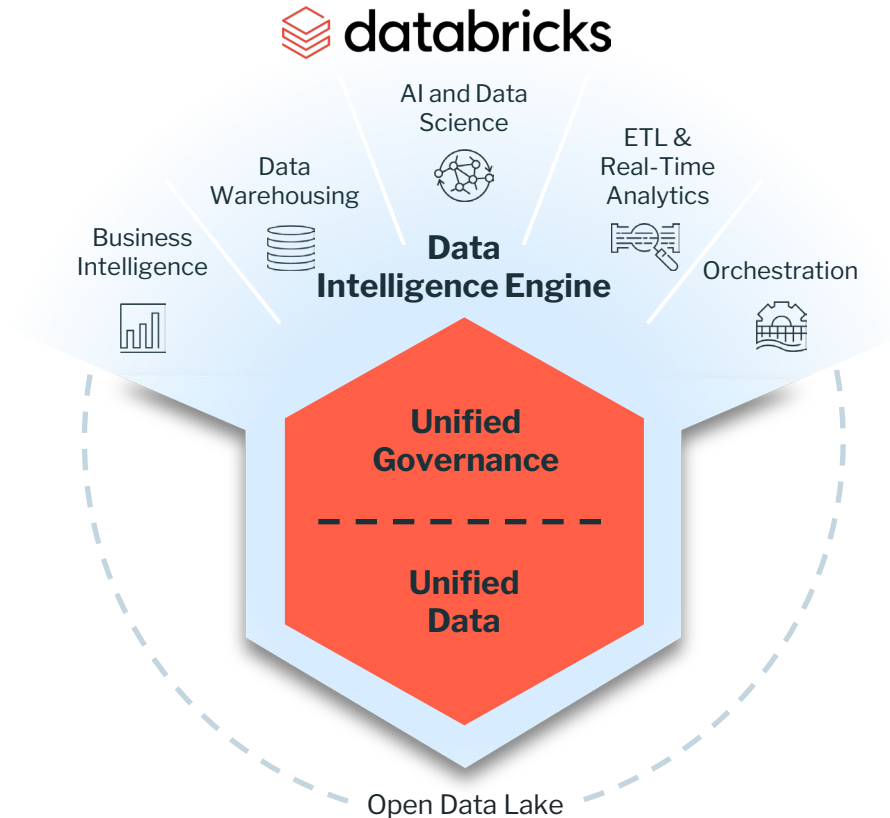
# EXPANDING LOG ANALYTICS AND THREAT HUNTING NATIVELY IN DATABRICKS



---

Ed Walsh, ChaosSearch CEO  
6/12/2024

# THE LAKEHOUSE IS THE FOUNDATIONAL AI DATA ARCHITECTURE

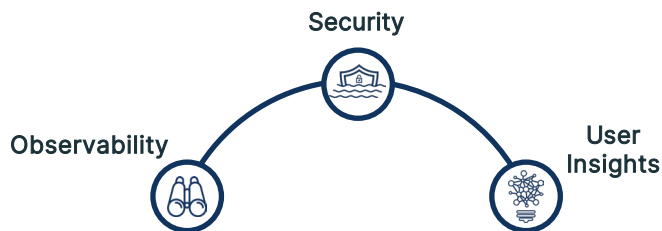


**In an AI-enabled world, a centralized data foundation is key for business success**

**The lakehouse has emerged as the core data architecture and the Databricks Data Intelligence Platform is built on a lakehouse architecture**

- Centralize data on the data lake with unified catalog and governance (Unity)
- Standardize open data format (Delta Lake) and query engine (Apache Spark™)
- Support multiple data sources and use cases, with AI focus via superior data engineering, data science experience

## BUT IT STILL HAS LIMITATIONS IN OPERATIONAL ANALYTICS



STREAMING DATA

??

Current lakehouse approach still has some limitations for operational analytics leading to separate systems such as Elasticsearch or OpenSearch to handle them.

This approach increases operational analytics' **cost**, **management toil**, reduces their **retention** and creates **data silos**.

It prevents the data lakehouse to from fulfilling its promise of centralization across use cases, and hinders SREs and Security users from leveraging the lakehouse's benefits.

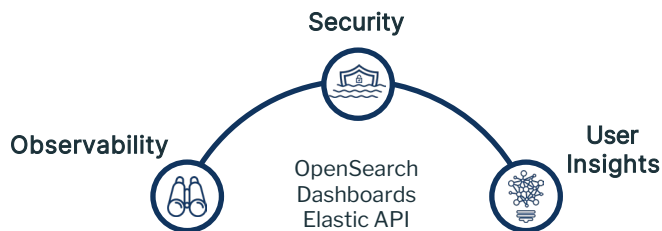
**ChaosSearch was built for  
operational analytics on the lake**

# LIVE SEARCH / ELK ON THE LAKE

A Fraction of the Cost. No Retention Limits.



## CHAOSSEARCH



- Reduced Time, Cost and Complexity
- Operational Data Lake
- Groundbreaking Database Innovation

Just a few of our customers:

**EQUIFAX**

**CISCO**



**CLOUD IMPERIUM**

**Klarna**

**blackpoint**

**REVINATE**



# KEY CHAOSSEARCH CAPABILITIES



## Efficient Text Search over telemetry looking for unknowns

Enables troubleshooting access patterns which are key for troubleshooting in Observability and threat hunting in Security, without reingestion.



## Optimized for live JSON with support of wide and dynamic schema and nested fields

Built for live ingestion and to handle the complexities of JSON, so you can seamlessly handle 3<sup>rd</sup> party sources and can give your developers flexibility, while allowing users to analyze data the way they want.



## Support of Opensearch and Elastic API familiar to SRE / SecOps

Native support of operational analytics tooling, so SREs, Eng. and Threat Hunters can use the tools they are used to, but now with all the benefits of the lake and none of the toil of Elasticsearch or OpenSearch.

# CHAOSSEARCH BRINGS LIVE SEARCH / ELK NATIVE ON DATABRICKS

Live Unified Data Lakehouse for AI-driven World

CHAOSSEARCH + databricks



STREAMING DATA



**Live Unified Data Lakehouse**  
Delta Lake with Unity Catalog

## Bringing log analytics to the lakehouse

Remove silos and bring new datasets, capabilities and use cases to your Databricks environment

Enable your Observability and Security teams with the tools and APIs they are used to

## Bringing Lakehouse and AI to Log Analytics

Bring the power of Databricks SQL, ML and AI to your Observability and Security teams

Bring the lakehouse to your Observability and Security teams for better governance, superior capabilities and dramatic savings

# KEY BENEFITS OF CHAOSSEARCH ON DATABRICKS



## Remove Silos with a Broader Lakehouse

Consolidate ELK workloads on Databricks, bring SREs, Engineers and Threat Hunters to the platform, and remove the need for an external data silo



## Add Capabilities to your Databricks Environment

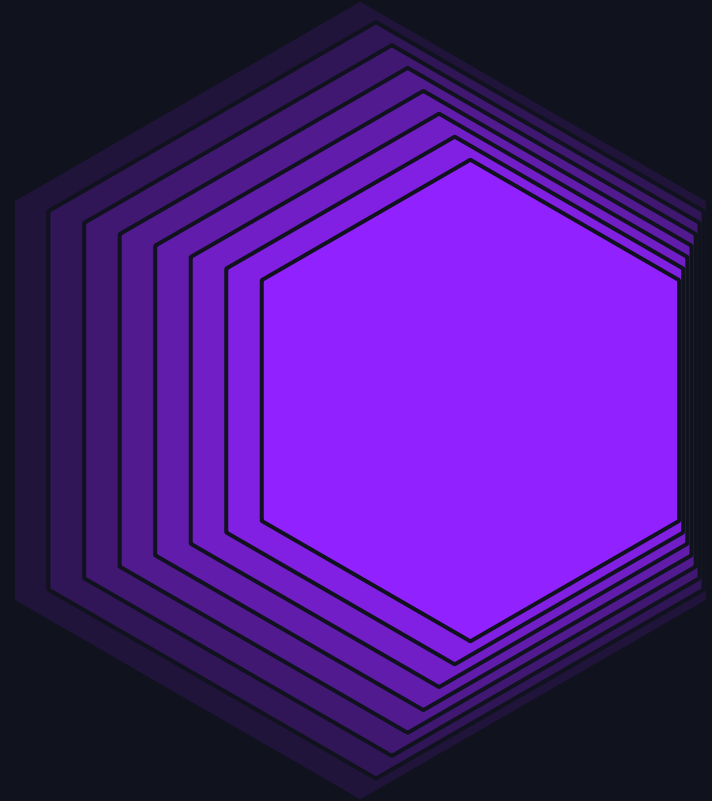
Extend live ingestion flexibility, live search / hunting capabilities and supported tooling (i.e. ELK tooling)



## Create a Unified Live Data Foundation for AI Future

Extend your Data Lakehouse, creating a unified source of truth across data and teams; and bring the AI/ML and SQL benefits of Databricks to additional live data sources

# EXPANDING LOG ANALYTICS AND THREAT HUNTING NATIVELY IN DATABRICKS



---

Ed Walsh, ChaosSearch CEO  
6/12/2024



# SET UP AS EASY AS 1-2-3



## Run ChaosSearch on Databricks

Create all-purpose compute with ChaosSearch's native libraries running on Databricks

- If you're an existing ChaosSearch customer, you can connect to an existing environment instead



## Set up Live Ingestion

Stream data to raw data buckets in your Databricks environment and configure live ingestion into Delta Tables on Unity Catalog



## Query in Chaos Console (OSD) or Notebooks

Query data leveraging ChaosSearch via OpenSearch Dashboards (OSD) in Chaos Console or Elastic API OR in Databricks via Notebooks, all leveraging Databricks Compute