

Sources | **Ingest** | **Transform** | **Query / Process** | **Serve** | **Analyse** | **Integrate**

Data Intelligence Platform
Orchestration, CI/CD, MLOps and DataOps

Batch and Streaming | **Data Science, AI / ML and GenAI Apps - Mosaic AI** | **AI/BI**

Data Engineering and Processing | **Data Warehousing**

Data Intelligence | **Operational DB** | **Applications**

Data and AI Governance - Unity Catalog
HMS Federation | Catalog and Lineage | Access Control

Data Management | **Collaboration**

Storage
Amazon S3

Sources
ETL: Files / Logs (semi-structured), Sensors and IoT (unstructured), RDBMS / DWH (structured), Business Apps (structured), Media (unstructured)
Federation / Sharing: AWS Glue, Amazon Redshift, Market places, Data Shares

Integrate
ID Provider
IAM Identity Center
Governance: Enterprise Catalog
AI Services: Anthropic, Amazon Bedrock, ...
Orchestration: External Orchestrator

Key:
3rd party (orange box)
Databricks (blue box)

Legend:
1 Integrate external metadata into Unity Catalog
2 Discover remote data in Unity Catalog
3 Check access permissions and read directly from the object storage (for any Databricks service and client, since all use Spark/Photon)