

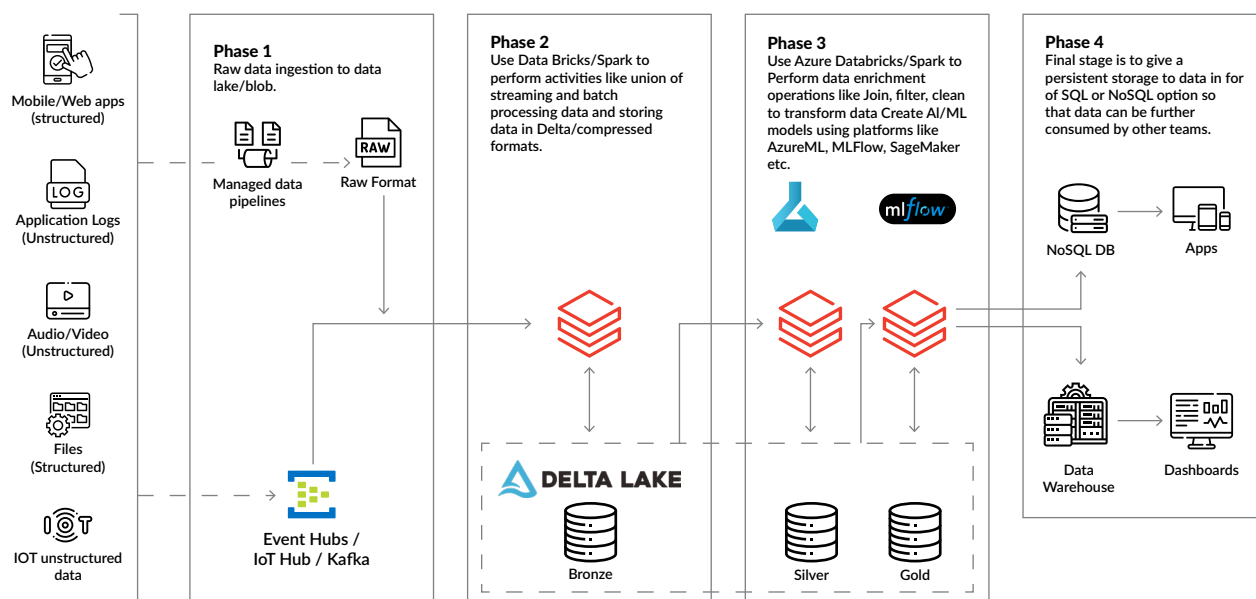
# Scalable Data Insights using Databricks



## Agivant's Scalable Data Analytics Platform Using Databricks Lakehouse Architectural Approach

With its distributed architecture and optimized data processing engine, Databricks can significantly improve the performance of massive data processing tasks. Spark's in-memory computing capabilities and advanced optimization techniques result in faster data transformations and analytics.

Databricks offers a unified platform integrating data engineering, data science, and machine learning capabilities. This integrated environment eliminates the need for switching between different tools, promotes collaboration across teams, and streamlines end-to-end data processing workflow.



## Agivant's AI Innovation Lab Has a Rich Set of Expertise in Implementing Complex Data Engineering Services Using Databricks Lakehouse Architectural Strategy



### Data Modeling

**Leverage a well-defined schema:** Design and enforce a schema for your data to ensure consistency and improve query performance.

**Partitioning and clustering:** Use appropriate partitioning and clustering strategies to optimize data retrieval and minimize the amount of data processed during queries



### Data Ingestion

**Batch and real-time ingestion:** Set up efficient pipelines for batch and real-time data ingestion to keep your Lakehouse current.

**Change data capture (CDC):** Utilize CDC techniques to capture incremental changes and update the Lakehouse accordingly.



### Data Governance

**Data encryption:** Encrypt data at rest and in transit to protect sensitive information.

**Access controls:** Implement fine-grained access controls to restrict data access based on roles and responsibilities.

**Auditing and monitoring:** Establish auditing and monitoring mechanisms to track data access, changes, and system performance.



#### Data Quality

**Data validation:** Apply data validation techniques to ensure the integrity and quality of the data stored in the Lakehouse.

**Data profiling:** Perform data profiling to understand your data's structure, completeness, and distribution.



#### Performance Optimization

**Caching:** Utilize caching techniques to speed up query performance for frequently accessed or computationally expensive datasets.

**Data skipping:** Leverage indexing or metadata-based techniques to skip unnecessary data during query execution.

**Data compression:** Apply appropriate data compression techniques to reduce storage costs and improve query performance.



#### Data Lifecycle Management

**Data retention policies:** Establish data retention policies to manage the lifecycle of your data, including archiving or deleting stale data.

**Data lineage:** Maintain a comprehensive record of data lineage to track data transformations and ensure data traceability.



#### Documentation / Collaborations

**Collaboration tools:** Use collaborative features of Databricks, such as notebooks and version control, to encourage teamwork and knowledge sharing.

**Documentation:** Document data pipelines, transformations, and custom logic to facilitate understanding and maintainability.



#### AI/ML

**Collaboration tools:** Use collaborative features of Databricks, such as notebooks and version control, to encourage teamwork and knowledge sharing.

**Documentation:** Document data pipelines, transformations, and custom logic to facilitate understanding and maintainability.

## Value to customer

Agivant AI Innovation Lab has a rich reusable library of best practices and key learnings to build highly scalable data architecture with Databrick as a core technology.

- ▶ Provides unified Analytics Platform for data engineering, data science, and analytics to improve collaboration and quality of insights.
- ▶ Implementation expertise in Microsoft Cloud Scale Analytics Reference Architecture using Databrick health lakehouse and OMOP standard data model.
- ▶ Healthcare organizations deal with large and diverse datasets. A health lakehouse powered by Apache Spark offers scalability and high-performance processing capabilities. It can efficiently handle the volume, velocity, and variety of healthcare data, ensuring timely analysis and insights.
- ▶ Robust security features to protect sensitive data and ensure compliance with data privacy regulations. It provides encryption at rest and in transit, fine-grained access controls, auditing capabilities, and identity and access management (IAM) systems integration.
- ▶ Collaborative features help share code snippets and leverage version control to ensure seamless collaboration and maximize productivity.
- ▶ Lakehouse architecture supports both real-time and batch processing. It can handle streaming data ingestion, enabling real-time analytics and insights. At the same time, it can process batch data, allowing for comprehensive and historical analysis.

**Agivant** = **AGI**lity x rele**VANT**



Agivant is a new age **AI-First Digital & Cloud Engineering** Services Company.



We enable **clients** with the power of **technology solutions** to help them achieve their trajectory of **growth and relevance, through agility.**



We partner with clients to help them in **Technology Innovation, Scale & Optimization and Engineering Enablement.**



We are a **result oriented team** that focuses in delivering **client success** through their **business growth** and **relevance** to the new world of **opportunities.**

## We are Agivant

Agivant is a new-age AI-First Digital and Cloud Engineering services company that drives Agility and Relevance for our client's success.

Powered by cutting-edge technology solutions that enable new business models and revenue streams, we help our clients achieve their trajectory of growth. Agility is a core muscle, an integral part of the fabric of a modern enterprise. To succeed in an ever-changing business environment, every modern organization needs to adapt and renew itself quickly. We help foster a more agile approach to business to reconfigure strategy, structure, and processes to achieve more growth and drive greater efficiencies.

### Our Services

Our portfolio of services helps leading enterprises adopt cutting-edge technology solutions to transform their products and services and evolve business models.



Platform Engineering



AI and Data



Cloud



Experience Engineering

