

Our Databricks Partnership

Unified data, analytics and AI that accelerates innovation

Businesses today collect a massive amount of data from multiple sources and from across domains and functions. They need to build and run big data jobs quickly and easily, with multiple processors running simultaneously. They also need to eliminate data redundancy, ensuring quality, reliability, and integrity of their data while providing analytics that help improve forecasting and outcomes.

Databricks can help businesses manage and analyze their big data using applications that require large-scale data processing, such as real-time analytics, machine learning (ML), and data warehousing.

Our Databricks Accelerator

We leverage our subject matter expertise and long-standing partnership with Databricks to provide a comprehensive suite of services to help expedite our clients' digital transformation journeys. Our Databricks Accelerator includes:

- ▶ **Data Architecture**
Using Databricks, we design data architectures that provide access to high quality, relevant and accessible data, while reducing costs and improving efficiencies.
- ▶ **Database Platform Migration**
We help our clients migrate their data assets from legacy platforms to Databricks and stay with them throughout their migration journey—from assessment to production.
- ▶ **Solutions Implementation**
Our Databricks experts recommend solution architectures based on our client's cloud environment and help implement solutions with Apache Spark's best practices.

Databricks Advantage

The Databricks Lakehouse Platform combines the best elements of data lakes and data warehouses to deliver reliability, strong governance, and performance of data warehouses with the openness, flexibility, and ML support of data lakes.

The system platform can verify the data, ingest it into in-house systems, process data and run business logic. This platform is powered by Apache Spark, a unified analytics engine that enables distributed computing on multiple compute nodes with implicit data parallelism and fault tolerance.

The Databricks Lakehouse Platform integrates easily with popular BI tools like Tableau and Power BI. This enables users to quickly visualize data or explore the results generated by ML models. Moreover, Databricks' ability to return complex queries in a few seconds enables teams to build live responsive analytics dashboards for business teams.

Our Expertise

- Develop solutions on Databricks platform using SparkSQL and PySpark which provide quick access to analytics on big data.
- Provide integration between existing cloud resources and services and build a solution architecture on our client's cloud platform, helping them connect from Databricks to their storages and resources.
- Integrate with BI tools such as Power BI and Tableau, helping businesses to analyze data and build reports rapidly.
- Orchestrate big data pipelines powered by Databricks and speed up the analytics lifecycle.
- Develop and manage ML lifecycles on Databricks so that the data science team can build powerful ML models and bring analytics from big data. The ML processes are powered by the Spark engine which reduces the time taken to train a particular model. An entire ML lifecycle can be automated on Databricks, from data ingestion to productionizing an ML model for consumption.

Addressing Industry Problems with Databricks

We have helped many businesses to leverage Databricks. Here's a glimpse of the type of problems we have helped our clients address:

Pharmaceutical

Challenges

Enable a mechanism that would remove inconsistencies and errors in data coming from various sources (i.e., IoT equipment, laboratory machines, and pharma lifecycle data from external vendors) and provide one version of the truth for timely decision-making pertaining to market strategy and products.



Our Solution

- Leveraged Databricks to ingest nearly 30 Terabytes of market scan data for analysis.
- Verified data using PySpark script that ran on Databricks cluster.
- Captured and fixed the inconsistencies and generated a comprehensive report.
- Used Tableau to build live dashboards for running real-time queries.

Challenges

Ingest medical data from IoT devices, bring it to the client's cloud system, and analyze that data for usage patterns, refills, maintenance use cases.



Our Solution

- Used Azure Data Factory pipelines to bring the data generated by medical devices to the source system.
- Leveraged Azure Data Factory to automate the entire process—from availability of new data to processing and loading it into existing tables to reflecting the insights in real time in a Tableau dashboard.

Challenges

Minimize the human effort generally needed in reviewing marketing material on a particular drug by running an ML model.



Our Solution

- Used Azure Data Factory pipeline to create an automated process on Databricks for new files or content coming into a source location, and projected the results in a Tableau dashboard in real time.
- Implemented a write-back mechanism on the Tableau dashboard to capture feedback from users and send it back to the database.

Databricks Implementation Options

Blue Altair's certified Databricks experts help clients expedite their data processes and analytics so business teams can gain accurate insights and make timely decisions.

Options	Source Systems	Workflows	Reports	Lakehouse Deployment	Duration	Estimated Cost
Starter	Up to 2 sources	2-4 simple transformations	2 reports	Non-production environment	10 business days	\$25K
Comprehensive	Up to 4 sources	Workflows for each source and associated transformations	Up to 10 reports	Production deployment Included	10-12 weeks	\$75K
Tailored	No Limit	Workflows for each source and associated transformations including custom transformations	No Limit	Production deployment Included	Custom Duration - Contact Blue Altair	Custom Pricing - Contact Blue Altair

About Blue Altair

Blue Altair is a niche, industry-recognized business and technology consulting firm that assists our clients with digital transformations. We offer Assessment and Strategy, Technology Implementation, and Managed Services in API Management and Integration; Data Management; Digital Application Development; and Data Science and AI. Our Client Success capability ensures a higher-than-industry rate of successfully delivered projects, with a primary focus on program and project management, business analysis, and quality assurance. Blue Labs is our innovation hub, where we use cutting-edge technology to build offerings that deliver accelerators and solutions. Our culture is the heart of our existence, and our core values are the key drivers for our handpicked, top-tier performers.

Blue Altair plays a pivotal role alongside enterprise's IT and business teams to ensure critical business pain points are addressed, and enterprise business goals are achieved.

Contact us to learn more about how Blue Altair can help implement Databricks in your organization.

info@bluealtair.com | www.bluealtair.com

1530 Wilson Blvd., Suite 650
Arlington, Virginia 22209 USA
+1 703 721 7275



United States | India | Australia | Argentina