

Blendata Enterprise

Introducing our **revolutionary big data platform** for elevated organizational data management.



Experience reduced processing time, faster time-to-achievement, and resource optimization.



Unlock operational efficiency and gain a competitive edge through data insights and automation.



Seamlessly connect with all your data for effortless handling of complex data management challenges.

Welcome to our unified and powerful solution.

Key Benefits

Effortless Data Management with Data Lakehouse:

Our streamlined big data platform is designed for effortless data management and time reduction. Ready for any data and any workloads with the Lakehouse concept, With simplified data management & analytics and all-in-one capabilities, it covers data ingestion, processing, analysis, and utilization. It offers low-code or no-code solutions, eliminating the need for coding expertise. Experience less effort and reduced time with our user-friendly platform.

High-Speed Distributed Architecture:

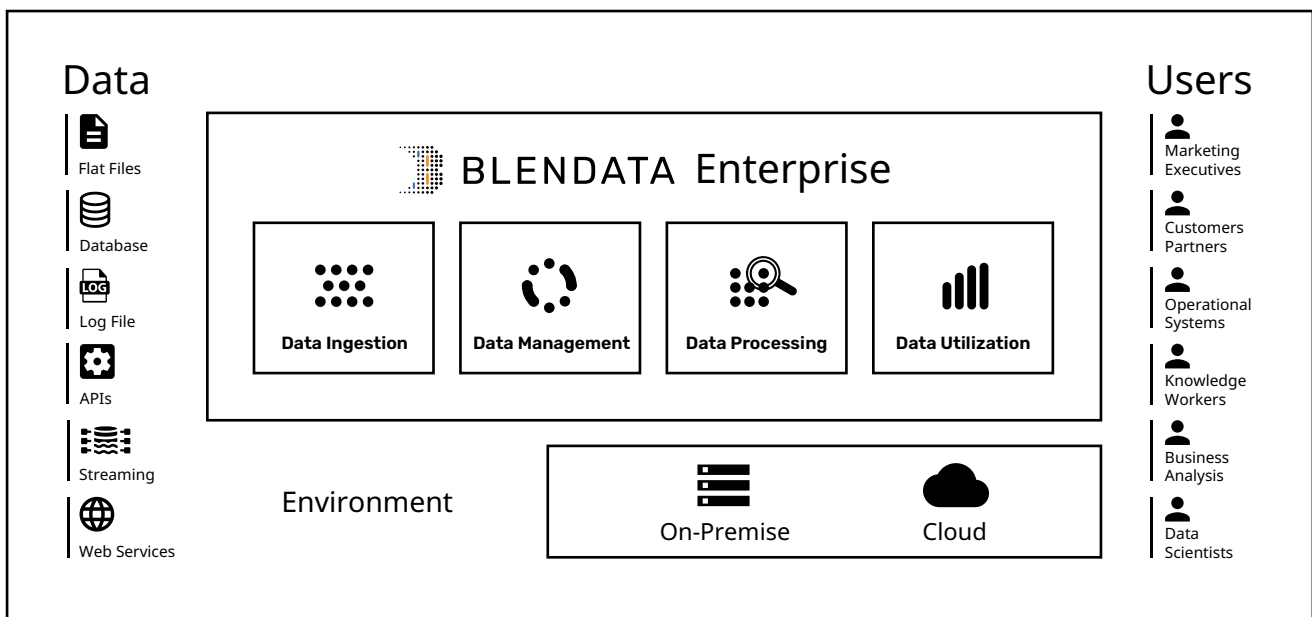
Our platform is built on a decoupling compute and storage architecture. Computational powered by Blendata's proprietary Distribute In-Memory parallel processing, The Optimized Apache Spark (1.5x faster than OSS), and logical storage supported by a famous open source Delta Lake, it eliminates the need for traditional high-cost scale-out, Coupling architecture setups.

Seamless Data Integration:

Connect to diverse data sources for structured, semi-structured, and unstructured data. Our platform supports multiple interfaces, including cloud storage sites like Amazon S3 and Azure Blob, as well as on-premise environments such as flat files, logs, Kafka, HDFS, and enterprise databases like Oracle, MS SQL Server, and MongoDB. Enjoy dynamic data ingestion with ease.

Flexible Data Access and Robust Security:

Our platform offers an open design, enabling organizations to leverage data usage and other platform functions effortlessly. Connect through APIs supporting REST API, ODBC/JDBC, or direct data access using open format files. Rest assured, all connections, in-transit and at-rest data, are safeguarded against unauthorized access. We employ SSL/TLS standard, columnar at-rest encryption and in-transit encryptions, OAuth 2.0 user/role-based access control authentication, and authorization protocols with Multi-factor authentication supported (MFA) to ensure the highest level of data security.



Key Features

Data Ingestion

Easy Data Integration:

Our platform offers a wide range of built-in connectors, allowing you to effortlessly ingest data from various sources. Whether it's flat files (CSV, JSON, Plaintext, Parquet, etc.), databases (Oracle, MS SQL, MySQL, etc.), cloud services, or streaming data, no coding is required.

Streamlined Data Ingestion:

Experience automated data ingestion from flat files, logs, databases, and cloud storage using our built-in connectors. Choose from methods like scheduled replication (batch or streaming), data virtualization, or changed data capture (CDC) to ensure efficient and timely data transfer.

On-the-Fly Data Transformation:

Transform your data during the ingestion process with ease. Our platform utilizes Spark SQL, providing a powerful set of transformation functions to handle large volumes of data seamlessly.

Automated Schema Discovery:

Whether you're working with files, databases, or streaming data through Kafka, our engine is designed to automatically discover the characteristics of your data and create the corresponding schema for each ingested table. Additionally, you have the flexibility to modify the schema using our user-friendly GUI.

Data Management

User & Role-Based Access Control:

Easily manage access rights by tailoring roles or users for each asset, such as data, visualizations, dashboards, and Blendata Enterprise functions. Take control of who can access and interact with your valuable resources.

Key Features

Data Management

Data Encryption with KMS Integration:

Ensure the security of your data at rest by encrypting it at the table or even column level. You can use your own encryption key or integrate with an external Key Management System (KMS) to synchronize the encryption process (customization required).

Data Lifecycle Policy:

Effectively manage data retention for each table with our data lifecycle policy. Apply stages to your data, such as hot, warm, or cold, to define the compression and storage type. This allows you to optimize storage utilization and streamline data management.

Multi-Storage Capability:

Connect your big data storage to multiple targets effortlessly. Enable data tiering across multiple storage devices or specify a specific storage device for each table. This flexibility enables efficient data storage and retrieval.

Advanced Data Management Techniques:

Benefit from a range of advanced data management techniques that enhance your data management capabilities. These techniques include data compaction, compression, lifecycle management (hot-warm-cold), skipping, predicated push-down, dynamic partition pruning, and physical file partitioning. Maximize the efficiency and performance of your data operations with these powerful features.

Data Processing & Analytics

SQL Analytics:

Easily prepare, transform, correlate, aggregate, and analyze your data using our SQL Analytics feature. The intuitive data exploration user interface allows data analysts and engineers to write queries in familiar ANSI SQL syntax. You can explore, manipulate, and schedule jobs on your data effortlessly.

Data Prep Workbench:

Empower non-technical users to prepare their data sets quickly with our Data Prep Workbench. It offers the same engine and capabilities as SQL Analytics but provides easier low/no-code options for data preparation. Get your data ready for analysis in minutes, even without extensive technical knowledge.

Notebook Interface for AI/ML:

Enjoy a first-class experience for data engineers and ML developers with our Notebook interface. It supports popular development languages such as Python, R, Scala, and SQL. Leverage the power of notebooks to streamline your data engineering and machine learning workflows.

Scheduler + Chain Job:

Simplify your data pipeline with our built-in job scheduler and chain job functionality. Schedule all your defined jobs with ease and connect them in a chain to streamline the flow of data from end to end. Take control of your data processing and ensure smooth execution of your data tasks.

Key Features

Data Utilization

Visualization & Dashboard:

Easily create meaningful visualizations by dragging and dropping query results into a wide range of charts. Group them together to build interactive dashboards, apply filters, and drill down for deeper insights. Share these data-rich applications with your colleagues.

Notification:

Set up SQL-based conditions to trigger alerts and receive data notifications via the web console, email, script, or API calls (customization required).

API:

Securely access your data through a REST API, with OAuth2 authentication for enhanced protection. For low-level data integration, you can also connect using Spark ODBC/Hive2server JDBC, ensuring authentication, authorization, and row-level security for ODBC/JDBC connections.

Export:

Seamlessly export processed data to various formats and different downstream destinations, including files, databases, or cloud storage.

Security and Compliance

Security at Rest:

With our platform, you can safeguard your data at rest through columnar encryption. Integration with Key Management Systems (KMS) allows for enhanced encryption and protection. Additionally, our platform employs data protection methods at the storage level, including disk encryption, erasure coding, replication, and more.

Security at Transit:

Our platform ensures secure data transmission through SSL/TLS protocols. Every transmission is encrypted, providing a secure channel for data transfer.

Security at Use:

Maintain control over your data with Blendata Enterprise's comprehensive security features. Implement row-level security to control access to specific rows of data. User and role-based access control allows you to define and manage access rights for different users and roles. Masking and hashing functionalities add an extra layer of data protection, ensuring sensitive information remains secure.

Compliance:

Our platform is designed with enterprise-grade security measures that are ready to meet any compliance requirements. We utilize global security methods such as encryption-at-rest, in-transit, and all communications between related services.