ORACLE

Extend your Oracle Fusion Analytics with a Data Lakehouse

Extend your analysis across all data with a
Lakehouse built on new OCI capabilities

Purpose statement

In this brief, explore how the combination of Oracle Fusion Analytics and a data lakehouse on Oracle Cloud Infrastructure delivers powerful enterprise advantages. Get a firsthand look at how extending Oracle Fusion Analytics with a data lakehouse can reveal new opportunities, optimize efficiencies, and increase customer satisfaction across your business.

Disclaimer

This document is for informational purposes only and is intended solely to assist you in planning for the implementation and upgrade of the product features described. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle. Due to the nature of the product architecture, it may not be possible to safely include all features described in this document without risking significant destabilization of the code.

ORACLE

Contents

ORACLE

The [recently announced](#) Oracle Cloud Infrastructure (OCI) services for building data lakehouses bring significant new value to Fusion Analytics customers. To understand what new possibilities this creates, it's useful to review the key capabilities of Oracle Fusion Analytics, the new data lakehouse capabilities, and how the combination of Oracle Fusion Analytics and a data lakehouse on OCI delivers powerful analytic advantages.

Oracle Fusion Analytics: The system of insight for fusion cloud applications

Oracle Fusion Analytics delivers comprehensive data and analytics necessary for application insights in a single cloud service. It is powered by the [Oracle Autonomous Database](#) and [Oracle Analytics Cloud](#) (OAC), delivering robust capabilities for data management and advanced analytics. Purpose-built for Fusion Application customers, Fusion Analytics delivers rapid time-to-value thanks to a fully managed data pipeline, a single data model for all your functional areas, and best practice KPIs and dashboards.

Fully managed service

Fusion Analytics takes complete care of extracting and transforming data from Fusion Applications and landing the data in the Autonomous Database in the form of a data model that is optimized for analytics use cases. The service continuously monitors the data pipeline and ensures that, while new data is loaded into the warehouse, users experience no service interruption. Fusion Analytics is certified with Fusion Cloud Applications before each new release, ensuring that both packaged and custom content continues to work and perform whenever an upgrade of Fusion Application occurs.

Common data model for cross-functional analysis

Additionally, Fusion Analytics features a data and semantic model with conformed dimensions for finance, HR, operations, marketing, sales, and service. This single layer on top of all operational data enables tracking and measurement and provides a single set of cross-functional KPIs that can be shared across different business functions.

Comprehensive analytics

In addition to providing hundreds of prebuilt KPIs, Fusion Analytics offers support for metric aggregation aligned with complex hierarchies and based on customer-configurable aggregation buckets. Fusion Analytics gives customers a head start with a wide variety of ready-to-go content, and includes the comprehensive analytics capabilities of OAC that allow customers to create custom KPIs, leverage natural language processing, and perform advanced analytics with analysis powered by machine learning (ML).

Fusion Analytics is tailor made to enable Fusion Application customers to gain insights from their application data in a comprehensive cloud service. However, this doesn't mean the service is restricted to Fusion Application analysis. The service is also fully extensible and customizable; allowing customers to import external data into the Autonomous Database, expand the base semantic model, and add additional content.

Data extensibility

Fusion Analytics manages the pipeline for all data sourced from Fusion Applications and protects the corresponding target schema in the Autonomous Database instance to ensure consistency and continuity of the cloud service. Customers can access the database instance to load data from other sources and applications. Whether it's data from other on-premises applications, including Oracle E-Business Suite and Siebel, or data from third-party applications, such as Salesforce and Workday, customers have full flexibility to create custom schemas in the Autonomous database instance and populate these schemas with their choice of data extract, transform, and load (ETL) tooling.

All of the Autonomous Database's capabilities for loading, transforming, and making sense of data are readily available—including features for discovering anomalies, outliers, and hidden patterns for understanding data

ORACLE

dependencies. The built-in analytics functions for detecting anomalies, clustering data, sampling data, and training and running ML models are computed directly in the database. This eliminates the need to move data to dedicated analytics and ML systems. The Autonomous Database features SQL and PL/SQL for data preparation, model building, evaluation, and deployment, as well as Auto-ML capabilities and a built-in Python environment for data science use cases. To uncover hidden data relationships and make sense of location data, the Autonomous Database provides built-in graph and spatial analytics features.

Tight integration between the Autonomous Database and OAC provides powerful analytic benefits while reducing effort and time to value. For example, data scientists can use the Autonomous Database for ML and advanced analytics and the related insights can immediately be available in OAC for business users thanks to the seamless integration the database and OAC. Additionally, the Database's built-in analytics functions can be added to a data flow in OAC and, using OAC, users can easily analyze graph data in the Autonomous Database.

Semantic extensibility

As data warehouses grow and cover more subject areas, the data can become complex and difficult to analyze. A semantic model is a logical layer that simplifies access to the underlying database tables, enables enterprise governance and facilitates a single version of the truth—which can dramatically increase user satisfaction and adoption. Fusion Analytics comes with a semantic model that can be extended to include non-Oracle data sources so that access to the data for business users remains simple and easy to understand.

The extensibility framework of Fusion Analytics enables customers to semantically enrich the packaged semantic model of Fusion Analytics and augment it with semantics for data sourced from non-Fusion Applications. In addition to supporting Fusion Application flex fields extensions, the framework supports a variety of common semantic model extension use cases. It's possible to extend prebuilt subject areas, attributes, and metrics, create custom ones, define new hierarchies and session variables, and more. The security model of Fusion Analytics is integrated with the security configuration of the Fusion Application through synchronization of Fusion job roles and user mappings. The extensibility framework allows customers to secure new subject areas and data with custom duty and application roles. All the while, Fusion Analytics ensures that all extensions are preserved and will keep working from release to release.

Analyze all data—introducing the Data Lakehouse

The data needed to inform the business comes from many sources within and outside the organization, and in a variety of semi-structured and unstructured formats: weather data, online transactions from mobile devices, IoT sensors, social media posts, videos, or survey data. Incorporating this data into analysis may reveal new product or revenue opportunities, diminish fraud risk, optimize operational efficiencies, or increase customer satisfaction.

Insights from beyond the warehouse—diving into the data lake

A data lake stores data in its native (raw) format, which may include structured data from relational databases (rows and columns), semi-structured data (CSV, logs, XML, JSON), unstructured data (emails, documents, PDFs) and binary data (images, audio, video).[1]

This data is often big data—entering into the date lake (i.e. object storage) in large volumes, with high velocity and in a variety of formats. Because of the size, scope, and complexity, it is not easy to manage nor well suited for movement.

ORACLE

How can this variably structured, large volume data be incorporated into analyses?

Data Lakehouse on OCI
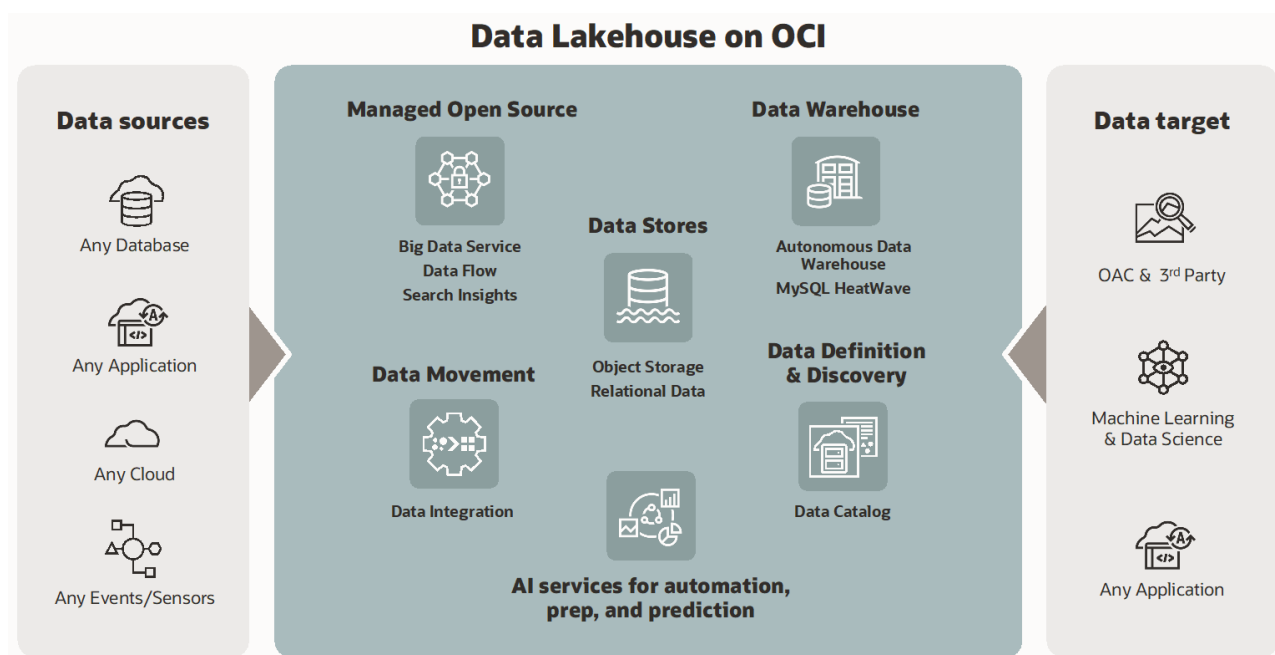
Enter the data lakehouse. A data lakehouse is a modern data platform which leverages a data lake ("lake") and data warehouse ("house") to expand the analytic opportunity across an organization's data estate ("lakehouse"). It combines the power and richness of data warehouses with the breadth and flexibility of the most popular open-source data technologies to manage semi-structured and unstructured data. The data lakehouse architecture enables you to store, understand, and analyze all data. With a lakehouse on OCI, Fusion Analytics customers can now extend their analysis and find new insights from semi- and unstructured data such as invoices, forms, text, audio, and video.

As discussed in The Future of the Data Lakehouse event, customers are building lakehouses using OCI to get the most value out of their data.

Object storage

OCI Object Storage enables customers to store any type of data in its native format. With built-in redundancy, OCI Object Storage is ideal for building modern applications that require scale and flexibility, as it can be used to consolidate multiple data sources for analytics, backup, or archive purposes.

Data catalog

OCI Data Catalog helps data professionals across the organization search, explore, and govern data using an inventory of enterprise-wide data assets. It automatically harvests metadata across an organization's data stores and provides a common metastore for data lakes. OCI Data Catalog simplifies the definition of business glossaries and curated information about data assets located in OCI and other locations so consumers can easily find needed data.

AI services

A data lakehouse can be built from the ground up on OCI to work with the latest AI frameworks and prebuilt AI services. New AI services for OCI help developers with a range of complex tasks, from language to computer vision to time-series forecasts:

ORACLE

- **OCI Language:** Performs text analysis at scale to understand unstructured text in documents, customer feedback interactions, support tickets, and social media. With built-in, pre-trained models, OCI Language eliminates the need for ML expertise and empowers developers to apply sentiment analysis, key-phrase extraction, text classification, named entity recognition, and more into their applications.

- **OCI Speech:** Provides automatic speech recognition through prebuilt models trained on thousands of native and non-native language speakers for real-time speech recognition. OCI Speech enables developers to easily convert file-based audio data containing human speech into highly accurate text transcriptions and can be used to provide in-workflow closed captions, index content, and enhance analytics on audio and video content.

- **OCI Vision:** Provides pre-trained computer vision models for image recognition and document analysis tasks. It also enables users to extend the models to other industry and customer-specific use cases, such as scene monitoring, defect detection, and document processing with their own data. OCI Vision can be used to detect visual anomalies in manufacturing, extract text from forms to automate business workflows, and tag items in images to count products or shipments.

- **OCI Anomaly Detection:** Delivers business-specific anomaly detection models that flag critical irregularities early, which enables faster resolution and less operational disruption. OCI Anomaly Detection provides REST APIs and SDKs for several programming languages, which developers can use to easily integrate anomaly detection models into business applications. It is built on the patented MSET2 algorithm, which is used worldwide in highly sensitive situations like nuclear reactor health monitoring, and can be used for fraud detection, predicting equipment breakdown, and receiving data from multiple devices to predict failures.

- **OCI Forecasting:** Delivers time-series forecasts through ML and statistical algorithms without the need for data science expertise. OCI Forecasting helps developers quickly create accurate forecasts for their critical business metrics, including product demand, revenue, and resource requirements. These forecasts all have confidence intervals and explainability to help people make the right business decisions.

- **OCI Data Labeling:** Helps users build labeled datasets to train AI models. Users can assemble data, create and browse datasets, and apply labels to data records through user interfaces and public APIs. The labeled datasets can be exported and used for model development across many of Oracle's AI and data science services, including OCI Vision and OCI Data Science, for a consistent model-building experience.

Data integration

OCI Data Integration enables developers to simplify complex data extract, transform, and load processes (ETL/E-LT) into data lakes and warehouses for data science and analytics with Oracle's modern, no-code data flow designer.

Managed open source

Oracle Big Data Service is a Hadoop-based data lake for storing and analyzing large amounts of raw customer data. Aa a managed service, Oracle Big Data Service comes with a fully integrated stack that includes both open source and Oracle value-added tools that simplify your IT operations. Oracle Big Data Service makes it easier for enterprises to manage, structure, and extract value from organization-wide data.

OCI Data Flow is a fully managed Apache Spark service with no infrastructure for customer IT teams to deploy or manage. OCI Data Flow lets developers deliver applications faster because they can focus on application development without getting distracted by operations.

Oracle Fusion Analytics with a Lakehouse on OCI: Key benefits for Oracle Fusion Analytics customers

The already robust value of Fusion Analytics gets even better with a lakehouse deployed using OCI and new lakehouse services—empowering customers to deliver greater impact more efficiently and with broader use cases. Specifically, there are five core benefits of a lakehouse deployed on OCI for Fusion Analytics customers:

ORACLE

1. **Query more data:** Structured data (in the Autonomous Database), semi-structured, or unstructured (in Object Storage). Autonomous Data Lake Accelerator, a component of Oracle Autonomous Database, can seamlessly consume object storage data, scale processing to deliver fast queries, autoscale the database compute instance when needed, and reduce the impact on the database workload by isolating object storage queries from the database compute instance.

   Learn more: Cloud data lakehouse - process enterprise and streaming data for analysis and machine learning

2. **Query a variety of data:** Customers can now query data in Fusion Analytics and Object Storage, including data which may come from other clouds (AWS, Azure), streaming data, and unstructured data, with broad support for various data formats (Parquet, Avro, ORC, CSV, JSON, XML).

   Learn more: Analyze data from external object storage sources using Oracle Cloud Infrastructure Data Flow

3. **Query data in place (No data movement):** Fusion Analytics queries are now federated across a lakehouse on OCI, enabling interactive queries on data in both Autonomous Data Warehouse and Object Storage. By querying the data in place, you no longer have to move the data.

   Learn more: Data Lakehouse

   See also: Cloud data lakehouse - process enterprise and streaming data for analysis and machine learning

4. **Catalog your data:** Data lakehouse services use Data Catalog for unified data definition. Data Catalog helps data professionals across the organization search, explore, and govern data using an inventory of enterprise-wide data assets. It automatically harvests metadata across an organization's data stores and provides a common metastore for data lakes. Data residing in Fusion Analytics, and across your data estate is indexed, discoverable and ready for analysis with Fusion Analytics.

   Learn more: Data Catalog

   See also: The key technologies of data lakehouses on OCI

5. **Support Open-Source Workloads:** Provide the flexibility and openness to empower your data team to leverage open-source offerings such as Hadoop and Spark. Customers can easily migrate existing or build new open-source data lakes with our fully managed services like Oracle Big Data Service and Oracle Data Flow. Spark, HIVE, Hbase, and many more services can be easily deployed and scaled on OCI. Data Flow is a serverless Spark service that enables our customers to focus on their Spark workloads with zero infrastructure concepts.

   Learn more: Oracle Cloud Infrastructure makes it easy to build a modern lakehouse architecture that is open, flexible, and intelligent

Fusion Analytics and OCI services for building lakehouses together comprise a set of data and analytic services running on the Oracle Cloud that are designed to handle the most complex data use cases across all of an organization's data—Oracle or non-Oracle, structured and unstructured—in an integrated and efficient manner. Fusion Analytics and lakehouse services are native to OCI and designed to be used together. Fusion Analytics is built on top of Autonomous Database and OAC and provides a set capacity for these included OCI services. Fusion Analytics customers can expand the capacity of its data warehouse on the Autonomous Database to accommodate large volumes of third-party data. To further take advantage of data lakehouse services, customers can simply provision instances of the desired services from the OCI console and start using them in combination with the Fusion Analytics service.

For Fusion Analytics customers, a data lakehouse on OCI allows for expanded analysis across a data estate (across Fusion Analytics and data in Object Storage) to broaden insights. Below are a few sample use cases that demonstrate the benefits of Fusion Analytics in combination with the data lakehouse on OCI.

ORACLE

Oracle Fusion Analytics and Lakehouse use cases

- **ERP Analytics—Revenue and Demand Forecasting**: There are many businesses for which changes in the weather can mean changes in revenue. Weather data is stored in OCI Object Storage. Using OCI Data Flow and AI Services, this weather data can be prepared and mined, weather pattern changes can easily be correlated with the historical revenue time series data from Fusion ERP Analytics to build models which can predict future revenue. Similarly, social media insights can be analyzed or used in a model to predict the future sales for specific products.

- **HCM Analytics—Employee Survey Analysis**: Customers can leverage the OCI services to deploy a lakehouse to track and analyze sentiment and engagement based on employee surveys. Sentiment analysis uses trained sentiment classifiers to provide sentiment information for documents or topics within documents. Employee survey data is first loaded into Object Storage using Oracle Data Integrator, OCI Data Integration, or other data integration tool of choice. By running sentiment analysis with OCI Language service on top of the survey data in Object Storage, customers can derive sentiment metadata. The resulting sentiment data in Object Storage together with the relational HR data in Fusion HCM Analytics can subsequently be analyzed in combination. This can be accomplished by linking the sentiment data in Object Storage from the Autonomous Database and enabling users to query the combined data with OAC.

- **SCM Analytics—Field Engineering and Product Reliability**: High-cost capital equipment is often monitored in the field, and failure rates are closely analyzed to improve the reliability of the products. The monitoring involves multiple data streams, including weather data, to capture the environmental conditions, telemetry from the equipment in use, and images to capture the visual properties. Product companies can collect such data in OCI Object Storage and analyze it using OAC in combination with manufacturing data (information regarding where the product was manufactured, what time, on which equipment, etc.) and component sourcing data (such as components used, from which suppliers, and from which lot) in Fusion SCM Analytics to gain insights regarding product reliability.

- **CX Analytics—Marketing Attribution**: Customer information comes into an organization from a variety of touchpoints along the journey from prospect to paying customer. Interactions occur across various channels encompassing social (LinkedIn, Twitter, etc.), blogs, articles, display ads, search, and web properties, and once a form gets filled the "unknowns" become a "known" lead. As the lead gets qualified, there are further touchpoints including emails, webinars, events, and meetings. Which of these channels and touchpoints are statistically significant in the successful conversion of prospects to customers? To address this question, you can marry semi-structured data from logs, click-streams, and search engines, stored and transformed in OCI Object Storage, with structured data in marketing automation and CRM platforms, provided by Fusion CX Analytics.

The future of Oracle Fusion Analytics and Data Lakehouse

Recent announcements share how to connect your enterprise with AI-driven analytics and use new AI Services to prepare data at scale with ML-powered language, speech, vision, and data labeling services. New AI services for decision making, including anomaly detection and forecasting, have been introduced—with more to come!

Customers now have powerful new capabilities to achieve competitive advantage and maximize their investments in Oracle Cloud Applications and Oracle Cloud Infrastructure. Stay tuned for future innovations that will empower customers across all industries to capitalize on insights from across their data estate using Fusion Analytics, OCI services for building lakehouses, and AI Services.

ORACLE

AUTHORS: Joey Fitts and Stefan Schmitz, Oracle

Connect with us

Call +**1.800.ORACLE1** or visit **oracle.com**. Outside North America, find your local office at: **oracle.com/contact**.

blogs.oracle.com   facebook.com/oracle   twitter.com/oracle

ORACLE