

Fact Sheet: Acuvate Data Platform

INTRODUCTION

The Acuvate Data Platform is targeted as an Enterprise Data Platform, and it means that it can scale and handle all volumes and types of data within an enterprise (from small to large). It is supported on various cloud flavors and On Premise since it is normally installed as part of the customer's infrastructure (either Cloud or On Premise). However, installation, integration, and ongoing support are fully offered by Acuvate, with ongoing support being SLA (Service Level Agreement) based.

CAPABILITIES

Acuvate Data Platform comes with the below capabilities:

- Offered on various Cloud flavors (MS / AWS / GCP) and On Premise:**
 Overall, our Acuvate Data Platform is platform-neutral (a number of components are Databricks-based), but some components we select, such as IOT Gateway, are unique for each platform.
- Energy extensions supporting Timeseries data:**
 Able to load and manipulate.
- Data engineering:**
 Acuvate Data Platform allows users to ingest, process, clean, and transform large volumes of structured and unstructured data using Apache Spark. It supports ETL (extract, transform, load) processes and optimizes data pipelines for performance and scalability.
- Data analytics:**
 With Acuvate Data Platform, users can perform advanced analytics, including real-time stream processing, SQL queries, and interactive data exploration. It offers visualization tools and integration with popular BI tools like Power BI for creating dashboards and reports.
- Machine learning:**
 Acuvate Data Platform provides an environment for developing, training, and deploying machine learning models. It offers built-in algorithms, support for popular ML libraries, and tools for hyperparameter tuning, feature engineering, and model evaluation.
- Collaboration:**
 The platform features a collaborative workspace where data engineers, data scientists, and business analysts can work together on shared notebooks using different programming languages like Python, Scala, R, and SQL. Version control, commenting, and access controls facilitate teamwork and knowledge sharing.
- Integration with Enterprise Single Sign On (SSO) services:**
 Supporting both OIDC (OpenID Connect) and SAML (Secure Access Markup Language).
- Integration with Acuvate Edge services:**
 Based on EdgeXFoundry Edge Platform services.
- Scalability capabilities**
 - Workspace APIs:**
 For account owner to create workspaces (organize notebooks, libraries and experiments into folders) at scale.
 - SCIM (System for Cross-domain Identity Management)**
 Allows customers to automatically onboard and off-board users.
 - Clusters and Jobs APIs**
 Deploy thousands of clusters and jobs at scale reliably and consistently backed by a massive scale, low-latency API endpoint.
 - Multiple Workspaces**
 Quickly spin up 1000's of new workspaces as needed in the same account with the right access and security policies applied.

Fact Sheet: Acuvate Data Platform

USE CASES

There are multiple use cases for Industries and especially for Energy & Manufacturing:

- For Energy & Manufacturing, all Timeseries data can be streamed (real-time) concurrently from many sources (factories & plants globally) into the Acuvate Data Platform, then stored in native format and for AI-ML to utilize this data as an important source for Predictive Maintenance.**
- Enterprise Data Platform for small and large companies:**
 Scalability is an important aspect to support demand for rapidly growing companies.
- HSSE: Bring (CCTV) camera data into the Acuvate Data Platform and use AI-MV to look for HSSE issues such as Lack of PPE protection, Emission detection, Oil Spillage detection, etc.**
- Having all information to manage automatic warehouses, including driving information in edge devices for time critical data.**

BUILDING BLOCKS

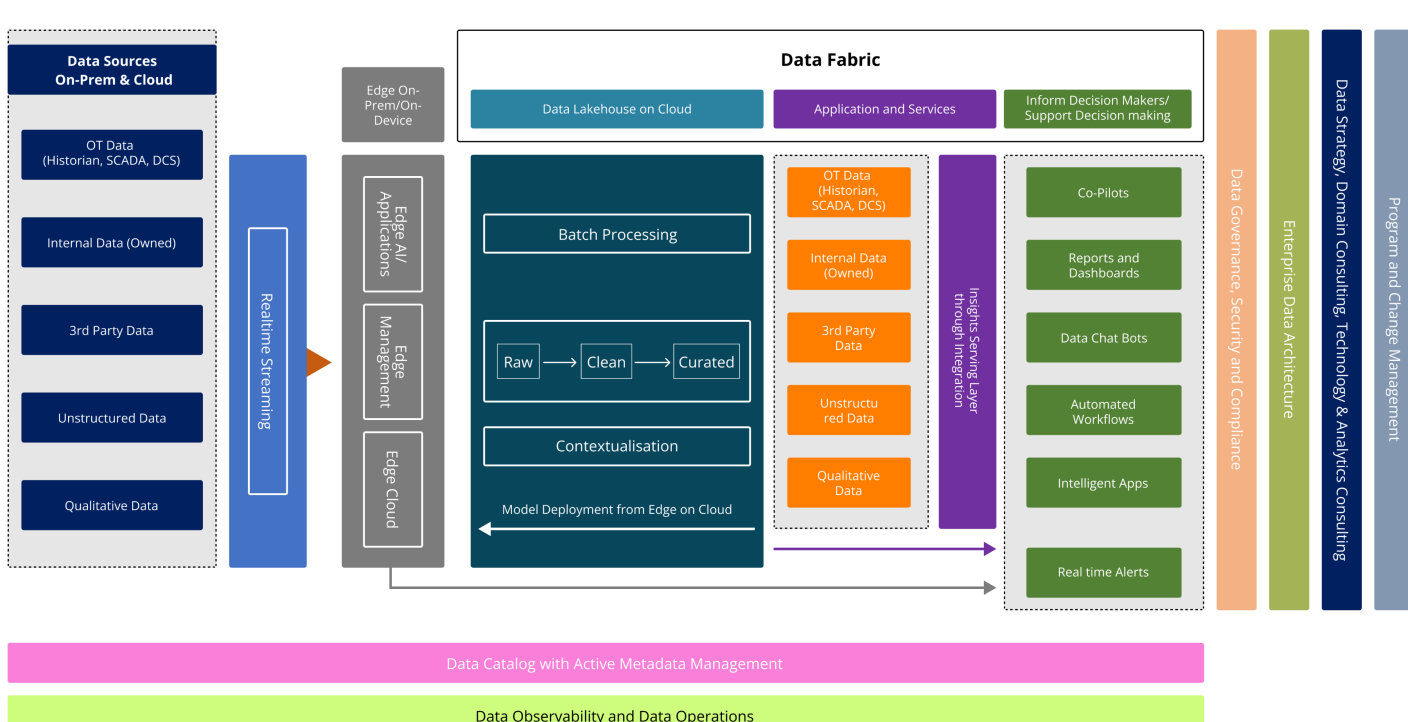
Acuvate Data Platform works by integrating various open-source technologies and proprietary components to provide a unified data analytics platform. Some key components include:

- Apache Spark**
 At the core of Acuvate Data Platform is Apache Spark, a powerful open-source distributed computing framework for processing large-scale data. Spark enables parallel data processing using resilient distributed datasets (RDDs) and supports multiple programming languages, including Python, Scala, and R.
- Delta Lake**
 Delta Lake is an open-source storage layer that brings ACID transactions (atomicity, consistency, isolation, durability) and other data reliability features to big data workloads. It sits on top of existing data lake storage systems (e.g., Amazon S3, Azure Data Lake Storage) and enables versioning, schema enforcement, and data indexing. This makes it easier to manage and maintain data consistency, quality, and performance in data pipelines.
- MLFlow**
 MLflow is an open-source platform for managing the complete machine learning lifecycle, including experimentation, reproducibility, and deployment. It integrates with Databricks to enable tracking of experiments, packaging of code, and sharing of results within the collaborative workspace. MLflow also helps manage model deployment and monitoring in production environments.
- Koalas**
 Koalas is an open-source library that brings pandas API compatibility to Apache Spark, enabling data scientists familiar with pandas to work with large-scale distributed data using the same API they are accustomed to. This simplifies the transition between single-node and distributed data processing, allowing users to leverage Spark's capabilities without learning a new API.
- RTDIP (Realtime Data Ingestion Platform)**
 ([Installation - Real Time Data Ingestion Platform \(rtdip.io\)](#)) supports a number of Energy specific features. RTDIP provides functionality to process and query real time data. The RTDIP SDK is central to building pipelines and querying data, so getting started with it is key to unlocking the capability of RTDIP.

Fact Sheet: Acuvate Data Platform

SUPPORT

- Contact our Advisors** for online access to information on how to exploit these services to solve your business problem(s) advisors@acuvate.com, and you will be contacted within 24 hours.
- Integration:**
 Integrating our solutions with company components, such as data sources and reporting tooling already in place.
- Operational support:**
 Once installed, we offer Operational Support & Management services including all elements you expect such as: Alert management, Preventive Management, Housekeeping, Performance Management, Evergreening, etc., and SLA based.



LICENSING

Contact Acuvate Sales for your Acuvate Data Platform licensing proposal.