DataRiver
open source data management

The MOMIS Data Preparation platform for Industrial Data Spaces

The MOMIS Data Preparation platform is based on a cloud-edge microservices architecture designed for industrial data management, offering advanced functionality for seamless data collection, integration and analysis. With integrated support for Data Spaces, the platform improves collaborative access to data across organisational boundaries, enabling users to securely share and manage data while maintaining compliance with pre-defined standards. The advanced data preparation capabilities provided by the platform allow users to easily and intuitively define pipelines for ingesting, transforming, storing and analysing their data. Data can be organised within Data Spaces, ensuring that it is always available and accessible to the owner and its customers.

"An edge-cloud platform for Data Preparation pipelines in Industrial environments."

Laboratory	DATARIVER
Specialization Area	Digital
Contacts	Andrea Livaldi, Mirko Orsini
Keyword	Big Data, Data Preparation, ETL, Microservices

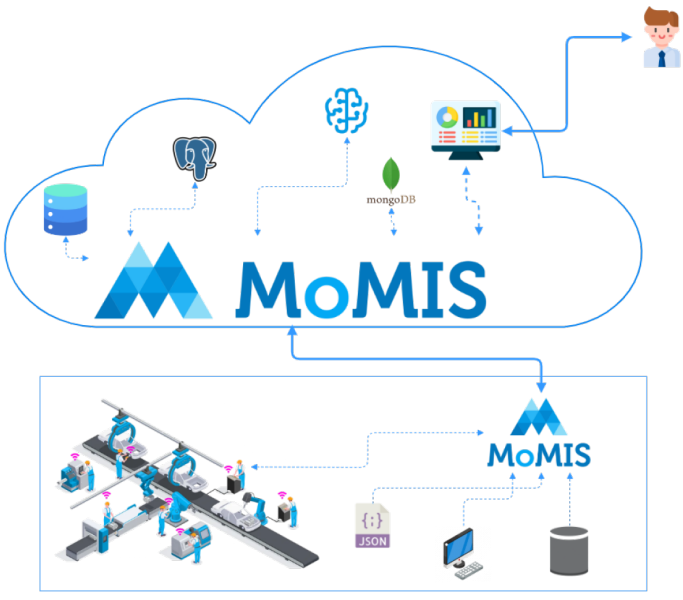

Fig. 1: MOMIS Data Preparation Platform

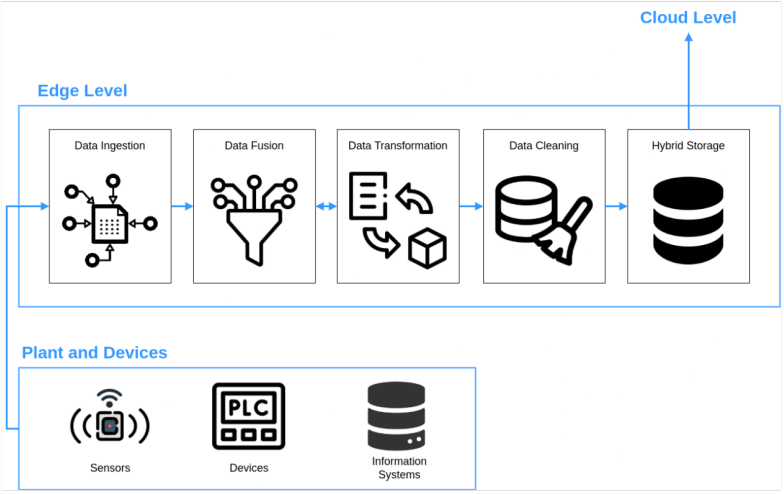

Fig. 2: MOMIS Data Preparation Platform: services overview

Description

The platform is designed for all companies that need to collect and integrate their data into structured and searchable databases. The different services of the platform offer advanced data integration and preparation functions:

•

Data ingestion: allows to connect and read structured and unstructured data from different data sources.

•

Data Transformation, implements structural transformation operations which enable the user to allign different data models into a predefined one.

•

Data Fusion, which allows to blend different data flows through SQL-like joins or unions.

•

Data Cleaning: a service that implements context-specific preprocessing operations optimized for AI algorithms input's data.

•

Hybrid Storage: end building block which allows to save preprocessed data into predefined format into different data storage systems.

Thanks to a unified interface for pipeline definition, these microservices can be configured, switched on and off to extend the functionality of the platform or lighten the computational load when not needed. The microservices are plug and play and can be extended as required. The platform can be extended with customised services through the use of standardised REST APIs. This allows users to deploy multiple instances of MOMIS in different environments, coordinated by a single user interface. Finally, the user-friendly designer also allows non-technical users to design and deploy data preparation pipelines.

Innovative aspects

The microservice architecture of the MOMIS Data Preparation Platform enables the creation and deployment of highly specialised pipelines for industrial data preparation. The platfrom allows data to be ingested from multiple heterogeneous sources and materialised in a unified storage, applying transformations to incoming data and applying artificial intelligence algorithms to discover anomalies.

The edge-cloud architecture of MOMIS helps companies deploy services close to production lines, where latencies are a constraint, or in the public/private cloud, allowing the most advanced analysis algorithms to be run on larger historical datasets.

Potential applications

The MOMIS data preparation platform allows Managers, Production and Sales Managers, IT Managers, Security and Logistics Managers to gain:

- understanding data and timely monitoring of business processes and services provided to customers
- Optimise strategic decisions, control activities and production processes
- Analysing company performance against competitors and market trends
- Make future forecasts by comparing with market and competitor data

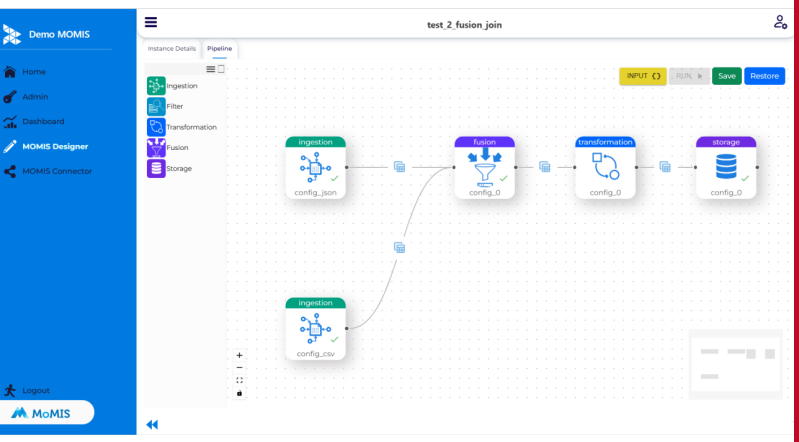

Fig. 3: MOMIS pipeline designer interface

Application example

Adopted in the "DATHA" project, which aims at increasing the competitiveness of enterprises through the strategic use of data, facilitating decision-making processes and enhancing the value of business information. The DATHA project is realised thanks to the European funds of the Emilia-Romagna Region.

MOMIS Data Preparation Platform is used in a variety of Industrial projects including National and Regional projects.
Among them, **DATHA** (DAta Transformation and Homogenization plAtform) project aims to improve business processes by facilitating the management and understanding of data flows within enterprises and focusing on the realization and optimization of the data economy. In order to do so, European Community has defined **Data Spaces**: virtual spaces where context-specific data can be collected from different institutions and companies.
In this project, **MOMIS** serves as a layer that collect data from manufacturing production lines and convert them into formats defined for manufacturing-specific data spaces. This objective is achieved leveraging: the **Ingestion** service for collecting raw data from devices and sensors, the **Transformation** service for adjusting the data structure accordingly to the one defined in the Data Model, the **Cleaning** service for remove invalid values and standardize the data and finally the **Hybrid Storage** for historicise the adapted data into cloud Data Spaces.
Once the data is saved in cloud it can be used to develop AI models for making prediction relevant for the manufacturing context, such as Anomaly Detection and Warehouse automation.

Involved partners	DataRiver
Implementation Time	6 months
Technology Readiness Level	TRL7 - System prototype demonstration in operational environment
Exploitation	The MOMIS Data Preparation Platform is an advanced data integration and analytics platform, useful for companies of the Ceramics, Mechanical, Logistics sectors, including Consulting Services and Facility Management, Global Services, as well as the Biomedical and the Pharmaceutical industries, enabling them to get a **clear and unified view of the company data.**

DATARIVER

Website http://www.datariver.it

Director Mirko Orsini

Published on 20/12/2024

DataRiver Srl is an innovative SME accredited as an Industrial Research Lab of the High Technology Network of the Emilia-Romagna Region. DataRiver is one of the founding partners of Clust-ER – Service Innovation and Clust-ER – Life Sciences and Wellbeing of the Emilia-Romagna Region.

Founded in 2009 as a Spin-Off of the University of Modena and Reggio Emilia, the company is Associate Member of the European Big Data Value Association, and is also a technology provider of the Industry 4.0 Competence Center BI-REX (Big Data Innovation & Research EXcellence). It is also self-certified as Contract Research Organization (CRO) at Italian Medicine Agency (AIFA).

DataRiver provides innovative services and consulting in the areas of Industry 4.0, Big Data Integration & Analytics, IoT, Location Intelligence, Semantic Search and Data Cleaning to both manufacturing and service companies as well as public administrations.

The services offered include the creation of web platforms for the collection and integration of heterogeneous and distributed data, as well as data analysis through Artificial Intelligence and Machine Learning techniques to help improve decision-making, production and predictive processes.

In the health sector, DataRiver deals with data management, data integration and data analysis to design and implement pathology and rare disease registries, investigational drug and medical device clinical trials, and remote patient support and monitoring programs for pharmaceutical, biomedical and homecare companies, and Scientific Institutes for Research, Hospitalisation and Health Care as well as hospitals.