DATHA - DAta Transformation and Homogenization plAtform

DATHA is a software platform that provides innovative tools for data preparation and interchange, enabling business professionals without data science expertise to explore, combine, and transform raw data from production lines into “refined data” in a simple and intuitive way. Furthermore, DATHA offers businesses new approaches to field testing to facilitate decision-making processes, generating value and increasing market competitive advantage. Based on Edge Cloud technologies, the platform enables the processing of information from production lines in Italy’s Data Valley and the cross-sector sharing of data and models. DATHA provides these capabilities in a dynamic and highly scalable way, enabling users to build their own pool of data-oriented services and tailor them to each specific industrial context.

"Cloud Edge platform for data preparation and the interchange of data and models"

Laboratory	DATARIVER
Specialization Area	Digital, Mechatronics and Materials
Contacts	Luca Magnotta, Mirko Orsini
Keyword	Data Preparation, Data Integration, Data Marketplace, Data Interoperability

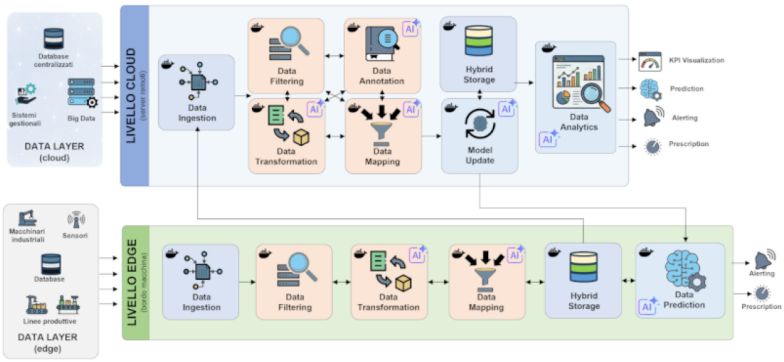


Fig. 1: DATHA - Architecture of the Edge and Cloud Data Preparation Platform

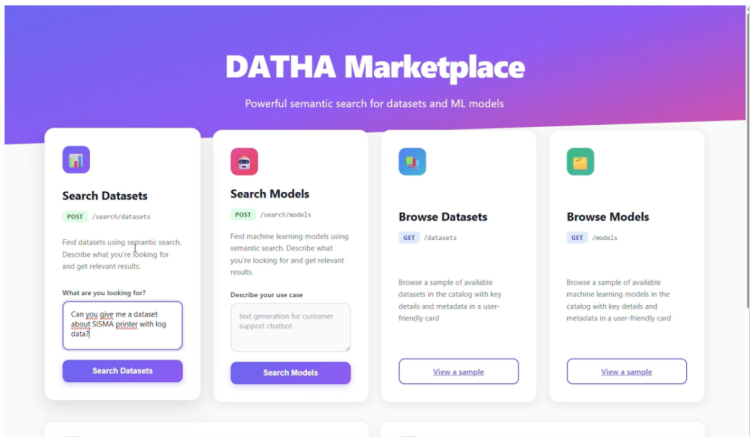


Fig. 2: Main screen of the DATHA Marketplace (DATHA's Enterprise Data Marketplace) for searching for data and models using a natural language interface

Description

The DATHA project is funded by the Emilia-Romagna Region (PR-FESR 2021-2027, Action 1.1.2). The solution developed—the DATHA Platform—promotes the strategic use of information in manufacturing SMEs by simplifying access to the big data generated by supply chains that adopt IoT technologies in edge-cloud environments.
As a self-service data preparation and analytics environment, it is designed to enable users, regardless of their technical expertise, to create custom analytics workflows. The platform consists of two integrated modules:

- **Data Preparation Platform:** An intelligent microservices infrastructure for Big Data preparation and analytics in the context of Industry 4.0 and 5.0. It manages the entire lifecycle of raw data through acquisition from heterogeneous sources, transformation, cleaning, and semantic fusion of data streams. The processed data is then analyzed using machine learning to enable predictive maintenance, anomaly detection, and real-time KPI calculation.
- **Enterprise Data Marketplace:** An ecosystem that enables the cataloging, sharing, and leveraging of industrial assets. Datasets and predictive models are searched for using natural language thanks to RAG technologies, overcoming the limitations of keyword searches and transforming knowledge into digital assets that can be shared across supply chains.

In this way, DATHA democratizes business analysis, making it accessible even to operators and managers.

Innovative aspects

The DATHA platform is an innovative solution for data preparation, analysis, and interoperability. Consisting of a Data Preparation Platform and a Data Marketplace, it is based on MOMIS Industrial IoT technology and utilizes a microservices architecture. Its key innovative features are:

- **Advanced Semantic Integration:** It standardizes diverse formats and protocols in real time into a common semantic model.
- **Low-Code Accessibility:** Thanks to LLMs, users can configure data preparation services and create indicators using natural language, which are automatically translated into operational logic.
- **Enterprise Data Marketplace:** It enables the secure sharing and reuse of datasets and machine learning models both within and across organizations. Unlike traditional standalone solutions, which require expensive teams of data scientists, DATHA democratizes analytics by enabling non-technical users to transform raw data into “refined data.” In addition, the cloud-edge infrastructure ensures horizontal scalability, reduces latency, and eliminates data fragmentation directly on the shop floor.

Potential applications

The DATHA platform is designed to be accessible even to users without specific IT skills (such as data engineers or data scientists), making it suitable for small businesses and SMEs. It is designed to operate in a wide range of sectors—including manufacturing, automotive, agri-food, and IT services—and to facilitate cross-sector data sharing and analysis. The DATHA Marketplace also enables strategically valuable applications: the creation of an ecosystem for the exchange of data and models between different organizations, while respecting intellectual property rights.

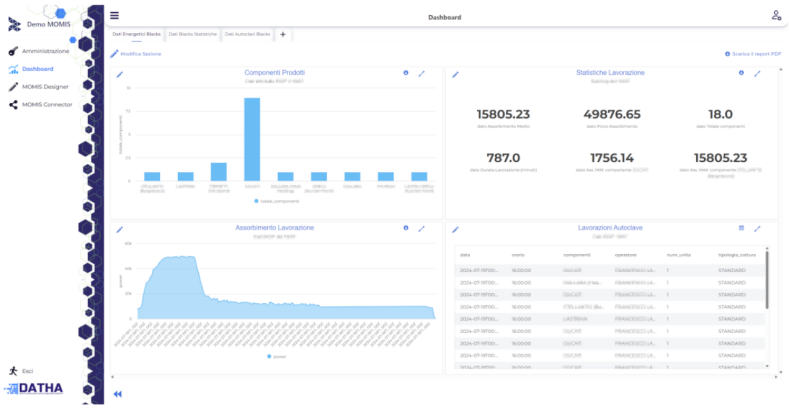


Fig. 3: DATHA - Production process monitoring dashboard for the manufacturing and automotive sectors.

Involved partners

CIRI ICT - University of Bologna
BI-REX Competence Center
AIRI – University of Modena and Reggio Emilia
Quix S.r.l.
Imola Informatica S.P.A.
Blacks S.r.l.
Toschi Vignola Srl

Implementation Time

3 person/month

Technology Readiness Level

TRL7 - System prototype demonstration in operational environment

Exploitation

The DATHA platform has been used in various real-world settings for the continuous monitoring of production processes in the manufacturing and automotive sectors.

Application example

DATHA Platform for optimizing the manufacturing processes of advanced composite components for high-precision industries such as aerospace, automotive, and motorsports

The use case, developed as part of the DATHA project and used in a manufacturing environment for advanced composite materials, addressed the challenge of optimizing complex planning processes characterized by non-homogeneous batches and often urgent, small-scale orders. Historically, raw production data was recorded by MES and ERP systems but remained unused and available only for occasional manual checks. Through the integration of the DATHA platform, the company digitized the entire process, transitioning from manual data extraction to automated, continuous real-time data collection that links the physical parameters of the machinery to energy consumption monitoring. The platform has made it possible to correlate energy efficiency with batch characteristics and identify the machining configurations with the highest energy consumption, providing a solid quantitative basis for more informed strategic decisions.
These results enable the automation roadmap to be planned with greater insight, preserving the value of advanced craftsmanship and transforming it into a scalable, efficient, and data-driven industrial model.

DATARIVER

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.

Website http://www.datariver.it

Director Mirko Orsini

Published on 11/06/2026