



AI-DAPT

Newsletter #3

May 2025

As we continue to push forward with innovation in AI and data pipeline automation, we are excited to report on the beginning of the project technical implementation phase, unveiling the technological backbone behind our vision: the AI-DAPT Reference Architecture. Designed to support explainable, adaptive, and human-centric AI workflows, this architecture lays the foundation for a trustworthy and scalable AI platform.

Inside AI-DAPT: From Vision to Platform

AI-DAPT is built on the ambition to reshape how artificial intelligence and data workflows are designed, automated, and trusted in real-world environments. Following an intensive requirements phase and the definition of the project's research agenda, the platform is now entering its **technical implementation phase**, the transition from vision to platform.

At the heart of this transition lies the **AI-DAPT Reference Architecture**, a modular, multi-layered framework that provides the technical foundation for **explainability, automation, human-in-the-loop control, hybrid AI models, and continuous adaptation**.

Tackling the Real Challenges in AI: AI-DAPT was conceived to address two central problems in today's AI landscape:

- A surplus of **unusable or non-actionable data** that must be transformed into quality inputs.
- The existence of **legacy scientific models** that, while trustworthy, need to be integrated into modern AI systems.

To solve this, AI-DAPT promotes a **data-centric mentality** along two axes: (i) **Data Axis**, Automating data definition, harvesting, annotation, cleaning, synthesis, and observability. (ii) **AI/Model Axis**, Enabling **hybrid AI pipelines** that blend machine learning with science-based (first-principles) models. This approach is embodied in the AI-DAPT Reference Architecture, a robust blueprint that harmonizes research objectives with operational needs.

A Collaborative Technical Foundation:

Developed collaboratively by all technical partners, the architecture was designed through an iterative, co-design process to ensure alignment with real-world needs across industries and user roles. Structured using the C4 architecture model, it promotes clarity and modularity from high-level system views to detailed service interactions.

Built for Adaptability and Integration:

Designed to serve the diverse needs of **data scientists, business users, and data engineers**, the platform ensures that: (a) **Human-in-the-loop control** remains integral at all stages; (b) **Explainability and transparency** are built-in rather than bolted on; (c) AI systems remain **observable, adaptive, and secure** as they evolve over time.

The result is a platform that doesn't just automate AI pipelines, but also adapts to new data, integrates legacy models, and respects the domain-specific constraints of industrial use cases.



AI-DAPT High Level Architecture

The AI-DAPT Reference Architecture is built around **eight integrated service layers**, each playing a key role in enabling trustworthy, adaptive, and explainable AI pipelines. This architectural framework is formally presented in **Deliverable D4.1 “AI-DAPT Reference Architecture and API Documentation”**, which defines the system’s structure, components, and interactions. While the deliverable will be made publicly available as soon as it is approved by the European Commission, readers can already explore some more detail in our recent [blog post](#).

The eight layers of the AI-DAPT architecture are:

1. **Data-AI Insights Services** – synthetic data generation, data valuation.
2. **Data Lifecycle Management Services** – data documentation, reconciliation, feature engineering.
3. **AI Lifecycle Management Services** – model observability, explainability, security, and tracking.
4. **Data Pipeline Services** – harvesting, annotation, cleaning.
5. **Data-AI Execution Services** – pipeline execution and experimentation environments.
6. **Data-AI Pipeline Monitoring Services** – real-time oversight and adaptive AI mechanisms.
7. **Platform Management Services** – authentication, notification, interoperability.
8. **Scalable Storage Services** – efficient, modular storage with tools like **MinIO**, **MongoDB**, and **Virtuoso**.

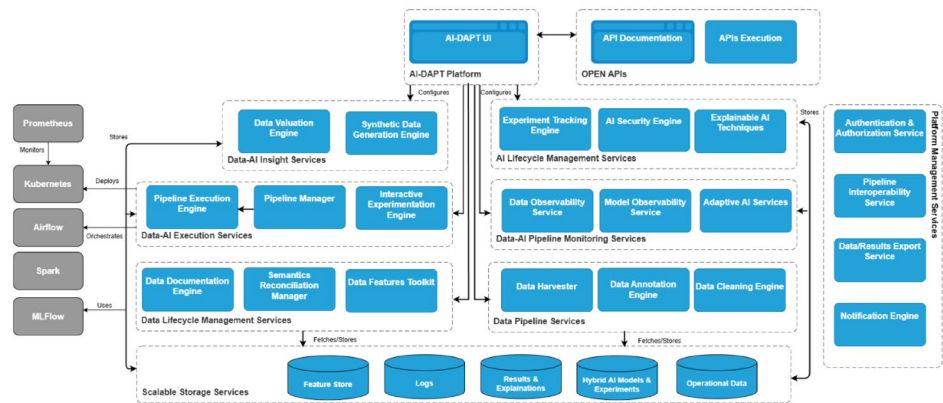


Figure 1: AI-DAPT High Level Architecture

The architecture leverages powerful open-source tools including **Apache Airflow** for orchestration, **MLflow** for model tracking, **Kafka** for streaming, and **Apache Spark** for scalable processing.

From User Journeys to Technical Workflows: Building on the user roles identified, the architecture supports tailored experiences for each of AI-DAPT’s key users: (i) **Data Scientists** use the platform to design, test, and optimize data and AI pipelines; (ii) **Business Users** engage through intuitive interfaces to access explainable insights; (iii) **Data Engineers** – manage robust infrastructure, pipeline orchestration, and data quality. For them, AI-DAPT defines **three main user journeys**, namely “Data Ingestion & Sculpting”, “Explainable AI (XAI)” and (iii) “XAI Pipeline Insights and Results”, each representing a distinct user role and their interaction with the platform, being implemented as **technical workflows**, orchestrated using tools like **Apache Airflow**.



Highlights and What’s Next!

Over recent months, AI-DAPT has actively contributed to major community events, including the **ADR Partnership Awareness Day**, and **DATA Week 2025**, where we contributed to the workshop on “Synthetic data: legal, ethical and technical implications”.

This group photo was taken during the latest General Assembly of the AI-DAPT project in Milan, where partners came together to align on progress and plan the next steps in advancing trustworthy AI.

Looking ahead, we’re excited to invite you to join us at **ADRF25**, taking place in Stavanger, Norway, on the **23rd and 24th of September**. We’ll be contributing to two workshops. The first, organised in collaboration with the HORIZON-CL4-2021-HUMAN-01-01 cluster of projects, is titled **“Efficient and trustworthy AI frameworks: making the best of cloud-edge European data for scalable real-world industry deployments.”** The second workshop is promoted by the [HELEN network](#). Don’t miss the chance to register by **6th June** for [early bird access](#).



<https://www.ai-dapt.eu/>

[@AI-DAPT](#) [@AI-DAPT](#)

[/company/ai-dapt/](#)