Newsletter #2

January 2025

As we enter the second year of AI-DAPT, the project's Research Agenda is now well-defined, paving the way for trustworthy and adaptive AI pipelines with a strong focus on data quality, synthetic data generation, hybrid AI models, and Explainable AI for transparency. We are eager to bring our innovative platform to life and test it across four demonstrators in Health, Robotics, Energy, and Manufacturing. Stay tuned to see how AI-DAPT is shaping the future of AI innovation!

Demonstrators Overview

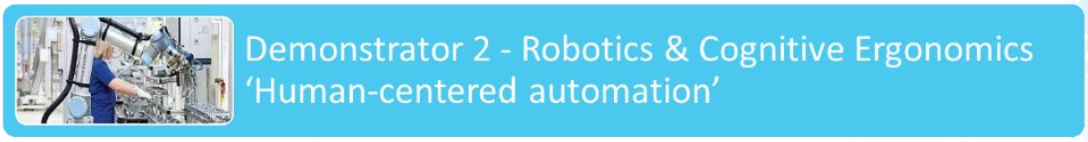
To demonstrate the AI-DAPT project's applicability, the platform will be tested and validated in four different real-life demonstrators. Each focusing on a specific domain: Health, Robotics, Energy, and Manufacturing.

- ***Demonstrator 1:* Personalized Medicine Based on Non-invasive Glucose Monitoring**

  Diabetes mellitus is a chronic condition requiring reliable glucose monitoring. Current methods are invasive, but Photoplethysmography (PPG) offers a non-invasive alternative using optical technology to detect blood volume changes. Demonstrator 1, involving MCS Data Labs & CHARITE from Germany, integrates PPG with AI, gathering blood glucose data to both detect and manage diabetes. This demonstrator targets an 80% reduction in model training time for detecting diabetes.

- ***Demonstrator 2:* Robotics & Cognitive Ergonomics towards Human-Centred Automation**

  Wearable technology has become a key data source for AI, including detecting stress levels. Demonstrator 2, led by Intellimech & MADE-CC from Italy, focuses on optimizing human-machine collaboration by monitoring work environments and employees' emotions and mental states. Goals include a 10% improvement in worker well-being, 25% better product quality through reduced waste, and 5% faster task completion.



- ***Demonstrator 3:* Cross-vector Residential Demand-Response Through Smart Heating**

  Household energy management is critical amid rising costs of energy bills and inefficient building stock. Demonstrator 3, driven by ZENITH & DOMX, in Greece, is dedicated enhancing personalized load and price forecasts in energy consumption, aiming for 50% error reduction, 30% peak-load reduction, and 5% increased peak-load flexibility by accurately profiling energy demand and consumer behavior.

- ***Demonstrator 4:* Predictive Maintenance of Production Assets**

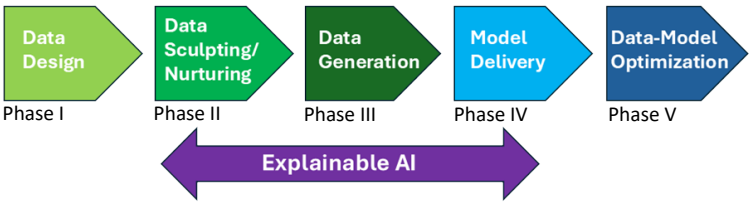
  Efficient equipment maintenance is essential in manufacturing. Demonstrator 4, involving OHS & BIBA in Germany, uses AI to streamline maintenance tasks, reduce downtime, and extend equipment lifespan. Key targets include 20-30% faster repairs, 15-20% longer equipment life, and 90-95% less effort in managing maintenance contracts.

AI-DAPT Research Agenda

The AI-DAPT Research Agenda is designed to improve the lifecycle of automated data and AI pipelines. It has been developed through extensive research, collaborative stakeholder consultations, and iterative refinement, addressing critical challenges and identifying open questions in the field. By focusing on key innovation areas, the agenda ensures that AI-DAPT remains at the cutting edge of AI development.

Supported by extensive literature review and the technology radar available in our website, this agenda is organized around 5 phases of the Data/AI pipeline lifecycle, each addressing unique and crucial challenges. The 1st phase (I), **Data Design**, explores strategies to effectively identify, collect, and structure data to meet the requirements of AI models. In the 2nd phase (II), **Data Nurturing**, the focus is on enhancing data quality and reliability using AI techniques for annotation, cleaning, and preparation, while ensuring transparency and fairness in data curation. The 3rd phase (III), **Data Generation**, addresses the limitations of real-world data by investigating synthetic data generation techniques that can ensure both utility and fairness. The 4th phase (IV), **Model Delivery**, focuses on hybrid science-guided AI models, for improved accuracy and robustness in industrial and real-world applications, as well as optimized training, evaluation, and deployment processes. Finally, the **Data-Model Optimization** phase (V) focuses on adaptive methodologies for bias assessment and correction, real-time detection of data drift and model degradation, and frameworks that enable continuous refinement of models in dynamic environments.

Explainable AI (XAI) ties all these phases together. By embedding transparency, interpretability, and trustworthiness into every stage of the pipeline lifecycle, it allows the understanding of AI insights, which is critical for both users and stakeholders. With explanations at every point of decision-making, this becomes both inclusive and accountable in its coverage across the pipeline, while always keeping the human-in-the-loop.



The agenda, released as deliverable D1.2 is not a static blueprint but a living, dynamic document that moves and is continuously updated to incorporate advancements in the AI landscape. This flexibility guarantees that AI-DAPT maintains its leadership in innovative pipeline solutions with the ability to meet new challenges and leveraging new opportunities.

As AI-DAPT progresses, the Research Agenda will continue to guide the project toward impactful outcomes. By fostering collaboration and embracing innovation, we invite you to join us in shaping the future of automated AI pipelines and unlocking the vast potential of data and AI technologies.

Pipelines in AI-DAPT Demonstrators

Core data pipelines typically involve Phases I, II, and III for preparing training data, or Phases I, II, and V for ongoing data collection. Common AI pipelines include training models (Phases I-IV) or inference on new data using trained models (Phases I, II, IV, V, or just IV, V if an independent data pipeline is in place). During the project's operational stage, all phases can be combined into Data/AI pipelines, either scheduled or event-triggered.

Exploratory Data Pipelines

These pipelines include:

- **Phase I: Data Design** – Harvest raw data, analyse properties, and assess quality.
- **Phase II: Data Nurturing for AI** – Preprocess data, fix errors, and engineer features.
- **Phase III: Data Generation** – Generate synthetic data when available data is scarce.

Exploratory AI Pipelines

These pipelines train and validate (hybrid) AI models, assuming prepared training data. They include:

- **Phase I: Data Design (partial)** – For data ingestion.
- **Phase IV: Model Delivery** – Experiment with models and assess performance. Additional **XAI pipelines** may explain input feature effects and extract insights for demonstrators.

MLOps Pipelines

MLOps pipelines are established workflows suitable for production environments, supporting scheduled or event-triggered operations. Examples include:

- Delivering predictions on new data (Phases I, IV, V).
- Collecting and appending new data for model retraining while monitoring data/model drift (Phases I, II, IV, V).

New Collaborations

AI-DAPT has recently become a member of the Horizon Europe Legal and Ethics Network (HELEN), which is dedicated to examining the ethical and legal implications of innovations in European research. HELEN consists of 14 EU-funded projects and focuses on addressing the challenges brought about by technological advancements across multiple sectors. The network encourages collaboration among its members and seeks to create a framework that aligns technological progress with ethical and legal standards, ensuring that innovations are secure and beneficial to society.

http://www.helen-hub.eu/ https://www.linkedin.com/groups/10015073/

The project has received funding from the European Union's Horizon 2023 research and innovation program under the Grant Agreement No. 101135826