AI Platform for Integrated Sustainable and Circular Manufacturing

Deliverable

D5.3 System Integration and Validation Report – 1st version

Actual submission date: 08/02/2024

Co-funded by
the European Union

Project Number:	101058585		
Project Acronym:	Circular TwAIn		
Project Title:	AI Platform for Integrated Sustainable and Circular Manufacturing		
Start date:	July 1st, 2022	**Duration:**	36 months
D5.3 System Integration and Validation Report – 1st version			
Work Package:	WP5		
Lead partner:	Nissatech		
Author(s):	Nenad Stojanovic (Nissatech), Vincenzo Cutrona (SUPSI), Maurizio Megliola (GFT), Irene Zattarin (GFT), Nikos Kefalakis (INTRA), Christos Stavrogiannis (INTRA), Sotiris Koussouris (Suite5), Iason Tzanetatos (CORE), Giovanni Di Oro (NOVA)		
Reviewers:	Mahdi Mohammadian (Polimi), Santiago Muíños Landín (AIMEN)		
Due date:	31/12/2023		
Deliverable Type:	R-Document, report	**Dissemination Level:**	PU - Public
Version number:	1.0		

Revision History

Version	Date	Author	Description
0.1	1/12/2023	Nissatech	ToC
0.2	14/12/2023	GFT, SUPSI, INTRA, NOVA, CORE, Suite5	Initial content
0.7	21/12/2023	GFT, SUPSI, INTRA, NOVA, CORE, Suite5	Detailed content
0.8	10/01/2024	GFT, SUPSI, INTRA, NOVA, Suite5	Completed content for T5.2-T5.5
0.82	12/01/2024	INTRA	Added content related to Product DT (T5.3)

0.85	19/01/2024	CORE	Provided input for the Petrochemical Pilot Section
0.9	26/01/2024	Nissatech	Preparing all sections for internal review
0.99	07/02/2024	Nissatech	Version ready for submission
1.0	08/02/2024	ENG	Final coordinator review before submission

Table of Contents

List of Figures

List of Tables

Definitions and acronyms

AAS	*Asset Administration Shell*
AI	*Artificial Intelligence*
CA	*Certificate Authorities*
CDS	*Circular Data Space*
CEAP	*Circular Economy Action Plan*
CSR	*Certificate Signing Requests*
CTDA	*Circular Twain Data Agent*
DAPS	*Dynamic Attribute Provisioning Service*
DAT	*Dynamic Attribute Tokens*
DME	*Dimethyl Ether*
DoA	*Description of Action*
DPP	*Digital Product Passport*
DS	*Data Space*
DT	*Digital Twin*
EC	*European Commission*
EDC	*Eclipse Dataspace Components*
EO	*Ethylene Oxide*
EoL	*End of Life*
EU	*European Union*
GA	*Grant Agreement*
HSM	*Hardware Security Modules*
IAM	*Identity and Access Management*
IDP	*Industrial Data Platform*
IDS	*International Data Space*
IDSA	*International Data Space Association*
JSON	*JavaScript Object Notation*
KPI	*Key Performance Indicator*
OPC UA	*Open Platform Communications Unified Architecture*
PDDL	*Planning Domain Definition Language*
PPE	*Personal Protection Equipment*
PC	*Project Coordinator*
TC	*Technical Coordinator*
THB	*Trial Handbook*
TPM	*Trusted Platform Module*
UC	*Use Case*
WEEE	*Waste of Electronic and Electrical Equipment*
WP	*Work Package*
XML	*eXtensible Markup Language*

Disclaimer

Executive Summary

In Circular TwAIn, WP5 is devoted to the development of the methods and tools to support data collection, AI-based processing and DT-based modeling, enabling Pilots' circular scenarios.

The present deliverable reports on the activities undertaken in the context of Task 5.6 ("System Integration, Testing, and Technical Validation"), and it aims to verify the adherence of the proposed technologies and tools adopted or developed to the requirements collected on data for the circular value chain and to validate other WP5 task outcomes with respect to Pilots expectations.

The first issue to be solved was the absence of a reference standard methodology to be applied in the task; for this reason, a Verification and Validation Methodology has been created starting from the approach proposed in AI REGIO[1] due to common AI technologies and partners involved, and it will be pursued for the whole task duration (i.e., also the activities to be carried on during the second phase of the Project have been detailed). The rationale behind such methodology is the need to assess in a holistic way the goodness of the requirements collected, and the technical dimension considering the Pilots needs and constraints.

The main results from active tasks (T5.1–T5.5) are presented in Section 3 (requirements and conceptual models) and Section 5 (application in the pilots), showing on one hand the functional capacity of the modules to implement various types of scenarios, but also the suitability to efficiently realize the requirements from the pilots. The focus was on the WEEE pilot, with a clear plan to transfer the knowledge and experience to the other two pilots.

Finally, the structure of this document and, with some adaptations, the methodology defined and followed here, has been practiced also in the D5.3 twin deliverable (D4.3 "*System Integration and Validation Report - 1st version*"). In fact, in the other Circular TwAIn's WP dedicated to technical developments (i.e., WP4,) there is a parallel task on Integration, Testing and Technical Validation (i.e., T4.6).

[1] https://www.airegio-project.eu/

1. Introduction

1.1 Purpose and Scope

This deliverable D5.3 "*System Integration and Validation Report – 1st version*" is aimed to report the results of the technical verification and validation of the Circular TwAIn WP5 outcomes.

In this first version the focus is on the verification of the applicability of the technologies selected (to be developed or adopted) to Pilots' UCs. Accent will be given also to possible scenarios opened by such technologies, given as suggestion to Pilots' or as suitable future applications enabled by Circular TwAIn.

The second version of the deliverable, expected at M36, will include the final validation (i.e., proved by the actual execution of Pilots' UCs) of the technologies in scope for WP5.

1.2 Structure of the document

Apart from the Introduction and the Conclusion, the document is structured as follows:

- Section 2 defines the Methodology adopted to execute the validation of the Circular TwAIn's WP5 outcomes: the first phase of the methodology has been followed for the 1st version of the validation, while the second phase will be the subject of the 2nd and final validation;
- Section 3 reports the analysis of the requirements relevant to WP5 in the context of data, AI and Circularity;
- Section 4 provides the details about the technologies developed and adopted in WP5, considering the outputs of the 5 development tasks of WP5 (i.e., from T5.1 to T5.5);
- Section 5 maps the requirements identified and the technologies proposed to all Pilots' domains (WEEE, BATTERY, PETRO), providing insights on the actual (M18) deployment expectation.

2. Verification and Validation Methodology

One of key objectives of this deliverable is to undertake the verification and validation of the outcomes from Circular TwAIn's WP5. This involves scrutinizing selected components of the Data and Digital Twins and developed technologies to ensure their compliance with the established Project requirements. In D2.2, various requirements were identified pertaining to challenges and enhancements within the realm of a Circular Data Space (CDS). These fundamental requirements are subsequently contextualized within each Pilot, as outlined in the Project Trial Handbooks[2] and other deliverables, such as the deliverable D5.1. However, there is a lack of direct correlation between the WP5 outcomes and the overarching requirements, both general and Pilot-specific, that they are anticipated to fulfil.

Figure 1 provides an overview of the methodology defined and adopted to verify and validate the outputs of Circular TwAIn's WP5. The proposed methodology unfolds in six steps organized into two phases, the first one more oriented towards the verification, the second towards the validation.

The initial phase (left side of Figure 1) encompasses a preliminary verification of the technologies developed or adopted in WP5 based on the requirements identified in other WPs (mostly in WP2 and reported in D2.2) and follows with their validation with respect to the tools needed to enable Pilots' UCs execution. The second phase will be carried out during the second period of the Project (M19-M36) and it will include the final validation of the assets developed under WP4 with respect to Pilots' feedback. Its take aways will be reported in D4.6 (due at M36).

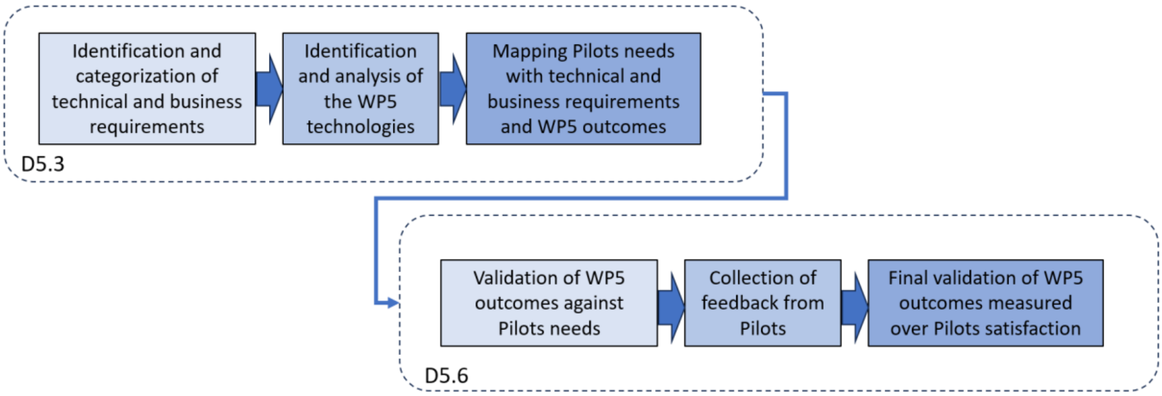

Figure 1: Verification and Validation methodology

The first step of this initial phase involves mapping requirements related to Data Space Components, Data Ontologies and all the technologies involved in this work package. This process aims to segregate, where feasible, the pertinent technical and business requirements, aggregating them into macro-categories. These macro-categories cover both technical and business requirements in three application domains: Data, AI, and Circular Economy. The activities performed are reported in Section 3.

[2] The Trial Handbook is a confidential "living" document and the central information source for all tasks regarding the pilots, defining in detail the entire process carried out throughout the experiment and the outcomes and results of the activities performed.

The second step, detailed in Section 4, focuses on the identification of the set of technologies to be developed or adopted in WP5 for constructing and support the AI and Digital Twin adoption. All these components undergo detailed description and validation against the aforementioned requirements, ensuring their suitability to address the identified needs in terms of data, AI and Circularity.

The third and final step (reported in Section 5) of the initial phase consists in a preliminary Pilot validation: for each Pilot, a subset of requirements has been chosen in relationship with the needs in Project documentation.

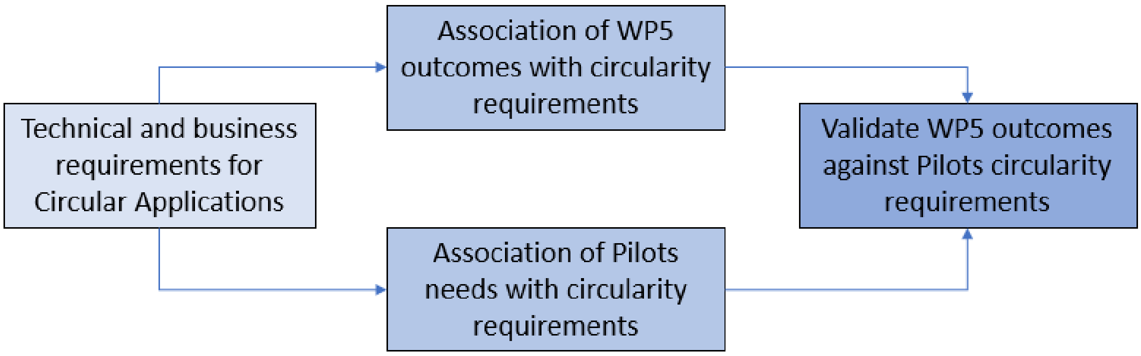

Figure 2: Detail of the Verification and Validation Tasks on the first phase

The second phase will involve the selection of suitable WP5 outcomes able to cover Pilots' needs, the validation of the selected architecture for each Pilot (i.e., deployment and integration) including insights on the iteration needed to achieve the desiderata, thanks to their feedback. Further details will be expounded upon in Deliverable 5.6, representing the second iteration of the current document.

3. Technical, Business and Circularity Requirements for Data, Data Space and Circular Applications

In this section the requirements identified in WP2 and reported in D2.2 ("*User Scenarios, Requirements and Performance Indicators - 1st version*"), especially the ones related with the objectives of WP5 (i.e., data, AI) and the application domain (i.e., Circularity), have been further analysed and better described highlighting the technology context. Furthermore, in the timeframe between the D2.2 submission in (M9) and the time this deliverable is written (M18), some extra requirements have been identified and reported here.

3.1 Requirements from a Circular system

The following requirements list summarizes the information collected in D2.2.

3.1.1 Technical requirements

- Information structured under a **common ontology and semantics**: structure the information with under appropriate ontology and semantics is fundamental to adopt a clear terminology and information sharing, foster the collaboration between the stakeholders and implement effective and scalable circular economy practices.
- **Material traceability**: to close the loop of the materials and create circular value chains it is fundamental to track the material flow to avoid losses along the chain, to control the material flow and create closed-loop material flows, and to be able to certify the origin of the materials.
- **Info traceability** is fundamental to be able to track the material flow, the products and the waste along the entire value chain. By tracing the material flow the companies can identify the opportunities of waste reduction, improve the usage of the resources and create new circular business models.
- **Information and material sharing** among the different actors of the value chains: to create closed loop material flow, it is mandatory to guarantee a way to exchange both info and material flow between the different actors of the value chains.

3.2 Data Requirements

Data requirements encompass a series of features that data must exhibit in the context of a Circular Data Space. In the context of a CDS, understanding and defining data requirements become crucial for effective information management. The following requirements list summarizes the information collected in D2.2:

- **Data Quantity:** Quantity is a critical data requirement in the context of a CDS as it determines the volume of information flowing within the system. The continuous circular nature of the Data Space implies a dynamic and ongoing exchange of data points. Being able to envisage and manage the quantity of data generated, processed, stored and recycled is essential for maintaining system efficiency and preventing data overload. It allows for the optimization of infrastructural resources but also impacts the gathering of insights through data-driven prediction systems, ensuring the seamless flow of data within the circular framework.

- **Data Quality:** Quality is a fundamental feature that data must exhibit. High-quality data for circular economy should be intrinsically good, contextually appropriate for circularity needs, clearly represented and accessible to the data consumer. There is a positive correlation between optimal data quality and maximising the economic value of data, especially when the aim is to exploit this data in the long term through data recycling.
- **Data Confidentiality:** Confidentiality refers to the property of certain sensitive information to be protected from unauthorised access or disclosure. It ensures that only authorized individuals or entities have access to certain data, preventing unauthorized parties from viewing, altering, or distributing the information. Data confidentiality is vital for safeguarding sensitive information, protecting privacy, and ensuring legal compliance. Confidentiality requirements for a Data Space involve the implementation of measures to safeguard sensitive information and restrict access to authorized entities.
- **Data Portability:** Portability refers to the ability of individuals or organizations to move and transfer their data seamlessly between different systems, platforms, or services. It ensures that users can easily access, share, or migrate their personal or business data across various applications, devices, or environments while maintaining its integrity and usability. In a circular context, data crosses different stages in the product lifecycle, therefore with the possibility of involving several systems in the process. Data portability ensures all systems to be able to preserve data structure and integrity, enabling data circularity.
- **Data Standardization:** Standardization is a property of data that involves establishing and adhering to a consistent format, structure, and representation for information. It ensures uniformity in the way data is collected, stored, processed, and exchanged within a system or across different systems. Standardizing data includes defining common data formats, units of measurement, coding conventions, and other specifications, making it easier to integrate and analyse information. In a circular economy, where resources are reused, recycled, and repurposed, standardized data ensures consistency in reporting, tracking, and sharing information about product lifecycles, materials, and environmental impact.
- **Data Interconnectedness:** Interconnectedness refers to the data property of being linked together and able to exchange information between various datasets, systems, or devices. In a highly interconnected data environment, different components, platforms, or sources are integrated to enable efficient communication and data flow. This property ensures different stakeholders, such as manufacturers, suppliers, consumers, and recyclers to exploit data for their needs.
- **Data Recycling:** Recycling is a property that involves the systematic reuse and repurposing of data throughout its lifecycle. Like the principles of material recycling in a circular economy, data recycling emphasizes extracting enduring value from existing datasets. This practice contributes to sustainability efforts by minimizing data redundancy and waste. In a Data Space within a circular economy framework, efficient data recycling enables stakeholders to derive insights from previously collected information for a more resource-conscious approach to data management.

3.3 Requirements from an Artificial Intelligence system

Overall, the requirements for an Artificial Intelligence (AI) system (i.e., a computer-based system designed to analyse large amounts of data, detect patterns, and learn from experience to make predictions and decisions) can vary depending on the specific use case and application; as such, below a non-exhausting list of generic system requirements along with their description is presented:

- **Data availability**: Typically, AI systems require large amounts of high-quality data to train on, thus the data must be relevant to the specific task or problem the AI system is being developed for.
- **Algorithms availability**: AI systems need powerful algorithms that can analyse and interpret complex data patterns, make decisions based on this data, and learn from new data as these becomes available.
- **Computing power**: AI systems require significant computing power to process large amounts of data and run complex algorithms. This may involve usage of specialized hardware such as GPUs or TPUs.
- **Accuracy**: AI systems must be accurate and reliable in their predictions and decisions; this in turn, requires rigorous testing and validation to ensure that the AI system is functioning as intended.
- **Transparency and explainability**: AI systems should be transparent and explainable, meaning that users should be able to understand how the system arrived at its decisions and what factors were considered in the process. This is especially important in high-stakes applications such as healthcare or finance.
- **Security and privacy**: AI systems must be secure and protect user privacy, especially when dealing with sensitive data (such as personal information).
- **Scalability**: AI systems should be able to scale up or down to accommodate changes in data volume or user demand without sacrificing performance or accuracy.
- **Integration**: AI systems must be able to integrate with other systems and applications in a seamless and efficient manner, to ensure that they can be used effectively within larger workflows and processes.

The following requirements list summarizes the information collected in D2.2.

3.3.1 Technical requirements

The role of Artificial intelligence in the scope of Circular TwAIn, considers applications on material, product, process and plant whilst providing a framework for human interpretable insights.

Hardware requirements:

From a functional point of view, it is crucial for the selected hardware platforms to be able to handle the processing requirements in an efficient manner in terms of processing speed. The optimal operation and collaboration of the various systems (i.e., robotic systems, etc.) is heavily dependent on the hardware platform. Moreover, to facilitate the continuous operation and accuracy of DTs, the hardware platform should be able to re-train the deployed models. When it comes to the hardware requirements for an AI system, these again will mostly depend on the specific application and the complexity of the tasks it needs to perform. In general, AI systems require significant processing power to analyse and learn from large datasets; this can be achieved through the use of high-performance processors, such as Graphics Processing Units (GPUs), Tensor Processing Units (TPUs), or Field Programmable Gate Arrays (FPGAs). In addition, AI systems require also large amounts of memory to store and process data efficiently. This can be achieved through the use of high-capacity Random Access Memory (RAM) or Solid-State Drives (SSDs). Similarly, significant amounts of storage is required from AI systems to store large datasets and trained models. Typically, this can be achieved through the use of hard disk drives or Network-Attached Storage (NAS) devices.

Programming:

In a similar way as the hardware requirements, the developed software should be scalable and be structured in a way that supports the deployment on various systems. Moreover, the software should facilitate the communication with other intelligent systems indirectly, utilizing the implemented Data Spaces as an intermediary actor. In respect to the programming requirements for an AI system, these will depend on the specific application and the type of AI system being developed. Overall, developing an effective AI system requires expertise in programming, machine learning (ML), and software engineering, which can be a complex task, but with the right skills and tools, AI systems can provide significant benefits in a wide range of applications. Under this context, a non-exhausting list of generic programming requirements for an AI system may include:

- Programming languages: AI systems can be developed using a variety of programming languages, including Python, Java, etc[6]; where the choice of language will depend on the specific requirements of the Project, since each language has its own strengths and weaknesses.
- Data processing and manipulation: Since AI systems require large amounts of data to train and learn from, efficient methods for data processing and manipulation are also required, which can be achieved through the use of libraries such as Pandas, NumPy, and SciPy.
- Machine learning libraries: Machine learning libraries such as TensorFlow, PyTorch, and Scikit-learn provide pre-built algorithms and functions that can be used to develop AI systems more efficiently.
- Deployment and integration: AI systems may need to be deployed and integrated with other systems and applications, which requires knowledge of software deployment and integration, as well as experience in developing APIs and microservices.

Avoidance of Data inconsistency: in practical terms, there are multiple data sources, depending on the system(s) in question, that fulfil some specific operations. The data that can be mined from each system will vary and the nature of the data is defined based on the operation of each system. In order to handle the inconsistencies between the systems, each AI model will be trained to fit a specific use case, thus the corresponding data. Data inconsistency is considered as another major challenge for AI systems, as it can lead to inaccurate predictions and poor performance. To address this challenge, some technical requirements for an AI system may include:

- Data pre-processing techniques: such as data cleaning, normalization, and feature scaling, which can help to reduce data inconsistencies and improve the accuracy of predictions.
- Data augmentation techniques: such as data synthesis and data sampling can be used to increase the amount of training data and improve the diversity of the data set, reducing the risk of overfitting and improving the accuracy of the model.
- Data quality assurance techniques: such as data profiling and data validation, can be used to identify and address data inconsistencies, reducing the risk of errors and improving the accuracy of the model.
- Data governance policies and procedures can be also put in place to ensure that data is consistent, accurate, and reliable, reducing the risk of errors and improving the performance of the AI system.

Sufficient data for training:

- Transfer learning: since such techniques, can be used to train an AI system on a related task or data set and then transfer the learned knowledge to the target task or data set, reducing the amount of training data required.
- Active learning: since such techniques can also be used to selectively choose which data points to label and use for training, reducing the amount of labelled data required.

Well-identified goals: the goals of the AI systems are to produce consistently accurate results, and for the human-machine collaborative systems. The systems, depending on the scope of their applications, must be able to accurately detect the parts in question, perform the disassembly operations, provide collaborative capabilities to human actors, and continuously stay accurate on their tasks by periodically training on newly acquired data samples.

- Evaluation metrics: Appropriate evaluation metrics (e.g., KPIs) should be established to measure the performance of the AI system against the intended goals. These metrics should be objective, relevant to the problem, and based on real-world performance.
- Domain expertise: The goals of an AI system should be defined in collaboration with domain experts who have a deep understanding of the problem space and the desired outcomes.
- Well-defined problem statement: A well-defined problem statement is essential to set clear goals for an AI system. This includes specifying the task, the target performance metric, and the intended user outcomes.
- User feedback: User feedback can be also used to refine and adjust the goals of the AI system over time. This feedback can be obtained through user testing, surveys, or other methods of user engagement.
- Managing exceptions in AI systems, typically refer to situations where the system fails to operate as expected. In order to overcome such failures, a non-exhausting list of technical requirements for managing exceptions in an AI system is reported.
- Robust error handling: The AI system should be designed to handle exceptions gracefully and provide appropriate error messages or feedback to users.

- Monitoring and logging: The system should be able to monitor its own behaviour and log any exceptions that occur, allowing developers to analyse and debug issues.
- Exception analysis: An AI system should be designed to analyse exceptions and identify the root causes of errors. This can involve techniques such as root cause analysis and fault tree analysis.
- Testing: Rigorous testing should be conducted to ensure that the system can handle a wide range of exceptions and edge cases. This can include techniques such as unit testing, integration testing, and end-to-end testing.

Managing limitations: while AI systems provide advanced capabilities to the deployed systems, some limitations apply. While the AI systems will be assessed for their accuracy and performance, it is expected that in some very rare cases, specifically on extremely rare circumstances where the system has not been exposed before, it is possible for a system to produce an inaccurate result. Constant training of the AI systems ensures a smooth operation, and explainability techniques as well as casual analysis can further assist with preventing and correcting accuracies.

Transparency: The AI system should be transparent, with clear documentation and explanations of how it works, what data it uses, and how it makes decisions. This can help users and developers understand the system's limitations.

Validation and verification: The system should be tested and validated to ensure that it performs as expected and that its limitations are well understood. This can involve techniques such as validation testing, sensitivity analysis, and stress testing.

Continuous monitoring and evaluation: The system should be continuously monitored and evaluated to detect any changes in its behaviour or performance. This can involve techniques such as drift detection, model retraining, and model updating.

Human oversight: AI systems should be designed with human oversight in mind, allowing humans to intervene when necessary and ensuring that the system's decisions are fair, transparent, and aligned with ethical and legal frameworks.

4. WP5 Outcomes / Identification of the technologies developed or adopted

4.1 Task 5.1

In this section we provide an update of the requirements for the Data4AI Platform.

Table 1: Roles in Data4AI Platform

Role	Description
Data Quality engineer	Owner of the data quality pipelines in an industry process. Responsible for ensuring that only data with desired (predefined) quality will be used in analyses
Process Owner	Business process owner Interested in having the industry facility properly operating all the time, with minimal running and maintenance costs

Table 2: Requirements for Data4AI Platform

ID	Description	Priority	Stakeholders	Integration	Validation
R1.1	Enabling to upload a dataset	H	Process owner	Completed	Completed Successfully
R1.2	Enabling to define a set of criteria	H	Process owner	Completed	Completed Successfully
R1.3	Enabling to upload the information about the errors/anomalies which happened in the past	H	Process owner	Completed	Completed Successfully
R1.4	Providing visual presentation of results	H	Process owner	Completed	Completed Successfully
R2.1	Enabling to select the data source	H	Data Quality engineer	Completed	Completed Successfully
R2.2	To define a set of criteria for data quality	M	Data Quality engineer	Completed	Completed Successfully
R2.3	Enabling to check / validate created pipeline	H	Data Quality engineer	Completed	Completed Successfully
R2.4	Provide automatic reporting about the data quality	H	Data Quality engineer	Completed	Completed Successfully
R2.5	Provide report about the data preprocessing results	H	Data Quality engineer	Completed	Completed Successfully
R2.6	Enable reconfiguration of the pipeline	M	Data Quality engineer	Completed	Completed Successfully

ID	Description	Priority	Stakeholders	Integration	Validation
R2.7	Enable to check / validate created pipeline in the circularity context	M	Data Quality engineer	Completed	Completed Successfully
R2.8	Enable automatic validation of the pipeline (based on defined KPIs)	M	Data Quality engineer	Completed	Completed Successfully

4.1.1 Description of the results (technologies developed/adopted)

In Figure 3 we present a high-level view on the Data quality pipeline.

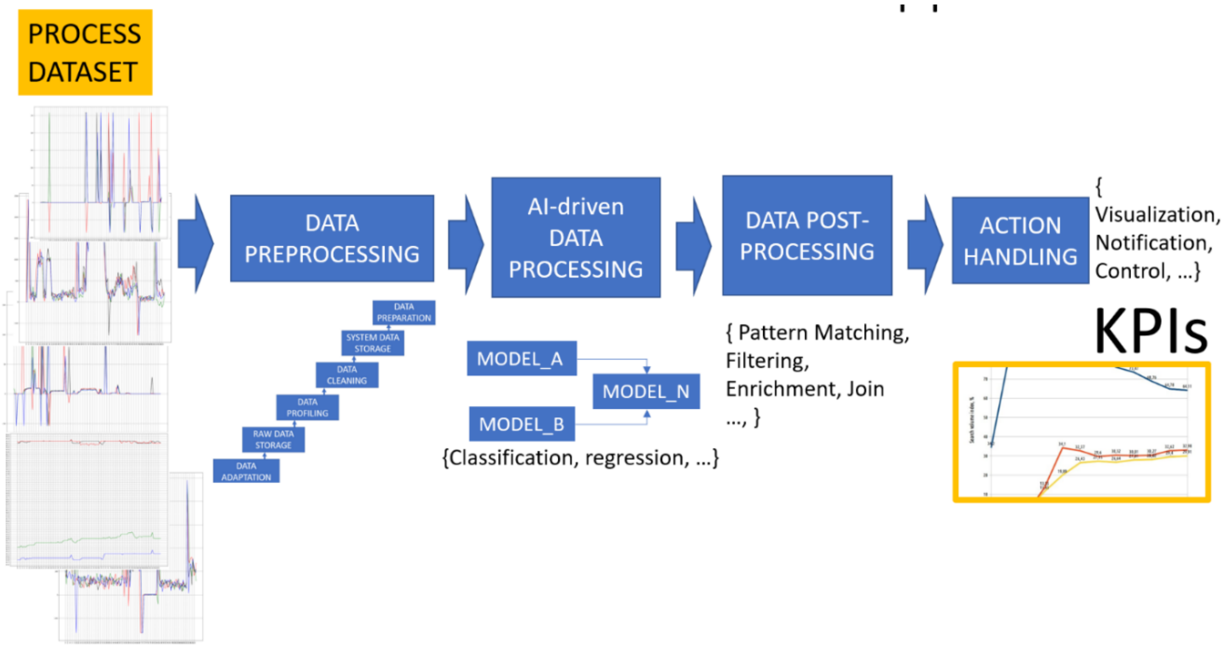

Figure 3: Data processing pipeline for AI applications

We describe the steps in the pipeline briefly:

- Data Preprocessing is a processing pipeline which transforms raw data in the well-formed data (valid structure) that can be processed by various data analysis methods
- AI-driven Data Processing is data analysis which can be done within or outside Data4AI Platform
- Data Postprocessing enables preparation of the data for output (e.g., filtering)
- Action Handling is related to the delivery of the output to other (control, notification, visualization) systems

As depicted in the Figure 4, Data Preprocessing ensures the Data Quality from the syntax point of view (it is in the valid form and can be processed automatically), whereas Data/Process behaviour understanding provides a semantic context for data quality, i.e., for interpreting data in a way suitable for creating AI applications. In Figure 4 we describe the Data Preprocessing pipeline.

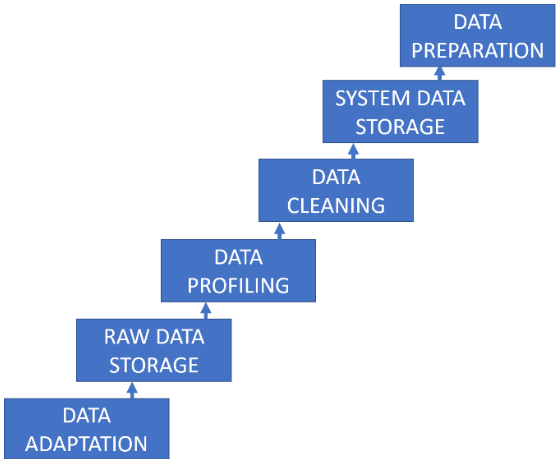

Figure 4: Data Preprocessing pipeline

We briefly describe the particular steps:

- Adapter – Reads the raw data from the files (properly prepared) and writes data into the raw data storage
- Raw data storage – Stores the raw data from the provided dataset into the previously defined format (d2twin, etc.)
- Profiling – Data Inspection (overview profiling of the raw data stored in the raw data storage)
- Data cleaning – Data cleaning according the info provided from data profiling (removing irrelevant data from the raw data)
- System data storage – Stores cleaned data after the profiling is done
- Data preparation – Getting data from the system data storage and preparing the data for the analytics algorithms

4.1.2 Mapping to the requirements provided in D2.2

Table 3 provides a mapping between the technical requirements presented in D2.2 and the current outputs of task T5.1, presenting who the former are fulfilled by the latter and presenting a justification for each case.

Table 3: Mapping to the requirements provided in D2.2

Requirement	Fulfilment by T5.1	Justification
AI systems require significant processing power to analyse and learn from large datasets.	Fulfilled by the infrastructure of the Cloud platform	The cloud infrastructure is able to scale and to provide the necessary resources like Graphics Processing Units (GPUs), Tensor Processing Units (TPUs).
AI systems require also large amounts of memory to store and process data efficiently.	Fulfilled by the infrastructure of the Cloud platform	The cloud infrastructure is able to scale horizontally to meet the requirement for increased memory availability.
Significant amounts of storage is required from AI systems to store large datasets and trained models.	Fulfilled by the infrastructure of the Cloud platform	The cloud infrastructure is able to scale horizontally to meet the requirement for increased data storage capacity.
High-power supplies are required from AI systems in order to support their processing requirements.	Fulfilled by the infrastructure of the Cloud platform	This is handled by the provider of the cloud-based platform infrastructure.
Developed software should be scalable and be structured in a way that supports the deployment on various systems	Fulfilled by the components developed in D5.2	Developed system is composed of different modules that can be easily extended and is based on containers for multi-system deployment support.
Software should facilitate the communication with other intelligent systems indirectly, utilizing the implemented Data Spaces as an intermediary actor.	Fulfilled by the Data Ingestion Services implemented	Data ingestion is supporting the collection of data from the Data Pipelines or data spaces directly.
Support for Data processing and manipulation.	Fulfilled by the Data Mapping Services	Data mapping allows to further manipulate data and make it interoperable.
Support for Multiple Machine learning libraries.	Fulfilled by the AI Model Repositories	Different models and libraries are available through the repositories.
Ease in Deployment and integration.	Fulfilled by the Cloud Base Platform and the Experimentation Composer	Cloud platform is offered as a Service, while the Experimentation Composer is easy to be deployed using docker containers.
Data pre-processing techniques.	Data4AI Platform service	N/A
Data augmentation techniques.	Data4AI Platform service	Specific blocks are provided to support data augmentation.
Data quality assurance.	Data4AI Platform service	N/A

Data governance policies and procedures.	Data4AI Platform service	N/A
Robust error handling.	Data4AI Platform service	Error handling is done in the backend and the appropriate messages are presented to the user.
Monitoring and logging.	Data4AI Platform service	Logs are kept in the Graylog.

4.2 Task 5.2

4.2.1 Requirements (internal for the task)

In task T5.2, the prototype of the AI for Circularity and Resilience toolkit has been implemented. The first version of the prototype includes the different parts of the AI Engine (as seen in the next sub-section) and a rich catalogue of pre-defined analytics that will be provided to the end-users in order to design, train and execute their AI pipelines.

The main requirements that led to the development of the modules and the integration of them into the first version of the AI for Circularity and Resilience toolkit prototype are listed in Table 4, in the form of users stories.

Table 4: Status of User Stories

User Story			**Integration Status**	**Validation Status**
As a	**I want to**	**so that**		
User	Be able to login into the AI Engine using my credentials	I can work in a protected environment	Completed	Completed Successfully with end users
User	be able to see the AI Pipelines of other users in my organisation	I can collaborate by using the artefacts created by my colleagues	Completed	Completed Successfully with end user
User	Select a static data source (e.g. file) for uploading my data	I can bring the data into the AI toolkit platform	Completed	Completed Successfully with end user
User	Select a dynamic data source (e.g. API, Kafka stream) for uploading my data	I can bring updated as well as real-time data into the AI toolkit platform and	Completed	Completed Successfully with end user
User	Be able to map my data to a specific industry data model	I provide context to the data to help the people that will work with them	Completed	Completed Successfully with end user
User	Be able to design my AI pipeline by selecting from a	I can build a complex pipeline without writing code, and by making	Completed	Completed Successfully

	list of ready-made AI/ML algorithms	sure no errors are made during the design process		with QA engineers
User	Train my AI model using part of the data I have uploaded on the platform	I can be ready to apply the model to the actual dataset that will be used in my analysis	Completed	Completed Successfully with QA engineers
User	Execute the trained AI model on the dataset	I can get the results of my analysis	Completed	Completed Successfully with QA engineers
User	Visualise the results of my analysis using different charts	I can have a visual representation of the outcomes of the AI pipeline	Completed	Completed Successfully with QA engineers
User	Export the results of my AI pipeline using an API	I can serve this output to other components	Completed	Completed Successfully with QA engineers
User	Deploy a local instance of the Experimentation Facility of the AI Toolkit	I can work with local resources for on my AI problem	Completed	Completed Successfully with QA engineers
User	Write my own analytics code / import own models in the Experimentation Facility of the AI Toolkit	I can create specific local experiments and run AI models that need to be executed locally and not in the cloud due to low latency requirements (e.g. image analysis)	Completed	Completed Successfully with QA engineers
User	Be able to deliver an AI pipeline designed on the Cloud platform to the Experimentation facility	I can run the AI pipeline locally, in case the overall pipeline has low latency requirements (imposed by data transfer limitations)	Pending for 2nd toolkit version	
User	Apply XAI methods on the models I have executed	I can get insights on the behaviour of certain models	Pending for 2nd toolkit version	

4.2.2 Description of the results (technologies developed/adopted)

As identified in deliverable D5.1, the overall AI Toolkit is composed of two core bundles which are shown in the next architecture figure (*Figure 5*).

Those are the following:

- The "XAI Pipeline Designer and Cloud Execution service", which is a cloud-based high fidelity AI pipeline designer and AI execution engine that will be used for retrieving data from the data pipelines, designing the AI pipelines, execution the AI models, and delivering the results to other entities (such as the Digital Twins), providing also explainability interfaces as required.
- The "Experimentation and Local Models Composer Interface" that will be used for drafting AI experiments that can be executed locally, and for executing those AI models that need to be deployed at the local/edge of each stakeholders, due to reasons of latency imposed by data transfers (in case of images or videos that need to be analysed and that would introduce a very big delay during data transfer to the cloud service), or due to data confidentiality reasons (in case no data is permitted to be transferred and analysed on the cloud environment).

Both of those components are fed with data (southbound) coming from the Data Spaces and/or directly from the Data Pipelines which are responsible for treating the data and make them usable for the AI to be executed.

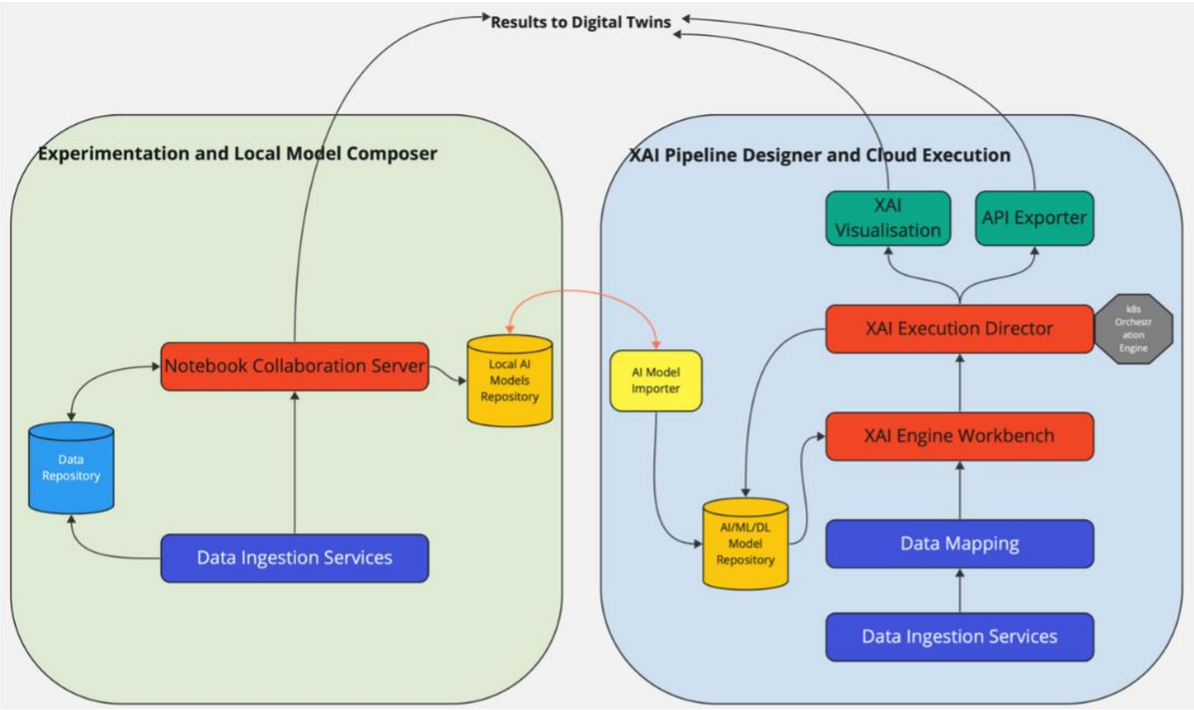

Figure 5: AI Toolkit Architecture

The results of both components are offered to the DT via APIs, either internal ones (in case the execution is placed inside the DT), or external ones, that communicate with the DT as the execution is taking place in an external to the DT's virtual environment.

During the previous period, development and integration work on the different modules as listed above has been performed, in order to be in a position to serve the end-users with the appropriate tools for satisfying their AI needs.

When it comes to the XAI Pipeline Designer and Cloud Execution services, the following modules have been developed and integrated.

The XAI Workbench, responsible for delivering a configurable framework that is used for the design of the different AI pipelines. *Figure 6* shows a screenshot of the XAI workbench interface with a complex model developed.

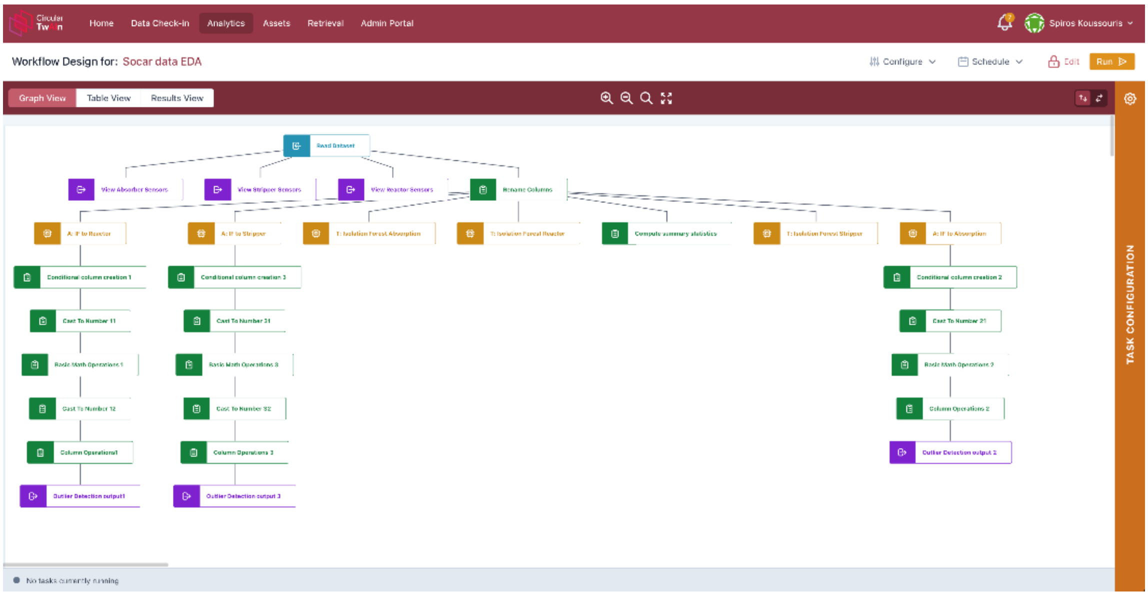

Figure 6: XAI Workbench Screenshot

The Workbench offers different data input blocks, data preparation blocks, data control blocks, AI/ML blocks, AI Evaluation blocks, ML Train blocks as well as Output blocks for saving and visualising of results.

The technologies used for the development of this framework include Node.JS for the backend and Vue.JS for the frontend.

The AI/ML libraries that are supported include those available in Spark, Scikit, TensorFlow and Python, (https://spark.apache.org, https://www.tensorflow.org, https://scikit-learn.org).

The AI/ML/DL Model Repository which is used as a storage facility for all the AI models and pipelines that are used in the XAI Workbench. The overall model repository is operating as an object storage based on Min.io (https://min.io/).

The XAI Execution Director is the component responsible for executing the AI pipelines and generating the corresponding results. This component has been developed by using Kubernetes as an execution engine that is ingesting the different blocks coming from the XAI workbench and is placing them in the correct order to run the overall pipeline functions (https://kubernetes.io/).

The XAI Visualisation component is developed using Vue.JS and is offering many different libraries for visualising the results of the output of each AI pipeline, allowing the user to configure the graphs to their liking. This component will be enriched as to provide explainability methods/interfaces in the second release of the prototype.

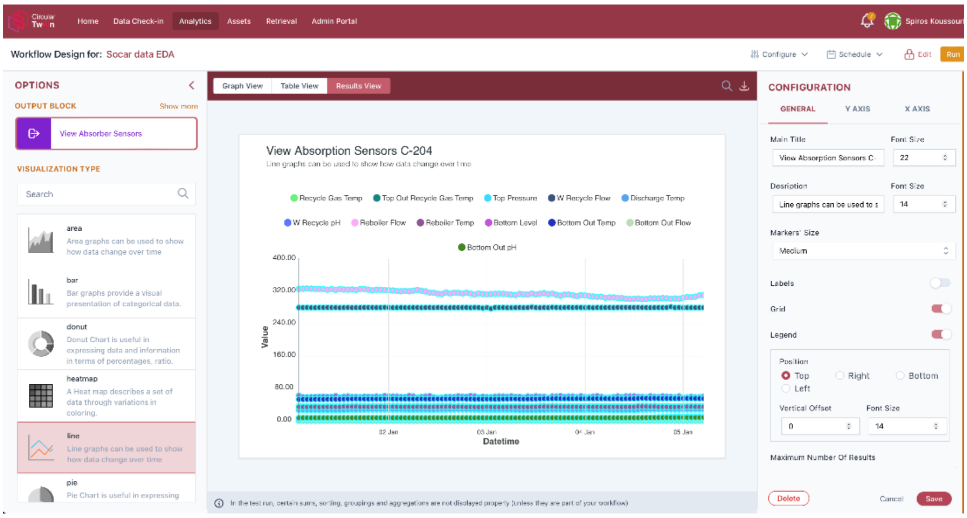

Figure 7: XAI Visualisation Dashboard Screenshot

The API Exporter enables the export of the AI/ML outputs through APIs, that need to be called by external entities. The user is able to construct such an API identifying the information that needs to be exposed, and is setting the necessary parameters to allow external entities to call them. This module is a custom Node.JS implementation, using VueJS for the frontend.

The Notebooks Collaboration Server that is part of the Experimentation and Local Models Notebook Interface is a deployment that is making used of a customised version (specifically customised for Circular TwAIn) of the CoCalc notebook server (https://cocalc.com).

4.2.3 Mapping to the requirements provided in D2.2, Section 5 (these are general requirements)

The following table (Table 5) provides a mapping between the technical requirements presented in D2.2 and the current outputs of task T5.2, presenting who the former are fulfilled by the latter and presenting a justification for each case.

Table 5: Mapping to the requirements provided in D2.2

Requirement	Fulfilment by T5.2	Justification
hardware platforms to be able to handle the processing requirements	Fulfilled by the infrastructure of the Cloud platform	The Cloud platform infrastructure is able to scale as necessary
the hardware platform should be able to re-train the deployed models	Fulfilled by the Cloud platform	The Cloud platform is able to train and re-train the AI models
AI systems require significant processing power to analyse and learn from large datasets	Fulfilled by the infrastructure of the Cloud platform	The cloud infrastructure is able to scale and to provide the necessary resources like Graphics Processing Units (GPUs), Tensor Processing Units (TPUs), and Field Programmable Gate Arrays (FPGAs)
AI systems require also large amounts of memory	Fulfilled by the infrastructure of the Cloud platform	The cloud infrastructure is able to scale horizontally to meet the

to store and process data efficiently		requirement for increased memory availability
Significant amounts of storage is required from AI systems to store large datasets and trained models	Fulfilled by the infrastructure of the Cloud platform	The cloud infrastructure is able to scale horizontally to meet the requirement for increased data storage capacity
high-power supplies are required from AI systems in order to support their processing requirements	Fulfilled by the infrastructure of the Cloud platform	This is handled by the provider of the cloud-based platform infrastructure
developed software should be scalable and be structured in a way that supports the deployment on various systems	Fulfilled by the components developed in D5.2	Developed system is composed of different modules that can be easily extended and is based on containers for multi-system deployment support
software should facilitate the communication with other intelligent systems indirectly, utilizing the implemented Data Spaces as an intermediary actor	Fulfilled by the Data Ingestion Services implemented	Data ingestion is supporting the collection of data from the Data Pipelines or data spaces directly
Support for AI different programming languages	Fulfilled by the XAI Workbench and the Notebooks Server	Different languages and frameworks are supported, while ready-made code blocks are provided
Support for Data processing and manipulation	Fulfilled by the Data Mapping Services	Data mapping allows to further manipulate data and make it interoperable
Support for Multiple Machine learning libraries	Fulfilled by the AI Model Repositories	Different models and libraries are available through the repositories
Ease in Deployment and integration	Fulfilled by the Cloud Base Platform and the Experimentation Composer	Cloud platform is offered as a Service, while the Experimentation Composer is easy to be deployed using docker containers
Data pre-processing techniques	Fulfilled by Task 5.1 services	N/A
Data augmentation techniques	Fulfilled by the XAI Workbench modules	Specific blocks are provided to support data augmentation
Data quality assurance	Fulfilled by Task 5.1 services	N/A
Data governance policies and procedures	Fulfilled by WP4 services	N/A
Support Transfer learning	Fulfilled by the AI Model Repositories	Already trained models can be used to solve similar problems

Support for Active learning	Fulfilled by the XAI Workbench modules	The different modules allow to choose for the exact data to be used as training
Evaluation metrics	Fulfilled by the XAI Visualisation component	Visualisations provide the ability to measure the performance of the AI output compared to real-world results
Robust error handling	Fulfilled by the XAI Execution Director	Error handling is done in the backend and the appropriate messages are presented to the user
Monitoring and logging	Fulfilled by the XAI Execution Director	Logs are kept in the backend
Exception analysis	Fulfilled by the XAI Execution Director	Analysis of exceptions is done utilising the logs of the backend
Transparency	Fulfilled by the overall system	Clear documentation (online) of the system is available
Human oversight	Fulfilled by the XAI Workbench	The human operator is able at any time to intervene to the execution of an AI pipeline

4.3 Task 5.3

4.3.1 Requirements (internal for the task)

The main requirements that led to the development of the modules and the integration of them into the first version of the Product DT infrastructure solution are listed in the following Table 6.

Table 6: Mapping to the requirements provided in D2.2

Req #	Requirement	Product DT validation Description	Integration Status	Validation Status
PMD T-01	Data Acquisition	Conduct testing to validate the efficiency and accuracy of data acquisition and exchange protocols. Ensure that the Product DT communicates seamlessly with other entities.	Completed	Completed
PMD T-02	Data Storing	Verify the robustness of the data storage system. Conduct tests to ensure proper data management, format representation, and security measures.	Completed	Completed
PMD T-03	Computation & Analytics	Evaluate the computational models and analytics employed by the Product DT. Validate their effectiveness in real-time and batch-oriented processing.	Pending	Pending

PMD T-04	Machine learning & Data Processing	Test the integration of machine learning solutions. Validate their impact on decision-aiding capabilities and overall enhancement of product intelligence.	Completed	Ongoing
PMD T-05	Interaction with other Materials DT	Test interfaces for interactions with Materials DT. Ensure seamless communication and data exchange.	Completed	Ongoing
PMD T-06	Interaction with Human DT	Test interfaces for interactions with Human DT. Ensure seamless communication and data exchange.	Pending	Pending
PMD T-07	Interaction with Process DT	Test interfaces for interactions with Process DT. Ensure seamless communication and data exchange.	Completed	Ongoing
PMD T-08	User Interface	Conduct usability testing for the developed user interface. Ensure that it provides effective monitoring capabilities and is user-friendly.	Pending	Pending
PMD T-09	Interaction with other Products	Simulate interactions between smart products. Validate the communication mechanism for interconnection and collaboration toward common objectives.	Completed	Pending
PMD T-10	Define a combined set of procedural rules	There must be a mechanism in place that allows the smart products to define a combined set of procedural rules.	Pending	Pending
PMD T-11	Integrate their individual contributions.	There must be a mechanism in place that allows the smart products to integrate their individual contributions, so the placing (where)/timing(when) of the next action/activity that needs to be done/executed can be determined properly.	Pending	Pending
PMD T-12	Visualize the impact of the individual decision-making processes.	There must be a mechanism in place that allows the smart products to visualize the impact of the individual decision-making processes, therefore, the next action/activity that needs to be done/executed can be determined properly.	Pending	Pending

4.3.2 Description of the results (technologies developed/adopted)

Integration/Testing Environment

In the current deployment of the Circular TwAIn testbed environment, which was introduced in D4.2 section 4.7.3, several tools and infrastructure components are seamlessly integrated to facilitate AI-enabled Digital Twins (DTs) for product and material integration with other Circular TwAIn components, such as Process DT, Product & Materials DS, XAI, and the industrial data platform. The environment is designed to support the efficient deployment, management, and monitoring of the Circular TwAIn components. The following elements contribute to the functionality of the environment.

Testbed Infrastructure:

The Circular TwAIn product & material testbed operates on Hetzner Cloud[3], leveraging the advantages of scalability, flexibility, and streamlined management. One of the VMs serve as the primary deployment and management host, responsible for orchestrating the deployment of dataspace components and supporting services. The other VMs are utilized to simulate and validate various aspects of the DS and DT environment, including different types of participants. In more details we have allocated (see Figure 8 below) the following publicly accessible cloud infrastructure @Hetzner:

- 3 x CPX41 (4vCPU-16GB RAM 160GB SSD): Hosting DS, DT, XAI, and industrial data platform services, emulating 3 participants/entities (data producers/consumers).
- 1 x CPX31 (4vCPU-8GB RAM 160GB SSD): Dedicated to CI/CD infrastructure and management services, including Jenkins, Keycloak, Portainer, and Harbor.
- 1 x CX11 (1vCPU, 2GB RAM, 20 GB SSD): Reserved for VPN/Firewall setup, facilitating secure partner access to the environment via VPN.

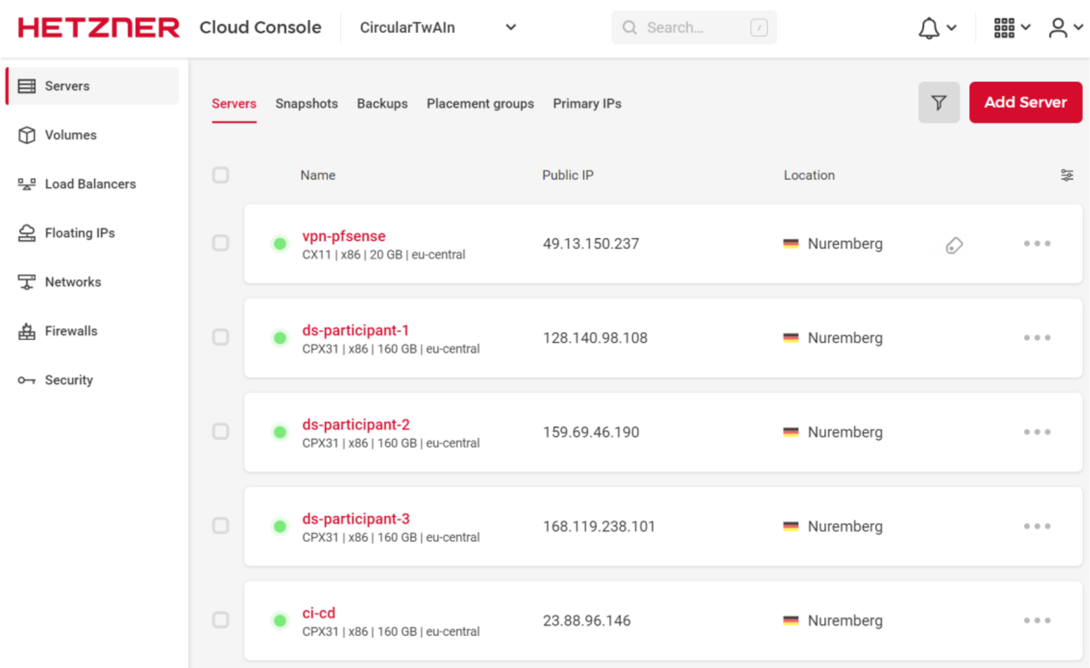

Figure 8: Circular TwAIn product & material testbed VMs

[3] https://www.hetzner.com/cloud

Deployment Automation:

Jenkins[4], an open-source automation server, is integral to automated infrastructure and component deployment. The integration with GitHub ensures real-time testing and deployment of updates to Circular TwAIn components (see Figure 9 below). The deployment process includes:

- Downloading the latest versions of DT/DS components.
- Building components as required.
- Deploying them to the testbed environment using new configurations.

Jenkins can be accessed through the following URL: https://jenkins.circulartwain.rid-intrasoft.eu/.

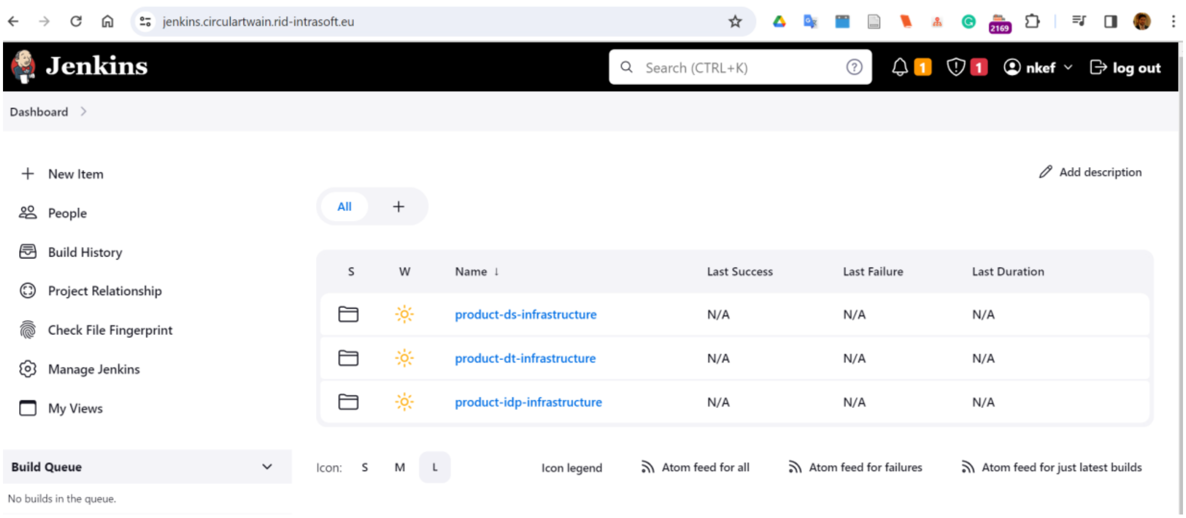

Figure 9: Jenkins dashboard – available pipelines

For artifact persistence, Jenkins uses Harbor[5] artifactory, an open-source registry ensuring compliance, performance, and interoperability across cloud-native platforms. Harbor facilitates in the deployment of artifacts across the different Docker environments that are available in the participants VMs. In Figure 10 we can find the Harbor consol listing the product & material projects. Harbor can be accessed through the following URL: https://harbor.circulartwain.rid-intrasoft.eu/.

[4] https://www.jenkins.io/

[5] https://goharbor.io/

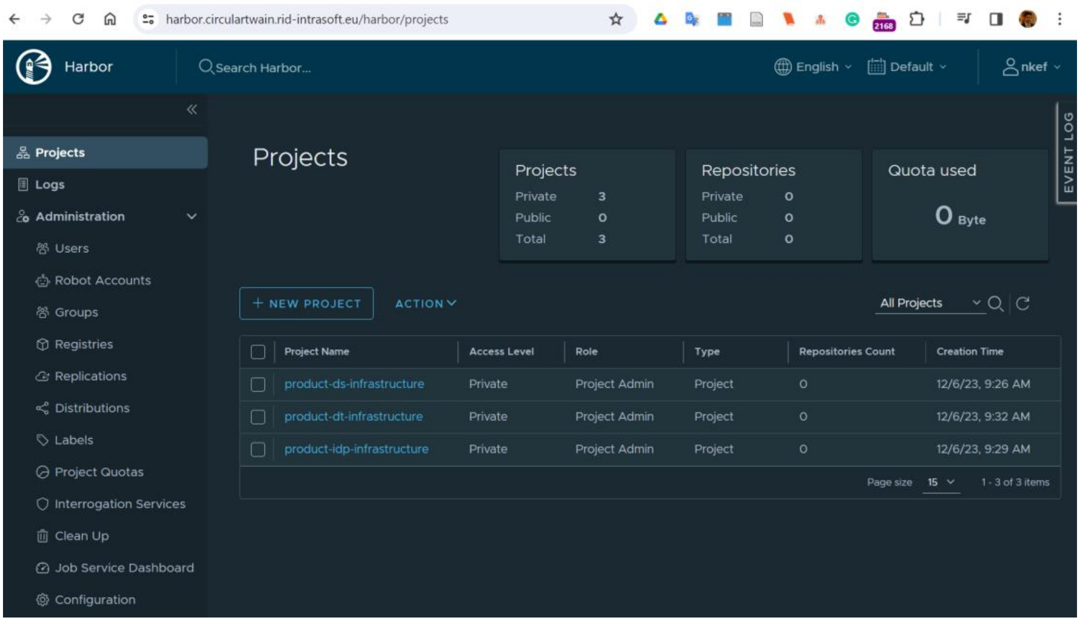

Figure 10: Harbor product projects list

Access Control:

For secure access to infrastructure and platforms lacking native authentication, a Single Sign-On (SSO) identity and access management solution, implemented by Keycloak[6], is deployed. It offers centralized management, support for standard protocols, and compatibility with LDAP and Active Directory. Keycloak is containerized using Docker, aligning with the overarching deployment strategy. Keycloak have been also integrated with infrastructure tools like Jenkins, Harbor, and Portainer (see Figure 11) to ensure that users can access these tools using single sign-on mechanism thru their GitHub account. Keycloak can be accessed through the following URL: [https://keycloak.circulartwain.rid-intrasoft.eu/.](https://keycloak.circulartwain.rid-intrasoft.eu/)

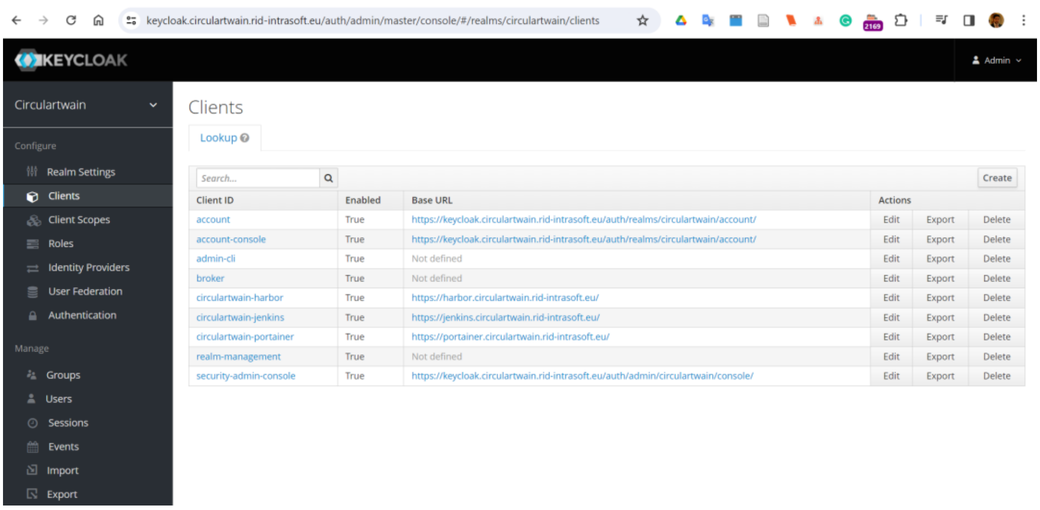

Figure 11 Keycloak management dashboard and client integration list.

[6] https://www.keycloak.org/

Enhancing Security and Facilitating Partner Access:

In the Circular TwAIn Project, PFsense is utilized for securing access to the environment (see Figure 12). Through the integration of OpenVPN, secure connections are established, ensuring a robust and protected infrastructure. The key components of this setup are listed below:

- PFsense: PFsense acts as the primary firewall and gateway, managing access to the Circular TwAIn environment.
- OpenVPN Integration: OpenVPN, integrated into PFsense, provides a secure and encrypted tunnel for partners to access the environment remotely.
- Hetzner's Firewall: The infrastructure is fortified by Hetzner's firewall, restricting access to trusted IP addresses. This feature ensures that only specified IPs, including dynamic ones, are allowed to interact with the integration environment.

One significant advantage is the compatibility with dynamic IP addresses, enabling partners to securely connect to the environment even with frequently changing IPs. The selective access control enforced by Hetzner's firewall ensures that only trusted IP addresses are permitted, minimizing the risk of unauthorized access. The encryption provided by OpenVPN guarantees that all communications between partners and the Circular TwAIn environment remain confidential, safeguarding sensitive data. In essence, this integrated security infrastructure not only ensures accessibility but also prioritizes the integrity and confidentiality of data exchanged within the Circular TwAIn environment.

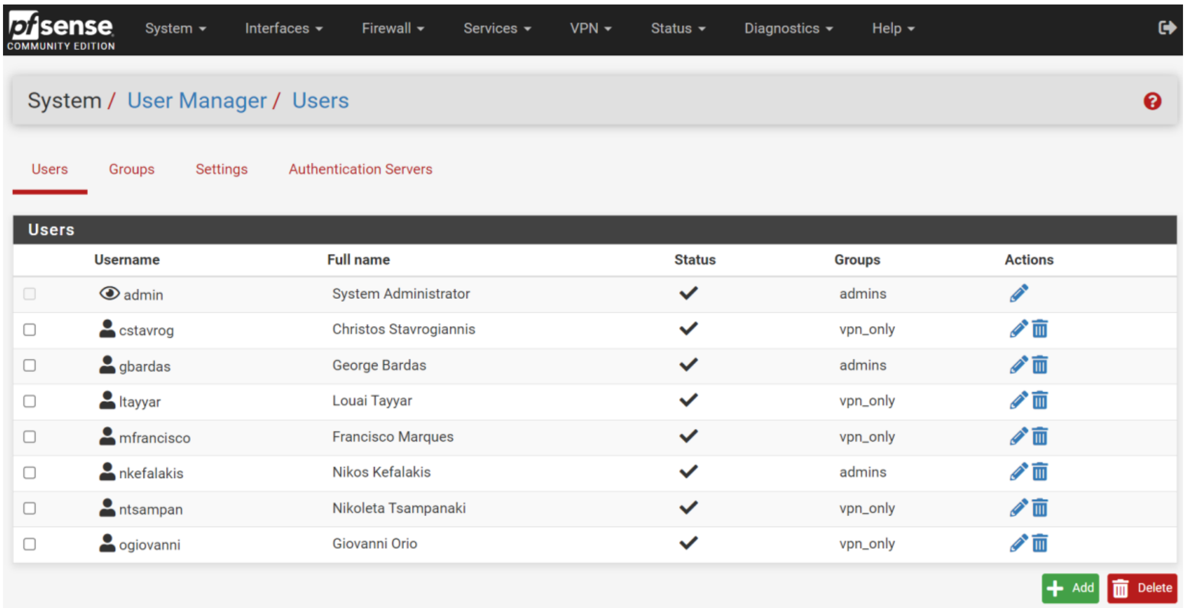

Figure 12: PFsense management console – available user list

Infrastructure Management & Monitoring:

The chosen deployment strategy (see deliverable D4.2 section 4.7) of Docker containerization is managed through the Community Edition (CE) of Portainer[7]. Portainer simplifies Docker ecosystem management with features such as:

[7] https://www.portainer.io/products-services/portainer-community-edition/

- A comprehensive UI encapsulating Docker CLI functionalities.
- Expert configurations and pre-validation checks.
- Access control management and LDAP authentication support.
- Detailed log viewer and remote process performance viewer.

In Figure 13 we can find the Portainer management console with all the managed environments and in Figure 14 we can find the list of containers deployed to the CI-CD VM. Portainer can be accessed through the following URL: https://portainer.circulartwain.rid-intrasoft.eu/.

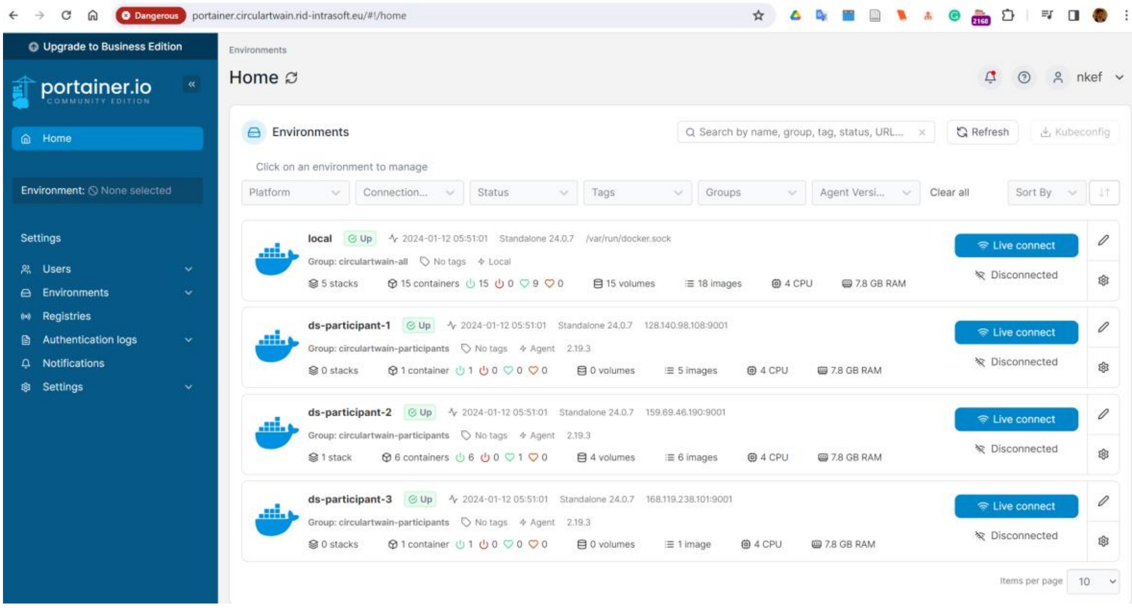

Figure 13 Portainer management console – available environments.

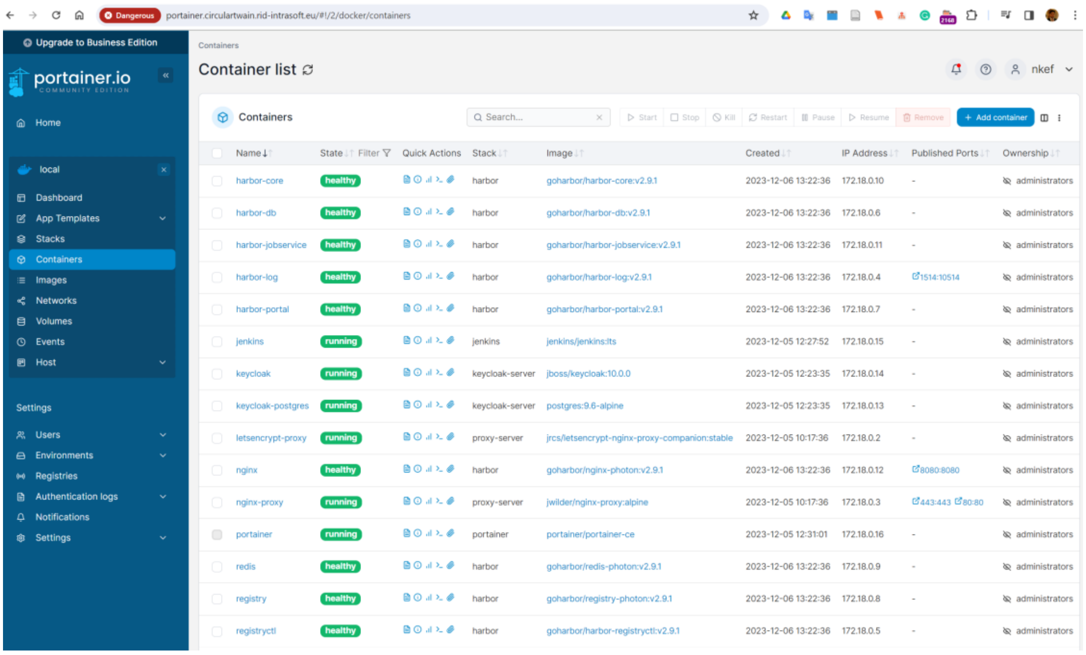

Figure 14 Portainer management console – deployed containers in ci-cd VM

Integration Architecture:

In Figure 15 we can find the Product & Materials integration architecture. The integration architecture is consisted of three main groups:

- DT infrastructure
- Industrial Data Platform which includes the
  - AI for Circularity Toolkit and
  - Data Preparation
- DS infrastructure

As shown in Figure 15 the architecture groups have been colour-coded and are described in detail in D4.2 and D5.2 deliverables. In Figure 15 we can also find the implementation technologies that are currently used for the solution which are:

- For DT infrastructure the following frameworks are used:
  - For Product/Material DT we are using Eclipse BaSyx[8]. More info can be found at Circular TwAIn "product-dt-infrastructure" GitHub repository[9]
  - For Process DT we are using NOVAAS which is integrated with product DT thru an MQTT Broker (EMQX)
- For the Industrial Data Platform the following technologies are used:
  - For AI for Circularity Toolkit:
    - S5 AI Toolkit utilizing its exposed API
    - TensorFlow[10] utilizing the Keras API. More info can be found at Circular TwAIn "product-idp-infrastructure" GitHub repository[11]
  - For Data Preparation, where more info can be found at Circular TwAIn "product-idp-infrastructure" GitHub repository[12]:
    - EMQX[13] for MQTT Broker
    - Apache Kafka[14] for managing data streaming
    - Elastic Logstash[15] for data routing and data preprocessing
    - Elasticearch[16] for data persistence
    - Elastic Kibana[17] for data visualizations from Elasticsearch

[8] https://eclipse.dev/basyx/
[9] https://github.com/Circular-TwAIn/product-dt-infrastructure
[10] https://www.tensorflow.org/
[11] https://github.com/Circular-TwAIn/product-idp-infrastructure
[12] https://github.com/Circular-TwAIn/product-idp-infrastructure
[13] https://www.emqx.io/
[14] https://kafka.apache.org/
[15] https://www.elastic.co/logstash
[16] https://www.elastic.co/elasticsearch
[17] https://www.elastic.co/kibana

- For the DS infrastructure we are using Eclipse EDC[18] more info can be found at Circular TwAIn "product-ds-infrastructure" GitHub repository[19].

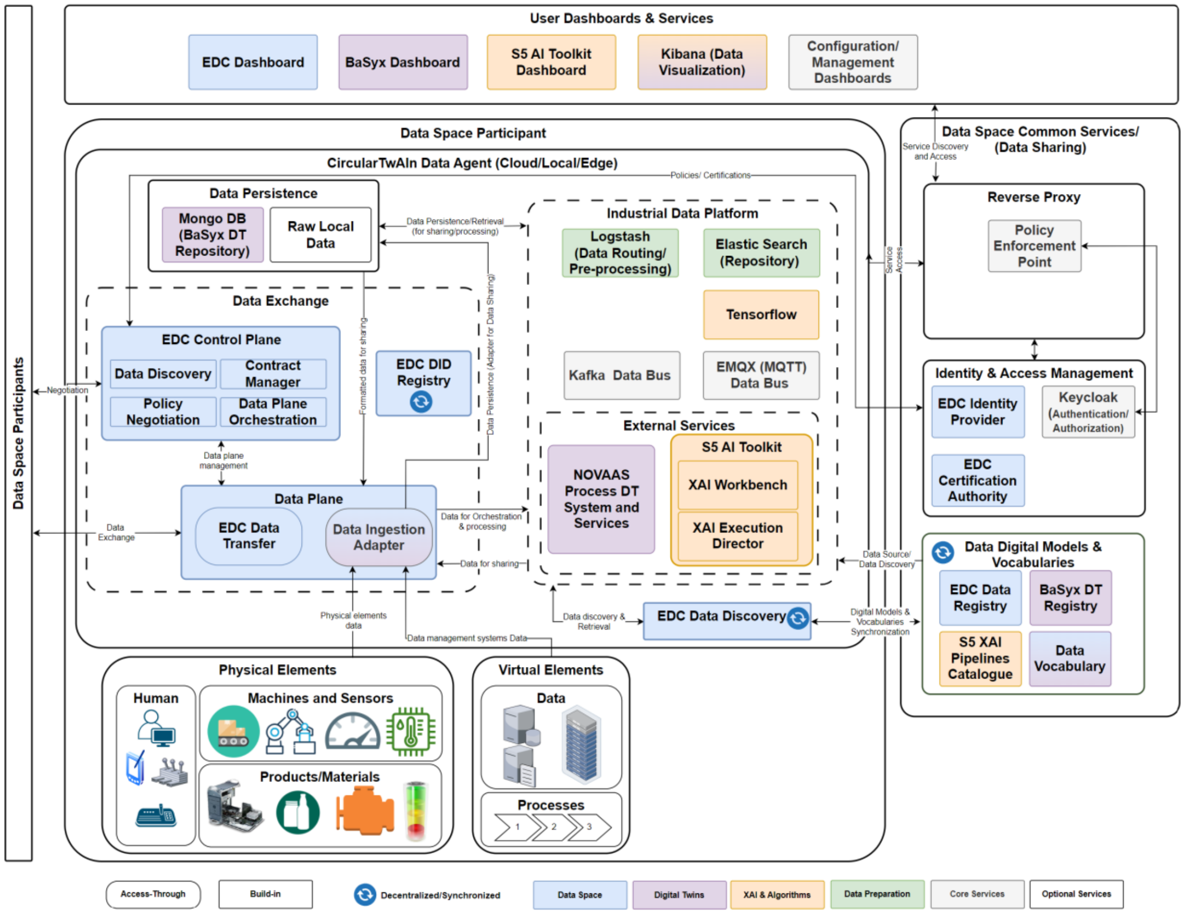

Figure 15: Product DT/DS integration architecture

Product & Material deployment and integration

For the Product & Material DT deployment and integration as we mentioned we are using Eclipse BaSyx. The Eclipse BaSyx flavour that is going to be used along with the installation instruction are available at the Circular TwAIn "product-dt-infrastructure" GitHub repository[20] also shown in Figure 16, where we can see the Jenkins pipeline executed for deploying the product DT infrastructure to DS participant 2 VM. The containers deployed in DS participant 2 VM are shown in Figure 18. There we can also see the EMQX MQTT broker deployed which is used for exchanging messages with NOVA's remote Process DT deployment (NOVAAS).

[18] https://eclipse-edc.github.io/docs/#/README
[19] https://github.com/Circular-TwAIn/product-ds-infrastructure
[20] https://github.com/Circular-TwAIn/product-dt-infrastructure

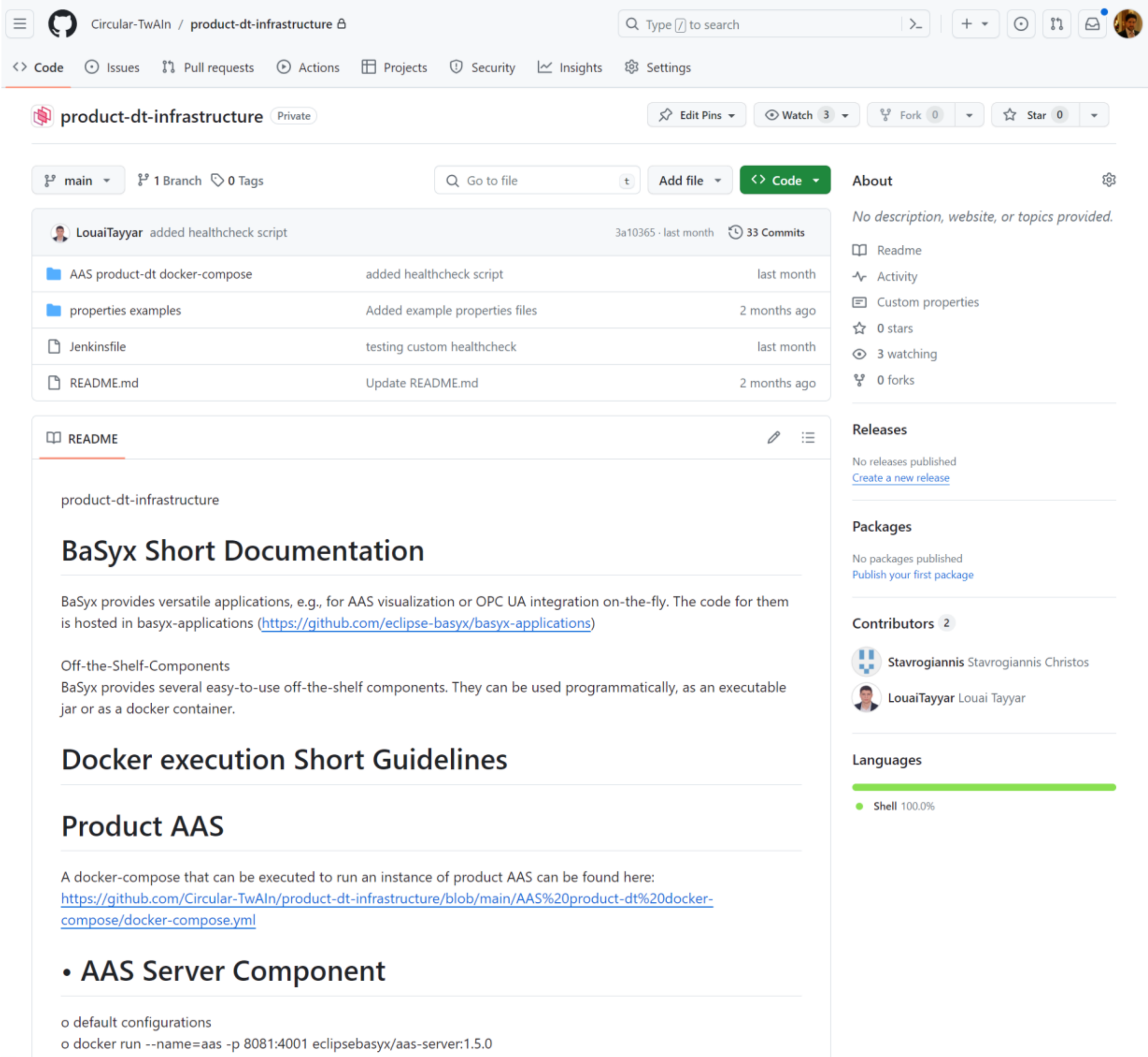

Figure 16: GitHub product-dt-infrastructure repository

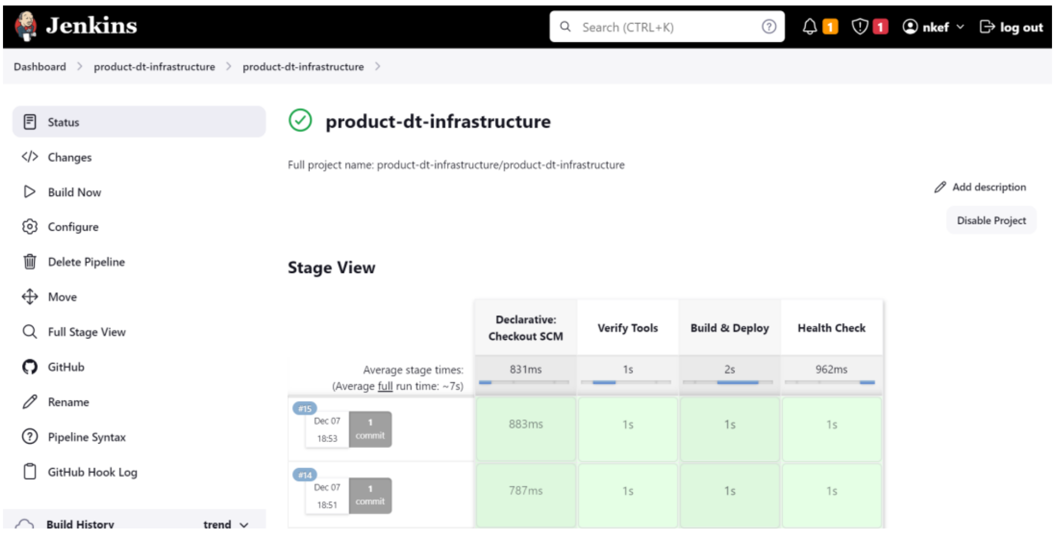

Figure 17: Jenkins dashboard showing the product-dt-infrastructure pipeline

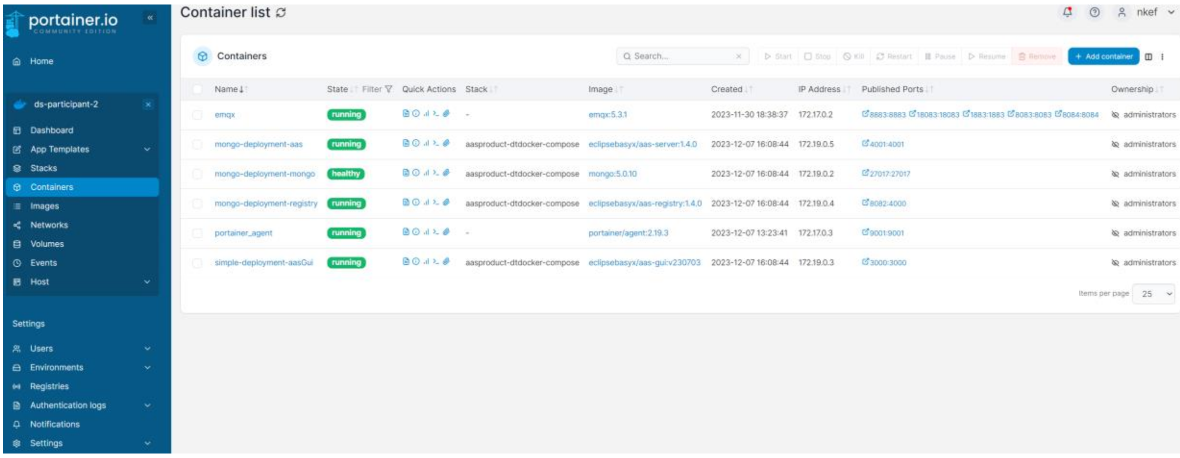

Figure 18: Portainer management console – deployed containers in “ds-participant-2” VM

4.3.3 Mapping to the requirements provided in D2.2, Section 5

Requirements from a Circular system:

- **Technical**:
    - Information structured under a **common ontology and semantics**: structure the information with under appropriate ontology and semantics is fundamental to adopt a clear terminology and information sharing, foster the collaboration between the stakeholders and implement effective and scalable circular economy practices.
    - **Material traceability**: to close the loop of the materials and create circular value chains it is fundamental to track the material flow to avoid losses along the chain, to control the material flow and create closed-loop material flows, and to be able to certify the origin of the materials.
    - Products and the waste along the entire value chain. By tracing the material flow the companies can identify the opportunities of waste reduction, improve the usage of the resources and create new circular business models.
    - **Information and material sharing**: among the different actors of the value chains: to create closed loop material flow, it is mandatory to guarantee a way to exchange both info and material flow between the different actors of the value chains.
- **Business**:
    - **Close-loop material flow**: close the loop of the material is fundamental to keep it inside the value chain to reduce the usage of virgin materials, the environmental impact, increase maximise the use of resources and reduce the waste.
    - **Waste reduction**: it is fundamental to reduce the environmental impact and it can be achieved by the increase of material utilization, reduction of production scraps, reduction of material consumption, reuse and recovery of post use material.

Table 7: Mapping to the requirements provided in D2.2

Requirement	Product DT Mapping	Fulfilment by T5.3	Justification
Common ontology and semantics	The Product DT employs a standardized ontology and semantics to structure information consistently. This ensures a shared understanding of terminology and facilitates seamless information exchange among stakeholders.	Use of a common Product DT registry and Data Vocabulary components. Use of standardized AAS Sub models.	These components provide standardized models and semantic knowledge for modelling and exchanging product and material data. It enables the consistent representation and interpretation of information, supporting circularity assessments and interoperability. Moreover, utilizing AAS enables the structuring of information under a common ontology and semantics. AAS provides a standardized way to represent and model information, fostering clear terminology and facilitating information sharing among stakeholders.
Material traceability	The Product DT incorporates features for material traceability, enabling the tracking of materials throughout the product lifecycle. This is essential for creating closed-loop material flows and certifying material origins.	Use of Digital Nameplate AAS Sub model.	AAS enables the inclusion of detailed information about the material composition of the product, its origin, and its journey throughout the lifecycle. This supports material traceability, helping to create closed-loop material flows and certify the origin of materials. The Digital Nameplate AAS sub model includes essential information for uniquely identifying the product. The sub model contains information related to the product's lifecycle, including its origin, manufacturing details, and relevant historical data. This supports traceability, allowing stakeholders to track the product from its creation to its current state.
Info traceability	The Product DT facilitates information traceability, allowing companies to track material flow,	Use of Digital Nameplate AAS Sub model.	AAS captures information related to the material flow, product lifecycle, and waste along the entire value chain.

	products, and waste throughout the entire value chain.		This information traceability aids in identifying opportunities for waste reduction, optimizing resource usage, and fostering the creation of new circular business models. The Digital Nameplate AAS sub model includes essential information for uniquely identifying the product. The sub model contains information related to the product's lifecycle, including its origin, manufacturing details, and relevant historical data. This supports traceability, allowing stakeholders to track the product from its creation to its current state.
Information and material sharing	The Product DT establishes a framework for the seamless exchange of both information and material flow among different actors in the value chain. This promotes collaboration and supports the creation of closed-loop material flows.	Use of a common Product DT registry and Data Vocabulary components. Use of AAS implemented by Eclipse BaSyx	AAS provides a standardized framework for information representation, making it feasible to exchange both information and material flow seamlessly among different actors in the value chain. This promotes effective collaboration and supports closed-loop material flows.
Close-loop material flow	The Product DT contributes to closing the loop of material flows within the value chain, reducing reliance on virgin materials. This helps in maximizing resource use, minimizing environmental impact, and promoting sustainable practices.	Use of AAS implemented by Eclipse BaSyx.	AAS, coupled with Eclipse BaSyx, supports the concept of a closed-loop material flow. By structuring the information in a way that facilitates circular practices, the system aims to keep materials within the value chain, reducing reliance on virgin materials and minimizing environmental impact.
Waste reduction	The Product DT supports waste reduction initiatives by enhancing material utilization, minimizing production scraps, reducing material consumption, and enabling the reuse and		Through comprehensive material and information traceability, the solution helps identify opportunities for waste reduction. It promotes the efficient utilization of materials, reduces production scraps, minimizes material consumption,

	recovery of post-use materials.		and encourages the reuse and recovery of post-use materials.

The integration of AAS and Eclipse BaSyx in the Product DT solution aligns with the technical and business requirements of a circular economy system. It promotes transparency, traceability, and collaboration, contributing to the effective implementation of circular practices and the reduction of environmental impact.

4.4 Task 5.4

4.4.1 Requirements (internal for the task)

In task 5.4, a prototype of an AI-enabled process DT has been implemented. The process DT can be seen as a higher-level monitoring and control system aimed at orchestrating an Industry 4.0 (I4.0) compliant production system, i.e., a system where real world components (such as sensors, machines or more in general physical assets) are digitally represented by Asset Administration Shells (AASs).

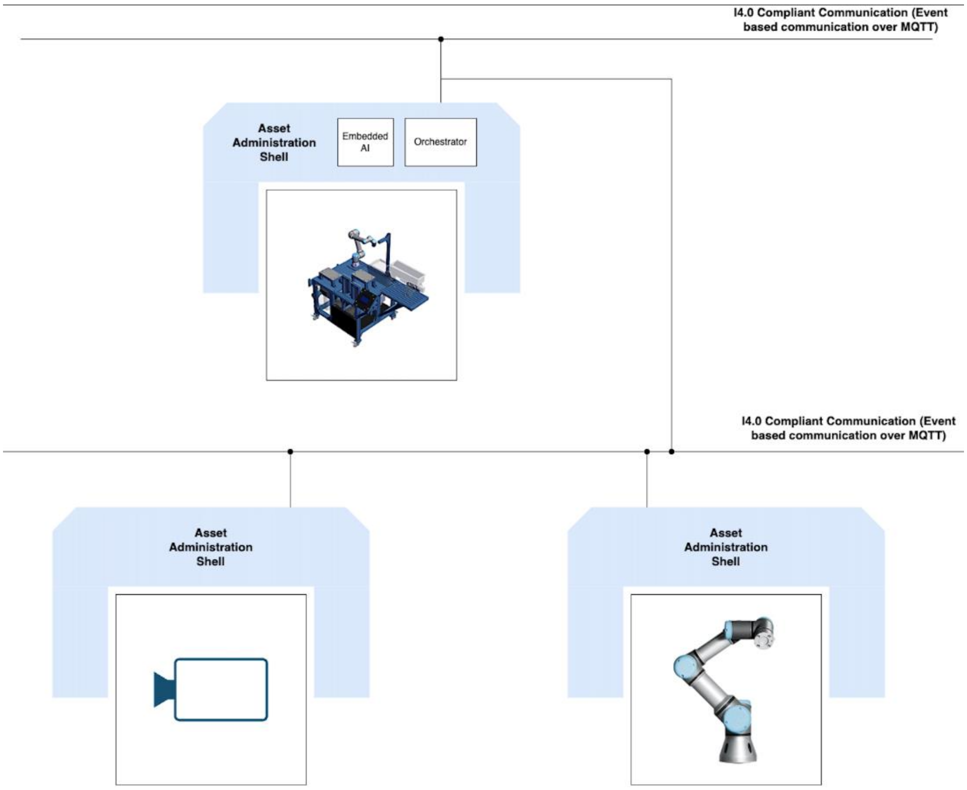

Figure 19: The process DT extends the AAS by providing orchestration and AI processing capabilities. It acts as a logically nested group of multiple sub-components

To clearly define the scope of the process DT and identify the main requirements or features to be supported a system concept has been defined (see deliverable D5.2 AI-enhanced Digital Twins Implementations for Product Process and Personae 1st version). The system concept allows the identification of the scope, boundaries, and responsibilities of the process DT. The system concept is the result of the combined analysis of the three project's pilots and helps in understanding the essential features of the system. *Table 8* shows the list of a required features for the process DT.

Table 8: List of required features for the process DT

Feature	ID	Integration Status
Dynamic data collection and extraction from different data sources at pilot's system	PDT_F_ALL_2	Completed
Capture and monitor real-time data from sensors and IoT devices attached to physical assets. This includes tracking parameters like temperature, humidity, pressure, energy consumption, as well as, visual inspection systems, robot operations, etc.	PDT_F_ALL_3	Completed
Process and/or pre-process data extracted from pilot's system	PDT_F_ALL_4	Completed
Support the execution of AI models both embedded and as a service	PDT_F_ALL_5	Completed
Synchronous and asynchronous event-based communication channels	PDT_F_ALL_8	Completed
Data representation and harmonization using the AAS standard.	PDT_F_ALL_9	Completed
Interact with external systems such as Human Digital Twin, Product Digital Twin, AASs, XAI and Data Space components to provide process data (gathered from the system) as well as the result of the execution of AI-models.	PDT_F_ALL_10	In progress
Provide easy-to-use and model-based orchestration mechanisms based on well-known standards and methodologies such as Finite State Machines (FSM), Business Process Model and Notation (BPMN), State Charts, etc.	PDT_F_ALL_11	In progress
Support the disassembly process by identifying components suitable for reuse/disassembly, tracking/assess the quality and condition of products as well as delivering a plan for disassembly.	PDT_F_ALL_12	In progress
Facilitate collaboration and communication among stakeholders involved in the circular manufacturing process. This can include sharing information, exchanging insights, and coordinating activities between designers, manufacturers, suppliers, and customers.	PDT_F_ALL_13	In progress
To provide process data to measure and assess the environmental and economic impacts of circular manufacturing practices in selected processes. This involves assessing resource consumption, waste generation, and carbon footprint, and generating meaningful reports and insights.	PDT_F_ALL_14	In progress

4.4.2 Description of the results (technologies developed/adopted)

The Circular TwAIn Process components used in the demonstrations are either developed from scratch or based on existing software components. In this deliverable the prototypes and software components that were fully developed for showcasing the capabilities of the

Circular TwAIn platform in general and process DT in particular are summarised and presented. The main outcome of the task is:

The design, configuration and deployment of a digital process made-up of Digital Twins for automating certain tasks within the considered pilots while deeply using Artificial Intelligence (AI).

Considering this, the following list of components have been developed to deliver the expected functionalities, namely:

- **Asset Administration Shell**: it is a standardized and digital representation of a physical asset in the context of Industry 4.0. It serves as a comprehensive, dynamic, and machine-readable description of an asset, encompassing not only its geometric and technical attributes but also information about its lifecycle, usage, and context within an industrial environment. The concept of AAS is designed to facilitate seamless communication, interoperability, and data exchange between diverse components of the industrial ecosystem. The Asset Administration Shell contributes to the creation of a digital twin for the physical asset. This digital representation facilitates efficient data exchange, decision-making, and optimization of industrial processes, supporting the goals of Industry 4.0 for intelligent, connected, and data-driven manufacturing environments. Taking into account the information collected from the pilots and included in the deliverable D3.2 the following physical assets are planned to be virtualized using an AAS:
    - Characterization and Assessment environment
    - Robot
    - Cameras

It is important to consider that the list of physical assets can be extended as the implementation of the solution in each of the pilot will move forward. This is practically in line with the design and development strategy (Feature-Driven Development) adopted in the Task 5.4 – in particular – and Work Package (WP) 5 in general (see deliverable D.5.2 – AI-Enhanced Digital Twins Implementations for Product Production and Personae).

Adoption of the AAS concept is the first step towards the design and implementation of more advanced functionalities and software applications. As a matter of fact, the AAS provides a solid foundation for the deployment of AI-based applications. This led to the second component that has been designed and developed, namely: Process Digital Twin (Process DT). The Process DT will play a fundamental role in exploiting data from a system populated by AASs. The process DT is composed by three core services as presented below:

- **Process Digital Twin**
    - **AI-embedded engine** for executing Visual Inspection algorithms. The AI-embedded engine has been implemented to execute You Only Look Once (YOLO) version 5 models. It is an object detection model that belongs to the YOLO family of real-time object detection algorithms. YOLO is known for its speed and accuracy in detecting objects in images and video frames. It divides the input image into a grid and predicts bounding boxes and class probabilities directly.
    - **Real-time planner of disassembly operations** for generating a list of tasks to be executed during the process. The planner is based on the Planning Domain Definition Language (PDDL) that is a computer language used in

artificial intelligence (AI) for describing the various aspects of a planning problem. It was designed to provide a standard and expressive way to represent planning domains and problems, particularly in the context of automated planning and scheduling systems. PDDL has been influential in the field of automated planning and continues to be widely used in research and applications where automated planners are employed to generate plans for complex systems. The way the plan is defined and solved is still under discussion and two possible solutions have been envisioned that could be applied taking into account where the solution will be deployed. Saying that, one possibility requires the use of ROS2 Planning System while the other considers the use of generative AI tools and models for creating a more interactive plan generation process.

- **Behaviour trees (BT) visualization and execution** for modelling the behaviour of the process DT and enabling the dynamic service composition. BT are a hierarchical and modular artificial intelligence (AI) framework commonly used in the field of game development, robotics, and other areas where decision-making processes need to be structured in a clear and flexible manner. Behaviour Trees provide a graphical representation of the decision-making logic and offer a more intuitive way to design complex behaviours compared to traditional finite state machines. Since the BT is a hierarchical tree structure it can be easily represented using eXtensible Markup Language (XML), see Figure 20.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<root BTCPP_format="4">
    <BehaviorTree ID="ProcessDTSupervisor">
        <Sequence>
            <WaitForReport name="WaitForPCReady" action="pc_ready" report="{report}"/>
            <Execute name="SnapFrontCamera" topic="front_camera" action="take_snapshot"/>
            <WaitForReport name="WaitForFrontSnap" action="take_snapshot" report="{report}"/>
            <Execute name="SnapTopCamera" topic="top_camera" action="take_snapshot"/>
            <WaitForReport name="WaitForTopSnap" action="take_snapshot" report="{report}"/>
            <GeneratePlan plan="{plan}"/>
        </Sequence>
    </BehaviorTree>
    <!-- Description of Node Models (used by Groot) -->
    <TreeNodesModel>
        <Action ID="Execute" editable="true">
            <input_port name="topic"/>
            <input_port name="action"/>
        </Action>
        <Action ID="GeneratePlan" editable="true">
            <output_port name="plan"/>
        </Action>
        <Action ID="WaitForReport" editable="true">
            <input_port name="action"/>
            <output_port name="report"/>
        </Action>
    </TreeNodesModel>
</root>
```

Figure 20: Example of a Behaviour Tree represented in XML format

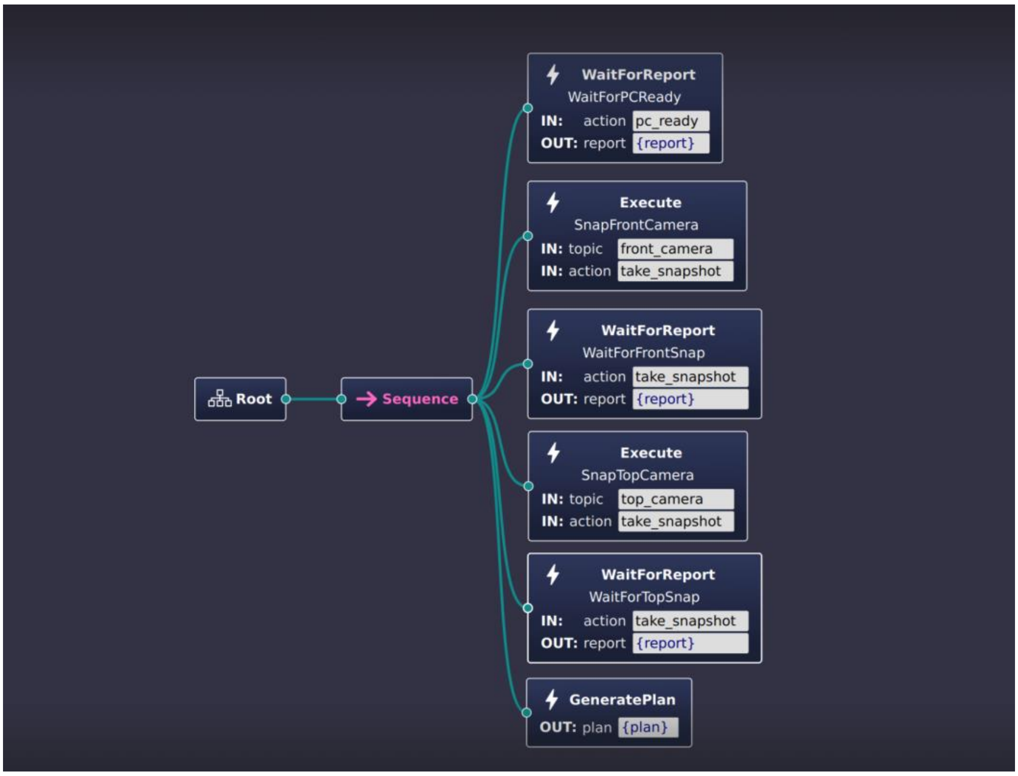

Figure 21: Graphical Representation of the Behaviour Tree for monitoring and tracking the execution of the plan

4.4.2.1 Dependencies

Until now the necessary AASs and the Process DT have been developed on the top of **NOVA Asset Administration Shell (NOVAAS)**. NOVAAS is an open-source implementation of the AAS concept using software2.0 development strategies. Both, the backend and frontend are using **Node-Red as development environment**. Therefore, NOVAAS is the baseline for building both an AAS for any physical asset that needs to be part of the Circular TwAIn system as well as for the process DT. This technological choice is driven by the need of standardizing, facilitating interoperability, ensuring consistency, supporting scalability and evolvability, and promoting data management and collaboration. For creating instances of AAS the **AASX package explorer** is used. Currently NOVAAS only supports version 2 of the AAS meta-model specification but a plan to migrate to version 3 is in place. Looking at the core services of the process DT:

- the AI-embedded engine requires the **TehnsorFlow.js JavaScript library** to define, train, and run machine learning models (**YOLO v5**) directly in a Node.js environment.
- The Real-time planner of disassembly operations is still under definition (as stated in section 4.2.2 and two main paths have been delineated. The first one requires the use of the ROS2 PlanSys software stack for solving PDDL model information.
- Behaviour trees execution and visualization requires the installation of the **Robotic Operating System (ROS)** and the **behaviourTree.cpp library** for the execution part while the visualization part is built using the **Rete.js JavaScript library** that allows to build interactive diagrams and visualizations in web applications.

4.4.3 Mapping to the requirements provided in D2.2, Section 5

Table 9 provides a mapping between the technical requirements presented in D2.2 and the current outputs of task T5.4, presenting who the former are fulfilled by the latter and presenting a justification for each case.

Table 9: Mapping to the requirements provided in D2.2

Requirement	Fulfilment by Task 5.4	Justification
Connectivity	This requirement consists at lower level in ensuring the data exchange and data flow between the process DT and all the physical resources that are part of it while at the upper level in ensuring the data exchange and sharing of process data between all the actors' and stakeholders in a circular economy ecosystem. Partially fulfilled by the Process DT.	The process DT will be capable to consume data from a system composed of AASs and orchestrate the process. As for the upper-level data exchange and sharing, the usage of certified IDS connector (such as true connector, eclipse DS components, etc) provides the right mechanisms for data exchange and sharing. This connector will be used as a service and further discussion is needed to understand if and which process data should be directly shared within a data space.
Decision Support or Actuation	The process DT makes sense in a fully digitalised system, where data can be gathered from the process asset, processed and results can be used to trigger actions, notifications etc. All these mechanisms are currently in place within the process DT framework.	The process DT is capable to receive events from the process resources/assets, process internally the events and trigger actions thanks according to the embedded behavioural model.
Common Ontology and Semantic	Fulfilled by the adoption of the AAS standard.	The process DT is an AAS that deeply relies on AI for further processing data gather from a system composed by AASs.
Material/Info traceability	Fulfilled by through the data exchange between process and product DT.	The process DT is sending the results of each process activity performed on the product to the product DT for creating/updating the Digital Product Passport (DPP).
Waste Reduction	Fulfilled by the data provided by the Process DT.	Data from the process is gathered and transmitted by the process DT for further analysis. Moreover, the process DT allows for embedding algorithms for processing process data for evaluating KPIs for waste and productivity.

Developed software should be scalable and be structured in a way that supports the deployment on various systems	Fulfilled by the process DT's components.	Developed system is composed of different modules that can be easily extended and is based on containers for multi-system deployment support.
Easy integration and deployment		
Algorithms availability	Task 5.4 and process DT allows the integration of algorithms for processing unstructured data.	YOLO v5 models can be executed from within the process DT to help in object detection and classification.
Support for active learning and Evaluation metrics	Process DT has been designed to include mechanisms for supporting the update of the learning models during the lifecycle.	As a part of the process DT the XAI service can be used to highlight how the algorithms are detecting and classifying the provided data.
Human oversight	The human is a critical part of the process for this reason the process DT is designed to send notifications to the human and to ask for permissions.	Due to the complexity of the system under study there is a need to include humans and ask them feedback instead of a fully automated system.
Programming	Fulfilled by the adoption of next-generation software development strategies.	Next-generation software development involves a set of tooling that aids in the development of software applications such as low-code development environment and generative AI.
Sufficient data for training	Fulfilled in task 5.4 by the adoption of simulation.	The generation and use of synthetic data through simulation is deeply used in task 5.4 to train process DT embedded algorithms while giving time to the pilot to gather and share real data from the process.
Transparency	Fulfilled by the adoption of next-generation software development strategies.	Process DT provides embedded documentation about API and description of the implemented logic.
Monitoring and logging	Fulfilled by the process DT.	Generated logs are kept in the backend. The IDE used for developing the process DT has its own highly dynamic logging mechanism.
Robust error handling	Fulfilled by the process DT.	Error handling is done in the backend and messages are

		generated. The usage of the flow-based programming facilitates the faster resolution time as well as code testing.
Exception analysis	Fulfilled by the process DT.	Analysis of exceptions requires faster code testing. The low-code programming helps a lot in this direction by allowing for faster testing. Moreover, the use of AI for testing is part of Task5.4 development strategy.
Well-identified goals	The process DT as generic as possible however specific part needs to be developed according to the identified goals.	In task 5.4 an agile development strategy (FDD) has been used to enable a very focused development where only features that makes senses according to the identified goal in each pilot are listed in the backlog for future development.

4.4.4 General Test Cases

The following test cases have been identified for an initial validation of the process DT functionalities, interactions with other Circular TwAIn components (intercomponent testing), and capabilities. These are generic test cases, i.e., test cases that are intended to validate process DT functionalities that are common to all the pilots where the process DT is going to be deployed.

Table 10: Generic test case

Test ID	**PDT-01**	
Title	Create and Install the AAS Metamodel description of the process DT	
Pre-Requisite	AASX package explorer installed Working with AAS metamodel v2 AASX archive containing process description in JSON format	
Expected Outcome	Process DT up and running and the process described in terms of submodels and submodel elements	
Process	**Actions**	**Result**
	Identify the data that the process DT should communicate to the outside Model this data in terms of submodel and submodel elements Creates the basic events that will wrap the data to be communicated to the other components of the Circular TwAIn system	Successfully tested

	Save the description as aasx w/ JSON Include the aasx archive in the process DT code repository	

Table 11: Generic test case

Test ID	PDT-02	
Title	Monitoring of a system composed by AASs	
Pre-Requisiste	MQTT broker infrastructure service is up and running	
Expected Outcome	Data streamed from AAS to the MQTT broker and available to Process DT	
Process	**Actions**	**Result**
	The user login within the backend of each AAS The user select the specific event or events to be published to the MQTT broker The user check if the MQTT client is properly configured Events are streamed to the MQTT broker and their structure follows AAS's standard directives Streamed events reaches the process DT	Successfully tested

Table 12: Generic test case

Test ID	PDT-03	
Title	Generation of Synthetic unstructured data	
Pre-Requisite	Simulator up and running (NVIDIA OmniVerse) Access to the original training data and required pre-trained models	
Expected Outcome	Generation of data using different conditions such as orientation, light, shadow, pose, etc	
Process	**Actions**	**Result**
	Ingest training data Create new data with different conditions Generated data are stored locally and used for training the AI-models	Successfully tested

Table 13: Generic test case

Test ID	PDT-04
Title	Deploy a pre-trained AI-model within the embedded AI engine
Pre-Requisite	The AI model has been pre-trained using Google Colab

Expected Outcome	Unstructured data are processed by the newer model	
Process	**Actions**	**Result**
	Export the AI-model in tensorFlow.js format Download and Store the archive containing the model files (configuration, model, weights, etc) Import the archive within the process DT Configure the AI engine to point of the model files	Successfully tested

Table 14: Generic test case

Test ID	**PDT-05**	
Title	Working with Behaviour trees (create the process TD behaviour)	
Pre-Requisite	Orchestrator service up and running Behaviour tree in XML format	
Expected Outcome	Next action to be performed	
Process	**Actions**	**Result**
	Model the expected behaviour of the process DT Import the xml file in the behaviour tree executor Send data about the process to compute the next actions	Successfully tested

Table 15: Generic test case

Test ID	**PDT-06**	
Title	Integration between behaviour tree and process DT	
Pre-Requisite	Orchestrator service up and running Behaviour tree in XML format imported Process DT up and running ROS middleware up and running	
Expected Outcome	Next action to be performed	
Process	**Actions**	**Result**
	The process DT submits to the behaviour tree service the current status of the process The behaviour tree service execute the model and compute the next action o be performed	In progress

	The process DT receives an event with the next action to be performed	

Table 16: Generic test case

Test ID	**PDT-07**	
Title	Integration between process DT and Product DT	
Pre-Requisite	Process DT up and running Product DT up and running MQTT broker infrastructure service is up and running	
Expected Outcome	Product DT updates product related data based on available knowledge	
Process	**Actions**	**Result**
	The process DT monitors product/process interactions The process DT creates an event with data related with the product gathered form the process The product DT uses this data to fill the necessary information for the DPP	In progress

Table 17: Generic test case

Test ID	**PDT-08**	
Title	Learn & Update	
Pre-Requisite	Process DT up and running	
Expected Outcome	Updated Neural Network for object detection and classification	
Process	**Actions**	**Result**
	Create a knowledge enhanced data structure Store the knowledge enhanced data structure Use the knowledge enhanced data structure for improving the Neural Network classifier	In progress

Table 18: Generic test case

Test ID	**PDT-09**
Title	Integration between Process DT and Human DT (HDT)
Pre-Requisite	Process DT up and running Human DT up and running MQTT broker infrastructure service) is up and running

Expected Outcome	Process DT and HDT are capable to exchange messages by using the communication infrastructure deployed	
Process	**Actions**	**Result**
	The process DT monitors product/process interactions (including human tasks) The HDT uses the information to update the worker's profile as well as to refine its own predictions	In progress

4.5 Task 5.5

4.5.1 Requirements (internal)

In the following tables, we have included comments about the status for each of the foreseen requirements.

Table 19: Mapping to requirements

Requirement (technical or business)	KPI	Validation criteria	Status
Human Digital Twin	The human DT should be able to represent the human workers in at least one of the Project use cases considering experience and skills.	Humans will be represented with the DT. The information contained within should be enough to represent them.	Humans can be described by means of a meta-model that allows users to describe any relevant characteristics (e.g., height, age, etc.). While the meta-model is general enough to also account skills, within CT we foreseen an extension of the meta-model to include specific classes describing skills.
AI modules deployment	The CT framework should support the deployment of both local and cloud AI modules.	The AI modules can be deployed on cloud or on premises.	From the HDT perspective, AI modules are "functional modules" producing some data related to factory entities (in this case, humans). Modules can be "activated" to make the HDT aware of their existence. It is not relevant where the module is deployed, given that it is never

			queried by the HDT (the communication happens by means of a data middleware, i.e., MQTT.
Real-time data sharing	Workers' inputs (e.g., instructions, sensors signals, …) should be transmitted to the HDT in real-time.	The worker's inputs are transmitted to the HDT in real-time.	The HDT manages the near real-time communication using a lightweight protocol for data exchange (i.e., MQTT).
Interoperability with external systems	The HDT should be easily integrated with other software systems.	The HDT is integrated with the other relevant software systems of the framework.	The HDT can be integrated with other systems in different ways: 1) by invoking its REST API. The API provides information about the factory entities, their characteristics, and the active functional modules; 2) by exchanging data on the data middleware using the MQTT protocol; 3) by querying historical data about humans by means of a specific REST API.
Interoperability with Product/Process DT	The HDT should be able to interact with the Product/Process DT.	The HDT can interacts with the Product/Process DT.	The interaction with other DTs is foreseen via REST APIs.
Human centric sustainable manufacturing	Enable human centric sustainable manufacturing.	The HDT enables human centric sustainable manufacturing.	

Table 20: List of requirements

Req #	Requirement	Description	Priority	Status
HDT-01	Data Protection	Data shall be collected and shared in a secure way (e.g., encrypted, …).	Critical	Supported: data transfer with HTTPS, API key auth
HDT-02	Data Storing	TBD: Define how and which data shall be stored.	Critical	HDT supports storage, data from sensors, from AI modules

				and workers characteristics (e.g., workers skills)
HDT-03	Data Anonymization	No personal, private or confidential information shall be stored.	Critical	Information stored in HDT is expected to be anonymized in advance.
HDT-04	Interaction with human worker	The HDT shall be able to exchange information with the human worker.	Critical	Foreseen by using dashboards.
HDT-05	Interaction with Product DT	The HDT shall be able to interact with the Product DT.	Preferred	The interaction with other DTs is foreseen via REST APIs.
HDT-06	Interaction with Process/Production DT	The HDT shall be able to interact with the Process/Production DT.	Preferred	The interaction with other DTs is foreseen via REST APIs.
HDT-07	User Interface	The HDT shall present a human readable interface.	Critical	Foreseen by using dashboards.
HDT-08	Real-time data sharing	Workers' inputs (e.g., instructions, sensors signals, …) shall be transmitted to the HDT in real-time.	Critical	MQTT server is available.
HDT-09	Trusted human-robot interaction	AI Algorithms used in the HDT should provide explainable outputs.	Preferred	TBD
HDT-10	Human-robot collaboration	The HDT should facilitate human-robot collaboration.	Preferred	Representation of workers' skills in HDT can meet these requirements.
HDT-11	Worker safety	The HDT should increase worker's safety.	Preferred	
HDT-12	Worker well-being	The HDT should increase worker's well-being.	Preferred	
HDT-13	Train AI by demonstrating a task	The HDT should enable a worker to program a robot by demonstrating a given task.	Preferred	TBD

4.5.2 Description of the results (technologies developed/adopted)

During the development done in the task, we have adopted Clawdite as the main technology for the HDT. Indeed, thanks to its flexibility and the underlying data model, Clawdite already satisfies many of the requirements listed in section "Requirements (internal)": the interaction with other DTs is guaranteed by the multiple agents and gateway pluggable to Clawdite; an orchestrator is available to activate new functional modules, acting as AI models providing new evidence from the available data; finally, its IIoT middleware serves as a mean for exchanging data in real-time (event-based). From the detailed analysis reported in D5.2, three main macro-blocks were missing to let Clawdite perfectly met the logical HDT design (provided as part of D5.2):

1. A new digital model for representing human profiles (i.e., workers' skills and knowledge)
2. New data connectors for interacting with Data Spaces
3. Additional agents to interact with other DTs.

In the following, we describe the possible scenario we foreseen to exploit the HDT, which represents a main result of the activities done in RP1. In our view, the HDT can support process optimization by suggesting the best workers to assign to each task in the process, leading to many advantages:

1. Workers with the right level of skills can complete the task in less time, optimizing the cycle time, and possibly reducing the machine time (energy saving).
2. High-skilled workers are less prone to errors, reducing the number of defective products (waste reduction and less reworking)
3. Assigning hard tasks to low-skilled workers may increase their stress level, negatively impacting their job satisfaction as well as their performance. However, if the skill level is not very far from the entry requirements, low-skilled workers can improve their skills by learning how to solve new more complex tasks.

In the following, we present a possible information flow that can support the above scenario:

1. The HDT receives the process plan as defined by the Process DT. The plan consists of a list of tasks, each one with the ID of the worker that should perform it. The plan should provide also tasks order, their dependencies, and the average time required to complete the task execution.
2. Depending on the information available to the HDT, different interactions are possible (not mutually exclusive):
    a. To send back to the process DT information about the workers involved in the plan, specifying:
        i. Whether the workers can perform the assigned task (based on their skills)
        ii. How long it would take them to perform the task (based on their experience)

   b. To send back to the process DT estimations about the final product quality and the estimated completion time, given the current plan, and based on workers skills and experience.
   c. To suggest different workers for certain tasks, depending on different strategies. For instance, if a worker is a trainee, the HDT can suggest:
      i. To assign a certain task to the trainee so to practice more and acquire new skills.
      ii. To assign a trainer and perform a shadowing activity, if the trainee has never performed the task.
3. The Process DT uses the new evidence provided by the HDT to validate the original plan or refine it. In case of a new plan, it can be sent again to the HDT and re-evaluated (iterative evaluation).
4. When a process is confirmed, executed and completed, the process DT may send to the HDT the information about the actual completion time (per worker, per task). Also, if available, the Product DT can provide the HDT with information about the final products' quality. The HDT exploits these new pieces of information to update the worker's profile and refine its own estimations:
   a. e.g., if the product is of good quality, the HDT can infer that the worker performed the task correctly. The worker's ability level for that task can be updated (likely increased).
   b. e.g., by counting the number of processed products, the HDT could infer the average speed of a specific worker in performing a certain task, thus updating the worker ability.

Concerning predictions and estimations (e.g., to estimate the final product quality, or to suggest a better worker for a task), they can be AI-driven. Depending on the underlying AI models, suggestions/predictions can also be explained by means of XAI.

Regardless of the prediction models, the HDT needs to represent the following information to support the presented flow:

1. **Human worker's skills**. To optimize the production process by, for instance, assigning very demanding tasks to high-skilled workers, information about worker's skills should be available. In our view, a human masters a skill with a certain level of confidence (a score on a predefined scale, e.g., 1-5). The skill level can be exploited to assess, for instance, the probability of a worker to complete a certain task requiring a specific skill level (the assumption is that different people can be more/less able to perform the same task).
2. **Relevant tasks**. The HDT should be aware of the tasks that workers perform, along with the entry requirements needed to fulfil each task. In this way, the HDT can be exploited for supporting activities such as job rotation and skill improvement. Each task should describe:
   - the list of skills required to complete it, along with the minimum threshold required for each skill.
   - the list of workers that should not cope with the task, because of external factors (e.g., health conditions).

3. **Manufacturing plans.** This information is provided by the Process DT and it is represented as a list of tasks, each one with its assigned worker (we assume a task can be assigned to only one worker to avoid adding too much complexity). For each task, additional information can be provided (e.g., estimated completion time, on which product the human is working, ...).
4. **Actual execution time**: This information is provided by the Process DT as soon as a plan is completed.
5. **Product quality**. Details about the quality of the final product are provided by the Product DT.

In this period, we started addressing the first point, i.e., skill representation. We leveraged the results observed in other EU projects where the O*NET Content Model (https://www.onetcenter.org/content.html) has been used to support workers profiling and training programmes (e.g., the STAR project, https://star-ai.eu/).

The Content Model is the conceptual foundation of O*NET and identifies the most important types of information about work. Figure 22 depicts the foundational blocks of the content model.

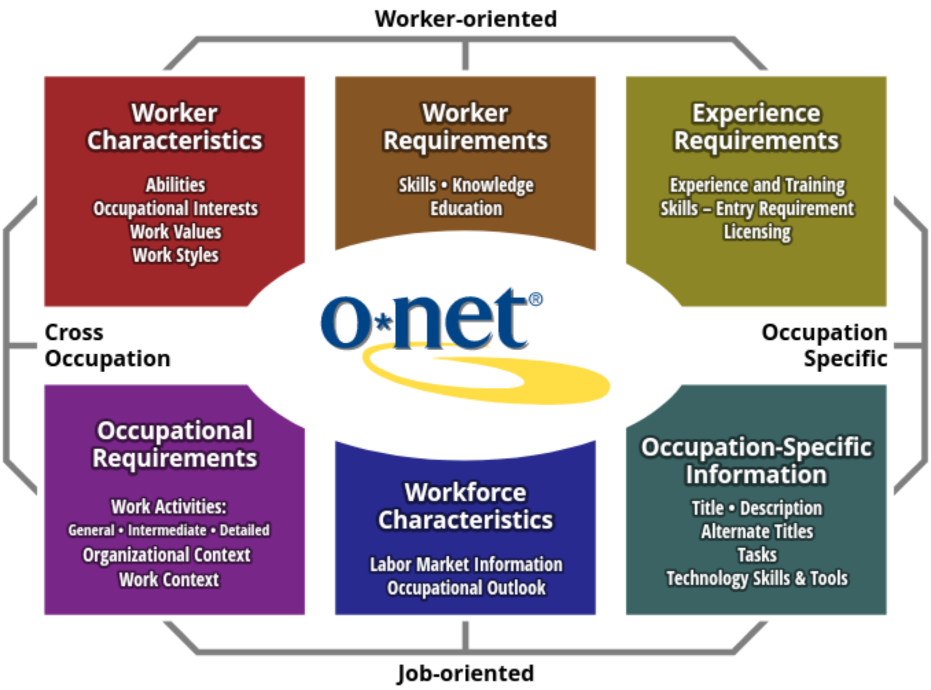

Figure 22: The O*NET Content Model overview

While the model is comprehensive and integrates many aspects into a theoretically and empirically sound system, for pursuing the project purpose (i.e., supporting the scenario presented above), some aspects are more relevant than other. As part of RP1 activities, we started extracting the data model parts that are relevant for the project so that to pave the next integration with Clawdite data model. To this end, it should be noted that Clawdite data model already implements the concept of “skill”, which however is very abstract and resembled to a basic "worker characteristic”. We consider the existing data model not rich enough to support the presented scenario, thus we decided to extend the data model with some new classes providing a more fine-grained information. A simplified overview of the O*Net data model is provided in Figure 23 (we mapped classes to the different Content Model areas by using the same colours).

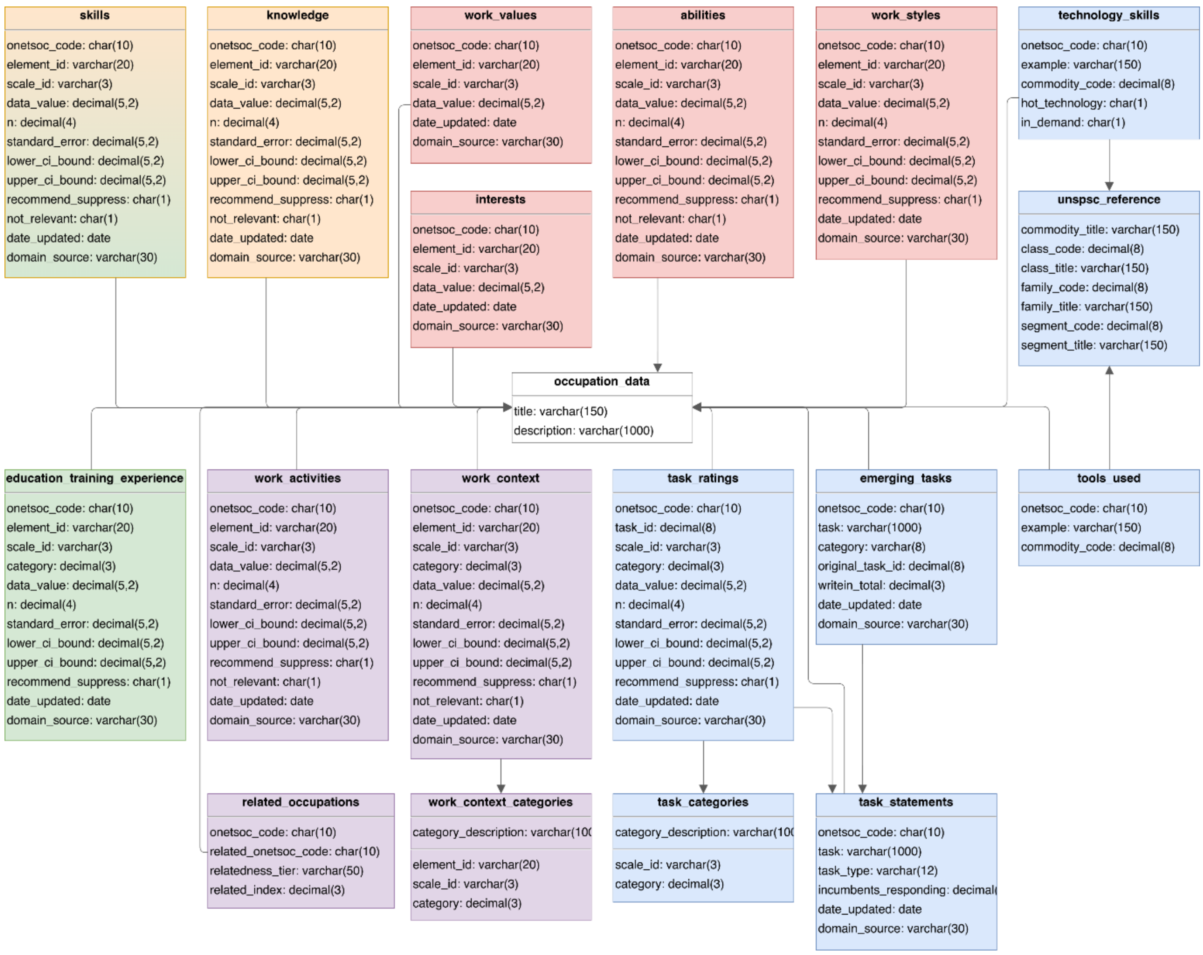

Figure 23: An excerpt of the O*Net data model, representing relevant classes to include in the HDT data model

The data model revolves around the concept of occupation data, which in turn is connected to different occupation aspects. Classes reported in Figure 23 represent the core entities to be reported within the Clawdite data model as well: skills and education/training (experience); skills and knowledge (worker requirements); work values and styles, interests, and abilities (worker characteristics); work context, activities, and related occupations (occupational requirements); tasks, technology skills and tools (occupation-specific information). Classes related to the area Workforce Characteristics are not considered, at least in this first version of the HDT (they are broader and focused on factors external to the company (i.e., labour market and occupational outlook).

5. Mapping of Circular TwAIn Pilots' needs and WP5 outcomes

In this section we provide a detailed overview of how are outcomes from WP5 covering the requirements from pilots.

In following subsections, each pilot is analysed by WP5 tasks.

5.1 WEEE

5.1.1 Pilot requirements vs WP5 Outcomes

The Table 21 summarizes the requirements coming from each Use Case (UC) in the context of WEEE.

Table 21: WEEE pilot requirements (as reported in deliverable D2.2)

Requirement (technical or business)	KPI	Validation criteria
Use Case a:		
Assistance in the identification process	20% reduction in the time of IT equipment testing and evaluation operations	To validate the requirement, an AI-based helper must be implemented to assist the operators in the recycling facility
Use of digital tools in the classification of desktop PCs	- 20% reduction in the time of IT equipment testing and evaluation operations - Increase by 10% the amount of equipment classified	To validate the requirement, an AI-based tool must be implemented to classify WEEE waste
Collection of product and component information	20% reduction in the time of IT equipment testing and evaluation operations	To validate the requirement, a library with information about WEEE components must be compiled
Use Case b:		
Assistance in the diagnosis of desktop pc components	- 20% reduction in the time of IT equipment testing and evaluation - 20% more value generated from recycled/reused components	To validate the requirement, an AI-based helper must be implemented to assist the operators in the recycling facility
Gather Product and components SoH criteria information	20% more value generated from recycled/reused components	To validate the requirement, a library with information about WEEE components' SoH criteria must be compiled
Installation of systems to perform diagnostic activities without operator support	20% reduction in the time of IT equipment testing and evaluation	To validate the requirement, the AI-based tools developed in Use Case a must be implemented in the facility, working autonomously.
Use Case c:		
Adaptation of automation for	- Automation of more than 25% of the operations required for disassembly	To validate the requirement, the process of automatically

disassembly to the Revertia´s process	- 30% reduction in the time required to dismantle IT equipment	disassembling process must be fit the Revertia's waste processing processes
Collaborative robotic application	- Automation of more than 25% of the operations required for disassembly - 30% reduction in the time required to dismantle IT equipment	To validate the requirement, a robotic application to collaborate with the operators must be implemented
Integration of AI implementation in disassembly process	- Automation of more than 25% of the operations required for disassembly - 30% reduction in the time required to dismantle IT equipment	To validate the requirement, an IA-based module must be implemented in the robotic application to help in the disassembly process
Use Case d:		
Implementation of collaborative Robot in the recycling facility	- Automation of more than 25% of the operations required for disassembly - 30% reduction in the time required to disassemble computer equipment	To validate the requirement, the elements developed in Use Case c must be implemented in the final recycling facility
Adaptation of workers to the robot	- Automation of more than 25% of the operations required for disassembly - Increase reuse by 20% of components - 30% reduction in the time required to disassemble computer equipment	To validate the requirement, the application must fit the workers necessities
Integration of robot in the production line	- Automation of more than 25% of the operations required for disassembly - Increase reuse by 20% of components - 30% reduction in the time required to disassemble computer equipment	To validate the requirement, the elements developed in Use Case c must be integrated in the final recycling facility
Set the scope of the collaborative robot	Automation of more than 25% of the operations required for disassembly	To validate the requirement, the scope of the robotic agent developed must be clearly defined
Use Case e:		
Gathering of WEEE waste materials' historical value for AI configuration	- 20% increase of the economic impact of the recycling/reusing process - Reduce the environmental impact of the recycling/reusing process by a 20%	To validate the requirement, a library with information about WEEE components' materials must be compiled
Up to date the database on a regular basis	- 20% increase of the economic impact of the recycling/reusing process	To validate the requirement, an automated tool to update the data of the materials library

	- Reduce the environmental impact of the recycling/reusing process by a 20%	
Use of AI decision-support tool by workers	- 20% increase of the economic impact of the recycling/reusing process - Reduce the environmental impact of the recycling/reusing process by a 20%	To validate the requirement, a support decision system based in AI technologies must be implemented an in use by the operators of the recycling facility
Adaptation of waste processing facilities	- 20% increase of the economic impact of the recycling/reusing process - Reduce the environmental impact of the recycling/reusing process by a 20%	To validate the requirement, the support decision system must be integrated in the recycling processes of Revertia

5.1.2 Task 5.1

Data4AI is used for supporting the quality control of used components, which is a critical aspect of any manufacturing process that involves monitoring and testing products to ensure they meet the desired specifications and standards. In this pilot, quality control system is important for supporting the decision: to either remanufacture the computer or recycle the components of the computer.

New method for creating data-driven profiles of the computer is developed. These profiles correspond to the normal behaviour of the measured parameters. The goal is to develop different behavioural models which will reflect the condition of a used computer (bad, moderate, less used). Consequently, for each model, there will be a different process for the remanufacturing, including the impact on the automatization. For example, if a computer is in a bad condition, it should be analysed by a human technician and not by the cobot associated to this work position.

This work can be extended by analysing also the information about the usage of the used computers, like for how long and in which condition and even by whom (which task), a computer was used.

5.1.3 Task 5.2

In this pilot, the role of AI will be focusing on the supporting the processes of dis-assembly of the components. The equipment must comply with minimum technical specifications (meet obsolescence criteria). Then those components that are susceptible to disassembly must be replaced by components of similar characteristics in order to obtain a second-hand product. This will be supported by training the AI system to identify visually the different components and to suggest which of those are candidate for disassembly. Moreover, the AI system, apart from the visual recognition of components, will be able to provide prediction on some specific attributes of the components at hand, by analysing past historical data of components that have been already inspected during the past disassembly and remanufacturing processes.

The visual analysis and recognition steps will be done by AI models that will run on the Experimentation Composer, as low latency is necessary in order to directly process the images acquired during the process.

The predictions on the models characteristics will be done by ingesting data in the ERP of the WEEE pilot and making use of the cloud-based platform.

5.1.4 Task 5.3

In the context of the WEEE pilot the product DT has to fulfil the following requirements:

Table 22: Mapping to requirements

Requirement (technical or business)	**KPI**	**Validation criteria**	**Validation Status**
Product digital twin	The product digital twin should be able to represent the WEEE products, components and value, taking into account quality and uncertainty.	A number of products will be represented with the DT. The information contained within should be enough to represent them.	Done
Re-construction of missing product life-cycle data	This AI module should work seamlessly with the product DT, and estimate environmental and monetary indicators for a product.	The indicators will be compared with experts' opinions to evaluate the performance.	Ongoing
Re-construction of missing product life-cycle data	This AI module should work seamlessly with the product DT and estimate environmental and monetary indicators for a product.	The indicators will be compared with experts' opinions to evaluate the performance.	Ongoing

<u>Product DT role for WEEE pilot demo</u>

The Product DT in the WEEE pilot is connecting with various infrastructures, facilitating communication, and providing a comprehensive digital representation of products for participants in the circular economy. For this first iteration the product DT infrastructure main role is to instantiate and manage the product DT enhancing its information throughout the pc case acceptance and processing within REVERTIA's WEEE treatment. Currently product DT infrastructure collects the Open-AudIT data generated from the process DT through the MQTT broker to instantiate the product DT. The following infrastructure have been setup to accommodate the process.

- **Use of Eclipse BaSyx for AAS infrastructure (for Product DT):** Eclipse BaSyx serves as the foundation for implementing the product Asset Administration Shell (AAS) infrastructure, enabling the representation of the Product Digital Twin. It facilitates the interaction and exchange of information related to the product.

- **Integration of Product DT with DS infrastructure and DPP 4.0 (Ongoing):** The Product DT is integrated into the Data Space (DS) infrastructure, fostering seamless communication and interaction. Integration with the Data Space is vital for the exchange, modelling, and processing of product-related data. This integration is achieved using the DPP 4.0 standard.
- **Integration with Product DT and XAI infrastructure for each of the participants (Ongoing):** The Product DT is connected to the eXplainable Artificial Intelligence (XAI) infrastructure for each participant. This integration ensures that the reasoning behind decisions made by AI components related to the product is transparent and understandable.
- **MQTT server deployment for Process DT integration (Done):** MQTT server deployment facilitates the integration of the Product DT with the Process DT. This ensures that the digital representation of the product is synchronized with the information generated during the processes involved in the WEEE pilot.

The following integration steps are planed:

- **Integration of Process DT with Product DT:** The Product DT is closely integrated with the Process DT. This integration ensures that the state of the product, as represented in the Product DT, aligns with the ongoing processes, providing a comprehensive view of the product lifecycle.
- **Investigate the development of an appropriate UI for searching and validating product DTs:** User Interface (UI) development is explored to create a user-friendly interface for searching and validating Product DTs. This ensures that stakeholders can efficiently interact with and validate the digital representation of products in the context of the WEEE pilot.

5.1.5 Task 5.4

At the current state of the implementation the work done under the scope of the task 5.4 has been focused on designing and developing a system capable to fulfil the following list of requirements.

Table 23: Mapping requirements gathered from WEEE pilot

<table>
<tr><td>Requirement

(technical or business)</td><td>Fulfilment by Task 5.4</td><td>Justification</td></tr>
<tr><td colspan="3">Use Case a:</td></tr>
<tr><td>Assistance in the identification process</td><td rowspan="3">The overall WEEE process has been digitalized using the AAS concept. Specifically, two cameras have been installed and their related AAS attached to them. The acquired images are processed by the process DT for identify the internal components as well as to estimate (where possible) their quality</td><td rowspan="3">The process DT has all the mechanisms implemented to receive MQTT events form the two camera AAS, process the events using a pre-trained AI model for object identification and classification. The result of the AI algorithm execution is a list with all the identified components, probability of the identification, and an estimation of the quality (where possible)</td></tr>
<tr><td>Use of digital tools in the classification of desktop PCs</td></tr>
<tr><td>Collection of product and component information</td></tr>
</table>

Use Case b:		
Assistance in the diagnosis of desktop pc components	An assessment and characterization environment has been designed and provided with all the toolset for gathering product information. Moreover, a set of scripts have been also implemented to have analogic data about the performance of some hardware components such as hard drive, cpu, etc..	As part of the process DT the characterization and assessment environment has been designed and developed. The process DT triggers the assessment of a product and receives the full report of the task.
Gather Product and components SoH criteria information		
Installation of systems to perform diagnostic activities without operator support		
Use Case c:		
Adaptation of automation for disassembly to the Revertia´s process	The AI-enabled process DT has been designed to facilitate the human-robot collaboration by computing a disassembly plan where human and robots are the main executors. Interfacing the process DT with the HDT and with the robot AAS a disassembly plan is generated and actions are passed to the responsible actors.	The process DT uses the behaviour tree to model the behaviour of the system and triggers all the necessary actions. The plan generation and execution is part of the behaviour tree. As for the plan generation there are still some open discussion about how it is generated. Generative AI techniques and/or already implemented planners are going to be considered.
Collaborative robotic application		
Integration of AI implementation in disassembly process		

Generally speaking the UC a, b and partially c have been prioritized, i.e., have been selected for passing from design to development and deployment.

Figure 24, shows how the initially system topology looks like in the WEEE pilot.

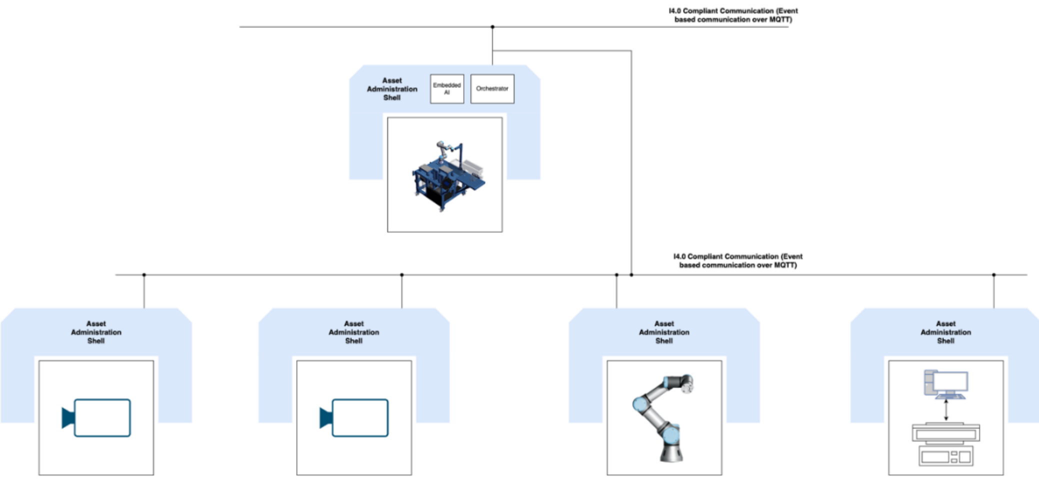

Figure 24: How the process DT environment changes for the WEEE pilot

5.1.5.1 Use Case a: Computer-vision driven product identification for the disassembly of IT equipment

The purpose of this UC is the employment of AI techniques for processing unstructured data (images) from a camera to automatically identify the hardware configuration of a pc. To deliver this feature/capability two camera AAS, XAI and process DT are used (see Figure 24). The AASs are used to digitalize the available physical assets. The AAS connects to the camera to take a snapshot of the product. The images are then sent to the process DT as MQTT events. The process DT is in turn responsible to gather the snapshots and apply pre-trained AI object detection and classification models. Furthermore, AI models for quality estimation are under development and will enable to improve more the automation of the whole system by delivering (where it is possible) an estimation of the quality of the identified components based on metrics that are still under discussion. To facilitate the comprehension on how the models are delivering their results heatmaps can be also generated by the XAI component.

The whole system resides within the company edge.

Dependencies

For the realization of the UC the following dependencies need to be included:

a) **Robotic Operating System (ROS):** it is an open-source middleware framework designed to develop and control robots. Despite its name, ROS is not an actual operating system; instead, it provides a set of software libraries and tools that facilitate the development of robotic applications. The ROS middleware is responsible to manage the communication with the physical assets (cameras) and it used by the AAS to retrieve data from the assets and to expose it by using the AAS standard.

b) **EigenCAM:** The XAI component uses the EigenCAM package for YOLOv5 to include state-of-the-art methods for Explainable AI for computer vision. The XAI component will be used as a service by the process DT for diagnosis model predictions.

These dependencies are specific to the WEEE pilot and are required to allow the correct functioning of the process DT.

Table 24: Specific test case

Test ID	**PDT-WEEE-IDENT-1**	
Title	Integration between AAS and the Camera	
Pre-Requisite	Camera connected to the pc through USB ROS middleware up and running Camera AAS up and running with ROS installed	
Expected Outcome	The Camera AAS receives a snapshot of the product	
Process	Actions	Result
	The user triggers the “take a snapshot” operation from the AAS UI A request is sent to the camera suing the ROS middleware A snapshot is sent to the AAS using the ROS middleware	Successfully tested

Table 25: Specific test case

Test ID	PDT-WEEE-IDENT-2	
Title	Apply AI detection algorithms	
Pre-Requisite	Process DT up and running Camera AASs up and running Process DT embedded AI model ready Behaviour tree service up and running MQTT broker infrastructure service is up and running	
Expected Outcome	A report containing the list of detected objects with type and quality estimation (where possible)	
Process	Actions	Result
	The behaviour tree service ask to the camera AAS to take a snapshot A request is sent to the camera using the ROS middleware A snapshot is sent back to the AAS using the ROS middleware	In progress

Table 26: Specific test case

Test ID	PDT-WEEE-IDENT-3	
Title	Camera AAS UI	
Pre-Requisite	Camera AASs up and running	
Expected Outcome	The current snapshot is shown on the UI	
Process	Actions	Result
	Press the take a snapshot button The current snapshot is shown on the AAS UI	In progress

5.1.5.2 Use Case b: Characterization and Assessment of Components and Subcomponents

The Characterization and Assessment feature is aimed to initially characterize and assess the product, i.e. to identify the hardware configuration of the products as well as to perform stress tests to be further analyzed and processed. In the WEEE pilot a set of tools have been developed and installed to assist humans during this activity. To facilitate the integration of such tools and the data exchange an AAS has been designed and developed (i.e., Characterization and Assessment AAS). The Characterization and Assessment AAS is the component that has been developed to create a comprehensive and standardized digital representation of the characterization and Assessment Environment while serving as a framework for integrating diverse data – related to the asset – within the Circular TwAIn platform. It enables the real-time monitoring of the characterization and assessment of the products by providing timely and accurate information about the status and configuration of the products to be processed.

The Characterization and Assessment AAS resides within the company edge. Data from the AAS can be extracted using a request/reply message exchange pattern (MEP) or using event-based communication where the AAS publish an MQTT event as soon as the assessment and characterization process has been concluded.

Dependencies

The AAS assumes the existence of an external portable ssd with Ubuntu pre-installed and configured to include:

a) **Open AudIT**: this tool allows to sniff the Local Area Network (LAN) to retrieve the exact configuration of any machine connected to the network (aka discovery process). It is a database of queriable information. Queries can be executed by using the provided REST Application Programming Interface (API).

b) **Sysbench:** it is a versatile and widely-used benchmark tool that evaluates the performance of various system parameters on Unix-like systems. It is often employed to test the CPU, memory, file I/O, and database performance of a system. The tool is written in Lua and is extensible.

c) **Lm-sensors package:** it is a collection of user-space utilities that allow monitoring various hardware sensors on a computer, including temperature, voltage, and fan speed sensors.

and has been designed to expose both the result of the discovery process – produced by the Open AudIT tool – as well as the result of the execution of the developed stress test scripts.

Table 27: Specific test case

Test ID	**PDT-WEEE-ASS-1**	
Title	Preparation of the environment	
Pre-Requisite	Ubuntu installed on an ssd Open AudIT installed and its service up and running Test Stress libraries installed Node-Red up and running with REST API and logic implemented for stress tests	
Expected Outcome	Environment up and running Open-AudIT discovery report and test stress assessment report	
Process	**Actions**	**Result**
	n/a	Successfully tested

Table 28: Specific test case

Test ID	**PDT-WEEE-ASS-2**
Title	Integration between AAS and the Characterization and Assessment environment

Pre-Requisite	Environment up and running Characterization and Assessment AAS up and running	
Expected Outcome	The AAS receives the Open-AudIT and test stress assessment reports	
Process	**Actions**	**Result**
	Login into Open-AudIT Setup Environment Execute the Assessment	Successfully tested

Table 29: Specific test cases

Test ID	**PDT-WEEE-ASS-3**	
Title	Characterization and Assessment AAS UI	
Pre-Requisite	Characterization and Assessment AAS up and running	
Expected Outcome	The UI shows the commands for interacting with the Characterization and Assessment Environment Message logs are shown The result of the Open-AudIT execution is presented on the UI The Metamodel is automatically generated depending on the result of the stress test	
Process	**Actions**	**Result**
	n/a	Successfully tested

The BATTERY pilot requires almost the same features/functionalities as the WEEE pilot, this means that - in terms of design – the same architecture and list of components is used. However, some specific parts needs to be implemented and particularized such as the connectors to the physical assets, the neural network model, specific algorithms and scripts for product characterization and assessment, etc…

5.1.6 Task 5.5

At the current state of the implementation, most of the activities done in the context of task 5.5 have been focused on the design of a system that could contribute to achieve some of the requirements listed in Table 23. Indeed, some of the KPIs can be achieved by jointly exploiting different DTs (e.g., a combination of Process DT and Human DT), We found that the HDT can help the fulfilment of the following KPIs:

- Collaborative robotic application (UC C): the Process DT will try to optimize the disassembling process in such a way to automatize more than 25% of the operations; however, the specific operations to automatize may depend on the specific skills and abilities of the worker, which are provided by the HDT.
- Adaptation of workers to the robot (UC D): different workers have different necessities. To let the Process DT achieving this KPI, the HDT should provide the

updated list of worker necessities, which may depend on external factors (e.g., worker's career path).

- Use of AI decision-support tool by workers (UC E): explaining AI decisions to humans may require adapting the explanation details level to the human knowledge. In this case, the HDT can be exploited to provide the set of explanations that can either be disclosed to the worker or can be understood by the worker (e.g., because of their limited knowledge on a specific topic).

As part of the design activities, partners involved in T5.5 prepared a questionnaire (available in Appendix A) for the WEEE pilot implementation partners to understand to which extent the HDT can be exploited to fulfil the KPIs above. Indeed, as of today, how to exploit the HDT within the WEEE pilot is still under discussion.

From the collected answers, it is clear that the size of the WEEE pilot is very small (2 human operators involved) and does not account for many optimizations. However, new operators are trained in different ways, i.e., with hands-on training (practice sessions for developing new skills), on-the-job training (some duties performed under close supervision), and shadowing (observe and learn). The HDT can support the workers training and upskilling by suggesting which is the best strategy to adopt for a given operator.

Concerning the decision-making processes, all of them are not consensual (a single operator takes the final decision, with no agreement from other co-workers). Particularly interesting is the role of the disassembly worker, who is going to take decisions about the steps to take for disassembling a computer, as well as the classification of computer's components (check the status of the waste). In this view, the HDT can serve as a proxy to decide whether to trust the operator decision (e.g., by considering if a worker usually misclassifies a certain computer component). This is crucial because a wrong decision may produce delays in the disassembly process, thus interfering with the work of other co-workers in the facility. About the final quality of the product, the Sales assistant covers a crucial role in decision-making. Indeed, the Sales assistant is the person in charge of deciding the destiny of the waste (either determine if it will be recycled, remanufactured, or reconditioned). Wrong decisions will lead to economic losses and waste of materials. In this case, the HDT can provide a confidence score about the Sales assistance decision, so that to recommend a review of the decision when the confidence score is below a certain threshold (e.g., by asking another assistant to provide another decision).

Data about workers are collected within the pilot in different ways:

- Knowledge: the level of worker training is acquired every year (last collection: 2022).
- Expertise: it is mainly based on the operations performed in the recent past.
- Taken decisions: registration of values of catches, treatments, reconditioning, etc. Finally, about process optimization and job allocation, tasks are assigned depending on the operator experience and skills, with a direct impact on the tasks assigned and the potential results achieved by the operators. Also, operators are assigned to a limited set of tasks, thus re-allocations within the same process should be addressed wisely.

While the role of the HDT within the pilot is still under discussion, from the collected responses we can assume that the work carried out in T5.5 can be mapped and implemented in WP6 in different ways, thus demonstrating the value of the developed system.

5.2 BATTERY

5.2.1 Technical requirements vs WP5 Outcomes

Table 30: BATTERY pilot requirements (as reported in deliverable D2.2)

Requirement (technical or business)	KPI	Validation criteria
Technical: battery pack/module disassembly	KPI1: Reduction of battery disassembly time by 10%, compared to fully manual disassembly.	To perform manual disassembly (Packs: Raeeman, Modules: Polimi) on the same battery models to create a benchmark
	KPI2: Achievement of battery disassembly tasks automation of 20% (currently 0%).	
Business: higher added value of the EoL product	KPI3: Increase by 25% of the fraction of reusable batteries over the total collected.	To estimate the number of battery modules reusable without cell level disassembly to create a benchmark
Technical: faster testing protocols	KPI4: Reduction of testing time by 15%.	To perform standard testing cycles on the same battery cell models to create a benchmark
Business: more efficient decision making	KPI5: Increase of profitability rate with respect to standard recycling by 20%.	The evaluated economic profile of Use Case 5 is compared with the standard recycling profitability rate by Cobat

5.2.2 Task 5.1

For the BATTERY pilot very similar methods as in WEEE pilot will be used.

5.2.3 Task 5.2

For the BATTERY pilot, the use of the AI toolkit is expected as following.

Computer vision models will be used to feed robotic operations for managing the process of battery disassembly procedures. These will be executed on the “Experimentation Composer” as these operations need to run at the locally and connect to the robotic machinery.

To aid these procedures, ML models will be used that will be also executed in the “Experimentation Composer” facility, while the training of those models will happen in the Cloud platform. The local execution of those models is required, as they need to be fed by the outputs of the computer vision models in order to help the operations of the disassembly in real-time.

The Cloud based platform, aside for providing support for training the above mentioned models, will be responsible for executing AI models that are able to ingest historical and test/real-life data for aiding in the characterization of the LIBs state-of-health. Moreover, optimisation algorithms will be used to aid the reconfigurable mechanical recycling technologies. Finally, prediction models will be used to derive to different scenarios that are able to support decisions that can result in highest profitability of the different LIBs circular economy scenarios.

5.2.4 Task 5.3

In the Circular TwAIn Battery pilot, the Product Digital Twin (DT) will be used for persisting the characteristics of batteries along with synthetic data derived from performance tests of battery cells. This enables a comprehensive and data-driven approach to the optimization of the current circular value chain for automotive battery systems. The battery pilot aims to enhance recovery processes by integrating circular economy nodes, from preliminary tests to recycling treatments, thus ensuring a sustainable and efficient approach to end-of-life battery management. The Product DT is particularly instrumental in achieving this by facilitating the seamless storage and retrieval of battery data throughout its lifecycle. The use cases, including computer-vision-driven collaborative robotics, machine learning-aided automated disassembly, AI tools for state-of-health characterization, optimized mechanical recycling, and a holistic decision-support system, leverage the Product DT to enhance decision-making processes, support automation, and contribute to key performance indicators (KPIs) such as reduced disassembly time, increased automation, higher reuse fraction, reduced testing time, and improved profitability rates.

5.2.5 Task 5.4

At the current state of the implementation the work done under the scope of the task 5.4 has been focused on designing and developing a system capable to fulfil the following list of requirements see (Table 31Table 23).

Table 31: Mapping requirements gathered from BATTERY pilot

Requirement **(technical or business)**	***Fulfilment* by Task 5.4**	**Justification**
Technical: battery pack/module disassembly	The process DT includes functionalities and features based on AI for identify *components* and subcomponents of the battery as well as for generating an disassembly plan for aiding the human during the process.	One of the main objectives of the process DT is to support the human expert during the disassembly of battery pack/module.
Technical: faster testing protocols	As part of the task we are studying the possibility *to* implement an assessment and characterization environment (the same as in WEEE pilot) to assess and characterize the state-of-health of the LIBs while improving the assessment and characterization process currently in use.	Characterization and assessment of the battery pack is the initial step for assessing the state-of-health of the battery before proceeding with the disassembly. A better characterizatio*n* is directly connected to an overall better use of the battery. All this information are also needed to populate the DPP and more in general in a circular environment.

5.2.6 Task 5.5

For the BATTERY pilot, the mapping between the HDT capabilities and the pilot needs is still to be fixed. The pilot is investigating the opportunity to involve one or two human operators in the process (which is, as of today, designed as a fully automated process). By inspecting the collected answers to the questionnaire, human operators are mainly required to make decisions about whether to intervene in the process, whether to reuse a battery cell or recycling it. In this case, the HDT can be exploited to assign a score to the operator decisions (e.g., operators that are expert on a certain topic may take much more informed decisions than non-expert operators). For this pilot, if humans are involved in the process, they are going to be assigned at most a couple of tasks, based on their skills, mastership, and expertise. Thus, also here the HDT represents a good candidate for suggesting job-reallocation to the Process DT.

5.3 PETROCHEMICAL

5.3.1 Pilot requirements vs WP5 Outcomes

The Table 32Table 21 provides an overview of the requirements coming from each Use Case (UC) in the context of the Petrochemical Pilot.

Table 32: PETROCHEMICAL pilot requirements (as reported in deliverable D2.2)

Requirement **(technical or business)**	**KPI**	**Validation criteria**
Use case A:		
Acquisition of required physical sensor data for AutoML tool **Acquisition of laboratory data for EO and CO2** **Installation of Sampling Point for EO and CO2**	Collection of physical sensor data Pre-processing of the gathered laboratory data Installation of sampling point for CO2 and EO	The delivery of datasets required for DT.
Use case B:		
Modelling of the EO stripper **Construction of the DME simulation**	Prediction EO content in the stripper bottom Prediction of flow rate of the stripper bottom Prediction of column bottom temperature Prediction of DME product flow rate	To validate the requirement, real process data will be compared with modelling results.
Use case C:		

Identification of abnormal process conditions	>90% accuracy in detecting anomalies related to the EO stripping process	To validate the requirement, when a process failure happens the predictive model will be checked whether it predicts correctly the failure or not
Use case D:		
AutoML tool development **AutoML tool testing**	AI module that describes EO Stripper unit with a maximum 5% prediction error. At least >=3 different machine learning algorithm will be implemented inside autoML tool.	To validate the requirement, the predictions of the AutoML tool will be compared with the laboratory results
Use case E:		
C-205 Process optimization	To reduce the LS consumption of EO stripper % 5 reductions in CO2 emission caused by the steam consumed in E-208, strategic level	To validate the requirement, the consumption of LS will be monitored and compared with the previous state.

5.3.2 Task 5.1

Data4AI Platform is used for variation analysis of the input data, as presented in Figure 25

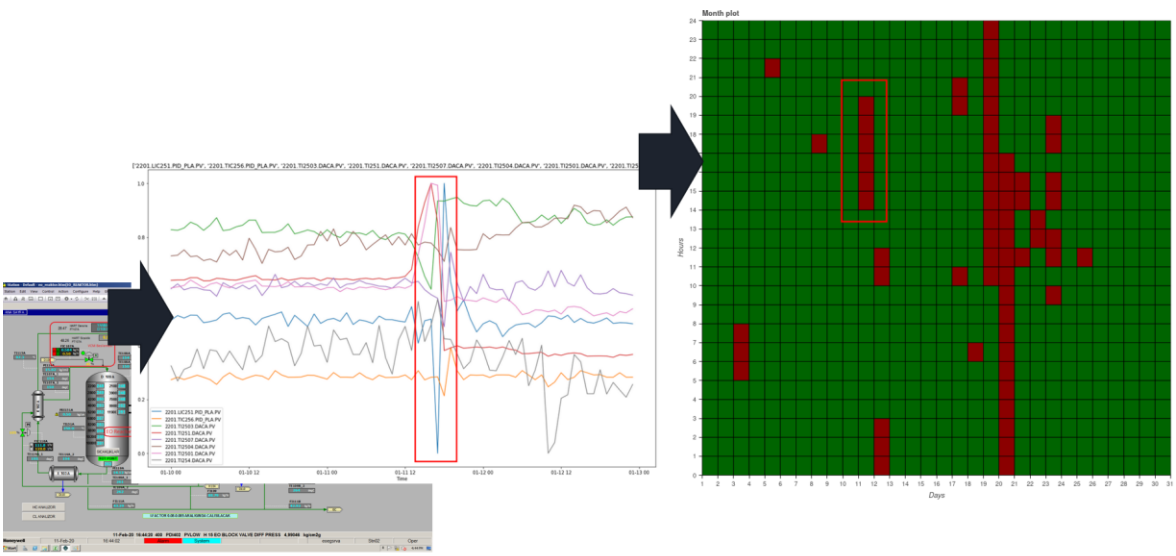

Figure 25: An example of the analysis performed on input data.

Each red section represents a period with variations (calculated using various methods).

Especially important feature is the root cause analysis, illustrated in Figure 26.

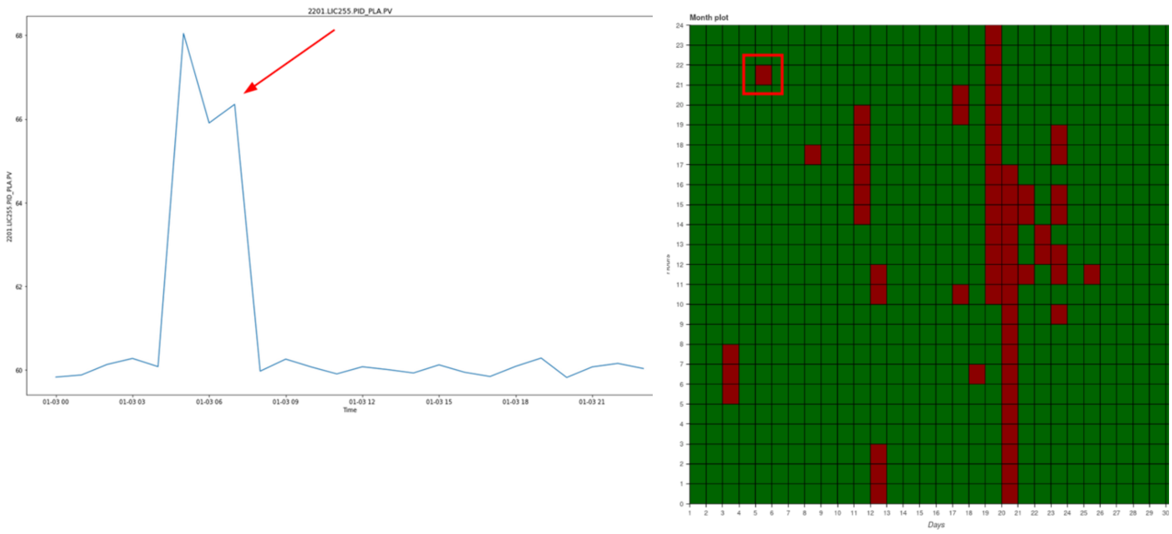

Figure 26: An example of the Root causes: 2201.ANT251.DACA.PV, 2201.FI252.DACA.PV

5.3.3 Task 5.2

In Petro pilot, the roles of the AI tool:

- Evaluation of waste CO2 for alternative fuel production (circular aspect)
- Optimization of steam consumption in EO stripper
- Anomaly detection

The role of AI in the circularity framework in Petro pilot will be to evaluate the production of dimethyl ether (DME), an alternative to diesel from CO2. The AI to be developed will enable the value that can be created by extending the life cycle of CO2 to be realized by using historical process data.

5.3.4 Task 5.3

The EO recovery process of the Petrochemical pilot is composed of 2 main assets, the EO Stripper and the EO Absorber. The real-world assets can be considered as actors within the manufacturing process. In particular, the EO reactor unit acts as a supplier, where EO and by-products of CO2 and H2O are produced and sent to the EO recovery unit after quenching. The EO recovery unit, acting as operator, separates the outlet gases and the EO is recovered by absorbing and stripper operations. The assets of EO purification act as customers; the purification unit further refines the product of the EO recovery unit.

The AI tool will be used to understand process unit failures of EO recovery stripper and absorber, to identify abnormal process conditions, to avoid unplanned shutdowns and successfully manage unpredictable feed and demands. This tool will be developed by using data analytics and AI.

5.3.5 Task 5.4

Digital Twin of the EO recovery unit will be used to optimize the steam consumption of the EO stripper. For this purpose, historical process data will be gathered from site sensors and controllers. Laboratory results will also be used. These historical data will be used to prepare the data-driven Digital Twin and a first-order physical model of the process. Auto-ML will be used to generate models of the process to optimize steam by manipulating the temperature and pressure of the stripper.

5.3.6 Task 5.5

As done for the WEEE pilot, the same questionnaire has been provided also to the PETRO pilot to investigate how to apply the HDT. For this pilot, there are three main types of operators involved:

1. Foreman: The role is to control, coordinate and supervise the operations in the control room, panel and plant site.
2. Process control technician: The role is to ensure the continuity of the process by monitoring the values related to the operation in the control room, taking the necessary actions and to adjust the production parameters according to their set values when required.
3. Field Equipment technician: Carrying out field operations, periodic control, value taking, sampling, product loading activities.

The three figures are usually trained by taking courses related to health&safety and sector-related orientation trainings. New operators start performing their duties under close supervision. To this end, the HDT seems not much relevant from the upskilling perspective.

Concerning critical decisions, process control technicians and foremen take decisions (also consensual) to change parameters values (in a safety range ensuring not to go beyond operational limits) so that to guarantee the continuity of production and production safety. These decisions are not frequent (once a week on average), but they can be supported by the HDT by, for instance, acting as a third judge in suggesting if and when to change parameters (learning from past decisions).

Finally, as per the WEEE pilot, operators are assigned to a limited set of tasks, mainly based on their experience and skills; thus, the HDT can be exploited to suggest job re-allocation based on operator skills.

Additional inputs

The manufacturing process of the Petrochemical pilot is composed of 3 main assets, the EO Reactors, the EO Stripper and the EO Absorber. The real-world assets can be considered as actors within the manufacturing process. In particular, the EO reactor unit acts as a supplier, where EO and byproducts of CO2 and H2O are produced and sent to the EO recovery unit after quenching. The EO recovery unit, acting as operator, separates the outlet gases and the EO is recovered by absorbing and stripper operations. Moreover, CO2 byproduct is removed from the reactor outlet gasses by means of CO2 absorption and a CO2 stripper system. The assets EO purification unit and steam generation unit act as customers; the purification unit further refines the product of the EO recovery unit, whereas the steam generation unit supplies steam to the production plant.

To further simplify the process mapping to better aid in the development of the Use Cases, a further breakdown of the process has been performed. As shown in the figure below, in this mapping, the process is described using 4 components, namely the EO Reactors, the EO Stripper, the EO Absorber and asset with code name “E 208” that contributes to the operation of the EO Stripper.

6. Conclusion

The present deliverable is reporting on the activities undertaken in the context of Task 5.6 ("*System Integration, Testing and Technical Validation*") and it aimed to verify the adherence of the proposed technologies and tools adopted or developed to the requirements collected on data for the circular value chain and to validate other WP5 tasks outcomes with respect to Pilots expectations.

The first issue to be solved was the absence of a reference standard methodology to be applied in the task; for this reason, a Verification and Validation Methodology has been created starting from the approach proposed in AI REGIO[21] due to common AI technologies and partners involved, and it will be pursued for the whole task duration (i.e., also the activities to be carried on during the second phase of the Project have been detailed). The rationale behind such methodology is the need to assess in a holistic way the goodness of the requirements collected, and the technical dimension considering the Pilots needs and constraints.

Main results from active tasks (T5.1 – T5.5) are presented in Section 3 (requirements and conceptual models) and Section 5 (application in the pilots), showing on one hand the functional capacity of the modules to implement various types of scenarios, but also the suitability to efficiently realize the requirements from the pilots. Focus was on WEEE pilot, with the clear plan to transfer the knowledge and experience on other two pilots.

[21] https://www.airegio-project.eu/

7. Appendix A

To investigate the applicability of the HDT to the different pilots, T5.5 defined a questionnaire to be filled out by pilot implementation partners. Here we report the details about the questionnaire. The discussions about the collected answers are provided in sections X, Y, and Z.

The Human Digital Twin (HDT) developed in Circular TwAIn will model workers' skills and expertise and aim to enhance human-machine interaction and workers' engagement in industrial settings. The following questions will help us in identifying relevant workers' characteristics to be modeled, and interactions with the manufacturing environment and processes.

Please reply to the following questions with reference to the production process you are considering for the pilot. In case of multiple processes, duplicate the set of questions and answer each set referring to only one process at a time.

1. How many human operators are involved in the process? Which are their roles? Please provide a detailed description of the operations performed by human workers (manual, cognitive, etc.).
2. How are new operators trained (training material, shadowing, ...)?
3. Please provide a detailed description of the decisions taken by human workers in your pilot processes. Please report also on the frequency of the decisions, and their nature (e.g., is the decision agreed by multiple people?)
4. How do the decisions influence the process/product and/or other workers?
5. Does the worker have an influence on the final product (e.g.: quality)? If so, how?
6. Which kind of data do you collect about human operators (knowledge, expertise, skills, experience, taken decisions, ...)? Which standards do you use (e.g., taxonomies, ontologies, etc.). If there is no data available, how would you collect such data?
7. Are operators assigned to a single task (or a limited set of tasks)? Or do they perform every task in the process?
8. Are tasks assigned to the operator depending on their experience or skills? If not, why (e.g., lack of knowledge about the expertise)?

Co-funded by
the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101058585