AI Platform for Integrated Sustainable and Circular Manufacturing

Deliverable

D4.2 Data Space Implementations for Materials/ Products and Process/Production - 1st version

Actual submission date: 15/09/2023

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
D4.2 Data Space Implementations for Materials/ Products and Process/Production - 1st version			
Work Package:	WP4		
Lead partner:	NETCOMPANY-INTRASOFT SA (INTRA)		
Author(s):	Nikos Kefalakis (INTRA), Christos Stavrogiannis (INTRA), John Soldatos (INTRA), Gabriele De Luca (ENG), Friedrich Volz (IOSB), Luis Usatorre (TECN)		
Reviewers:	Gabriele De Luca (ENG), Sergio Gusmeroli (POLIMI)		
Due date:	30/06/2023		
Deliverable Type:	OTHER	**Dissemination Level:**	PU
Version number:	1.0		

Revision History

Version	Date	Author	Description
0.1	27/04/2023	INTRA	First release of the template
0.2	23/05/2023	INTRA	Added SOTA analysis
0.3	30/05/2023	IOSB	Added EDC, EDC extension, Data Models for Process & Production and DS PoC testbed related contributions
0.4	09/06/2023	INTRA	Added Data Space Architecture
0.5	13/06/2023	TECN	Added Data exchanged types, DS mapping to pilot scenarios
0.6	14/06/2023	ENG	Added TRUE Connector related contributions

0.7	26/06/2023	INTRA	Added EDC deployment information for data catalogues, data exchange and IAM
0.8	17/07/2023	INTRA	Added Data Models for products/materials, Introduction Executive summary, and conclusions.
0.9	24/07/2023	INTRA	Updates throughout the document. Preparation and release for internal review.
0.10a	24/07/2023	ENG	Internal quality review 1
0.10b	25/07/2023	POLIMI	Internal quality review 2
0.11	29/08/2023	TECN	Addressed internal reviewers' comments related to Use Cases
0.12	11/09/2023	IOSB	Addressed internal reviewers' comments related to EDC Extension
0.13	12/09/2023	ENG	Addressed internal reviewers' comments related to TRUE Connector
0.14	14/09/2023	INTRA	Address internal reviewers' comments related to EDC and DPP. Performed updates throughout the document. Preparation of the document for submission.
1.0	15/09/2023	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*
AM	*Additive Manufacturing*
API	*Application Programming Interface*
CA	*Consortium Agreement*
CI	*Collaborative Intelligence*
CIM	*Content Information Management*
CLI	*Command-Line Interface*
CM	*Circular Manufacturing*
CRUD	*Create Update Delete*
CT	*Circular TwAIn*
CTDA	*Circular TwAIn Data Agent*
DAPS	*Dynamic Attributes Provisioning Service*
DDMV	*Data Digital Models & Vocabularies*
DL	*Deep Learning*
DLCP	*Digital Lifecycle Passport*
DNP	*Digital Nameplate*
DoA	*Description of Action*
DPP	*Digital Product Passport*
DPP	*Digital Product Passport*
DS	*Data Space*
DSCS	*Data Space Common Services*
DSP	*Data Space Participant*
DSS	*Decision Support System*
DT	*Digital Twin*
DTDL	*Digital Twin Definition Language*
EC	*European Commission*
EDC	*Eclipse Dataspace Components*
ERP	*Enterprise Resource Planning*
ETSI	*European Telecommunications Standards Institute*
EU	*European Union*
FA³ST	*Fraunhofer Advanced Asset Administration Shell Tools*
GA	*Grant Agreement*
GPU	*Graphic Processing Unit*
GUI	*Graphical User Interface*
HTTPS	*Hyper Text Transfer Protocol Secure*
IAM	*Identity and Access Management*
IDS	*International Data Space*
IDSA	*International Data Spaces Association*
IDTA	*Industrial Digital Twin Association*
IoT	*Internet of Things*
ISG	*Industry Specification Group*
IT	*Information Technology*
KPI	*Key Performance Indicator*
LCA	*Life Cycle Assessment*
ML	*Machine Learning*
MVD	*Minimum Viable Dataspace*
MVDS	*Minimum Viable Data Space*
NGSI-LD	*Next Generation Service Interface – Linked Data*
PDP	*Policy Decision Point*
PEP	*Policy Enforcement Point*
PLM	*Product Lifecycle Management*
RA	*Reference Architecture*
SDK	*Software Development Kit*
SLA	*Service Level Agreement*
TRUE	*TRUsted Engineering*
UML	*Unified Modelling Language*
W3C	*World Wide Web Consortium*

WEEE	*Waste Electrical and Electronic Equipment*
WG	*Working Group*
WP	*Work Package*
XAI	*eXplainable Artificial Intelligence*

Disclaimer

This document has been produced in the context of Circular TwAIn Project. The Circular TwAIn Project is part of the European Community's Horizon Europe Program for research and development and is as such funded by the European Commission. All information in this document is provided 'as is' and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability with respect to this document, which is merely representing the authors' view.

Executive Summary

Circular TwAIn Project aims to lower the barriers to adopt and leverage trusted AI technologies in the manufacturing and process industry's circular value chains. The Project envisions achieving end-to-end sustainability, from eco-friendly product design to the maximum exploitation of production waste across the circular chain. This deliverable aims to develop and implement an innovative data space architecture to enable sustainable by design products and optimize processes in a circular economy. It focuses on the work performed for tasks T4.3 and T4.4, which involve the design and implementation of data spaces for materials/products and process/production lifecycle assessment, respectively. The document presents the existing platforms and building blocks, describes the architecture and core elements, and provides examples of usage scenarios and testbeds.

The primary focus of the document is on the design and description of the Circular TwAIn Data Space Architecture. This architecture is centred around the Circular TwAIn Data Agent (CTDA) which includes the data exchanges and data processing capabilities of the data space participant and are designed to facilitate data sharing and consumption within the data space. The architecture also highlights some common data space services like the Identity & Access Management, Data Digital Models & Vocabularies, and Application Discovery. Lastly, the User Dashboards & Services layer provides user-friendly interfaces to interact with the data space.

The document also presents the CT Data Space Core Elements and their implementations. This includes the design of specific data models for materials and products, as well as for processes and production lifecycle assessment. The deliverable details the functionalities and features of Data Catalogues, Publication and Discovery, and Data Exchange, emphasizing the need for data sovereignty. It also explains the role of Identity and Access Management in ensuring secure data sharing, provides the dataspace participant onboarding process for different implementations and highlights the importance of Applications & Data Orchestration in enabling efficient data processing.

The document also details the criteria and rationale behind the technology choices made for the project. Moreover, it defines the core components deemed essential for the basic functionality and successful running of the Circular TwAIn system. The document introduces the CT DS testbed environment concept along with the required supported tools. The document also provides information on the versioning repository and its governance for hosting infrastructure deployment/configurations, extensions, new developments and documentation which are provided under the CT GitHub organization. The CT GitHub organization will be constantly maintained and updated throughout the Project duration.

The final part of the deliverable provides practical use-case scenarios and testbeds for the data space, showing how the architecture and its core elements can be applied to real-world circular economy scenarios. This includes the mapping of the data spaces to pilot scenarios and an overview of the data space PoC testbed.

Overall, the deliverable provides a comprehensive overview of the work performed under tasks T4.3 and T4.4 for the Circular TwAIn Project, demonstrating the potential of the data space architecture in supporting circular economy use-cases and the exchange/modelling of product, material, process, and production data.

1 Introduction

1.1 Scope and Objective of this Deliverable

The scope of this deliverable extends to the execution of tasks T4.3 and T4.4 of the Circular TwAIn Project. The document is intended to elaborate on the design, implementation, and utilisation of a data space that facilitates the sharing and utilisation of data relating to materials, products, and process/production lifecycle assessments, thereby enhancing circularity and sustainability in these domains. The objective is to provide insight into the tasks' progress and results, explaining the intricate details of the Circular TwAIn Data Space Architecture and how it aids the Circular TwAIn Project's broader goals. The deliverable aims to illustrate the complex mechanisms behind the data models, catalogues, and exchange, while maintaining an understanding of the human-centric and sustainable approach the Project promotes.

1.2 Deliverable Structure

The rest of the deliverable is structured as follows:

- Section 2 provides an overview of existing data spaces and building blocks, including platforms and implementation of data space building blocks.
- Section 3 provides a detailed description of the Circular TwAIn Data Space architecture.
- Section 4 provides an elaboration on the core elements and their implementations within the Circular TwAIn data space, including data models for materials and products and for process/production, data catalogues, publication and discovery, data exchange, identity and access management, and applications & data orchestration.
- Section 5 provides usage scenarios and testbeds for the data space, including mapping of data spaces to pilot scenarios and a data space proof-of-concept testbed.
- Section 6 provides the conclusion of the deliverable.

2 Existing Data Spaces & Building Blocks

2.1 Minimum Viable Dataspace definition

The Minimum Viable Data Space (MVDS) is an amalgamation of essential components required to establish a data space that enables secure and sovereign data exchange. It is designed to expedite the work of developers and experimenters by reducing implementation time and avoiding complex details that might delay the initial release. The MVDS allows for a functioning initial version that guarantees secure and sovereign data exchange. It serves as a foundation for the development team to iterate, identify, and respond to the requirements of the data space, based on their assumptions. In Circular TwAIn we have considered the International Data Space IDS and Eclipse Dataspace Components (EDC) ones.

2.2 IDS MVDS

The International Data Spaces Association (IDSA) offers a unique solution known as the IDS Minimum Viable Data Space (MVDS) [28], considered the current best practice (see Figure 1 below). The IDS MVDS comprises the following key components:

- Two or more IDS Connectors: These serve as the data sources and sinks, allowing for the exchange of data between different parties within the data space.
- Certificate Authority (CA)[1]: This entity provides X.509 certificates, which are ensuring the authenticity and integrity of data exchanges. It's important to note that this doesn't involve the process of certification.
- Dynamic Attributes Provisioning Service (DAPS)[2]: This is responsible for managing dynamic attributes and handling dynamic access tokens, for access control and data security within the data space.

A list of available open-source components on IDS GitHub[3].

[1] https://github.com/International-Data-Spaces-Association/IDS-testbed/blob/master/InstallationGuide.md#certificate-authority
[2] https://github.com/International-Data-Spaces-Association/IDS-testbed/blob/master/InstallationGuide.md#daps
[3] https://github.com/International-Data-Spaces-Association/idsa/blob/main/overview_repositories.md

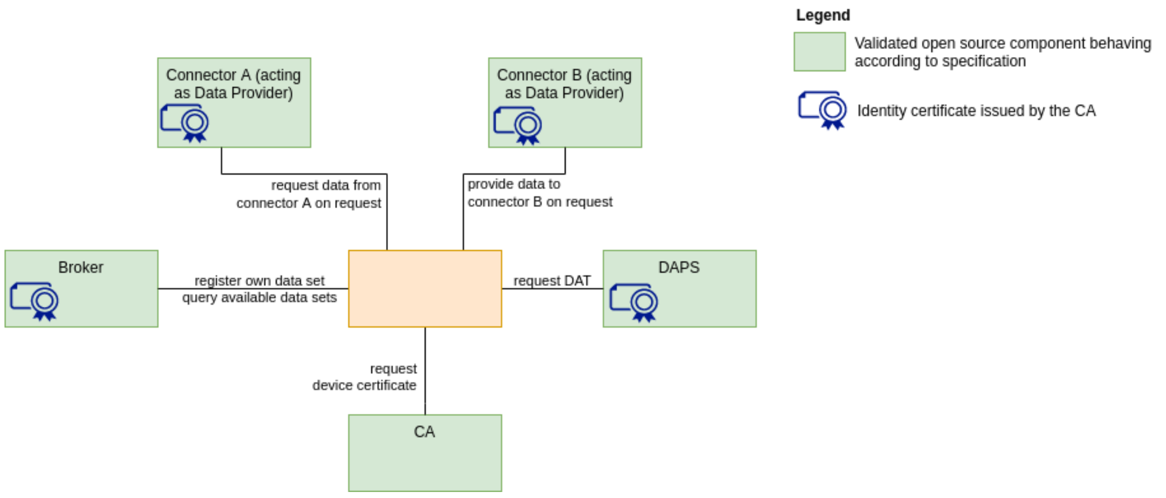

Figure 1: Minimal IDS MVDS setup with essential and already available components [28]

More documentation on the IDS MVDS can be found under the IDS Knowledge base[28] and for the IDS-testbed on IDS GitHub[4].

Table 1: IDS MVD Platform Information

Maturity Level	TRL 7
License Type	Apache-2.0 license

2.3 TRUE Connector MVDS

Engineering MVDS [3], a customization of IDS MVDS, consisted of the following components:

- Two TRUE Connectors [4]: Those 2 connectors serve as the data consumer and provider, allowing for the exchange of data between different parties within the data space.
- Certificate Authority (CA): This entity offers X.509 certificates that guarantee the integrity and validity of data transfers.
- Dynamic Attributes Provisioning Service (DAPS): Component responsible for managing dynamic attributes and handling dynamic access tokens, for access control and data security within the data space.

[4] https://github.com/International-Data-Spaces-Association/IDS-testbed/blob/master/README.md

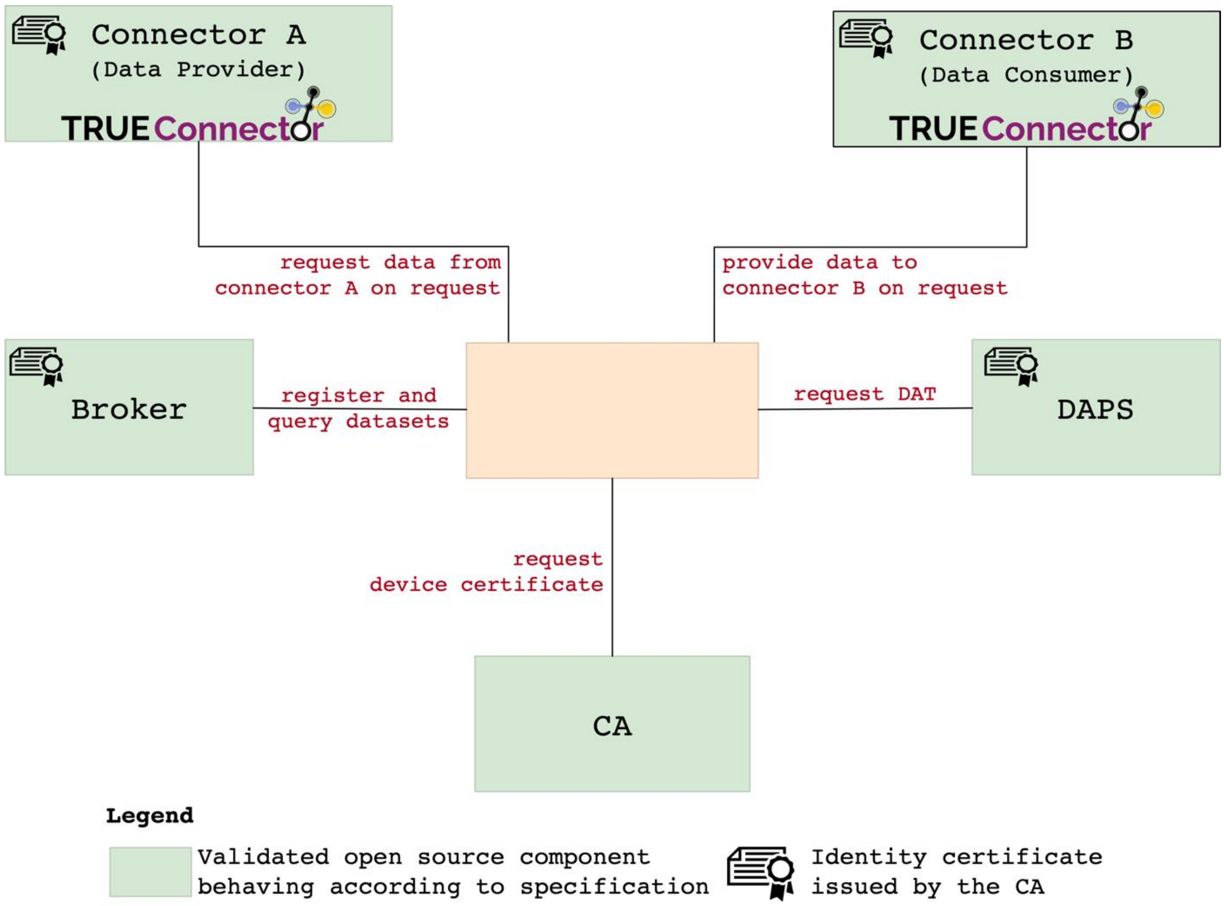

Figure 2: TRUE Connector MVDS

For more detailed overview of the proposed customization please refer to the following[5]

Table 2: TRUE Connector MVDS Platform Information

Maturity Level	TRL 7
License Type	AGPL 3.0 license

2.4 EDC MVD

The Eclipse Minimum Viable Dataspace (MVD) is a demonstration model of a dataspace that utilizes the Eclipse Dataspace Components (EDC). It aims to showcase the EDC's abilities and offers a tangible example of how a dataspace can be implemented. Additionally, EDC MVD provides a practical example of how decentralization can be implemented within a dataspace, since a fully decentralized dataspace can be challenging to conceptualize. Information about Eclipse MVD can be found at Eclipse-EDC GitHub[6].

The initial setup of the EDC MVD includes the following components:

[5] https://github.com/Engineering-Research-and-Development/true-connector-mvds
[6] https://github.com/eclipse-edc/MinimumViableDataspace

- Three EDC Connectors for three different companies (company1, company2, and company3) which also runs the EDC data dashboard (see Figure 3 below).
- A Registration Service, for managing the access and identity of the participants within the data space.
- An HTTP Nginx Server, employed to serve DID (Decentralized Identifier) Documents, for identity verification in decentralized networks.
- An Azurite blob storage service, which is a local emulation environment for Azure's blob storage, used for storing unstructured data in the cloud.

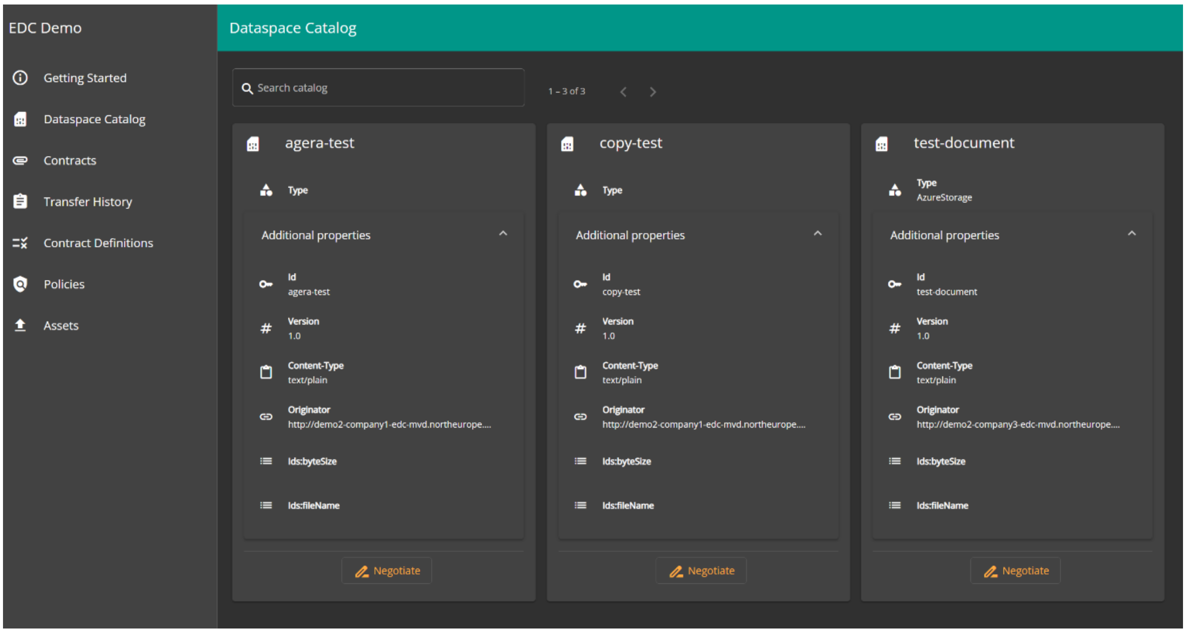

Figure 3: Viewing dataspace catalogue in EDC data dashboard[7].

Table 3: Platform Information

Maturity Level	TRL 7
License Type	Apache-2.0 license

[7] https://github.com/eclipse-edc/MinimumViableDataspace/blob/main/docs/developer/edc-data-dashboard/view-catalog.md

3 CT Data Space Architecture with DTs & XAI

Figure 4 below depicts an instantiation of the Circular TwAIn Functional Architecture diagram for the data space domain that includes DT and XAI elements. This comprehensive functional component diagram architecture supports the exchange, modelling, and processing of product, material, process, and production data within a circular economy data space. It facilitates data publishing, discovery, and exchange among untrusted actors, promoting collaboration and innovation in circular economy practices. The architecture ensures secure data sharing, interoperability, and transparency, empowering stakeholders to make informed decisions and drive sustainable circular economy initiatives.

For designing the functional component diagram, presented in Figure 4, various inputs and aspects have been taken into consideration which includes:

- The updated, based on the different tasks feedback and Project evolution, Circular TwAIn RA which is going to be reported in deliverable D3.4 “Conceptual Framework and Reference Architecture - 2nd version”.
- The DS building blocks reported in deliverable D3.2 “Data Space and Digital TwAIns Design - 1st version”.
- The data pre-processing and analytics toolkit architectures and offered functionalities reported in deliverable D5.1 “Circular TwAIn Data4AI Platform and AI Toolkit - 1st version”.
- The Circular TwAIn DT architectures reported in “D5.2 AI-enhanced Digital Twins Implementations for Products Production and Personae - 1st version”.
- The ontologies and industrial data platform architecture reported in deliverable “D4.1 Circular TwAIn Industrial Data Platform, Standards Ontologies - 1st version”.
- Finally other reference architectures, standards and background work, which have also been presented in the above mentioned deliverables, like the IDSA RA [22], OPEN DEI Building Blocks [21] and Simpl’s Architecture preparatory work [20].

The presented architecture view covers all the different aspects of Circular TwAIn (i.e., Data Spaces, Digital Twins, XAI & Algorithms, and data preparation) the components have been colour coded to identify their role. Although Figure 4 architecture view highlights the data space aspect contributing to the exchange and modelling of product, material, process, and production data it clearly specifies the relationships between the XAI, DTs and data preparation components. The architecture supports a modular approach and as such can be either used as a complete full-fledged solution integrating all the different Circular TwAIn aspects or by highlighting individual aspects as presented in deliverable D5.2 as part of the Product and Material DT.

The Figure 4 functional component diagram architecture is consisted of the following layers:

- Data Space Participant (DSP) which represents a group of components and systems applied to a Data Space Actor.
- Data Space Common Services (DSCS) which provides essential services and functionalities that are shared across the data space environment.

- User Dashboards & Services which provides various dashboards and services for users to interact with the circular economy data space.

In the sections below a detailed analysis of the DS functional component diagram can be found.

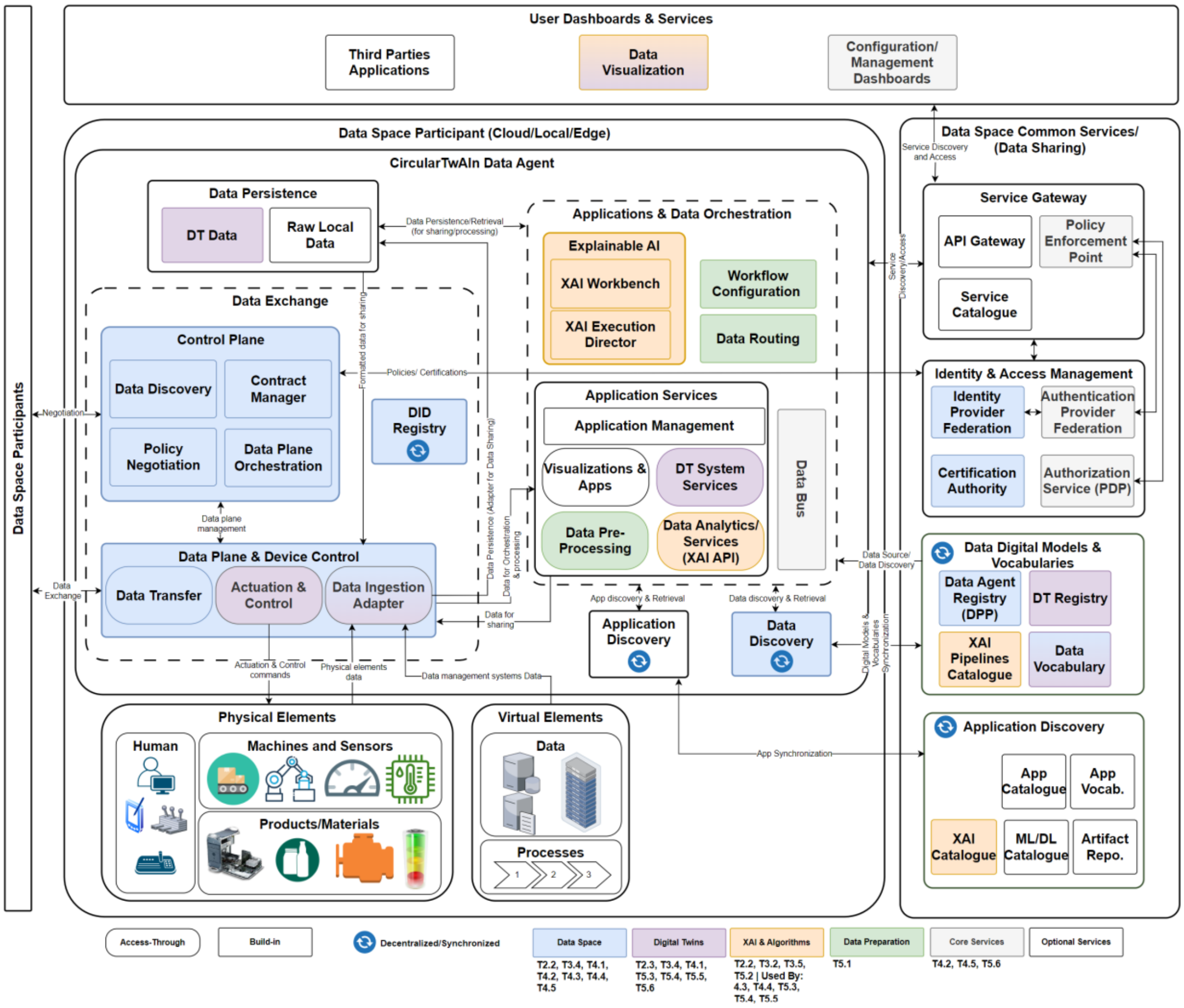

Figure 4: Circular TwAIn Data Space Functional Component Diagram

3.1 Data Space Participant (DSP)

The DSP (see Figure 4 above) represents a group of components and systems applied to a Data Space Actor. It can be hosted in the Cloud or at the Edge, enabling data exchange and collaboration among untrusted actors within the data space environment. By integrating the DSP into the architecture, participants within the data space can seamlessly collaborate, share data, and make informed decisions based on the exchanged information.

3.1.1 *Circular TwAIn Data Agent (CTDA)*

The CTDA is facilitating the exchange and processing of product, material, process, and production data within Circular TwAIn ecosystem. It consists of various sub-components and functionalities that enable seamless data management and collaboration within the circular economy context. These components include:

- Data Exchange: The Data Exchange component group (see Figure 4 above) is enabling the secure and efficient exchange of product, material, process, and production data. It facilitates data discovery, negotiation, transfer, and ingestion, while adhering to established policies and contracts.

- **Control Plane:** Responsible for managing the policies and contracts for data sharing, data discovery, and agreements for data exchange within the circular economy data space. It ensures secure and governed data exchange among participants.
  - **Data Discovery:** Enables decentralized data discovery, allowing data space participants to find relevant product, material, process, and production data for circular economy use cases. It promotes the discovery of available data sources for modelling and analysis.
  - **Contract Manager:** Manages data sharing agreements and contracts, ensuring compliance with defined policies, permissions, and privacy requirements for circular economy data exchange.
  - **Policy Negotiation:** Facilitates negotiation and agreement on data usage policies, promoting fair and transparent data exchange practices within the circular economy context.
  - **Data Plane Orchestration:** Orchestrates the data plane components for efficient data transfer, actuation, control, and ingestion. It optimizes data flow and processing within the circular economy data space.
- **Data Plane**: Handles the actual data transfer, actuation, control, and data ingestion.
  - **Data Transfer**: Facilitates data exchange among different data space participants, enabling the flow of product, material, process, and production data. It supports the exchange of information related to circular economy practices, such as material composition, production parameters, and environmental impact.
  - **Actuation & Control**: Allows data space participants to control physical elements, such as machines and sensors, in response to circular economy data insights. It enables the implementation of circular economy strategies, such as optimizing resource usage and reducing waste.
  - **Data Ingestion**: Facilitates the ingestion of data from physical and virtual elements, such as databases and sensors, into the circular economy data space. It ensures seamless integration of data from diverse sources for comprehensive analysis and modelling.
- **DID Registry:** The DID (Decentralized Identifier) Registry component operates in a decentralized manner, providing a unique identification mechanism for participants, data sources, and entities within the circular economy data space. It ensures the integrity and authenticity of digital identities associated with different actors.

- **Data Persistence**: Manages storage for "Raw Local Data" and "DT Data" storing and managing both raw local data and data in the form of digital twins (DT data). It ensures the availability of historical data for analysis, auditing, and decision-making purposes. This component supports the long-term storage and retrieval of product, material, process, and production data.

- **Application & Data Orchestratio**ns: The Applications & Data Orchestration group within the CTDA (see Figure 4 above) includes several sub-components and functionalities to support the seamless integration and management of applications and data. This group comprises components such as Explainable AI, Workflow Configuration, Data Routing, Application Services, and Data Bus. These components enable the execution of various tasks such as data pre-processing, data analytics, and visualization of data and applications. In more details the Applications & Data Orchestration is comprised from the following components:
    - **Workflow Configuration:** Configures and manages workflows for processing and analysing product, material, process, and production data. It enables the creation of automated data processing pipelines for circular economy use cases.
    - **Data Routing:** Routes data to appropriate destinations within the CTDA ensuring seamless data flow and interoperability. It directs data to specific applications, analytics modules, or storage systems for further processing or analysis.
    - **Data Bus:** Facilitates the exchange of data among different components and services within the circular economy data space. It provides a centralized communication channel for data sharing and collaboration, enabling real-time data updates and synchronization.
    - **Application Services:** encompass various applications that can be utilized within the circular economy context, including visualizations, digital twin system services, data pre-processing, and data analytics (both local and remote) etc.
        - **Application Management:** Manages various applications, including visualizations, DT system services, data pre-processing, and local/remote data analytics. These applications support circular economy decision-making, optimization, and resource management. They provide tools for data exploration, modelling, simulation, and scenario analysis in a circular economy context.
        - **Visualizations and Applications:** Enables the visualization and interpretation of product, material, process, and production data, providing insights into circular economy performance, resource usage, and environmental impact. It supports data-driven decision-making and promotes transparency in circular economy practices.
        - **DT System Services:** Provides the appropriate interfaces to interact with a complete DT system, allowing users to access and manipulate digital twins of products, materials, processes, and production systems. It supports the modelling and simulation of circular economy scenarios, facilitating experimentation and optimization.

        - **Data Pre-Processing:** Performs data cleansing, transformation, and normalization tasks to ensure data quality and consistency within the circular economy data space. It prepares data for further analysis and modelling, enhancing the accuracy and reliability of circular economy insights.
        - **Data Analytics** (Local and Remote Services): Enables advanced data analytics capabilities within the circular economy data space. It includes local and remote services that leverage machine learning, artificial intelligence, and statistical techniques to extract valuable insights from product, material, process, and production data. These insights can inform circular economy strategies, such as waste reduction, resource optimization, and sustainable product design.
    - **Explainable AI (XAI)**: The Explainable AI components group utilizes the XAI Pipelines Catalogue and XAI Catalogue from the Data Space Common Services vertical layer to facilitate the explainability and interpretability of AI algorithms applied to the circular economy data. The XAI group is consisted of:
        - **XAI Workbench:** Provides a user-friendly environment for developing, testing, and deploying explainable AI pipelines.
        - **XAI Execution Director:** Executes the deployed XAI pipelines, generating explainable outputs and visualizations. It facilitates the communication of AI-driven circular economy insights to stakeholders, promoting trust and understanding of the decision-making process.
- **Application Discovery and Data Discovery:** The Application Discovery component (see Figure 4 above) runs a local copy of the application discovery system, working in a decentralized fashion. It enables participants to discover and access relevant applications for data analysis, optimization, and decision-making in the circular economy context. Similarly, the Data Discovery component runs a local copy of the data discovery system, facilitating the decentralized discovery and access of relevant data sources within the circular economy domain.

3.2 Data Space Common Services

The Data Space Common Services component (see Figure 4 above) provides crucial functionalities for the overall management, security, interoperability, and standardization of the circular economy data space. It ensures secure access, identity management, standardized data models, and the discovery and reuse of applications and AI models. These services collectively support the seamless exchange, integration, and analysis of product, material, process, and production data within the circular economy context.

3.2.1 Service Gateway

The Service Gateway group (see Figure 4 above) ensures a unified and secure approach to accessing services and functionalities. It serves as an entry point for the data space participants and external actors to interact with the circular economy data space. It provides a standardized API interface and enforcing access control policies, it enables seamless interaction and collaboration among data space participants and external actors, fostering a well-connected and efficient circular economy environment. The Service Gateway group includes the following key components:

- **API Gateway:** The API Gateway is a central entry point that provides a unified interface for accessing the functionalities and services available within the circular economy data space. It manages and exposes APIs that allow data space participants to interact with different components and services seamlessly. The API Gateway also handles API request routing, load balancing, and caching to optimize performance and reliability.
- **Policy Enforcement Point (PEP):** The Policy Enforcement Point is responsible for enforcing the access control policies defined for the circular economy data space. It acts as a gatekeeper, evaluating access requests and determining whether they comply with the defined policies and rules. The PEP ensures that only authorized data space participants and external actors can access specific resources and functionalities. It enforces policies related to data sharing, data discovery, and other data space interactions, enhancing data governance and security.
- **Service Catalogue:** The Service Catalogue is a repository that maintains a comprehensive list of services and functionalities available within the circular economy data space. It provides a centralized overview of the capabilities and offerings, making it easier for participants to discover and access the services they require. The Service Catalogue includes detailed information about each service, such as its purpose, API specifications, input/output data formats, and access requirements. It facilitates efficient service discovery and integration within the overall data space ecosystem.

3.2.2 Identity & Access Management

The Identity & Access Management (IAM) group (see Figure 4 above) is responsible for managing user identities, providing secure authentication mechanisms, and enforcing fine-grained access control. This way it creates a trusted environment where data space participants can confidently collaborate, share data, and engage in circular economy initiatives while protecting sensitive information and maintaining data integrity. The IAM group includes the following key components:

- **Identity Provider:** The Identity Provider is responsible for authenticating and verifying the identities of data space participants and external actors. It acts as a central authority that issues and manages identity credentials, such as usernames and passwords, or digital certificates, to enable secure access to the data space. Through the Identity Provider, participants can securely log in to the data space and assert their identity, ensuring that only authenticated users are granted access to the system.
- **Authentication Provider Federation:** The Authentication Provider Federation is responsible for establishing trust relationships between different identity providers across multiple domains. It enables users from various organizations or external systems to use their existing identities to access the circular economy data space.

- **Certification Authority:** The Certification Authority is responsible for issuing and managing digital certificates used for secure communication and data exchange within the circular economy data space. Digital certificates serve as cryptographic credentials that verify the authenticity and integrity of participants' data and communications. By leveraging digital certificates, the data space can ensure secure data transmission, data integrity, and non-repudiation, providing a robust foundation for data trustworthiness.
- **Authorization Service (PDP):** The Authorization Service, also known as the Policy Decision Point (PDP), is responsible for enforcing access control policies based on the authentication and authorization decisions made by the Identity Provider and other security components. The PDP evaluates access requests against predefined policies to determine whether a user or actor is authorized to perform specific actions or access certain resources within the data space.

3.2.3 Data Digital Models & Vocabularies

The Data Digital Models & Vocabularies (DDMV) group (see Figure 4 above) is responsible for managing and standardizing the representation of data, digital models, and vocabularies used within the data space. This group ensures semantic interoperability, enabling seamless data exchange and understanding among different participants and systems. The main components of the DDMV include:

- **DT Registry:** is a centralized or decentralized repository that stores and manages Digital Twin (DT) models and instances within the circular economy data space. It provides a repository for digital twin descriptions, enabling stakeholders implementing CTDA to access and interact with digital twins of products, materials, processes, and production systems.
- **Data Vocabulary:** Defines a common vocabulary and ontology for describing product, material, process, and production data within the circular economy data space. It ensures semantic interoperability and facilitates data integration and exchange among different stakeholders and especially thru the Data Exchange of CTDA. By adhering to a shared data vocabulary, data space actors use consistent and agreed-upon terminology when exchanging data, reducing ambiguity and facilitating seamless data interpretation and processing.
- **Data Agent & DPP** (Digital Product Passport) **Registry**: The Data Agent and DPP Registry provides a centralized or decentralized repository for storing and managing Data Agent and Digital Product Passport (DPP) information. DPPs serve as comprehensive digital records that contain relevant data about products, materials, or assets throughout their lifecycle. So, this component enables the registration, discovery, and access to DPPs, allowing data space participants to access essential information about the provenance, composition, and sustainability aspects of products and materials within the circular economy.
- **XAI Pipelines Catalogue**: Maintains discoverable pipelines for explainable AI (XAI) within the circular economy data space. It allows stakeholders to find and reuse XAI pipelines for interpreting and understanding circular economy data insights. The XAI Pipelines Catalogue is mainly used from the Explainable AI component of the CTDA.

3.2.4 Application Discovery

The Application Discovery group (see Figure 4 above) enhances the data space's usability by providing a centralized or decentralized location for discovering and accessing diverse applications and services. By maintaining a standardized vocabulary and utilizing the Artifact Repository and ML/DL Catalogue, it ensures seamless application integration and deployment across the data space. Data space participants can leverage these discovered applications and services to process, analyse, and visualize data, enabling informed decision-making and promoting circular economy initiatives. The main components of the Application Discovery group include:

- **App Catalogue**: serves as a centralized or decentralized repository that stores information about available applications and services within the data space. It enables stakeholders to discover and utilize applications relevant to their circular economy use cases, fostering innovation and collaboration. The catalogue is used by the Application Services of the CTDA to discover and access available applications.
- **Artifact Repository:** The Artifact Repository is a storage system that houses artifacts such as software components, libraries, and dependencies required for the deployment and execution of applications within the data space. The App Catalogue interacts with the Artifact Repository to access and retrieve artifacts related to applications and services. This interaction ensures that the applications can be properly deployed and executed by data space participants, promoting seamless application integration and usage.
- **Application Vocabulary:** The Application Vocabulary component defines and maintains a standardized set of terms and concepts used to describe applications and services within the data space. It establishes a common language to ensure consistency in describing different tools and functionalities. Data space participants may use the Application Vocabulary when discovering and selecting applications or providing new Applications from and to the App Catalogue. This interaction ensures that participants can understand the capabilities and functionalities of each application, enabling informed decisions on which tools to utilize for specific circular economy tasks.
- **ML/DL Catalogue:** Maintains a catalogue of machine learning (ML) and deep learning (DL) models within the circular economy data space. It enables stakeholders to discover, evaluate, and utilize ML/DL models for data analysis, prediction, and optimization in circular economy use cases.
- **XAI Catalogue:** Maintains a catalogue of explainable AI (XAI) algorithms within the circular economy data space. It allows the Explainable AI component and Data Analytics Services of CTDA to discover, evaluate, and utilize XAI algorithms for generating interpretable circular economy data insights. The XAI Catalogue and App Catalogue provide common functionality and can be merged if required.

3.3 User Dashboards & Services

The User Dashboards & Services component (see Figure 4 above) serves as an essential interface for users to access and interact with the circular economy data space. It empowers users to explore, visualize, and analyse data, make informed decisions, and effectively manage the data space environment to drive sustainable circular economy practices. The main components of the User Dashboards & Services group include:

- **Third-Party Applications Dashboards:** This dashboard enables users to access and interact with third-party applications that complement circular economy use cases. Users can leverage specialized tools and services from external providers to enhance their circular economy initiatives. These applications may offer advanced analytics, optimization algorithms, or domain-specific features that contribute to sustainable practices.
- **Data Visualization Dashboards**: The data visualization dashboard provides visual representations of product, material, process, and production data within the circular economy data space. Users can explore and analyse data using charts, graphs, and interactive visualizations. This facilitates the understanding of circular economy performance, resource utilization, and environmental impact. Visualizations aid in data-driven decision-making, allowing users to identify opportunities for improvement and monitor progress towards circular economy goals.
- **Configuration/Management Dashboards:** The configuration/management dashboard offers a user-friendly interface for users to configure and manage the circular economy data space environment. It provides controls for access management, policy settings, system configurations, and monitoring functionalities. Users can customize the data space according to their specific requirements, ensuring secure and efficient data exchange, data governance, and overall administration of the circular economy data space.

4 CT Data Space Core Elements and Implementations

In this section we provide a list of core elements of the CT data space architecture. These elements are focusing on delivering a functional version of the data space with essential features and capabilities. While it may not include all the advanced features of the full architecture, the primary goal is to demonstrate the core principles of data exchange, data persistence, data processing, and data discovery in a circular economy context.

4.1 Data Models

In a Circular Data Space (DS) on manufacturing, two type of data can be exchanged, in two different formats and in two different ways.

On the one hand, data can be related to the a-material (either raw material, components or products) or to the b-manufacturing process (temperatures, pressures, times…)

On the other, data can be a-static -data at rest- or b-dynamic -data in motion-. Usually, data for material is at rest, as it does not change between processes. But process data can be either at rest (nominal values) or in motion (real values).

At last, data can be transferred a-periodically or b-on real time (streaming). Static data are usually transferred periodically while dynamic data can be either transferred periodically (on data batches) or on streaming (real time).

In D3.1 two different types of formats/ontologies is proposed for static data sharing within the circular DS on manufacturing:

- Asset Administration Shell (AAS)
- Digital Product Passport (DPP)

AAS files are related with the manufacturing (or de-manufacturing) process, while DPP is related with the material (or waste material) characteristics and properties. Depending on the requirements one type of file or the other will be transferred. In both cases, this type of files can be transferred through IDS RA so the trust and confidentiality of data is guaranteed.

For dynamic data sharing (real time, streaming), and for the time being, there is not a solution that guarantees trust and confidentiality. Connectors are a time bottleneck on this type of data transaction and need to be further developed.

4.1.1 For Materials and products

In Circular TwAIn we have adopted Digital Twins which are primarily used as an internal format and systems for data space participants. They enable the participants of the circular environment to perform internally asset optimization, process optimization, product/material tracking, sustainability assessment and product/material lifecycle management. More information on the DTs usage can be found in deliverable D5.2 “AI-enhanced Digital Twins Implementations for Products Production and Personae - 1st version”.

For data discovery and sharing within the data space, and the concept related to the circular economy, we need to offer comprehensive information and traceability about a physical product. This would include its design specifications, materials used, manufacturing processes, environmental impact, usage instructions, maintenance guidelines, and end-of-life options throughout the lifecycle of a physical product to support circular economy practices. This is partially missing from the DTs format and systems which is covered from other systems/standards like the DPP (Digital Product Passport). Below you can find a list of elements that are missing from DTs, which could be introduced as part of a combined solution:

- **Comprehensive Product Information:** While Digital Twins focus on real-time data and simulations of physical assets, Digital Product Passports provide a holistic record of a product's entire lifecycle. This includes design specifications, materials, manufacturing processes, environmental impact, and end-of-life options. Introducing a DPP element can enhance the solution's ability to provide comprehensive product information, allowing stakeholders to make informed decisions related to sustainability, compliance, and circular economy practices.
- **Circular Economy Metrics:** DPPs often include key metrics related to the circular economy, such as product recyclability, reuse potential, and the environmental impact of materials used. By incorporating circular economy metrics into the solution, businesses can assess the sustainability of their products and processes, identify areas for improvement, and support circular economy goals.
- **Supply Chain Transparency:** DPPs facilitate supply chain transparency by providing information about the origin of materials, suppliers' practices, and transportation details. By integrating supply chain transparency into the existing DTs solution, organizations can enhance traceability, ensure ethical sourcing, and promote responsible manufacturing practices.
- **Consumer Engagement:** Digital Product Passports can also include consumer-oriented features, such as QR codes or NFC/RFID tags on products, allowing consumers to access relevant product information and environmental impact data. Introducing these engagement features in the DTs solution can promote consumer awareness and drive sustainable purchasing decisions.
- **End-of-Life Management:** Digital Product Passports provide guidance on proper end-of-life management options, such as recycling, repurposing, or disposal. Integrating end-of-life management features into the DTs solution can support circular economy practices by encouraging responsible product disposal and resource recovery.
- **Granular Information Control:** Finally, a DPP allows for granular control over the information shared within the data space. Stakeholders can define specific attributes, parameters, or data points that are discoverable and accessible by others. This level of granularity enables businesses to protect sensitive information while still providing essential details relevant to the asset or product's performance. Granular information control enables the data space participant to apply:
    - Confidential Data Protection: Some data within the DPP might be classified as confidential or proprietary, and businesses may want to limit its exposure to a select group of authorized users.
    - Intellectual Property Management: Intellectual property (IP) protection is a critical consideration in industrial settings. A DPP allows businesses to safeguard proprietary design specifications, manufacturing processes, and other IP-related information.

- Compliance with Regulations: Some industries are subject to strict regulatory requirements, and data sharing must adhere to specific guidelines. A DPP enables businesses to align with regulatory standards by implementing data space sharing policies and ensuring that sensitive information is not exposed beyond necessary boundaries.
- Selective Data Sharing: Digital Twins may provide real-time data that is crucial for monitoring and optimizing industrial processes. By coupling DTs with a DPP, businesses can selectively share specific, non-confidential data points to a data space while keeping sensitive information restricted to authorized personnel within the company.
- Data Space Governance: DPPs offer a structured and controlled approach to data management within a data space. By integrating DPPs with DTs, businesses can establish data governance policies, define data ownership, and enforce compliance, ensuring data integrity and security across the entire lifecycle.

The aggregation and the exchange of product information among companies, users and other actors is seen as a key part for the creation of a circular economy. However, there is currently neither a common understanding of which information is required in a DPP, nor which technical solution should be used to operate the DPP.

DPP Data Requirements

Chapter III of the proposed Ecodesign for Sustainable Products Regulation [25] outlines the general and technical requirements for creating, accessing, and sharing digital product passports [26]. The general requirements for digital product passports, which will also be taken into consideration for Circular TwAIn, include complying with the following conditions:

- It shall be connected through a data carrier to a unique product identifier.
- The data carrier shall be physically present on the product, its packaging or on documentation accompanying the product.
- The data carrier and the unique product identifier shall comply with standard (‘ISO/IEC’) 15459:2015.
- All information included in the product passport shall be based on open standards, developed with an interoperable format and shall be machine-readable, structured, and searchable.
- The information included in the product passport shall refer to the product model, batch, or item.
- The access to information included in the product passport shall be regulated in accordance with the essential requirements set out in Article 10 and the specific access rights at the product group level shall be identified in the applicable delegated act adopted pursuant to Article 4.

<u>DPP & Batteries Criteria:</u>

From 2026, every industrial and electric vehicle battery must come with a digital product passport. The required information links to safety requirements and the targets for recycled content in batteries [26]. Based on the battery requirements already identified in Circular TwAIn we will consider providing the following information for data exchange thru the dataspace's environment:

- Material sourcing
- Carbon footprint
- Percentages of recycled materials used.
- Battery durability
- Repurposing and recycling guidelines [27]

DPP Implementation

The aggregation and the exchange of product information among companies, users and other actors is seen as a key part for the creation of a circular economy. However, there is currently neither a clear common understanding of which information is required in a DPP, nor which technical solution should be used to operate the DPP [23]. As mentioned in deliverable D3.2 "Data Space and Digital TwAIns Design - 1st version" in Circular TwAIn we consider following the DPP4.0 initiative[8], which proposes to use the AAS standard to document and make available relevant product information. The idea is to combine two standards: the digital nameplate (DNP4.0) via IEC 61406 (Identification Link) and the Asset Administration Shell (AAS) according to IEC 63278. This will enable the use of the standardized AAS model and its sub-models to represent the product-relevant data and the standardized AAS API to provide and/or consume it.

Potential for Value-Added Services based on DPP4.0

The implementation of DPP4.0 for products (e.g., batteries) within a data space can provide access to valuable information, enabling various value-added services. The DPP allows standardized APIs and accurate data semantics, facilitating scalable services for batteries [23]. Potential services derived from this approach include:

- Carbon Footprint Calculation Service: The DPP calculates the carbon footprint of batteries by considering their bill of materials and individual component life cycles. This service can offer rapid feedback to developers during product design, aiding in the selection of sustainable battery options for comparable assemblies.
- Product Data as a Service: Manufacturers can offer detailed product data beyond legal requirements, including additional information on primary battery components.
- DPP4.0 Viewer Service: Consumers and users can visualize DPP4.0 contents in a user-friendly manner, enhancing transparency and understanding. This can

[8] https://dpp40.eu/

positively influence behaviour when purchasing or recycling batteries, tailored to specific use cases.

- Bill of Material Service: The DPP allows easy addition of battery bill of materials and documentation of component changes. Engineering, production, or service employees can keep the DPP up to date throughout the battery's lifetime. Component removal information can also highlight recycling or repair possibilities.
- DPP4.0 DS Repository: DPP4.0 decentralized or centralized repository service enables providers to store and share battery product data in a decentralized or centralized manner depending on the DS implementation and the provider’s preferences.
- Other Potential Services: The AAS integration allows the inclusion of sub-models for disassembly, service notifications, or spare part services. Considering the entire battery lifecycle (cradle-to-grave), IIoT data can be used for monitoring or anomaly detection services. The expansion of AI applications will be also explored.

DT (AAS) and DPP (DPP4.0) usage in a Circular DS environment.

Below we provide a circular economy dataspace scenario using the recycling of e-mobility Li-Ion battery packs to illustrate the distinctions between the usage of DT (using AAS) and DPP (using DPP4.0) in a circular economy scenario.

Digital Product Passport (DPP), specifically DPP4.0 in this scenario, are utilized to document and make available relevant product information, in this case, information related to car batteries. DPP4.0 contains detailed data about the product, including material sourcing, carbon footprint, percentages of recycled materials used, battery durability, repurposing, and recycling guidelines. DPP4.0 is using the Asset Administration Shell (AAS) standard for documenting and providing this product information and by combining the digital nameplate (DNP4.0) it identifies the AAS representation for providing standardized product data.

Digital Twins (DTs) implemented using the AAS standard, serve as digital representations of physical entities within the circular economy dataspace, including car batteries. DTs contain a broader scope of information about physical entities, not limited to a specific product and they represent various assets and entities in the dataspace, potentially including machinery, components, and products like car batteries.

The recycling company uses DPP4.0 to access Digital Product Passports for individual car batteries it receives for recycling. It retrieves detailed data about each battery, including its carbon footprint, material composition, recycling guidelines, and more. This information helps the recycling process and guides decisions regarding reuse and recycling. Simultaneously, the same recycling company uses Digital Twins implemented with AAS to access information about the broader ecosystem. It can obtain data about the recycling machinery, transportation logistics, and other assets involved in the recycling process, not just information about individual batteries. This broader view helps optimize the entire recycling operation.

In this scenario, the DPP offers a laser-focused view on the product specifications and its sustainable aspects based on what the data owner (DS data producer) would like to share focusing on the product traceability. While the DT provides a broad perspective on the battery's lifecycle, integrating real-time data and contextual information from various stages of the circular economy. By combining the focused product information from the DPP with the broader lifecycle view from the DT, stakeholders can make informed decisions on how best to repurpose, recycle, or dispose of e-mobility Li-Ion battery packs.

4.1.2 For Process / Production

Regarding the data sharing and exchange of process and production data, Deliverable D3.1 proposed a DT approach where the DT is synchronized with the process. Since processes can have a vast number of APIs and communication protocols, it is not recommended to connect data space connectors directly to these APIs. Additionally, the semantics in a process are usually only defined on process level and not with global ontologies, which means that a DT referencing common semantics is more suitable.

Since the DT synchronized with the process has a defined API and a subset of communication protocols, the integration of the DT into data space connectors can be realized more efficiently. In section 4.3.1 below, the approach to share process data with data space connectors used in Circular TwAIn are described.

4.2 Data Catalogues, Publication and Discovery

4.2.1 Data models and Ontologies for Materials and products

As also mentioned in D3.2 “Data Space and Digital TwAIns Design - 1st version” the successful implementation of Digital Product Passports (DPPs) in a value chain requires stakeholders to adopt a unified vision, information models, and practices. This entails using standards for modelling DPPs to ensure interoperability and trustworthiness. For this, a DPP registry is necessary, ensuring the findability and accessibility of a DPP. The technical implementation of the registry could either be a Data Space component or a centralized DPP registry, depending on whether it's intended for a single organization or all Data Space participants. The registry should offer functionalities for registering, finding, ranking, and interacting with DPPs. It can also store 'abstract' DPPs for material types. Importantly, only the DPP metadata will be centralized, while the DPPs themselves will be stored in a distributed manner. If DPPs are implemented based on the Asset Administration Shell (AAS) standard, the existing AAS registry could be extended, promoting development speed and synergy.

4.2.2 Data Catalogues and Discovery Technologies

EDC

The IDS Meta Data Broker is a service that facilitates the publishing and searching of metadata among International Data Spaces Participants. It serves as a specialized IDS Connector, similar to an App Store, to ensure interoperability and interactions within the IDS ecosystem.

The role of an IDS Meta Data Broker can be summarized as follows:

- Optional Component.
- Offers Index Service.

- Enables Link-Based Discovery.
- Multiple Brokers: International Data Spaces can have multiple IDS Meta Data Brokers operating concurrently.

The EDC offers a set of modules for cloud environments (AWS, Azure) which allows users to control these modules "out-of-the-box." The EDC framework is easily customizable and can be extended to support more environments as per specific requirements.

In the context of data cataloguing and discovery, the EDC provides the following capabilities:

- **Data Cataloguing**: The EDC enables participants to catalogue their data assets by providing metadata about the data, such as its structure, format, access permissions, and other relevant information. This metadata helps in describing and categorizing the data, making it easier to search and discover.
- **Data Discovery**: The EDC facilitates the discovery of data across different participants in a dataspace. It enables users to search for specific datasets based on various criteria, including metadata attributes, access permissions, and data availability. This allows organizations to find and access the data they require for their specific use cases.

Detailed information related to the Eclipse Dataspace Connector can be found in the GitHub repository [8].

Implementation

Source code and detailed information on data cataloguing related to the MVD (described in 4.3.1) can be found in the GitHub repository [14].

EDC Extension

The EDC supports the publication of its data resources to the IDS Metadata Broker via a connector extension[9]. However, if the participants know each other and the connector URI of the other participant, the data catalogue can be directly requested from the provider EDC. A broker is therefore only needed to discover data resources from unknown data space participants.

TRUE Connector

The TRUE Connector supports the publication of its data resources to the IDS Metadata Broker via proxy endpoint in charge of creating a valid IDS message which will add IDS related elements and forward them to Broker. Furthermore, the resource catalogue can be required directly to the provider through the SelfDescription API endpoints.

[9] https://github.com/sovity/edc-extensions/tree/main/extensions/ids-broker-client

4.3 Data Exchange

4.3.1 EDC

The EDC backend supports HTTPS/REST as communication protocol out-of-the-box, meaning that DTs implemented with AAS, that provide such an endpoint, can be directly integrated. If the DT only provides communication protocols like OPC UA, the backend service of the EDC must be extended to allow connections to such a process DT.

The EDC has its own information model that is converted to the IDS information model for communication with other IDS connectors. Internally, EDC Assets describe the data resources to be shared over the EDC. In this case, this would be process data. Such EDC Assets must be created with the EDC API. Unfortunately, even though the IDS information model is fully supported by the EDC, the compatibility with other IDS connectors is limited. This is a result of the split data- and control-plane, which is currently only found in the EDC.

Implementation

The EDC Connector is designed to implement the International Data Spaces standard (IDS) and other relevant protocols [2]. The EDC are designed in an extensible way in order to support alternative protocols and integrate in various ecosystems. EDC is offered under the Apache-2.0 license, and it is available through the Eclipse Dataspace Components organizations on GitHub[10]. Further documentation can be found at the EDC Docs[11].

To ease the deployment, familiarity, and usage of the EDC ecosystem a Minimum Viable Dataspace (MVD) setup is offered. The MVD is an example implementation of a dataspace that utilizes the Eclipse Dataspace Components (EDC) [9]. The primary goal is to showcase the capabilities of the EDC, provide palpable examples of dataspace concepts through a specific implementation, and can serve as a starting point for building a customized dataspace. The MVD enables developers to assess and use the EDC and provides the ability to realize the requirements of a fully operational dataspace. Finally, it offers a concrete manner of illustrating how practical implementation of decentralization can be achieved.

Dependencies

- Java SE 11+
- Gradle 5.0+
- Docker Engine 17.12.0
- Docker-Compose 2.4

Availability

- Code for EDC can be found in the respective repository in GitHub [7].
- Code for EDC-MVD can be found in the respective repository in GitHub [9].

Installation Guidelines

[10] https://github.com/eclipse-edc
[11] https://eclipse-edc.github.io/docs/

In order to run the above instances of the EDC and the EDC-MVD you need to run the below command(s):

```
export MVD_UI_PATH="/path/to/mvd-datadashboard"

docker compose --profile ui -f system-tests/docker-compose.yml up –build
```

The docker-compose file that will be executed can be found in the respective repository in GitHub [12].

Documentation

Detail documentation on how to run an instance of the EDC-MVD can be found in the GitHub repository of EDC Minimum Viable Dataspace under the ReadME.md file of the main repository [10]. More details on the exact example and usage of the MVD can be found in the system-test specific repository [13]. Moreover, a Postman Collection exists, to run and to issue requests to an MVD instance of your choice. You will need to adapt the environment variables accordingly to match the target MVD instance. The detailed Postman collection can be found in the corresponding repository [11].

Moreover, examples that cover the basic functionality can be found in the repository containing some basic samples which aim to teach the very fundamentals about using the EDC framework [5]. It is highly recommended for users new to the Project, to follow these samples first before moving on to any of the other scopes.

Extension for AAS

The integration between AAS-compliant digital twins and IDS connectors is an important topic for the Circular TwAIn Project, as the DTs of products and processes should be shared via a circular data space. To make this integration as easy as possible for domain experts and software developers, without requiring them to have in-depth knowledge of both standards, we decided to use, adapt and extend our background component “EDC Extension for Asset Administration Shell”[12] . Besides the fact that the component has been developed by us and thus customization and extension would be much easier, it has some important advantages compared to similar approaches. These include extensibility for different implementations of the AAS standard, support for both DT providers and DT consumers, a graphical user interface that guides the integration process without having to implement it yourself, etc. A detailed comparison can be found in our paper [29].

The extension is continuously being improved and updated for the Circular TwAIn architecture. The code can be found in the Circular TwAIn Github to support the pilots applying AAS with the EDC.

On the data provider side, AAS files (AASX, JSON) can be uploaded. Alternatively, URLs are entered into the extension to automatically share the AAS over the EDC. When changes are made to the AAS of a product or process, the extension automatically updates the corresponding resources in the EDC. Depending on the use-case, changes can happen

[12] https://github.com/Circular-TwAIn/EDC-Extension-for-AAS

frequently, in which case manual processing is not feasible and the extension or similar solutions are required.

On the consumer side, the extension also simplifies the request of AAS data and provides a user interface for both provider and consumer. With the GUI, consumers can graphically browse the data catalogue of other connectors and select suitable AAS data sources.

Main considerations when designing the extension include:

- supporting all AAS files / formats
- supporting different implementations of AAS services, if the AAS is already a running service
- providing graphical interface
- reduce user effort

Additional details on the extension components and implementation can be found in Circular TwAIn Deliverable D3.2. In theory, the extension could be decoupled from the EDC by interacting with the EDC API to create assets and contracts but the EDC provides convenient access to several internal components for extensions.

Lastly, it should be mentioned that the EDC receives frequent updates from the open-source community with breaking changes. The extension is currently in the process to be updated for support of the latest EDC version 0.2.0.

4.3.2 TRUE Connector

The TRUE Connector is integrable with several existing IDS services and totally configurable in terms of data format (multipart/mixed, multipart/form, http-header) and protocols (HTTP, HTTPS, Web Socket over HTTPS, IDSCPv2). This ensures a wide flexibility for the data sharing and exchange, so that the TRUE Connector can be easily extended to interact with AAS. The specific Data APP and the related functionalities have been already described in D3.2 making able to manage AAS-based scenario designing and implementing a dedicated extension. On the other hand, the already existing FIWARE extension enables the data exchange through Orion Context Broker, putting at the disposal of the data space the context information coming from a lot of data sources using different protocols supported by the IDAS Agents (i.e., MQTT, OPC UA, OMA LWM2M, LoRaWAN, Sigfox, etc.).

Implementation

The TRUE Connector can act as a provider or consumer to exchange any kind of data and files, but in order to support AAS-based information exchange, it can be extended through a dedicated Data APP. With that customized Data App the use will be able to manage AAS files, persist AASX files or get files from the remote backend system, such as crawl AAS data like submodules and elements. It could be possible even to create IDS mapping in terms of IDS artifacts. Once AASX data has been parsed, the mapping can be defined. An additional enhancement is to associate Usage Control Rules to the created artifacts, so that the user can add or edit a policy, enabling the sovereign data exchange. Can be applied an elaboration of received data with the aim to make data available for a specific AAS implementation (i.e., FAST). For this purpose, the FAST protocol and similar technologies can be supported.

4.4 Identity/Access and Contract/Policy Management

4.4.1 Circular TwAIn Dataspace Onboarding Process

In the domain of the Circular TwAIn Project, onboarding a new dataspace participant signifies the initiation of a technical-business relationship with a legal entity aiming to partake in the Project's circular data space. Given the diversity of sectors and geographies, the onboarding process tends to vary across participants. However, a cornerstone for the onboarding process in Circular TwAIn lies in the Data Sovereignty Agreements, as introduced in D2.2, which stem from the larger Industrial Agreements (IAs).

Data Sovereignty Agreements play a pivotal role in ensuring the confidential and trustworthy sharing of data. They guarantee the implementation of robust confidentiality and trust models which ensures the security and privacy of shared data, especially related to circularity data such as waste, CO_2 emissions, and energy consumption, which are vital for Circular TwAIn's AI application modules. Data Sovereignty Agreements encompass various levels, including materials, products, and processes, all aligned with the principles of the Industrial Data Space (IDS) paradigm and its data sovereignty principles. Below are the key technical considerations that are addressed from the two technical solutions, namely EDC and TRUE Connector, proposed in Circular TwAIn:

- Identity and Access Management (IAM): Ensuring that only authenticated and authorized entities have access to data.
- Sovereign Data Transfer: This pertains to the transfer of data in a manner that preserves its sovereignty.
- Policy Negotiation: Before data is shared or accessed, there would be a negotiation phase where both parties agree on a set of policies.
- Dynamic policy agreement: Automated tools to negotiate and finalize data sharing policies between parties.
- Policy enforcement: Tools to ensure that once policies are agreed upon, they are strictly enforced.
- Contract and Policy Management: Keeping track of all agreements and ensuring compliance.

While the technical approach provides the foundation for the Data Sovereignty Agreements, it is equally vital to align these technical strategies with the unique business contexts identified by the Circular TwAIn pilots. In this deliverable we focus on the technical approach on implementing the Data Sovereignty Agreements utilizing the methodology proposed from the EDC and TRUE connector options. The business context will be specified by the pilots which will validate this methodology in the second version of the deliverable.

Below we can find the methodology and implementation offered by the EDC and TRUE Connector which are proposed to be adopted in the Circular TwAIn Project.

4.4.2 EDC

The Eclipse Dataspace Connector offers a framework for connecting and managing data in a way that respects individual sovereignty. It enables seamless integration with identity, data catalogue, and transfer technologies, ensuring compliance, policy enforcement, and control capabilities across different organizations.

Onboarding of a new data space participant using EDC

Leveraging the capabilities of the Eclipse Dataspace Components (EDC) the onboarding process ensures a robust, decentralized, and verifiable approach to integrating new participants into the Circular TwAIn dataspace. The onboarding process for a new dataspace participant in the Circular TwAIn Project, implemented via the Eclipse Dataspace Components (EDC), can be understood as a sequence of activities that ensure a new participant's credentials and identity are verified, and they are successfully integrated into the dataspace. Below we are also using REVERTIA as a new onboarding participant example in the context of the Waste Electrical and Electronic Equipment (WEEE) pilot. Here is a detailed explanation of each onboarding step depicted also in Figure 5 below:

1. Obtain Necessary Verifiable Credentials (VCs): Before the formal onboarding process begins, participants need to gather and possess the necessary Verifiable Credentials (VCs) that adhere to the onboarding policy of the Circular TwAIn dataspace.
    - Process:
        - The potential participant identifies and obtains the necessary VCs, which may include company registration documents, certifications, or other documents attesting to their role, identity, or expertise.
        - These credentials validate the participant's status as a genuine entity that meets the specific requirements set by the dataspace.
    - REVERTIA Example: Gathering the essential VCs for REVERTIA to participate to a CT dataspace may include:
        - Certifications of handling electronic waste.
        - Licensing for re-manufacturing electronic components.
        - Business licenses for both domestic and commercial operations.
        - Certifications or evidence of sustainable and eco-friendly processes.
2. Control Plane Sends Onboarding Request: Initiate the onboarding process by sending a formal request.
    - Process:
        - The participant, through its Control Plane, sends an onboarding request to the Circular TwAIn dataspace's registration service.
        - This request typically contains the participant's Decentralized Identifier (DID) and other relevant data to assist in their identification and verification.
    - REVERTIA Example: to initiate REVERTIA's integration into the dataspace:

            - REVERTIA, using its Control Plane, sends an onboarding request containing its Decentralized Identifier (DID) and a description of its operations to the Circular TwAIn dataspace's registration service.

3. Retrieval and Verification of Credentials: Ensure the potential participant's credentials are valid and align with the dataspace's requirements.
    - Process:
        - The registration service retrieves the participant's DID document. This document contains information that aids in the validation of the participant's identity.
        - Subsequently, the registration service fetches the Verifiable Credentials from the participant's Identity Hub.
        - These credentials are then cross-verified for their authenticity, ensuring they have not been tampered with and are issued by a recognized entity.
    - REVERTIA Example: to authenticate REVERTIA's qualifications and expertise in the WEEE domain.
        - The registration service extracts REVERTIA's DID document.
        - It then fetches the Verifiable Credentials from REVERTIA's Identity Hub, ensuring they are genuine and correspond to the provided DID.
4. Evaluation Against Onboarding Policy: Assess if the participant meets the onboarding requirements set by the Circular TwAIn dataspace.
    - Process:
        - The onboarding policy is a set of conditions or criteria that a participant must meet to join the dataspace.
        - The claims contained in the participant's Verifiable Credentials are scrutinized against this policy.
        - This evaluation ensures that only entities that align with the values, goals, and requirements of the dataspace are allowed to participate.
    - REVERTIA Example: to determine if REVERTIA aligns with the objectives and requirements of the Circular TwAIn dataspace.
        - The claims in REVERTIA's Verifiable Credentials are assessed against the dataspace's onboarding policy.
        - Circular TwAIn WEEE dataspace onboarding policy for REVERTIA could emphasize on:
            - REVERTIA's capability to drive automation in WEEE treatment.
            - Its commitment to the circular economy, ensuring maximum reuse and responsible recycling.
            - Its methods to reduce manual and repetitive tasks.

   - Its potential to promote Circular Economy Business Models.

5. Successful Onboarding and Dataspace Entry: Formalize the participant's entry into the dataspace.
   - Process:
     - If the evaluation in step 4 concludes that the onboarding policy has been met, the participant is formally added to the list of dataspace participants.
     - To provide evidence of this successful onboarding, a new Verifiable Credential is generated. This VC attests to the participant's new status as a member of the Circular TwAIn dataspace.
     - This VC is then pushed to the participant's Identity Hub, serving both as a confirmation of their membership and a credential for future interactions within the dataspace.
   - REVERTIA Example: to finally formalize REVERTIA's membership in the Circular TwAIn dataspace.
     - Once REVERTIA meets the onboarding policy conditions, it's added to the dataspace participants list.
     - REVERTIA is then provided with a new Verifiable Credential, affirming its status in the Circular TwAIn dataspace. This VC can also help REVERTIA in future interactions, collaborations, or verifications within the dataspace.

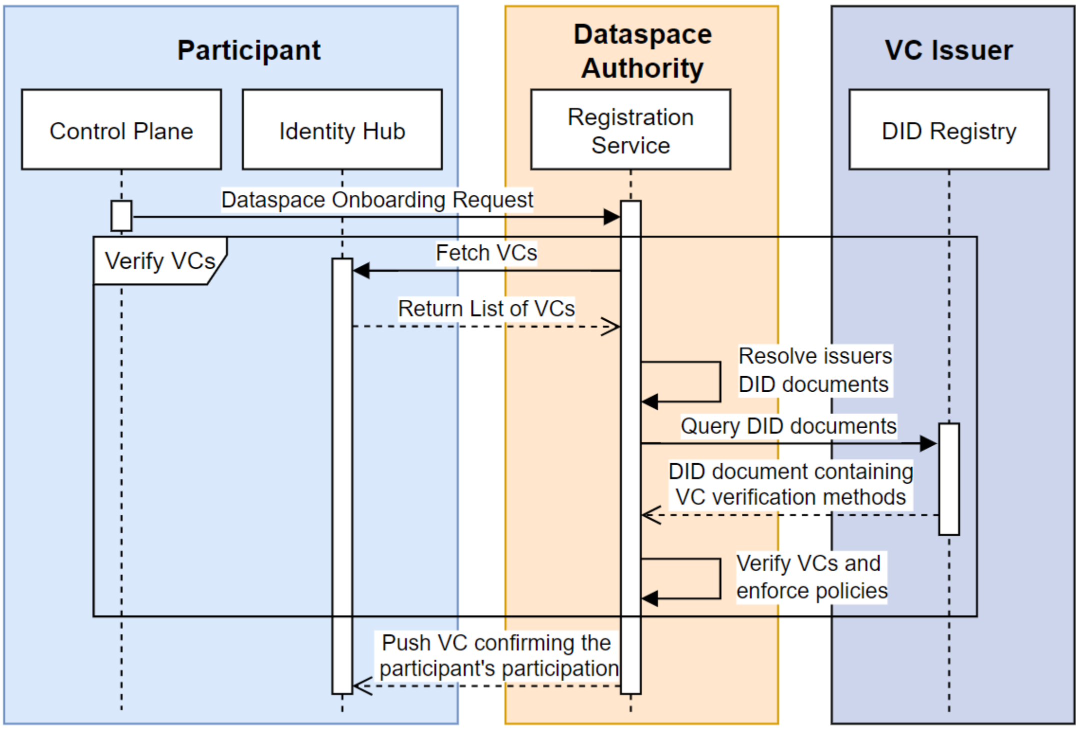

Figure 5: Onboarding new EDC dataspace participant.

Decentralized Identifiers

Decentralized Identifiers (DIDs): These are identifiers unique across the EDC ecosystem. They can be generated and governed by their respective owners without relying on external entities, embodying the principle of self-controlled identities.

Verifiable Credentials: Within EDC, and aligning with the W3C Standard, a verifiable credential is essentially a declaration made by an entity about itself. This declaration is endorsed by another participant and may be validated by a third party in the EDC system.

Identity Hubs: Within the EDC environment, entities holding verifiable credentials require a mechanism to store these credentials both digitally and securely. Identity hubs in EDC offer this storage solution, ensuring credentials are not only stored safely but are also accessible for verification when needed**.**

How the Eclipse Dataspace Connector (EDC) leverages DIDs and VCs to implement decentralized identity management can be found at the EDC Identity Management documentation[13].

Identity Hub - Get claims

Upon receiving an IDS request, an EDC participant must authenticate the sender's identity and enforce the appropriate access rules for that sender. To implement these rules, the participant requires the claims associated with that sender. Document details on the process of retrieving these claims from IdentityHub can be found at EDC documentation[14].

Implementation

Identity providers: includes OAuth2-based implementations as well as distributed identity systems [6].

Policies: Policy model used in EDC, as in TRUE Connector (see section below), is based on the Open Digital Rights Language (ODRL)[15]. EDC policy definition is described in the EDC usage control documentation available on GitHub[16]

Minimum Viable Dataspace

Regarding the MVD there are a set of test resources which are provided in order to run MVD locally [13]. Each EDC Connector has its own set of Private and Public keys in PEM and Java KeyStore formats, e.g., system-tests/resources/vault/company1. The following command are used respectively to:

- generate a private key.
- generate corresponding public key.
- create a self-signed certificate.

[13] https://eclipse-edc.github.io/docs/#/submodule/Publications/Identity%20Management/DID_EDC?id=decentralized-identifiers-and-the-eclipse-dataspace-connector
[14] https://eclipse-edc.github.io/docs/#/submodule/IdentityHub/docs/developer/decision-records/2022-07-01-get-claims/
[15] https://www.w3.org/TR/odrl-model/
[16] https://github.com/eclipse-edc/Connector/blob/main/docs/developer/architecture/usage-control/policies.md

```
openssl ecparam -name prime256v1 -genkey -noout -out private-key.pem

openssl ec -in private-key.pem -pubout -out public-key.pem

openssl req -new -x509 -key private-key.pem -out cert.pem -days 360
```

The generated keys are imported into keystores, (i.e., "company1-keystore.jks") which are located in the "system-tests/resources/vault/company1/" directory. The keystore is protected with the password "test123". In the case of MVD local instances, a filesystem based vault is used, and its keys are managed using a java properties file, (i.e., "company1vault.properties" located in the "system-tests/resources/vault/company1/" directory. The file system vault is should only be used for testing purposes. The Web DIDs can be found in the "system-tests/resources/webdid" folder. The "publicKeyJwk" section of each "did.json" file was generated by converting the corresponding public key to JWK (JSON Web Key) format. For instance, the public key of the company1 connector was converted to JWK using the command:

```
docker run -i danedmunds/pem-to-jwk:1.2.1 --public --pretty < system-
tests/resources/vault/company1/public-key.pem > key.public.jwk
```

TRUE Connector

The TRUE Connector supports the entire contract negotiation flow, once the agreement between the two parties is reached, defining the rules to exchange the resources offered, starts the interaction with the Usage Control framework. In particular, the TRUE Connector integrates Platoon Usage Control, an open-source module, using an in-memory database with persisting database on file system or alternatively some more resilient databases can be used like PostgreSQL. Platoon Usage control supports usage policies written in the IDS Usage Control Language based on ODRL. The policy patterns supported by the Data Usage Control module are the following ones:

- Allow the Usage of the Data: provides data usage without any restrictions.
- Prohibit the Usage of the Data: prohibits data usage.
- Interval-restricted Data Usage: provides data usage within a defined time interval. (interval defined with start end end date)
- Duration-restricted Data Usage: allows data usage for a defined time period. (duration calculated from contract start date)
- Role-restricted Data Usage: allows data usage for a defined role.
- Purpose-restricted Data Usage Policy: allows data usage for defined purpose.
- Restricted Number of Usages: allows data usage for n times.
- Personal Data: filter out the contents of the data according to the data subject´s consents. To apply this rule, the Usage Control module interacts with CaPe.

TRUE Connector is even compatible with MYDATA Usage Control, and switching is done by changing a property in the configuration file and integrating MYDATA Usage Control service into docker-compose.yml file.

The TRUE Connector is able to interact with the several Identity Providers (Fraunhofer AISECv2, Orbiter, Omejdn, ATOS Trust Centre) and DAPS Services for requiring and validating a token.

TRUE Connector does not provide technical enforcement of usage contracts out of the box. It tries to establish legally binding contracts for each data resource between the participants to enable data sovereignty.

To have a basic setup and start TRUE Connector instance with both Provider and Consumer, alongside Usage Control, the first step is to download it from [4].

You can either use the provided certificates out-of-the-box or create your own self-signed ones following the next steps:

Generate private key and certificate with the following openssl command:

```
openssl req -x509 -newkey rsa:2048 -keyout consumer-key.pem -out consumer-cert.pem -sha256 -days 365 -subj "/C=IT/ST=Italy/L=Lecce/O=Engineering Ingegneria Informatica SpA/OU=R&D/CN=TRUEConnector" -addext "subjectAltName=DNS:be-dataapp-consumer,DNS:ecc-consumer,DNS:uc-dataapp-consumer"
```

- keyout - key name
- out - certificate name
- subj - information about the certificate owner like Company name, Country etc.
- subjectAltName - host names and/or IP address of the consumer components and host machine

Afterwards you will be prompted to insert the password.

To generate a KeyStore with the previous key and cert use the command:

```
openssl pkcs12 -export -out consumer-keyStore.p12 -inkey consumer-key.pem -in consumer-cert.pem -name true-connector-consumer
```

- out - KeyStore name
- inkey - private key
- in - certificate

Since the TRUE Connector uses the Java programming language it is advised to use the .jks format for the KeyStores and TrustStores. To convert the consumer KeyStore from .p12 to .jks use the following keytool command:

```
keytool -importkeystore -destkeystore true-connector-consumer-keystore.jks -srckeystore consumer-keyStore.p12 -srcstoretype PKCS12 -alias true-connector-consumer
```

- destkeystore - name of the new .jks KeyStore
- srckeystore - name of the .p12 KeyStore

Here are the commands for the same process for the provider:

```
openssl req -x509 -newkey rsa:2048 -keyout provider-key.pem -out provider-cert.pem -sha256 -days 365 -subj "/C=IT/ST=Italy/L=Lecce/O=Engineering Ingegneria Informatica
```

```
SpA/OU=R&D/CN=TRUEConnector"  -addext  "subjectAltName=DNS:be-dataapp-provider,DNS:ecc-provider,DNS:uc-dataapp-provider"


openssl pkcs12 -export -out provider-keyStore.p12 -inkey provider-key.pem -in provider-cert.pem -name true-connector-provider


keytool -importkeystore -destkeystore true-connector-provider-keystore.jks -srckeystore provider-keyStore.p12 -srcstoretype PKCS12 -alias true-connector-provider
```

At the end we need the TrustStores for the consumer and provider respectively in which we will add both certificates:

```
keytool -import -keystore true-connector-consumer-truststore.jks  -file provider-cert.pem -alias true-connector-provider


keytool -import -keystore true-connector-consumer-truststore.jks  -file consumer-cert.pem -alias true-connector-consumer


keytool -import -keystore true-connector-provider-truststore.jks  -file consumer-cert.pem -alias true-connector-consumer


keytool -import -keystore true-connector-provider-truststore.jks  -file provider-cert.pem -alias true-connector-provider
```

These commands can be also used to add certificates from other services e.g. DAPS, Broker, other providers, to the TrustStores.

After creating the KeyStores and TrustStores, you have insert their name, passwords, aliases and private key passwords in the .env:

```
#Consumer SSL settings

CONSUMER_SERVER_SSL_ENABLED=true

CONSUMER_KEYSTORE_NAME=true-connector-consumer-keystore.jks

CONSUMER_KEY_PASSWORD=password

CONSUMER_KEYSTORE_PASSWORD=password

CONSUMER_ALIAS=true-connector-consumer

#TRUSTORE (used also by IDSCP2)

CONSUMER_TRUSTORE_NAME=true-connector-consumer-truststore.jks
```

```
CONSUMER_TRUSTORE_PASSWORD=password


#Provider SSL settings
PROVIDER_SERVER_SSL_ENABLED=true
PROVIDER_KEYSTORE_NAME=true-connector-provider-keystore.jks
PROVIDER_KEY_PASSWORD=password
PROVIDER_KEYSTORE_PASSWORD=password
PROVIDER_ALIAS=true-connector-provider
#TRUSTORE (used also by IDSCP2)
PROVIDER_TRUSTORE_NAME=true-connector-provider-truststore.jks
PROVIDER_TRUSTORE_PASSWORD=password
```

For the most recent changes, updates and commands, please refer to [4]. After which you need to start the script which will prepopulate volumes with the next command:

```
sudo ./prepopulate_be_dataapp_data_provider.sh
```

and then start docker containers with the next command:

```
docker-compose up &
```

The basic setup comes pre-configured with the following:

- Secure https communication between all components (dataApp - ECC, ECC-ECC, and ECC-dataApp), using self-signed certificate
- multipart form format of the message between all components
- DapsInteraction disabled
- Disabled Usage control
- Disabled Clearing House
- Disabled validate protocol in Forward-To header
- Disabled hostname validation
- Disabled CheckSum validation

For full MVDS, which is composed out of TRUE Connector with both Provider and consumer, Platoon Usage Control, Omejdn as Identity Provider and MetaData Broker, refer to [3].

Figure 6 depicts the whole contract negotiation flow with included external components such as Usage Control and Clearing House.

Before the start of the negotiation process, a DescriptionRequestMessage is sent to identify the actors and potentially deny access if the Dynamic Attribute Token (DAT) is not valid. Initially, a DescriptionRequestMessage is sent by the consumer without payload. If DAT is invalid, the Provider sends RejectionMessage with an optional reason. However, if DAT is valid, SelfDescriptionResponse is sent to the Consumer. If the incoming message is assumed trustworthy, the Provider answers with a DescriptionResponseMessage. During the establishment phase of the negotiation, this message contains the currently valid SelfDescription of Provider in JSON-LD, including its provided IDS Resources and respective ContractOffers. Note that the connector is not obliged to provide ContractOffers for any/all resources but can also only announce their existence. The usage conditions might be sensitive too and do not need to be supplied. However, the provisioning of ContractOffers eases their usage and therefore should be in the interest of a Data Provider.

Contract Request Message is an initial message sent in the Contract Negotiation flow. It can contain requestedElement, if we know what artifact we are requesting, or without it, if we need to get the whole self-description document, and then analyze it and get the element we are looking for. If everything goes well, the Provider will send to Consumer response with "ids:ContractAgreement".

When a Contract Agreement request is sent, the Provider validates the contract agreement, where the contract agreement is uploaded in Usage Control, after which it is registered in Clearing House. In the end, response will be MessageProcessedNotificationMessage, after which the consumer checks the MessageProcessedNotificationMessage and original message are ContractAgreements messages and repeats processes in Usage Control and Clearing House. If potential problems occur in the Clearing House, or in Usage Control (e.g., contract is not valid), Rejection messages will be returned.

After whole whole Contract Negotiation process is done, consumer can send ArtifactRequestMessage.

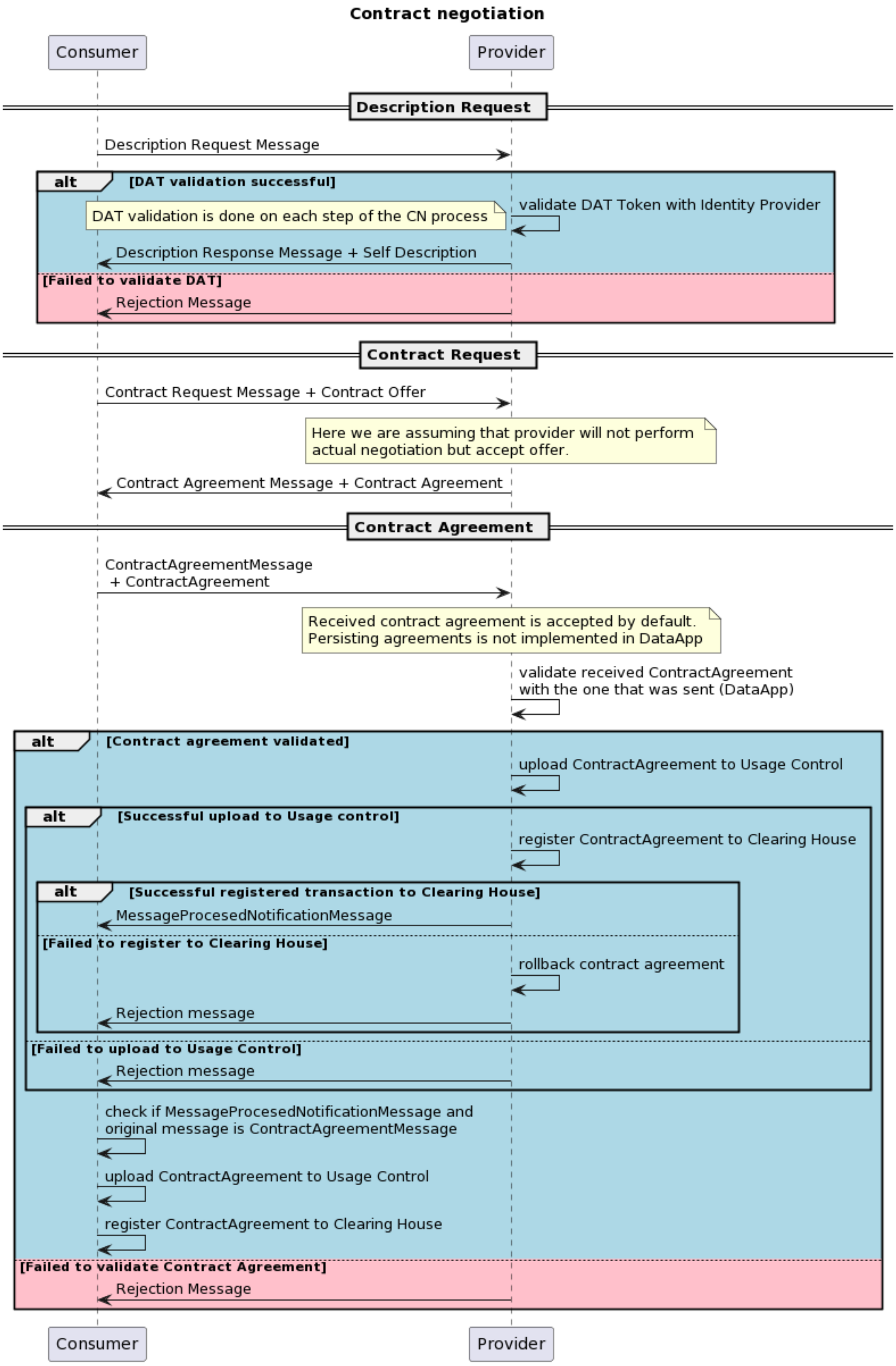

Figure 6: TRUE Connector Contract Negotiation Diagram

4.5 Technology Selection and Minimum Required Components

In this section, we provide an overview of the reasoning behind selecting the different technologies to implement the CT dataspace and data models. We also provide a summary of the minimum required components for a DS and map them with the identified technologies.

DPP4.0

DPP4.0 emerges as a frontrunner in the arena of product information documentation and accessibility. Its core strength lies in its ability to combine standards like the digital nameplate (DNP4.0) via IEC 61406 and the Asset Administration Shell (AAS) as per IEC 63278. DPP4.0 was selected for data sharing due to its compatibility to AAS, ensuring interoperability when instantiating the products traceability data.

EDC (MVD)

The selection of the Eclipse Dataspace Components (EDC) was based on several key factors. Firstly, its affiliation with the Eclipse Foundation (EF) ensures robust governance structures and facilitates collaborative community building. Secondly, EDC's versatility extends beyond IDS, accommodating various data protocols, including IDS(A) and alternatives like WEB-DID. Lastly, the requirements stemming from initiatives like Catena-X, Eona-X, IDSA, Gaia-X, and MDS align with EDC's capabilities, enabling a decentralized dataspace implementation with extensible shared service interfaces. The key focus of the EDC is the contract negotiation between data providers and consumers to realize legally binding contracts which are a requirement for data exchange. Moreover, the EDC also supports data exchange outside the connector to support big data. EDC emerged as the optimal choice for CT Project's dataspace due to its alignment with Project requirements. Its adaptability and versatility make it a promising foundation for data exchange and integration efforts. Moving forward, CT Project intends to leverage EDC for further development, customized deployments, and new integrations, reinforcing its role in realizing the Project's objectives and facilitating seamless data exchange.

EDC Extension for AAS

The EDC Extension for AAS simplifies the sharing and usage of AAS with the EDC and is the only available solution to increase user experience for AAS in EDC. Without the Extension, the usage of the EDC in the pilots would require knowledge of technical EDC details on how to publish AAS digital twins with the EDC.

TRUE Connector (MVDS)

The TRUE Connector is one of the most popular open-source IDS Connector implementations. The certification process is ongoing, making the solution one of the first certified IDS components. The connector is also compliant with a wide range of existing IDS infrastructural components, making it easily integrable for making a Data Space and through its flexibility to be part of existing ones.

An overview of the core elements and the technologies used is provided in Table 4.

Table 4: CT DS Core elements and technology mapping

Impl. Name	CT Usage	CT Example	Background/ Foreground/ Extension	Requirement Mapping
Element: Data Models & Vocabularies				
DPP 4.0	Sharing and traceability of product data	Provide detailed information about individual car batteries within the circular economy dataspace.	Background	Data storage, Common ontology and semantics, Material traceability, Info traceability
TRUE Connector IDS Info model 4.2.7	Description of the metadata used for supporting the data exchange	CT Data Space Participants can interact one each other using the IDS Info Model and a specific vocabulary for the payload.	Background	Data storage, Common ontology and semantics
Element: Data Catalogues, Publication & Discovery				
EDC	EDC stores Data Resource Metadata in Catalogue and publishes to IDS Broker	Battery Passport is offered with EDC	Background	Connectivity, Data storage, Common ontology and semantics
EDC Extension	Simplifies import of AAS into Catalogue	Battery Passport Provider does not require technical details and shares Battery AAS with Extension	Foreground	Data storage, Common ontology and semantics
TRUE Connector MVDS	TRUE Connector MVDS stores Data Resource Metadata in self-description document and publishes to IDS Broker (part of the MVDS)	CT Providers can customize the self-description document to create a catalogue of resources.	Foreground	Connectivity, Data storage, Common ontology and semantics
Element: Data Exchange				
EDC	EDC is responsible for data-sovereign data exchange to other connectors	Recycling Center and Battery Provider exchange Battery AAS with EDC	Background	Data monetization,

				Data business model, Data sovereignty
EDC Extension	Extension simplifies the data exchange by offering graphical interface	Recycling Center requests the Battery AAS by using Extension	Foreground	
TRUE Connector MVDS	TRUE Connector (part of the MVDS) is responsible for data-sovereign data exchange to other connectors	CT Participants can exchange data using an under-certification component interoperable with existing solutions like DSC, GEC Connector, Boost4.0 Connector, etc.	Foreground	Data monetization, Data business model, Data sovereignty
Element: Identity & Access Management				
EDC	EDC supports X509 identity offered by IDS IdentityProvider	Battery Providers each have unique identity to provide authorization / authentication		Data filtering & aggregation, Data transaction, Trustworthy
TRUE Connector MVDS	TRUE Connector MVDS supports X509 identity offered by IDS Identity Provider (part of the MVDS)	CT Providers can be registered and identified in order to participate to the Data Space	Foreground	Data filtering & aggregation, Data transaction, Trustworthy

4.6 EDC, IDS and True Connector Relation

As already mentioned above in Circular TwAIn we will work towards using multiple solutions for implementing a dataspace environment (e.g., TRUE Connector and EDC). In Figure 7 below we can see a possible high-level relationship of the different implementations and standards applied in separate federations and exchanging data thru the different connectors, which in our case are the EDC [30] and the TRUE Connector.

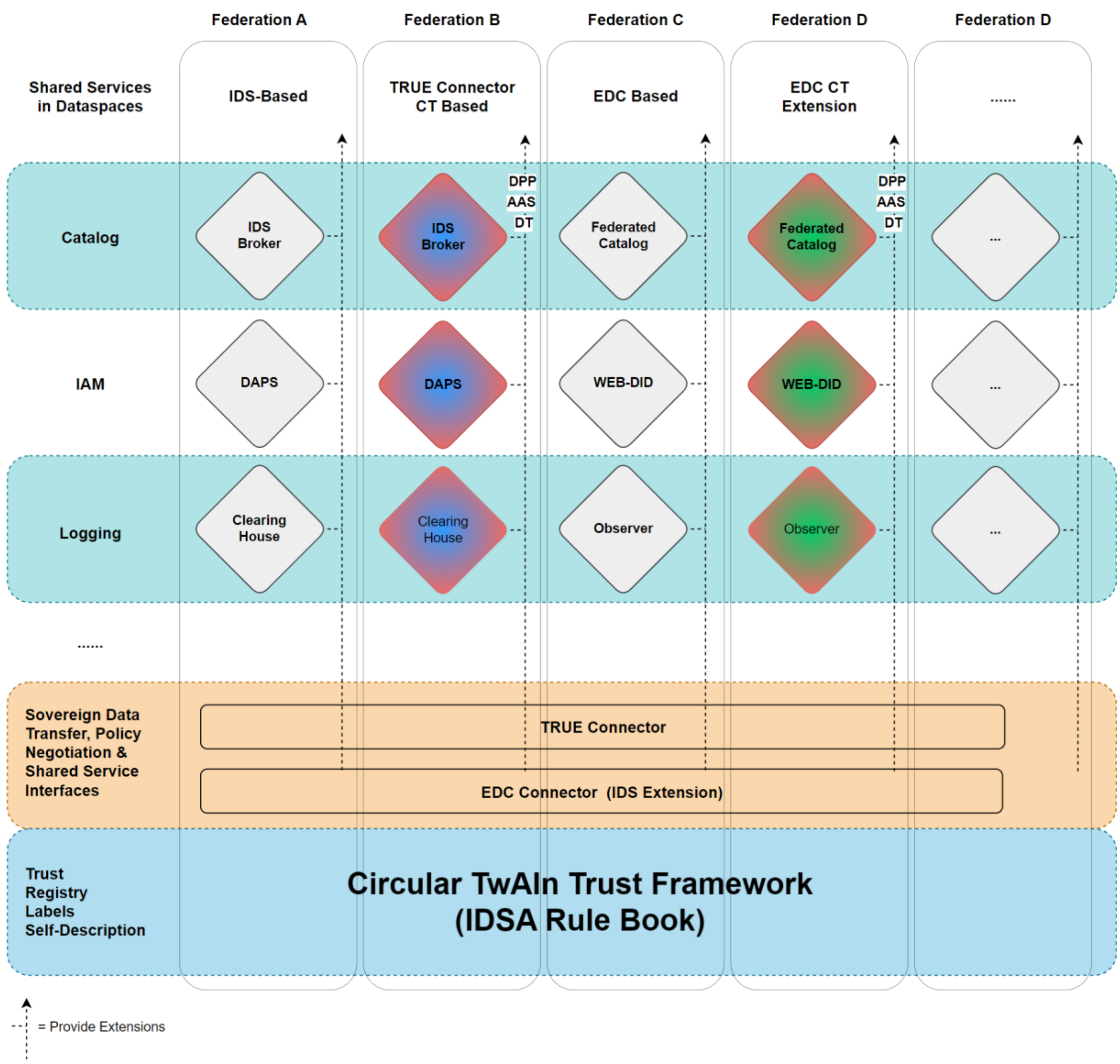

Figure 7: IDS TRUE EDC relations over CT

4.7 Integration, Packaging, Availability and Testbed Tools

Within T4.3 and T4.4 we provide the design and relevant implementation for a CT MVDS based on the already available implementations of other third-party components but also extensions and implementations provided within the context of CT based on the pilot requirements.

In this section we present the code/artifacts availability, deployment strategy and how it can be utilized by the different components. It also provides information about possible tools that can be utilised for the platform management, monitoring and access control for a CT MVDS.

4.7.1 Code Management

To establish a consistent and centralized platform for managing the code and extensions related to the Circular TwAIn (CT) dataspace GitHub is utilized as the primary tool. A dedicated organization under the name "Circular TwAIn" has been set up (see Figure 8 below) on GitHub (Circular TwAIn GitHub Organization[17]). For the Circular TwAIn GitHub Organization governance, we have agreed that each CT technology provider will have an assigned organization owner, responsible for maintaining and managing repositories. The CT GitHub repositories will host:

- New developments, where new repositories will be created under the organization.
- Extensions of existing projects tailored for Circular TwAIn requirements.
- Links to third-party components used unaltered. A README file will serve as the introductory guide, pointing to external tools and resources.

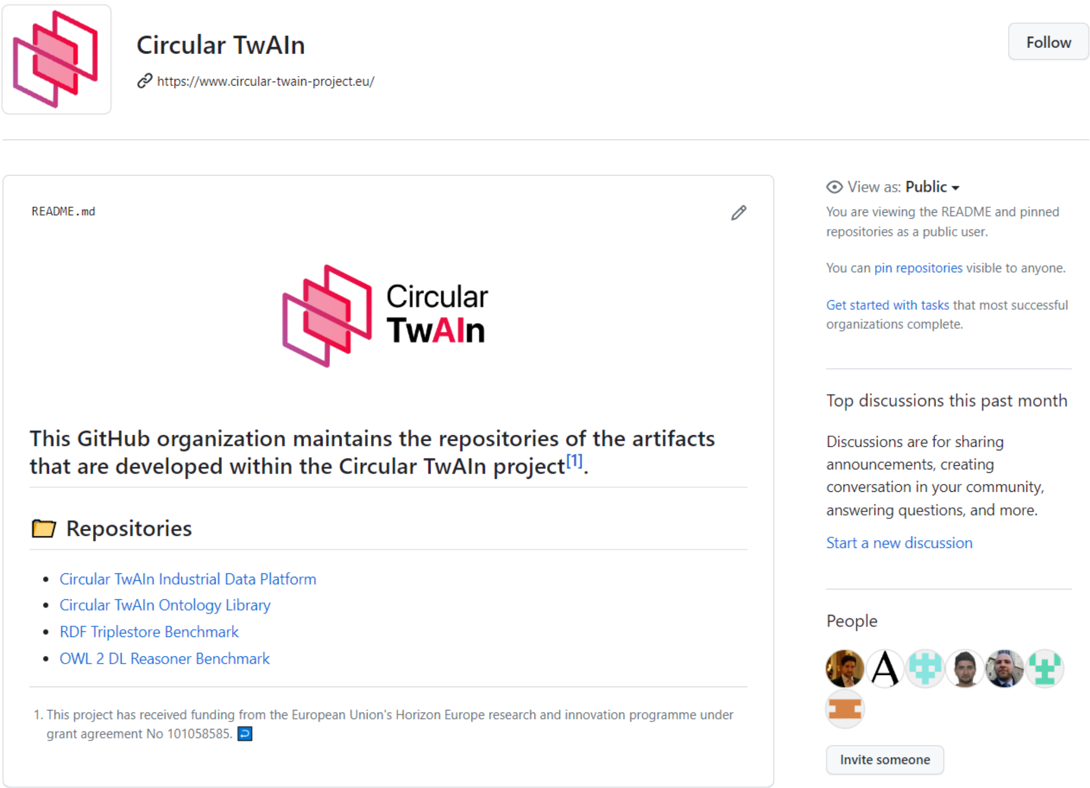

Figure 8: Circular TwAIn GitHub Organization welcome page

4.7.2 Deployment and Packaging Strategy

To implement an efficient packaging and deployment strategy for swift and scalable application rollouts recognized as an industry-standard solution is Docker[18] containers and Docker Compose respectively.

- Advantages: Quick deployment cycle, container isolation, and enhanced security features, including policy enforcement as outlined in IDS RAM.
- Docker Compose[19]: For orchestrating multiple Docker containers, Docker Compose is employed. With its YAML configuration, setting up interconnected services is straightforward. Comprehensive documentation is accessible at Docker Docs.

[17] https://github.com/Circular-TwAIn
[18] https://docs.docker.com/
[19] https://docs.docker.com/compose/gettingstarted/

4.7.3 Testbed infrastructure and Tools

To ensure the successful integration and validation of the Circular TwAIn dataspace components a dedicated testbed environment is planned to be established. This section provides a comprehensive overview of the infrastructure and tools earmarked for this task.

Deployment Infrastructure

For the testbed environment, virtual hosts will be procured and deployed on a public cloud provider known as Hetzner Cloud[20]. This choice offers several advantages, including scalability, flexibility, and ease of management. One of these VMs will serve as the primary deployment and management host, responsible for orchestrating the deployment of dataspace components and supporting services. The other VMs will be utilized to simulate and validate various aspects of the dataspace environment, including different types of participants.

Deployment Automation

For automating the required infrastructure and components deployment, especially the ones that configurations require frequent updates, appropriate deployment automation tools are going to be used like Jenkins[21]. Jenkins is a leading open-source automation server. Its wide range of plugins and integrations allows it to seamlessly automate tasks related to downloading, building, and deploying software components.

To ensure that updates to the Circular TwAIn dataspace components are captured and tested in real-time, Jenkins will be integrated with GitHub. This means that whenever there is a new commit, especially involving setup and configuration changes, Jenkins will be triggered. Upon detecting a change, Jenkins will:

- Download the latest version of the dataspace components.
- Build the components as required.
- Deploy them to the testbed environment using the new configurations.

Access Control

In order to offer secure access to the infrastructure and more specifically for the platforms and services that does not implement authentication the option of a SSO (Single Sign On) identity and access management will be offered. One of the Most commonly used Open-Source identity and access management software is the Keycloak[22] which offers enhanced security through Single Sign-On (SSO) functionality. Some of the Keycloak features are:

- Centralized management

[20] https://www.hetzner.com/cloud
[21] https://www.jenkins.io/
[22] https://www.keycloak.org/

- Support for standard protocols such as OpenID Connect, OAuth 2.0, and SAML 2.0.
- Seamless integration with platforms through adapters.
- Compatibility with LDAP and Active Directory.

Keycloak will be containerized using Docker, aligning with the overarching deployment strategy. Directions on how to install Keycloak using docker can be found at the Keycloak on Docker page[23]. Moreover, all the information related with the Keycloak functionalities, deployment and usage can be found at the Keycloak's documentation[24].

Infrastructure Management & Monitoring

Since the preferred deployment strategy is the docker containerization to facilitate the ecosystem management and monitoring there are various offerings one of which is the Community Edition (CE) of Portainer[25].

Portainer CE Simplifies Docker ecosystem management, by offering a user-friendly interface, thus making Docker's complexities more manageable.

Features:

- 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.

[23] https://www.keycloak.org/getting-started/getting-started-docker
[24] https://www.keycloak.org/documentation
[25] https://www.portainer.io/products-services/portainer-community-edition/

5 Data Space Usage Scenario / Testbed

5.1 Data Spaces Mapping to Pilot Scenarios

In Table 5 we specify whether the pilot and use cases are foreseen to use data spaces or not. This is indicated in the data space usage column. We can already anticipate that all use cases will create a Data Space.

Table 5: Data spaces mapping to pilot scenarios

PILOT	USE CASE	CIRCULAR SCENARIO	DATA SPACE USAGE	DATA TRANSMITTED (Type, format, frequency…)
WEEE	USE CASE A: Computer-vision driven product identification for the disassembly of IT equipment	After implementing this UC, the expectation is to have a vision system in the sorting area of the recycling facility to auto identify the arriving items and provide support to the operators when deciding what to do with said items. This relates to the scenario "EoL -> Operations"	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
WEEE	USE CASE B: Characterization and assessment of components and subcomponents	After implementing this UC, the expectation is to have a digital helper classifying and characterising the different elements identified in the shorting station, evaluating them and providing a profile of them for the operator. This profile will be used decide the best way to approach the recycling process to apply. This relates to the scenario "From End of Life to Manufacturing"	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
WEEE	USE CASE C: Real time planning of the disassembling operations	After implementing this UC, the expectation is to provide an AI solution to handle complex WEE waste elements automatically. Providing a human or robotic operator with the information to successfully disassemble them. This relates to the scenario "From End of Life to Manufacturing"	Yes	See details in D2.2, sections 4.1.4 and 4.1.5

WEEE	USE CASE D: Collaborative robotics for the support of manual operations	After implementing this UC, the expectation is to provide a robotic solution to help human operators by automating some heavy or tedious processes. This relates to the scenario "From End of Life to Manufacturing"	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
WEEE	USE CASE E: Market-oriented holistic decision-support-system for WEE de- and re-manufacturing	After implementing this UC, the expectation is to provide an IA-based decision-support-system able to study the market and providing a human operator wide array of information to decide the best use of the components retrieved. This relates to the scenario "From End of Life to Manufacturing"	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
BATTERY	UC A: Computer-vision driven collaborative robotics for the disassembly of LIB packs	After implementing this UC, the expectation is to provide a computer-vision driven system supporting and improving the efficiency of the collaborative dismantling of the battery packs at the facility of the battery dismantler. This relates to the scenario "From End of Life to Manufacturing"	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
BATTERY	UC B: Machine learning aided automated disassembly of LIB modules	After implementing this UC, the expectation is to provide an AI system able to support the module disassembly optimizing the disassembly parameters according to the battery model, end-of-life state and of the cells to be detached. This relates to the scenario "From End of Life to Manufacturing"	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
BATTERY	UC C: AI tool for the characterization of the LIBs state-of-	After implementing this UC, the expectation is to provide a mutual data and model	Yes	See details in D2.2, sections 4.1.4 and 4.1.5

	health combining historical and testing data	driven AI supporting methodology to certify the state-of-health and remaining useful lifespan of LIB cells. This relates to the scenario “From End of Life to Manufacturing”		
BATTERY	UC D: AI tool for optimized mechanical recycling of degraded LIBs	After implementing this UC, the expectation is to provide an AI driven optimization tool to drive reconfigurable mechanical recycling technologies in the selective isolation of material streams, in function of the LIB cells materials composition and architecture and of market quotations of target materials. This relates to the scenario “From End of Life to Manufacturing”	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
BATTERY	UC E: Market oriented holistic decision-support-system for the LIBs de- and re-manufacturing	After implementing this UC, the expectation is to provide an AI-based decision-support-system able to guides operational choices in function of the most profitable expected scenario, merging data available on the battery in terms of materials composition and state-of-health with data characterizing the market quotations. This relates to the scenario “From End of Life to Manufacturing”	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
PETROCH	UC A: Data acquisition and representation for AI framework	After implementing this UC, the expectation is to provide historical process data for AI framework. This relates to the scenario “From Operations to design”.	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
PETROCH	UC B: Developing hybrid circular twin of the process	After implementing this UC, the expectation is to provide data-driven Digital Twin and	Yes	See details in D2.2, sections 4.1.4 and 4.1.5

		first-order physical model of the EO stripping and also to provide DME simulation. This relates to the scenario "From Operations to design".		
PETROCH	UC C: Use of data analytics, AI and model verification to understand process unit failures	After implementing this UC, the expectation is to provide AI tool for detecting process anomalies with >90% accuracy. This relates to the scenario "From Operations to design".	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
PETROCH	UC D: Developing AutoML module for Process Industry	After implementing this UC, the expectation is to provide AutoML module for Process Industry. This relates to the scenario "From Operations to design".	Yes	See details in D2.2, sections 4.1.4 and 4.1.5
PETROCH	UC E: Generation of a tool for process optimisation	After implementing this UC, the expectation is to reduce energy consumption of the EO stripping process by using AI app. This relates to the scenario "From Operations to design".	Yes	See details in D2.2, sections 4.1.4 and 4.1.5

Below we can find a template (Table 6) for the DS implementation characteristics that will be used for each use case. This table will be applied and provided in next version of the deliverable.

Table 6: DS implementation characteristics per use case

TYPE OF IDS CONNECTOR in each use case	Eclipse Dataspace Components Connector (EDC TRUE Connector (TRU) Others
REASONS for the SELECTION of the TYPE OF IDS CONNECTOR	To be explained
STAKEHOLDERS in the DS	Improvement of D2.2., Section 4.1.2/4.2.2/4.3.2, indicating the stakeholders in each use case.

IDS RA MODULES Required in each use case	DAPS, Broker, CH …
DS technical requirements that will be analysed AND HOW in each use case (see D2.2- 5.2.1)	Connectivity: Data filtering and aggregation: Data storage: Data transaction Decision support or actuation
DS BUSINESS requirements that will be analyzed AND HOW in each use case (see D2.2- 5.2.2)	Data monetization Data business model Data sovereignty Trustworthy

5.2 Data Space PoC Testbed

The IOSB testbed in Karlsruhe, Germany is a test ground for AAS and EDC components in typical factory and sensor environments before usage in the pilots. It includes factory equipment like PLC, power sensors, miniature size robotic arm and conveyor belts. The software components can be installed and tested on a workstation PC connected to the industrial ethernet switch.

In the IOSB testbed, two EDC connectors for provider and consumer and a Omejdn DAPS were set up to simulate the sharing of battery data stored in an AAS. The goal is the implementation of a minimal data space that can be extended for the pilots. For the AAS, the FA³ST Service was used to load the battery model and connect to external systems for live data. The EDC was built with the previously mentioned EDC Extension for AAS to simplify the sharing process with an UI.

Figure 9 shows the EDC Extension UI to share an AAS over the EDC.

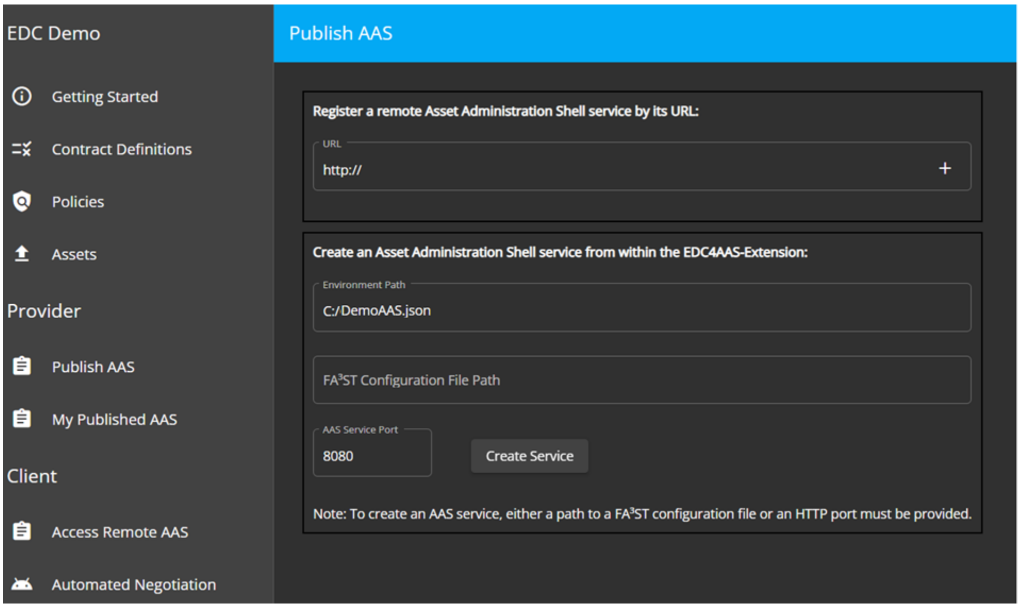

Figure 9: Extension UI to share an AAS over the EDC

After the AAS is shared, a valid contract between provider and customer must be negotiated. Figure 10 below shows the customer UI to select relevant AAS data. After that the negotiation window will be opened to send out a Contract Request to the provider. If the provider accepts, the data will be transmitted to the endpoint the consumer stated. A broker was not yet set up, but publicly available IDS brokers can be used with the EDC as stated in 4.2.2.

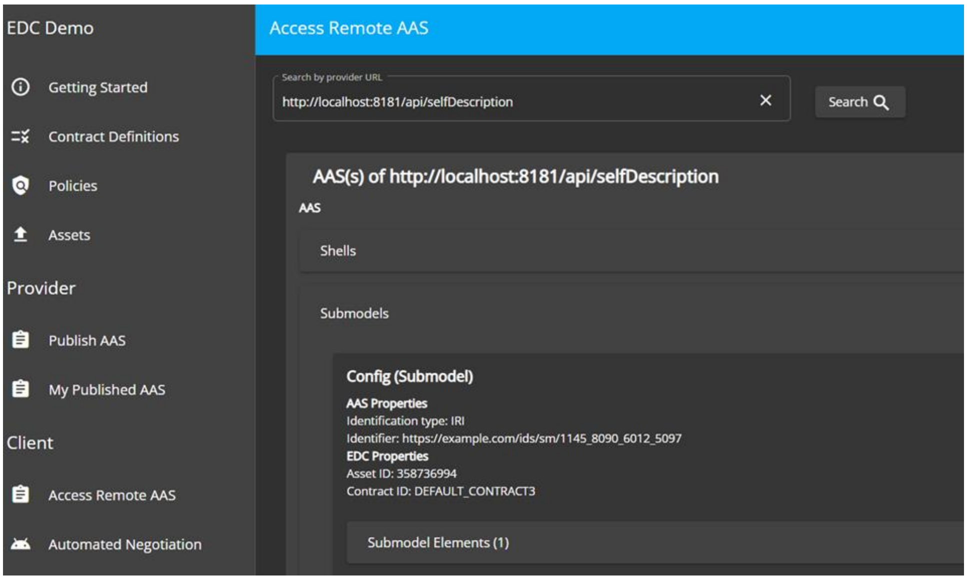

Figure 10: Customer UI to select relevant AAS data

Our next steps will be the adaptation of components like the EDC Extension for AAS for the TRUE connector and a solution for the incompatibility between TRUE and EDC connector. Additionally, components like the IDS Vocabulary Provider can be used to provide circularity ontologies, which must be referenced in the AAS. An IDS Clearing House could be beneficial for the pilots in case they require traceability of data and the creation of bills according to data usage.

6 Conclusion and Future Outlook

The work carried out under Task T4.3 and Task T4.4, as detailed in this deliverable, demonstrates significant strides towards realising the vision of the Circular TwAIn Project. The novel designs and implementations used of the data spaces for both materials/products and process/production lifecycle assessments have laid the groundwork for a transformative approach to circular manufacturing value chains.

Through our work on identifying data models and ontologies, the provision and consumption of materials/products data in the data space have been enabled. This facilitates the development of sustainable by design products and promotes other essential optimizations. Furthermore, we have effectively designed and built upon necessary data space building blocks, including those for data sharing and exchange, and data sovereignty. The attention given to the connectivity layer has ensured the effective creation, management, and sharing of digital twins' data.

The Circular TwAIn Data Space Architecture detailed in this deliverable embodies the confluence of Digital Twins (DTs) and Explainable AI (XAI) technologies. The architecture provides a robust platform to enable the exchange, modelling, and processing of data related to product/material and process/production in a circular economy data space. By facilitating data publishing, discovery, and exchange among untrusted actors, it promotes collaboration and boosts innovation.

As we move forward, the focus will be on:

- Establishing a DS testbed environment for component integration and validation.
- Instantiating and validating DPPs based on CT use cases.
- Identifying Lifecycle assessment scenarios and applying/validating XAI technologies introduced in other tasks.
- Work toward the collaboration of different DS connectors (e.g., TRUE and EDC)
- Constantly updating GitHub repository with new developments, extensions and documentation based on the different integration and validation needs.
- Further refining and testing the designed systems.

The data models, catalogues, and data exchange mechanisms will be subject to continual enhancement, while the identity and access management and applications & data orchestration will be developed to enhance user-friendliness and functionality.

A significant part of the upcoming work involves testing of these systems under the pilot scenarios. Our goal is to validate them, identify areas of improvement, and ensure they effectively support the pilot scenarios and the objectives of the Circular TwAIn Project. Lastly, we anticipate further engagement with pilot owners and technology provides to gather feedback and ensure the continued alignment of our work with the pilot and the other Circular TwAIn technical enabler needs.

References

[1] Alonso, Á., García-Pozo, A., Cantera, J. M., De La Vega, F. M., & Hierro, J. J. (2018b). Industrial Data Space Architecture Implementation Using FIWARE. Sensors, 18(7), 2226. https://doi.org/10.3390/s18072226

[2] Delgado, M. T. (2023). *Eclipse dataspace components*. projects.eclipse.org. https://projects.eclipse.org/projects/technology.edc

[3] TRUE Connector MVDS: https://github.com/Engineering-Research-and-Development/true-connector-mvds

[4] Engineering TRUE Connector https://github.com/Engineering-Research-and-Development/true-connector

[5] Eclipse-Edc. (2023, January 11). Samples/basic/README.md at main · eclipse-edc/Samples. GitHub. https://github.com/eclipse-edc/Samples/blob/main/basic/README.md

[6] Eclipse-Edc. (n.d.-a). Connector/extensions/common/iam/oauth2/oauth2-core at main eclipse-edc/Connector. GitHub. https://github.com/eclipse-edc/Connector/tree/main/extensions/common/iam/oauth2/oauth2-core

[7] Eclipse-Edc. (n.d.-a). GitHub - eclipse-edc/Connector: EDC core services including data plane and control plane. GitHub. https://github.com/eclipse-edc/Connector

[8] Eclipse-Edc. (n.d.-b). Connector/extensions/control-plane/api/management-api/catalog-api at main · eclipse-edc/Connector. GitHub. https://github.com/eclipse-edc/Connector/tree/main/extensions/control-plane/api/management-api/catalog-api

[9] Eclipse-Edc. (n.d.-b). GitHub - eclipse-edc/MinimumViableDataspace: Guidance on documentation, scripts and integration steps on using the EDC Project results. GitHub. https://github.com/eclipse-edc/MinimumViableDataspace/tree/main

[10] Eclipse-Edc. (n.d.-c). GitHub - eclipse-edc/MinimumViableDataspace: Guidance on documentation, scripts and integration steps on using the EDC Project results. GitHub. https://github.com/eclipse-edc/MinimumViableDataspace/tree/main#readme

[11] Eclipse-Edc. (n.d.-c). MinimumViableDataspace/deployment/data/MVD.postman_collection.json at main · eclipse-edc/MinimumViableDataspace. GitHub. https://github.com/eclipse-edc/MinimumViableDataspace/blob/main/deployment/data/MVD.postman_collection.json

[12] Eclipse-Edc. (n.d.-c). MinimumViableDataspace/system-tests/docker-compose.yml at main · eclipse-edc/MinimumViableDataspace. GitHub. https://github.com/eclipse-edc/MinimumViableDataspace/blob/main/system-tests/docker-compose.yml

[13] Eclipse-Edc. (n.d.-e). MinimumViableDataspace/system-tests at main · eclipse-edc/MinimumViableDataspace. GitHub. https://github.com/eclipse-edc/MinimumViableDataspace/tree/main/system-tests#readme

[14] Eclipse-Edc. (n.d.-f). MinimumViableDataspace/extensions at main · eclipse-edc/MinimumViableDataspace. GitHub. https://github.com/eclipse-edc/MinimumViableDataspace/tree/main/extensions

[15] FIWARE Foundation. (2021). FIWARE for Data Spaces [Position Statement]. https://www.fiware.org/wp-content/uploads/FF_PositionPaper_FIWARE4DataSpaces.pdf

[16] FIWARE. (n.d.). GitHub. https://github.com/FIWARE

[17] FIWARE. (n.d.-b). GitHub - FIWARE/tutorials.Getting-Started: :blue_book: FIWARE 101: An Introduction to the FIWARE Platform. GitHub. https://github.com/FIWARE/tutorials.Getting-Started/tree/master

[18] FIWARE. (n.d.-b). Step-by-Step for NGSI-v2. https://fiware-tutorials.readthedocs.io/en/latest/#list-of-tutorials

[19] Fiware-Orion. (n.d.). https://fiware-orion.readthedocs.io/en/master/

[20] European Commission, "Preparatory work in view of the procurement of an open source cloud-to-edge middleware platform: Architecture Vision Document", 24 March 2022, Available: https://ec.europa.eu/newsroom/dae/redirection/document/86241

[21] L. Nagel and D. Lycklama, "Design Principles for Data Spaces - Position Paper", IDSA 2021, Available: http://doi.org/10.5281/zenodo.5105744

[22] Sebastian Steinbuss, "IDS Reference Architecture Model Version 4", International Data Spaces Association 2022, Available: https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-ram-4/

[23] Electro and Digital Industry Association, "Implementing Level 2 of the ZVEI-ShowCase PCF@Control Cabinet", Whitepaper-2023, Available: https://www.zvei.org/fileadmin/user_upload/Themen/Industrie/PCF%40ControlCabinet/White-Paper_Implementing-Level-2-of-the-ZVEI-Show-Case.pdf

[24] Kai Garrels, Sten GrünerElectro, et al., "DPP 4.0: An Architecture Proposal for a DPPSystem to implement the EU Digital Product Passport for Industrial Products", Electro and Digital Association 2023, Available: https://www.zvei.org/fileadmin/user_upload/Themen/Industrie/PCF%40ControlCabinet/20230411_Discussion_Paper_DPP40.pdf

[25] Directorate-General for Environment, "Proposal for Ecodesign for Sustainable Products Regulation.", 30 Mar. 2022, Available: https://ec.europa.eu/environment/publications/proposal-ecodesign-sustainable-products-regulation_en

[26] Chris Stretton, "Digital product passports (DPP): what, how, and why?", Circularise, April 2022, Available: https://www.circularise.com/blogs/digital-product-passports-dpp-what-how-and-why

[27] European Commission, "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL concerning batteries and waste batteries, repealing Directive 2006/66/EC and amending Regulation (EU) No 2019/1020", Access: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0798

[28] IDSA, "Knowledge-Base: Minimum Viable Data Space (MVDS)", Available: https://docs.internationaldataspaces.org/knowledge-base/mvds, accessed 07/2023.

[29] Volz, F., Sutschet G., Stojanovic L., Usländer T, On the role of Digital Twins in Data Spaces. Sensors 2023, 23, 7396. Submitted for review, pending publication.

[30] Markus Spiekermann, “Eclipse Dataspace Connector – How to build Data Spaces”, Eclipsecon, 2021.

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