



Milestone 2.10 - Interoperability demonstrators with real-world use of data and different Levels of Service - 2nd round)

Annex A - Interoperable NAP demonstrator of a marketplace or clearing house
NAP

Status: final
Version: 1.2
Date: 11 March 2025

Legal disclaimer

The sole responsibility for the content of this document lies with the authors. It does not necessarily reflect the opinion of the European Union. The European Commission is not responsible for any use that may be made of the information contained therein. All images are provided by the respective partners (unless otherwise noted) and are approved for reproduction in this publication.

Document information

Project acronym	NAPCORE
Full project title	National Access Point Coordination Organisation for Europe
Grant Agreement No.	MOVE/B4/SUB/2020-123/SI2.852232
Activity no. and title	WG2, Task 2.4, Work item 2.4.3 - Cross border Marketplace Demonstrator
Author(s)	Ed Ooms, Arthur Keesen, National Access Point for Mobility Data (NTM)
Co-author(s)	
Related to Milestone no.	2.10
External Milestone	yes

Document history

Version	Date	created/ modified by	Comments
0.1	14.11.2024	Ed Ooms, Arthur Keesen, NTM	Initial draft
1.0	16.12.2024	Ed Ooms, Arthur Keese, NTM	Final draft
1.1	09.01.2024	João Montenegro	Review
1.2	11.03.2025	Hannah Walther	Final version after SCOM written procedure

Document status:

☐	To be revised by partners involved in the preparation of the document
☐	For review/ approval by the Core Alignment Team and Peer Reviewers
☐	For approval by the NAPCORE Steering Committee
☒	Final
☐	Updated/revised final version (e.g. when a second version of Milestone is published)

Abbreviations

Abbreviation	Meaning
EDC	Eclipse Data space Connector
DATEX	Data Exchange for Traffic Telematics
DEFLOG	Data Exchange for Logistics
DSSC	Data Spaces Support Centre
IDSA	International Data Spaces Association
IDS-RAM	International Data Space Referential Architecture
MS	Member States
NAP	National Access Point
NAPCORE	National Access Point Coordination Organisation for Europe
NDSS	NDW Data Space Society
NDW	Dutch National Data Portal for Road Traffic
NTM	Dutch National Access Point for Mobility Data
OTM	Open Trip Model
WG2	Working Group 2

Terms and Concepts

NAPCORE	NAPCORE (National Access Point Coordination Organisation for Europe) is an EU project aimed at improving the interoperability and coordination of NAPs across Europe.
NAPs	NAPs are digital platforms mandated by EU regulations to centralize and provide access to mobility-related data, such as traffic and travel information, to support Intelligent Transport Systems (ITS).
Federated data	Data managed, stored and provided by the original data holders
Centralized data	Data sent to a centralized location that provides, stores and sometimes aggregates or e.g. repairs data.
Serialization	The process of converting data/information into a format that can be easily stored, transmitted, or reconstructed later. This is essential for transferring and interpreting data between systems. E.g. Information can be serialized in XML, CSV or JSON format.

Table of contents

1. Introduction.....	6
1.1. Reading Guide.....	6
1.2. Context and background.....	6
1.3. Problem statement	7
1.4. Goals of the marketplace demonstrator.....	8
1.5. Business Value	8
1.6. Objectives and scope	8
2. Background.....	9
2.1. NAPCORE NAP Reference Architecture	9
2.2. International Data Space Reference Architecture	10
2.3. Data Space Support Centre Framework.....	10
3. Functional design	11
3.1. Actors & Roles.....	12
3.2. Definition of use cases	12
4. Functional architecture	14
4.1. IST architecture.....	14
4.1.1. Providing datasets.....	14
4.1.2. Publishing datasets	15
4.2. SOLL architecture	15
5. Technical Architecture	18
5.1. Introduction	18
5.2. Technical components	19
5.2.1. Eclipse Dataspace Connector	19
5.3. Scalability.....	22
5.4. Data Exchange Layer	22
5.5. Infrastructure Layer	22
6. Conclusions.....	23
6.1. Feedback to NAP's Reference Architecture and LoS Framework	23
6.2. Lessons learnt	23
6.3. Recommendations for Implementation	24
6.3.1. Growing the dataspace	24
6.3.2. Long-term Success Factors.....	24
7. Document References.....	26

List of tables



Table 1 - Chapters and their content	6
Table 2 - Actors & Roles in Use Cases.....	12
Table 3 - Technical Components	19

List of figures

Figure 1 - Cross border data for Road Works	7
Figure 2 - NAPCORE Reference Architecture - Functional view	9
Figure 3 - Demonstrator Use Cases	11
Figure 4 - 'Road Traffic Management' data provision in Germany and the Netherlands	14
Figure 5 - IST - Publish dataset on NAP	15
Figure 6 - Soll - publish dataset on NAP	16
Figure 7 - Functionality for onboarding.....	17
Figure 8 - DSSC Control plane & Data plane.....	18
Figure 9 - Eclipse Data Connector	20
Figure 10 - EDC Negotiation Connector	20
Figure 11 - EDC Secure Gateway	20
Figure 12 - Overview of Demonstrator Data space.....	21



1. Introduction

Working Group 2 (WG2) is concerned with Interoperability and the level of service of the National Access Points (NAPs) of Member States (MS). WG2 aims to enhance the compatibility of NAPs in the European Union.

This document describes the development of a cross-border ‘marketplace demonstrator’ as part of NAPCORE Work item 2.4.3 which is part of task 2.4.

In this document, the functional operation of the demonstrator is described in detail. Based on the functional operation, it is explained which technical solution has been developed.

Different frameworks will be examined to give a better understanding of how the demonstrator relates to functionality as described in different Digital Data Space documentation:

- International Data Space Referential Architecture (IDS-RAM [1]);
- Data Spaces Support Centre (DSSC [4]).

1.1. Reading Guide

In this document, abbreviations are used. These abbreviations are explained in the abbreviations chapter at the beginning of this document. In the Terms and Concepts chapter terms and concepts used in this document are explained to give a better understanding. These terms and concepts are printed in italics.

If an external document or site is used. You can find a list of Document References. A Document Reference is shown in brackets: e.g.[1].

Table 1 describes the way this document is organized.

Table 1 - Chapters and their content

§1 - Reasoning for the marketplace demonstrator	
§1.1 / §1.5	Introduction of context and background of the marketplace demonstrator
§2 - Use Cases	
§3 - Reference Architectures & Frameworks	
§3.1	Relation to NAPCORE Reference Architecture
§3.2	Relation to IDS-RAM
§3.3	Relation to DSSC
§4 - Functional Architecture	
§4.1	Current situation
§4.2	Desired situation
§5 - Technical Architecture	
§5.1 / §5.2	Technical components & interaction between components
§5.3	Scalability
§6 - Recommendations for implementation	

1.2. Context and background

NAPCORE coordinates NAPs across Europe. One of its purposes is to work on interoperability between NAPs, their data publications and the data being provided by the data publishers.

NAPCORE exists of different working groups with their own tasks. Each working group has its own subject and task groups carry out work for specific subjects. In task 2.4.3. the *NAPCORE* demonstrator for a cross-border marketplace is developed.

The demonstration must show that it is possible to use cross-border data, published on different *NAPs*, in an interoperable manner.

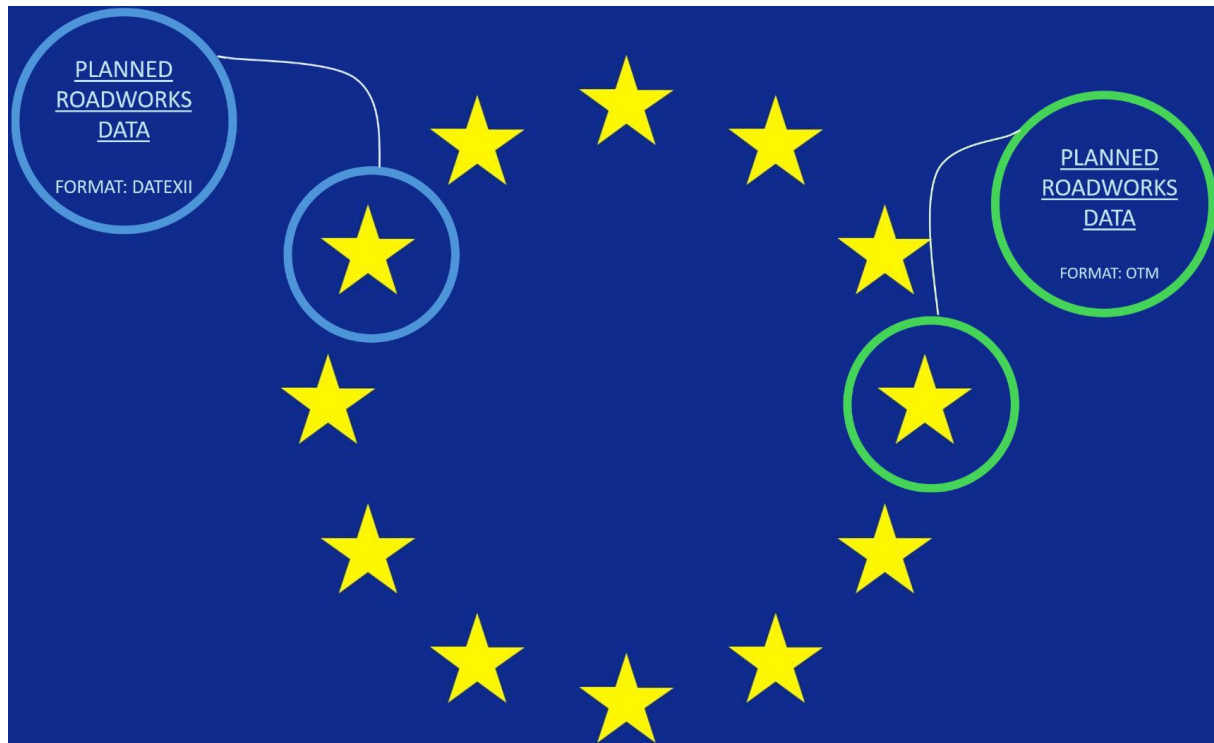


Figure 1 - Cross border data for Road Works

1.3. Problem statement

The global marketplace has seen significant growth in recent years, with cross-border logistics and traffic becoming increasingly common. This means that the mobility domain needs to take cross-border traffic into account, resulting in a higher need for cross-border data. However, cross-border data exchange comes with its own set of challenges, including regulatory and technical barriers.

This Demonstrator creates a cross-border planned roadworks data exchange between the Netherlands and Germany. The goal is to make the open dataset for planned roadworks in Germany available to logistics service providers in the Netherlands, so they can adjust their planning of both routing and deliveries based on this information.

Figure 1, in subchapter 1.2, shows data can be published on *NAPs* by different member states. Member states can choose to have a *NAP* that contains data (*Centralized data*). They can also choose to publish the data only and point to the location where data can be accessed (*Federated data*).

In the case of federated data, the data owner decides how the data is structured and *serialized*. In the case of centralized data, the data publisher may be able to change the way the data is *serialized*. This leads to some challenges.

At the level of the *NAPs* terms of use are available for the published dataset. To be able to negotiate the contract between data publisher and data consumer, extra conditions need to be specified and available. This could be made easier if both participants are in a dataspace. To use (e.g. cross border) data the data provider and data consumer need to agree on:



- Make agreements with several parties that provide the data:
 - What certification is needed to make secure data transactions?
 - If identification is needed; what method is used?
 - If payments/credits are required; how does the payment obligation work?
 - What has been arranged for auditing?
 - Etcetera.
- The manner of deserializing data (for example, DATEXII) and reserializing data (OTM) to be able to interpret the data in an interoperable way.

1.4. Goals of the marketplace demonstrator

The following specific goals can be listed:

- Explore solutions to overcome the challenges of European-wide metadata and data sharing;
- Promote collaborative metadata and data sharing practices among different member states and their NAPs;
- Make relevant metadata and real-time data accessible to all interested parties through any NAP;
- Present a potential solution, by leveraging existing technology to facilitate sharing of metadata among NAPs in Europe;
- Provide a technical solution to how this type of interoperability can be achieved, for example in the form of suggested minimum requirements for the reference architecture;
- Provide input to the NLKF measuring the level of service.

1.5. Business Value

Currently, planned roadworks data for the Netherlands is available to logistics service providers through the DEFLog (Data Exchange for Logistics) marketplace. However, this data does not include planned roadworks for Germany.

There are a number of benefits to enabling logistical service providers to access planned roadworks data for both the Netherlands and Germany. This would allow them to:

- Improve their route planning and avoid delays caused by roadworks.
- Reduce their fuel consumption and emissions.
- Provide better customer service by being able to communicate more accurately about expected delivery times.
- Enhance cross-border logistics efficiency.

1.6. Objectives and scope

The primary goal of this demonstrator is to showcase the use of marketplace functionality that facilitates a secure and efficient exchange of data between Dutch data consumers and German data providers. It aims to make planned roadworks data from Germany accessible to logistical service providers in the Netherlands through the DEFLog marketplace. Additionally, the demonstrator seeks to identify challenges related to cross-border data sharing and monitoring agreements, particularly concerning potential breaches by either the data provider or the data user.

2. Background

To enhance interoperability, standards and reference architecture should be used. For this demonstrator two reference architectures are used and one methodology.

2.1. NAPCORE NAP Reference Architecture

In WG2, Task 2.3, work is done to develop a NAP reference architecture [6], whose functional view is depicted in Figure 2. This architecture is developed to help member states have an overview of which functions should be provided by the NAPs. Also, a translation to the physical world – the components/applications that need to be in place – is made.

In this document, references to this architecture will be made by using the codes in italics (e.g. *F3.6.5*) in the text.

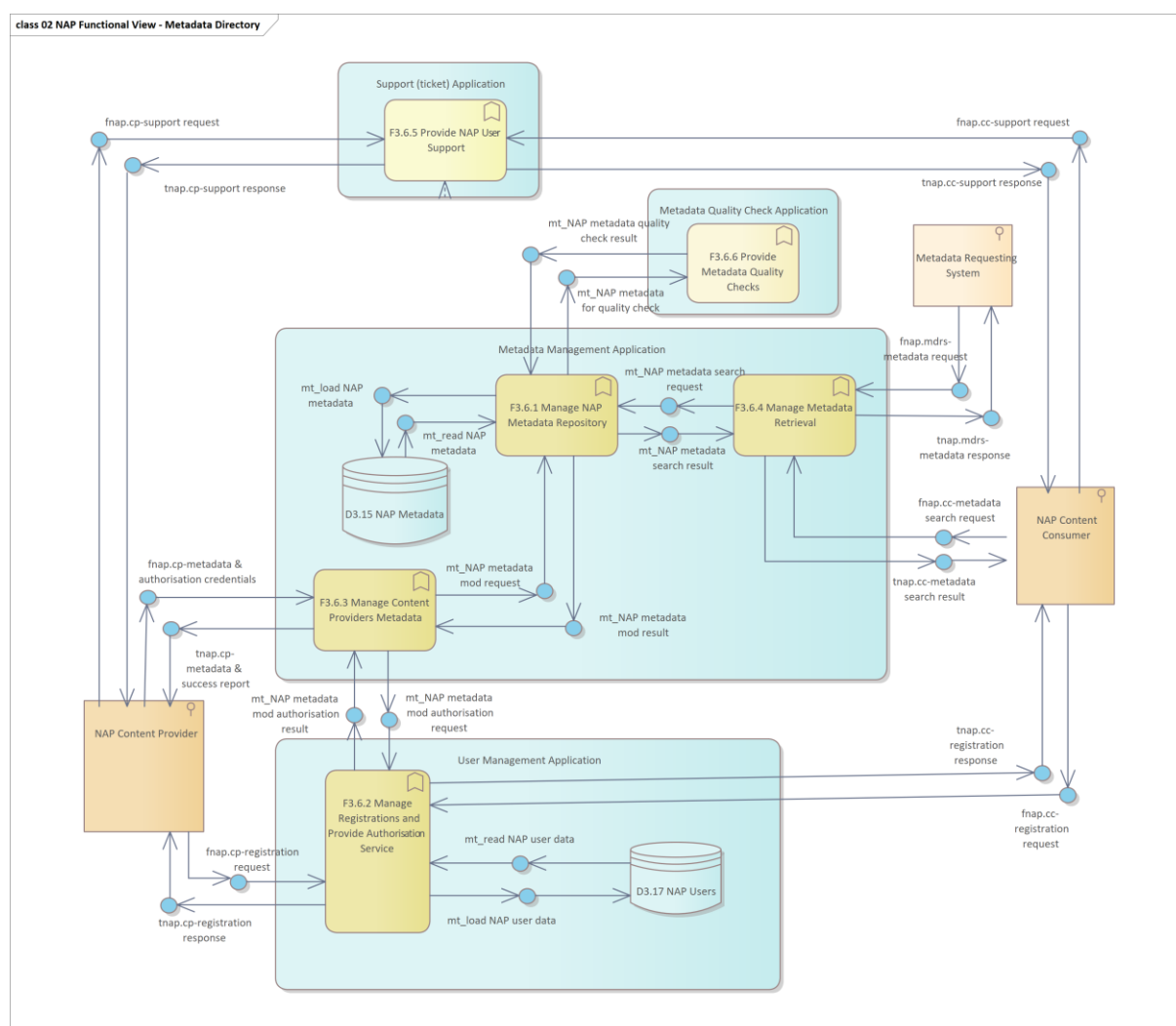


Figure 2 - NAPCORE Reference Architecture - Functional view

2.2. International Data Space Reference Architecture

The International Data Spaces Reference Architecture Model [1] (IDS-RAM) is a conceptual framework developed by the International Data Spaces Association (IDSA). It provides a standardized approach for enabling secure, interoperable, and sovereign data sharing among organizations while maintaining control over data usage. The IDS-RAM supports the creation of data ecosystems where participants can share and exchange data in a trusted and compliant manner.

In this document references to this architecture will be made by using footnotes (when references are made in diagrams) or in the text in italic (e.g. *source:IDS-RAM* [1]).

2.3. Data Space Support Centre Framework

The Data Space Support Centre [4] (DSSC) is an initiative of the European Union designed to help organizations develop and operate data spaces. These are ecosystems where stakeholders can securely share and collaborate on data while respecting privacy, intellectual property, and sovereignty.

Some terms in the solution architecture are taken from the DSSC.

In this document references to this framework will be made by using a reference in the text in italic (e.g. *source:DSSC* [4]).

3. Functional design

These use cases describe the desired (to be) functionality for the data consumer. In Figure 2 an overview of use cases is available. Each use case will be explained on the following pages.

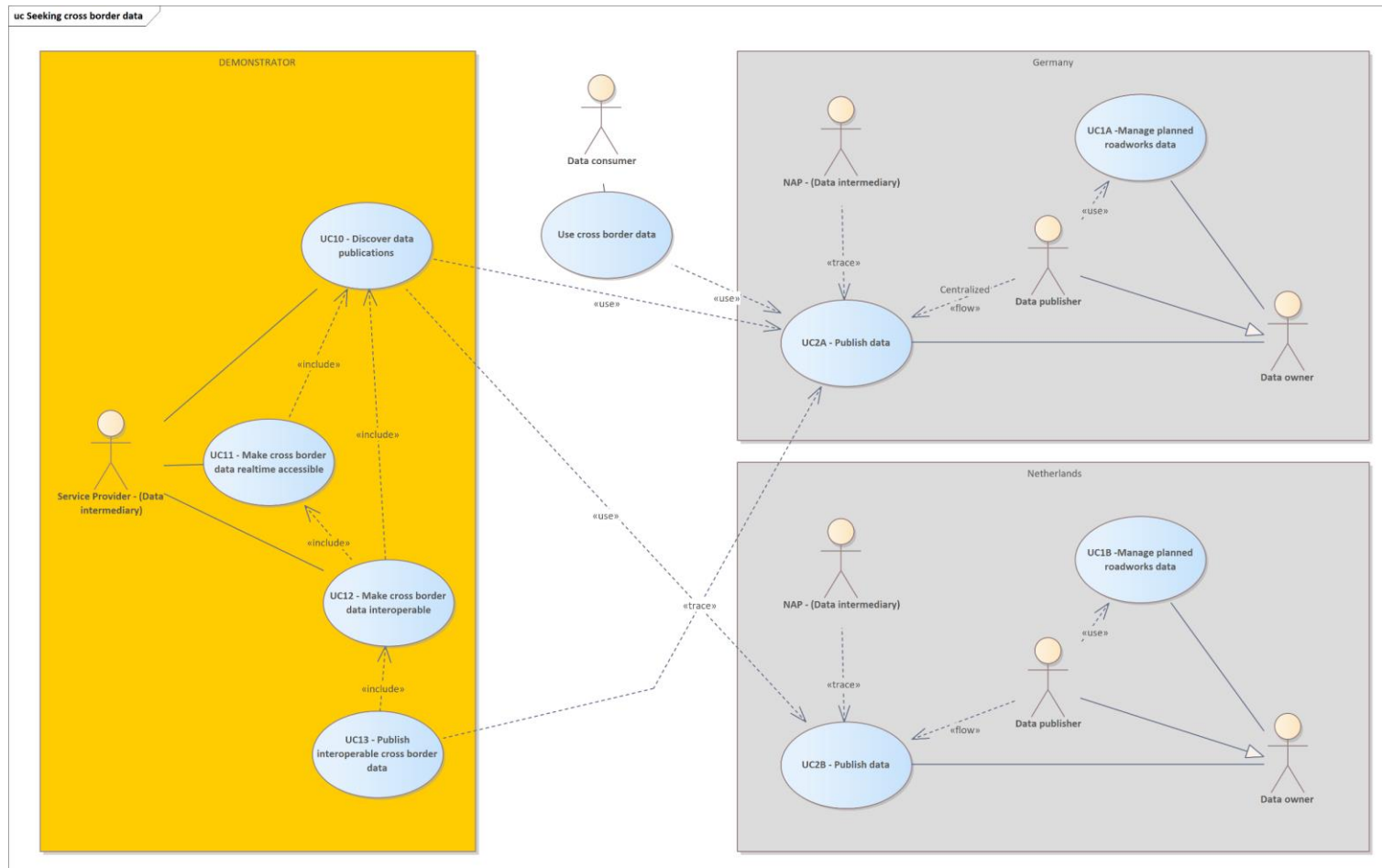


Figure 3 - Demonstrator Use Cases

3.1. Actors & Roles

Table 2 - Actors & Roles in Use Cases

Data owner¹	Legal entity or natural person executing control over data.
Metadata Broker¹	Stores and manages information about the data sources available in the International Data Spaces.
Data intermediary¹	Act as trusted entities and are commonly considered as "platforms". They assume a rather central role compared to the great number of data suppliers and customers, though multiple, especially competitive platforms of the same role may and shall exist.
Data consumer¹	The Data Consumer receives data from a Data Provider. From a business process modelling perspective, the Data Consumer is the mirror entity of the Data Provider; the activities performed by the Data Consumer are therefore similar to the activities performed by the Data Provider.
Service provider¹	Receives data from a Data Provider (or another Service Provider) and either returns the calculation result to the same or directs it to an indicated Data Consumer (which then is a Service Consumer at the same time). The participant who receives processed data from a Service Intermediary could be again a Service Intermediary as data can be routed through an arbitrary number of instances of services in the IDS.
NAP²	Facilitate access, easy exchange and reuse of transport related data, in order to help support the provision of EU-wide interoperable travel and traffic services to end users.

3.2. Definition of use cases

UC1A – Manage planned roadworks data DE	
Abstract	The data owner ensures data is accurate, secure, and used appropriately in accordance with organizational policies and applicable regulations.
Precondition	Data of different roadwork organisations is available in Germany.
Postcondition	Data is centrally available at the NAP of Germany.

UC1B – Manage planned roadworks data NL	
Abstract	The data owner ensures data is accurate, secure, and used appropriately in accordance with organizational policies and applicable regulations.
Precondition	Data of different roadwork organisations is available in the Netherlands.
Postcondition	Data is centrally available to publish to the NAP of the Netherlands.

UC2A – Publish data DE	
Abstract	The data publisher is adding information about the dataset (metadata), to publish the data the German National Access Point.
Precondition	Data is available at a data provider (UC1A).
Postcondition	Data is available and published with dataset information at the National Access Point of Germany.

¹ Roles according to the IDS-RAM [3]

² Source: EU - National Access Points [5]

UC2B – Publish data NL	
Abstract	The data publisher is adding information about the dataset (metadata), to publish information about the dataset at the National Access Point of the Netherlands.
Precondition	Data is available at a data provider (UC1A).
Postcondition	Data is published with dataset information at the National Access Point of the Netherlands.

UC10 – Discover data publications	
Abstract	A service provider is discovering data publications.
Precondition	Data is published at a National Access Point (UC2A + UC2B).
Postcondition	An overview of dataset publications is available.

UC11 – Make cross border data real time accessible	
Abstract	A service provider is actively obtaining data in (near) real time.
Precondition	Data is available at a number of data providers (UC1A + UC2A).
Postcondition	Data is obtained and available to share at one or more NAPs

UC12 – Make cross border data interoperable	
Abstract	A service provider uses methods to transpose DATEXII information to OTM information.
Precondition	Data is available at a data provider (UC1A).
Postcondition	Data is published in a standardized manner (OTM) with dataset information at the National Access Point of the Netherlands.

UC13 – Publish interoperable data to NAP	
Abstract	A service provider publishes the transposed dataset (UC12) to the National Access Point.
Precondition	Datasets are available at the NAPs or at the Service Provider in OTM format.
Postcondition	Dataset is published at the National Access Point(s).

UC20 – Use Cross border dataset	
Abstract	A data consumer can use the cross border dataset.
Precondition	Datasets are published (UC13) in OTM format.
Postcondition	The data consumer can use cross border data published on one or more NAPs in a standardized format (OTM).

4. Functional architecture

4.1. IST architecture

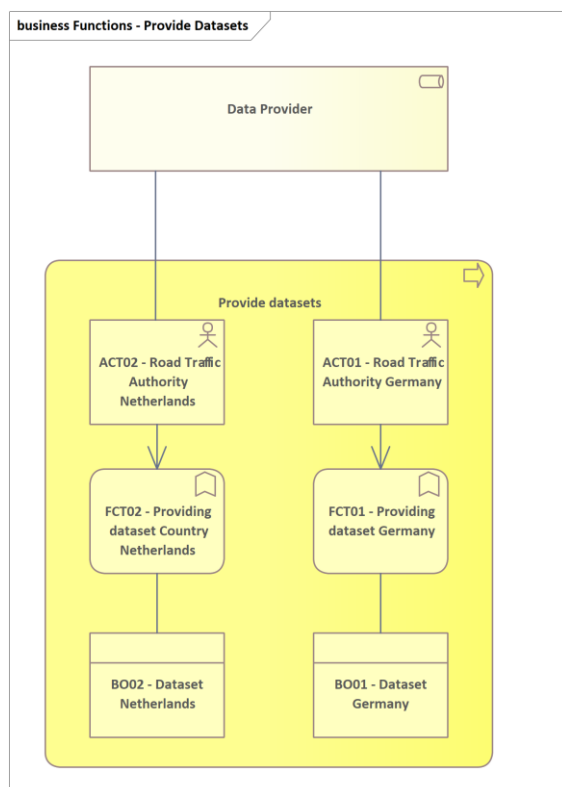
Currently, datasets are provided through data providers³. Their datasets should be published on NAPs. At the NAP there are (broadly) two ways of working:

1. The NAP can be used to consume data directly (Centralized data, like the German Mobilithek [7]);
2. The NAP refers to a location data can be consumed (Federated data, like the Dutch NTM[8]).

In case the data consumer wants to use cross-border data, the data consumer needs to know where to find the data for each dataset. Also, the data provider and the data consumer need to make agreements on:

- The level of trust needed between data provider and data consumer to consume the data
- The methods used for information exchange
 - Securing connections
 - Certifying connections through certificates
 - Authenticating users (or not)
 - Authorizing users (or not).
- Etcetera.

4.1.1. Providing datasets



For this demonstrator functions available in the IDS-RAM³ – are used as a starting point.

In this figure⁴ and ⁵, the Road Traffic Authority of Germany (ACT01) and the Traffic Data Portal (NDW) of the Netherlands (ACT02) provide two separate datasets (BO01/BO02) containing information about 'Road Traffic Management'.

Figure 4 - 'Road Traffic Management' data provision in Germany and the Netherlands

³ source:IDS-RAM[1]

⁴ ACT = Actor, FCT = Function, BO = Business Object, COL = Collaboration

⁵ Light yellow items are formal roles and functions taken from the International Data Space Reference Architecture Functions[1] and Roles [2]

4.1.2. Publishing datasets

Each dataset is published on different National Access Points.

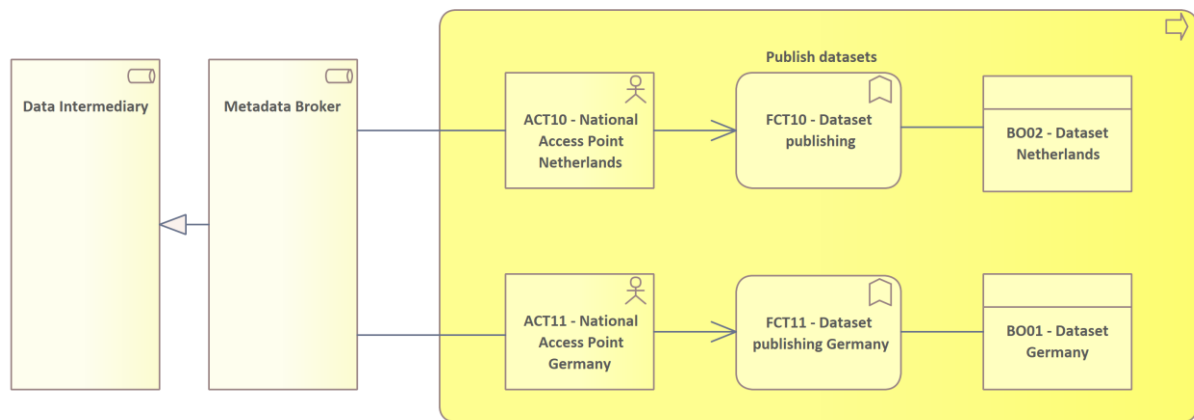


Figure 5 - IST - Publish dataset on NAP

4.2. SOLL architecture

In the target (SOLL) scenario, a set of Data Space functions will be implemented. As outlined in Sections 3.2 and 3.3, reference architectures and frameworks are utilized to work out the functions to be incorporated into this demonstrator. This approach enables organizations to leverage the ideas and solutions demonstrated here. Additionally, it provides a foundation for extending functionality to enhance their own data spaces.

REQ01 - Have a singular method to find published data provided by different data providers

Functionality is required (FCT20 in the green box in Figure 6) to allow data consumers to find cross-border data at the NAP of their choice. The 'no wrong door' principle should apply. Some kind of discovery service should be available to find datasets. This discovery service should result in an overview of all datasets published on all NAPs within the scope of the data space.

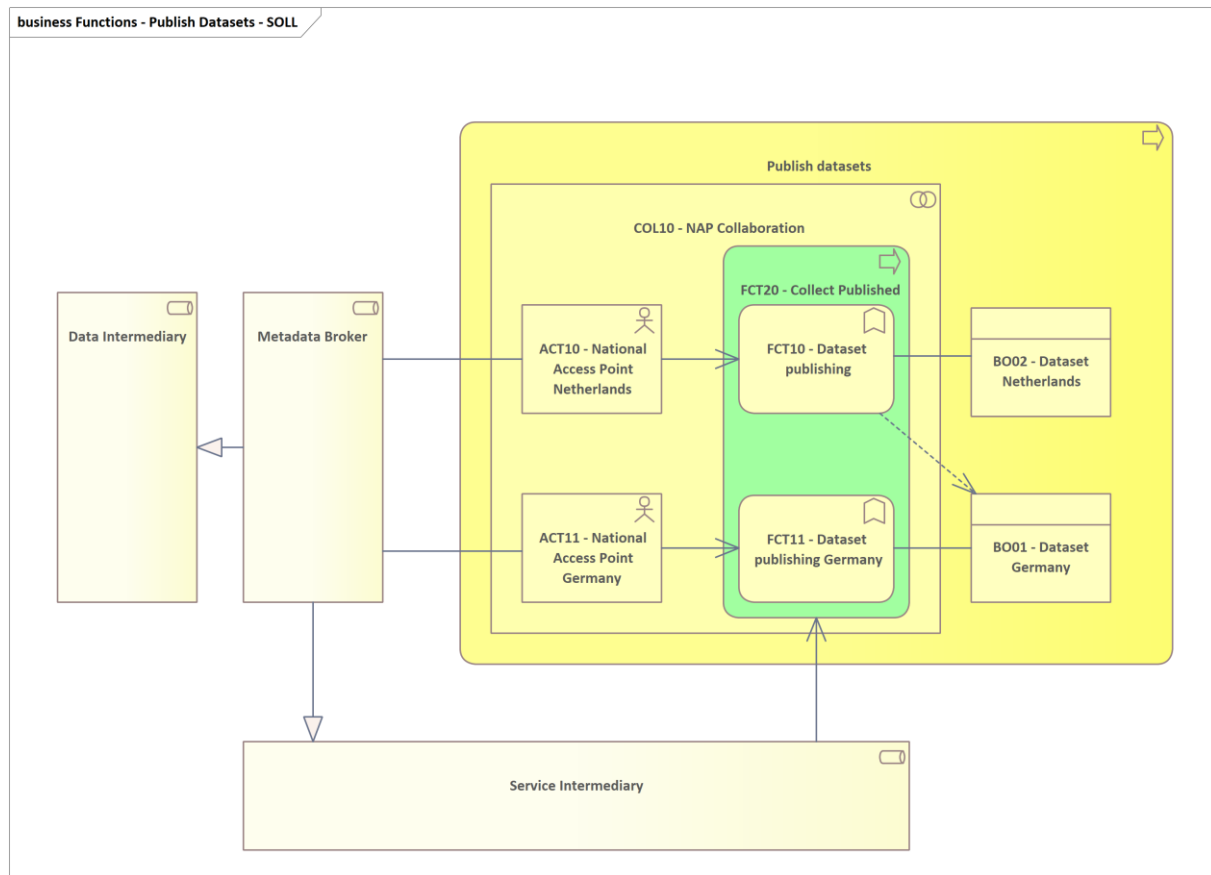


Figure 6 - Soll - publish dataset on NAP

REQ02 - Reduce the amount of point-to-point agreements to be made to retrieve the data

Dataspaces are about creating an environment where published data from different data providers can be accessed by a data consumer in a standardized manner. Agreements need to be made with one organisation (Service Intermediary) to determine what type of entity is consuming data (data consumer) and what the conditions are for authentication, authorization, securely exchanging data, etcetera. Once these agreements are made a policy will be set up to determine what information is available from different (future) data providers.

Dataspaces have many functions. In this demonstrator the following functions are in scope:

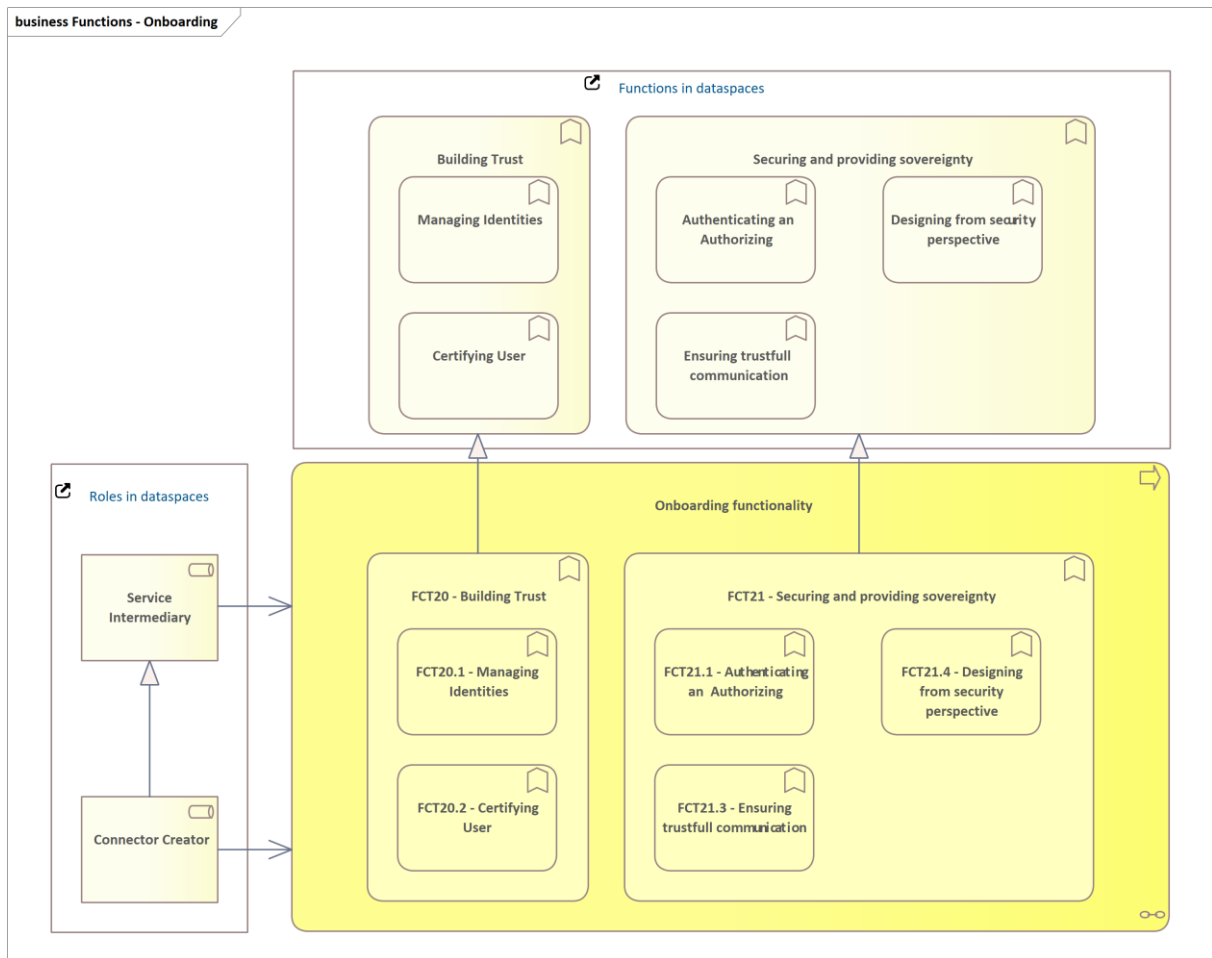


Figure 7 - Functionality for onboarding

5. Technical Architecture

5.1. Introduction

The data Spaces Support Centre (DSSC[4]) distinguishes three types of software layers; the control plane, the data plane and the identity provider. These terms will be used in the following chapters.

Control plane

The *control plane* is responsible for deciding how data is managed, routed and processed.

Data plane

The *data plane* is responsible for secure exchange of data between data provider > data space > data consumer and vice versa.

Identity Provider/Manager

The *data plane* uses standardized methods to identify the data consumer and provides proof of the identity of the data consumer to the *control plane*.

The terms 'control plane, data plane and Identity provider' are used in some images in this document.

Figure 8

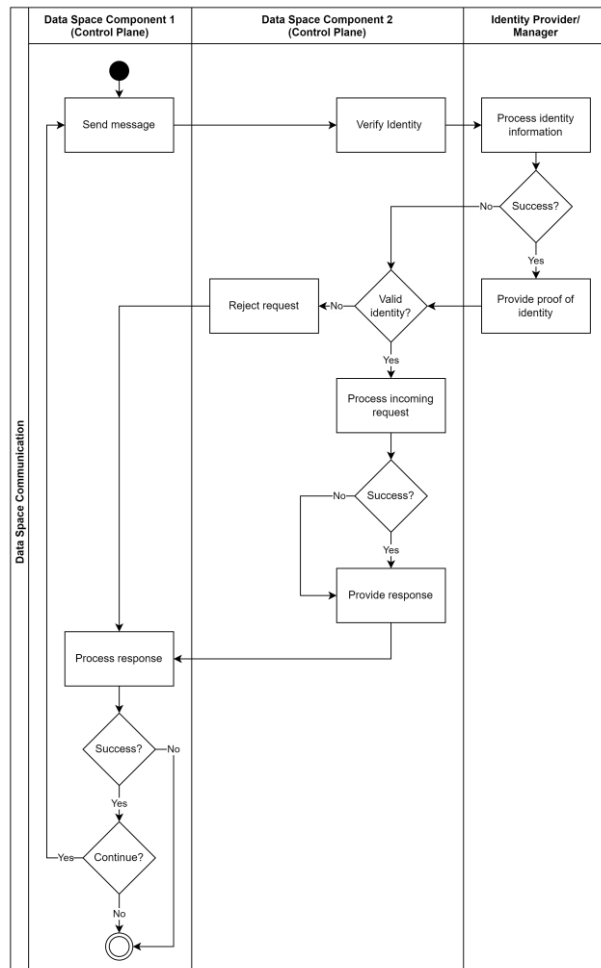


Figure 8 - DSSC Control plane & Data plane

5.2. Technical components

For the setup of this demonstrator, NTM aimed to make as much use as possible of standard components from existing data spaces. The Eclipse Tractus-X open-source project from the Catena-X data space was used for this demonstrator.

Due to the large number of active users, NTM is confident that reusing these off the shelf components ensures the solution will be scalable and easy to deploy in the future.

5.2.1. Eclipse Dataspace Connector

An important component of the demonstrator is the Eclipse Dataspace Connector (EDC). This connector is implemented for each data provider and each data consumer. In the table below a description is given for EDC, its components and component functions:

Table 3 - Technical Components

Component	Functions	Description
EDC	Framework to assure secure and compliant data exchange according to policies	The EDC is an open-source software project designed to enable secure and controlled data exchange within data spaces. It is a critical component for implementing data spaces that comply with frameworks such as GAIA-X and the European Data Spaces initiatives like IDS-RAM.
Catena-X data space	End-to-End Data Traceability Standardized Data Models Decentralized Architecture ⁶ Integration with Ecosystem Standards Modularity and Extensibility	The Catena-X data space is a collaborative ecosystem designed for the automotive industry. It facilitates secure, standardized, and interoperable data sharing among stakeholders across the automotive value chain, including manufacturers, suppliers, logistics providers, and service partners. Supports interoperability with related initiatives like Eclipse Tractus-X, which provides the technical backbone for the Catena-X Data Space. At the moment, 180 parties were connected to this data ecosystem.
Catana Tractus X	Standardized interfaces Policy enforcement Real time data provisioning Enriching provisioned data	Eclipse Tractus-X is an open-source project under the Eclipse Foundation, developed to create a collaborative data ecosystem for (originally) the automotive industry. It supports the principles of data sovereignty, transparency, and interoperability, enabling stakeholders across a value chain to securely share data and collaborate efficiently. It is aimed to digitize the end-to-end value chain.

⁶ Operates without a central authority, ensuring trust and security through decentralized governance and technologies

In Figure 9, an overview of the EDC application landscape is available:

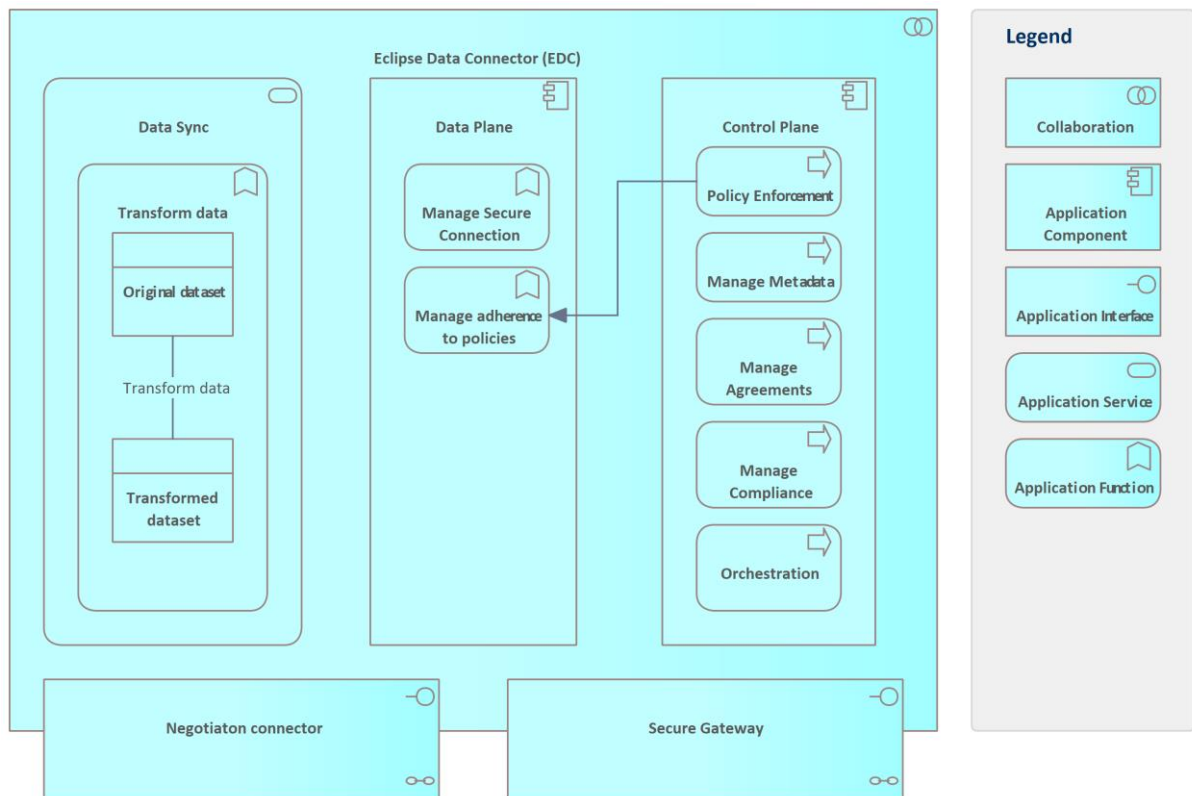


Figure 9 - Eclipse Data Connector

The connectors at the bottom of the above diagram are shown in more detail in Figures Figure 10 and Figure 11.

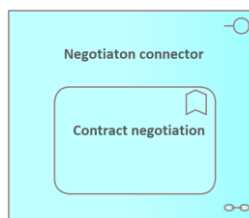


Figure 10 - EDC Negotiation Connector



Figure 11 - EDC Secure Gateway

Figure 12 shows the EDC framework in relation to the dataspace solution in its entirety. It shows the various parties involved and the components that need to be configured.

In blue are the reusable Eclipse Tractus-X components that need to be set up/configured, specifically the IAM Keycloak, Managed Identity Wallet, Discovery Services, and the Eclipse Data Space Discovery (EDC) for both data providers and data consumers.

In black, are additional components, specifically for this use case: an NDW data transformation pipeline, where data (purple block) from Mobilithek is retrieved twice a day and converted to OTM, and the main NAP function of the Dutch NAP, acatalogue for users to view the available datasets and metadata. All components are published open source in the NAPCORE Git repository [11].

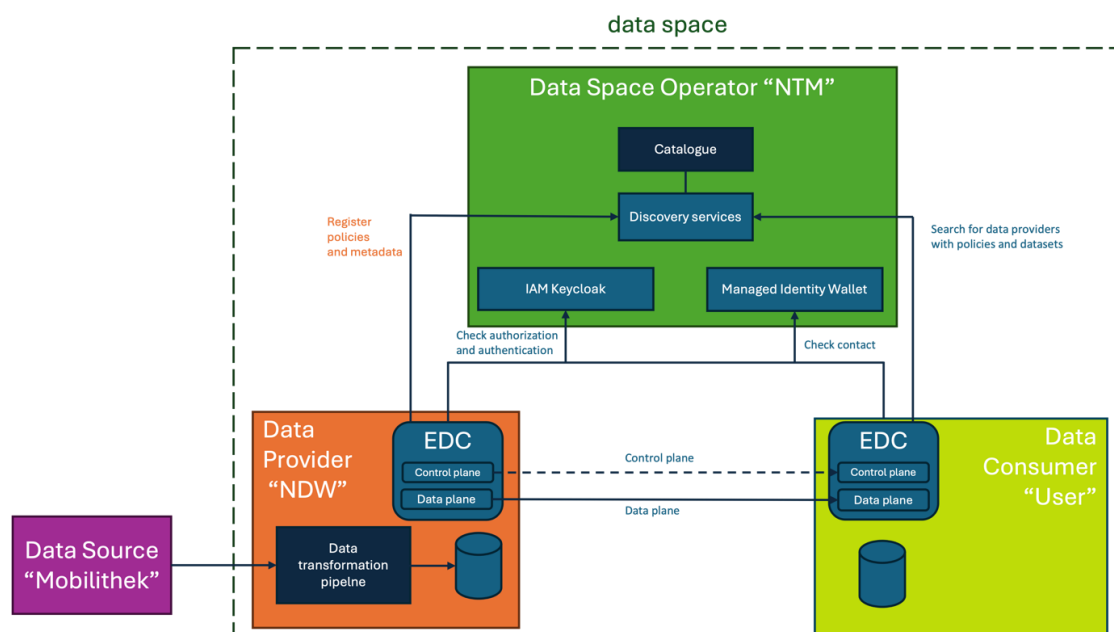


Figure 12 - Overview of Demonstrator Data space

Dataspace Role

For this demonstrator, NTM is carrying out a set of functions in the demonstrator Data Space. NTM has a responsibility for the authorization (trust) of users by ensuring the identification⁷ of parties and authentication through the wallet. NTM also runs a discovery service which – for this demonstrator - is configured to ‘harvest’ Mobilithet for publications regarding road maintenance.

Also, NTM hosts a frontend for the demonstrator dataspace Catalogue, allowing data customers to find datasets that are published (cross-border) in the dataspace.

Data Provider & Data Publisher

NDW has the roles of the data provider and data publisher (*source: IDS-RAM Roles[2]*), making data regarding to German roadworks available in OTM at the Dutch NAP⁸. Therefore Subsequently, the data will be made available for end users at the NDW EDC implementation.

Data Customer Role

The end user is a data customer (*source: IDS-RAM Roles[2]*) requesting access through the data space. Upon approval, the data customer could, through their own EDC, exchange contract information and retrieve data accordingly from the data provider.

The key aspect of this data space setup, according to the principles of federated data sharing, is that data providers, on behalf of the data owners, always remain in control of access to their data (data sovereignty - *source: IDS-RAM Functions[3]*) because they can manage this within their own connector.

Data providers can revoke access from users and terminate contracts through the EDC connector. The registration and establishment of contracts occur via the EDCs, over which the operator has no control. The operator can, however, revoke access for participants to the entire data space.

⁷ The open source identity and access management solution IAM Keycloak [10]

⁸ Publication discovered at the MobilityDCAT-AP[9] API at Mobilithet and data fetched.

As part of the demonstrator, a template with an example configuration for the EDC connector is made available [11]. Parties looking to set up an EDC are supported in this way to easily for configure tasks such as registering and managing agreements. This approach minimizes implementation effort.

Additionally, a proxy EDC connector is possible, where a third party handles hosting and/or maintenance. This solution remains fully aligned with the developments of Catena-X.

Since the exchange of data-and therefore any potential logging of dataset queries - occurs via the data plane between the EDCs, the operator (NTM) does not have standard access to view this logging. However, data space conditions can require that data providers make logging available to the data space operator. This can also be directly configured in the template EDC connector.

The direct 1-to-1 forwarding of data by users can be legally addressed within a data space, as the exchange of data via the EDC involves a contract with terms of use specific to that data. These terms can state that direct 1-to-1 forwarding without modifications is not permitted.

5.3. Scalability

The components being used in this demonstrator are all scalable. Now only one published dataset is made available in the demonstrator dataspace to proof seamless operation of all dataspace components. In future any other published dataset can be onboarded to the dataspace by carrying out onboarding procedures and configure the components accordingly.

5.4. Data Exchange Layer

The data exchange layer facilitates secure data transfer between participants through the Eclipse Dataspace Connector (EDC) implementations. This layer comprises two primary components: the EDC Consumer and EDC Provider instances.

The EDC Consumer component manages the data consumption process, handling contract negotiations and policy verification whilst ensuring secure data receipt. Conversely, the EDC provider component publishes data offers, enforces access policies, and manages data transfer operations. Together, these components ensure reliable and secure data exchange whilst maintaining data sovereignty.

5.5. Infrastructure Layer

The foundation of the architecture rests upon Kubernetes Service, which provides the containerisation and orchestration capabilities essential for modern cloud-native applications. In this demonstrator the Azure Kubernetes Service is used and manages the deployment, scaling, and operation of application containers across a cluster of hosts, ensuring high availability and efficient resource utilisation.

Kubernetes is a portable, extensible, open-source platform for managing containerized workloads and services. Therefore, it can be used on other cloud platforms as well. This approach ensures consistency across environments whilst providing the flexibility to scale individual components as needed. The microservice architecture also allows scaling and updating of components, whilst the distributed security model ensures robust protection of data assets.

The separation of data and metadata management provides additional flexibility, allowing for efficient data discovery without compromising data security or sovereignty. This architectural decision supports the dataspace's goal of facilitating data exchange whilst maintaining strict control over access and usage.

6. Conclusions

6.1. Feedback to NAPs Reference Architecture and LoS Framework

It would be advisable to consider the extent to which NAPs should perform certain functions within a data space. This can be approached through a path of gradual development. Starting with simpler implementations (e.g. providing a secure connector) and progressively expanding toward a comprehensive dataspace allows for smoother adaptation. This phased approach helps market participants become accustomed to the concept of entrusting their data securely within the dataspace framework.

It is also recommended to adopt widely accepted architectures and frameworks. Within the scope of the dataspace domain, selecting a single architecture or framework would be beneficial. This choice would help standardize the terminology and technology used, fostering consistency and facilitating seamless integration within the dataspace ecosystem.

Within NAPCORE, a clear understanding should be established regarding how the concept of being a dataspace should be interpreted at the European level and which functions the dataspace should fulfil. It is also advisable to describe de facto standards to facilitate the development of a dataspace. Once this has been clarified, it should be incorporated into the NAP Reference Architecture[6] and the LoS Framework.

6.2. Lessons learnt

The demonstration revealed several key insights:

Use of Open-source Tools: The use of solutions like Catena-X and Tractus-X significantly accelerated the implementation of a dataspace, providing reliable components for secure data exchange. Catena-X offers a comprehensive framework that includes standardised protocols and tools for data sharing. This meant we didn't have to build these fundamental elements from scratch. The platform's architecture ensures that data remains secure and that organisations maintain control over their shared information.

Tractus-X complements this by providing open-source components that handle crucial aspects such as:

- Identity and access management
- Data encryption and security measures
- Standardised data formats and exchange protocols
- Consent management mechanisms
- Audit logging capabilities

By building upon these existing solutions, we could focus on adapting and configuring the components to meet our specific requirements, rather than developing the entire infrastructure ourselves. This approach not only saved considerable development time but also ensured we were working with battle-tested solutions that follow industry best practices for secure and reliable data exchange.

Standardisation Across Regions: A key finding from our project is that significant differences exist between how Germany and the Netherlands implement DATEXII. While both countries technically use the same base version of DATEXII, their practical implementations vary considerably. These variations became apparent during development, particularly in how each country interprets and uses specific DATEXII elements. A notable example is the different approach to location referencing using alertC.

These differences made it impossible to perform a straightforward, one-to-one conversion to the desired OTM format, which was specifically designed around the Dutch implementation of DATEXII.

Standardisation of OTM: During the software development of the Demonstrator, a translation was needed from the DATEXII version used in Germany for roadworks data to a Dutch-compatible OTM format. For the requested OTM5, it appeared that no established standard existed yet, which caused problems when converting the data source. These issues were resolved by making our own assumptions about the correct structure for OTM5.

Scalability and Interoperability: By using standardised components like the EDC, we have established a scalable and flexible dataspace that supports interoperability between different sectors and regions. The solution's scalability manifests in several important ways.

First, new participants can be easily added to the network without complex technical adjustments.

Additionally, the system is designed to incorporate additional data services as they become necessary or available. The infrastructure can seamlessly handle increasing data volumes as more organisations join and share information. Furthermore, the architecture supports new use cases without requiring major modifications to the existing setup. Finally, the underlying infrastructure can be expanded based on demand, allowing for organic growth as the network's requirements evolve.

This scalable design ensures that the dataspace can grow alongside the needs of its participants, whether that growth comes from increased usage, new participants, or evolving requirements. This future-proof approach helps protect the initial investment in the infrastructure while providing confidence that the system can adapt to future needs.

6.3. Recommendations for Implementation

6.3.1. Growing the dataspace

Based on experiences from this successful implementation, it is recommended to adopt a phased approach to grow the dataspace implementation. Begin with the determination of core functionality and gradually expand this functionality based on participant needs and feedback. This approach helps manage complexity and ensures stable operation. Start small to allow participants to gain trust in the dataspace.

Establishing strong governance and work from the 'security by design' is crucial in the way of working from the start. This includes clear definitions of roles, responsibilities, well-documented procedures, and regular review processes. These elements help ensure smooth and trustful operation and facilitate rapid resolution of any issues that arise. Pay special attention to a feedback loop and the processes behind this loop for participants in the dataspace. Regular communication with participants helps ensure their needs are met and any issues are quickly identified and addressed. Establish clear communication channels and support procedures and maintain detailed documentation of all aspects of the marketplace implementation and operation.

6.3.2. Long-term Success Factors

For long-term success, maintain focus on continuous improvement of both technical infrastructure, security and operational processes. Regular reviews of system performance, security measures, and participant feedback help identify areas for improvement. This ongoing refinement helps ensure the marketplace continues to meet participant needs effectively.

Investing in participant support and community building helps drive adoption and effective use of the marketplace. Regular training sessions, user groups, and knowledge-sharing opportunities help build a strong participant community and facilitate effective use of the marketplace.



7. Document References

Nr	Date	Version	Description	Documentation
1	02.12.2024	4.0	International Data Space Reference Architecture	IDS-RAM
2	20.11.2024	4.0	Functions in dataspace (IDS-RAM)	Functions - IDS-RAM v4.0
3	20.11.2024	4.0	Roles in dataspace (IDS-RAM)	Roles - IDS-RAM v4.0
4	12.09.2023	-	DSSC	Control plane vs. Data plane
5	02.12.2024		Definition National Access Point	EU National Access Point
6	17.04.2024	1.1	Harmonization of EU NAP architectures and first layout of potential NAP federation	Reference Architecture
7	03.12.2024	-	German National Access Point	Mobilithek
8	03.12.2024	-	Dutch National Access Point	Nationaal Toegangspunt Mobiliteitsdata
9	04.04.2024	1.0.1	MobilityDCAT-AP standard	MobilityDCAT-AP recommendation
10	04.12.2024	26.0.7	Keycloak - Identity and Access Management	IAM Keycloak
11	10.12.2024	-	Github repository NAPCORE – Demonstrators	Demonstrators



Subannexes

Subannex 1 Process Overview and Phases

The process for the dataspace is divided into three main components: **Company Onboarding**, **Data Provisioning**, and **Data Consumption**. Each component follows specific steps to establish a reliable and secure dataspace environment. In the Demonstrator only step 3 is supported.

Company Onboarding

During this phase, the dataspace operator initiates a formal onboarding process for new participants:

- **Invitation and Registration:** Potential companies receive an invitation from the operator to participate in the dataspace. They complete a registration form, which is assessed to validate the identity and reliability of the participant.
- **Validation Process:** A validation process is carried out to check if the participant meets the conditions for joining, such as technical standards and security requirements.
- **Access to the Dataspace:** Only verified and approved participants are granted access to the dataspace, contributing to a secure and trusted environment.

Data Provisioning

After the onboarding phase, data providers focus on making data available within the dataspace:

- **Data Asset Registration:** Data providers identify and register their datasets within the dataspace, including key metadata such as origin, ownership information, and accessibility criteria. The data is registered as a data asset, which represents a specific dataset in the provider's backend.
- **Policy Rules and Contracts:** Policies are established that define who can access the dataset, how long the data can be retained, and under what conditions it can be used. These policies are enforced by the Eclipse Dataspace Connector (EDC).
- **EDC Configuration:** The EDC is configured to ensure secure data exchange. This includes settings for access management, defining contract terms, and establishing data transfer guidelines.

Data Consumption

In the consumption phase, data customers gain access to data within the dataspace:

- **Data Search and Viewing Contracts:** Data customers can search for datasets through a central catalogue hosted outside the dataspace. Here, they can view contract terms and decide on access to specific datasets.
- **Data Requests and Contract Negotiations:** Once a dataset is identified, the data customer can submit a contract request. The EDC supports contract negotiations by indicating if a consumer meets the conditions for access. Each dataset has policies that define the applicable terms.
- **Data Transfer:** Once a contract is concluded, the consumer is granted access to the dataset via a secure endpoint. Data exchange is controlled by the EDC, ensuring that data is accessible only to authorised users.

Subannex 2 Technical Architecture and Process Overview

Dataspace Process Steps with Mobilithek Dataset

These are the detailed steps of the demonstration process, with Mobilithek as the primary data service:

1. **Data Collection and Conversion**
The data provider, NDW, retrieves the roadworks dataset from Mobilithek every 12 hours. The data is then converted to the OTM format (a logistics standard) and published within NDW's own environment. Although technically this dataset is open, for the purpose of the demonstration, we treat it as a closed dataset that can only be accessed via a contract.
2. **Registration and Publication**
NDW registers the dataset and the associated access endpoint on the Dutch marketplace (NTM) and updates the metadata so that users clearly understand the status and content of the dataset.
3. **Searching for Data**
A data consumer searches for the desired dataset on the NTM marketplace and submits a subscription request. In doing so, the consumer gains access to the contract terms associated with the dataset.
4. **Contract Validation and Endpoint Access**
NTM checks if there is an existing contract that grants access to the dataset. If such a contract exists, the endpoint is shared with the consumer. In the Proof of Concept (PoC), we assume that a valid contract is already in place.
5. **Data Retrieval and Validation**
The consumer retrieves the data via the NDW endpoint. NDW then validates with NTM whether a valid contract exists for both the dataset and the consumer. If everything checks out, the data is provided.

Subannex 3 Explanation of technical terms

I EDC: Eclipse Dataspace Connector (EDC)

The EDC is the core of the dataspace solution. This component manages and controls data exchange between the parties involved. The EDC distinguishes between a **Control Plane** and a **Data Plane**:

- **Control Plane:** Manages access and access control. This layer handles user authentication, policy management, and contract negotiations.
- **Data Plane:** Responsible for the actual data transfer. This separation enhances security by ensuring that data is only transmitted once all access conditions have been verified and approved.

Technical Components and Functions of the EDC

1. **Data Assets:** A data asset in the EDC is a representation of a dataset located in the provider's backend. Each asset includes a Data Address, which points to the dataset's storage location. This address enables secure data discovery and retrieval.



2. **Policy Management:** Policies determine who can view or use datasets. Examples include geographic restrictions, usage purposes, or sector-specific limitations.
3. **Contract Negotiations and Conclusions:** During contract negotiations, the dataspace interprets policies that act as guidelines. Once both parties agree to the terms, a formal contract is drawn up. This contract contains all agreements regarding data usage, such as data deletion requirements, restrictions on data sharing with third parties, and compliance with legal regulations.

Subannex 4 In-Depth Insight into the Benefits of the Eclipse Dataspace Connector

The Eclipse Dataspace Connector offers significant advantages to organisations aiming to securely and reliably exchange data. Here are some of the main benefits in more detail:

Data Sovereignty and Access Control

With the EDC, data owners retain full control over their data, even after it has been shared. The Policy Engine, part of the Control Plane, ensures that the owner can define who gets access to the data, for what purposes, and under what conditions. This prevents unauthorised access and reduces the risk of data breaches.

Why is this important?

- Reduces the risk of data breaches and unauthorised use.
- Allows for specific usage conditions, such as time-limited access or restrictions on how data may be used.

Secure Data Exchange

The EDC uses advanced security techniques, such as encryption during transmission (Data Plane) and strict authentication and authorisation methods (Control Plane). This ensures that data cannot be intercepted or misused by unauthorised parties during the exchange.

Why is this important?

- Protects sensitive data against cyberattacks and data theft.
- Ensures confidential communication between partners within a dataspace.

Interoperability and Standardisation

The EDC is designed to be compatible with international standards for data exchange and management, such as **GAIA-X**, a European initiative for secure data infrastructures. This means the connector can be easily integrated with other systems and platforms adhering to the same standards.

Why is this important?

- Facilitates collaboration between organisations, even if they use different technologies.
- Reduces technical barriers when establishing new data connections.



Open Source and Modular Architecture

The EDC is open-source, meaning the source code is freely available and can be modified and improved by the community. Additionally, the EDC has a modular architecture, allowing organisations to select and customise specific components to meet their needs.

Why is this important?

- Enables customisation and allows organisations to tailor functionalities without being locked into a proprietary solution.
- Reduces costs, as no licence fees need to be paid for commercial software.

Compliance with Laws and Regulations

The EDC provides built-in functionalities for compliance with laws and regulations, such as the **General Data Protection Regulation (GDPR)**. The Control Plane, for example, can ensure data minimisation, rights management, and logging of data access for audit purposes.

Why is this important?

- Ensures that organisations remain within legal boundaries when sharing data.
- Simplifies compliance with audits and regulatory requirements.

Scalability and Performance

The EDC is designed to scale as demand grows. The Data Plane can handle large amounts of data without loss of performance, and the Control Plane efficiently manages policy control, even in complex data scenarios.

Why is this important?

- Supports growing data volumes and an increasing number of partners within a dataspace.
- Ensures seamless operation, even in large-scale implementations.

Subannex 5 Policies and Contracts in Detail

Access and Usage Policies

The policies within the dataspace not only define who has access but also how the data can be used. These policies are established in two primary types:

- **Access Policies:** Determine which users can view or download a dataset. For example, datasets may be restricted to specific geographic regions, such as users only in the Netherlands.
- **Usage Policies:** Regulate the use of data once access has been granted. This may involve data deletion after a specified period, restrictions on reuse, or specific purposes for which the data may be used.

Contract Negotiations and Agreements

During contract negotiations, the dataspace establishes guidelines derived from policies. Once both parties accept the conditions, a formal contract is produced. This contract binds participants to the agreed conditions, such as data retention requirements, usage limitations, and adherence to regulatory standards.

Practical Examples of Policies

When sharing data in a digital space like Catena-X, it's important to set clear rules about how others can use your information. Think of these rules as a set of instructions that come with your data, much like the care instructions that come with a new piece of clothing. Here are some common ways you can control how your data is used:

The **Default Policy** is the simplest approach, making your data freely available to all approved users within the network. This is like putting a book in a public library – anyone who is a member can access it.

With a **Geographic Access Restriction**, you can share your data only with users in certain parts of the world. For instance, you could choose to share your data only within the European Union. This is similar to how some streaming services only work in certain countries.

Time-Limited Usage means you can share data temporarily. For example, you can set a rule that requires others to remove your data from their systems after 30 days. Think of it like lending someone a book with a strict return date.

Purpose-Restricted Usage lets you specify exactly what others are allowed to do with your data. For instance, you might allow your data to be used only for research projects. This is like giving someone permission to use your car, but only for driving to work.

Each of these rules helps ensure your data is used responsibly and in the way you intended. You can choose the rule that best matches your needs and comfort level with sharing.

Subannex 6 Data Plane and Control Plane in the Eclipse Dataspace Connector

Data Plane

The **Data Plane** within the EDC is the component responsible for the actual **data exchange** between parties. It handles the efficient and secure transfer of data, ensuring that information is transmitted in a controlled manner from one organisation to another.

Key Functions of the Data Plane:

1. **Data Transfer:** The Data Plane manages the physical transmission of data, ensuring that it arrives at the receiving party securely and without any loss of integrity.
2. **Data Encryption:** During the transfer, encryption is used to ensure that the data can only be read by the authorised party. The EDC supports various encryption methods to guarantee secure data transmission.

3. **Data Integrity:** The Data Plane implements mechanisms such as checksums and other verification techniques to ensure that data has not been altered or corrupted during transmission.
4. **Bandwidth Management and Optimisation:** In scenarios involving large volumes of data, the Data Plane manages bandwidth efficiently to ensure smooth data transfer without network congestion.

Characteristics of the Data Plane:

- **Efficiency:** The focus is on processing data exchanges quickly and efficiently, especially when handling large amounts of data in real-time.
- **Security:** Encryption and data integrity checks are crucial to ensure that only the intended recipient can access the data.
- **Flexibility:** Supports various data formats and protocols to enable data exchange between different platforms.

Control Plane

The **Control Plane** is the component responsible for managing and controlling data exchanges within the EDC. It focuses on **policy enforcement, authorisation, and monitoring**, ensuring that the Data Plane operates according to defined guidelines and agreements.

Key Functions of the Control Plane:

1. **Authentication and Authorisation:** Controls who has access to which data and under what conditions. Users and systems must be verified and authorised before they gain access to data.
2. **Policy Enforcement:** Enforcing rules and policies for data exchange, including access limits, usage conditions, and restrictions on data storage and utilisation.
3. **Auditing and Monitoring:** Keeps track of data exchanges, records who accessed the data, and monitors any deviations from policy. This is essential for reporting and regulatory compliance.
4. **Contract Management:** Manages agreements between parties on the conditions of data exchange, including legal and technical terms regarding who can do what with specific data.
5. **Configuration and Orchestration:** Provides a central interface to configure and manage different data connectors, allowing access rules and policies to be easily adjusted and distributed.

Characteristics of the Control Plane:

- **Policy-Driven:** All decisions are based on pre-defined policies set by the data owner.
- **Security and Compliance:** Ensures that data exchange complies with required regulations (such as GDPR) and internal security standards.
- **Traceability:** Through auditing and logging, the Control Plane provides visibility into all actions, contributing to transparency and accountability.