



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

Annex B - Metadata demonstrator with real-world use of data and different Levels of Service

Status: final  
Version: 1.4  
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 2.4.4 - Interoperable NAP demonstrator with Metadata.                                |
| Author(s)                | Ligia Conceição, PT-Armis; Sara Linhas, PT-Armis; Peter Lubrich, DE-Federal Highway Research Institute (BAST) |
| Co-author(s)             | João Montenegro, PT-Armis                                                                                     |
| Related to Milestone no. | 2.10                                                                                                          |
| External Milestone       | yes                                                                                                           |

**Document history**

| Version | Date       | created/ modified by                          | Comments                                   |
|---------|------------|-----------------------------------------------|--------------------------------------------|
| 1.0     | 09.09.2022 | Ligia Conceição, Armis                        | 1 <sup>st</sup> Draft                      |
| 1.1     | 27.02.2024 | João Montenegro, Armis<br>Peter Lubrich, BAST | Revision – 2 <sup>nd</sup> draft           |
| 1.2     | 04.12.2024 | João Montenegro & Sara Linhas, Armis          | Final Draft                                |
| 1.3     | 09.01.2025 | João Montenegro                               | Final draft (after CAT review)             |
| 1.4     | 11.03.2025 | Hannah Walther                                | Final version after SCOM written procedure |

**Document status:**

|                                     |                                                                                      |
|-------------------------------------|--------------------------------------------------------------------------------------|
| <input type="checkbox"/>            | To be revised by partners involved in the preparation of the document                |
| <input type="checkbox"/>            | For review/ approval by the Core Alignment Team and Peer Reviewers                   |
| <input type="checkbox"/>            | For approval by the NAPCORE Steering Committee                                       |
| <input checked="" type="checkbox"/> | Final                                                                                |
| <input type="checkbox"/>            | Updated/revised final version (e.g. when a second version of Milestone is published) |

## **Abstract**

This report provides insights on NAP interoperability regarding the use of metadata across NAPs. It goes specifically about the limitations of the current NAP Architecture regarding metadata extraction and dissemination. Recommendations to improve the general scenario of MS's NAPs interoperability and harmonisation are tackled in this document.



## Abbreviations

| Abbreviation | Meaning                                                    |
|--------------|------------------------------------------------------------|
| EC           | European Commission                                        |
| EU           | European Union                                             |
| CKAN         | Comprehensive Knowledge Archive Network                    |
| DCAT-AP      | Data Catalogue Vocabulary-Application Profile              |
| DR           | Delegated Regulation                                       |
| MS           | Member States                                              |
| NAP          | National Access Point                                      |
| NAPCORE      | National Access Point Coordination Organisation for Europe |
| SWG          | Sub-working Group                                          |
| WG           | Working Group                                              |
| WP           | Working Programme                                          |



## Table of contents

|                                                              |           |
|--------------------------------------------------------------|-----------|
| <b>1. Introduction.....</b>                                  | <b>6</b>  |
| 1.1. Objectives and scope.....                               | 6         |
| 1.2. Methodology & Structure of the document .....           | 6         |
| <b>2. Context and background .....</b>                       | <b>7</b>  |
| 2.1. National Access Points (NAPs).....                      | 7         |
| 2.2. Metadata .....                                          | 8         |
| <b>3. Solution design .....</b>                              | <b>9</b>  |
| 3.1. Purpose and goals .....                                 | 9         |
| 3.2. Relation to other projects .....                        | 9         |
| 3.3. Technologies for Implementation .....                   | 10        |
| 3.4. Requirement Analysis.....                               | 11        |
| 3.5. Definition of use-cases .....                           | 13        |
| <b>4. Detailed architecture framework.....</b>               | <b>14</b> |
| 4.1. Requirements / user needs .....                         | 14        |
| 4.2. Software stack architecture per module.....             | 14        |
| 4.3. Functional architecture .....                           | 15        |
| 4.4. User Interface .....                                    | 17        |
| 4.5. Considerations for the implementation.....              | 18        |
| <b>5. Results.....</b>                                       | <b>19</b> |
| 5.1. Use case 1 (Overview of NAP Datasets) .....             | 19        |
| 5.2. Use case 2 (Filter the available data by category)..... | 20        |
| <b>6. Conclusions.....</b>                                   | <b>21</b> |
| 6.1. Feedback to other NAPCORE WG2's tasks .....             | 21        |
| 6.2. Lessons learnt and recommendations .....                | 24        |

## List of tables

|                                                                                                       |    |
|-------------------------------------------------------------------------------------------------------|----|
| Table 3.1 - NAP analysis in terms of abilities to extract metadata April 2023 .....                   | 11 |
| Table 4.1 - Software stack Architecture according to Demonstrator Functionality Layers .....          | 14 |
| Table 6.1 - Essential KPIs for the extraction and processing of the metadata in the Demonstrator .... | 21 |
| Table 6.2 - Essential for future insurance of metadata quality .....                                  | 22 |
| Table 6.3 - NRA Modules importance .....                                                              | 24 |

## List of figures

|                                                                                     |    |
|-------------------------------------------------------------------------------------|----|
| Figure 3.1 – High level architecture .....                                          | 9  |
| Figure 3.2 – NAP analysis in terms of API documentation and DCAT-compliance .....   | 12 |
| Figure 4.1 – Demonstrator Initial Mockup .....                                      | 17 |
| Figure 5.1 – Demonstrator Interface in displaying basic Metadata of a Dataset ..... | 20 |
| Figure 5.2 – Demonstrator Map Display in Use Case 2 .....                           | 20 |



## 1. Introduction

Working Group 2 (WG2), Interoperability and level of service of National Access Points (NAPs), aims to enhance the compatibility of Member States' (MS) NAP data and features as well as harmonisation of levels of service. In order to achieve more interoperability, this document describes the development of a Metadata Demonstrator as part of Task 2.4, NAP interoperability demonstrators, of WG2 and more specifically of Work item 2.4.4 - Interoperable NAP demonstrator with Metadata (in short "Metadata Demonstrator"). This report compiles details on the solution design, the architecture framework, results and conclusions about the Metadata Demonstrator.

### 1.1. Objectives and scope

This report intends to outline the comprehensive scope and objectives of the Metadata Demonstrator, which primary aim is to demonstrate the feasibility of metadata consolidation from multiple National Access Points (NAPs) into a unified metadata portal.

By achieving this, the document aims to provide a step-by-step analysis of the implementation of the Demonstrator, highlighting its design approach, and shedding light on the current limitations within the metadata systems of the NAPs. Ultimately, the document serves as a critical resource for stakeholders and decision-makers interested in advancing the harmonization and utilization of metadata from different National Access Points for more effective and integrated use on a wide scale.

### 1.2. Methodology & Structure of the document

Considering the large scope of this subject, chapter 2 starts with the explanation of concepts related to the main research questions, followed by a brief exposition of the metadata characteristics (data standards per data category, as mentioned in DRs), challenges and methods related to the harmonization of said metadata. The conceptualization finishes by presenting several related projects. Chapter 3 presents the solution design, including the applied technologies, a requirement analysis, and possible use cases. Chapter 4 presents the architecture of the different modules of the Demonstrator, as well as the user interface and limitations identified while implementing. Chapter 5 addresses the results of the different modules. Finally, Chapter 6 comprises the feedback for improvement of the works carried out in other WG2 tasks (i.e., NAP's Levels of Service and Reference Architecture) and outlines the main learnings coming from the development of the demonstrator.

## 2. Context and background

### 2.1. National Access Points (NAPs)

The Directive 2010/40/EU on Intelligent Transport Systems (ITS) and its Delegated Regulations<sup>1</sup> mandate the establishment of a National Access Point (NAP) for mobility data within each European Member State. Presently, there are over 30 operational National Access Points, spanning across nearly all EU Member States and extending beyond, wherein mobility-related data is disseminated and made accessible for purposes such as travel information services.

It has been observed that the existing National Access Points exhibit significant disparities in their structural configurations and data access interfaces. Additionally, there exists considerable diversity in data formats and standards employed across Europe. In pursuit of achieving better alignment and cohesion, the National Access Point Coordination Organisation for Europe (NAPCORE) initiative was initiated.

NAPCORE receives co-financing through a Programme Support Action under the European Commission's Connecting Europe Facility. NAPCORE has been established as a coordinating mechanism aimed at enhancing the interoperability of National Access Points, which serve as the foundation for the exchange of mobility data across Europe. NAPCORE seeks to enhance interoperability by standardizing and aligning mobility data, thereby increasing accessibility and expanding the availability of mobility-related data through coordinated data access and improved harmonization of European National Access Points. Furthermore, NAPCORE empowers National Access Points and National Bodies by defining and implementing common procedures and strategies, reinforcing the position and role of National Access Points, and facilitating steps towards the development of pan-European solutions that enhance the utilization of EU-wide data.

NAPCORE has been conceived in the spirit of collaboration and consultation and encompasses 36 participants, including 33 Beneficiaries representing 26 EU Member States and 3 associated partners. Additionally, there are 37 Implementing Bodies involved in this endeavour. While the initial duration of the Programme Support Action extends until the conclusion of 2024, the overarching objective is to establish a sustainable and forward-looking organizational platform.

#### ***NAP Typologies***

It has been recently proposed a differentiation of NAP types according to the characteristics of the data presented in them:

- Metadata directory as term for metadata portals,
- Data platform as term for data portals.

#### *Metadata directory*

The metadata directory contains descriptions of the data and services important for their identification, assessment and subscription by consumer, i.e. metadata. It provides facilities to authenticate data provider and functionality to insert a metadata records. The NAP provides search

---

<sup>1</sup> European Union. Directive 2010/40/EU of the European Parliament and of the Council of 7 July 2010 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport. Official Journal of the European Union, L 207/1, 6 August 2010. Available online: <https://eur-lex.europa.eu/eli/dir/2023/2661>

and discovery services of the hosted records to end user and via machine readable to other portals. The NAP is not involved in the data exchange process between data providers and data consumers.

Key features: Storage & provision of metadata records of data, services and providers; search/discovery functionality; data provider authentication.

#### *Data platform*

The data platform contains descriptions of the data and services important for their identification, assessment and subscription by consumer, i.e. metadata, together with the data and/or services themselves. It provides facilities to authenticate data provider and data consumer and functionality to insert a data and metadata records to NAP internal database. The NAP provides search and discovery services of the hosted records to end user and via machine readable to other portals and distribution interface for the data dissemination. The NAP is an intermediary between data provider and consumer, the data however remain in the ownership of the data provider.

Key features: Storage & provision of metadata records of data, services and providers, storage & provision of data and services; search/discovery functionality; data provider and data user authentication.

It is noted that some NAPs offer elements from both mentioned typologies, i.e., they have a hybrid typology.

The Metadata Demonstrator will mostly focus on the development of a solution that is capable of taking advantage of the metadata directory NAP type.

## **2.2. Metadata**

One of the main functions of the NAPs is to enable access and exchange of mobility data. A detailed and standardized data set description, the so-called Metadata, facilitates such access by enabling discoverability of data sets within an NAP. Metadata describes the administration, organisation, and content of a dataset and of a data service. Metadata datasets are therefore crucial elements to make NAPs accessible and searchable

Regarding the actual usage of metadata profiles in individual NAPs, previous NAP reviews (among others in WG 3) revealed a quite heterogeneous metadata landscape across all reviewed NAPs. Thus, NAPs are based on different metadata schemes, and/or provide different ways to access the metadata.

Consequently, efforts were identified to harmonise the metadata infrastructures among the individual NAPs. A concrete solution by NAPCORE for such harmonisation is a recent metadata specification “mobilityDCAT-AP”, as a common metadata approach specific to NAPs. mobilityDCAT-AP provides a structured, interoperable and harmonised approach to describing and exchanging metadata about datasets and about access for such datasets related to mobility, and in particular related to Intelligent Transport Systems (ITS). It is recommended that the metadata management of National Access Points (NAPs) in Europe, or any other mobility data portals, is based on mobilityDCAT-AP in order to harmonise their data descriptions and ease the exchange of metadata in the mobility data ecosystem.

mobilityDCAT-AP is built as a mobility-related extension of DCAT-AP, a well-known metadata scheme for data portals in Europe. mobilityDCAT-AP is adapting and refining this scheme to provide a DCAT-AP-conformant representation of metadata specific to mobility data.

Further details about the goals and concepts behind mobilityDCAT-AP, and links to its documentations and preparatory actions can be found on the NAPCORE website<sup>2</sup>. mobilityDCAT-AP is an important baseline for the information model of the metadata aggregator, as foreseen for Demonstrator.

### 3. Solution design

#### 3.1. Purpose and goals

The main goal of this Demonstrator is to showcase an approach for a harmonized metadata approach in individual NAPs, and for cross-border accessibility of NAP metadata. This way, it will foster the overarching goal of NAP interoperability (see NAPCORE WG 2) in terms of metadata. In particular, it will show how heterogeneous metadata from individual NAPs can be harvested, processed and displayed in an integrated IT solution. The main functionality will be the accessibility of metadata from multiple NAPs, i.e., a data seeker would be able to discover NAP data offerings from multiple NAPs via one online platform. This way, the solution is somewhat similar to the EU Open Data portal<sup>3</sup>, which also harvests several other data portals.

The high-level architecture for the foreseen IT solution can be seen in Figure 3.1.

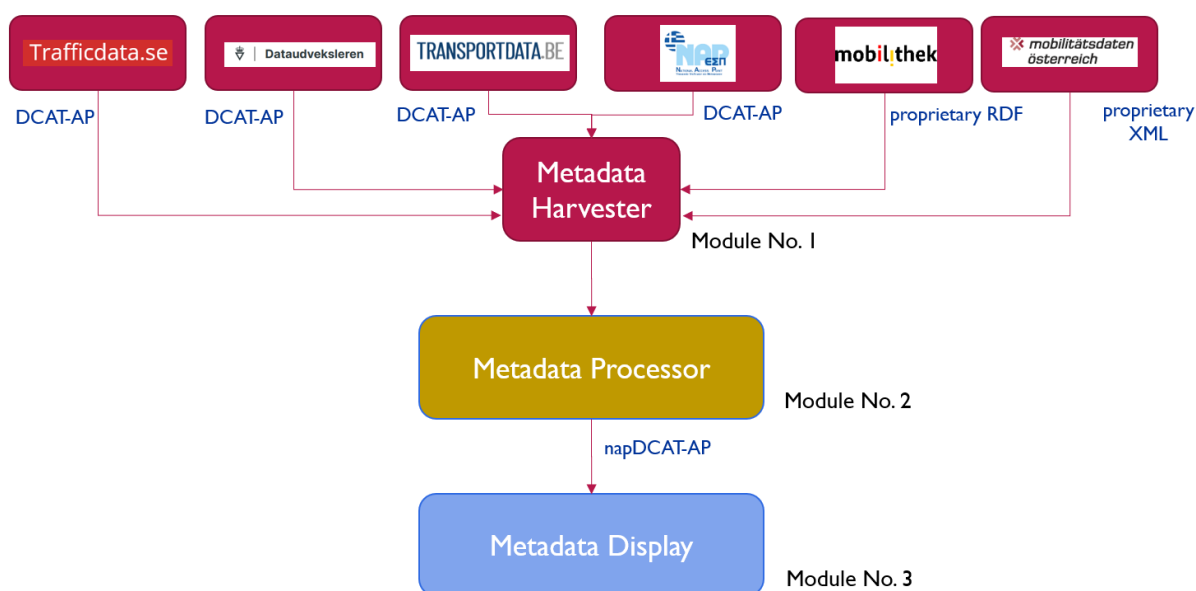


Figure 3.1 – High level architecture

Another goal of this Demonstrator is to showcase the usage of mobilityDCAT-AP for real-life NAP datasets.

#### 3.2. Relation to other projects

##### *Dependencies, relations to other projects or inputs to this demonstrator*

The main input for the Metadata Demonstrator is the mobilityDCAT-AP specification (see section 2.2), elaborated by NAPCORE sub-working group 4.4.

<sup>2</sup> <https://napcore.eu/release-of-the-mobilitydcat-ap/>

<sup>3</sup> <https://data.europa.eu/en>

### ***Outcomes and future relations to other projects from this demonstrator***

#### *NAPCORE WG2 Task 2.1 Level's of service of NAPs*

Provide an example on how the different NAP's Levels of service affect the automatization and dissemination of NAP data (through their metadata availability and structure) on a European scale. Additionally, with regard to the NAP LoS KPI Framework (NLKF), developed under task 2.1, the Metadata Demonstrator may provide suggestions for possible new KPIs, as well as adjustments to the existing ones that focus on metadata.

#### *NAPCORE WG2 Task 2.3 NAP Reference Architecture*

Understand how the current NAP architectures are “compatible” to this solution, and identify the NAP architectures limitations. The outcomes of this Metadata Demonstrator will help build the NAP Reference Architecture.

The Demonstrator identifies key architectural components that allow for NAPs to be compatible to the proposed solution. Therefore, it provides a relevant input to the development of the NAP reference architecture, since it proposes a set of functionalities and interfaces that should be adopted by the NAPs with the view to enhancing metadata interoperability,

### **3.3. Technologies for Implementation**

#### ***CKAN – Open Source Data Management System***

CKAN stands out as an open-source tool tailored to meet the specific needs of developers when it comes to metadata extraction from data providers<sup>4</sup>. Its primary function is to facilitate the publication, sharing, and organization of datasets, streamlining access to valuable information. CKAN boasts several features that make it a top choice for developers seeking efficient metadata extraction solutions:

- **User Management:** CKAN offers robust user management capabilities with varying access levels, including basic, member, editor, and organization admin privileges. This fine-grained control ensures that developers can tailor access permissions to their specific requirements.
- **API Tokens:** CKAN provides API tokens, ensuring secure authentication and data importation. These tokens serve as secure keys for accessing and managing data within the platform.
- **Rich Architecture:** CKAN's architecture is comprehensive and includes datasets, resources, a Datastore, and Datapusher for efficient data management. Additionally, it features visualization plugins designed to aid developers in interpreting data effectively.
- **Regular Data Updates:** CKAN offers a Windows Service (WS) option for automated, regular data updates on the platform. This service seamlessly interacts with both the EMEL API and CKAN's APIs, ensuring data integrity through rigorous data checks.

Developers can harness CKAN's APIs to interact directly with the platform. These APIs empower developers to retrieve and modify information related to datasets, resources, and other platform components. By submitting well-structured requests, it is possible to access and update data in a controlled and automated manner. A crucial aspect of using CKAN is its focus on data integrity. The Windows Service is meticulously designed to verify data consistency before initiating updates. This proactive approach minimizes the risk of data loss or corruption, thereby guaranteeing that the data remains reliable and accurate throughout its lifecycle.

In summary, CKAN's user management capabilities, API support, rich architecture, and data integrity safeguards make it a standout solution for developers looking to extract metadata from diverse data

---

<sup>4</sup> <https://ckan.org/>

providers. Its open-source nature and comprehensive feature set position CKAN as a powerful tool for efficient and secure data management and extraction.

### ***Chimera: Composable Semantic Transformation Pipelines***

Chimera is a software solution to solve data interoperability issues. Among others, it enables metadata converting to harmonize heterogeneous dataset descriptions against a reference metadata profile<sup>5</sup>.

This solution was used in a previous project for a tool named Asset Manager with the goal of consolidating metadata coming from different sources using Semantic Web technologies<sup>6</sup>. The presented solution implements an automated ingestion pipeline harmonising metadata from three different European NAPs in a single metadata catalogue.

## **3.4. Requirement Analysis**

The Metadata Demonstrator will serve as a prototype to merge and harmonise real-life metadata sets from different operational NAPs. The harmonisation will be based on the above mentioned mobilityDCAT-AP specification. A prerequisite for the Demonstrator is that NAP metadata sets are accessible in machine-readable format and reusable for post-processing. Standard technologies are preferred, such as REST APIs for the data access, and DCAT (or dialects) for the data structures. Some NAPs do already offer such features, other do not. Thus, only a few NAPs will be utilised for the Metadata Demonstrator.

Considering the different characteristics of the NAPs, a survey was done to analyze the characteristics of the NAPs in terms of their abilities to extract metadata.

*Table 3.1 - NAP analysis in terms of abilities to extract metadata April 2023*

| <b>NAP</b> | <b>Homepage</b>                                                                                                                                                                                       | <b>Metadata Extraction</b>                                            |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| Austria    | <a href="http://www.mobilitydata.gv.at/">http://www.mobilitydata.gv.at/</a>                                                                                                                           | Metadata API being implemented                                        |
| Belgium    | <a href="https://www.transportdata.be/en/">https://www.transportdata.be/en/</a>                                                                                                                       | Obtain metadata download link through CKAN API                        |
| Bulgaria   | <a href="https://www.mtc.government.bg/en/category/294/national-access-points-transport-related-data">https://www.mtc.government.bg/en/category/294/national-access-points-transport-related-data</a> | No information provided                                               |
| Croatia    | <a href="https://www.promet-info.hr/en">https://www.promet-info.hr/en</a>                                                                                                                             | No information provided                                               |
| Cyprus     | <a href="http://www.traffic4cyprus.org.cy/trafficapp/?wp=index-gr">http://www.traffic4cyprus.org.cy/trafficapp/?wp=index-gr</a>                                                                       | No information provided                                               |
| Czechia    | <a href="https://data.gov.cz/datov%C3%A9-sady">https://data.gov.cz/datov%C3%A9-sady</a>                                                                                                               | No information provided                                               |
| Denmark    | <a href="https://du.vd.dk/">https://du.vd.dk/</a>                                                                                                                                                     | Aggregates All Datasets Metadata in one file, >> Through HTTP request |
| Estonia    | <a href="https://web.peatus.ee/">https://web.peatus.ee/</a>                                                                                                                                           | No information provided                                               |
| Finland    | <a href="https://finap.fi/#/">https://finap.fi/#/</a>                                                                                                                                                 | No information provided                                               |
| France     | <a href="https://transport.data.gouv.fr/">https://transport.data.gouv.fr/</a>                                                                                                                         | Metadata machine extracted in other format                            |

<sup>5</sup> <https://github.com/cefriel/chimera>

<sup>6</sup> <https://ceur-ws.org/Vol-2939/paper6.pdf>



|                |                                                                                                                                                                                                                             |                                                    |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|
| Germany        | <a href="https://service.mdm-portal.de/">https://service.mdm-portal.de/</a>                                                                                                                                                 | Metadata provided in proprietary format            |
| Greece         | <a href="http://www.nap.gov.gr/">http://www.nap.gov.gr/</a>                                                                                                                                                                 | Obtain metadata download link through CKAN API     |
| Hungary        | <a href="https://napportal.kozut.hu/">https://napportal.kozut.hu/</a>                                                                                                                                                       | Metadata machine extracted through na API          |
| Ireland        | <a href="https://data.gov.ie/">https://data.gov.ie/</a>                                                                                                                                                                     | Obtain metadata download link through CKAN API     |
| Italy          | <a href="http://www.cciss.it/">http://www.cciss.it/</a>                                                                                                                                                                     | No information provide                             |
| Latvia         | <a href="https://lvceli.lv/en/road-network/statistical-data/transport-sector-open-data/">https://lvceli.lv/en/road-network/statistical-data/transport-sector-open-data/</a>                                                 | No information provided                            |
| Lithuania      | <a href="https://www.visimarsrutai.lt/gtfs/">https://www.visimarsrutai.lt/gtfs/</a>                                                                                                                                         | No information provided                            |
| Luxembourg     | <a href="https://data.public.lu/en/">https://data.public.lu/en/</a>                                                                                                                                                         | No information provided                            |
| Macao          | <a href="https://news.transport.gov.mt/data/">https://news.transport.gov.mt/data/</a>                                                                                                                                       | No information provided                            |
| Netherlands    | <a href="https://ndovloket.nl/">https://ndovloket.nl/</a>                                                                                                                                                                   | No information provided                            |
| Norway         | <a href="https://transportportal.atlas.vegvesen.no/no/">https://transportportal.atlas.vegvesen.no/no/</a>                                                                                                                   | Metadata machine extracted through an API          |
| Poland         | <a href="https://dane.gov.pl/dataset/1739,krajowy-punkt-dostepowy-kpd-multimodalne-usugi-informacji-o-podrozach">https://dane.gov.pl/dataset/1739,krajowy-punkt-dostepowy-kpd-multimodalne-usugi-informacji-o-podrozach</a> | No information provided                            |
| Portugal       | <a href="https://nap-portugal.imt-ip.pt/">https://nap-portugal.imt-ip.pt/</a>                                                                                                                                               | No information identified                          |
| Romania        | <a href="https://pna.cestrin.ro/">https://pna.cestrin.ro/</a>                                                                                                                                                               | No information provided                            |
| Slovakia       | <a href="https://odoprave.info/">https://odoprave.info/</a>                                                                                                                                                                 | No information provided                            |
| Slovenia       | <a href="https://www.ncup.si/en/multimo%20dal">https://www.ncup.si/en/multimo%20dal</a>                                                                                                                                     | Existing API does not provide metadata information |
| Spain          | <a href="https://nap.mitma.es/">https://nap.mitma.es/</a>                                                                                                                                                                   | No information provided                            |
| Sweden         | <a href="https://www.trafficdata.se/">https://www.trafficdata.se/</a>                                                                                                                                                       | Obtain metadata download link through CKAN API     |
| Switzerland    | <a href="https://opentransportdata.swiss/">https://opentransportdata.swiss/</a>                                                                                                                                             | Metadata machine extracted through an API          |
| United Kingdom | <a href="https://data.gov.uk/">https://data.gov.uk/</a>                                                                                                                                                                     | Obtain metadata download link through CKAN API     |

Following this analysis, the Metadata Demonstrator will consider only NAPs which offer certain extraction, i.e., access and reuse features for metadata. Additionally, it was analysed if applicable NAPs provide proper documentation to access the metadata (e.g., an API documentation) and if the metadata is provided in a DCAT-compliant format. To obtain such information, extensive research was conducted, i.e., documentation related to all National Access Points, if available, was thoroughly investigated. Figure 3.2 describes the outcomes of this analysis.

| Member State                | AT | BE | BG | HR | CY | CZ | DK | EE | FI | FR | DE | GR | HU | IE | IT | LV | LT | LU | MT | NL | NO | PL | PT | RO | SK | SI | ES | SE | CH | UK |
|-----------------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Documented API for metadata |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| Metadata in dcat format     |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |

Figure 3.2 – NAP analysis in terms of API documentation and DCAT-compliance

It is noted that an enhancement of individual NAPs in order to satisfy the above-mentioned requirements is not part of the NAPCORE project. Thus, not all NAPs currently qualify to be utilised by



the Demonstrator; we included only the ones that fulfilled both conditions highlighted in Figure 3.2, which were: Belgian, Danish, German, Greek and Swedish. This is, however, not an issue, as the Proof-of-Concept (POC) deployment of the Demonstrator (see chapter 5) is harvesting only selected NAPs.

### 3.5. Definition of use-cases

The working group will focus on two use cases, described below. These use cases were chosen because they ensure that the demonstrator addresses real user needs. This clarity in the selection process helps readers of this document to understand the value proposition of both the conceptual and technical solutions. The overall use case can be described as: *“As a Service Provider I wish to view the metadata of all NAPS so that I can get a grasp of the available traffic data characteristics”*. This was broken down in two more specific use cases.

#### **Use case 1: OVERVIEW OF NAP DATASETS**

As a Service Provider I would like to grasp which countries have an easily available metadata description of the dataset, organized by geographical area, so that I can design a product that is capable of extracting data across borders.

#### **Use case 2: FILTER THE AVAILABLE DATA BY CATEGORY**

As a Service Provider I would like to have access to up-to-date metadata that is organized according to a certain category, for example Heavy Vehicle Parking, so that I can observe what are the characteristics of that dataset (format of the data, updated timeframe).

Overview of the solution design **Primary objective:**

- To demonstrate metadata concentration from several NAPs in the ONE metadata portal.
- To show a proof-of-concept for the new metadata specification mobilityDCAT-AP, as a backbone of the metadata Demonstrator.

**Secondary objectives:**

- To showcase what metadata and in which form, what keywords shall be used what data naming conventions, language choices and descriptions shall be used to help understanding and findability of the services.
- To showcase data/metadata visualisation, search and all the processes around harvesting metadata and computing and displaying quality.
- To showcase technical and lingual problems of concentrating metadata

**Prerequisites:**

- All harvested portals must support a DCAT-compliant metadata structure that can be mapped to mobilityDCAT-AP
- All harvested portals must have metadata exposed either as a link or as API (static metadata download or SPARQL query)
- An IT solution must be built that enables reading, storing, harmonising metadata from different NAPs, and exposing such harmonised metadata in one common GUI

**Use-cases:**

- Overview of NAP Datasets
- Filter the available data by category

## 4. Detailed architecture framework

### 4.1. Requirements / user needs

Designing a detailed architecture framework for a Demonstrator to extract metadata from National Access Points (NAP) in an automatic way involves defining the system's components, interactions, and data flows. Below is a detailed framework for such a Demonstrator. This framework includes various components, such as data extraction, processing, storage, and user interfaces:

- **Data Extraction Module:** Responsible for extracting metadata from NAP sources.
- **Processing Engine:** Processes and transforms raw metadata into a structured format.
- **Metadata Storage:** Stores processed metadata efficiently.
- **User Interface:** Provides a user-friendly interface for interaction.
- **Automation Scheduler:** Schedules and triggers automatic extraction processes.

The Interaction between the different components works as the following:

- Data flows from NAP sources to the Data Extraction Module.
- Extracted data moves to the Processing Engine.
- Processed metadata is stored in the Metadata Storage.
- Users interact with the User Interface to view and manage metadata.
- The Automation Scheduler triggers regular metadata extraction.

Further explanation of how the different modules would work and be implemented according to this general architecture will be described further in the following sub-sections.

### 4.2. Software stack architecture per module

A software stack architecture is organized into layers, as it can be seen in Table 4.1, that handle different aspects of application development. The frontend layer is responsible for the user interface (UI), including presentation logic and client-side scripting. It encompasses elements like web pages, mobile app screens, and interactive components, often leveraging frameworks like React or Angular for efficient development. On the backend, the application logic resides, managing user requests, processing data, and interacting with the database layer. Backend languages such as Python, Java, or Node.js are used to build server-side logic, along with frameworks like Django or Express.js for streamlined development. This layer also exposes APIs for communication between the frontend and backend, facilitating data exchange and application functionality.

Meanwhile, the database layer handles data storage and access, utilizing databases like MySQL, PostgreSQL, or MongoDB. It stores structured or unstructured data and manages operations such as querying, inserting, updating, and deleting records. Object-Relational Mapping (ORM) or Object-Document Mapping (ODM) libraries abstract away complexities, providing a higher-level interface for interacting with databases. Together, these layers form a cohesive software stack architecture, enabling the development of robust and scalable applications with clear separation of concerns. This modular approach allows for easier maintenance, scalability, and flexibility in adapting to evolving business requirements and technological advancements.

Table 4.1 - Software stack Architecture according to Demonstrator Functionality Layers

| Layer    | Framework        | Programming Language | Purpose                                                    |
|----------|------------------|----------------------|------------------------------------------------------------|
| Frontend | ASP.NET Core MVC | C#                   | ASP.NET Core MVC is used to create the frontend of the web |



|             |                                        |     |                                                                                                                                              |
|-------------|----------------------------------------|-----|----------------------------------------------------------------------------------------------------------------------------------------------|
|             | User Interface: HTML, CSS, Razor Views |     | application. Razor Views are used to generate dynamic HTML content, and CSS is used for styling.                                             |
| Application | ASP.NET Core Web API                   | C#  | This layer handles the application logic and provides API endpoints for communication between the frontend and backend.                      |
| Backend     | ASP.NET Core                           | C#  | The backend layer contains the business logic and data processing code. It interacts with the database layer for data retrieval and storage. |
| Database    | Microsoft SQL Server                   | SQL | Purpose: Microsoft SQL Server is used as the relational database management system (RDBMS) for storing and managing structured data          |

### 4.3. Functional architecture

Functional architecture refers to the structural design of a software system based on its functional components and how they interact to fulfil specific tasks or requirements. In the context of an application with three modules - Extractor, Processor, and Display - the functional architecture outlines how data flows through these modules to achieve the application's objectives. The Extractor module retrieves data from a source, the Processor module manipulates and processes the extracted data according to predefined rules or algorithms, and the Display module presents the processed data to users in a comprehensible format. This architecture ensures a clear separation of concerns, with each module handling distinct responsibilities in the overall functioning of the application.

#### **Functionalities Module no. 1: "Metadata Harvester"**

The metadata harvester module is a crucial component designed to fulfil the specific objective of retrieving Metadata records from individual National Access Points (NAPs), whether on a periodic basis or based on triggers. Its primary functions encompass extracting metadata records in their original format, conducting validity checks, and attaching "Meta-Metadata" to enhance the overall metadata management process.

Core Functions of the Metadata Harvester Module:

**Periodic or Trigger-Based Retrieval:** The module is engineered to enable the retrieval of Metadata records from various NAPs. It can operate either on a scheduled periodic basis (such as daily, weekly, or monthly) or in a trigger-based manner, responding to specific events or requests. This flexibility ensures that the most up-to-date metadata is consistently acquired.

- **Original Format Preservation:** One of the fundamental tasks of the module is to extract metadata records from NAPs and save them in their original format. This is crucial for maintaining data fidelity and ensuring that metadata remains compatible with the source system.
- **Validity Checks:** The module incorporates mechanisms to perform validity checks on the retrieved metadata. These checks address critical questions such as the availability of data,

correctness of the XML schema, and adherence to defined standards. Any anomalies or issues identified during these checks are flagged for further action.

- **Meta-Metadata Attachment:** To enhance the management and traceability of metadata records, the module attaches "Meta-Metadata" to each retrieved record. This includes essential information such as the timestamp of retrieval (indicating when the data was collected), the URL of retrieval (identifying the source NAP), and the result of the validity check (indicating whether the metadata is in compliance with predefined standards).

#### Key Benefits of the Metadata Harvester Module:

- **Automation:** By operating on a periodic or trigger-based basis, the module automates the process of metadata retrieval, reducing manual effort and ensuring the consistency of metadata updates.
- **Efficiency:** The module's ability to save metadata in its original format ensures that data is readily usable without the need for extensive reformatting or transformation.

In summary, the metadata harvester module is a critical component of a metadata management system, serving to retrieve, validate, and enhance metadata records from NAPs. Its functions align with the objective of ensuring data availability and correctness making it an indispensable tool for organizations working with diverse data sources and standards.

#### ***Functionalities Module no. 2: "Metadata Processor"***

The metadata mapping and processing module of the Demonstrator serves a critical role in harmonizing metadata records sourced from different portals while adhering to the mobilityDCAT-AP standard. Its primary objective is to analyse and transform metadata records to ensure consistency and compliance with the designated metadata scheme.

The following bullets provide a detailed description of its functions and processes, with the main goal to allow a mapping of the metadata structure to the mobilityDCAT-AP specification:

- **Structure Mapping:** The module begins by mapping the metadata structure of harvested records to conform to the mobilityDCAT-AP specification. This involves aligning metadata elements with the prescribed format and schema.
- **Element Analysis:** It then thoroughly analyzes the elements within the saved metadata records to understand the completeness and quality of the data. This includes assessing which RDF fields (metadata elements) are populated and the manner in which they are filled.
- **Applying mobilityDCAT-AP scheme:** The module applies the official mobilityDCAT-AP scheme to ensure uniformity and compliance with the designated metadata specification. This step is crucial for harmonizing individual metadata records according to a common metadata scheme.

#### ***Functionalities Module no. 3 "Metadata Display"***

The core purpose of this module is to make metadata from various NAPs visible and easily searchable. Users can access metadata records from different sources through a unified interface, simplifying the discovery process.

To enhance user experience and streamline data discovery, the module may incorporate aligned search and filter functions. This means that users can apply consistent search criteria and filters across metadata records from different NAPs, regardless of the original source.



The module may also implement a standardized structure for result pages, irrespective of the original NAP source. This ensures a consistent and user-friendly presentation of metadata records, making it easier for users to interpret and compare information.

A noteworthy feature of this module is the ability to offer geographical visualization. It allows users to visualize metadata records on maps, providing a spatial context to the data. This feature can be particularly valuable for geospatial datasets, facilitating location-based exploration.

#### Key Benefits of the Metadata Display Module:

- **Enhanced Data Discovery:** By making metadata from multiple NAPs visible and searchable, the module improves the accessibility of valuable information, supporting data discovery and decision-making.
- **User-Friendly Experience:** Optional features like aligned search and result page structures enhance the user experience by providing consistency and ease of use.
- **Geographical Insight:** The geographical visualization feature adds a layer of spatial understanding, which can be especially valuable for datasets with geographic components.

In conclusion, the metadata display module is a crucial component of the demonstrator, focused on ensuring a consistent and streamlined usage of metadata across multiple NAPs, while offering functionalities to enhance the visibility and accessibility of metadata records from multiple NAPs. Optional features like aligned search, result page structure, and geographical visualization contribute to a more comprehensive and user-friendly experience.

## 4.4. User Interface

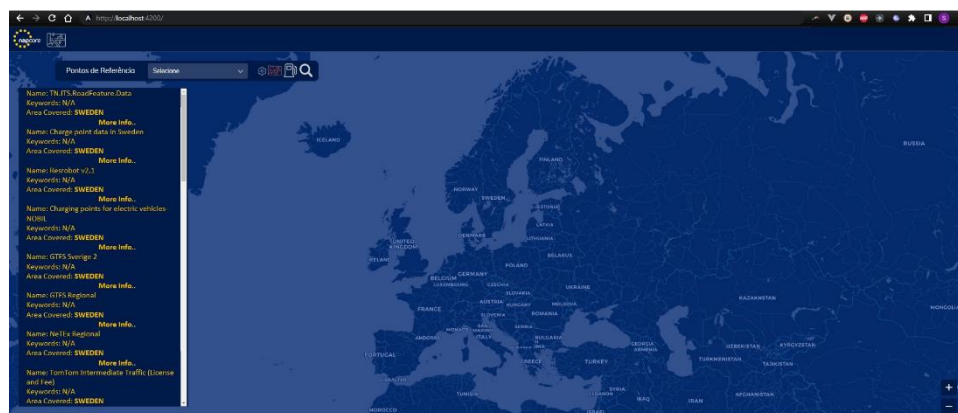


Figure 4.1 – Demonstrator Initial Mockup

The user interface mockup with geographic visualization and a standardized result page plays a pivotal role in enhancing the user experience and facilitating efficient data discovery. The geographic visualization component enables users to gain valuable insights by visually representing metadata records on a map, providing a spatial context that aids in understanding data distribution across geographical locations. This feature is particularly beneficial for datasets with geographic relevance, allowing users to explore data in a more intuitive and insightful manner.

Simultaneously, the standardized result page ensures consistency in the presentation of metadata records retrieved from multiple National Access Points (NAPs). Users benefit from a uniform layout and structure, simplifying the comparison and evaluation of metadata records originating from diverse sources. This consistency streamlines the process of locating relevant information and minimizes cognitive load, ultimately optimizing the effectiveness of data discovery within the Demonstrator. Together, these interface elements contribute to a user-friendly and comprehensive experience,



aligning with the Demonstrator's goal of harmonizing and making metadata records from various sources more accessible and usable.

#### **4.5. Considerations for the implementation**

The following bullet points point out what needs to be considered when implementing the solution proposed.

The costs associated with implementation:

- Presence of an external API in NAP
- Identification of fields for addition or extraction in the view

Service level considerations:

- Message throughput
- Uptime

Determining the implementation path involves evaluating:

- Frequency of metadata updates
- Choice between purchasing an Interchange as a service and integrating it with a NAP
- Preference for tight or loose integration

A source code, along with concise implementation guidelines and information about the mapping tool developed in this Demonstrator, is currently being prepared to be accessible on Github by request.

Implementation may also necessitate establishing agreements with the various NAP operators, when the metadata published on the Demonstrator's User Interface touch on re-usage or publications rules, which might be stipulated by the affected NAP operators.

## 5. Results

The following results section presents an analysis of the processes and outcomes for a proof-of-concept (POC) implementation for some aspects proposed in architecture framework (see chapter 4). In particular, the POC involves a metadata extraction from selected National Data Access Points (NAPs), and the utilization of such extracted metadata for enhancing data discoverability.

By analysing the use cases with metadata extraction and processing, it is possible to highlight the challenges such as time-consuming processes, potential failures due to changes in NAPs access, and discrepancies in metadata structure. Moreover, it is possible to validate the effectiveness of metadata in facilitating data discoverability even through full-text search functionalities, as well as the importance of open access to dataset metadata for efficient data discovery across platforms.

### 5.1. Use case 1 (Overview of NAP Datasets)

The process of extracting metadata from NAPs has successfully identified a vast pool of over 300 datasets, as depicted in Figure 5.1. However, the extraction and processing phase is inherently time-consuming, marked by several challenges:

- The extraction process may encounter failures, particularly when there are alterations in NAPs access, such as changes in URLs within GET requests or incorrect API access.
- The initial validation step demands continuous comparison with each NAP interface. If discrepancies arise, questions like whether it's due to an incorrect number of datasets or issues with the Extraction URL need to be addressed. For instances where no metadata is present, determining whether it's non-existent or if the extraction process failed for a specific metadata dataset becomes crucial.
- The presence of missing fields and the mention of different resources in NAP metadata necessitate a meticulous field-by-field comparison.
- Empty fields in the metadata of NAPs, such as the Quality Indicator, pose additional challenges. This prompts the exploration of proposing a Quality Indicator for the metadata structure to enhance its completeness and accuracy.

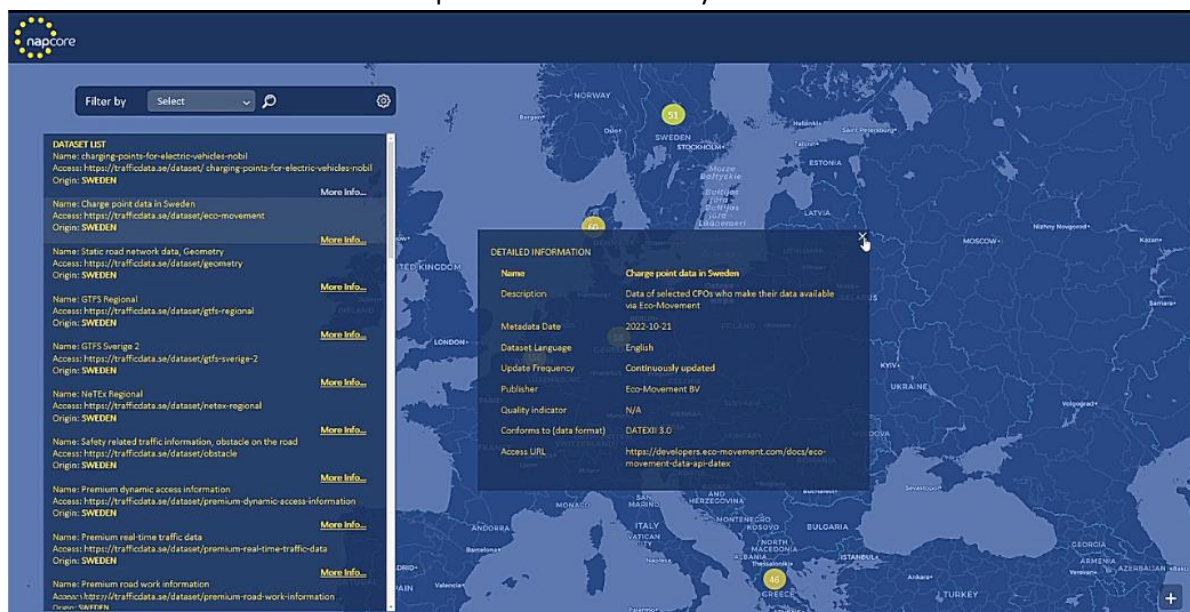


Figure 5.1 – Demonstrator Interface in displaying basic Metadata of a Dataset

## 5.2. Use case 2 (Filter the available data by category)

Figure 5.2 shows the GUI result of a test search query for NAP datasets. Approximately 10 datasets were obtained as a result of this test search query. This query concerns Alternative Fuels (AF) datasets. The user is allowed to see the characteristics of each queried dataset after clicking on the dataset of interest to it. Although these characteristics are not depicted in the Figure, they were shown in greater detail during a demonstration at the NAPCORE Mobility Data Days 2023 event in Budapest.

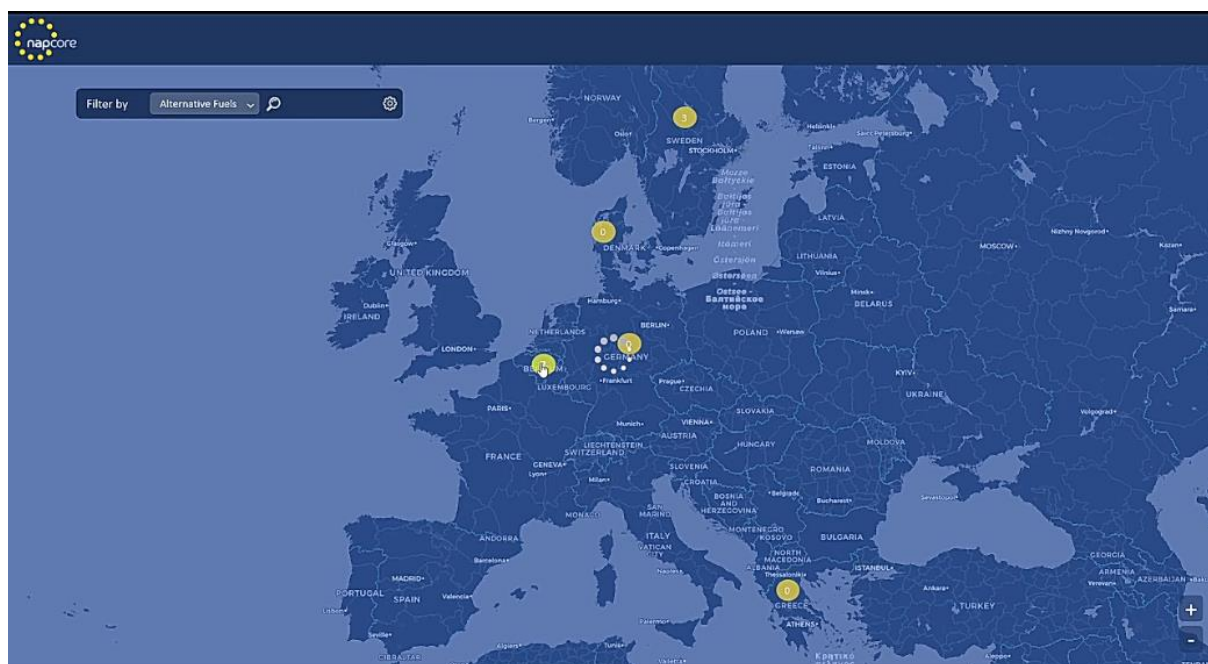


Figure 5.2 – Demonstrator Map Display in Use Case 2

The full-text search, utilizing a text search index tailored for a specific term or phrase within the Title and Description fields, yielded satisfactory outcomes.

This process served to validate the flexibility of metadata in enhancing data discoverability. The ability to use data discoverability extends to both metadata repositories and various NAP types within data platforms, if there is publicly accessible access to the metadata of the datasets. This underscores the importance of making dataset metadata openly available to facilitate efficient data discovery across diverse platforms.

## 6. Conclusions

Considering the outputs of the current analysis, the implementation of the Demonstrator has allowed to draw conclusions that guide us in refining and enhancing the approaches to improve data discoverability via an efficient metadata extraction from National Access Points (NAP). While subchapter 6.1 provides feedback to other NAPCORE tasks (namely, 2.1 and 2.3), which are the ones capable of further elaborating on interoperability and harmonisation aspects of the NAPs, subchapter 6.2 focuses on the main lessons learnt and general recommendations.

### 6.1. Feedback to other NAPCORE WG2's tasks

#### Task 2.1 – NAPs Levels of Service

The work methodology and the outcomes of the Metadata demonstrator provide relevant inputs for activities related to the Levels of Service (LoS) of the NAPs. The LoS is the main focus of task 2.1, which developed the NAP LoS KPI Framework (NLKF), a “tool” for estimating the LoS of NAPs, based on several features translated into KPIs. When comparing the features pertaining to the NLKF with the requirement analysis from this report (section 3.4 **Fehler! Verweisquelle konnte nicht gefunden werden.**), it is possible to identify the minimum requirement values for certain features (KPIs) so that the NAP could be integrated into this demonstrator. Such KPIs and respective “minimum values”, along with the main reasons why they are required, are listed in Table 6.1. It is worth mentioning that this analysis considered the 2024 version of the NLKF<sup>7</sup>.

Table 6.1 - Essential KPIs for the extraction and processing of the metadata in the Demonstrator

| KPI                                                                      | Value                                                                                                     | Reason                                                                                                                                                         |
|--------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.1 – Online Availability                                                | 1: Yes                                                                                                    | Maybe we can even erase this one, but there might be possible to extract metadata if previous access had been provided. But it would be impossible to find it. |
| 1.12 – Procedure for publication input of metadata or data on to the NAP | 1: Data resource metadata and/or providers' data is added by the NAP operator                             | Someone needs to have been able to input metadata at some point in time.                                                                                       |
| 1.13 – Metadata output and access restrictions                           | 3: Consumers need to register to get access to specific metadata only for non-publicly available datasets | In this demonstrator example, only publicly available metadata for extraction is used.                                                                         |
| 3.1 – Search functionalities                                             | 2: Discovery services available (search/filtering) based on harmonized metadata (free text or keywords)   | The means to Validate the results of the demonstrator entailed extraction and manual validation.                                                               |
| 3.2 – Search Results                                                     | 1: Yes                                                                                                    | The means to Validate the results of the demonstrator entailed extraction and manual validation.                                                               |
| 3.3 – Machine-readable metadata                                          | 1: Provision of machine-readable metadata in a self-describing format (JSON, XML, ...)                    | Essential functionality for metadata automatic extraction                                                                                                      |

<sup>7</sup> NAP LoS KPI Framework 2024 (<https://github.com/NAPCORE/NAP-LoS-KPI-Framework.git>)

| KPI                                     | Value                                                                                                                                                                                                                                                     | Reason                                                                                                                                                          |
|-----------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 6.1 – Metadata Specification            | 1: The NAP considered suggestions from the mobilityDCAT-AP, but the available metadata deviates – to some extent – from that specification (e.g., NAPs that had partial or full adoption of the EU-EIP specification – the Coordinated Metadata Catalogue | To achieve metadata extraction, a minimum standardisation of metadata is required.                                                                              |
| 7.3 – Operational procedure information | 1: Yes                                                                                                                                                                                                                                                    | If the extraction processes are not fully described with some sort of documentation it is not possible to understand how the extraction of metadata can be made |

Cross-checking with the results of the 2<sup>nd</sup> NAP LoS Self-assessment, it is possible to verify which NAPs adhere to these minimum values and, and hence, hypothetically, offer the necessary conditions to integrate this demonstrator. Regarding the NAPs included in this demonstrator, the German, Greek and Danish responses to the self-assessment indicate that they fulfil all those conditions, while the Belgian and the Swedish did not achieve the minimum value for one of the aforementioned KPIs each. In the case of Belgium, their NAP does not reach value 1 for “Operational procedure information”, however, due to direct contact with the Belgian NAP operators, it was possible to overcome the initial lack of information regarding operational procedure. As for the Swedish NAP, although they replied that their NAP does not provide machine-to-machine readable metadata (i.e., they replied 0 to KPI 3.3), automatic metadata was possible, which may indicate that there might have been a mistake while filling in the NLKF tool or a change within their system.

In addition to the NAPs used in the demonstrator, two other NAPs – which also provided feedback to the 2<sup>nd</sup> LoS self-assessment Workshop – also offered the necessary features highlighted in Table 6.1: the French and the Finish. The French NAP could not be integrated in this demonstrator since the format for automated metadata extracting was incompatible with the one used for this demonstrator (see Table 3.1). As for the Finish NAP, even though it is pointed out that their metadata is fully described, a documented API for harvesting metadata could not be found at the time of the analysis (see Figure 3.2). It is important to underline that the technical work of this demonstrator started in early 2023, and such NAPs might have been upgraded (developed new functionalities) in the meantime between the development of the work and the realization of the 2<sup>nd</sup> LoS self-assessment Workshop (March 2024).

In that sense, in addition to the necessary values for the extraction and processing of metadata, we built a list of essential features for future insurance of metadata quality, which would allow for the proper development of the activities demonstrated in this report. Table 6.2, presented in a similar format to Table 6.1, contains those essential features, along with associated values and justification.

*Table 6.2 - Essential for future insurance of metadata quality*

| KPI                                                             | Value                                           | Reason                                                       |
|-----------------------------------------------------------------|-------------------------------------------------|--------------------------------------------------------------|
| 1.12 - Procedure for publication of metadata or data on the NAP | 3: The data providers are inputting data and/or | Some validation of metadata is done. (The data providers are |



|                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                             |
|----------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                        | metadata + content of the (meta)data is verified manually in some random samples by the NAP operator (like technical or logical check of correct input type or completeness of the mandatory fields)                                                                                                                                                                                                                                    | inputting data and/or metadata + content of the (meta)data is verified manually in some random samples by the NAP operator (like technical or logical check of correct input type or completeness of the mandatory fields)) |
| 1.16 - Indication of data modification | 1: Webserver uses “Last updated time” information                                                                                                                                                                                                                                                                                                                                                                                       | Information to decide accurate scheduling of metadata extraction procedures                                                                                                                                                 |
| 3.3 - Machine-readable metadata        | 3: Provision of machine-readable metadata in a self-describing format complying to commonly agreed standard mobilityDCAT-AP                                                                                                                                                                                                                                                                                                             | Harmonized format to be extracted                                                                                                                                                                                           |
| 4.2 – Content and metadata             | 1: Established common responsibilities and procedures for the NAP content and metadata maintenance and updates established on a regular basis between NAP operators, data suppliers and data publishers: Keeping the data up-to date by systematically assessing data quality. For data and metadata –resources – at least once a year; checking if links are still functioning (to and from datasets) – at least once every six months | Continuous insurance of metadata quality                                                                                                                                                                                    |
| 6.1 – Metadata Specification           | 2: Full compliance/adoption with the MobilityDCAT-AP specification                                                                                                                                                                                                                                                                                                                                                                      | Standardized metadata Specification will help in the processing of metadata                                                                                                                                                 |
| 6.2 – Harvesting functionalities       | 1: a. Provision of metadata for harvesting interface of other NAPs (i.e., machine-to-machine readable metadata as Linked Data)                                                                                                                                                                                                                                                                                                          | Harvesting functionalities is a nice to have to decrease metadata extraction time.                                                                                                                                          |

To enhance and make the NLKF tool more complete, we recommend verifying if the requirements and pre-requisites highlighted in section 3 are represented as features in the current version. Moreover, the feedback provided in this section may allow for refining the features listed in Table 6.1 and Table 6.2.

### **Task 2.3 – NAP architecture**

As it can be seen in Table 6.3, the main functionalities needed from the NAPs for this demonstrator were: machine-to-machine access interface (Availability), the availability to provide metadata to consumers and to provide metadata search (Accessibility). According to NAP Reference Architecture M2.9 (first results of the NAP Reference Architecture work), these functionalities are related to metadata, data, and service: Metadata structure and content and the Metadata provision to user.



Some issues regarding the Metadata quality checks might provide improvements in future iterations. In Table 6.3, it is possible to relate the crucial functionalities with the Physical View and the Organisational View component essential for the retrieval of metadata in the Demonstrator approach according to the terminology presented in M2.9 developed by the NAPCORE Working Group (WG) 2, Task (T)2.3.

Table 6.3 - NRA Modules importance

| Functional View                                                                                                                                | Physical View                                                                                                                                                                                                                    | Organisational View                                                                                                                                                                                                                                         |
|------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Metadata Requesting System (a machine) retrieves metadata from NAP via a dedicated API. The Data Requesting System (machine)                   | NAP Metadata Management component that handles the management of the metadata, its access and retrieval, authorisation of NAP Content Providers (in cooperation with NAP User Management component) and metadata quality checks. | NAP Content Manager that ensures that the quality guidelines/criteria of content (data and metadata) are met, data formats are correct, and content is processed in compliance with the Service Level Agreement contracted with the NAP Technical Operator. |
| Data Provision System (machine) represent an addition to Data Platform NAP for machine-to-machine exchange of traffic and travel data with NAP |                                                                                                                                                                                                                                  | NAP Content Provider which is responsible for collecting, providing, and updating traffic and travel data and metadata                                                                                                                                      |

## 6.2. Lessons learnt and recommendations

One notable limitation of the Demonstrator stems from the time latency observed during the extracting and processing phases. This latency hinders real-time updates of metadata datasets, creating a lag between the availability of new data and its incorporation into the Demonstrator. Addressing this delay is crucial to ensure that the Demonstrator reflects the most current and relevant information.

The absence of an interface for National Access Point Management System (NAP MS) control poses a challenge, as the manual addition of each NAP access API URL can be cumbersome. Developing a user-friendly interface for managing NAPs would streamline this process, allowing for more dynamic and flexible control over the Demonstrator's data sources.

The Demonstrator currently focuses on two modules for metadata extraction. While effective for extracting metadata from a single file encompassing all NAP datasets and using CKAN API to obtain and process datasets, this limited scope may overlook other potential sources. Expanding the range of extraction modules would enhance the Demonstrator's adaptability to diverse data structures.

### Observed NAP Limitations:

The lack of harmonized metadata across NAPs introduces challenges during the integration process. Field-by-field validation becomes a necessity, especially when dealing with disparate formats. The adoption of the DCAT-AP format has proven advantageous, simplifying the identification of gaps and facilitating a more standardized approach to metadata integration.

Still, even a DCAT-AP-compliant representation of metadata in a NAP does not resolve all challenges. The main reason is the heterogenous provision of metadata information across individual NAPs, in



particular the population and usage of individual metadata fields. For example, some NAPs describe the data contents by a free-text description, other NAPs use a (sometimes complex) set of keywords, and other a predefined list of content categories. The mobilityDCAT-AP specification is resolving such issues by clear stipulations on the population and usage of metadata fields. Thus, it is recommended that any future deployment of new NAP systems, or revisions of existing NAP systems, consider a native implementation of the mobilityDCAT-AP specification within their internal information models, so a post-mapping (as proposed in the Demonstrator architecture) becomes reluctant.

Additionally, the ease of scaling the Demonstrator, particularly beyond the proof-of-concept (POC) stage, is contingent upon well-documented and readily available metadata extraction APIs. A robust extraction API ensures scalability, allowing the Demonstrator to adapt to an expanding ecosystem of NAPs seamlessly.

The registration process introduces additional steps when extracting metadata, an aspect not utilized in this POC. Moreover, the difference in the extraction processes between data and metadata from the same datasets raises considerations for streamlining these operations to enhance efficiency.

As for opportunities for Improvement of the Demonstrator:

Efforts should be directed towards minimizing the latency in the extracting and processing phases to enable real-time updates. Utilizing efficient algorithms and optimizing processing workflows can contribute to achieving this goal. Developing an intuitive user interface for NAP MS control would empower administrators to manage NAP access API URLs seamlessly. This enhancement can enhance the overall usability and flexibility of the Demonstrator. Consideration should be given to expanding the metadata extraction modules, ensuring the Demonstrator is capable of handling a variety of extraction scenarios. This includes exploring additional sources and methods for extracting metadata.

Encouraging NAPs to adopt standardized metadata formats, such as mobilityDCAT-AP, can significantly alleviate integration challenges. Collaborative initiatives to promote adherence to common standards will contribute to a more harmonized and interoperable environment.

A comprehensive scalability plan, including well-documented extraction APIs, should be established to facilitate the Demonstrator's growth beyond the POC stage. This involves anticipating the needs of a larger ecosystem and preparing the infrastructure accordingly.

Streamlining the registration process and harmonizing the metadata extraction processes for data and metadata from the same datasets will enhance the overall efficiency of the Demonstrator. Simplifying these operations reduces complexities and improves user experience.

In conclusion, the observed limitations serve as guideposts for future enhancements, positioning the Demonstrator as well as other tools to evolve into more robust and adaptable functionalities for automated metadata extraction from National Access Points. By addressing these challenges and embracing opportunities for improvement, we pave the way for tools that anticipate and accommodate the complexity of a dynamic data landscape.