



# Across

Automated zero-touch cross-layer provisioning framework for 5G and beyond vertical services

## D1.2: Data Management Plan

|                          |                                                     |                       |                   |
|--------------------------|-----------------------------------------------------|-----------------------|-------------------|
| <b>Lead beneficiary:</b> | UBITECH                                             | <b>Contributors:</b>  | CTTC, ICP, TID    |
| <b>Reviewers:</b>        | George Agapiou (WINGS), Dimitris Giannopoulos (UoP) |                       |                   |
| <b>Type:</b>             | DMP                                                 | <b>Dissemination:</b> | PU                |
| <b>Document version:</b> | 1.0                                                 | <b>Due date:</b>      | M06, 30 June 2023 |

**Abstract:** Deliverable D1.2 – Data Management Plan (DMP) describes plans for organizing, documenting, storing, and sharing data in line with FAIR principles (Findability, Accessibility, Interoperability, and Reusability). It considers issues such as data protection and confidentiality, data preservation and curation, methodologies and standards applied, and provides measures to ensure reproducibility of research outputs. The Data Management Plan is delivered in M6 and is constantly updated through the project lifespan.

## Project information

|                            |                                                                                             |
|----------------------------|---------------------------------------------------------------------------------------------|
| <b>Project title:</b>      | Automated zero-touch cross-layer provisioning framework for 5G and beyond vertical services |
| <b>Project acronym:</b>    | ACROSS                                                                                      |
| <b>Grant Agreement No:</b> | 101097122                                                                                   |
| <b>Type of action:</b>     | HORIZON JU Research and Innovation Actions                                                  |
| <b>Call:</b>               | HORIZON-JU-SNS-2022                                                                         |
| <b>Topic:</b>              | HORIZON-JU-SNS-2022-STREAM-A-01-07<br>Real-time Zero-touch Service Technologies             |
| <b>Start date:</b>         | 1 January 2023                                                                              |
| <b>Duration:</b>           | 36 months                                                                                   |

## Document information

|                             |                                                                                                            |
|-----------------------------|------------------------------------------------------------------------------------------------------------|
| <b>Associated WP:</b>       | WP1                                                                                                        |
| <b>Associated Task(s):</b>  | T1.2                                                                                                       |
| <b>Main authors:</b>        | Dimitrios Klonidis (UBITECH)<br>Dimitris Manolopoulos (UBITECH)                                            |
| <b>Contributors:</b>        | Lluís Gifre (CTTC)<br>David Cosgrave (ICP)<br>Angela Burgaleta Ledesma (TID)<br>Ioannis Markopoulos (NOVA) |
| <b>Reviewers:</b>           | George Agapiou (WINGS)<br>Dimitris Giannopoulos (UoP)                                                      |
| <b>Type:</b>                | DMP                                                                                                        |
| <b>Dissemination level:</b> | PU                                                                                                         |
| <b>Due date:</b>            | 30/06/2023                                                                                                 |
| <b>Submission date:</b>     | 27/06/2023                                                                                                 |

## Document revision history

| Version      | Date      | Changes                                        | Contributor(s)                 |
|--------------|-----------|------------------------------------------------|--------------------------------|
| <b>0.1</b>   | 10/3/2023 | ToC                                            | Dimitris Manolopoulos          |
| <b>0.2</b>   | 30/3/2023 | Abstract, Executive summary & Introduction     | Dimitris Manolopoulos          |
| <b>0.3</b>   | 30/3/2023 | Creation of Section 2.1                        | Dimitris Manolopoulos          |
| <b>0.4</b>   | 31/3/2023 | Creation of Section 2.2, Section 2.3           | Dimitris Manolopoulos          |
| <b>0.5</b>   | 7/4/2023  | Creation of Section 3.1 – 3.2                  | Dimitris Manolopoulos          |
| <b>0.6</b>   | 11/4/2023 | Creation of Section 3.3 – 3.4                  | Dimitris Manolopoulos          |
| <b>0.7</b>   | 19/4/2023 | Creation of Section 4 – 5                      | Dimitris Manolopoulos          |
| <b>0.8</b>   | 20/4/2023 | Creation of Section 6 – 7                      | Dimitris Manolopoulos          |
|              | 21/4/2023 | Populated table 1                              |                                |
|              | 24/4/2023 | Populated table 2                              |                                |
| <b>0.9</b>   | 25/4/2023 | Populated table 3 & 4                          | Dimitris Manolopoulos          |
|              | 26/4/2023 |                                                |                                |
|              | 27/4/2023 |                                                |                                |
|              | 28/4/2023 | Finish Table 4 & 5                             | Dimitris Manolopoulos          |
| <b>0.9.1</b> | 4/5/2023  | Tables 6 – 9                                   | Dimitris Manolopoulos          |
| <b>0.9.2</b> | 5/6/2023  | Comments and corrections                       | Lluis Gifre (CTTC)             |
| <b>0.9.3</b> | 6/6/2023  | Comments and corrections                       | David Cosgrave (ICP)           |
| <b>0.9.4</b> | 9/6/2023  | Comments and corrections                       | Angela Burgaleta Ledesma (TID) |
| <b>0.9.5</b> | 15/6/2023 | Peer review                                    | George Agapiou (WINGS),        |
| <b>0.9.6</b> | 16/6/2023 | Peer review                                    | Dimitris Giannopoulos (UoP)    |
| <b>0.9.7</b> | 18/6/2023 | Comments and corrections                       | Dimitrios Klonidis             |
| <b>0.9.8</b> | 21/6/2023 | Implementation of comments from Peer Reviewers | Dimitris Manolopoulos          |
| <b>1.0</b>   | 22/6/2023 | Final version for Submission                   | Dimitris Manolopoulos          |

**Disclaimer**

The content of this document reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

While the information contained in the documents is believed to be accurate, the authors(s) or any other participant in the ACROSS consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the ACROSS Consortium nor any of its members, their officers, employees, or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the ACROSS Consortium nor any of its members, their officers, employees, or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

**Copyright message**

© ACROSS Consortium. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation, or both. Reproduction is authorised provided the source is acknowledged.

## Table of contents

|                                                                                                 |    |
|-------------------------------------------------------------------------------------------------|----|
| List of acronyms and abbreviations .....                                                        | 6  |
| List of tables .....                                                                            | 7  |
| Executive summary .....                                                                         | 8  |
| 1 Introduction .....                                                                            | 9  |
| 2 Data summary .....                                                                            | 9  |
| 2.1 Types and formats of data the project will generate/collect .....                           | 9  |
| 2.2 Purpose of the data collection/generation and relation to the objectives of the project ... | 14 |
| 2.3 Reusability of existing data .....                                                          | 18 |
| 3 FAIR data .....                                                                               | 23 |
| 3.1 Making data findable, including provisions for metadata .....                               | 23 |
| 3.2 Making data openly accessible .....                                                         | 23 |
| 3.3 Making data interoperable .....                                                             | 27 |
| 3.4 Increase data re-use (through clarifying licenses) .....                                    | 28 |
| 4 Allocation of resources .....                                                                 | 29 |
| 5 Data security .....                                                                           | 30 |
| 6 Ethical aspects .....                                                                         | 30 |
| 7 Conclusions .....                                                                             | 31 |

## List of acronyms and abbreviations

| Abbreviation | Description                                                |
|--------------|------------------------------------------------------------|
| <b>DMP</b>   | Data Management Plan                                       |
| <b>EC</b>    | European Commission                                        |
| <b>API</b>   | Application Programming Interface                          |
| <b>FAIR</b>  | Findability, Accessibility, Interoperability, Re-usability |
| <b>GDPR</b>  | General Data Protection Regulation                         |
| <b>ICT</b>   | Information Communication Technologies                     |
| <b>ORDP</b>  | Open Research Data Pilot                                   |
| <b>HE</b>    | Horizon Europe                                             |
| <b>RDM</b>   | Research Data Management                                   |

## List of tables

|                                                                                          |    |
|------------------------------------------------------------------------------------------|----|
| Table 1: Research data outputs types and format.....                                     | 10 |
| Table 2: Project objectives with related dataset collection/generation. ....             | 14 |
| Table 3: Data reused or intended to be reused by ACROSS. ....                            | 18 |
| Table 4: ACROSS origin, expected size, data utility of datasets collected/generated..... | 19 |
| Table 5: Findable data including provisions for metadata.....                            | 23 |
| Table 6: Provisions for openly accessible data .....                                     | 24 |
| Table 7: ACROSS provisions for making data interoperable.....                            | 27 |
| Table 8: ACROSS data licences and provisions for data re-use.....                        | 28 |
| Table 9: Resources allocated for data management and making data FAIR in ACROSS. ....    | 29 |
| Table 10: Data security provisions in ACROSS. ....                                       | 30 |
| Table 11: Ethical aspects related to data sharing in ACROSS .....                        | 30 |

## Executive summary

### What is the focus of this Deliverable?

The D1.2 - Data Management Plan (DMP) is a living document, intended to describe the data management life cycle for all data sets that are collected, processed or generated by ACROSS. It also intends to formalize the means of production of software as well as the goals and structures that will ensure the software produced by ACROSS remains accessible and reusable in the short, medium and long term.

### What are the deliverable contents?

The ACROSS project designs and implements an end-to-end service deployment and management platform for next-generation networks and services, aiming at unprecedented levels of automation, performance, scalability, and energy efficiency. Thus, it deals with any type of data and datasets generated or aggregated through its various activities. These types of data and datasets will be made available, in line with the guidelines for the Open Data, Software and Code Guidelines in Horizon Europe<sup>1</sup>. The DMP also includes the dataset metadata specification that are used in the data registry, following an appropriate relevant standard<sup>2</sup>. It specifies the recommended licensing schemes as suggested by HEU. The DMP describes both existing and planned data sets.

### Conclusions and recommendations

The ACROSS DMP as depicted in D1.2 constitutes the basic tool that is used to manage data in the ACROSS project.



## 1 Introduction

Proper Research Data Management (RDM) is mandatory for any Horizon Europe project generating or reusing research data. It is a key part of Horizon Europe's open science requirements. In Horizon Europe, beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles, and should at least do the following:

- Prepare a Data Management Plan (DMP) and keep it updated throughout the course of the project.
- Deposit data in a trusted repository and provide open access to it ('as open as possible, as closed as necessary').
- Provide information (via the same repository) about any research output or any other tools and instruments needed to re-use or validate the data.

Open access to research data refers to the right to access and re-use digital research data under the terms and conditions set out in the Grant Agreement. Openly accessible research data can typically be accessed, mined, exploited, reproduced and disseminated under defined and clearly specified terms and conditions (either free of charge or at a cost for the user).

The DMP is not a fixed document; it evolves and gains more precision and substance during the lifespan of the project.

This first version of the DMP is delivered in month 6 (M6) of the project. As the project evolves, potential updates of this plan are examined on a biannual basis (M6, M12, M18, M24, M30, M36) and confirmed at a plenary level. Any potential changes/updates will be reported through the periodic reports on M18 and M36.

## 2 Data summary

### 2.1 Types and formats of data the project will generate/collect

The ACROSS project facilitates a broad range of data collection mechanisms that relate with the overall platform environment, the individual enabling modules that provide the trusted and intelligent ZTP features, the testing and validation outcomes from the deployed testing cases, and the outcomes of the supporting studies and analyses. The types of research data outputs from the above categories will at least include:

- (a) Requirements and specifications datasets at system level and at module and software component level (WP2);
- (b) Software tool datasets for the management and orchestration of the multi-domain orchestration processes (WP3 - T3.1);
- (c) Software tool datasets for the management and orchestration of the distributed edge domain instances (WP3 - T3.2);
- (d) Software tool datasets for the development and deployment of platform modules including: the integration fabric interfaces, the end-user interface, the northbound APIs, the application porting process, the telemetry mechanism and the linked data storage and analytics (WP3). The AI-based event processing engines, the secure and trusted data governance and sharing processes, and the ZTP processes across the various layer and instances (WP4);

- (e) Collected datasets from the creation and collection of events over the TID realistic network environment for the training of the AI algorithms and datasets of the processed outcomes after AI training (WP4);
- (f) Design information datasets for the testbed deployment and the testing scenarios including methodologies on the execution of the scenarios for the collection of KPI parameters (WP5);
- (g) Collected datasets from the performance validation of the functional use cases per use case scenario (WP5);
- (h) Collected datasets from the deployment and evaluation of the use cases in industrial environments (WP5);
- (i) Datasets from market and societal analyses of the developed solutions (WP6).

Table 1: Research data outputs types and format.

| Research data outputs | Type                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Format                                                                                                                                                                                                                                                                                                                                                                             |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (a)                   | <ul style="list-style-type: none"> <li>• Text documents such as functional and non-functional requirements, technical specifications, use case scenarios, and user stories.</li> <li>• Diagrams and schematics such as system architecture diagrams, flowcharts, and sequence diagrams.</li> <li>• Models such as data models, object models, and state machine models.</li> <li>• Tables and spreadsheets such as requirement traceability matrices, test plans, and test cases.</li> <li>• Code artifacts such as code snippets, code modules, and libraries.</li> <li>• Configuration files such as network configuration files, software configuration files, and system configuration files.</li> <li>• Input and output data formats for software modules or components.</li> <li>• Constraints and assumptions, including hardware and software requirements, limitations, and dependencies.</li> </ul>           | Typically, may include .csv, .docx, .xlsx, flowcharts, and other similar formats. They could also be stored in databases, software repositories, or other digital formats that allow for easy management and retrieval.                                                                                                                                                            |
| (b)                   | <ul style="list-style-type: none"> <li>• <b>Configuration data:</b> This includes data related to the configuration of network devices, virtual network functions (VNFs), and network services.</li> <li>• <b>Performance data:</b> This includes data related to the performance of network devices, VNFs, and network services such as latency, throughput, and packet loss.</li> <li>• <b>Monitoring data:</b> This includes data related to the monitoring of network devices, VNFs, and network services such as CPU utilization, memory utilization, and network traffic.</li> <li>• <b>Event data:</b> This includes data related to events occurring in the network such as alarms, notifications, and faults.</li> <li>• <b>Service data:</b> This includes data related to the management and orchestration of network services such as service topology, service templates, and service catalogue.</li> </ul> | Can vary depending on the specific software tool and the type of data being stored. Some common formats include structured data formats such as JSON or XML, as well as unstructured formats like plain text or binary data. Additionally, some software tools may use proprietary formats specific to that tool or vendor. The format used will depend on the requirements of the |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     | <ul style="list-style-type: none"> <li>• <b>Resource data:</b> This includes data related to the management of network resources such as virtual machines, storage, and network connectivity.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | software tool and the type of data being stored.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| (c) | <ul style="list-style-type: none"> <li>• <b>Configuration data:</b> This includes information about the hardware and software configuration of the edge domain instances, such as network topology, device types, operating systems, and applications.</li> <li>• <b>Performance data:</b> This includes information about the performance of the edge domain instances, such as CPU utilization, memory usage, and network bandwidth.</li> <li>• <b>Log data:</b> This includes log files generated by the edge domain instances, which can provide information about system events and errors.</li> <li>• <b>Event data:</b> This includes data related to events that occur on the edge domain instances, such as the detection of a new device or the occurrence of a security breach.</li> <li>• <b>Monitoring data:</b> This includes data collected by monitoring tools deployed on the edge domain instances, such as metrics related to resource usage, network traffic, and application performance.</li> <li>• <b>Metadata:</b> This includes additional information about the edge domain instances and their associated resources, such as location, ownership, and security attributes.</li> </ul>                                                                                                                                                                                                                                                                 | <p>JSON, XML, or YAML.</p> <p>It may also be stored in databases, spreadsheets, or other file formats. Additionally, the data may be transmitted between systems using APIs, protocols like REST or gRPC, or other messaging formats like MQTT or AMQP.</p>                                                                                                                                                                                                                                                                                                     |
| (d) | <ul style="list-style-type: none"> <li>• <b>Configuration data:</b> This includes settings, parameters, and options used to configure the various platform modules.</li> <li>• <b>Telemetry data:</b> This includes data collected from the network devices and applications, such as performance metrics, event logs, and alarms.</li> <li>• <b>Analytics data:</b> This includes data used for data analysis and visualization, such as network traffic data, user behaviour data, and application performance data.</li> <li>• <b>API data:</b> This includes data related to the northbound APIs, such as API documentation, API usage data, and API performance data.</li> <li>• <b>Linked data:</b> This includes data that links different data sources together, such as semantic web data and linked data from various sources.</li> <li>• <b>AI-based event processing data:</b> This includes data used for training and deploying AI-based event processing models, such as training data, model configuration data, and model performance data.</li> <li>• <b>Secure and trusted data governance and sharing data:</b> This includes data related to security and access control, such as user profiles, access policies, and audit logs.</li> <li>• <b>Zero Touch Provisioning (ZTP) data:</b> This includes data used for automated device provisioning and configuration, such as device templates, network topology data, and configuration scripts.</li> </ul> | <p>May vary depending on the specific tool and the type of data. However, some common formats for software-related data include code files (e.g., Java, Python, C++, etc.), configuration files (e.g., YAML, JSON, XML, etc.), database files (e.g., SQL, NoSQL, etc.), log files (e.g., text, CSV, JSON, etc.), and various types of binary files (e.g., executables, libraries, etc.). The specific format may also depend on the tool and its purpose, as well as any relevant standards or conventions that apply to the particular domain or industry.</p> |
| (e) | <ul style="list-style-type: none"> <li>• Network traffic data, such as packet header information, flow statistics, and traffic volume and patterns.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | The data formats could include structured data                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                        |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     | <ul style="list-style-type: none"> <li>• Network performance data, such as latency, packet loss, and jitter.</li> <li>• Network configuration data, such as device and link configurations and settings.</li> <li>• Security data, such as intrusion detection alerts and logs.</li> <li>• System logs and event data, such as system messages and alerts, device failures, and user actions.</li> </ul> <p>The datasets of the processed outcomes after AI training may include:</p> <ul style="list-style-type: none"> <li>• Trained AI models, including deep learning neural networks and other machine learning algorithms.</li> <li>• Analytical data generated by the trained models, such as predictions, recommendations, and anomaly detection results.</li> <li>• Evaluation metrics, such as accuracy, precision, recall, and F1-score, used to measure the performance of the trained models.</li> <li>• Training and testing data sets, which may include annotated data for supervised learning and unannotated data for unsupervised learning.</li> </ul> | <p>formats such as CSV, JSON, or XML, or unstructured formats such as log files, images, or videos. The format could also depend on the specific AI algorithms used for the training, such as deep learning or machine learning, and the software tools used for the data processing and analysis.</p> |
| (f) | <ul style="list-style-type: none"> <li>• <b>Network topology data:</b> This includes information on the structure and configuration of the test network, including the devices, links, protocols, and services being used.</li> <li>• <b>Test scenario data:</b> This includes information on the specific scenarios being tested, such as the traffic patterns, application workloads, and user behaviours that are being simulated.</li> <li>• <b>KPI measurement data:</b> This includes information on the KPIs being measured during testing, such as latency, throughput, packet loss, and availability.</li> <li>• <b>Testbed configuration data:</b> This includes information on the hardware and software components being used in the testbed, including the servers, switches, routers, and other networking equipment.</li> <li>• <b>Methodology data:</b> This includes information on the testing methodologies being used, such as the test procedures, tools, and techniques being employed to collect and analyse the data.</li> </ul>                  | <p>May vary depending on the specific tool or platform being used to manage the testbed and execute the testing scenarios. However, common formats for the data may include spreadsheets, databases, log files, and JSON or XML formatted files.</p>                                                   |
| (g) | <ul style="list-style-type: none"> <li>• <b>Network performance metrics:</b> These metrics measure various aspects of network performance, such as latency, packet loss, throughput, jitter, and bandwidth utilization.</li> <li>• <b>Application performance metrics:</b> These metrics measure the performance of specific applications running on the network, such as response time, transaction rate, and user concurrency.</li> <li>• <b>System performance metrics:</b> These metrics measure the performance of the underlying hardware and software systems, such as CPU usage, memory usage, and disk I/O.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                           | <p>May vary depending on the specific metrics being measured and the tools used to collect them. For example, network performance metrics may be collected using tools such as packet sniffers or network</p>                                                                                          |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|     | <ul style="list-style-type: none"> <li>• <b>User behaviour data:</b> This data captures user interactions with the system and can include information such as user sessions, clickstream data, and user preferences.</li> <li>• <b>Log data:</b> This data captures system logs generated by various components of the system, such as network devices, servers, and applications, and can be used for troubleshooting and analysis.</li> </ul>                                                                                                                                                                                                                                                                            | <p>performance monitors and stored in formats such as CSV, JSON, or XML. Application performance metrics may be collected using application performance monitoring (APM) tools and stored in formats such as log files or time-series databases. System performance metrics may be collected using system monitoring tools and stored in formats such as system logs or performance counters. User behaviour data may be collected using web analytics tools and stored in formats such as clickstream data or session logs. Log data may be stored in various formats such as text files, log files, or syslog.</p> |
| (h) | <ul style="list-style-type: none"> <li>• <b>Environmental data:</b> such as temperature, humidity, air quality, noise level, etc.</li> <li>• <b>Equipment data:</b> such as power consumption, production rate, equipment uptime, maintenance records, etc.</li> <li>• <b>Process data:</b> such as workflow diagrams, process flow rates, process deviations, etc.</li> <li>• <b>Quality data:</b> such as quality control parameters, test results, product yield, etc.</li> <li>• <b>Safety data:</b> such as worker safety records, safety incident reports, safety inspection reports, etc.</li> </ul>                                                                                                                | <p>May vary depending on the type of data, but they could include structured data in databases or spreadsheets, unstructured data in log files or text documents, and multimedia data such as images or videos.</p>                                                                                                                                                                                                                                                                                                                                                                                                  |
| (i) | <ul style="list-style-type: none"> <li>• <b>Market research data:</b> This could include data on market trends, customer needs, market size, and potential competitors. It may be collected through surveys, interviews, focus groups, or secondary research.</li> <li>• <b>Stakeholder usage data:</b> This could include data on how stakeholders are using the developed solutions, such as which features are most commonly used and how often customers are interacting with the solution.</li> <li>• <b>User feedback data:</b> This could include data on user satisfaction, user preferences, and user suggestions for improvement. It may be collected through surveys, interviews, or feedback forms.</li> </ul> | <p>.csv, .docx, .xlsx</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

|  |                                                                                                                                                                                                                                                                                                                                                                                                                                         |  |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
|  | <ul style="list-style-type: none"> <li>• <b>Economic impact data:</b> This could include data on the economic impact of the developed solutions, such as cost savings or increased revenue for businesses using the solution.</li> <li>• <b>Social impact data:</b> This could include data on the social impact of the developed solutions, such as improved quality of life for end-users or reduced environmental impact.</li> </ul> |  |
|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|

## 2.2 Purpose of the data collection/generation and relation to the objectives of the project

The project objectives associated with data collection/generation are depicted in the following table:

Table 2: Project objectives with related dataset collection/generation.

| Objective                                                                                                                                                                                                                                                                                                                                                                                         | Dataset collection/generation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>Objective 1:</b> Design a future-proof service orchestration framework built upon iterative (i) deep infrastructure visibility, (ii) unsupervised secure reasoning via AI, and (iii) self-tuning zero-touch capabilities, acting upon massive spontaneous events originating from across the end-to-end geo-distributed cloud continuum.</p> <p><b>Related research data outputs:</b> a</p> | <ul style="list-style-type: none"> <li>• <b>User requirements dataset:</b> This dataset should include information about the users of the ACROSS platform, their needs, preferences, and expectations.</li> <li>• <b>Technical requirements dataset:</b> This dataset should include information about the technical requirements of the ACROSS platform, such as the hardware and software requirements, performance metrics, security and privacy considerations.</li> <li>• <b>Business requirements dataset:</b> This dataset should include information about the business requirements of the ACROSS platform, such as the target markets, pricing, and revenue models.</li> <li>• <b>Existing standards dataset:</b> This dataset should include information about the existing standards related to service orchestration, such as the ETSI ZSM.</li> <li>• <b>ACROSS architecture dataset:</b> This dataset should include information about the design of the ACROSS architecture, such as the components, interfaces, protocols, and data models.</li> <li>• <b>Test cases dataset:</b> This dataset should include information about the test cases that will be used to demonstrate and validate the ACROSS platform, such as the scenarios, data sets, performance metrics, and evaluation criteria.</li> <li>• <b>Impact dataset:</b> This dataset should include information about the impact of the ACROSS platform on the standards, open-source platforms, and the market, such as the adoption rate, user feedback, and market share.</li> </ul> |
| <p><b>Objective 2:</b> Design &amp; implement a service deployment &amp; orchestration framework empowered by (i) an ensemble of highly distributed domain</p>                                                                                                                                                                                                                                    | <ul style="list-style-type: none"> <li>• User, technical, and business requirements for ACROSS.</li> <li>• Data on existing NFV orchestration systems, including OSM and OpenStack.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |



|                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>orchestrator instances across the various edge sites, (ii) orchestrated - in an end-to-end fashion - by a logically centralised cloud-managed multi-domain orchestrator, through an integration fabric of standardized interfaces.</p> <p><b>Related research data outputs a, c, d</b></p>                                        | <ul style="list-style-type: none"> <li>• Data on disaggregated O-RAN deployments, P4 whiteboxes/SmartNICs, and in-band network telemetry.</li> <li>• Standardized interface data, including TOSCA, YANG, GSMA GST/NEST, and TMF Open APIs, and open telemetry APIs (P4.org, IETF In-situ).</li> <li>• Realistic test cases for multi-domain orchestration, including end-to-end service orchestration.</li> <li>• Data on existing integration fabrics and standards, including ETSI ZSM.</li> <li>• Data on service APIs, including service onboarding, zero-touch orchestration, real-time policy management, AI, telemetry and analytics, and data generation/augmentation as a service APIs.</li> <li>• Data on the performance and scalability of ACROSS, including deployment time, resource utilization, and service delivery metrics.</li> <li>• Data on the impact of ACROSS on OPEX reduction and automation levels, compared to existing solutions.</li> <li>• Data on ACROSS adoption and user satisfaction, including feedback from stakeholders and market analysis.</li> </ul> |
| <p><b>Objective 3:</b> Design and implement an “ultra-sensing” open end-to-end telemetry infrastructure, featuring zero-touch telemetry reporting and collection at the speed of the underlying hardware, along with standardised telemetry reports and visualisation mechanisms.</p> <p><b>Related research data outputs: e</b></p> | <ul style="list-style-type: none"> <li>• Telemetry data collected from O-RAN cells, whitebox switches, SmartNICs, and edge/core NFs and service containers</li> <li>• Energy consumption data collected from programmable hardware</li> <li>• Network traffic data collected from Kubernetes netdata and prometheus agents</li> <li>• Data on monitoring accuracy and compute capacity efficiency resulting from offloading telemetry reporting into programmable hardware</li> <li>• Cost data related to the adoption of cheaper programmable whitebox hardware</li> <li>• Data on ACROSS reactivity to events, as measured by telemetry visibility into service microbursts</li> <li>• Telemetry reports and visualizations generated by ACROSS to enable real-time analytics and intelligence through AI</li> <li>• User feedback on the effectiveness and usability of ACROSS telemetry reporting and visualization mechanisms.</li> </ul>                                                                                                                                               |
| <p><b>Objective 4:</b> Develop an “ultra-instinct” unsupervised artificial intelligence stratum that will turn deep end-to-end telemetry data (Objective-3) into automated and proactive actionable decisions to empower the control loops of the cross-domain orchestration platform</p>                                            | <ul style="list-style-type: none"> <li>• Telemetry data from across the network, including performance metrics, traffic patterns, and network topology information.</li> <li>• Synthetic data generated to simulate different network conditions and scenarios, including different types of traffic, device failures, and security attacks.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

|                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(Objective-2) in a fully automated fashion (Objective 4).</p> <p><b>Related research data outputs: e</b></p>                                                                                                                                                                                                                                                                                   | <ul style="list-style-type: none"> <li>• Data from policy engines and other sources that can inform the AI system about business goals, customer requirements, and other contextual information.</li> <li>• Historical data on past network issues and outages, which can be used to train the AI system to recognize and respond to similar issues in the future.</li> <li>• Real-time data from sensors and other sources that can be used to detect anomalies and predict potential issues before they occur.</li> <li>• Data on energy usage and other resource consumption, which can be used to optimize the network and reduce costs.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <p><b>Objective 5:</b> Develop a broad range of zero-touch operations as programmable “hooks” in response to events originating from (i) raw infrastructure &amp; telemetry data (Objective 3) and/or (ii) produced intelligence (Objective 4), both translated into actionable decisions applied to the ACROSS orchestrator (Objective 2).</p> <p><b>Related research data outputs: f, g</b></p> | <ul style="list-style-type: none"> <li>• <b>Raw infrastructure and telemetry data:</b> This dataset can be collected from ACROSS infrastructure components, including servers, switches, routers, and other network devices. The telemetry data will include metrics such as CPU utilization, network bandwidth usage, memory usage, and storage usage.</li> <li>• <b>Produced intelligence data:</b> This dataset can be generated from ACROSS AI engines, which will analyse the raw infrastructure and telemetry data to identify patterns, anomalies, and other events that require attention.</li> <li>• <b>Stakeholder-driven zero-touch operations dataset:</b> This dataset can be generated by collecting policies created by ACROSS stakeholders. The policies will define zero-touch operations that can be injected through the ACROSS UI /NBIs.</li> <li>• <b>Device-driven zero-touch operations dataset:</b> This dataset can be generated by monitoring infrastructure and telemetry events and identifying trigger events that initiate zero-touch operations.</li> <li>• <b>Intelligence-driven zero-touch operations dataset:</b> This dataset can be generated by monitoring the ACROSS AI engine and identifying trigger events that initiate zero-touch operations.</li> <li>• <b>Composite zero-touch operations dataset:</b> This dataset can be generated by modelling chains of primitive zero-touch operations to realize complex tasks in an automated fashion.</li> <li>• <b>Zero-touch operations and classes dataset:</b> This dataset can be generated by mapping all zero-touch operations and classes into real test cases with tangible impact on ACROSS stakeholders, standards, and open-source communities.</li> <li>• <b>OPEX reduction dataset:</b> This dataset can be generated by comparing personnel costs traditionally spent for manual infrastructure onboarding and configuration, service maintenance, troubleshooting, and security before and after automation.</li> </ul> |



|                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                                                                                                                                                                                                                                                                         | <ul style="list-style-type: none"> <li>• <b>Automated edge infrastructure provisioning time dataset:</b> This dataset can be generated by monitoring the time it takes to provision edge infrastructure with ACROSS automation compared to existing solutions.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <p><b>Objective 6:</b> Design and implement security &amp; trust mechanisms to vertically fortify the ACROSS orchestration platform, focusing on (i) device integrity, (ii) secure data acquisition &amp; dissemination, and (iii) secure AI against attacks (e.g., DDoS heavy hitters, adversarial attacks).</p> <p><b>Related research data outputs:</b> h</p>        | <ul style="list-style-type: none"> <li>• <b>Device integrity logs:</b> Records of the state of various devices and their software components to detect any unauthorized changes or tampering.</li> <li>• <b>Network traffic logs packet captures and aggregated traffic flows:</b> Data collected from network traffic to detect any suspicious activities, including potential DDoS attacks.</li> <li>• <b>Adversarial sample datasets:</b> Datasets that include examples of data designed to fool AI classifiers and test the robustness of the system against adversarial attacks.</li> <li>• <b>Anomaly detection datasets:</b> Datasets used to train AI models to detect anomalies in telemetry data that could indicate a security breach or attack.</li> <li>• <b>Malware and virus signature databases:</b> Datasets containing signatures of known malware and viruses used by security systems to detect and prevent them.</li> <li>• <b>Threat intelligence feeds:</b> Datasets that provide information on known and emerging threats to help improve the system's defence against attacks.</li> <li>• <b>Security event logs:</b> Records of security events and incidents detected by the system or reported by users to help identify and respond to security incidents.</li> <li>• <b>Privacy-preserving datasets:</b> The aim is to use synthetic data not associated with physical people. All the data created in the Data Generation Lab (Network Digital Twin) are not associated with people, so it is not necessary to anonymise it.</li> </ul> |
| <p><b>Objective 7:</b> Design, deployment, and validation of test cases, that exercise the entire range of zero-touch operations (see Obj. 5) across multiple geo-distributed vertical sectors over the ACROSS platform and real testbeds, while aligning each test case with realistic aspects of automation.</p> <p><b>Related research data outputs:</b> f, g, h</p> | <p>Objective 7 focuses on the design, deployment, and validation of test cases that exercise zero-touch operations across multiple geo-distributed vertical sectors over the ACROSS platform and real testbeds. It mentions the need for a Data Generation environment with bindings to both edge and core cloud environments, which provides realistic data augmentation/generation as a service and quality assurance for AI-based orchestration. Therefore, this environment could potentially generate datasets for testing and validation purposes, e.g.:</p> <ul style="list-style-type: none"> <li>• Description details of Use case scenarios for validation datasets</li> <li>• Deployment details dataset</li> <li>• Validation results datasets</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <p><b>Objective 8:</b> Maximise the impact and adoption of ACROSS through relevant standardisation and open-source</p>                                                                                                                                                                                                                                                  | <ul style="list-style-type: none"> <li>• Dataset of contributions made by ACROSS to relevant standards bodies and open-source communities,</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |

|                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>activities that will render ACROSS an exemplar orchestration platform with true market potentials.</p> <p><b>Related research data outputs:</b> i</p> | <p>including the nature of the contribution, the standard or community, and the impact of the contribution.</p> <ul style="list-style-type: none"> <li>• Dataset of ACROSS extensions made to commercial components, such as multi-domain end-to-end orchestrator and AI-based zero-touch engine, demonstrating the adoption and impact of ACROSS in the market.</li> <li>• Dataset of ACROSS leadership positions within standards bodies and open-source communities, including the duration and impact of each position.</li> <li>• Dataset of ACROSS contributions to whitepapers, studies, or specifications, including the impact of each contribution.</li> <li>• Dataset of ACROSS source code contributions to open-source projects, including the impact of each contribution.</li> <li>• Dataset of ACROSS inputs to chapters for work or study items, including the impact of each input.</li> <li>• Dataset of ACROSS acknowledged proofs-of-concept, demonstrating the practical applications and impact of ACROSS.</li> <li>• Dataset of ACROSS's leverage of its contributions to achieve or maintain leadership positions in standards bodies and open-source communities, including the impact of each leverage action.</li> </ul> <p>These datasets could be used to evaluate the impact and adoption of ACROSS and to identify areas for further improvement and development.</p> |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

### 2.3 Reusability of existing data

ACROSS uses and will reuse existing data where needed throughout the project life. Existing data will be used in various tasks and could be integrated with new data where appropriate. In such cases references to authors and institutions will be made available. Table 3 below shows datasets (including literature/journals) used or intended to be used:

*Table 3: Data reused or intended to be reused by ACROSS.*

| WP  | Data sources reused                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| WP2 | Requirements and specifications datasets at system level and at module and software component level <b>(a)</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| WP3 | <ul style="list-style-type: none"> <li>• Software tool datasets for the management and orchestration of the multi-domain orchestration processes <b>(b)</b></li> <li>• Software tool datasets for the management and orchestration of the distributed edge domain instances <b>(c)</b></li> <li>• Software tool datasets for the development and deployment of platform modules including: the integration fabric interfaces, the end-user interface, the northbound APIs, the application porting process, the telemetry mechanism and the linked data storage and analytics <b>(d)</b></li> </ul> |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| WP4 | <ul style="list-style-type: none"> <li>• Collected datasets from the creation and collection of events over the TID realistic network environment for the training of the AI algorithms and datasets of the processed outcomes after AI training <b>(e)</b></li> <li>• The AI-based event processing engines, the secure and trusted data governance and sharing processes, and the ZTP processes across the various layer and instances <b>(d)</b></li> </ul>                                    |
| WP5 | <ul style="list-style-type: none"> <li>• Design information datasets for the testbed deployment and the testing scenarios including methodologies on the execution of the scenarios for the collection of KPI parameters <b>(f)</b></li> <li>• Collected datasets from the performance validation of the functional use cases per use case scenario <b>(g)</b></li> <li>• Collected datasets from the deployment and evaluation of the use cases in industrial environments <b>(h)</b></li> </ul> |
| WP6 | Datasets on market and societal analysis of the developed solutions <b>(i)</b>                                                                                                                                                                                                                                                                                                                                                                                                                    |

- What is the origin of the data?
- What is the expected size of the data?
- To whom might it be useful ('data utility')?

Answers to the above questions are integrated in Table 4:

*Table 4: ACROSS origin, expected size, data utility of datasets collected/generated*

| Dataset         | Origin | Expected Size | Data Utility                                                                                                                       |
|-----------------|--------|---------------|------------------------------------------------------------------------------------------------------------------------------------|
| Project plan    | WP1    | A few MB      | Defines the work breakdown structure, schedule, resources, and overall project plan for ACROSS.                                    |
| Risk Register   |        | A few MB      | Identifies, assesses, and tracks project risks and provides mitigation strategies to avoid or minimize potential negative impacts. |
| Meeting Minutes |        | A few GB      | Documents the discussions and actions taken during project meetings to ensure accountability and communication among stakeholders. |

|                                                                                                                                                                                                                                                                             |                                          |           |                                                                                                                                                  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| Progress Reports                                                                                                                                                                                                                                                            |                                          | A few MB  | Provides regular updates on the status of the project, including achievements, challenges, and milestones, to stakeholders and funding agencies. |
| Requirements and specifications datasets at system level and at module and software component level                                                                                                                                                                         | WP2                                      | 10-100 MB | Used to define and specify the ACROSS platform architecture and ecosystem.                                                                       |
| Software tool datasets for the management and orchestration of the multi-domain orchestration processes                                                                                                                                                                     |                                          | 50-500 MB | Used for managing and orchestrating multi-domain orchestration processes.                                                                        |
| Software tool datasets for the management and orchestration of the distributed edge domain instances                                                                                                                                                                        |                                          | 50-500 MB | Used for managing and orchestrating distributed edge domain instances.                                                                           |
| Software tool datasets for the development and deployment of platform modules including: the integration fabric interfaces, the end-user interface, the northbound APIs, the application porting process, the telemetry mechanism and the linked data storage and analytics |                                          | 50-500 MB | Used for developing and deploying ACROSS platform modules and functionalities.                                                                   |
| Software tool datasets for the management and orchestration of the multi-domain orchestration processes                                                                                                                                                                     |                                          | < 500 GB  | High <sup>1</sup>                                                                                                                                |
| Software tool datasets for the management and orchestration of the distributed edge domain instances                                                                                                                                                                        | WP3/developed by the consortium partners | < 300 GB  | High                                                                                                                                             |
| Software tool datasets for the development and deployment of platform modules including: the integration fabric interfaces, the end-user interface, the                                                                                                                     |                                          | < 1 TB    | High                                                                                                                                             |

<sup>1</sup> When "High" is indicated under utility for a dataset, it suggests that the data is highly valuable, relevant, and beneficial for the intended purpose within the work package. A "High" utility rating indicates that the dataset is of substantial value and plays a crucial role in the successful execution of the tasks and goals within the specific work package.

|                                                                                                                                                                         |                                                                                       |                                  |                                                                                                                                   |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| northbound APIs, the application porting process, the telemetry mechanism and the linked data storage and analytics                                                     |                                                                                       |                                  |                                                                                                                                   |
| Events dataset for AI training                                                                                                                                          | WP4/ Collected from TID realistic network environment                                 | Large                            | Used for training AI algorithms                                                                                                   |
| Processed outcomes dataset after AI training                                                                                                                            |                                                                                       |                                  | Used for evaluating the performance of AI algorithms                                                                              |
| Performance validation datasets                                                                                                                                         | WP4/ Collected from the deployment of the use cases in industrial environments        |                                  | Used for evaluating the performance of the use cases                                                                              |
| Dataset of triggering events for ZTP and required correction/mitigation/optimization actions                                                                            | WP4/ Collected from knowledge owned by network and/or infrastructure operators/owners |                                  | Used for designing the network and infrastructure automation workflows the ZTP component of the ACROSS platform needs to support. |
| Dataset of relevant security validations, attestations and enforcements                                                                                                 | WP4/ Collected from security and infrastructure experts                               |                                  | Used for designing the security and trust measures to be featured by the ACROSS platform.                                         |
| Market and societal analysis datasets                                                                                                                                   | WP4/ Collected from market and societal analysis of the developed solutions           | A few MB                         | Used for understanding the impact of the developed solutions on the market and society                                            |
| Testbed design information datasets for the deployment and testing scenarios, including methodology for the execution of scenarios for the collection of KPI parameters | WP5                                                                                   | To be specified at a later stage | Provides information on how the testbed was designed and how the testing scenarios were executed to collect KPI parameters        |
| Collected datasets from the performance validation of functional use cases per use case scenario                                                                        |                                                                                       | To be specified at a later stage | Provides data on the performance of functional use cases in the testbed                                                           |

|                                                                                                   |     |                                  |                                                                                                                         |
|---------------------------------------------------------------------------------------------------|-----|----------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| Collected datasets from the deployment and evaluation of the use cases in industrial environments |     | To be specified at a later stage | Provides data on the performance of use cases in real-world industrial environments                                     |
| Dataset of the ACROSS multi-domain, multi-tenant testbed                                          |     | To be specified at a later stage | Provides information on the testbed used for integration and validation of the ACROSS platform                          |
| Methodology, scenarios and use cases for validating the ACROSS platform                           |     | To be specified at a later stage | Provides information on how the ACROSS platform was validated                                                           |
| Market and societal analysis datasets                                                             | WP6 | A few MB                         | Provide insights into the market potential and societal impact of the developed solutions                               |
| Dissemination and communication datasets                                                          |     | A few MB                         | Track the dissemination and communication activities of the project, including publications, conferences, and workshops |
| Exploitable market potential datasets                                                             |     | A few MB                         | Identify potential commercialization opportunities for the developed solutions                                          |
| Impact evaluation datasets                                                                        |     | A few MB                         | Evaluate the impact of the project and its solutions on the target industries and society as a whole                    |
| User feedback datasets                                                                            |     | A few MB                         | Collect feedback from end-users on the usability, functionality, and effectiveness of the developed solutions           |

|                                         |         |          |                                                                                                              |
|-----------------------------------------|---------|----------|--------------------------------------------------------------------------------------------------------------|
| Business models and strategies datasets |         | A few MB | Develop and test various business models and strategies for the commercialization of the developed solutions |
| Deliverable reports                     | All WPs | A few GB | Technical deliverables as they are described in the DoA and DoW.                                             |

### 3 FAIR data

#### 3.1 Making data findable, including provisions for metadata

ACROSS will not only provide FAIR data where applicable (data should be “as open as possible and as close as necessary”) but will mostly support the community in the uptake of research data sharing and practices, in alignment with FAIR principles. The following table provides information for making ACROSS data findable, including provisions for metadata to promote reuse.

*Table 5: Findable data including provisions for metadata*

|                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                              |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g., persistent, and unique identifiers such as Digital Object Identifiers)? | YES – DOI, unique and persistent URI.<br>Collected services' information will be available at the dedicated portfolio/catalogue system URL. Other aspects of the persistent identifiers could be implemented by periodical snapshots of the database that contains the data. |
| What naming conventions do you follow?                                                                                                                                                                                                 | ACROSS-[WP]-[title]-[ver]-[DDMMYYYY].[ext]                                                                                                                                                                                                                                   |
| Will search keywords be provided that optimize possibilities for reuse?                                                                                                                                                                | YES                                                                                                                                                                                                                                                                          |
| Do you provide clear version numbers?                                                                                                                                                                                                  | YES – accessed with unique and persistent URI                                                                                                                                                                                                                                |
| What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how?                                                                               | Metadata for General Research Data will follow the Dublin Core and DataCite Metadata Schema.                                                                                                                                                                                 |

#### 3.2 Making data openly accessible

By default, ACROSS will openly provide data produced following the principle “as open as possible, as closed as necessary”, to comply with ethical or security requirements and avoid related conflicting issues. The following table summarizes current considerations.

Table 6: Provisions for openly accessible data

|                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.</p> | <p>Metadata that will be created in the project includes information about the datasets used and produced, such as:</p> <ul style="list-style-type: none"> <li>• Dataset title and description</li> <li>• Creator and contributor information</li> <li>• Date of creation and modification</li> <li>• Spatial and temporal coverage</li> <li>• Data format and size</li> <li>• Data access and use conditions</li> <li>• Persistent identifiers (e.g., Digital Object Identifiers) if applicable</li> <li>• Provenance information (e.g., data sources, processing steps, quality control)</li> </ul> <p>If metadata standards do not exist in the discipline, the project team will create a metadata schema that includes the above elements and any additional relevant information. The metadata schema will be documented and shared with relevant stakeholders to ensure interoperability and compatibility with existing metadata standards and tools. The team may also consult with domain experts and metadata specialists to ensure that the metadata schema meets best practices for data discovery, accessibility, and reuse.</p> |
| <p>How will the data be made accessible (e.g. by deposition in a repository)?</p>                                                                                                                                                                                              | <p>The specific details of data accessibility will depend on the individual work packages and the nature of the data produced. However, as a general principle, the project should aim to make data as widely accessible as possible while also protecting any sensitive or confidential information.</p> <p>One common approach to data accessibility is deposition in a trusted repository. The choice of repository will depend on factors such as the type of data, the intended user community, and any relevant disciplinary standards or practices. Ideally, the chosen repository should provide a persistent and citable identifier (such as a DOI) and should adhere to relevant data management best practices, such as providing long-term preservation and metadata standards.</p> <p>In addition to deposition in a repository, the project may also consider other means of data</p>                                                                                                                                                                                                                                            |



|                                                                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                | <p>accessibility, such as making data available through a project website, data portal, or through integration with existing data infrastructures. The project should ensure that any data accessibility options selected are sustainable and meet the needs of both the data producers and users.</p> <p>Examples of repositories to be used:</p> <p><b>GitHub:</b> GitHub is a web-based hosting service for version control using git. It provides a platform for open-source software development and collaborative software projects.</p> <p><b>GitLab:</b> like GitHub, GitLab is also a web-based platform used for version control and collaborative software development. It provides a hosting service for Git repositories and supports continuous integration/continuous deployment (CI/CD) pipelines.</p> <p><b>Zenodo:</b> a general-purpose open-access repository that allows researchers to deposit data, software, and other research outputs. It is operated by CERN and offers a variety of features including persistent identifiers (DOI), versioning, and integration with other services such as GitHub.</p> |
| <p>What methods or software tools are needed to access the data?</p>           | <p>All three platforms (GitHub, GitLab, Zenodo) support Git repositories, which are commonly used for version control and collaboration in software development projects. Users can access the data through Git commands or through web-based interfaces provided by each platform. Zenodo is also specifically designed for research outputs and provides features such as persistent identifiers and versioning to ensure the data is discoverable and citable.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <p>Is documentation about the software needed to access the data included?</p> | <p>The consortium partners will take the necessary steps to include documentation about the software needed to access the data. This can help ensure that others can easily understand and use the software. The documentation will include information on any dependencies, installation instructions, usage instructions, and troubleshooting tips. It is also useful to include examples and tutorials to help users get started.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

|                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Is it possible to include the relevant software (e.g., in open-source code)?</p>       | <p>Yes, it is possible to include the relevant software in open-source code if the software meets the open-source criteria. By making the software open source, it can be freely accessed, used, modified, and distributed by anyone, including other researchers who may want to build upon it or reuse it for their own purposes. This can help to promote transparency, reproducibility, and collaboration in research.</p> <p>However, if the software is not intended to be open source, then it may not be possible to include it as such. In this case, other options for sharing the software may need to be considered, such as providing access to the software through a restricted access platform or sharing the software with collaborators directly.</p> |
| <p>Where will the data and associated metadata, documentation and code be deposited?</p>  | <p>Dedicated repository in Zenodo, GitHub, GitLab (or else) . In case of contributing to open-source software, the existing open-source code repositories.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| <p>Have you explored appropriate arrangements with the identified repository?</p>         | <p>For the technical WPs: WP2 – WP5 a GitLab repo is run and maintained by the WP2, WP3 leader UBITECH.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <p>If there are restrictions on use, how will access be provided?</p>                     | <p>If there are restrictions on use of the data, access will be provided according to the specified terms and conditions set by the data owner or the repository where the data is deposited (e.g., access will be given using user authentication and authorization, handling users' verification and access levels.). This may include requiring users to agree to a license or providing access only to authorized individuals or organizations. The project team will work with the repository and data owner to ensure that access is provided in accordance with any restrictions. Any such restrictions will be clearly communicated in the associated metadata and documentation to ensure that users are aware of them before accessing the data.</p>          |
| <p>Is there a need for a data access committee?</p>                                       | <p><b>YES</b> – there are GDPR issues associated</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <p>Are there well described conditions for access (i.e., a machine-readable license)?</p> | <p>Conditions for access will be described when the access policies are finalized.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <p>How will the identity of the person accessing the data be ascertained?</p>             | <p>The method of ascertaining the identity of the person accessing the data will depend on the</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

|  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | <p>specific policies and procedures established by the repository or access control mechanism in use. For example, some repositories may require users to create an account and provide identification information such as name, affiliation, and email address. Other repositories may use a single sign-on mechanism that authenticates users through an existing identity provider, such as a university or research institution. In some cases, access to sensitive or restricted data may require additional verification steps, such as providing proof of employment or obtaining approval from a data access committee.</p> |
|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

### 3.3 Making data interoperable

Table 7: ACROSS provisions for making data interoperable

|                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Are the data produced in the project interoperable, that is allowing data exchange and reuse between researchers, institutions, organisations, countries, etc. (i.e., adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating recombinations with different datasets from different origins)?</p> | <p>The project aims to adhere to relevant standards for data formats to ensure interoperability between different datasets and facilitate data exchange and reuse between researchers, institutions, organizations, countries, etc. Wherever possible, open-source software applications will be used to ensure compatibility and facilitate recombination with different datasets from different origins. In addition, efforts will be made to ensure that the data produced by the project can be integrated with existing datasets and infrastructure, particularly those related to the telecommunications industry.</p>                                                                                                                                                                                                                                                |
| <p>What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?</p>                                                                                                                                                                                                                                                                   | <p>To make the data produced in the project interoperable, the project will follow relevant vocabularies, standards, and methodologies such as:</p> <p><b>OpenAPI:</b> The OpenAPI specification is a widely used standard for describing RESTful APIs. Using OpenAPI, we can provide a machine-readable description of the API, which makes it easier for other researchers and organizations to integrate with our data.</p> <p><b>Resource Description Framework (RDF):</b> RDF is a standard model for data interchange on the web. It provides a framework for representing data and metadata in a machine-readable format, enabling data interoperability between different systems.</p> <p><b>Dublin Core:</b> The Dublin Core Metadata Initiative provides a set of metadata standards for describing resources on the web. Using Dublin Core, we can provide a</p> |

|                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                                                                             | <p>standardized set of metadata for our data, making it easier for other researchers to find and use our data.</p> <p><b>Schema.org:</b> Schema.org provides a set of structured data vocabularies for use in web pages. By using Schema.org, we can provide structured data for our web pages, making it easier for search engines to understand and interpret our data.</p> <p><b>Linked Data:</b> Linked Data is a set of best practices for publishing and connecting structured data on the web. By following Linked Data principles, we can make our data more discoverable and easily integrable with other datasets.</p> <p><b>FAIR data principles:</b> The FAIR (Findable, Accessible, Interoperable, Reusable) data principles provide a framework for making research data more findable, accessible, interoperable, and reusable. By following the FAIR principles, we can ensure that our data is maximally interoperable and reusable by other researchers and organizations.</p> |
| <p>Will you be using standard vocabularies for all data types present in your data set, to allow interdisciplinary interoperability?</p>                                    | <p><b>Yes</b>, although it may evolve dynamically during the project lifetime to ensure an ontology alignment within the EOSC.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <p>In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?</p> | <p><b>Yes</b>, if uncommon or project specific ontologies or vocabularies are used.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

### 3.4 Increase data re-use (through clarifying licenses)

Table 8: ACROSS data licences and provisions for data re-use.

|                                                                                                                                                                                                                                                     |                                                                                                                                                                                                                                                                                                                                 |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>How will the data be licensed to permit the widest re-use possible?</p>                                                                                                                                                                          | <p>Probably one of Creative Commons license options, but this will be decided when a full picture of the data will be available. The license for each dataset will be one providing the widest re-use possible.</p> <p>For software, possibly Apache or GNU. More information will be specified in a following DMP version.</p> |
| <p>When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.</p> | <p>The data will be made available for re-use as soon as the final/publishable version of the data is available.</p>                                                                                                                                                                                                            |

|                                                                                                                                                                               |                                                                                                                                                                                                |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why. | This will also be made clear as soon as there is a good view of the datasets achieved.                                                                                                         |
| How long is it intended that the data remains re-usable?                                                                                                                      | Data is intended to remain re-usable for as long as it is allowed by project resources and infrastructure. This will be reviewed as the project progresses.                                    |
| Are data quality assurance processes described?                                                                                                                               | Processes are not described, but effort will be put in providing quality data. The following data quality metadata will be provided: accuracy, relevance and consistency for the Stakeholders. |

## 4 Allocation of resources

*Table 9: Resources allocated for data management and making data FAIR in ACROSS.*

|                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| What are the costs for making data FAIR in your project?                                                                                                                     | <p>The costs for making data FAIR can vary depending on several factors, such as the size and complexity of the data, the level of documentation and metadata required, the type of repository used for data deposition and preservation, and any costs associated with open access publishing. Some potential costs to consider when making data FAIR in a project include:</p> <p><b>Data management and curation costs:</b> These can include costs associated with data storage, backups, and security, as well as costs for developing and maintaining data documentation and metadata.</p> <p><b>Repository costs:</b> If using a repository for data deposition and preservation, there may be costs associated with repository fees, data storage fees, and fees for data retrieval or access.</p> <p><b>Software and tools:</b> Depending on the complexity of the data and the analysis methods used, software and tools may be required to support data analysis, visualization, and dissemination. These tools may be open source or proprietary and may have associated costs.</p> <p><b>Staff time:</b> Developing and implementing data management plans, creating documentation and metadata, and preparing data for sharing and publication can be time-intensive tasks, and may require additional staff time or resources.</p> |
| How will these be covered? Note that costs related to open access to research data are eligible as part of the HEU grant (if compliant with the Grant Agreement conditions). | This is under consideration and further information will be specified in a following DMP version.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

|                                                                                                                                                  |                                                                   |
|--------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|
| Who will be responsible for data management in your project?                                                                                     | Project Coordinator NOVA                                          |
| Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)? | Resources for long term preservation have not yet been discussed. |

## 5 Data security

Table 10: Data security provisions in ACROSS.

|                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)? | Data collected/generated are stored in a ACROSS cloud workspace, where every file stored is maintained and encrypted using AES 256-bit encryption in geographically diverse areas. HTTPS protocol is used for secure communication between endpoints as a standard. It is the usual HTTP which runs on top of encrypted sockets (SSL/TLS) on the transport layer of the network stack (TCP/IP). Data will be moved to Zenodo repositories for long term preservation.                                                                                                |
| Is the data safely stored in certified repositories for long term preservation and curation?                                       | <b>YES</b> – For the Zenodo repositories all files uploaded to Zenodo are stored in CERN's EOS service in an 18 petabytes disk cluster. Each file copy has two replicas located on different disk servers. For each file they store two independent MD5 checksums. One checksum is stored by Invenio, used to detect changes to files made from outside of Invenio. The other checksum stored by EOS, is used for automatic detection and recovery of file corruption on disks. For ACROSS data periodical database snapshots will be made and stored independently. |

## 6 Ethical aspects

Table 11: Ethical aspects related to data sharing in ACROSS

|                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA). | Task 1.3 is dedicated for this and it defines how research will be executed in the project regarding the ethics issues during the implementation of the Project (including, but not limited to, confidentiality, integrity, validity, objectivity, accuracy, transparency, trustworthiness, authenticity, respect for autonomy, reciprocity and equity), in collaboration with the project partners and the independent ethics committee that will be established and operate during the project implementation to closely monitor and consult the consortium with regards to any activity involving ethics issues. The task will also examine legal and regulatory issues related to project implementation and potential barriers that arise from them. |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|                                                                                                                        |    |
|------------------------------------------------------------------------------------------------------------------------|----|
| Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data? | No |
|------------------------------------------------------------------------------------------------------------------------|----|

### Other issues

Do you make use of other national/funder/sectorial/ departmental procedures for data management?  
If yes, which ones?

- NO

## 7 Conclusions

This deliverable is the first version of the ACROSS Data Management Plan and contains an initial description of datasets collected/generated in this first stage of the project. The described datasets may be of value for the project and will be exploited by the different tasks through the course of the project. The document will be updated as the list of datasets is enriched with new information or datasets. Datasets use, sharing, preservation and dissemination aspects will be specified in all cases. All this updated information will be included in the future versions and revisions of the current document.

---

<sup>1</sup> <https://open-research-europe.ec.europa.eu/for-authors/data-guidelines>

<sup>2</sup> <https://fairsharing.org/search?fairsharingRegistry=Standard&q=metadata>