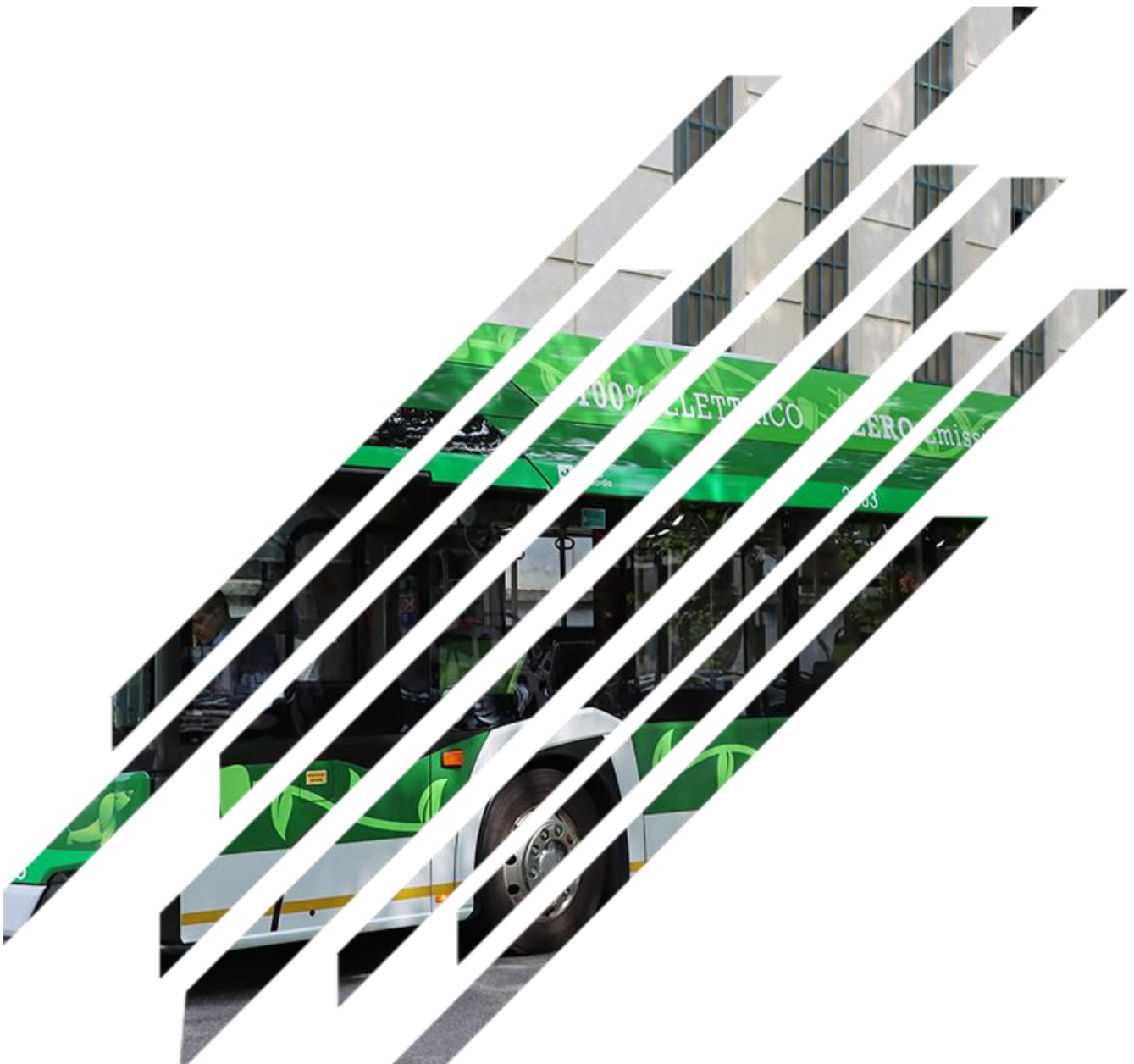


Smart Public transport Initiatives for climate  
Neutral cities in Europe

# D7.2 - SPINE Data Management Plan

## Version 1





## Document Summary Information

Grant Agreement No	101096664	Acronym	SPINE
Full Title	Smart Public transport Initiatives for climate Neutral cities in Europe		
Start Date	01/01/2023	Duration	48 months
Project URL	<a href="https://www.spine-project.eu/">https://www.spine-project.eu/</a>		
Deliverable	SPINE Data Management Plan Version 1		
Work Package	7		
Contractual due date	<u>30.06.2023</u>	Actual submission date	29.06.2023
Nature	DPM, Data Management Plan	Dissemination Level	PU-Public
Lead Beneficiary	MOBY.X		
Responsible Author	Giannis Tsouros		
Contributions from			

## Disclaimer

The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the documents is believed to be accurate, the author(s) or any other participant in the SPINE 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 SPINE 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 SPINE 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

© SPINE Consortium, 2023-2026. This deliverable contains original unpublished work except where clearly indicated otherwise.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.



This project has received funding from the Horizon Europe research and innovation programme under the Grant Agreement No 101096664.



## Revision history

Version	Issue Date	% Complete <sup>1</sup>	Changes	Contributor(s)
v0.1	3/3/2023	5	Initial Deliverable Structure	Giannis Tsouros (MOBYx)
V0.5	14/4/2023	20	Updated Deliverable Structure and Population of some sections	Giannis Tsouros (MOBYx)
V1.0	15/5/2023	90	First Complete version – for internal review	Giannis Tsouros (MOBYx)
V1.5	16/6/2023	95	Reviewed by George Tsakiris (KNT) and Sissi Koronaïou (INLE)	Giannis Tsouros (MOBYx)
V2.0	26/06/2023	100		Giannis Tsouros (MOBYx)

## Glossary of terms and abbreviations used

Abbreviation / Term	Description
<b>APCs</b>	Article Processing Charges
<b>CC BY</b>	Creative Commons Attribution (CC BY)
<b>DMP</b>	Data Management Plan
<b>DPO</b>	Data Protection Officer
<b>DT</b>	Digital Twins
<b>FAIR</b>	Findable, Accessible, Interoperable, Reusable
<b>GDPR</b>	General Data Protection Regulation
<b>KPIs</b>	key performance indicators
<b>LL</b>	Living Lab
<b>PII</b>	Personally identifiable information

<sup>1</sup> According to SPINE's Quality Assurance Process



## Table of Contents

<b>Document Summary Information</b> .....	2
<b>Disclaimer</b> .....	2
<b>Copyright message</b> .....	2
<b>Revision history</b> .....	3
<b>Glossary of terms and abbreviations used</b> .....	3
Executive Summary.....	5
1 Introduction.....	6
<b>1.1 Mapping SPINE Outputs</b> .....	6
<b>1.2 Deliverable Overview and Report Structure</b> .....	8
2. Data Summary.....	9
2.1 Purpose of Data Collection.....	9
2.2 Relation of Data to the project's objectives.....	9
2.3 Data sources in relation to Living Labs.....	10
2.3.1 Primary Data.....	10
2.3.2 Secondary Data.....	11
3 Fair data.....	13
3.1 Making data findable, including provisions for metadata.....	13
3.2 Making data openly accessible.....	14
3.3 Making data interoperable.....	14
4 Allocation of Resources.....	15
5 Data Security and Protection.....	16
5.1 Storage of sensitive data.....	16
5.2 Provisions for sharing of data amongst partners.....	16
5.3 Adherence to the General Data Protection Regulation.....	16
5.2.1 Implementing Ethical and GDPR Requirements in the SPINE Project.....	17
5.2.2 Data Anonymisation Guidelines.....	17
6 Ethical Aspects.....	19
6.1 Informed Consent.....	19
6.2. Consent forms and data anonymization.....	20
6.3 Appointment of Data Protection Officers.....	20
6.4 Exchanging, archiving and preservation of data.....	20
7 Open Science and Open Access.....	21
7.1. Open access publishing guidelines.....	21
7.2. Monitoring open science compliance.....	22
7.3 Ensuring access to datasets, databases, and repositories.....	23
8 Conclusions.....	24
9 References.....	25
Annex A: Consent form template.....	26
Annex B: Data anonymisation guidelines.....	28
Annex C: SPINE's Data Protection Officers.....	29

### List of Tables

Table 1: Adherence to SPINE's GA Deliverable & Tasks Descriptions.....	6
--	---



## Executive Summary

This Data Management Plan (DMP) is an important document for SPINE. It outlines the protocols and procedures that will be followed to manage the data generated and collected throughout the project's lifecycle. The DMP is designed to ensure the best practices of data management, aligning with the overarching goals and objectives of the project.

The DMP begins with a comprehensive Data Summary (Section 2), detailing the purpose of data collection and its relation to the project's objectives. It describes the data sources, both primary and secondary, that are associated with the project's Living Labs and twinning cities.

The plan adheres to the FAIR data principles, covering aspects related to making data findable, accessible, and interoperable (Section 3). This includes provisions for metadata and guidelines for open access to promote transparency and collaborative research.

Section 4 outlines the Allocation of Resources, detailing the commitment and role of different stakeholders within the project, ensuring the successful implementation of the data management strategy. This section is closely tied with the tasks of learning from Living Labs and identifying barriers, ensuring the appropriate resources are allocated to achieve the project's objectives.

Data Security and Protection (Section 5) is a key focus of the DMP, highlighting the strategies for storage of sensitive data, provisions for data sharing among partners, and adherence to the General Data Protection Regulation.

Ethical considerations are addressed in Section 6, covering aspects such as informed consent, guidelines for participation in Living Labs, consent forms and data anonymization, and appointment of Data Protection Officers.

Lastly, the DMP emphasizes the commitment of SPINE towards Open Science and Open Access (Section 7). It provides guidelines for open access publishing and monitoring open science compliance, as well as ensuring access to datasets, databases, and repositories.

In summary, this DMP is a living document that will be updated as SPINE progresses, to reflect the evolving data management needs and to align with any updates in the data management policies and guidelines. It is an important instrument for SPINE to achieve its objectives in an efficient, ethical, and transparent manner.



# 1 Introduction

## 1.1 Mapping SPINE Outputs

Purpose of this section is to map SPINE's Grant Agreement commitments, both within the formal Deliverable and Task description, against the project's respective outputs and work performed.

Table 1: Adherence to SPINE's GA Deliverable & Tasks Descriptions

SPINE GA Component Title	SPINE GA Component Outline	Respective Document Chapter(s)	Justification
DELIVERABLE			
D7.2 SPINE Data Management Plan, Version 1	<b>Initial Data Management Plan with definition of data and characterization procedures</b>	Chapter 2	<i>Chapter 2 (Data Summary) would cover the definition of data. Specifically, Section 2.1 (Purpose of Data Collection) and Section 2.2 (Relation of Data to the project's objectives) would discuss why and how data is collected and used in the project. Additionally, Section 2.3 (Data sources in relation to Living Labs) and the sub-sections under it (2.2.1 Primary Data, 2.2.2 Secondary Data) would likely discuss the characterization procedures of data.</i>
	and <b>related EC policies.</b>	Chapters 3,5,6,7	<i>The EC policies would be discussed in Chapter 3 (Fair data) through Sections 3.1, 3.2, and 3.3 that discuss the principles of making data findable, accessible, and interoperable respectively, which are core principles of the EC's open data and fair data policies.</i>  <i>Additionally, EC policies on data security and protection would be covered in Chapter 5 (Data Security and Protection), with specific sections on storage of sensitive data, data sharing provisions, and adherence to the General Data Protection Regulation (GDPR).</i>  <i>Chapter 6 (Ethical Aspects) and Chapter 7 (Open Science and Open Access) would also address related EC policies in terms of ethical considerations and open science compliance.</i>



TASKS			
<b>Task 7.4 Objective 1</b>	To create, maintain and update a data management plan	Section 2: Data Summary, Section 3: Fair data, Section 4: Allocation of Resources, Section 5: Data Security and Protection, Section 6: Ethical Aspects, Section 7: Open Science and Open Access	These sections outline the plan for data management, including the purpose and sources of data, how data will be made findable, accessible, and interoperable, how resources will be allocated for data management, how data security and protection will be ensured, and how ethical aspects and open science principles will be adhered to.
<b>Task 7.4 Objective 2</b>	To ensure compliance with the FAIR principles	Section 3: Fair data	This section specifically addresses how data will be made findable, accessible, interoperable, and reusable in accordance with the FAIR principles.
<b>Task 7.4 Objective 3</b>	To ensure data security and compliance with GDPR	Section 5: Data Security and Protection	This section outlines the provisions for data security, including storage of sensitive data, sharing of data amongst partners, and adherence to the General Data Protection Regulation.
<b>Task 7.4 Objective 4</b>	To manage ethical aspects of data collection and use	Section 6: Ethical Aspects	This section outlines the ethical considerations regarding informed consent, guidelines for participation in Living Labs, consent forms and data anonymization, and appointment of Data Protection Officers.
<b>Task 7.4 Reporting and Updating</b>	Produces two further versions (2 and 3) of the DMP in M18 and M30, respectively. Ensures that the DMP remains a "living document" by having partners revise the data templates based on recent and relevant information.	Section 2.5 & Section 5.1	Section 2.5 provides an overview of the reporting and updating process for the DMP, while Section 5.1 outlines the specific steps for updating the DMP throughout the project, including the production of two additional versions at M18 and M30, and the revision of data templates by partners to keep the DMP current and relevant.



## 1.2 Deliverable Overview and Report Structure

This deliverable is structured into seven key sections:

1. **Introduction:** This section provides a general overview of the deliverable and its objectives, setting the context for the subsequent sections.
2. **Data Summary:** This section describes the types of data collected and generated in the SPINE project. It details the purpose of data collection and its relevance to the project's objectives. The data sources, including primary and secondary data, associated with Living Labs are also discussed.
3. **FAIR Data:** This section outlines the project's approach to making data Findable, Accessible, Interoperable, and Reusable (FAIR). The provisions for metadata, data accessibility, and interoperability are discussed.
4. **Allocation of Resources:** This section explains the allocation of resources for data management within the project. It shows the connection to other tasks within the project, particularly those concerned with learning from Living Labs and identifying barriers to innovation.
5. **Data Security and Protection:** This section focuses on the mechanisms in place for data protection, including the storage of sensitive data, provisions for data sharing, and adherence to the General Data Protection Regulation (GDPR).
6. **Ethical Aspects:** This section covers ethical considerations related to data management, including informed consent, guidelines for participation in Living Labs, consent forms, data anonymization, and the appointment of Data Protection Officers.
7. **Open Science and Open Access:** This section outlines the project's commitment to open science and open access. Guidelines for open access publishing, monitoring open science compliance, and ensuring access to datasets are included here.





## 2. Data Summary

This section will provide an overview of the data that will be produced or collected during SPINE, including its origin (i.e., whether it is primary or secondary data), nature (e.g., qualitative or quantitative), and scope (e.g., its scale, relevance to the project's objectives, etc.). This section will also identify the data formats and standards that will be used and discuss the expected size and complexity of the datasets.

### 2.1 Purpose of Data Collection

The primary purpose of data collection and generation in SPINE is to inform, guide, and support the project's process into generating meaningful impact in the pilot cities and to succeed in the overall project objectives. The data will provide evidence-based insights into current public transport systems, user experiences, urban mobility patterns, and potential areas for improvement.

SPINE aims to transform urban mobility by introducing innovative solutions to public transport systems, with a particular focus on safety, efficiency, and user satisfaction. The collection of primary data, particularly from the Living Labs (via, for example the Citizen app), will provide valuable user perspectives and real-time feedback on implemented measures. This will allow for the iterative refinement of solutions based on direct user experience and needs, thus ensuring their relevance and effectiveness.

Secondary data from partner cities and existing databases will provide broader context, enabling the project to consider wider urban mobility trends, socio-economic factors, and existing infrastructure. This will support the project's aim to deliver solutions that are not only innovative but also pragmatic and adaptable to different urban contexts. In summary, data collection and generation in SPINE is integral to achieving its objectives. The data will ensure that the project's outputs are evidence-based, user-centered, and adaptable to the diverse needs of European cities.

### 2.2 Relation of Data to the project's objectives

The primary purpose of data collection and generation in SPINE is to facilitate the co-creation and co-management of innovative, affordable, inclusive, and resilient mobility solutions. This aligns with the project's objectives as follows:

**Objective 1 (O1):** The data will support the analysis of urban, suburban, peri-urban forms, and the public realm within which public transport systems operate. Information collected on existing conditions, transport policies, trends, discussions, and perceptions of current public transport services (R1.1-R1.4) will guide the development of innovative strategies and new approaches of significant impact.

**Objective 2 (O2):** Data collection is essential to prototyping, demonstrating, and evaluating innovative mobility solutions in the four Living Labs (LLs). Feedback and acceptance rates from participants will validate the demonstrations (R2.3), and the data gathered will contribute to long-term sustainability assessments.

**Objective 3 (O3):** The data generated will support the deployment of digital tools such as Digital Twins (DT), simulations, AI, behavioural models, and data-driven decision-making for public transport



mobility interventions. The data will be used in the SPINE Urban Mobility Data Space, Digital Twins of PT networks, and other tools (R3.1-R3.5).

**Objective 4 (O4):** The data collected in the Twinning Cities will aid in the replication and demonstration of solutions and business models from the four Leader Cities LLs (R4.1-R4.3). User feedback and acceptance rates will validate these demonstrations, with data contributing to long-term sustainability assessments.

**Objective 5 (O5):** Data will be essential for fostering dissemination, transferability, replication, and up-scaling of innovative solutions. It will guide the establishment of a Collaborative Network of LLs and the development of guidelines for scaling-up/transferability across Europe (R5.1-R5.2).

**Objective 6 (O6):** The collected data will contribute to updating the European transport policy framework, especially on the local and regional level (R6.1-R6.2). It will provide insights for documenting pathways and solutions and aligning the outcomes with recent advancements in spatial and mobility planning, SUMP, and SULP.

## 2.3 Data sources in relation to Living Labs

### 2.3.1 Primary Data

Primary data will be collected directly from the Living Labs (LLs) within the project. This includes qualitative data derived from discussions and feedback from participants, as well as quantitative data from surveys or other measurement tools. A list of potential primary datasets from LLs and Twinning cities may include the following:

1. **User-generated data:** Living Labs often involve active participation from users, such as cyclists, pedestrians, or public transport users. These users generate data through their interactions with the transportation system, including feedback, preferences, opinions, and behaviour patterns (e.g., *the tracking data from the Citizen app*).
2. **Sensor data:** Living Labs may deploy various sensors throughout the transportation infrastructure to collect objective and quantitative data. This can include traffic flow data, vehicle speed and occupancy data, air quality measurements, noise levels, or other environmental parameters relevant to transport and mobility (e.g., *Parking availability data from parking sensors*).
3. **GPS and geolocation data:** By utilising GPS technologies and geolocation services, Living Labs can track the movements of individuals, vehicles, or other modes of transportation. This data helps analyse travel patterns, route choices, travel times, and other location-based information (e.g., *Telematics data from an existing service which will be integrated into SPINE solutions*).
4. **Surveys and questionnaires:** Living Labs often employ surveys and questionnaires to gather insights from participants. These can include demographic information, preferences, satisfaction levels, attitudes towards specific transportation options, or feedback on proposed interventions or prototypes (e.g., *user acceptance questionnaire for a new service in a LL*).



5. **Qualitative data:** Living Labs also collect qualitative data through methods such as interviews, focus groups, or observations. These methods capture in-depth insights into participants' experiences, perceptions, and suggestions related to transport and mobility (e.g., *Interview transcripts regarding safety in PT in a LL*).
6. **Co-creation artifacts:** Living Labs emphasize the active involvement of stakeholders in the design process. Therefore, they generate co-creation artifacts like sketches, prototypes, design concepts, or collaborative workshop outputs. These artifacts provide valuable data on the creative ideas and contributions of participants (e.g., *Co-created solutions design and prototypes*).
7. **Usage data from experiments/pilots:** Living Labs may implement SPINE solutions to test them in real-world settings. Data on the usage, performance, and impacts of these interventions are collected to assess their effectiveness/impact, identify potential improvements, and inform decision-making (e.g., *Impact assessment datasets*).

### 2.3.2 Secondary Data

Secondary data includes information sourced from external datasets, databases, and repositories. Each partner city and Living Lab will provide relevant secondary data, such as existing public transport usage statistics, urban planning data, demographic data, and other relevant socio-economic data.

Secondary data may include some or all of the following, based on availability and scope of collecting such datasets:

**Mobility data:** Information related to transportation modes, usage patterns, traffic flows, and public transit systems. This data can help understand the current mobility landscape and identify potential areas for improvement.

**Urban infrastructure data:** Details on existing infrastructure, such as roads, pedestrian walkways, bike lanes, and public transport networks. This data is crucial for designing and implementing smart mobility solutions.

**Environmental data:** Information on air quality, noise pollution, and greenhouse gas emissions within the partner cities. This data can be used to assess the environmental impact of mobility solutions and monitor improvements over time.

**Socio-economic data:** Demographic information, including population size, age distribution, income levels, and employment rates. This data can help identify specific user groups that may benefit from SPINE interventions and ensure that mobility solutions are tailored to local needs.

**Policy and regulatory data:** Information on existing mobility policies, regulations, and incentives at the city, regional, and national levels. This data is essential for understanding the regulatory context in which SPINE operates and for identifying opportunities to align with or influence relevant policies.

**Stakeholder data:** Contact details and background information of key stakeholders, such as local authorities, transport operators, and community organizations, involved in the Living Labs. This data will help facilitate effective communication and collaboration throughout the project.



**Public opinion data:** Surveys, interviews, or focus group data collected from citizens and stakeholders to gauge their perception of current mobility challenges, needs, and preferences. This data can help inform the design and implementation of SPINE interventions and ensure that they are aligned with local expectations and priorities.

Each Living Lab and partner city may have access to different types and amounts of secondary data, depending on their local context and data availability. It is essential to work closely with the Living Labs and partner cities to identify and access relevant data sources and ensure that they are used effectively to inform and support the SPINE's activities.



### 3 Fair data

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

All data that SPINE collects and generates, including primary data from Living Labs (LLs), secondary data sources required to be hosted internally, and other relevant data sources, will be stored and catalogued in the respective city partners' dataspace. This approach underpins SPINE's commitment to a decentralized data storage and exchange system, aligning with the International Data Spaces (IDS) Dataspace architecture.

Each data set within these localized dataspace will be labelled with standardized metadata, using a schema that aligns with both the IDS and FAIR data principles. This metadata will include a unique identifier, the data source, the date of collection/generation, the data format, and a brief description of the data content.

This decentralized approach offers several key advantages. Firstly, it supports better data sovereignty as each city partner retains control over their data. Secondly, it fosters improved scalability, allowing the system to expand seamlessly with the addition of new data sources. Lastly, it enhances data security by reducing the risk of a single point of failure, and it improves the speed of data access and processing by minimising data transit distances.

The proposed naming convention strategy for the primary and secondary data collected from Living Labs and partner cities can help ensure consistency, readability, and ease of use throughout SPINE is presented below.

##### LL\_City\_DataType\_Source\_Version\_Date

1. **LL:** An abbreviation for "Living Lab" to indicate that the data is related to a Living Lab.
2. **City:** A short abbreviation or code for the partner city (e.g., 'ANT' for Antwerp, 'LPL' for Las Palmas, 'HER' for Herakleion, etc.).
3. **DataType:** A brief description of the type of data (e.g., 'Mobility', 'UrbanInfra', 'Environment', 'SocioEcon', 'Policy', 'Stakeholder', 'PublicOpinion').
4. **Source:** An abbreviation or code representing the data source (e.g., 'Gov' for government, 'NGO' for non-governmental organization, 'Pvt' for private sector, 'Uni' for university, etc.).
5. **Version:** A version number to track updates and modifications to the dataset (e.g., 'v1', 'v2', 'v3', etc.).
6. **Date:** The date when the dataset was last updated, in the format 'YYYYMMDD' (e.g., '20230501' for May 1, 2023).

##### LL\_City\_EventType\_SessionID\_DataType\_Version\_Date

1. **LL:** An abbreviation for "Living Lab" to indicate that the data is related to a Living Lab.
7. **City:** A short abbreviation or code for the partner city city (e.g., 'ANT' for Antwerp, 'LPL' for Las Palmas, 'HER' for Herakleion, etc.).



2. **EventType:** A brief description of the event type (e.g., 'Workshop', 'FocusGroup', 'Interview', 'Survey', 'Observation').
3. **SessionID:** A unique identifier for the specific session or event (e.g., 'S01', 'S02', 'S03', etc.).
4. **DataType:** A brief description of the type of data (e.g., 'Transcript', 'Summary', 'Findings', 'Audio', 'Video', 'Images').
5. **Version:** A version number to track updates and modifications to the dataset (e.g., 'v1', 'v2', 'v3', etc.).
6. **Date:** The date when the dataset was last updated, in the format 'YYYYMMDD' (e.g., '20230501' for May 1, 2023).

Examples:

- **LL\_BOL\_Workshop\_S01\_Transcript\_v1\_20230501:** Transcript of a workshop session (Session 1) in the Bologna Living Lab, version 1, updated on May 1, 2023.
- **LL\_TAL\_FocusGroup\_S02\_Findings\_v2\_20230415:** Findings from a focus group session (Session 2) in the Tallinn Living Lab, version 2, updated on April 15, 2023.

### 3.2 Making data openly accessible

Data generated by SPINE will be made openly accessible to the maximum extent possible, taking into account privacy, ethical, and commercial considerations. This will be facilitated through the SPINE Urban Mobility Data Space, which will provide a user-friendly interface for accessing and downloading data. Where necessary, data will be anonymized or aggregated to protect individual privacy before being made accessible. Accessible data will be provided in commonly used file formats to ensure easy use by external researchers and stakeholders.

For information regarding open science and the strategy towards dissemination of generated knowledge and datasets, please refer to Section 7 of this Deliverable.

### 3.3 Making data interoperable

To ensure that data sets can be easily combined and used together, all data will be stored in standardized, machine-readable formats. Where applicable, data will be coded using common data standards and vocabularies to ensure consistency across different data sets. Additionally, the metadata provided for each data set follow a standardized schema, ensuring that it can be understood and used by both humans and machines. This approach will ensure that data from SPINE can be easily used in combination with other data sources, facilitating cross-analysis and the development of innovative solutions.



## 4 Allocation of Resources

This section details how resources are allocated for data management in the project. Managing data effectively is a critical part of the project, ensuring that the data collected is reliable, accessible, and can contribute to SPINE's objectives.

### 4.1 Budget for Data Management and Data collection

SPINE has allocated a specific part of the budget for data management. This covers the costs of data collection, storage, processing, analysis, and dissemination. It also includes the resources necessary for the secure and ethical management of data. This is drafted for each partner and foreseen in the GA for data collection purposes.

### 4.2 Human Resources

Data management tasks are assigned to dedicated members of the project team who have the necessary skills and expertise. These tasks include data collection, data processing, data analysis, and data dissemination.

### 4.3 Infrastructure Resources

The project has allocated resources for the necessary infrastructure for data storage, processing, and analysis. This includes both physical infrastructure (like servers) and digital infrastructure (like cloud storage and processing services).

### 4.4 Link with Tasks T2.6 and T4.9

Tasks T2.6 (LLs lessons learned and barriers' identification) and T4.9 (LLs lessons learned and twinning assessment) play a critical role in the data management of the SPINE project. These tasks involve consolidating lessons learned, collecting generated knowledge, and documenting the barriers to implementing non-qualified solutions and the challenges faced when implementing the innovations.

The resources allocated for these tasks will significantly contribute to the data management of the project as they will help in identifying, collecting, processing, analysing, and storing valuable data from the project's Living Labs. This data will be crucial for understanding the effectiveness of the project's innovative mobility solutions and their potential for scalability and replicability.

### 4.5 Continuous Monitoring and Updating

The allocation of resources for data management is continuously monitored and updated throughout the project's lifecycle. This ensures that the resources are efficiently used and that the project's data management needs are met effectively.



## 5 Data Security and Protection

SPINE places great importance to data security and protection, especially as it involves the collection and processing of sensitive data. This section outlines the measures in place to ensure the secure storage of data, provisions for data sharing among partners, and adherence to the General Data Protection Regulation (GDPR).

### 5.1 Storage of sensitive data

In SPINE, we ensure the secure storage of all data, placing particular emphasis on the handling of sensitive information such as personal opinion data, attitudinal data, and GPS tracking data (location). The project will retain this personal data strictly for its duration, securely storing it in an encrypted format within a secure environment.

This information, coded to enhance its privacy, will never be shared with any party outside of the SPINE consortium without explicit consent from the individual in question. Each piece of sensitive data is further safeguarded through password access, underscoring our commitment to data security.

SPINE strictly upholds the principle of "user consent", respecting that it is the end-user who determines what personal or private data can be utilized. All data related to individual users will be collected and subsequently anonymized to eliminate any possibility of single-user identification.

### 5.2 Provisions for sharing of data amongst partners

Within the Consortium Agreement, special provisions are made to ensure the secure handling of sensitive/personal information, data, and code, as well as the protection of confidentiality through non-disclosure to third parties.

### 5.3 Adherence to the General Data Protection Regulation

The GDPR [3] is a comprehensive data protection regulation that applies to organisations operating within the European Union (EU) or processing the personal data of EU citizens. The regulation establishes several key requirements for organisations, including:

- a. Lawful, fair, and transparent processing:** Organizations must process personal data lawfully, fairly, and transparently. This involves obtaining valid legal grounds for processing, such as consent, contractual obligations, or legitimate interests, and providing clear and accessible information about how the data will be used.
- b. Purpose limitation:** Personal data must be collected for specified, explicit, and legitimate purposes, and not further processed in a manner that is incompatible with those purposes.
- c. Data minimisation:** Organisations should collect and process only the minimum amount of personal data necessary to achieve the specified purposes.
- d. Accuracy:** Personal data should be accurate and, where necessary, kept up-to-date. Inaccurate data should be corrected or deleted without delay.





**e. Storage limitation:** Personal data should be retained only for as long as necessary to fulfil the specified purposes. Organisations must implement appropriate retention policies and securely delete data when it is no longer needed.

**f. Integrity and confidentiality:** Organisations must ensure the security and confidentiality of personal data by implementing appropriate technical and organisational measures to protect against unauthorized access, disclosure, alteration, or destruction.

**g. Accountability:** Organisations are responsible for demonstrating their compliance with the GDPR by maintaining records of processing activities, conducting data protection impact assessments, and appointing a Data Protection Officer (DPO) if required.

### 5.2.1 Implementing Ethical and GDPR Requirements in the SPINE Project

To ensure the SPINE project's compliance with ethical principles and GDPR requirements, several measures should be undertaken, including:

- a. Developing and implementing a Data Management Plan (DMP) that outlines the project's data collection, processing, storage, and sharing practices.
- b. Obtaining informed consent from participants and providing clear information about their rights and the project's objectives.
- c. Ensuring anonymity and confidentiality of participant data through the use of anonymization techniques and secure data storage and transmission methods.
- d. Conducting risk assessments and implementing measures to minimize potential risks or harm to participants.
- e. Adhering to GDPR principles, such as data minimization, purpose limitation, and transparency, throughout the project.
- f. Appointing a Data Protection Officer (DPO) to oversee data protection activities and ensure compliance with GDPR requirements.
- g. Regularly reviewing and updating the project's ethical and data protection practices to address emerging concerns or changes in the regulatory landscape.

### 5.2.2 Data Anonymisation Guidelines

Data anonymisation is an essential aspect of ensuring privacy and compliance with GDPR requirements within the SPINE project. Anonymizing data involves removing personally identifiable information (PII) from datasets so that individuals cannot be identified, either directly or indirectly, through the remaining data. This section outlines the guidelines for data anonymisation in the SPINE project.

1. **Identifying Personally Identifiable Information (PII):** Project partners must identify all PII in their datasets. PII includes, but is not limited to, names, addresses, email addresses, phone numbers, identification numbers, and other unique identifiers. Partners should also consider



indirect identifiers that could be used in combination with other data points to re-identify individuals.

2. **Selecting Anonymisation Techniques:** There are several anonymisation techniques available, such as data masking, pseudonymisation, generalisation, and aggregation. The choice of technique depends on the nature of the data, the intended use of the anonymised data, and the desired level of privacy protection. Project partners must evaluate and select the most appropriate anonymisation technique(s) for their datasets.
3. **Assessing Re-identification Risks:** After applying the chosen anonymisation techniques, project partners must assess the risk of re-identification. This involves evaluating the likelihood that individuals can be re-identified through the remaining data or by combining the anonymised data with other publicly available datasets. If the re-identification risk is deemed too high, additional anonymisation techniques should be applied to further protect privacy.
4. **Data Minimisation:** Project partners should adhere to the principle of data minimisation, which involves collecting and processing only the minimum amount of data necessary to achieve the project's objectives. This helps reduce the amount of PII in the datasets and thus simplifies the anonymisation process.
5. **Data Retention:** Anonymised data should be retained only for as long as necessary to fulfil the project's objectives. Project partners must establish clear data retention policies and ensure that anonymised data is deleted or securely archived when it is no longer required.
6. **Documentation:** All anonymisation processes, techniques, and decisions must be thoroughly documented. This documentation should be retained as part of the project's records and may be required for compliance audits or other regulatory purposes.
7. **Training and Awareness:** Project partners should ensure that all team members who handle personal data are trained in data anonymization techniques and understand the importance of data privacy and compliance with GDPR requirements.



## 6 Ethical Aspects

### 6.1 Informed Consent

To ensure that SPINE adheres to ethical principles and safeguards the rights and well-being of participants in Living Labs, the following guidelines have been established:

**Informed Consent:** Obtain informed consent from all participants before initiating any research activities. Participants should be provided with clear and concise information about the purpose, procedures, and expected outcomes of the research. They should also be informed of their rights, including the right to withdraw from the study at any time without negative consequences.

**Anonymity and Confidentiality:** Protect the anonymity and confidentiality of participant data by using anonymisation techniques, such as pseudonymisation, encryption, and aggregation. Access to personal data should be restricted to authorised personnel only, and data should be securely stored and transmitted.

**Non-discrimination:** Ensure that the selection of participants is fair and unbiased, and that no individual or group is discriminated against based on factors such as age, gender, ethnicity, socioeconomic status, or disability.

**Gender-sensitive analysis:** Incorporate a gender-sensitive approach in the research design and implementation, recognizing and addressing the different roles, needs, and experiences of men, women, and people of other gender identities. Gender-sensitive analysis seeks to understand and account for gender differences and inequalities, helping to ensure that the research does not inadvertently reinforce or perpetuate gender stereotypes or biases.

**Respect for autonomy and dignity:** Treat all participants with respect and dignity, ensuring that their autonomy is upheld throughout the research process. This includes respecting their choices, preferences, and decisions regarding their involvement in the study.

**Risk minimisation:** Identify and assess potential risks to participants and implement measures to minimize or mitigate these risks. Researchers should monitor the progress of the study and promptly address any emerging concerns or unforeseen issues.

**Data protection and privacy:** Comply with GDPR and other applicable data protection regulations, ensuring that participant data is collected, processed, and stored in a manner that respects their privacy and protects their personal information.

**Transparent communication:** Maintain open communication with participants, providing them with updates on the progress of the research, as well as any relevant findings or outcomes. Participants should have the opportunity to ask questions, voice concerns, and provide feedback throughout the study.

**Ethical review and oversight:** Obtain ethical approval from relevant institutional review boards or ethics committees before initiating the research. Conduct regular reviews and evaluations of the research process to ensure continued adherence to ethical principles and guidelines.



## 6.2. Consent forms and data anonymization

Consent forms are an essential component of our research process. These forms provide detailed information about the research project, including its objectives, the data to be collected, the intended use of data, and the rights of participants. We obtain written consent from all participants before initiating any research activities. Moreover, we anonymise the data collected from participants to protect their privacy and ensure compliance with GDPR regulations.

## 6.3 Appointment of Data Protection Officers

To oversee and manage data protection and privacy concerns, we appoint Data Protection Officers (DPOs) for the SPINE project. DPOs are responsible for monitoring compliance with GDPR and other applicable data protection regulations (See Annex C). They act as the point of contact between the project and relevant data protection authorities, ensuring that the project adheres to the highest ethical standards and maintains open communication with stakeholders.

## 6.4 Exchanging, archiving and preservation of data

SPINE will utilise the SPINE Dataspace developed in T3.4 for data exchange, conforming to International Data Spaces (IDS) standards. The SPINE Dataspace is a decentralized system architecture that merges data from various connectors. The exchange of data within this system ensures seamless interoperability and secure access to data across the project. Connectors are developed for streaming and queuing message exchanges, linking to Living Labs data infrastructures, supporting data capture, and facilitating event-based data flows.

Archiving data is a responsibility primarily shouldered by the Data Providers, in our case, the Living Labs (LLs). However, the SPINE Urban Mobility Dataspace Reference Architecture, as defined in ST3.4.1, provides support by outlining the archiving process across multiple layers - business, functional, process, information, and system. Each layer contributes to an extensive data archiving structure that enables appropriate cataloguing, storage, and retrievability of data. In cases where the LLs lack persistent storage infrastructure, SPINE is considering strategies to assist with data archiving to ensure no crucial data is lost.

With respect to data preservation, though this responsibility predominantly belongs to the Data Providers, the SPINE Dataspace's multi-layered approach is designed to support the preservation of data's long-term integrity and accessibility. This approach provides a firm foundation to help sustain data in a decentralised system. The System Layer is particularly vital for data preservation, as it breaks down the logical software components, considering multiple facets such as integration, configuration, deployment, and extensibility of these components. This layer ensures that data is preserved in a sustainable format and is safely stored in a secure and reliable system. For LLs with limited capacity for data preservation, SPINE is exploring solutions to provide support while still maintaining the integrity and accessibility of the data.



## 7 Open Science and Open Access

### 7.1 Open access publishing guidelines

The concept of open access publishing has gained significant importance in recent years, driven by the need to promote the dissemination of knowledge and enhance transparency in scientific research. Open access publishing enables free, unrestricted online access to scholarly literature, ensuring that research findings are widely available to both the scientific community and the general public. This section provides an overview of open access publishing guidelines relevant to SPINE.

#### 2.1.1. Definition of Open Access Publishing

Open access publishing refers to the practice of making peer-reviewed scholarly research freely available online, without any financial, legal, or technical barriers to access. It allows readers to view, download, copy, distribute, print, search, or link to the full texts of published articles without requiring a subscription or payment. Open access publishing is crucial for fostering collaboration, accelerating innovation, and maximizing the impact of research.

#### 2.1.2. Open Access Publishing Policies in the European Union and SPINE

The European Union strongly supports open access publishing as a means of improving the visibility and accessibility of research outcomes. Under the Horizon Europe framework, all projects receiving funding are required to ensure that any peer-reviewed publications resulting from the project are made openly accessible. This mandate extends to the SPINE project and its consortium members.

To comply with this requirement, the SPINE project is committed to ensuring that all publications resulting from the project are made available through open access channels, either via the gold or green open access models. The gold open access model involves publishing articles directly in open access journals, while the green open access model allows authors to deposit a version of their published articles in an open access repository.

#### 2.1.3. Guidelines for Selecting Appropriate Open Access Journals and Repositories

To facilitate open access publishing, SPINE partners should follow these guidelines when selecting appropriate open access journals and repositories:

- Ensure that the chosen journal or repository is reputable, indexed in relevant databases, and complies with the principles of the Directory of Open Access Journals (DOAJ).
- Verify that the journal or repository supports the Creative Commons Attribution (CC BY) license, which allows for the widest possible dissemination and re-use of the published content.
- Choose a journal or repository with a clear and transparent policy on publication fees, waivers, and embargo periods.

#### 2.1.4. Embargo Periods, Licensing, and Copyright Considerations

Embargo periods refer to the time between the publication of an article in a subscription-based journal and its availability through an open access repository. While SPINE encourages immediate open access, some journals may impose an embargo period, typically ranging from 6 to 12 months.



Project partners should strive to negotiate the shortest possible embargo periods to ensure timely access to their research findings.

Regarding licensing, SPINE recommends adopting the Creative Commons Attribution (CC BY) license for all publications. This license allows others to freely distribute, adapt, and build upon the work, provided that the original authors are appropriately credited.

Project partners should also be aware of copyright considerations, including retaining the right to deposit their articles in open access repositories and ensuring compliance with the publisher's copyright and self-archiving policies.

#### 2.1.5. Support for Publication Fees and Waivers

Publication fees, also known as Article Processing Charges (APCs), are often associated with open access journals to cover the costs of peer review, editorial services, and online hosting. SPINE consortium acknowledges that these fees can be a barrier to open access publishing and is committed to supporting partners in securing funding for APCs, where possible. In cases where funding is not available, project partners are encouraged to negotiate fee waivers or discounts with publishers, or to consider alternative open access options that do not incur APCs

## 7.2. Monitoring open science compliance

To effectively monitor open science compliance, SPINE has established a set of key performance indicators (KPIs) that will be used to assess the level of adherence to open science principles. These KPIs include:

1. Number of open access publications produced by the project partners.
2. Percentage of datasets made available through open access repositories.
3. Degree of compliance with FAIR (Findable, Accessible, Interoperable, and Reusable) data principles.
4. Number of collaborations and partnerships established with external stakeholders and researchers.

The SPINE project will adopt a continuous monitoring approach, with partners required to report their progress in meeting the KPIs every three months. This reporting process will be facilitated through regular steering meetings where partners will present updates on their open science compliance efforts. The data collected during these meetings will be analysed and discussed to identify potential areas of improvement and to ensure that the project remains on track in terms of open science compliance.

In addition to the internal monitoring process, the SPINE project will also engage with external stakeholders to gather feedback on the project's open science initiatives. This feedback will be collected through surveys, workshops, and other engagement activities and will be used to further improve the project's open science compliance strategies.

Overall, the monitoring of open science compliance in the SPINE project is aimed at fostering a culture of openness and collaboration among project partners and ensuring that the project's



research outputs are widely accessible to the scientific community and other stakeholders. By establishing a robust monitoring process and engaging with both internal and external stakeholders, the SPINE project aims to achieve its open science objectives and maximise the impact of its research.

### 7.3 Ensuring access to datasets, databases, and repositories

To maximise the impact and reach of the SPINE project's research, it is crucial to ensure open access to the datasets, databases, and repositories generated or utilised throughout the project. This section outlines the strategies and procedures that will be put in place to guarantee the accessibility of these resources to the scientific community and other interested stakeholders.

1. Identification of relevant resources: The first step in ensuring access to datasets, databases, and repositories is to identify the resources that are generated or used within the SPINE project. This process will involve close collaboration between all project partners, who will provide information about the resources they are working with, their formats, and any associated metadata.
2. Adherence to FAIR data principles: To make the identified resources as accessible and reusable as possible, the SPINE project will adhere to the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles. This includes using standard data formats, providing comprehensive metadata, and employing persistent identifiers for all resources.
3. Selection of appropriate repositories: To ensure that the project's datasets, databases, and repositories are accessible to the wider scientific community, they will be stored in established and recognized open access repositories. The choice of repositories will depend on the specific requirements of each resource, such as subject area, format, and licensing.
4. Development of data sharing agreements: To comply with ethical and legal requirements, particularly those related to the GDPR, data sharing agreements will be developed for all resources that involve personal or sensitive information. These agreements will outline the terms and conditions for accessing and using the data, as well as any measures taken to protect the privacy of participants.
5. Documentation and training: To facilitate access to the project's datasets, databases, and repositories, the SPINE project will develop comprehensive documentation, including user guides and tutorials. Additionally, training sessions will be organised for project partners and external stakeholders to help them navigate and use the resources effectively.



## 8 Conclusions

Deliverable, D7.2 Data Management Plan, is a detailed guide to the data handling processes and practices to be employed within SPINE. It outlines the strategies for data collection, storage, accessibility, and protection, ensuring compliance with the FAIR principles and GDPR regulations. The DMP serves as a roadmap, guiding the consortium partners towards responsible data management, while facilitating collaboration and efficient use of data resources in the pursuit of the project's objectives.

This document is not static; it is a living document that will be updated and refined throughout the project lifespan at M18, M36, and M48, to reflect the evolving nature of the project and the lessons learned from the Living Labs and twinning cities. These updates will ensure the DMP's continued relevance and effectiveness, further ensuring that the project's data is managed in a manner that is both ethical and efficient.

The DMP works in concert with D7.3 Ethics Deliverable, which sets out the ethical standards and guidelines for the project. The principles outlined in D7.3 are integrated into the data management processes described in this DMP, reinforcing our commitment to ethical data handling. Future versions of the DMP will continue to align with the ethical standards set out in D7.3, ensuring that as our data management practices evolve, they remain rooted in our commitment to responsible and ethical conduct.





## 9 References

[1] FAIR Data Principles: Wilkinson, M. D., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

[2] Data Management in EU Projects: European Commission (2021). Guidelines on FAIR Data Management in Horizon 2020.

[3] **General Data Protection Regulation (GDPR)**: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).



## Annex A: Consent form template



### Participation Consent Form

#### **SPINE** “*Smart Public Transport Initiatives for climate-neutral Cities in Europe*”

I, the undersigned, hereby agree to participate in the SPINE project's local Living Labs, an initiative aimed at discussing and co-creating solutions to improve public transport in my area through a "think globally, act locally" approach. I understand that my participation is voluntary and that I may withdraw at any time without giving a reason.

The Living Labs will involve various activities and roles, fostering a diverse and active engagement of participants, including but not limited to:

1. Brainstorming sessions
2. Workshops
3. Interaction with digital tools
4. Games and simulations
5. Debates and discussions
6. Co-creation and evaluation of potential solutions
7. Engagement through various communication channels, both physical and virtual
8. Fluid roles, such as inventor, end-user, specialist, and planner

Please note that parts of the Living Labs' conference calls or live events may be recorded for record-keeping, analysis, and dissemination purposes. In such cases, you will be asked for your consent before any recording takes place.

#### **Data Protection and Confidentiality (GDPR Compliance)**

As a participant in the SPINE project's Living Labs, I understand that my personal data will be collected, processed, and stored in accordance with the General Data Protection Regulation (GDPR). I acknowledge that:



1. The data collected from me will be used solely for the purpose of this project and will not be shared with third parties without my explicit consent.
2. My personal data will be anonymized and aggregated with other participants' data for research and analysis purposes.
3. I have the right to access, rectify, or erase my personal data at any time by contacting the project team by sending an email request to the SPINE national Ethics and Data Protection Manager. In such case, the collect data should not be used for SPINE and must be erased, unless I agree that my personal data can be used for the project after being thoroughly anonymised.
4. All reasonable measures will be taken to ensure the security of my personal data, and any breaches will be reported promptly to the relevant authorities.
5. My personal data will be retained for the duration of the project, and upon its completion, it will be securely deleted or anonymized for future research purposes.

## Consent

I have read and understood the information provided above and agree to participate in the SPINE project's local Living Labs. I understand that my participation is voluntary, and I may withdraw at any time without giving a reason.

By signing this consent form, I confirm that I am at least 16 years of age (or have obtained the consent of a parent or guardian if under 18) and that I agree to the terms and conditions stated above.

Participant's Name (printed): \_\_\_\_\_

Participant's Signature: \_\_\_\_\_

Date: \_\_\_\_\_

Parent/Guardian's Name (printed, if applicable): \_\_\_\_\_

Parent/Guardian's Signature (if applicable): \_\_\_\_\_

Date (if applicable): \_\_\_\_\_

Please return the completed consent form to the SPINE project team before participating in the Living Labs. If you have any questions or concerns, please contact the project team at [Your Organisation's Contact Information].



## Annex B: Data anonymisation guidelines

Guideline	Dataset applicable to	Description of guideline	KPI to measure
1. Identifying PII		Identify all personally identifiable information in datasets, including direct and indirect identifiers	Number of PII data points identified
2. Selecting Anonymisation Techniques		Evaluate and select the most appropriate anonymisation techniques for the dataset based on data nature and intended use	Anonymisation techniques applied
3. Assessing Re-identification Risks		Evaluate the likelihood of re-identification after applying anonymisation techniques and adjust if necessary	Risk assessment conducted and documented
4. Data Minimisation		Collect and process only the minimum amount of data necessary to achieve the project's objectives	Amount of data collected and processed
5. Data Retention		Retain anonymised data only for as long as necessary and establish clear data retention policies	Maximum date of retention



## Annex C: SPINE's Data Protection Officers

<b>Partner</b>	<b>Data protection officer (DPO)</b>	<b>Contact</b>
<b>INLE</b>	Sissi Koronaiou	sissi.koronaiou@inlecomsystems.com
<b>ANTW</b>	Anne Barre	informatieveiligheid@antwerpen.be
<b>COBO</b>	Anna Lisa Minghetti	dpo-team@lepida.it
<b>TALL</b>	Merit Valgjärv	Merit.Valgjarv@tallinnlv.ee
<b>VALL</b>	Mónica Borrego	mborrego@ava.es
<b>ZILI</b>	Mgr. Róbert Alföldy	dualfi@dualfi.sk
<b>BARR</b>	Dr. António José Guimarães Madureira	dpo@cm-barreiro.pt
<b>COS</b>	Marin Miletić	marin.miletic@sibenik.hr
<b>HERA</b>	Karampatzaki Despoina	dpo@heraklion.gr
<b>GDYN</b>	Danuta Pietrzak	iod@gdynia.pl
<b>PALM</b>	Deborah Garcia	cinesi@cinesi.es
<b>SRM</b>	Tommaso Bonino	tommaso.bonino@srm-bologna.it
<b>AUVASA</b>	Mónica Borrego	mborrego@ava.es
<b>AKHE</b>	Stratoudaki Ioanna	asktelhr@gmail.com
<b>GUAG</b>	Ángel Placeres	angel.placeres@guaguas.com
<b>UAEGEAN</b>	Hortarias Nikolaos	dpo@aegean.gr
<b>CARNET</b>	Beatriz Gimeno	info.cit@upc.edu
<b>UA</b>	Koen Pepermans	privacy@uantwerpen.be
<b>AIM</b>	Mark Brackstone	Mark.brackstone@aimsun.com
<b>HU</b>	Anna Fredriksen	dataskydd@hh.se
<b>INSY</b>	Alina Tuktarova	alina@dipeeo.com
<b>CIVI</b>	Katerina Nikolopoulou	kn@civinet.gr
<b>ODRAZ</b>	Lidija Pavić-Rogošić	lidija@odraz.hr
<b>UG</b>	Andrzej Urban,	iod@ug.edu.pl
<b>CMO</b>	Floridea Di Ciommo	florideadc@gmail.com



<b>MOBY</b>	Ioannis Tsouros	j.tsouros@mobyx.co
<b>IST-ID</b>	Carlos Ribeiro	<a href="mailto:dpo@ist-id.pt">dpo@ist-id.pt</a>
<b>UNIZA</b>	Ghadir Pourhashem	Ghadir.Pourhashem@uniza.sk
<b>CINESI</b>	Deborah Garcia	cinesi@cinesi.es
<b>SAGULPA</b>	César García	cesargarcia@sagulpa.com
<b>KNT</b>	George Tsakiris	george.tsakiris@konnnecta.io
<b>PKA</b>	Mateusz Woźniak	iodo@pkagdynamia.pl
<b>YUNEX</b>	Manon Raap	manon.raap@yunextraffic.com
<b>EURNEX</b>	Celestino Sánchez	cesama@eurnex.eu
<b>EIP</b>	Liliana Andrei	liliana.andrei@eiproject.eu
<b>CITYWAY</b>	Myrtille Elmidoro	myrtille.elmidoro@cityway.fr
<b>PNO</b>	Jaap Pels	jaap.pels@pnoconsultants.com
<b>IBI</b>	Antulio Richetta	arichetta@IBIGroup.com
<b>ATOM</b>	Jurgen Sahtel	js@atommobility.com