Tec4MaaSEs

Technologies for Manufacturing as a Service Ecosystems

Deliverable 1.3

# Data Management Plan v1

WP1: Project Management

|  |  |
|---|---|
| Editor: | Armend Duzha |
| Lead beneficiary: | MAG |
| Version: | 1.0 |
| Status: | Final |
| Delivery date: | 30/06/2024 |
| Dissemination level: | PU (Public) |

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101138517.

# Deliverable Factsheet

| | |
|---|---|
| **Grant Agreement No.** | 101138517 |
| **Project Acronym** | Tec4MaaSEs |
| **Project Title** | Technologies for Manufacturing as a Service Ecosystems |
| **Start date** | 01/01/2024 |
| **Duration** | 36 months |

| | |
|---|---|
| **Deliverable Name** | D1.3 Data Management Plan v1.0 |
| **Related WP** | WP1 Project Management |
| **Due Date** | 30/06/2024 |

| | |
|---|---|
| **Author** | Armend Duzha (MAG) |
| **Contributor(s)** | Yiannis Mourtos (AEUB), Konstantinos Kaparis (UOM), Leif Stoermer (AIBEL), Maria Garcia (URA), Javier Coello, Iker Martinez de Zuazo (ERREKA), Patricia Casla, Jon Ander Sarasua, Itxaso Cascón (TEKNIKER), Isa Aksu (KAREL) |
| **Reviewer(s)** | Andreas Georgiou (UOM), Maunya Doroudi Moghadam (UIO) |
| **Approved by** | All partners |

**Disclaimer**

This document reflects the opinion of the authors only.

While the information contained herein is believed to be accurate, neither the Tec4MaaSEs consortium as a whole, nor any of its members, their officers, employees or agents make no warranty that this material is capable of use, or that use of the information is free from risk and accept no liability for loss or damage suffered by any person in respect of any inaccuracy or omission.

# Document History

| Version | Date | Author(s) | Organisation | Description |
|---|---|---|---|---|
| 0.1 | 08/03/2024 | Armend Duzha | MAG | ToC |
| 0.2 | 05/04/2024 | Armend Duzha | MAG | First complete draft |
| 0.3 | 26/04/2024 | Yiannis Mourtos | AUEB | Contribution to Section 2 |
| 0.4 | 29/04/2024 | Konstantinos Kaparis | UOM | Contribution to Section 2 |
| 0.5 | 10/05/2024 | Leif Stoermer | AIBEL | Contribution to Section 2 |
| 0.6 | 05/06/2024 | Maria Garcia | URA | Contribution to Section 2 |
| | | Javier Coello, Iker Martinez de Zuazo | ERREKA | |
| | | Patricia Casla, Jon Ander Sarasua, Itxaso Cascón | TEKNIKER | |
| 0.7 | 11/06/2024 | Isa Aksu | KAREL | Contribution to Section 2 |
| 0.8 | 18/06/2024 | Armend Duzha | MAG | Addition of the appendixes,<br>Final draft ready for internal review |
| 0.9 | 22/06/2024 | Andreas Georgiou | UOM | Peer review |
| 0.10 | 24/06/2024 | Maunya Doroudi Moghadam | UIO | Peer review |
| 0.11 | 28/06/2024 | Armend Duzha | MAG | Integration of comments from internal review |
| 1.0 | 30/06/2024 | Armend Duzha | MAG | Final version ready for submission |

# Executive Summary

This deliverable presents the first version of the Data Management Plan (DMP) for the Tec4MaaSEs project, which aligns with the EC guidelines for FAIR Data Management in Horizon Europe. It describes the data to be generated, collected and processed in the Tec4MaaSEs project and how these data are then managed and published.

The list of datasets expected to be generated, collected and processed during the project consists of pilot’s data and anonymous and aggregated statistics. These data are therefore subject to change, considering also the definition of the Tec4MaaSEs exploitation and business models. The publishing platforms selected are the project website, Zenodo for long-term archiving, and GitLab for open-sourced code.

Effective data management is a dynamic process that evolves with the project's advancements and discoveries. Therefore, the DMP is designed to be adaptable to integrate new data types, technological advancements, and changes in methodology that may arise during the project's lifespan.

The DMP emphasises ethical practices, respecting privacy, and ensuring GDPR compliance. The project prioritises transparency in data handling to ensure that all stakeholders, including participants and the broader scientific community, understand how and why data is used. The approach includes strict data security measures to protect the data's integrity and confidentiality while ensuring its availability and utility for the project's objectives.

# Table of Contents

## List of Tables

## Acronyms and Abbreviations

| Acronym | Description |
|---:|---|
| **DMP** | Data Management Plan |
| **DOI** | Digital Object Identifier |
| **EC** | European Commission |
| **EU** | European Union |
| **FAIR** | Findable, Accessible, Interoperable, Re-usable |
| **GDPR** | General Data Protection Regulation |
| **PCO** | Project Coordinator |
| **WP** | Work Package |
| **WPL** | Work Package Leader |

# 1 Introduction

This document serves as the Data Management Plan (DMP) for the Tec4MaaSEs project, presenting a comprehensive guide for managing data throughout the project's lifecycle. It outlines strategies and methods for handling various data types, collection and processing methodologies, data sharing and preservation strategies, and ethical and legal considerations, especially concerning privacy and security.

## 1.1 Purpose and Scope

The purpose of this document is to describe how data will be handled during the project lifetime. For all data generated, collected and processed during the Tec4MaaSEs project, a detailed description will be provided including the source, the standards and metadata used for data preservation and maintenance, as well as the process of how this data will be exploited and/or shared/made accessible for verification and re-use, in accordance with the EC Guidelines for FAIR Data Management in Horizon Europe [1]. The DMP also ensures adherence to legal and ethical standards, particularly the General Data Protection Regulation (GDPR) and related legislation.

This deliverable is a living document and will be periodically revised to align with the evolving needs and development with the Tec4MaaSEs project, ensuring compliance with current legislation and promoting a culture of transparency and collaboration in research. This is the first version of the document and the final version will be released in month 36.

## 1.2 Relation with other Deliverables

The DMP is vital to all work packages and their respective deliverables, underlining the ubiquity of data management in the Tec4MaaSEs project. It is a foundational and strategic reference, harmonizing data handling across various project stages. Beyond its overarching influence, the DMP is specialized in technical packages, ensuring high data generation, collection, and processing standards. It rigorously upholds ethical guidelines, particularly for sensitive data, and informs dissemination and communication activities by establishing clear protocols for data sharing. In exploitation activities, the DMP's guidelines enhance the accessibility and reusability of data, facilitating effective utilization of the project’s results. Similarly, for standardization activities, the DMP sets a consistent approach to data handling.

## 1.3 Structure of the document

The rest of the document is structured as follows:

- **Section 2** provides a comprehensive overview of the datasets to be used during the project, including information such as type and format, expected size and data utility.
- **Section 3** outlines the strategies to ensure the research data is findable, accessible, interoperable, and reusable.
- **Section 4** describes the resources required for effective data management.
- **Section 5** specifies the security measures to protect data integrity and prevent unauthorised access during and after the project end.
- **Section 6** presents the ethical principles and legal requirements.
- **Annex 1** provides the questionnaire used to collect the relevant information for creating the data management plan.

- **Annex 2** provides the data description template designed in compliance with the FAIR principles in order to collect, document, and share data systematically.

# 2 Data Summary

This section provides an overview of the existing or foreseeable datasets involved in the Tec4MaaSEs project. For each dataset and in accordance to the FAIR data management guideline [2], a description, name of the standards used for storage and metadata, and the chosen open access platform is provided.

## 2.1 Overview of datasets

During the project lifetime several datasets form various consortium members, representing different domains, will be produced. These may include, but are not limited to, publications, software artifacts, specifications, administrative data, and various other forms of data relevant to the project's activities. At this juncture, six months into the project duration, we provide a high-level overview of these datasets as presented in Table 1 below. However, it is important to notice that the specifics concerning these datasets, including their nature, extent, and utilisation, are subject to continuous refinement as the project unfolds.

**Table 1: Datasets to be generated/collected in Tec4MaaSEs**

| Name of the dataset | Owner(s) | Accessibility |
|---|---|---|
| Optimisation service datasets | AEUB | *Open Access + Restricted Access* |
| Analytics service datasets | UOM | *Open Access + Restricted Access* |
| Value Network 1 dataset(s) | KAREL | *Open Access + Restricted Access* |
| Value Network 2 dataset(s) | URA, ERREKA, TEKNIKER | *Open Access + Restricted Access* |
| Value Network 3 dataset(s) | AIBEL | *Open Access + Restricted Access* |
| Public deliverables | | *Open Access* |
| Scientific publications | | *Open Access* |
| Software components | | *Open Access* |

As the project evolves, this table will be updated and the final version of the DMP in month 36.

## 2.2 Description of datasets

For Tec4MaaSEs, we have developed a data description template (see Annex 2) following the FAIR principles described in Section 3 in order to collect, document, and share data systematically. This template caters to various data types relevant to the project, ensuring uniformity and adherence to high data management standards.

The project partners, who in many cases will serve as data providers, have engaged in preliminary identification of the datasets expected to be generated or utilised throughout the project's lifecycle. To do so, we have designed a specific questionnaire (see Annex 1) following the structure of the template for Horizon Europe projects. However, these initial identifications are contingent on their current understanding of the project's scope and may evolve as the project progresses and specific requirements become more defined. Therefore, the details provided in the subsequent subsections are considered provisional, and any modifications or updates will be diligently documented in future iterations of this deliverable.

### 2.2.1 Optimisation service datasets

Table 2: Optimisation service datasets

| | |
|---|---|
| Name of dataset | Optimisation service datasets |
| Responsible partner | AUEB |
| ID | Not available yet |
| Description | A brief description of the dataset |
| Purpose | The optimization related data will be generated and utilised in line with the goal of T3.4 needs for optimization tools and methods for re-configuring value networks on a regular basis or ad hoc when disruptive conditions occur. |
| Type | Numeric data (structured) |
| Format | CVS, JSON |
| Size | Not possible to estimate at the moment |
| Language | English |
| Source | AUEB |
| Version | Not available yet |
| Date | Not available yet |
| Accessibility | Open access + restricted (for certain dataset series) |
| Storage | Not available yet |
| Restrictions | To be defined |
| Standards & metadata | To be defined |
| Licensing | To be defined |
| Keywords | To be defined |

### 2.2.2 Analytics service datasets

Table 3: Analytics service datasets

| | |
|---|---|
| **Name of dataset** | Analytics service datasets |
| **Responsible partner** | UOM |
| **ID** | Not available yet |
| **Description** | A brief description of the dataset |
| **Purpose** | The analytics related data will be generated and utilised in line with the goal of T3.3 needs for analytics to support the decomposition of any service request submitted to the MaaS platform into subservices in order to offer recommendations for the potential resource providers. |
| **Type** | Numeric data (structured) |
| **Format** | CVS |
| **Size** | Not possible to estimate at the moment |
| **Language** | English |
| **Source** | AUEB |
| **Version** | Not available yet |
| **Date** | Not available yet |
| **Accessibility** | Open access + restricted (for certain dataset series) |
| **Storage** | Not available yet |
| **Restrictions** | To be defined |
| **Standards & metadata** | To be defined |
| **Licensing** | To be defined |
| **Keywords** | To be defined |

### 2.2.3 Value Network 1 Dataset(s)

Table 4: Value Network 1 Dataset(s)

| | |
|---|---|
| Name of dataset | Name of the dataset |
| Responsible partner | KAREL |
| ID | Not available yet |
| Description | A brief description of the dataset |
| Purpose | The data will be used for the validation of the Tec4MaSEs services |
| Type | Numeric |
| Format | XML |
| Size | Not sure yet (approx. 9 MB) |
| Language | - |
| Source | KAREL |
| Version | Not available yet |
| Date | Not available yet |
| Accessibility | Restricted |
| Storage | To be defined |
| Restrictions | To be defined |
| Standards & metadata | To be defined |
| Licensing | To be defined |
| Keywords | To be defined |

### 2.2.4 Value Network 2 Dataset(s)

Table 5: Value Network 2 Dataset(s)

| | |
|---|---|
| **Name of dataset** | CAD files of mould structural parts and cavity inserts |
| **Responsible partner** | URA, ERREKA |
| **ID** | Not available yet |
| **Description** | CAD files of the parts involved in manufacturing services requests (for additive manufacturing and machining) |
| **Purpose** | Provide requirements related to the manufacturing service request |
| **Type** | Structured |
| **Format** | STEP |
| **Size** | Not possible to estimate at the moment |
| **Language** | N/A |
| **Source** | CAD files are generated using CAD/CAM software |
| **Version** | Not available yet |
| **Date** | Not available yet |
| **Accessibility** | Restricted for real parts – Open access for test files including parts designed for testing purposes. |
| **Storage** | Real data are stored internally. To be defined for test data. |
| **Restrictions** | CAD files of real parts are subject to IPR and NDA. Test files can be provided. |
| **Standards & metadata** | ISO 10303: Standard for the Exchange of Product Model Data - STEP |
| **Licensing** | To be defined for test files |
| **Keywords** | CAD |

| Name of dataset | 2D drawings of mould structural parts and cavity inserts |
|---|---|
| Responsible partner | URA, ERREKA, TEKNIKER |
| ID | Not available yet |
| Description | 2D drawings of the parts involved in manufacturing services requests (for additive manufacturing and machining) |
| Purpose | Provide requirements related to the manufacturing service request |
| Type | Unstructured |
| Format | PDF |
| Size | Not possible to estimate at the moment |
| Language | N/A |
| Source | 2D drawings are generated using CAD/CAM software |
| Version | Not available yet |
| Date | Not available yet |
| Accessibility | Restricted for real parts – Open access for test files including parts designed for testing purposes. |
| Storage | Real data are stored internally. To be defined for test data. |
| Restrictions | 2D drawings of real parts are subject to IPR and NDA |
| Standards & metadata | - |
| Licensing | To be defined for test files |
| Keywords | 2D Drawing |

| | |
|---|---|
| **Name of dataset** | Plastic injection process parameters |
| **Responsible partner** | ERREKA |
| **ID** | Not available yet |
| **Description** | Process parameters for manufacturing a part using plastic injection moulding |
| **Purpose** | Provide requirements related to the manufacturing service request |
| **Type** | Structured |
| **Format** | PDF |
| **Size** | Not possible to estimate at the moment |
| **Language** | English |
| **Source** | Process parameters datasheets |
| **Version** | Not available yet |
| **Date** | Not available yet |
| **Accessibility** | Restricted for real parts – Open access for test files including simulated process parameters |
| **Storage** | Real data are stored internally. To be defined for test data. |
| **Restrictions** | To be defined |
| **Standards & metadata** | - |
| **Licensing** | To be defined for test files |
| **Keywords** | Plastic injection process parameter |

| | |
|---|---|
| **Name of dataset** | Manufacturing resources technical specification |
| **Responsible partner** | URA, ERREKA, TEKNIKER |
| **ID** | Not available yet |
| **Description** | Technical specifications of the manufacturing resources (for additive manufacturing, machining, and plastic injection moulding) involved in the pilot. |
| **Purpose** | Provide information of the capabilities of the available manufacturing resources to be used to provide additive manufacturing and machining/plastic injection moulding services |
| **Type** | Structured |
| **Format** | To be defined based on the projects needs |
| **Size** | Not possible to estimate at the moment |
| **Language** | English |
| **Source** | Technical datasheets |
| **Version** | Not available yet |
| **Date** | Not available yet |
| **Accessibility** | Open access to technical specification of the manufacturing resources involved in the pilot |
| **Storage** | To be defined |
| **Restrictions** | To be defined |
| **Standards & metadata** | - |
| **Licensing** | To be defined |
| **Keywords** | capability |

| | |
|---|---|
| **Name of dataset** | Production planning |
| **Responsible partner** | URA, ERREKA, TEKNIKER |
| **ID** | Not available yet |
| **Description** | Production planning |
| **Purpose** | Provide information of the availability of the manufacturing resources to be used to provide additive manufacturing, machining, and plastic injection moulding services |
| **Type** | Structured |
| **Format** | .xlsx |
| **Size** | Not possible to estimate at the moment |
| **Language** | English |
| **Source** | Internal planning datasheet |
| **Version** | Not available yet |
| **Date** | Not available yet |
| **Accessibility** | Restricted for real production planning information – Open access for test files including simulated production planning data. |
| **Storage** | Real data are stored internally. To be defined for test data. |
| **Restrictions** | To be defined |
| **Standards & metadata** | To be defined |
| **Licensing** | To be defined |
| **Keywords** | Capacity, follow up |

| | |
|---|---|
| **Name of dataset** | Manufacturing execution information |
| **Responsible partner** | URA, ERREKA, TEKNIKER |
| **ID** | Not available yet |
| **Description** | Manufacturing execution information |
| **Purpose** | Provide follow up information as well as manufacturing resources capacity usage. |
| **Type** | Structured |
| **Format** | .xlsx |
| **Size** | Not possible to estimate at the moment |
| **Language** | English |
| **Source** | Internal access database |
| **Version** | Not available yet |
| **Date** | Not available yet |
| **Accessibility** | Restricted for real manufacturing information – Open access for test files including simulated manufacturing execution data. |
| **Storage** | Real data are stored internally. To be defined for test data. |
| **Restrictions** | To be defined |
| **Standards & metadata** | To be defined |
| **Licensing** | To be defined |
| **Keywords** | Capacity, follow up |

### 2.2.5 Value Network 3 Dataset(s)

Table 6: Value Network 3 Dataset(s)

| | |
|---|---|
| **Name of dataset** | Name of the dataset |
| **Responsible partner** | AIBEL |
| **ID** | Not available yet |
| **Description** | A brief description of the dataset |
| **Purpose** | Intended use of the data |
| **Type** | Text, Image, Video, numeric data, geospatial data |
| **Format** | .txt, .pdf, .CVS, JSON, .XML, .rdf, |
| **Size** | Not possible to estimate at the moment |
| **Language** | Norwegian |
| **Source** | AIBEL |
| **Version** | Not available yet |
| **Date** | Not available yet |
| **Accessibility** | Restricted |
| **Storage** | Not available yet |
| **Restrictions** | To be defined |
| **Standards & metadata** | To be defined |
| **Licensing** | To be defined |
| **Keywords** | To be defined |

## 2.3 Public deliverables

The following table presents the list of public deliverables of the Tec4MaaSEs project.

Table 7: Tec4MaaSEs public deliverables

| ID | Deliverable Title | Partner | Expected date |
|---|---|---|---|
| D2.1 | Reference cases and actionable models for reconfigurable value networks and service decomposition v1 | UOM | *August 2024* |
| D2.3 | Governance framework v1 | MAG | *December 2024* |
| D2.5 | Tec4MaaSEs architecture and specifications v1 | ATC | *December 2024* |
| D3.1 | Digital Twins models and services for factories and value networks v1 | IOSB | *June 2025* |
| D3.3 | Analytics for resource-subservice matching and service composition v1 | UOM | *June 2025* |
| D3.5 | Context-driven optimized (re-)configuration services v1 | AUEB | *June 2025* |
| D3.7 | Tec4MaaSEs governance services v1 | MAG | *June 2025* |
| D4.1 | Tec4MaaSEs integrated platform v1 | ATC | *June 2025* |
| D5.1 | Pilot validation plan and assessment report | TEKNIKER | *June 2025* |
| D2.2 | Reference cases and actionable models for reconfigurable value networks and service decomposition v2 | UOM | *December 2025* |
| D2.4 | Governance framework v2 | MAG | *December 2025* |
| D2.6 | Tec4MaaSEs architecture and specifications v2 | ATC | *December 2025* |
| D3.2 | Digital Twins models and services for factories and value networks v2 | IOSB | *February 2026* |
| D3.4 | Analytics for resource-subservice matching and service composition v2 | UOM | *February 2026* |
| D3.6 | Context-driven optimized (re-)configuration services v2 | AUEB | *February 2026* |
| D3.8 | Tec4MaaSEs governance services v2 | MAG | *February 2026* |
| D4.2 | Tec4MaaSEs integrated platform v2 | ATC | *June 2026* |
| D4.3 | Tec4MaaSEs balanced scorecard toolkit | MAG | *June 2026* |
| D5.4 | Impact assessment and lessons learnt | ARC | *December 2026* |

## 2.4 Scientific publications

In this section, we are going to include all scientific publications produced during the project duration.

## 2.5 Software components

While Tec4MaaSEs project is committed to provide most of the software components as Open Source, the final decision will be made along with the IPR agreement in WP6. The current analysis is performed taking into account the list of exploitable components, defined in the DOA as open-source components.

# 3 FAIR Data

Within the Tec4MaaSEs project, the management and utilisation of data is guided by the FAIR data principles [2]. This approach enables the consortium to promote open science, enhance data sharing and re-use, and foster stakeholder collaboration. Integrating the FAIR principles into the data management approach guarantees that data generated, collected and processed by the Tec4MaaSEs project can be efficiently located, accessed, integrated, and repurposed, thus optimising the project's impact and efficiency.

## 3.1 Making Data Findable

The first principle of the FAIR data principles, Findability, plays a pivotal role in ensuring that data can be effortlessly and efficiently located by a diverse range of users. In this context, 'users' encompass researchers, data analysts, machines, and algorithms, thus enabling automated data discovery and analysis. This principle can be delineated into four sub-principles. Table 8 below explains how these sub-principles will be applied in during the Tec4MaaSEs project.

Table 8: Sub-principles of Findability and application in Tec4MaaSEs

| Sub-principle | Explanation | Application in Tec4MaaSEs |
|---|---|---|
| **F1 - Unique and Persistent Identifier** | Assigning a globally unique and persistent identifier to each dataset and its metadata, ensuring precise retrieval and preventing duplication. | Utilising systems like DOIs or UUIDs to assign distinct identifiers to datasets and metadata, facilitating accurate location and reference. |
| **F2 - Rich Metadata** | Including detailed metadata covering data origin, characteristics, and conditions enriches the data's contextual understanding. | Providing comprehensive descriptions including authorship, creation dates, locations, collection, and processing methodologies. |
| **F3 - Metadata-Data Linkage** | Ensuring metadata explicitly includes the data's identifier, creating a transparent and robust link between the dataset and its description. | Embedding the unique identifiers within the metadata to establish a transparent connection, simplifying association and reference. |
| **F4 - Data Registration and Indexing** | Registering or indexing data and metadata in searchable resources to enhance discoverability. | Incorporating datasets and metadata into searchable repositories or catalogues, ensuring ease of access for diverse users. |

### 3.1.1 Data discoverability

To bolster the discoverability of data within the Tec4MaaSEs project, a whole range of measures tailored to the classification of each dataset will be implemented. These measures aim to ensure that data is readily accessible to authorised users.

In open-access datasets, the focus will be on enhancing discoverability by incorporating search keywords and assigning unique DOI identifiers. This approach ensures that project members and relevant stakeholders can easily locate and reference these datasets.

In cases where datasets are classified as confidential, meticulous attention will be given to enriching metadata manually. This metadata will provide comprehensive insights into the data's nature, origin, and relevance.

However, access to such rich metadata will be restricted to approved users to maintain the confidentiality and security of sensitive data.

Tec4MaaSEs underscore its commitment to effective data management in alignment with the FAIR data principles by tailoring our approach to enhance data discoverability while safeguarding confidentiality. These measures ensure that data remains accessible to authorised individuals and entities while upholding stringent security and privacy standards.

### 3.1.2 Data identification mechanisms

To ensure robust data identification within the Tec4MaaSEs project, we will employ effective mechanisms that allow for precisely tracking and referencing datasets. These mechanisms have been carefully chosen to align with best practices and facilitate seamless data management.

Assigning Digital Object Identifiers (DOIs) to all Open Access datasets is central to this approach. DOIs are a cornerstone in the data identification strategy, providing each dataset with a unique, permanent digital footprint. This facilitates easy tracking and retrieval and enhances data citation practices.

Further augmenting our data identification process, we implement rigorous versioning and user registration mechanisms. These measures ensure precise version control and trackability of dataset alterations, thereby maintaining data integrity over time. Integrating user registration processes aids in monitoring data access and usage, fortifying our commitment to data security and GDPR compliance.

Table 9 below outlines the data identification mechanisms in Tec4MaaSEs, tailored to the specific nature of various datasets, ensuring robust and secure data management aligned with the project's objectives.

**Table 9: Data identification mechanisms in Tec4MaaSEs**

| Mechanisms | Nature of data | Example in Tec4MaaSEs context |
|---|---|---|
| **DOIs (Digital Object Identifiers)** | Open Access Data | Dissemination material, such as a publication or a publicly available project deliverable, would be assigned a DOI to facilitate academic referencing and public accessibility. |
| **Versioning System** | All Types of Data (Open Access, Confidential, etc.) | In the case of pilot datasets updated during different project phases, a versioning system would track changes, ensuring clarity on dataset evolution over time. |
| **User Registration Processes** | Confidential or Restricted Access Data | For confidential datasets, like internal performance metrics or sensitive pilot project data, a pilot registration system would be implemented to monitor and restrict access, aligning with data privacy and security protocols. |

### 3.1.3 Data naming convention

Adopting a systematic and coherent naming convention for data files is essential for ensuring efficient organisation and retrieval. The naming convention to be applied for all datasets shall include at least the following information:

- The acronym of the project.

- The name of the dataset;
- The version number;
- The date of creation
- The format

### 3.1.4 Data versioning control

Rigorous version control of datasets is a critical aspect to ensure that most current and accurate data is made available for reference. A standardised version control process will be implemented across all datasets. This process involves:

- **Version Number**: Each dataset will be marked with a unique version number. For example, the initial version of a dataset on optimisation might be designated as "Tec4MaaSEs_Optimisation_v1.0". For each new version of the dataset an incremental numbering will be applied.
- **Modification Date**: Alongside the version number, the date of the latest modification will be recorded. For instance, "Tec4MaaSEs_Optimsiation_v1.0_25-03-2024" would indicate that this version was updated on 25th March 2024.
- **Change Log Maintenance**: A change log will be maintained for each document, detailing the nature of updates and revisions made in each version.

### 3.1.5 Metadata creation

In the Tec4MaaSEs project, the metadata creation will be diligently aligned with recognised international standards tailored to the specific nature of each dataset. For instance:

- **ISO/IEC 11179 standard**: This standard, known for its comprehensive approach to metadata registry, will be employed for specific datasets where detailed, structured metadata is crucial. For example, an "Infrastructure Security Analysis" dataset could follow this standard to ensure metadata is robust and systematically organised.
- **Asset Administration Shell (AAS) standard**: The AAS standard will be utilised for datasets requiring a more manually intensive approach to metadata creation. This is particularly relevant for datasets with complex or concrete data structures. An example within Tec4MaaSEs could be the "Network Protocols Efficiency" dataset, where each data element would be meticulously documented per AAS guidelines.

This strategic approach to metadata creation ensures that all datasets within the Tec4MaaSEs project are accompanied by high-quality, standard-compliant metadata, facilitating effective data management and usability.

## 3.2 Making Data Openly Accessible

The principle of Accessibility under the FAIR data principles ensures that once data is located, it can be seamlessly accessed by human users and/or machines. The following strategies will be implemented within Tec4MaaSEs project to adhere to this principle:

- ***Standardized Communication Protocol:*** Utilizing a universally recognised protocol such as HTTP/HTTPS, ensures that data, retrievable by its unique identifier, is accessible in an automated manner and does not require specialised tools.
- ***Open, Free, and Universally Implementable Protocol:*** This approach promotes broad usage and accessibility, ensuring the protocol is available to all without restrictions or fees.

- ***Authentication and Authorization:*** The adopted protocols will support robust authentication and authorisation processes where necessary, to protect sensitive data.
- ***Metadata Accessibility:*** Consistent with the FAIR principles, even if the actual data is no longer available, the metadata will remain accessible to provide valuable context and support future research initiatives.

Table 10 below presents how these sub-principles will be applied in during the Tec4MaaSEs project.

**Table 10: Sub-principles of Accessibility and application in Tec4MaaSEs**

| Sub-principle | Explanation | Application in Tec4MaaSEs |
|---|---|---|
| **A1 - Retrieval by Identifier** | Data should be retrievable through a standardised protocol using its unique identifier. | Datasets will be accessible via HTTP/HTTPS, allowing easy retrieval by their unique DOIs. |
| **A1.1 - Open and Free Protocol** | The protocol for data access should be open, accessible, and universally implementable. | Tec4MaaSEs will ensure data accessibility through widely accepted protocols like HTTP/HTTPS without access restrictions. |
| **A1.2 - Authentication and Authorization** | Protocols should support identity verification and access level determination. | Tec4MaaSEs will implement authentication and authorisation procedures for confidential datasets to control access. |
| **A2 - Metadata Accessibility** | Metadata should remain accessible even if the data is no longer available. | Tec4MaaSEs will maintain metadata accessibility for all datasets, irrespective of availability. |

**Data Repositories**

Tec4MaaSEs consortium will ensure open access to all research datasets, including public deliverables and scientific publications. The repository that will be used for this purpose is Zenodo [3], an online platform that allows easily storage in various sizes and formats, and provides flexible licensing, access and re-use of research data.

Research data needed for validation of results presented in scientific publications will be uploaded in Zenodo as soon as possible, following the acceptance by the relevant committee. In case an embargo period should be applied, data will be deposited in an online repository (e.g. project website).

Confidential data that cannot become public due to privacy, security and/or confidentiality restrictions will be only made accessible to restricted consortium partners in the project repository (SharePoint), following a data sharing agreement. Moreover, when possible, anonymization, aggregation, minimization or sampling techniques may be applied to ‘real’ data used during the project to guarantee the preservation of personal or sensitive information.

Finally, for the open-source software component, the Tec4MaaSEs consortium has selected GitLab [4], an online repository which supports distributed source code development and management. The Tec4MaaSEs produced datasets could also be built on the GitLab platform to raise awareness, increase impact, and ensure long-term sustainability of the project’s results. GitLab will also host parts of some of the Tec4MaaSEs open-source code components that will be implemented during the project.

## 3.3 Making Data Interoperable

In the Tec4MaaSEs project, data interoperability is fundamental to ensure effective integration and analysis with other datasets. This aligns with the FAIR data principles, focusing on three key interoperability aspects as presented in below:

Table 11: Sub-principles of Interoperability and application in Tec4MaaSEs

| Sub-principle | Explanation | Application in Tec4MaaSEs |
|---|---|---|
| **I1. Standardised Language for Knowledge Representation** | Data and metadata should be represented using a standardised language that is widely understood and accepted. | Tec4MaaSEs will use common data formats like JSON or XML, ensuring compatibility with industry standards. |
| **I2. FAIR-Compliant Vocabularies** | Utilising standard vocabularies that adhere to FAIR principles to enhance data integration. | Employing standardised terminologies across datasets to facilitate seamless integration and analysis. |
| **I3. Qualified References to Other Data** | Data should include references to related datasets or metadata, fitting into a broader context. | Tec4MaaSEs datasets will reference related datasets, like cross-referencing a optimisation analysis dataset with a simulation dataset. |

## 3.4 Making Data Re-Usable

The principle of Reusability can be broken down into several key components to ensure data and metadata are not just utilised once but have potential for further applications. Table 12 below presents how Reusability's will be applied in during the Tec4MaaSEs project.

Table 12: Sub-principles of Reusability and application in Tec4MaaSEs

| Sub-principle | Explanation | Application in Tec4MaaSEs |
|---|---|---|
| **R1. Rich Descriptions** | Data and metadata should be comprehensively described to enhance understanding and applicability. | Datasets will include extensive context, quality, and relevance details. |
| **R2. Clear Usage License** | Data should be released with a clear license that informs users about their rights and obligations. | Data will have explicit licensing, such as Creative Commons, to clarify usage rights. |
| **R3. Detailed Provenance** | Providing detailed provenance information to add context to the data's collection and processing | Data will include comprehensive provenance information outlining how the data was collected and processed. |
| **R4. Community Standards** | Ensuring data meets domain-relevant community standards for improved reusability. | Tec4MaaSEs will adhere to internationally recognised cybersecurity data formatting and structuring standards. |

### 3.4.1 Data quality assurance process

In the Tec4MaaSEs project, data quality and integrity are critical for achieving its objectives. A comprehensive data quality assurance process has been established to ensure the data is accurate, reliable, and high-quality. This process encompasses a range of validation checks, error detection routines, and manual data review and verification procedures. The approach to maintaining data quality and integrity includes the following critical processes in Table 13, highlighting the steps taken to maintain the integrity and usefulness of the project's datasets:

Table 13: Tec4MaaSEs process for data integrity and usefulness maintenance

| Process | Explanation | Application in Tec4MaaSEs |
|---|---|---|
| **Validation Checks and Error Detection** | Systematic validation and error detection routines to identify and correct data inaccuracies. | Automated checks on datasets to ensure correct formatting and data consistency. |
| **Manual Data Review and Verification** | Manual processes to ensure data accuracy and reliability. | Expert review of datasets to validate optimisation and simulation reports. |
| **Regular Audits and Peer Reviews** | Conducting periodic audits and peer reviews to maintain high data quality standards. | Regular audits data related to ensure accuracy and completeness. |

### 3.4.2 Duration for which data will remain re-usable

This section focuses on the duration for which data will remain reusable, ensuring long-term accessibility and usability for ongoing and future research:

- **Minimum Preservation Period**: A defined minimum period of **five years** after the project end is set by the Tec4MaaSEs consortium during which data will be preserved, accessible, and usable, tailored to meet the needs of the research community.
- ***Reliable Data Repositories***: Data will be stored in trusted repositories (see Section 3.2) known for their long-term preservation capabilities and commitment to maintaining data in a usable format.
- **Regular Data Updates**: Acknowledging the dynamic nature of smart manufacturing, regular updates will be implemented to maintain the data's relevance and accuracy.

# 4 Allocation of Resources

## 4.1 Costs for long-term preservation

The project's funding will cover the essential costs of adhering to FAIR data principles, including data publication, updating, maintenance, and security. Costs related to scientific publications in Open Access journals shall be compliant with the rules specified in Article 6 and Annex 5 of the Grant Agreement.

The project anticipates a data preservation period of **five years** after the project end, utilising cost-effective platforms such as Zenodo for open repositories and SharePoint platform for secure storage. Each consortium partner will cover the expenses related to the data they generate, including storage and anonymisation. The long-term value of this data is particularly noteworthy, given its relevance and potential applications in future projects.

## 4.2 Roles and Responsibilities

Individual responsibilities on data management in the consortium are:

- **Project Coordinator (MAG)** – oversees the overall data management strategy and ensures compliance with the project's objectives. In addition, it is responsible for the data storage in SharePoint.
- **Scientific and Technical Manager (AUEB)** – oversees the technical datasets, focusing on their curation, archiving, and accessibility, especially in open repositories (e.g. Zenodo, GitLab).
- **Dissemination and Communication Manager (UOM)** – handles scientific publications and deliverables suitable for publication in the project website and Zenodo, ensuring adherence to Open Access policies.
- **Data Manager (MAG)** – creates and updates the DMP, coordinates and supports data providers, ensuring compliance with the DMP processes, and maintains the list of datasets up to date.
- **Work Package Leaders** – responsible for the quality control of the data generated within their respective Work Package.
- **All other partners** – responsible for the identification of their own dataset(s) suitable for publication. Moreover, each partner has to follow the policies set out in this DMP. Validation and registration of datasets and metadata is the responsibility of the partner that generates the data. Backing up data for sharing through Open Access repositories is the responsibility of the partner possessing the data. If a dataset is updated, the partner that possesses the data shall manage the different versions and ensure the latest version is available.

# 5 Data Security

## 5.1 Data Storage and Preservation

The Data Storage and Preservation infrastructure consists of several web-based platforms that together provide long-term open access to all generated or collected data of the project. The following list presents the platforms to be used during the project and describes their concepts for publishing, storage, and backup.

### 5.1.1 Project website

UOM has designed and setup a project webpage, which provides a general overview of the project objectives and its approach and will inform target audience on its development status. A dedicated section for resources will be integrated in order to publish deliverables as well as scientific publication. All documents will be published using the portable document format (PDF) and be enriched by using simple metadata information, such as the title and the type of the document. All information on the project website can be accessed without the need of creating an account. The webpage will be available during the project lifetime, and remain available for at least two years after the project end.

URL: http://www.tec4maases.eu/

### 5.1.2 Zenodo

Zenodo is an open repository which helps researchers to share research results in a wide variety of formats for all scientific disciplines. It was developed by the OpenAIRE+ project and is maintained by CERN using one of Europe's most reliably hardware infrastructures.

Zenodo not only supports the publication of scientific papers or white papers, but also the publication of any structured research data (e.g. using XML). Zenodo provides a connector to GitHub that supports open collaboration for source code and versioning for all kinds of data. All uploaded results are structured by using metadata, like for example the contributors' names, keywords, date, location, kind of document, license, and others. Considering the language of textual metadata items, English is preferred. All metadata is licensed under CC license (Creative Commons 'No Rights Reserved'). The property rights or ownership of a result does not change by uploading it to Zenodo.

All public results generated or collected during the Tec4MaaSEs project will be uploaded to Zenodo for long-term storage and open access.

URL: http://zenodo.org/communities/tec4maases

### 5.1.3 GitLab

GitLab is a well-established online repository which supports distributed source code development, management, and revision control. It is primarily used for source code data. It enables world-wide collaboration between developers and provides also some facilities to work on documentation and to track issues. GitLab provides paid and free service plans. Free service plan can have any number of public, open-access repositories with unlimited collaborators. Private, non-public repositories require a paid service plan. Many open-source projects use Gitlab to share their results for free. The platform uses metadata like contributors' nicknames, keywords, time, and data file types to structure the projects and their results. The service is hosted by ATC in Greece.

All source-code components that are implemented during the project and decided to be public will be uploaded to an open access GitLab repository.

## 5.2 Data Security Measures

The necessary data anonymity will be ensured. State-of-the-art encryption and pseudonymisation/ anonymisation techniques will be applied across all stages of data lifecycle following the principles set out in the GDPR. Strict access controls, two-factor authentication, and detailed user activity logs will protect sensitive resources/information.

## 5.3 Data Backup and Recovery

A comprehensive backup plan will be established, including scheduled full and incremental backups to capture changes regularly. Version control mechanisms will maintain a dependable historical record of datasets, allowing for easy restoration if required. In case of any disruptions, disaster recovery procedures will prioritise swift restoration of access to essential data.

## 5.4 Data Sharing

To ensure that the data exchanged between project partners is kept secure and free from any unauthorised access or tampering, the consortium will use the most reliable and advanced secure protocols, such as HTTPS and SFTP. Furthermore, to provide an extra layer of security, all sensitive data will be encrypted before being transmitted so that even if it falls into the wrong hands, it will be completely unreadable.

# 6 Ethical Aspects

Tec4MaaSEs is committed to ensure compliance with ethical principles and fundamental rights embedded in the regulatory framework of the European Union, including the Charter of Fundamental Rights of the European Union as well as the European Convention on Human Rights. The data processing activities within the project will be carried out in accordance with GDPR and ePrivacy Directive 2002/58/EC.

Moreover, the corresponding national data protection legislation should be taken into consideration, and all legal documents and certifications required for compliance with such legislation will be obtained. The Parties who provide or transfer to any other Party information containing Personal Data must have: (i) the authority and/or the authorisation to disclose the aforementioned information; (ii) obtained appropriate informed consents from all the data subjects involved, or from any applicable institution, and (iii) a confirmation that there is no restriction that would prevent any other Party from using the shared information. Data protection by design and by default will be at the core of the research and development work as well as the project outputs.

The personal data will be anonymized and homomorphic encryption will be used for pseudo-anonymization. Identifiable data will be dissociated from the rest of the data in a separate database. Personal data will be processed in Tec4MaaSEs using a neutral code with the aim to render data non-attributable to any natural person.

Due to the way the research will be carried out, consortium partners making decisions about the collection and processing of personal data, in order to achieve the various goals and objectives of the project, will be deemed joint data controllers. As required by the GDPR, an agreement outlining the allocation of the obligations and responsibilities of the joint data controllers within Tec4MaaSEs, will be created. A key aspect of this arrangement is regular and effective communication between data controllers to ensure a consistent and effective approach.

Tec4MaaSEs implementation does not anticipate or plan any transfer of personal data to third parties. Relevant legal foundations, appropriate safeguards and compliance measures will be identified and implemented if data sharing is deemed necessary at a later stage. These safeguards will be documented in the final version of the Data Management Plan.

## Conclusions

This first version of the Data Management Plan (DMP) provides a comprehensive strategy for responsible data management throughout the project's duration. Guided by a commitment to the FAIR data principles, the DMP prioritises data protection, robust security measures, and long-term value of data.

While the specific datasets to be used in the later stages of the project are still under development, the current plan proactively addresses the potential inclusion of personal data and the use of AI technologies. This demonstrates Tec4MaaSEs commitment to rigorous data management planning from the earliest phases.

In conclusion, this deliverable represents a thorough and proactive blueprint for managing complex datasets. It underscores the project's commitment to ethical data handling, secure research practices, adapting strategies as the project evolves, and contributing positively to advancing cybersecurity innovation.

# References

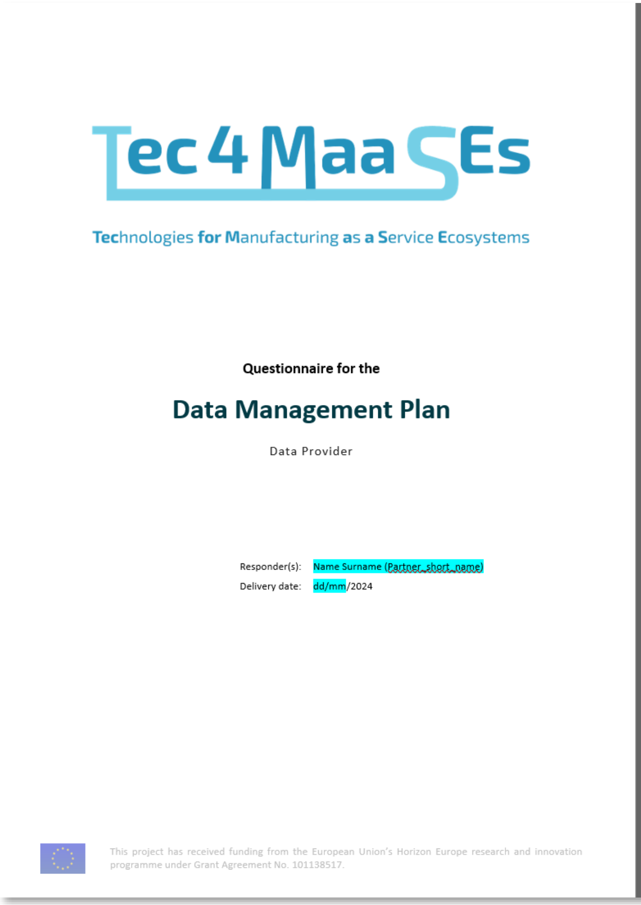
[1] Guidelines on Data Management in H2020, Version 3.0, 26 July 2016. Available at: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] FAIR Data Principles: https://www.force11.org/group/fairgroup/fairprinciples

[3] Zenodo: https://zenodo.org

[4] GitLab: https://about.gitlab.com/

# Annex 1: DMP Questionnaires



Technologies for Manufacturing as a Service Ecosystems

Questionnaire for the

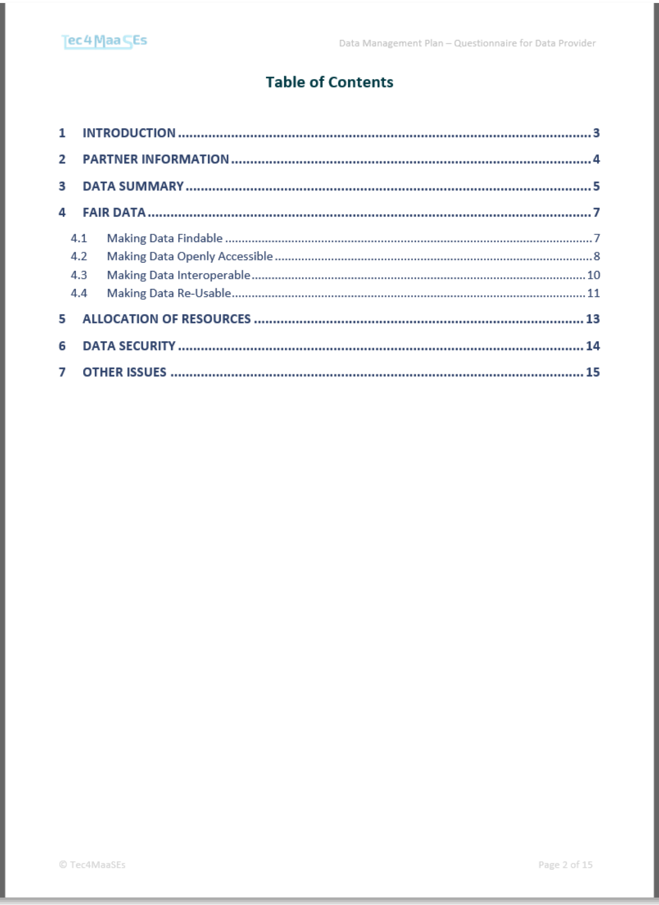
Data Management Plan

Data Provider

Responder(s): Name Surname (Partner_short_name)

Delivery date: dd/mm/2024

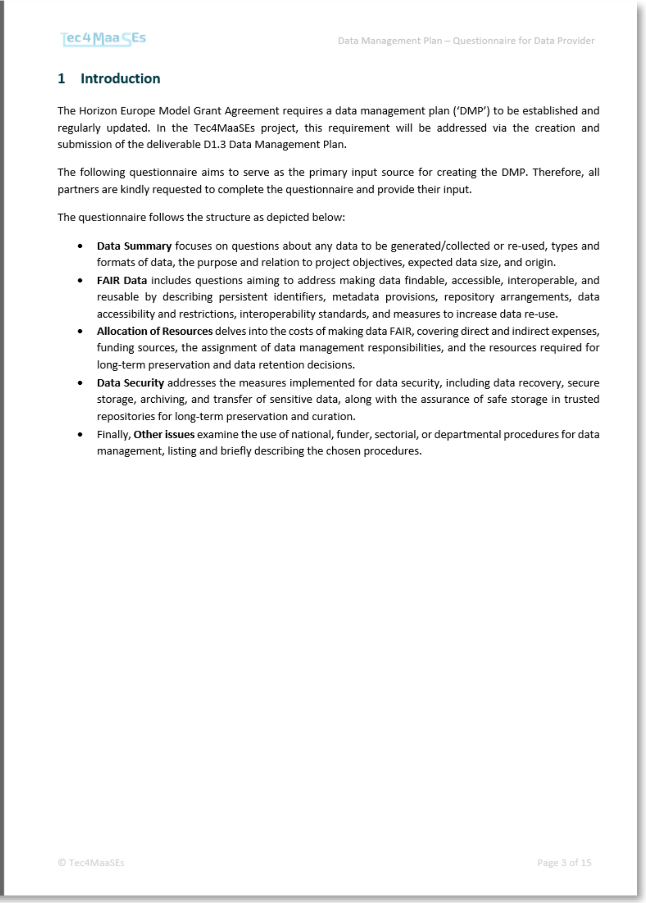
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101138517.



Data Management Plan – Questionnaire for Data Provider

Table of Contents

© Tec4MaaSEs Page 2 of 15



Data Management Plan – Questionnaire for Data Provider

1 Introduction

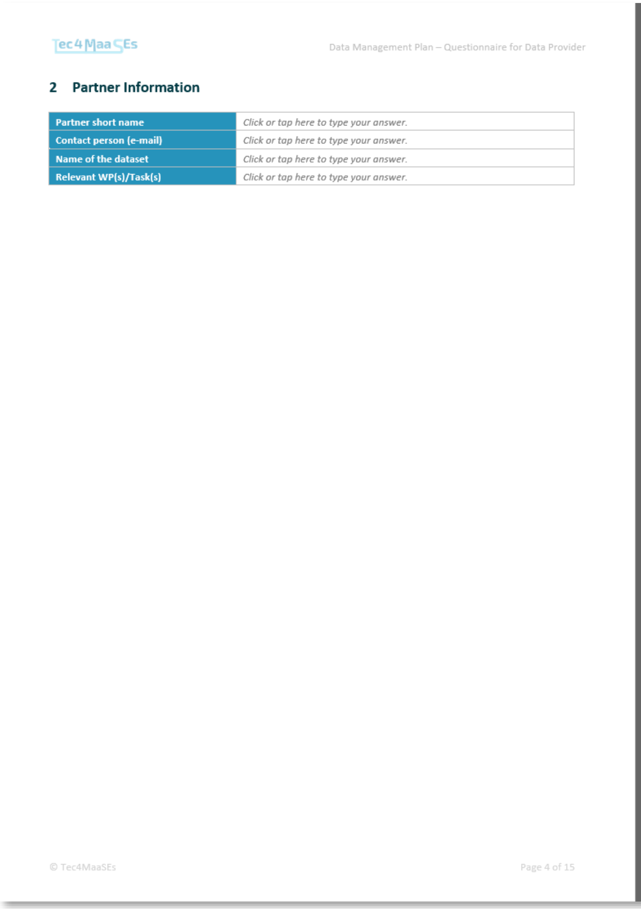
The Horizon Europe Model Grant Agreement requires a data management plan ('DMP') to be established and regularly updated. In the Tec4MaaSEs project, this requirement will be addressed via the creation and submission of the deliverable D1.3 Data Management Plan.

The following questionnaire aims to serve as the primary input source for creating the DMP. Therefore, all partners are kindly requested to complete the questionnaire and provide their input.

The questionnaire follows the structure as depicted below:

- **Data Summary** focuses on questions about any data to be generated/collected or re-used, types and formats of data, the purpose and relation to project objectives, expected data size, and origin.
- **FAIR Data** includes questions aiming to address making data findable, accessible, interoperable, and reusable by describing persistent identifiers, metadata provisions, repository arrangements, data accessibility and restrictions, interoperability standards, and measures to increase data re-use.
- **Allocation of Resources** delves into the costs of making data FAIR, covering direct and indirect expenses, funding sources, the assignment of data management responsibilities, and the resources required for long-term preservation and data retention decisions.
- **Data Security** addresses the measures implemented for data security, including data recovery, secure storage, archiving, and transfer of sensitive data, along with the assurance of safe storage in trusted repositories for long-term preservation and curation.
- Finally, **Other issues** examine the use of national, funder, sectorial, or departmental procedures for data management, listing and briefly describing the chosen procedures.

© Tec4MaaSEs Page 3 of 15

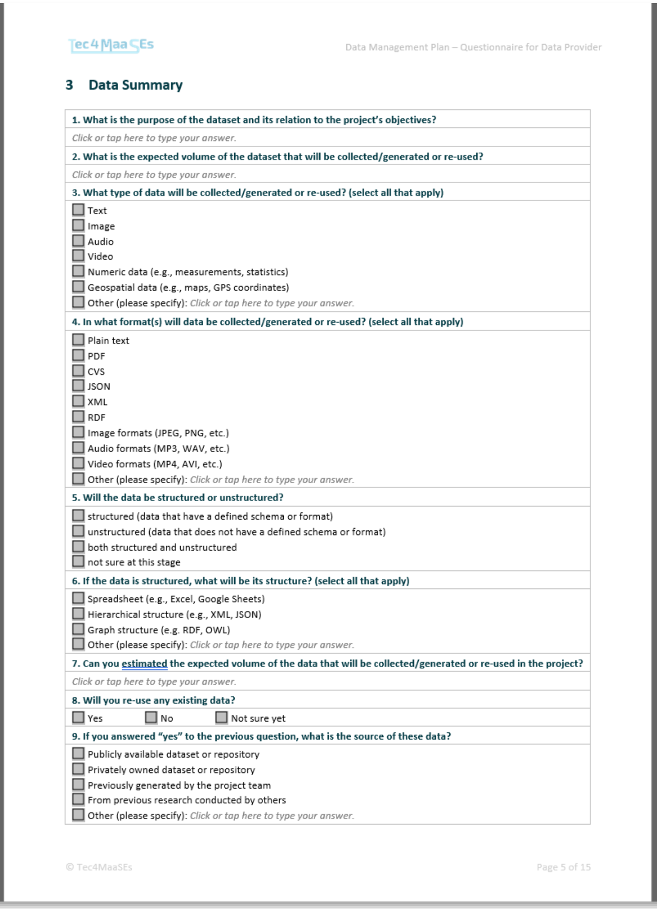


Data Management Plan – Questionnaire for Data Provider

## 2 Partner Information

| | |
|---|---|
| Partner short name | *Click or tap here to type your answer.* |
| Contact person (e-mail) | *Click or tap here to type your answer.* |
| Name of the dataset | *Click or tap here to type your answer.* |
| Relevant WP(s)/Task(s) | *Click or tap here to type your answer.* |

© Tec4MaaSEs Page 4 of 15



Data Management Plan – Questionnaire for Data Provider

## 3 Data Summary

**1. What is the purpose of the dataset and its relation to the project's objectives?**

*Click or tap here to type your answer.*

**2. What is the expected volume of the dataset that will be collected/generated or re-used?**

*Click or tap here to type your answer.*

**3. What type of data will be collected/generated or re-used? (select all that apply)**

☐ Text
☐ Image
☐ Audio
☐ Video
☐ Numeric data (e.g., measurements, statistics)
☐ Geospatial data (e.g., maps, GPS coordinates)
☐ Other (please specify): *Click or tap here to type your answer.*

**4. In what format(s) will data be collected/generated or re-used? (select all that apply)**

☐ Plain text
☐ PDF
☐ CVS
☐ JSON
☐ XML
☐ RDF
☐ Image formats (JPEG, PNG, etc.)
☐ Audio formats (MP3, WAV, etc.)
☐ Video formats (MP4, AVI, etc.)
☐ Other (please specify): *Click or tap here to type your answer.*

**5. Will the data be structured or unstructured?**

☐ structured (data that have a defined schema or format)
☐ unstructured (data that does not have a defined schema or format)
☐ both structured and unstructured
☐ not sure at this stage

**6. If the data is structured, what will be its structure? (select all that apply)**

☐ Spreadsheet (e.g., Excel, Google Sheets)
☐ Hierarchical structure (e.g., XML, JSON)
☐ Graph structure (e.g. RDF, OWL)
☐ Other (please specify): *Click or tap here to type your answer.*

**7. Can you estimated the expected volume of the data that will be collected/generated or re-used in the project?**

*Click or tap here to type your answer.*

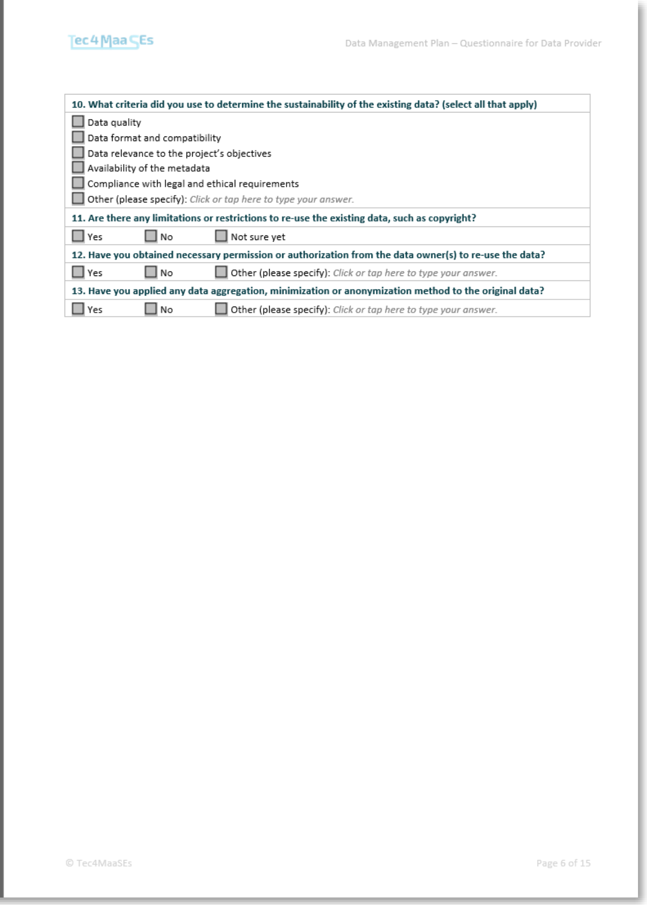
**8. Will you re-use any existing data?**

☐ Yes ☐ No ☐ Not sure yet

**9. If you answered "yes" to the previous question, what is the source of these data?**

☐ Publicly available dataset or repository
☐ Privately owned dataset or repository
☐ Previously generated by the project team
☐ From previous research conducted by others
☐ Other (please specify): *Click or tap here to type your answer.*

© Tec4MaaSEs Page 5 of 15



Data Management Plan – Questionnaire for Data Provider

**10. What criteria did you use to determine the sustainability of the existing data? (select all that apply)**

☐ Data quality
☐ Data format and compatibility
☐ Data relevance to the project's objectives
☐ Availability of the metadata
☐ Compliance with legal and ethical requirements
☐ Other (please specify): *Click or tap here to type your answer.*

**11. Are there any limitations or restrictions to re-use the existing data, such as copyright?**

☐ Yes ☐ No ☐ Not sure yet

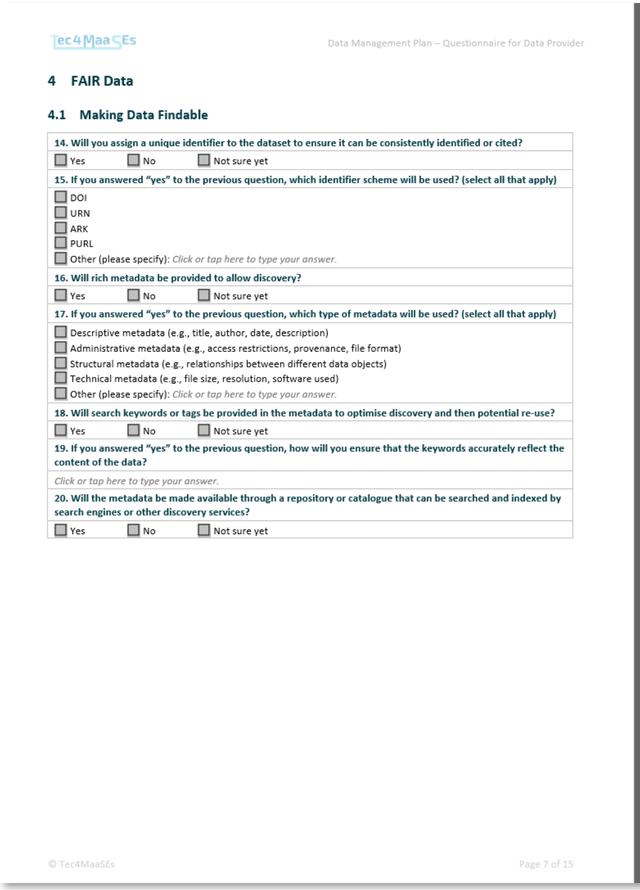
**12. Have you obtained necessary permission or authorization from the data owner(s) to re-use the data?**

☐ Yes ☐ No ☐ Other (please specify): *Click or tap here to type your answer.*

**13. Have you applied any data aggregation, minimization or anonymization method to the original data?**

☐ Yes ☐ No ☐ Other (please specify): *Click or tap here to type your answer.*

© Tec4MaaSEs Page 6 of 15



Data Management Plan – Questionnaire for Data Provider

## 4 FAIR Data

### 4.1 Making Data Findable

**14. Will you assign a unique identifier to the dataset to ensure it can be consistently identified or cited?**

□ Yes □ No □ Not sure yet

**15. If you answered "yes" to the previous question, which identifier scheme will be used? (select all that apply)**

□ DOI
□ URN
□ ARK
□ PURL
□ Other (please specify): *Click or tap here to type your answer.*

**16. Will rich metadata be provided to allow discovery?**

□ Yes □ No □ Not sure yet

**17. If you answered "yes" to the previous question, which type of metadata will be used? (select all that apply)**

□ Descriptive metadata (e.g., title, author, date, description)
□ Administrative metadata (e.g., access restrictions, provenance, file format)
□ Structural metadata (e.g., relationships between different data objects)
□ Technical metadata (e.g., file size, resolution, software used)
□ Other (please specify): *Click or tap here to type your answer.*

**18. Will search keywords or tags be provided in the metadata to optimise discovery and then potential re-use?**

□ Yes □ No □ Not sure yet

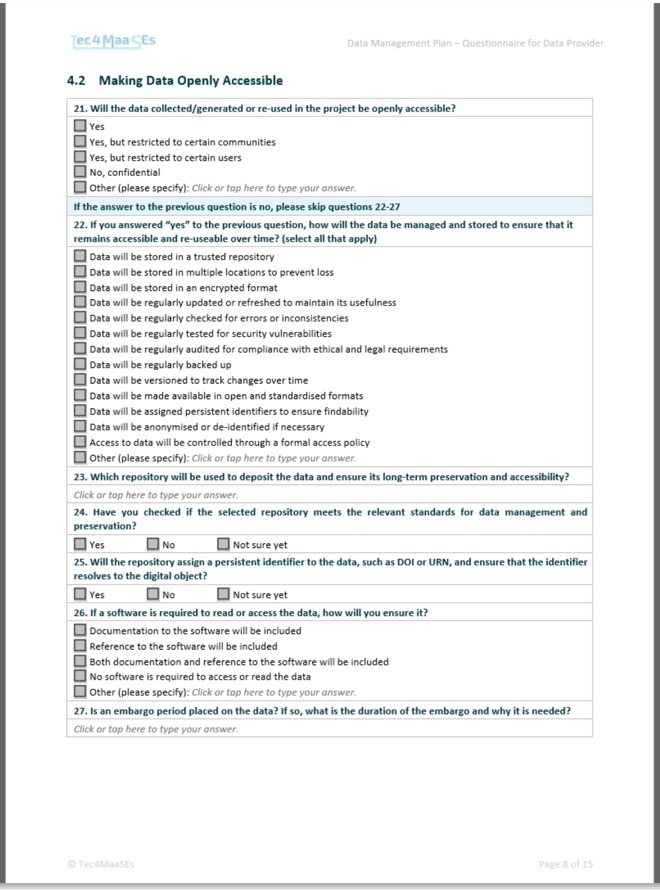
**19. If you answered "yes" to the previous question, how will you ensure that the keywords accurately reflect the content of the data?**

*Click or tap here to type your answer.*

**20. Will the metadata be made available through a repository or catalogue that can be searched and indexed by search engines or other discovery services?**

□ Yes □ No □ Not sure yet

© Tec4MaaSEs Page 7 of 15



Data Management Plan – Questionnaire for Data Provider

### 4.2 Making Data Openly Accessible

**21. Will the data collected/generated or re-used in the project be openly accessible?**

□ Yes
□ Yes, but restricted to certain communities
□ Yes, but restricted to certain users
□ No, confidential
□ Other (please specify): *Click or tap here to type your answer.*

**If the answer to the previous question is no, please skip questions 22-27**

**22. If you answered "yes" to the previous question, how will the data be managed and stored to ensure that it remains accessible and re-useable over time? (select all that apply)**

□ Data will be stored in a trusted repository
□ Data will be stored in multiple locations to prevent loss
□ Data will be stored in an encrypted format
□ Data will be regularly updated or refreshed to maintain its usefulness
□ Data will be regularly checked for errors or inconsistencies
□ Data will be regularly tested for security vulnerabilities
□ Data will be regularly audited for compliance with ethical and legal requirements
□ Data will be regularly backed up
□ Data will be versioned to track changes over time
□ Data will be made available in open and standardised formats
□ Data will be assigned persistent identifiers to ensure findability
□ Data will be anonymised or de-identified if necessary
□ Access to data will be controlled through a formal access policy
□ Other (please specify): *Click or tap here to type your answer.*

**23. Which repository will be used to deposit the data and ensure its long-term preservation and accessibility?**

*Click or tap here to type your answer.*

**24. Have you checked if the selected repository meets the relevant standards for data management and preservation?**

□ Yes □ No □ Not sure yet

**25. Will the repository assign a persistent identifier to the data, such as DOI or URN, and ensure that the identifier resolves to the digital object?**

□ Yes □ No □ Not sure yet

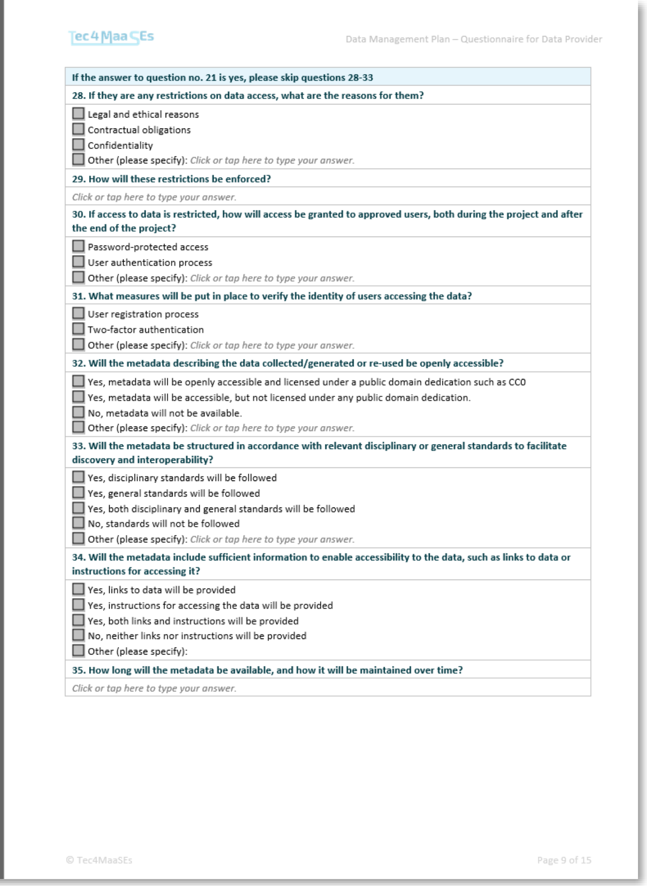
**26. If a software is required to read or access the data, how will you ensure it?**

□ Documentation to the software will be included
□ Reference to the software will be included
□ Both documentation and reference to the software will be included
□ No software is required to access or read the data
□ Other (please specify): *Click or tap here to type your answer.*

**27. Is an embargo period placed on the data? If so, what is the duration of the embargo and why it is needed?**

*Click or tap here to type your answer.*

© Tec4MaaSEs Page 8 of 15



Data Management Plan – Questionnaire for Data Provider

**If the answer to question no. 21 is yes, please skip questions 28-33**

**28. If they are any restrictions on data access, what are the reasons for them?**

□ Legal and ethical reasons
□ Contractual obligations
□ Confidentiality
□ Other (please specify): *Click or tap here to type your answer.*

**29. How will these restrictions be enforced?**

*Click or tap here to type your answer.*

**30. If access to data is restricted, how will access be granted to approved users, both during the project and after the end of the project?**

□ Password-protected access
□ User authentication process
□ Other (please specify): *Click or tap here to type your answer.*

**31. What measures will be put in place to verify the identity of users accessing the data?**

□ User registration process
□ Two-factor authentication
□ Other (please specify): *Click or tap here to type your answer.*

**32. Will the metadata describing the data collected/generated or re-used be openly accessible?**

□ Yes, metadata will be openly accessible and licensed under a public domain dedication such as CC0
□ Yes, metadata will be accessible, but not licensed under any public domain dedication.
□ No, metadata will not be available.
□ Other (please specify): *Click or tap here to type your answer.*

**33. Will the metadata be structured in accordance with relevant disciplinary or general standards to facilitate discovery and interoperability?**

□ Yes, disciplinary standards will be followed
□ Yes, general standards will be followed
□ Yes, both disciplinary and general standards will be followed
□ No, standards will not be followed
□ Other (please specify): *Click or tap here to type your answer.*

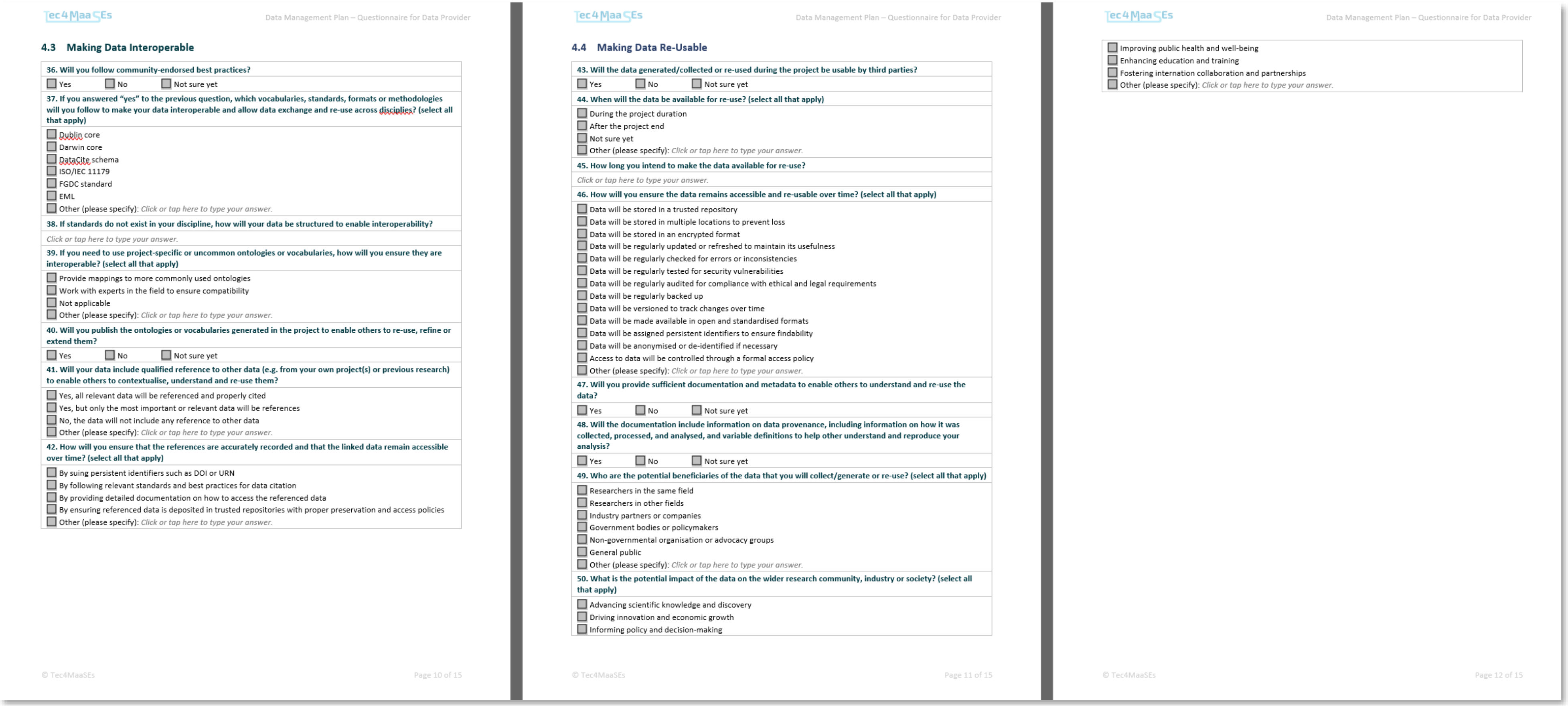
**34. Will the metadata include sufficient information to enable accessibility to the data, such as links to data or instructions for accessing it?**

□ Yes, links to data will be provided
□ Yes, instructions for accessing the data will be provided
□ Yes, both links and instructions will be provided
□ No, neither links nor instructions will be provided
□ Other (please specify):

**35. How long will the metadata be available, and how it will be maintained over time?**

*Click or tap here to type your answer.*

© Tec4MaaSEs Page 9 of 15



Data Management Plan – Questionnaire for Data Provider

## 4.3 Making Data Interoperable

**36. Will you follow community-endorsed best practices?**

□ Yes □ No □ Not sure yet

**37. If you answered "yes" to the previous question, which vocabularies, standards, formats or methodologies will you follow to make your data interoperable and allow data exchange and re-use across disciplies? (select all that apply)**

□ Dublin core
□ Darwin core
□ DataCite schema
□ ISO/IEC 11179
□ FGDC standard
□ EML
□ Other (please specify): *Click or tap here to type your answer.*

**38. If standards do not exist in your discipline, how will your data be structured to enable interoperability?**

*Click or tap here to type your answer.*

**39. If you need to use project-specific or uncommon ontologies or vocabularies, how will you ensure they are interoperable? (select all that apply)**

□ Provide mappings to more commonly used ontologies
□ Work with experts in the field to ensure compatibility
□ Not applicable
□ Other (please specify): *Click or tap here to type your answer.*

**40. Will you publish the ontologies or vocabularies generated in the project to enable others to re-use, refine or extend them?**

□ Yes □ No □ Not sure yet

**41. Will your data include qualified reference to other data (e.g. from your own project(s) or previous research) to enable others to contextualise, understand and re-use them?**

□ Yes, all relevant data will be referenced and properly cited
□ Yes, but only the most important or relevant data will be references
□ No, the data will not include any reference to other data
□ Other (please specify): *Click or tap here to type your answer.*

**42. How will you ensure that the references are accurately recorded and that the linked data remain accessible over time? (select all that apply)**

□ By suing persistent identifiers such as DOI or URN
□ By following relevant standards and best practices for data citation
□ By providing detailed documentation on how to access the referenced data
□ By ensuring referenced data is deposited in trusted repositories with proper preservation and access policies
□ Other (please specify): *Click or tap here to type your answer.*

© Tec4MaaSEs Page 10 of 15

Data Management Plan – Questionnaire for Data Provider

## 4.4 Making Data Re-Usable

**43. Will the data generated/collected or re-used during the project be usable by third parties?**

□ Yes □ No □ Not sure yet

**44. When will the data be available for re-use? (select all that apply)**

□ During the project duration
□ After the project end
□ Not sure yet
□ Other (please specify): *Click or tap here to type your answer.*

**45. How long you intend to make the data available for re-use?**

*Click or tap here to type your answer.*

**46. How will you ensure the data remains accessible and re-usable over time? (select all that apply)**

□ Data will be stored in a trusted repository
□ Data will be stored in multiple locations to prevent loss
□ Data will be stored in an encrypted format
□ Data will be regularly updated or refreshed to maintain its usefulness
□ Data will be regularly checked for errors or inconsistencies
□ Data will be regularly tested for security vulnerabilities
□ Data will be regularly audited for compliance with ethical and legal requirements
□ Data will be regularly backed up
□ Data will be versioned to track changes over time
□ Data will be made available in open and standardised formats
□ Data will be assigned persistent identifiers to ensure findability
□ Data will be anonymised or de-identified if necessary
□ Access to data will be controlled through a formal access policy
□ Other (please specify): *Click or tap here to type your answer.*

**47. Will you provide sufficient documentation and metadata to enable others to understand and re-use the data?**

□ Yes □ No □ Not sure yet

**48. Will the documentation include information on data provenance, including information on how it was collected, processed, and analysed, and variable definitions to help other understand and reproduce your analysis?**

□ Yes □ No □ Not sure yet

**49. Who are the potential beneficiaries of the data that you will collect/generate or re-use? (select all that apply)**

□ Researchers in the same field
□ Researchers in other fields
□ Industry partners or companies
□ Government bodies or policymakers
□ Non-governmental organisation or advocacy groups
□ General public
□ Other (please specify): *Click or tap here to type your answer.*

**50. What is the potential impact of the data on the wider research community, industry or society? (select all that apply)**

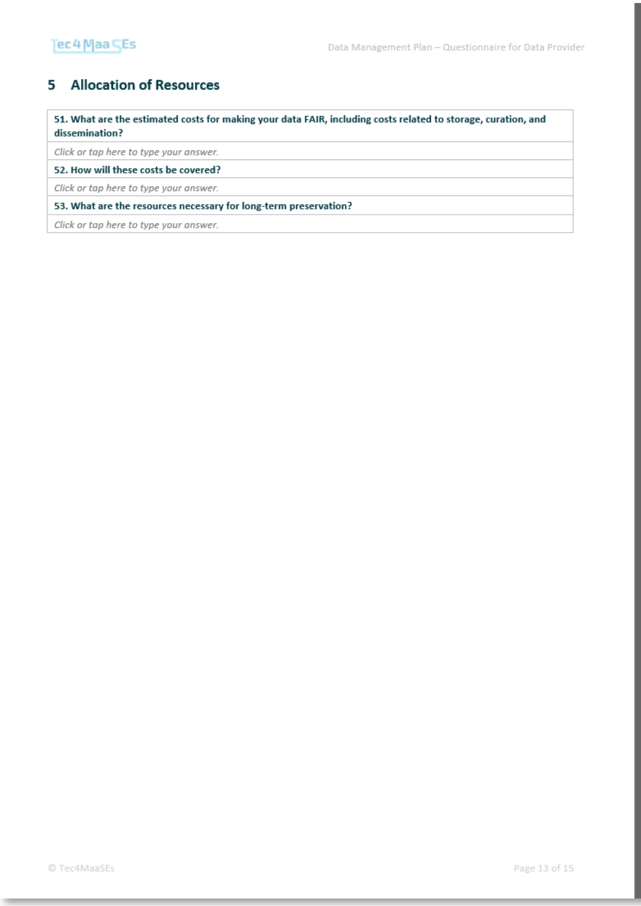
□ Advancing scientific knowledge and discovery
□ Driving innovation and economic growth
□ Informing policy and decision-making

© Tec4MaaSEs Page 11 of 15

Data Management Plan – Questionnaire for Data Provider

□ Improving public health and well-being
□ Enhancing education and training
□ Fostering internation collaboration and partnerships
□ Other (please specify): *Click or tap here to type your answer.*

© Tec4MaaSEs Page 12 of 15



Data Management Plan – Questionnaire for Data Provider

## 5 Allocation of Resources

**51. What are the estimated costs for making your data FAIR, including costs related to storage, curation, and dissemination?**

*Click or tap here to type your answer.*

**52. How will these costs be covered?**

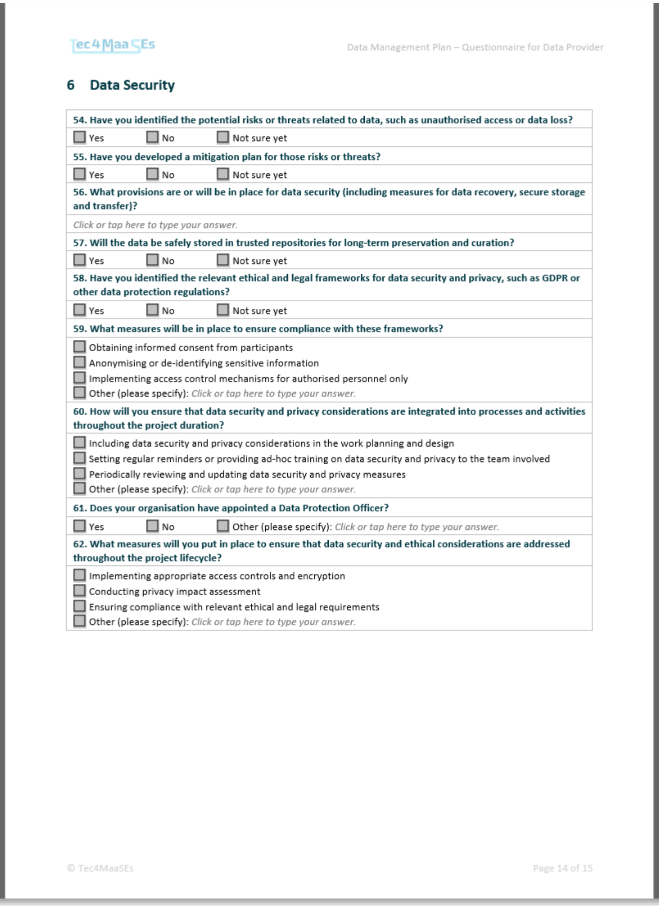
*Click or tap here to type your answer.*

**53. What are the resources necessary for long-term preservation?**

*Click or tap here to type your answer.*

© Tec4MaaSEs

Page 13 of 15



Data Management Plan – Questionnaire for Data Provider

## 6 Data Security

**54. Have you identified the potential risks or threats related to data, such as unauthorised access or data loss?**

□ Yes □ No □ Not sure yet

**55. Have you developed a mitigation plan for those risks or threats?**

□ Yes □ No □ Not sure yet

**56. What provisions are or will be in place for data security (including measures for data recovery, secure storage and transfer)?**

*Click or tap here to type your answer.*

**57. Will the data be safely stored in trusted repositories for long-term preservation and curation?**

□ Yes □ No □ Not sure yet

**58. Have you identified the relevant ethical and legal frameworks for data security and privacy, such as GDPR or other data protection regulations?**

□ Yes □ No □ Not sure yet

**59. What measures will be in place to ensure compliance with these frameworks?**

□ Obtaining informed consent from participants
□ Anonymising or de-identifying sensitive information
□ Implementing access control mechanisms for authorised personnel only
□ Other (please specify): *Click or tap here to type your answer.*

**60. How will you ensure that data security and privacy considerations are integrated into processes and activities throughout the project duration?**

□ Including data security and privacy considerations in the work planning and design
□ Setting regular reminders or providing ad-hoc training on data security and privacy to the team involved
□ Periodically reviewing and updating data security and privacy measures
□ Other (please specify): *Click or tap here to type your answer.*

**61. Does your organisation have appointed a Data Protection Officer?**

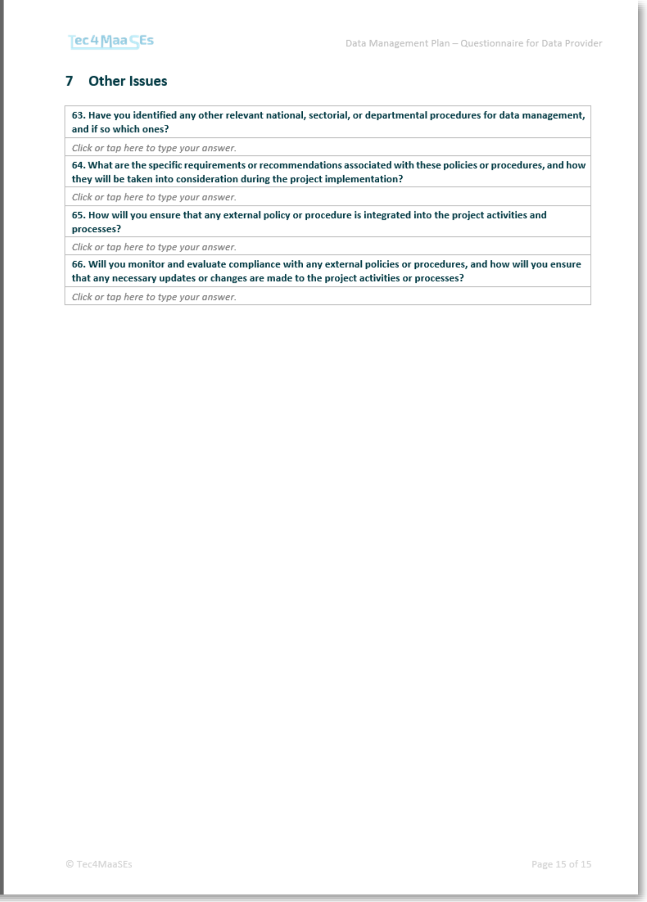
□ Yes □ No □ Other (please specify): *Click or tap here to type your answer.*

**62. What measures will you put in place to ensure that data security and ethical considerations are addressed throughout the project lifecycle?**

□ Implementing appropriate access controls and encryption
□ Conducting privacy impact assessment
□ Ensuring compliance with relevant ethical and legal requirements
□ Other (please specify): *Click or tap here to type your answer.*

© Tec4MaaSEs

Page 14 of 15



Data Management Plan – Questionnaire for Data Provider

## 7 Other Issues

**63. Have you identified any other relevant national, sectorial, or departmental procedures for data management, and if so which ones?**

*Click or tap here to type your answer.*

**64. What are the specific requirements or recommendations associated with these policies or procedures, and how they will be taken into consideration during the project implementation?**

*Click or tap here to type your answer.*

**65. How will you ensure that any external policy or procedure is integrated into the project activities and processes?**

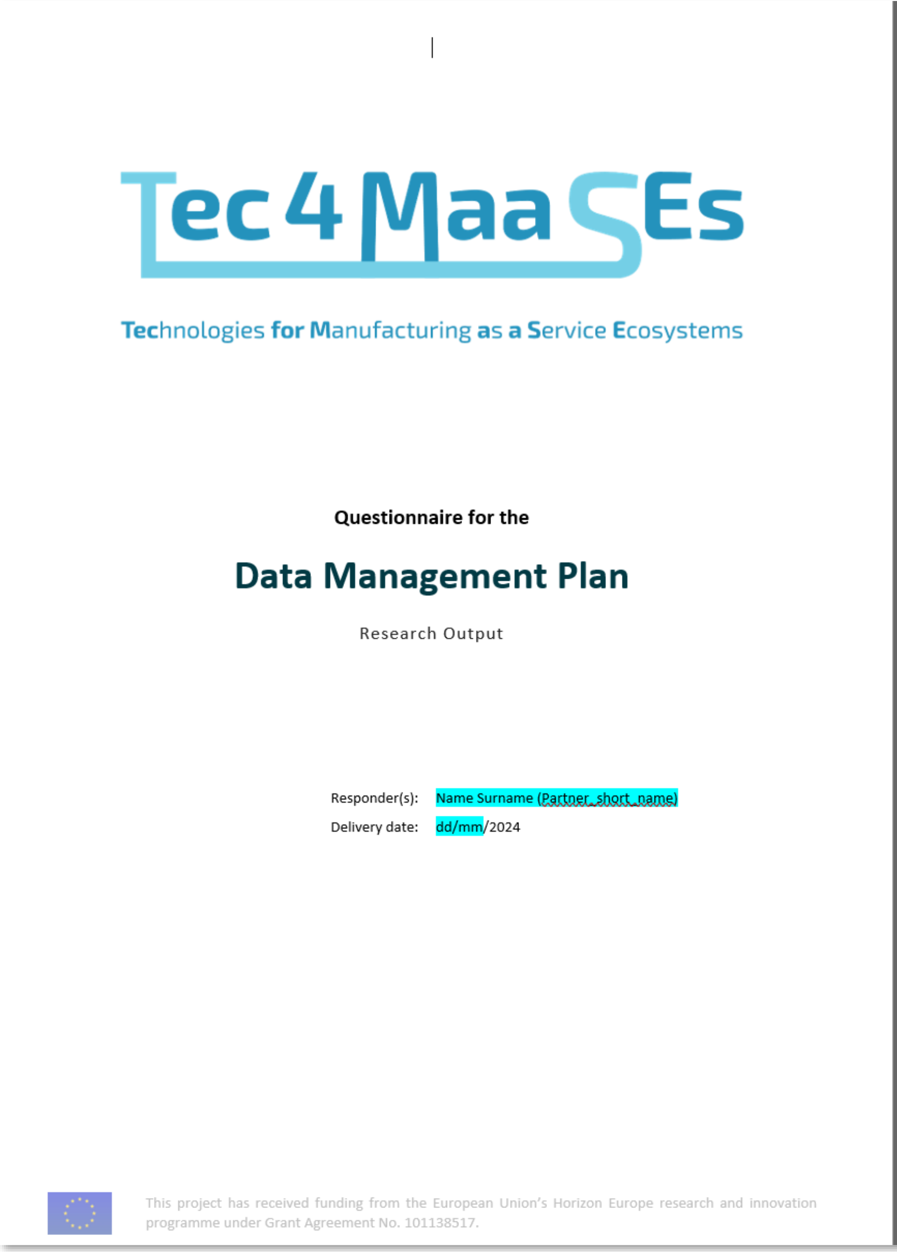
*Click or tap here to type your answer.*

**66. Will you monitor and evaluate compliance with any external policies or procedures, and how will you ensure that any necessary updates or changes are made to the project activities or processes?**

*Click or tap here to type your answer.*

© Tec4MaaSEs

Page 15 of 15



Tec4MaaSEs

Technologies for Manufacturing as a Service Ecosystems

Questionnaire for the

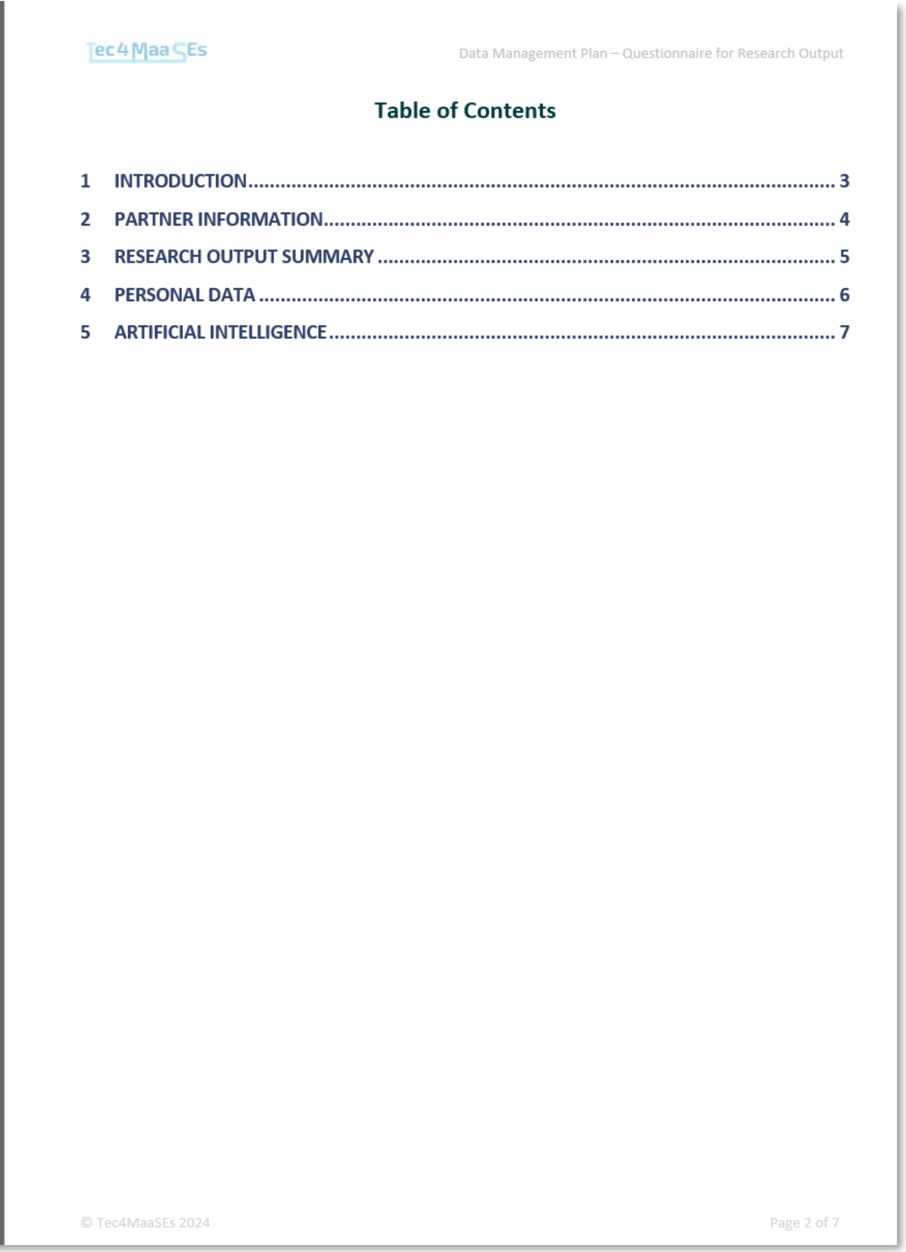
# Data Management Plan

Research Output

Responder(s): Name Surname (Partner_short_name)

Delivery date: dd/mm/2024

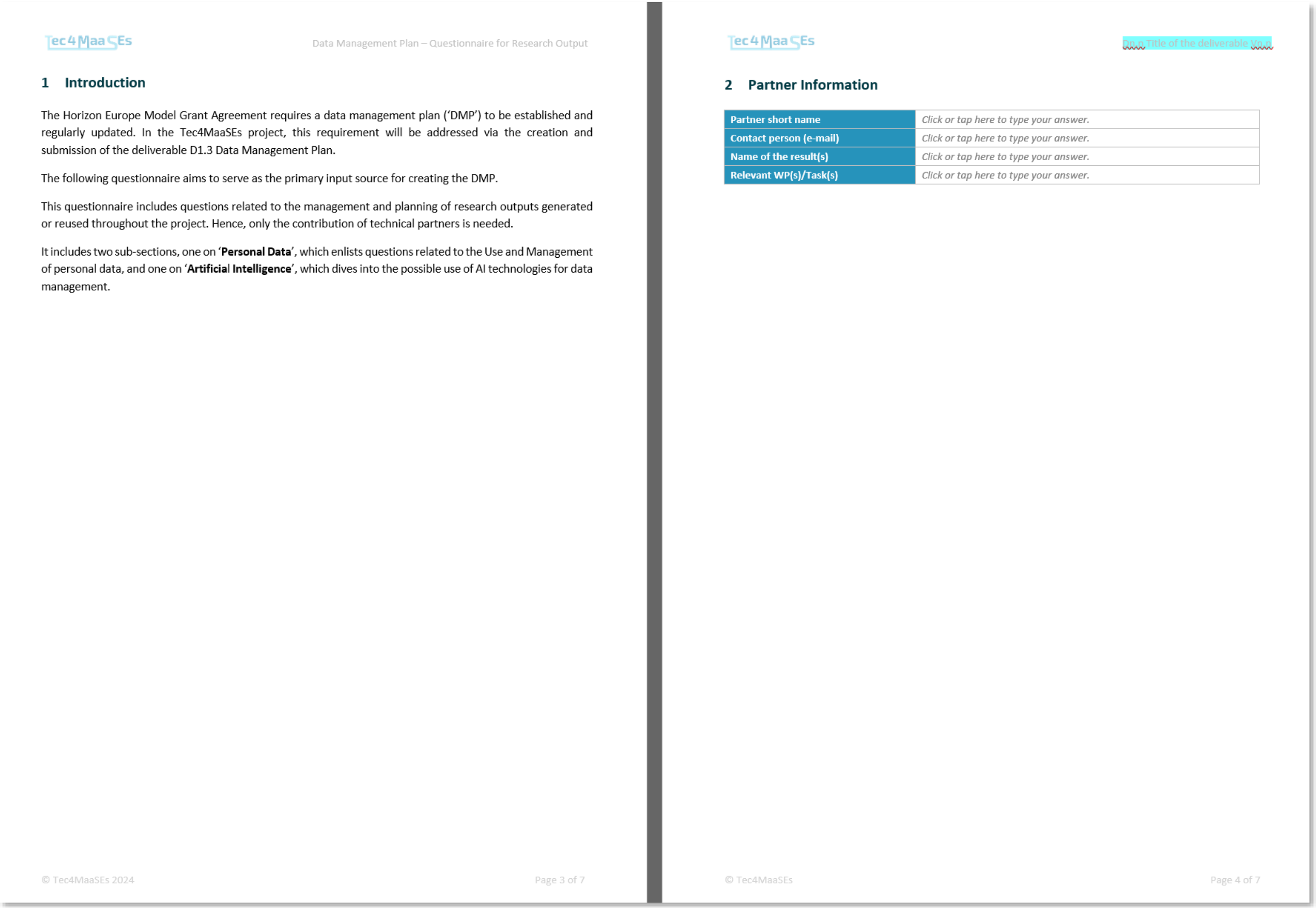
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101138517.



Tec4MaaSEs Data Management Plan – Questionnaire for Research Output

## Table of Contents

© Tec4MaaSEs 2024 Page 2 of 7



Tec4MaaSEs Data Management Plan – Questionnaire for Research Output

## 1 Introduction

The Horizon Europe Model Grant Agreement requires a data management plan ('DMP') to be established and regularly updated. In the Tec4MaaSEs project, this requirement will be addressed via the creation and submission of the deliverable D1.3 Data Management Plan.

The following questionnaire aims to serve as the primary input source for creating the DMP.

This questionnaire includes questions related to the management and planning of research outputs generated or reused throughout the project. Hence, only the contribution of technical partners is needed.

It includes two sub-sections, one on '**Personal Data**', which enlists questions related to the Use and Management of personal data, and one on '**Artificial Intelligence**', which dives into the possible use of AI technologies for data management.

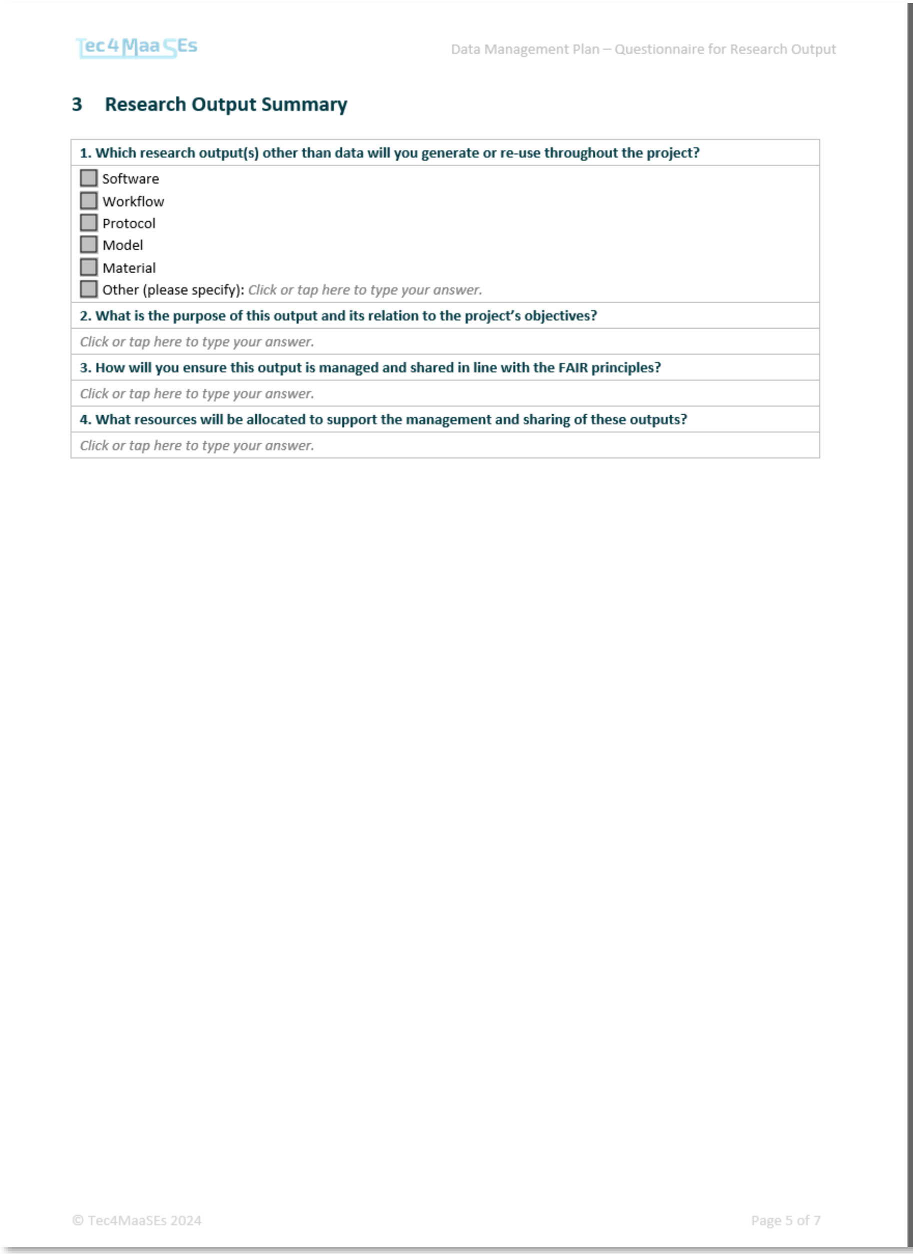
© Tec4MaaSEs 2024 Page 3 of 7

Tec4MaaSEs Dn.n Title of the deliverable Vn.n

## 2 Partner Information

| | |
|---|---|
| **Partner short name** | *Click or tap here to type your answer.* |
| **Contact person (e-mail)** | *Click or tap here to type your answer.* |
| **Name of the result(s)** | *Click or tap here to type your answer.* |
| **Relevant WP(s)/Task(s)** | *Click or tap here to type your answer.* |

© Tec4MaaSEs Page 4 of 7



Tec4MaaSEs Data Management Plan – Questionnaire for Research Output

## 3 Research Output Summary

**1. Which research output(s) other than data will you generate or re-use throughout the project?**

☐ Software
☐ Workflow
☐ Protocol
☐ Model
☐ Material
☐ Other (please specify): *Click or tap here to type your answer.*

**2. What is the purpose of this output and its relation to the project's objectives?**

*Click or tap here to type your answer.*

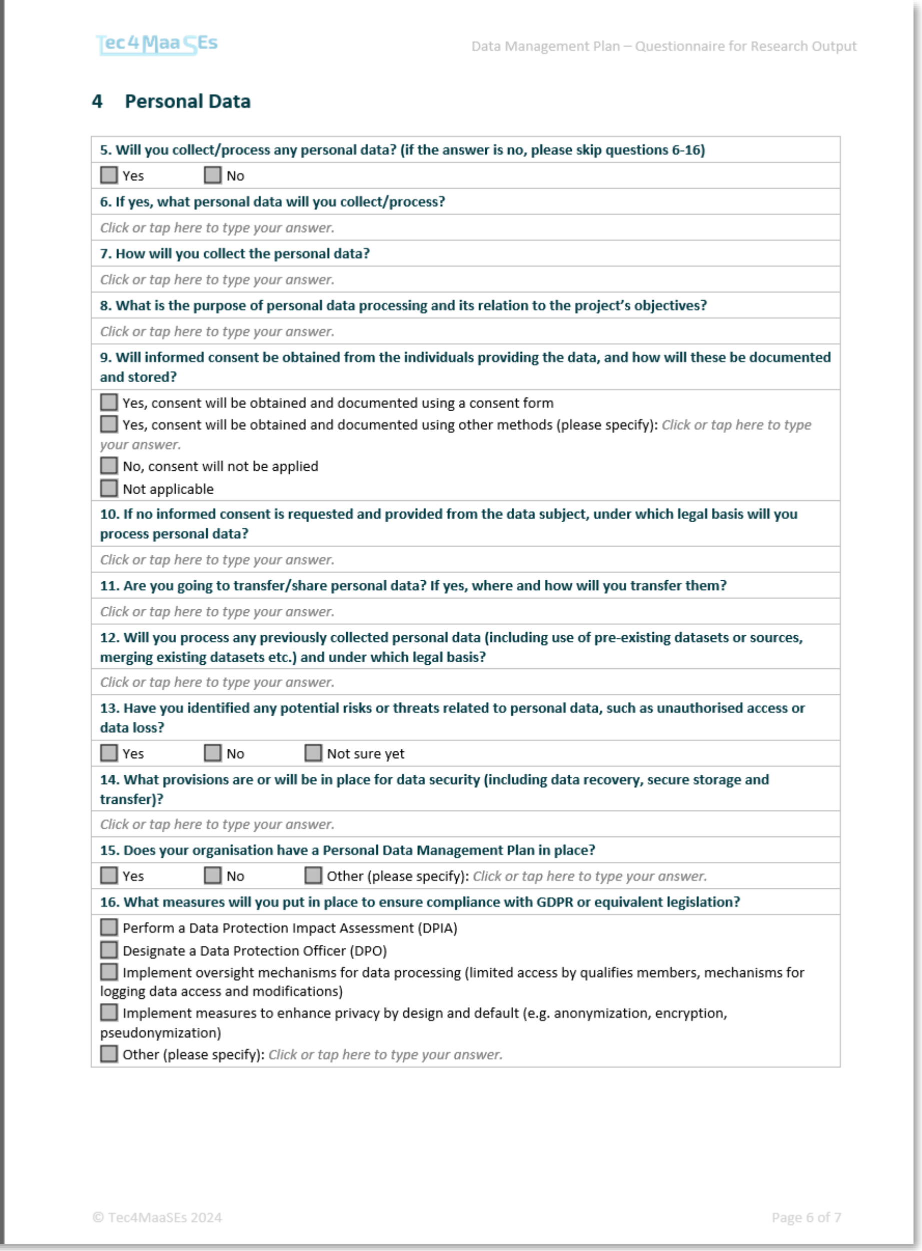
**3. How will you ensure this output is managed and shared in line with the FAIR principles?**

*Click or tap here to type your answer.*

**4. What resources will be allocated to support the management and sharing of these outputs?**

*Click or tap here to type your answer.*

© Tec4MaaSEs 2024 Page 5 of 7



Tec4MaaSEs Data Management Plan – Questionnaire for Research Output

## 4 Personal Data

**5. Will you collect/process any personal data? (if the answer is no, please skip questions 6-16)**

☐ Yes ☐ No

**6. If yes, what personal data will you collect/process?**

*Click or tap here to type your answer.*

**7. How will you collect the personal data?**

*Click or tap here to type your answer.*

**8. What is the purpose of personal data processing and its relation to the project's objectives?**

*Click or tap here to type your answer.*

**9. Will informed consent be obtained from the individuals providing the data, and how will these be documented and stored?**

☐ Yes, consent will be obtained and documented using a consent form
☐ Yes, consent will be obtained and documented using other methods (please specify): *Click or tap here to type your answer.*
☐ No, consent will not be applied
☐ Not applicable

**10. If no informed consent is requested and provided from the data subject, under which legal basis will you process personal data?**

*Click or tap here to type your answer.*

**11. Are you going to transfer/share personal data? If yes, where and how will you transfer them?**

*Click or tap here to type your answer.*

**12. Will you process any previously collected personal data (including use of pre-existing datasets or sources, merging existing datasets etc.) and under which legal basis?**

*Click or tap here to type your answer.*

**13. Have you identified any potential risks or threats related to personal data, such as unauthorised access or data loss?**

☐ Yes ☐ No ☐ Not sure yet

**14. What provisions are or will be in place for data security (including data recovery, secure storage and transfer)?**

*Click or tap here to type your answer.*

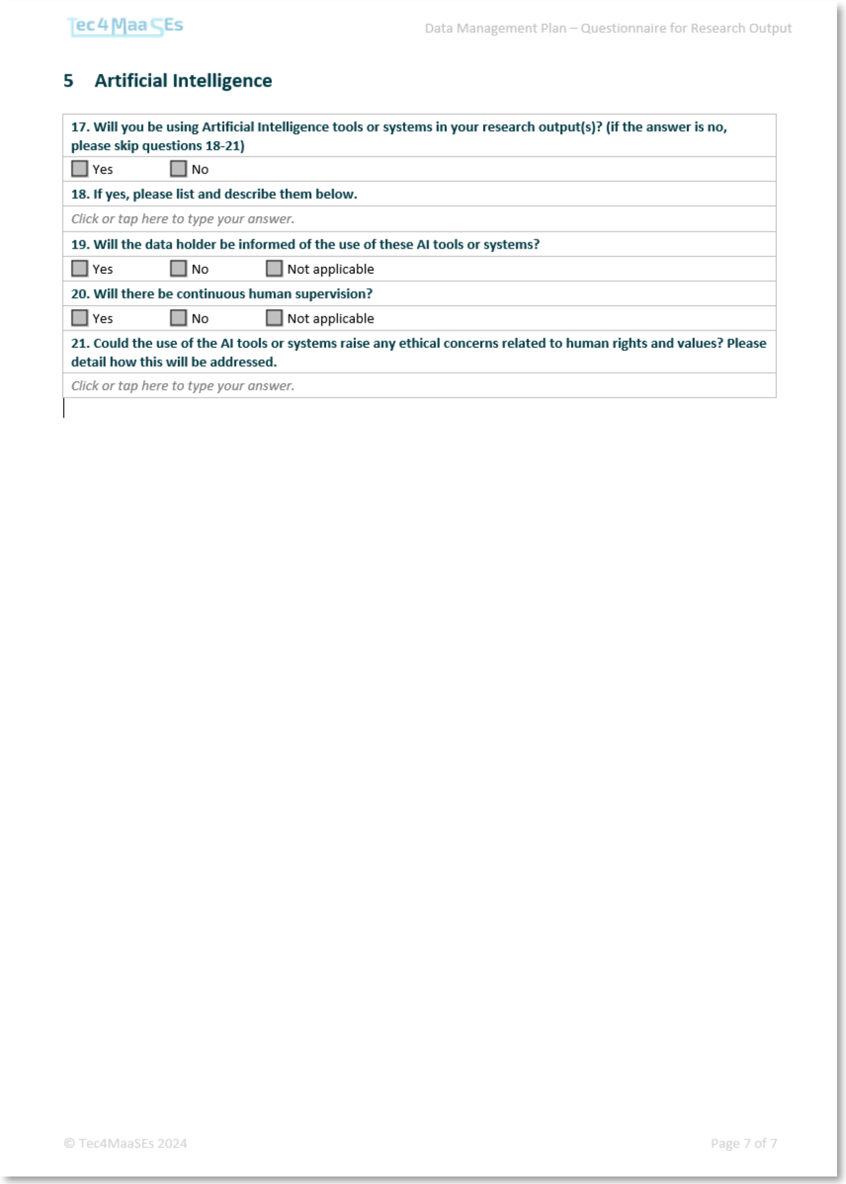
**15. Does your organisation have a Personal Data Management Plan in place?**

☐ Yes ☐ No ☐ Other (please specify): *Click or tap here to type your answer.*

**16. What measures will you put in place to ensure compliance with GDPR or equivalent legislation?**

☐ Perform a Data Protection Impact Assessment (DPIA)
☐ Designate a Data Protection Officer (DPO)
☐ Implement oversight mechanisms for data processing (limited access by qualifies members, mechanisms for logging data access and modifications)
☐ Implement measures to enhance privacy by design and default (e.g. anonymization, encryption, pseudonymization)
☐ Other (please specify): *Click or tap here to type your answer.*

© Tec4MaaSEs 2024 Page 6 of 7



## 5 Artificial Intelligence

| |
|---|
| **17. Will you be using Artificial Intelligence tools or systems in your research output(s)? (if the answer is no, please skip questions 18-21)** |
| □ Yes □ No |
| **18. If yes, please list and describe them below.** |
| *Click or tap here to type your answer.* |
| **19. Will the data holder be informed of the use of these AI tools or systems?** |
| □ Yes □ No □ Not applicable |
| **20. Will there be continuous human supervision?** |
| □ Yes □ No □ Not applicable |
| **21. Could the use of the AI tools or systems raise any ethical concerns related to human rights and values? Please detail how this will be addressed.** |
| *Click or tap here to type your answer.* |

## Annex 2: Data Description Template

| Making Data Findable | |
|---|---|
| **Name of dataset** | Name of the dataset |
| **Responsible partner** | Name of the partner responsible for the data generated/collected/used |
| **ID** | Unique identifier of the dataset, e.g. DOI. |
| **Description** | A brief description of the dataset |
| **Purpose** | Intended use of the data |
| **Type** | Type of the data (e.g. structured, unstructured) |
| **Format** | Format of the data |
| **Size** | Estimated size of the dataset |
| **Language** | Language(s) used in the dataset |
| **Source** | The source of data that have been generated/collected |
| **Version** | Unique identifier for each version of the dataset |
| **Date** | Release date (preferred format dd-mm-yyyy) |
| **Making Data Accessible** | |
| **Accessibility** | Open/Restricted |
| **Storage** | The repository where the data will be/has been stored. |
| **Restrictions** | Any limitation on data sharing and re-use. |
| **Making Data Interoperable** | |
| **Standards & metadata** | Metadata standards used (e.g. ISO/IEC 27001). |
| **Making Data Re-Usable** | |
| **Licensing** | License applied (e.g. Creative Commons 4.0) |
| **Keywords** | Specified keywords to increase the possibilities of re-use. |