Secure management of IoT devices lifecycle through identities, trust and distributed ledgers

D7.3 Data Management Plan v1

<table>
<thead>
<tr>
<th>Document Summary Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grant Agreement No</strong></td>
</tr>
<tr>
<td><strong>Acronym</strong></td>
</tr>
<tr>
<td><strong>Full Title</strong></td>
</tr>
<tr>
<td><strong>Start Date</strong></td>
</tr>
<tr>
<td><strong>Duration</strong></td>
</tr>
<tr>
<td><strong>Project URL</strong></td>
</tr>
<tr>
<td><strong>Deliverable</strong></td>
</tr>
<tr>
<td><strong>Work Package</strong></td>
</tr>
<tr>
<td><strong>Contractual due date</strong></td>
</tr>
<tr>
<td><strong>Actual submission date</strong></td>
</tr>
<tr>
<td><strong>Nature</strong></td>
</tr>
<tr>
<td><strong>Dissemination Level</strong></td>
</tr>
<tr>
<td><strong>Responsible author</strong></td>
</tr>
<tr>
<td><strong>Lead Beneficiary</strong></td>
</tr>
<tr>
<td><strong>Authors</strong></td>
</tr>
<tr>
<td><strong>Internal reviewers</strong></td>
</tr>
</tbody>
</table>
Revision history (including peer reviewing & quality control)

<table>
<thead>
<tr>
<th>Version</th>
<th>Issue Date</th>
<th>% Complete</th>
<th>Changes</th>
<th>Contributor(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>v1.0</td>
<td>05/12/2021</td>
<td>50%</td>
<td>Initial contents</td>
<td>Hui Song (SINTEF)</td>
</tr>
<tr>
<td>v2.0</td>
<td>10/12/2021</td>
<td>90%</td>
<td>Update and new content</td>
<td>Konstantinos Loupos (INLE)</td>
</tr>
<tr>
<td>v3.0</td>
<td>17/12/2021</td>
<td>100%</td>
<td>Partner inputs on potential data items, reviewing completed</td>
<td>All partners</td>
</tr>
<tr>
<td>v4.0</td>
<td>17/12/2021</td>
<td>100%</td>
<td>Final version</td>
<td>Hui Song (SINTEF)</td>
</tr>
</tbody>
</table>

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 authors(s) or any other participant in the ERATOSTHENES 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 ERATOSTHENES 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 ERATOSTHENES 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

© ERATOSTHENES Consortium, 2020-2025. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorised provided the source is acknowledged.
# Table of Contents

1 Executive Summary ................................................................. 6
2 Introduction ............................................................................... 7
  2.1 Mapping ERATOSTHENES Outputs ......................................... 7
  2.2 Deliverable Overview and Report Structure .............................. 8
3 Data Summary ........................................................................... 9
  3.1 ERATOSTHENES Data Lifecycle ............................................. 9
  3.2 Types of data assets ............................................................... 12
    3.2.1 IoT-related data ............................................................ 12
    3.2.2 Trust-related data ......................................................... 12
    3.2.3 Stakeholder-related data ............................................... 12
    3.2.4 Project management data .............................................. 13
  3.3 Data Analysis per WP ............................................................ 13
  3.4 Initial data assets ................................................................... 13
  3.5 DMP in ERATOSTHENES WPs and WP leaders’ responsibilities and Allocation of resources .......... 26
4 General Data Protection Regulation (GDPR) ................................. 27
  4.1 GDPR Compliancy ................................................................. 27
  4.2 General Data Protection Policy ............................................... 28
    4.2.1 Introduction .................................................................. 28
    4.2.2 Definitions ................................................................... 28
    4.2.3 Policy scope ................................................................ 28
    4.2.4 Personal data processing ................................................. 29
    4.2.5 Data protection legal roles .............................................. 30
    4.2.6 Notice and consent ......................................................... 32
    4.2.7 Rights of data subjects .................................................... 33
    4.2.8 Data protection documentation system .............................. 35
    4.2.9 Data protection assessment ............................................. 35
    4.2.10 Technical and organizational measures ............................ 36
    4.2.11 Data breach .................................................................. 36
    4.2.12 Data transfers to third countries ................................... 36
    4.2.13 Sanctions and damages ................................................ 36
    4.2.14 ERATOSTHENES Web-server Personal Data Protection and Privacy policy ....................... 36
    4.2.15 ERATOSTHENES Day-to-Day Data Usage and Related Processes ............................ 38
5 Making Data FAIR ....................................................................... 40
  5.1 Making data findable .............................................................. 40
  5.2 Making data openly accessible ............................................... 40
    5.2.1 Open Access to scientific publications ............................... 40
    5.2.2 How will the data be made available? .............................. 41
  5.3 Making data interoperable ...................................................... 41
  5.4 Increase data re-use ............................................................... 42
  5.5 Security Aspects of Repository Server (TEAMWORK) ................ 42
    5.5.1 Physical Security .......................................................... 42
    5.5.2 Network Security .......................................................... 42
    5.5.3 Content Security ........................................................... 42
6 Conclusions ............................................................................... 43
Annex I – Data Management Report............................................... 44
Annex II - Non-Disclosure Agreement (DNA) for the Advisory Board (AB) members ................................. 46
List of Figures
Figure 1 - ERATOSTHENES Project Data Lifecycle .......................................................... 10

List of Tables
Table 1: Adherence to ERATOSTHENES GA Deliverable & Tasks Descriptions ........................................... 7
Table 2: ERATOSTHENES Data Creation and Collection Indicators .................................................................. 10
Table 3: ERATOSTHENES Data Processing Indicators .................................................................................. 11
Table 4: ERATOSTHENES Publication and Utilization Indicators ................................................................. 11
Table 5: ERATOSTHENES Storage and Re-use Indicators ............................................................................ 11
Table 6: Initial data asset identified at the beginning of the project ................................................................. 22
### Glossary of terms and abbreviations used

<table>
<thead>
<tr>
<th>Abbreviation / Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>Advisory Board</td>
</tr>
<tr>
<td>CSV</td>
<td>Comma Separated Values</td>
</tr>
<tr>
<td>DMP</td>
<td>Data management plan</td>
</tr>
<tr>
<td>DPIA</td>
<td>Data Protection Impact Assessment</td>
</tr>
<tr>
<td>DPO</td>
<td>Data Protection Officer</td>
</tr>
<tr>
<td>FAIR</td>
<td>Finding Accessible Interoperability Reusability</td>
</tr>
<tr>
<td>GDPR</td>
<td>General Data Protection Regulation</td>
</tr>
<tr>
<td>GA</td>
<td>Grant Agreement</td>
</tr>
<tr>
<td>IdM</td>
<td>Identity Management</td>
</tr>
<tr>
<td>IoT</td>
<td>Internet of Things</td>
</tr>
<tr>
<td>KPI</td>
<td>Key Performance Indicator</td>
</tr>
<tr>
<td>TEE</td>
<td>Trusted Execution Environment</td>
</tr>
<tr>
<td>TLS</td>
<td>Transport Layer Security</td>
</tr>
<tr>
<td>MB</td>
<td>MegaByte</td>
</tr>
<tr>
<td>MBPS</td>
<td>Megabyte Per Second</td>
</tr>
<tr>
<td>WP</td>
<td>Work Package</td>
</tr>
</tbody>
</table>
1 Executive Summary

This deliverable presents consortium's plan on handling research data during and after the project end, types of data collected and processed and/or generated, methodology and standards to be applied, sharing or open access and how the data will be curated/preserved in line with the H2020 Guidelines and FAIR (findable, accessible, interoperable, and reusable) data management. The deliverable presents detailed information on the project data lifecycle, privacy, and the project’s policies for data collection, storage, access, sharing, protection, retention, and destruction.

In more detail, ERATOSTHENES data management plan consolidates and structures the project’s processes, procedures and activities to support an effective and efficient data management methodology associated to the project schedule, needs and scope. This begins with a detailed examination and structuring of the data present and relevant to ERATOSTHENES activities, WPs, tasks and pilots, as well as their lifecycle and all other related principles for complying to their FAIR treatment (conventions, openness, metadata, reusability etc.). Following the data management is positioned at a high operational layer, clarifying and positioning data management ownerships and responsibilities for each WP leaders, partners etc., whilst also defining the data categories and special treatment for each. ERATOSTHENES will fully comply and respects GDPR policies through an agreed holistic policy for handling all data within the project’s activities and across the project’s WPs and tasks. ERATOSTHENES intends to share all project data in the ZENODO platform. Data management reporting, consolidation and check lists have been added as annexes followed by the ERATOSTHENES non-disclosure agreement for the project Advisory Board (AB).
2 Introduction

The ERATOSTHENS Data Management Plan (DMP) is considered as a key element of good data management. This DMP presents a plan on handling research data during and after project end, including types of data collected and processed, methodology and standards to be applied etc. The DMP presents detailed information on the data lifecycle, privacy, and project's policies for the data collection and processing (including storage, access, sharing, retention, destruction) by ERATOSTHENS. In particular, as part of making research data findable, accessible, interoperable and re-usable (FAIR), this DMP includes information on:

- The handling of research data during and after the end of the project;
- What data will be collected, processed, and generated;
- Which methodology and standards will be applied;
- Whether data will be shared and / or made open access;
- How the data will be curated and preserved.

The DMP is aligned with internal ethics for adequate ethical standards and adequate data protection measures. ERATOSTHENES collects data (including personal data) from different data sources and we classify the data into following categories:

- Qualitative and quantitative research data, i.e., data from pilots.
- Administrative data, e.g., participants details, communications, identity management data.
- Data from public sources, e.g., legislation, government guidance, codes of practices, results of ethical horizon scanning.
- Open source data collected from publicly available sources.
- Publications and dissemination data, e.g., data related to open peer-reviewed publications, interviews, reports, proceedings, stakeholders, capacity programme, contact details for webinars/workshops, dissemination contacts and enquiries.

The DMP is a living document and will be updated regularly, as new insights on data use are to be expected throughout the project.

The structure of this DMP is oriented towards the template provided by the European Commission:


2.1 Mapping ERATOSTHENS Outputs

The purpose of this section, is to map ERATOSTHENS Grant Agreement commitments, both within the formal Deliverable and Task description, against the project’s respective outputs and work performed.

Table 1: Adherence to ERATOSTHENS GA Deliverable & Tasks Descriptions

<table>
<thead>
<tr>
<th>ERATOSTHENS GA Component Title</th>
<th>ERATOSTHENS GA Component Outline</th>
<th>Respective Document Chapter(s)</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELIVERABLE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D7.3 - Data Management Plan v1</td>
<td>DMP initial version (M4) including data definitions, management procedures, GDPR policies (T7.4).</td>
<td>Chapters 3-5</td>
<td>Data management plan as described in the DoA including all project data management procedures, templates, a GDPR policy and FAIR principles.</td>
</tr>
<tr>
<td>TASKS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Produce the Data Management Plan (DMP, D7.3) including and respecting GDPR policies and procedures for personal information protection. The plan will cover the rules of handling research data during and after the project, including the characterization of the data that will be collected, processed or generated. Special attention will be paid to Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and repealing Directive 95/46/EC (GDPR Data Protection Regulation).

<table>
<thead>
<tr>
<th>T7.4 Quality assurance, data management, and risk management</th>
<th>Chapters 3-5</th>
<th>The document the potential data assets that can be potentially generated from the ERATOSTHENES tasks, the methodology to make these data FAIR, and the way to protect the security and privacy of the data.</th>
</tr>
</thead>
</table>

### 2.2 Deliverable Overview and Report Structure

The rest of this document is structured as follows.

- Section 3 describes the initial data sets we have identified at the beginning of the project.
- Section 4 defines the internal methods we use to store, exchange and follow the data.
- Section 5 describes how we plan to make the data FAIR internally and externally.
- Section 6 concludes the report.
3 Data Summary

In this section of the data management plan (DMP), we define the concepts of data collection and the purpose of data collection in relation to the working structure of the ERATOSTHENES project. After consultation with the leaders of ERATOSTHENES WPs, we have described the following data points below:

- **Means** of data collection
- **Types** of data that will be collected (sensor information, source codes etc.)
- **Formatting** of the actual data
- **Data size** and growth rate predictions
- **Data reproduction** and re-usability (whenever applicable)
- **Data versioning** and control to align data following data modifications
- **Data handling** software and tools to generate/modify/process

What follows is a definition of the data that will be created/distributed for each of the ERATOSTHENES WPs. These are categorized as:

- Scientific data relating to data from the 3 ERATOSTHENES pilot sites.
- Source codes and algorithm development and/or training.
- ERATOSTHENES project management and project related documentation, reporting and management files.
- Dissemination and communication documents.

All ERATOSTHENES data files may include narrative texts, numbers, images, software codes, audio files, video files, internal/external reports. The structure of this chapter complies with the FAIR data management template of the EC (DMP component 1). What follows is a data summary for each of the ERATOSTHENES WPs regarding the following:

- **Data purpose**
- **Data format and types**
- **Re-used** Information
- **Data origin**
- **Data size**
- **Data utility**

In the chapters that follow, we define the data collection for each of the ERATOSTHENES WPs as well as the purpose of the data collected and how they relate to the project objectives. After the first round of analysis at the beginning of the project, we identified several potential data assets. These will be the initial input to the continual data management throughout the project: We will not only follow up the collection, storage and sharing of these data assets, but also use them as exemplars for identifying new data assets along the project work.

3.1.1 ERATOSTHENES Data Lifecycle

In this chapter we have examined the complete information life-cycle\(^1\) that will be taking part in ERATOSTHENES project. In this, we consider the diverse stages at which information will be made, overseen or utilized amid the total execution and later. What takes after, backed by Figure 1 underneath, is an examination of the information lifecycle as well as the ways to control, oversee and report the related information. In each of the chapters that take after, particular measurements for the information administration and control have been included and are afterward summarized into the information administration report format (Attach 1 – Information Administration Report) that will be utilized at different venture stages to control information administration compliance. The graph that takes after shows the ERATOSTHENES information lifecycle in a normal stream, not-excluding other stream of information between their life-cycle\(^2\).

---

\(^1\) Data Management Lifecycle and Software Lifecycle Management in the Context of Conducting Science, W. C. Lenhardt, S. Ahalt, B. Blanton, L. Christopherson, R. Idaszak, Journal of open research software

\(^2\) Data Lifecycle Overview, USGS Data Management
3.1.1.1  Data Creation/Collection

As implied by the title, this stage includes the data creation and/or collection as it relates to the various data provided by the ERATOSTHENES pilots as well as the project reports and other documents/spreadsheets. This includes the creation of the data by each of the respective owner and collection in an approach to be structured in appropriate formats and layouts to enable their processing by the other project components/modules. Specific metrics at this stage relate to the following:

Table 2: ERATOSTHENES Data Creation and Collection Indicators

<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Means of verification</th>
<th>Target Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format</td>
<td>Compliance with existing standards of data exchange</td>
<td>XLS, XML, etc.</td>
</tr>
<tr>
<td>Availability and Readability</td>
<td>Whole package of data available, non-corruption, whole percentage collected</td>
<td>100% received 100% accessible</td>
</tr>
<tr>
<td>Fit for Use</td>
<td>Data follow data compliance for proper processing and review</td>
<td>100% usable by intended beneficiary/ies</td>
</tr>
<tr>
<td>Consistency and Completeness</td>
<td>Data are consistent and complete for the intended purpose</td>
<td>Including 100% of information for the intended purpose</td>
</tr>
<tr>
<td>Relation</td>
<td>Data processing follows a precise relation to their purpose of collection</td>
<td>100% purpose precision</td>
</tr>
</tbody>
</table>

3.1.1.2  Data Processing and Analysis

This stage is related to the actual data processing by the various data processors that are the partners that will be having access to the data for processing or dissemination activities following the project needs and outcomes. During this stage we need to ensure that the suitable partners can perform data processing in a concise approach to fulfill the ERATOSTHENES needs. This stage includes all steps towards data verification, organization, transformation, integration and extraction for the intended use. Data analysis includes all the actions/methodology executed on the
actual data that describe existing facts, identify outlines, develop data clarifications etc. This stage is closely related to the processing stage previously described and forms as a consequent stage. Specific metrics at this stage relate to the following:

**Table 3: ERATOSTHENES Data Processing Indicators**

<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Means of verification</th>
<th>Target Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data logic</td>
<td>Data can be and are processed following a concise logic and approach</td>
<td>New and processed data follow precise data logic</td>
</tr>
<tr>
<td>Organization and Utility</td>
<td>Suitable content organization of data under processing</td>
<td>100% organized data</td>
</tr>
<tr>
<td>Validation</td>
<td>Ensuring that the data under processing are correct and relevant</td>
<td>100% validated and relevant data</td>
</tr>
<tr>
<td>Aggregation</td>
<td>Whenever multiple data need to be aggregated ensure that this is done in a concise approach</td>
<td>100% aggregate-able data</td>
</tr>
<tr>
<td>Transformation</td>
<td>Transformation of data to the proper format(s) for processing</td>
<td>Capability of data for transformation (if needed)</td>
</tr>
<tr>
<td>Calibration</td>
<td>Calibration of data for their intended purpose</td>
<td>Data properly calibrated</td>
</tr>
</tbody>
</table>

### 3.1.1.3 Data Publication and Utilization

In this stage, publication of data refers to the capability to share data openly to public whereas utilization includes the steps towards data sharing (internally to ERATOSTHENES). This implies that the data should be medium and agent independent, making sure that the transfer can be implemented in an automated or not approach. The purpose in this stage is to ensure that the data is shared with the appropriate controlling mechanisms to ensure protection of proprietary data as well as the data integrity itself [5]. This stage is closely linked with the next stage, i.e., data storage and archiving, as far as metadata are related to ensure data search-ability (as another feature of the FAIR data treatment). Specific metrics at this stage relate to the following:

**Table 4: ERATOSTHENES Publication and Utilization Indicators**

<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Means of verification</th>
<th>Target Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means-independent</td>
<td>Transferring of the data in a means-independent approach</td>
<td>100% means independent transferability</td>
</tr>
<tr>
<td>Security (a)</td>
<td>Data stored in a secure enough server</td>
<td>At least access control provided over a TLS protocol</td>
</tr>
</tbody>
</table>

### 3.1.1.4 Data Storage, Archiving and Re-Use

The storage and archiving stage is also a very critical stage as it relates to the data access, sharing, storage, archiving (including search capabilities) and re-use. An important factor here is the updated status of the data so that no newer versions exist (or is clearly indicated if newer versions does exist). This should also involve actions to ensure accidental data losses, corruption and unauthorized access. Data storage and archiving is also strongly linked to data re-usability that is also within the scope of the FAIR data treatment. Specific metrics at this stage relate to the following:

**Table 5: ERATOSTHENES Storage and Re-use Indicators**

<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Means of verification</th>
<th>Target Values</th>
</tr>
</thead>
</table>
### Up to date

| Ensuring that the stored data are up to date for the specific purpose and no later version exists | 100% updated |

### Meta Data

| Existence of meta data in stored files | Relevant metadata have been included into the archive per data set |

### Security (b)

| Access control provided | Access control setup |

### Security (c)

| Server is considered as safe enough (TLS connection protocol) | At least TLS connection configuration |

### Bandwidth

| Control of server bandwidth | Effective storage server bandwidth > 2 MBPS |

### Retention

| Properly setting expiration dates for all data after which the data will be deleted | Expiration date noted |

## 3.2 Types of data assets

The initial data assets indicate the following types of data that are most relevant to ERATOSTHENES.

### 3.2.1 IoT-related data

**Description.** The research, development and evaluation activities within ERATOSTHENES will involve IoT devices and systems. Data about the IoT system will be collected, such as the architecture and design of the system, the metadata and identity information of IoT devices, the runtime monitoring and feedbacks from the IoT systems, etc.

**Why are the data needed?** The various types of data from IoT devices will be used by ERATOSTHENES partners for the design, implementation and evaluation of trust and identity management techniques. Outside the project, the data can be useful for IoT system designers and researchers as reference for system design or benchmarking.

**Data format.** JSON and CSV files will be the common format for such data, but other formats will not be excluded.

### 3.2.2 Trust-related data

**Description.** We will also collect data related to the trust management of IoT devices and systems, such as the trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices, etc.

**Why are the data needed?** The data will be used internally for understanding, designing and implementing the trust management agent. The data will also be used to design the test cases for evaluating the trust management mechanisms. Externally, we expect the data to be useful for cyber-security researchers and practitioners in the domain of trustworthy IoT.

**Data format.** Due to the technical choices for the project implementation, part of these data will be in the format of blockchain records. Other possible format includes JSON or CSV documents.

### 3.2.3 Stakeholder-related data

**Description.** Part of the project work will involve empirical studies and training with potential stakeholder. The materials and results will be collected during and after the studies and training courses. No personal and sensitive data will be included in the data assets.

**Why are the data needed?** The data will be used for the design and development of ERATOSTHENES tools. They will also be the reference for the exploitation and dissemination activities of the project results. For external users, the data can be used for analyzing stakeholder involvement in IoT trust management, and for designing cyber-security training courses.

**Data format.** Data will be stored in their original document format, together with generated PDF files for external distribution.
3.2.4 Project management data.

Description. The project management team will reserve important management data such as deliverables, meeting agenda and minutes, presentations, and key decisions.

Why are the data needed? The data will be used as reference for the day-to-day operation of the project.

Data format. Data will be stored as original document files, together with PDF files for external distribution.

3.3 Data Analysis per WP

This chapter details the data summary from the viewpoints of all ERATOSTHENES WPs as far as the following are concerned:

- Purpose of data collection/generation and relation to ERATOSTHENES activities
- Types and formats of data collected/generated
- Information whenever data are being re-used (if re-used)
- Data origin
- Data expected size
- Data utility

WP1 - Methodological IoT Trust and Identity Framework

WP1 is focused on the elaboration and structuring of requirements, use cases, validation KPIs and the architecture design. This will set the baselines for technical developments and their validation within the project’s pilots. Its main objectives are the verification and extension of use cases to derive new models and related KPIs, to analyse existing Trust and IdM frameworks and platforms from different application domains, identify social and legal constraints, and to conduct user studies with the final aim of setting the basis for the technical developments within the project and their evaluation. Data involved in the work package will be:

Task 1.1: Research Agenda, IoT threat landscape and security challenges

- Purpose of data collection/generation and relation to ERATOSTHENES activities: Monitoring the state of the art of security and privacy to adjust the work agenda.
- Types and formats of data collected/generated:
  - Threats related data: Identified risks/threats that need to be managed. Format = docx or another common format.
  - Opportunity related data: Data derived from analyzing the state of the art that represent an opportunity for the project. Format = docx or another common format.
- Information whenever data are being re-used (if re-used): can be re-used.
- Data origin: State of the art and created by project partners during the project.
- Data expected size: <50 MB
- Data utility: support for internal R&D activities.

Task 1.2: User Driven Requirements and Pilot Scenarios definition

- Purpose of data collection/generation and relation to ERATOSTHENES activities: User input about potential requirements of the pilots
- Types and formats of data collected/generated:
  - Documents summarizing feedback
- Information whenever data are being re-used (if re-used): no re-use outside project
- Data origin: Potential end-users contacted by pilot partners
- Data expected size: ~1MB
- Data utility: Starting point for requirements and pilot definition

Task 1.3: Methodological Trust Framework Definition
• Purpose of data collection/generation and relation to ERATOSTHENES activities: Identify the technical requirements of the use-cases.
• Types and formats of data collected/generated:
  o Use case Requirements. Format = docx or another common format.
  o Key Performances Indicators (KPI). Format = docx or another common format.
• Information whenever data are being re-used (if re-used): no re-use outside project
• Data origin: Generated by partners during the project.
• Data expected size: <5MB
• Data utility: Use cases development and evaluation during the project.

Task 1.4: Reference Trust and Identity Management Architecture
• Purpose of data collection/generation and relation to ERATOSTHENES activities: Potential assets for the architecture
• Types and formats of data collected/generated:
  o Documents describing assets
• Information whenever data are being re-used (if re-used): no re-use outside project
• Data origin: Generated by partners
• Data expected size: ~5MB
• Data utility: Information about assets as starting point for architecture discussion

WP2 - Dynamic Trust Management and Agents
WP2 activities will focus on dynamic trust management and trust agents' development. The WP aims to develop the trust broker mechanisms to feed the TMB with contextual attributes for trust evaluation, establish a threat modelling approach that enables the dynamic and reflective analysis of end-to-end system threats, enable the automated deployment of trust agents on IoT devices, develop a trusted execution environment (TEE) on IoT/Edge Devices and develop an automated recovery process of Trust Manger and the IoT Devices Network Enrolment.

Across all tasks, this WP will generate project management data in the form of formal reports (e.g., task deliverables, meeting minutes, presentations). These formal reports will be in Microsoft (MS) office (or related) format (i.e., .doc/.docx, .pdf etc) and will be considered as internal or external depending on their nature and character as defined in the ERATOSTHENES Grant Agreement (see public, confidential, restricted reports). The formal reports will be stored in the ERATOSTHENES Web-Space (Teamwork).

In addition, the tasks will generate IoT and trust related data, including source code and technical documentation (hardware/software design documents):

Task 2.1 Trust Broker Mechanism:
• Purpose of data generation and relation to ERATOSTHENES activities: design & development of the ERATOSTHENES trust broker mechanism.
• Types and formats of data generated:
  o IoT-related data (architecture and design of the IoT system, metadata and identity information of IoT devices, runtime monitoring and feedbacks from the IoT systems). Format = Eclipse or other common modelling format.
  o Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices). Format = JSON, csv or another common format.
• Information whenever data are being re-used (if re-used): no re-use.
• Data origin: created by project partners during the project.
• Data expected size: <1GB
• Data utility: support for internal R&D activities.

Task 2.2 Threat Modelling for Continuous Risk Assessment
• Purpose of data generation and relation to ERATOSTHENES activities: security threat modelling to identify and prioritize critical security protection measures in the ERATOSTHENES trust broker.
• **Types and formats of data generated:**
  - IoT-related data (architecture and design of the IoT system, metadata and identity information of IoT devices, runtime monitoring and feedbacks from the IoT systems). Format = Eclipse or other common modelling format.
  - Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices). Format = JSON, csv or another common format.

• **Information whenever data are being re-used (if re-used):** no re-use.

• **Data origin:** created by project partners during the project.

• **Data expected size:** <1GB

• **Data utility:** support for internal R&D activities.

### Task 2.3 Automated Deployment of Trust Agents

**Purpose of data generation and relation to ERATOSTHENES activities:** design & development of automated deployment mechanism for the ERATOSTHENES trust agents.

**Types and formats of data generated:**
- IoT-related data (architecture and design of the IoT system, metadata and identity information of IoT devices, runtime monitoring and feedbacks from the IoT systems). Format = Eclipse or other common modelling format.
- Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices). Format = JSON, csv or another common format.

**Information whenever data are being re-used (if re-used):** no re-use.

• **Data origin:** created by project partners during the project.

• **Data expected size:** <1GB

• **Data utility:** support for internal R&D activities.

### Task 2.4 Trusted Execution on IoT/Edge Devices

**Purpose of data generation and relation to ERATOSTHENES activities:** design & development of trusted execution environment on resource-constraint IoT/Edge devices.

**Types and formats of data generated:**
- IoT-related data (architecture and design of the IoT system, metadata and identity information of IoT devices, runtime monitoring and feedbacks from the IoT systems). Format = Eclipse or other common modelling format.
- Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices). Format = JSON, csv or another common format.

**Information whenever data are being re-used (if re-used):** no re-use.

• **Data origin:** created by project partners during the project.

• **Data expected size:** <1GB

• **Data utility:** support for internal R&D activities.

### Task 2.5 Automated recovery process of trust agent

**Purpose of data generation and relation to ERATOSTHENES activities:** design & development of automated recovery process for trust agents.

**Types and formats of data generated:**
- IoT-related data (architecture and design of the IoT system, metadata and identity information of IoT devices, runtime monitoring and feedbacks from the IoT systems). Format = Eclipse or other common modelling format.
- Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices). Format = JSON, csv or another common format.

**Information whenever data are being re-used (if re-used):** no re-use.

• **Data origin:** created by project partners during the project.

• **Data expected size:** <1GB
• **Data utility**: support for internal R&D activities.

**Task 2.6 IoT Devices Network Enrolment Mechanism**

- **Purpose of data generation and relation to ERATOSTHENES activities**: design & development of network enrolment mechanism.
- **Types and formats of data generated**:
  - IoT-related data (architecture and design of the IoT system, metadata and identity information of IoT devices, runtime monitoring and feedbacks from the IoT systems). Format = Eclipse or other common modelling format.
  - Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices). Format = JSON, csv or another common format.
- **Information whenever data are being re-used (if re-used)**: no re-use.
- **Data origin**: created by project partners during the project.
- **Data expected size**: <1GB
- **Data utility**: support for internal R&D activities.

**Task 2.7: Interoperability layer with legacy infrastructure**

- **Purpose of data collection/generation and relation to ERATOSTHENES activities**: design & development of interoperability layer to integrate the ERATOSTHENES trust broker prototypes with legacy infrastructure.
- **Types and formats of data collected/generated**:
  - IoT-related data (architecture and design of the IoT system, metadata and identity information of IoT devices, runtime monitoring and feedbacks from the IoT systems). Format = Eclipse or other common modelling format.
  - Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices). Format = JSON, csv or another common format.
- **Information whenever data are being re-used (if re-used)**: no re-use.
- **Data origin**: created by project partners during the project.
- **Data expected size**: <1GB
- **Data utility**: support for internal R&D activities.

**WP3 - Decentralized Identity Management**

This Work Package is devoted to providing a horizontal Identity Management solution using distributed ledger technologies following a Self-Sovereign Identity approach. This Identity Management solution needs to cover the end user identification but also allows other stakeholders (service providers, edge components, IoT devices, etc.) to be integrated into the ERATOSTHENES project. Although the data used in this WP can be personal and even sensitive data, the fact that it is using a Self-Sovereign Identity approach, empowering the control of the end users over their own data will help in the protection of the security and privacy issues associated with the process of this data thanks to the use of cryptographic techniques and an application of decentralised ledger technologies.

**Task 3.1: Context-aware identity and access management**

- **Purpose of data collection/generation and relation to ERATOSTHENES activities**: Metadata associated to the communication (location, operative system, etc), that can help on the identification of the end user
- **Types and formats of data collected/generated**: JSON
- **Information whenever data are being re-used (if re-used)**: no re-use.
- **Data origin**: created by project partners during the project.
- **Data expected size**: N/A
- **Data utility**: support for internal R&D activities.
**Task 3.2: Self-Sovereign Identity (SSI) management models and mechanisms**

- **Purpose of data collection/generation and relation to ERATOSTHENES activities:** Identity information
- **Types and formats of data collected/generated:** Verifiable credentials and decentralized identifiers and cryptographic material associated
- **Information whenever data are being re-used (if re-used):** no re-use.
- **Data origin:** created by project partners during the project.
- **Data expected size:** N/A
- **Data utility:** support for internal R&D activities.

**Task 3.3: Distributed ledger-based and privacy-preserving IoT Identity Management**

- **Purpose of data collection/generation and relation to ERATOSTHENES activities:** Devices' identity information
- **Types and formats of data collected/generated:** Collected from IoT devices used in the experiments and pilots
- **Information whenever data are being re-used (if re-used):** no re-use.
- **Data origin:** created by project partners during the project.
- **Data expected size:** 1MB each
- **Data utility:** support for internal R&D activities.

**WP4 - Lifecycle Management, Inter-ledger Information Sharing and Recovery**

WP4 activities are focusing on the Lifecycle Management, Inter-ledger Cyber-Threat Information Sharing and Recovery Solutions. Its objectives are related to the design and develop the DLT-based Trust Framework enforcement and recovery framework for secure sharing of cyber-threat information in IoT networks., the trust network Smart Contracts implementation for storing and sharing trust scores in the IoT network, the inter-ledger platform for Cyber-threat information sharing to support the exchange of trust information and relevant security information among the stakeholders, the Federated threat analysis models for continuous assessment as the defence service for detecting different attacks and exploits inside IoT networks, the Intrusion detection capability for IoT-based context and networks and the Lifecycle Security of IoT Devices framework.

**Task 4.1: DLT-based Trust Framework enforcement and recovery system**

- **Purpose of data collection/generation and relation to ERATOSTHENES activities:** The storage of trust scores on a distributed ledger will enable the implementation of a recovery system that will be based on these values to reset the network in a stable state after an attack.
- **Types and formats of data collected/generated:**
  - Trust scores by devices, Cyber-threat information. Format: blockchain record.
- **Information whenever data are being re-used (if re-used):** can be re-used by trust managers and trust framework.
- **Data origin:** Collected from cloud services of manufacturers and vendors
- **Data expected size:** ~1MB
- **Data utility:** support for internal R&D activities.

**Task 4.2: Trust network Smart Contracts implementation in IoT context**

- **Purpose of data collection/generation and relation to ERATOSTHENES activities:** The development of smart contracts will allow the storage, management and query of the trust scores on a distributed ledger.
- **Types and formats of data collected/generated:**
  - Trust scores of devices. Format: blockchain record.
- Information whenever data are being re-used (if re-used): can be re-used
- Data origin: created by project partners during the project; Calculated in T2.1 and pushed to the ledger.
- Data expected size: ~1MB
- Data utility: support for internal R&D activities.

### Task 4.3: Inter-ledger platform for Cyber-threat information sharing

- **Purpose of data collection/generation and relation to ERATOSTHENES activities:** Generation, collection, anonymization and sharing of CTI data to improve security and cyber-response capabilities throughout the whole the infrastructure.
- **Types and formats of data collected/generated:**
  - Indicators of Compromise (IoC) for sharing
  - Cyber-threat information
  - Format: csv; JSON; records
- **Information whenever data are being re-used (if re-used):** can be re-used
- **Data origin:** created by project partners during the project. Forensic evidence of potential intrusions. Collected from cloud services of manufacturers and vendors.
- **Data expected size:** ~1MB each
- **Data utility:** support for internal R&D activities.

### Task 4.4: Federated threat analysis models for continuous assessment

- **Purpose of data collection/generation and relation to ERATOSTHENES activities:** design and development of Machine Learning models to perform security monitoring and anomaly detection for IoT devices and networks, and which can be executed on edge computing devices.
- **Types and formats of data collected/generated:**
  - Publicly available datasets containing network information and labelled with attacks. Format: CSV, PCAP.
- **Information whenever data are being re-used (if re-used):** initially, the development will rely on publicly available datasets and therefore they can be re-used.
- **Data origin:** publicly available and created by project partners during the project.
- **Data expected size:** ~GBs.
- **Data utility:** support for internal R&D activities.

### Task 4.5: Intrusion detection for IoT-based context and networks

- **Purpose of data collection/generation and relation to ERATOSTHENES activities:** collect public dataset to test the solution, collect network traffic data or event-related info into the internal database during validation/demonstration, data generated are reports and alerts
- **Types and formats of data collected/generated:** datasets format - pcap, csv; reports/alerts formats – syslogs, json.
- **Information whenever data are being re-used (if re-used):** only publicly available datasets can be re-used
- **Data origin:** created by project partners during the project. Forensic evidence of potential intrusions
- **Data expected size:** ~GB
- **Data utility:** support for internal R&D activities.

### Task 4.6: Lifecycle Security of IoT Devices

**Purpose of data collection/generation and relation to ERATOSTHENES activities:** Manufacturer Usage Description for IoT devices used in the experiments and pilots, as well as potential threats or expected behaviors

- **Types and formats of data collected/generated:**
- MUD profiles: Trust-related data (trust scores and reputations of devices, the potential cyber-threats, the stored public keys for relevant devices).
- Format: Specific data formats, JSON
  - **Information whenever data are being re-used (if re-used):** will be re-used
  - **Data origin:** created by project partners during the project.
  - **Data expected size:** ~1MB each
  - **Data utility:** support for internal R&D activities.

### WP5 - Framework integration, real-world pilots and cybersecurity exercises

#### The Pilot 1 – Automotive

In Pilot 1 there will be different interactions between different actors. In the first use case the main interaction will be between the car and the infrastructure being hampered because of an attacker. In this case will be collected a log with the network activity to provide feedback on how the communication is carried out between the vehicle and the infrastructure and how the attacker performs his malicious activity and how the other elements (included the victim) react to this situation. Also, will be provided an internal log of the victim which will allow to understand how it managed internally the whole situation.

In the second use case, the main type of output data will be very similar as it is in the first use case but in another context. In this situation the network information log will provide the interaction between the vehicle and the server which will perform a software update, and also will be found the attack from another actor. As it appears in the other use case, in the second one there will be an internal log of the victim which will allow to understand how to internally manage the attack and the software update process.

The main output of the whole Pilot 1 task will be formed by reports to evaluate the entire use case which will be used to build the deliverables.

#### The Pilot 2 – Personalized Healthcare

The remote patient monitoring production system contains person sensitive data, in particular, personal health data. Thus, TELLU is fully compliant with GDPR and other local and international legislation and laws. However, in this project test data are sufficient in order to evaluate and validate project results. Thus, in the context of this project there will be no personal or sensitive data involved. Nevertheless, all relevant requirements related to GDPR and other relevant legislations and laws will be put forward and evaluated for the technology providers in ERATOSTHENES. This may include full control of geographical location of data in terms of storage and processing, requirements to cope with security mechanisms such as data encryption, authorization and access control and mechanisms such as anonymisation/pseudonymization, etc. In general, the workflow and data used in the health use case is as follow (again it is pointed out that in the project only test data will be applied):

- **Collection:** Collecting measure and data from welfare sensors and medical devices, as well as third party systems and services (e.g., hospital systems and external welfare systems).
- **Usage:** Data are used for remote supervision of elderly and to follow up patients with chronical diseases such as chronic obstructive pulmonary disease, diabetes and kidney failure. In the future, it is expected that data will be further exploited for supporting diagnosis and predict events such as risk of a patient falling.
- **Sharing:** A sample of test data that represent typical real cases will be shared in ERATOSTHENES to experiment and validate the technologies provided.
- **Processing:** The data will be processed across the IoT, Edge and Cloud space. Important aspects include security, privacy and trustworthiness throughout the design, deployment and operation phases and to support patients in their daily life and care providers in providing high quality care.
- **Lifespan:** The test data sample will be available to the end of the project.

#### The Pilot 3 – Industry 4.0

The IoT technology is the core of the industry 4.0 revolution. Nowadays, at an industrial floor hundreds of connected devices are used to monitor, optimize and control various processes. It has introduced cybersecurity risks in our critical
industries. The pilot aims to increase the security by design in industrial IoT network and communication by introducing novel approaches on IoT Asset Identification and the use of disposable IDs to identify trustworthy entities in communication networks. The IoT asset management in industries are still done using legacy technologies with static identities and centralized identity generation and management. The objectives are to generate secure unique identities which are disposable and distributed. Consequently, IoT devices can be securely bootstrapped and become part of critical industrial network.

The pilot will produce device and firmware metadata. It will include firmware version, build data, device mac address, IP address, sensor data etc. The pilot will use data generated by other work packages. It will involve secure device identities, blockchain records, PUF based keys etc.

WP6 - Dissemination, Exploitation, and Standardization Activities

WP6 will ensure scale up through wide dissemination, exploitation actions and Capacity Building aiming at infrastructure sustainability, organisational development and human capital development through training on the practical use of the ERATOSTHENES Concepts, Capabilities and Offering. The communication Programme will incorporate innovative methods for sharing information with the public fostering better alignment between policy for safety and security and the resulting governance through which it is enacted, as resides within multiple individual sectors and domains. An exploitation plan for all project partners will be produced throughout the project cycle, which will constitute a live document, and will persist after the project ends.

This WP will not produce any technical components but will produce dissemination documents and will use data and documents generated from the other WPs. The data generated in the project can be grouped into source code, technical documentation and formal reports. Deliverables and other formal internal reports are another type of data that this WP will be creating and managing/controlling. These files will be MS office (or related) documents (i.e., .doc/.docx, .pdf etc) and will be considered as internal or external depending on their nature and character as defined in the ERATOSTHENES Grant Agreement (see public, confidential, restricted reports). The formal reports and deliverables will be stored in the ERATOSTHENES Web-Space (Teamwork). Additional data will be created in discussion forum, in this case the forum itself will maintain this type of unstructured information, that will be mainly generated from stakeholder open discussions (i.e., ERATOSTHENES user group forum, twitter, LinkedIn, etc.).

WP7 - Project Management

As the Project Management WP7, WP7 is responsible for the technical and administrative coordination of the project, including quality and ethical activities. These tasks include the creation and processing of various types of documents and files to ensure efficient and effective management of ERATOSTHENES. These types of data mainly consist of documents and spreadsheets or presentation files. This mainly includes documents and spreadsheet files for collection and progress/regular reporting (internal and/or to EC) by ERATOSTHENES. These are mainly MS Office documents (.doc/.docx, .xls/.xlsx, .pdf, etc.) that are distributed internally to the consortium or sent to the EC as the final version.

At the same time, WP7 is responsible for managing and processing meeting-related documents such as meeting minutes, and presentations. These files can also be MS Office (or similar) related documents (.doc/.docx, .xls/.xlsx, .pdf, .ppt/.pptx, etc.), all partners are expected to have access to them, but they are always considered internal documents of the ERATOSTHENES Consortium (not distributed outside of ERATOSTHENES).

Deliverables and different internal reviews are other form of reports that this WP may be developing and managing/controlling. These documents may be mostly MS office (or related) documents (i.e., .doc/.docx, .pdf etc) and may be taken into consideration as internal or external relying on their nature described within the ERATOSTHENES Grant Agreement (see public, confidential, confined reviews).

Patenting files will also be considered as internal documents so they should fall into the above category and type of data. Other files that will be created in the framework of WP7 consist of quality management and templates for all the above (and possibly more) purposes. These files will be most of the time consist of .doc, .docx, .pdf, .xls, .xlsx, .ppt, .pptx files created by the quality manager and used and shared by the ERATOSTHENES partners.
All related files in WP7 (Project Management) will be stored in the ERATOSTHENES web-space (TEAMWORK) where all partners have personalised login details and therefore we consider that access if fully controlled and safe. Links for exchanging these files internally will be circulated via email or the ERATOSTHENES Teamwork server itself.
3.4 Initial data assets

Table 6 list the potential data assets that we have identified in the beginning of the project. This list will be updated at later stages of the project and reported in the next version of the Data Management Plan – Final (D7.9, M40, January 2025). The data below have been structured per WP and task and are referenced with IDs (1st column). Apart from the actual definition of the data sets, we provide information regarding their metadata and followed standards (what metadata will accompany the data and if there are any standards used), Re-use prospects (potential of data re-use in other domains, sectors and research), diffusion level (internal/external to the consortium) and presence of PII (presence of personal identifiable information).

Table 6: Initial data asset identified at the beginning of the project

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>01.01</td>
<td>Technical specifications</td>
<td>Asset descriptions provided by the project partners</td>
<td>WP1</td>
<td>T1.1</td>
<td>KUL</td>
<td>XLSX</td>
<td>TBD</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Project-Internal</td>
<td>INT</td>
</tr>
<tr>
<td>01.02</td>
<td>Expert Interviews</td>
<td>Records or questionnaires</td>
<td>WP1</td>
<td>T1.2</td>
<td>INLE</td>
<td>Docx</td>
<td>N/A</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Requirements</td>
<td>INT</td>
</tr>
<tr>
<td>01.03</td>
<td>User Requirements</td>
<td>Text</td>
<td>WP1</td>
<td>T1.2</td>
<td>INLE</td>
<td>Xls</td>
<td>N/A</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Requirements</td>
<td>INT</td>
</tr>
<tr>
<td>02.01</td>
<td>Feedback and recommendation from other devices</td>
<td>Devices that used services from other devices evaluate these services and provide feedback and recommendations for them. The TMB may collect these data from the ledger.</td>
<td>WP2</td>
<td>T2.1</td>
<td>UPRC</td>
<td>Blockchain record</td>
<td>~1MB</td>
<td>JSON</td>
<td>Y</td>
<td>To be reused by developers and researchers on risk assessment</td>
<td>INT</td>
<td>N</td>
</tr>
<tr>
<td>02.02</td>
<td>Experiment results of automatic deployment and recovery</td>
<td>Deployment models; logs of deployment and recovery processes on different type of devices.</td>
<td>WP2</td>
<td>T2.3, T2.5</td>
<td>SINTE, T2.5, F</td>
<td>CSV or JSON</td>
<td>1MB each</td>
<td>DCAT</td>
<td>Y</td>
<td>Reference for researchers or developers to design recovery mechanisms</td>
<td>Partially EXT</td>
<td>N</td>
</tr>
<tr>
<td>02.03</td>
<td>Technical specifications</td>
<td>System architectures per use case or unstructured application case descriptions</td>
<td>WP2</td>
<td>T2.2</td>
<td>KUL</td>
<td>Docx</td>
<td>TBD</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Case studies</td>
<td>INT/EXT</td>
</tr>
</tbody>
</table>

3 Availability of metadata and followed standards
4 Availability of documentation at the distribution stage of the data. Are they accompanied by a report/documentation?
5 Potential reuse of data in other domains, work and efforts
6 Diffusion level [Internal/External] to the consortium
7 Presence of PII (Personal identifying information)

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Grant Agreement No 101020416.
<table>
<thead>
<tr>
<th>ID</th>
<th>Data</th>
<th>Data description and collection</th>
<th>WP</th>
<th>Task</th>
<th>Owner</th>
<th>Format</th>
<th>Size</th>
<th>Metadata/Standards</th>
<th>Doc</th>
<th>Reuse Prospects</th>
<th>Diffusion</th>
<th>PII</th>
</tr>
</thead>
<tbody>
<tr>
<td>02.04</td>
<td>Values of context-based reputation and network experience parameters</td>
<td>These values outline the parameters that influence the calculation of a device's initial trust score. Collected from end services/users as well as from device manufacturers and vendors.</td>
<td></td>
<td>T2.6</td>
<td>UPRC</td>
<td>Blockchain record</td>
<td>~1MB per device</td>
<td>JSON</td>
<td>Y</td>
<td>To be reused by developers and researchers on risk assessment</td>
<td>INT</td>
<td>N</td>
</tr>
<tr>
<td>02.05</td>
<td>DIDs and their respective public keys.</td>
<td>The address of a ledger account acts as a DID. Moreover, the ledger stores the DID's public key.</td>
<td></td>
<td>T2.6</td>
<td>UPRC</td>
<td>Blockchain record</td>
<td>~1MB per device</td>
<td>JSON</td>
<td>N</td>
<td>Reuse for Identity management and Access Control</td>
<td>INT</td>
<td>N</td>
</tr>
<tr>
<td>03.01</td>
<td>Metadata from device</td>
<td>Metadata associated to the communication (location, operative system, etc), that can help on the identification of the end user.</td>
<td></td>
<td>T3.1</td>
<td>Atos</td>
<td>json</td>
<td>N/A</td>
<td>JSON</td>
<td>N</td>
<td>N</td>
<td>INT</td>
<td>N</td>
</tr>
<tr>
<td>03.02</td>
<td>Identity Information</td>
<td>Verifiable credentials and decentralized identifiers and cryptographic material associated</td>
<td>WP3</td>
<td>T3.2</td>
<td>Atos</td>
<td>json</td>
<td>N/A</td>
<td>DID, VC W3C</td>
<td>N</td>
<td>N</td>
<td>INT</td>
<td>Y</td>
</tr>
<tr>
<td>03.03</td>
<td>Devices' identity information</td>
<td>Collected from IoT devices used in the experiments and pilots</td>
<td></td>
<td>T3.3</td>
<td>UMU</td>
<td>Specific data formats; JSON</td>
<td>1MB each</td>
<td>N/A</td>
<td>Y</td>
<td>IoT device data</td>
<td>INT</td>
<td>N</td>
</tr>
<tr>
<td>04.01</td>
<td>Cyber-threat information</td>
<td>Collected from cloud services of manufacturers and vendors</td>
<td></td>
<td>T4.1</td>
<td>INLE</td>
<td>Blockchain record</td>
<td>~1MB each</td>
<td>N/A</td>
<td>N</td>
<td>IoT device data</td>
<td>INT</td>
<td>N</td>
</tr>
<tr>
<td>04.02</td>
<td>Trusted scores of devices</td>
<td>Calculated in T2.1 and pushed to the ledger</td>
<td></td>
<td>T4.1</td>
<td>INLE</td>
<td>Blockchain record</td>
<td>~1MB each</td>
<td>N/A</td>
<td>N</td>
<td>Ledger applications</td>
<td>INT</td>
<td>N</td>
</tr>
<tr>
<td>04.03</td>
<td>Input data for threat analysis models</td>
<td>Publicly available datasets containing network information for IoT devices and networks and including attacks (e.g., UNSW-NB15, TON_IoT, BoT-IoT). Data from the project’s pilots.</td>
<td>WP4</td>
<td>T4.4</td>
<td>ATOS</td>
<td>CSV, PCAP files.</td>
<td>~GB</td>
<td>N/A</td>
<td>Y</td>
<td>Y</td>
<td>Ext</td>
<td>N</td>
</tr>
<tr>
<td>-----</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>----------------------</td>
<td>------------</td>
<td>--------------------</td>
<td>-----------</td>
<td>-----------------</td>
<td>---------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>04.04</td>
<td>Output data for threats detected during project’s pilots</td>
<td>Results of the threat analysis executed for the three project’s pilots.</td>
<td>T4.4</td>
<td>ATOS</td>
<td>CSV, JSON</td>
<td>1MB each</td>
<td>N/A</td>
<td>N</td>
<td>N</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>04.05</td>
<td>Network traffic</td>
<td>The IDS installed on a particular network reads, logs and analyses network traffic packets.</td>
<td>T4.5</td>
<td>ENG</td>
<td>pcap if offline-detection</td>
<td>N/A</td>
<td>N</td>
<td>N</td>
<td>Network management</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>04.06</td>
<td>ML models or dataset</td>
<td>ML model to update the detection capabilities, or a dataset to train the detection algorithm and create a new model</td>
<td>T4.5</td>
<td>ENG</td>
<td>Pickle or pcap or csv</td>
<td>N/A</td>
<td>N</td>
<td>N</td>
<td>ML applications</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>04.07</td>
<td>Alerts</td>
<td>Notifications to the network administration team of threats</td>
<td>T4.5</td>
<td>ENG</td>
<td>N/A</td>
<td>N/A</td>
<td>N</td>
<td>N</td>
<td>Network management</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>04.08</td>
<td>Indicators of Compromise (IoC) for sharing</td>
<td>Forensic evidence of potential intrusions</td>
<td>T4.3, T4.4, T4.5</td>
<td>UMU</td>
<td>csv; JSON; records</td>
<td>1MB each, ~GBs</td>
<td>JSON</td>
<td>Intrusion detection</td>
<td>INT</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>04.09</td>
<td>MUD profiles</td>
<td>Manufacturer Usage Description collected IoT devices used in the experiments and pilots.</td>
<td>T4.6</td>
<td>UMU</td>
<td>Specific data formats; JSON</td>
<td>1MB each</td>
<td>N/A</td>
<td>N</td>
<td>MUD applications</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>05.01</td>
<td>Logs and results of Pilot 1 activities</td>
<td>Logs from the network activity, internal log of the vehicle and evaluation report.</td>
<td>T5.3</td>
<td>IDIAD</td>
<td>Csv</td>
<td>TBD</td>
<td>CSV</td>
<td>N</td>
<td>Network management</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>05.02</td>
<td>Data structured according to the HL7-FHIR semantic datamodel. Wide spread standard for the eHealth domain</td>
<td>Test data for the eHealth pilot on remote patient monitoring</td>
<td>WP5</td>
<td>TELLU</td>
<td>JSON - FHIR</td>
<td>5 MB</td>
<td>HL7-FHIR</td>
<td>N</td>
<td>Test data</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>05.03</td>
<td>device and firmware metadata</td>
<td>Device MAC, IP, firmware version, build data</td>
<td>T5.5</td>
<td>DWG</td>
<td>csv, json records</td>
<td>2MB each</td>
<td>DCAT</td>
<td>Y</td>
<td>References for internal IoT</td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

8 (We assume that the information on the partner's data servers will be treated confidentially)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>05.04</td>
<td>Training materials, and results of training courses.</td>
<td>Documents used in the training courses, interview records from the course attendees, etc.</td>
<td></td>
<td></td>
<td></td>
<td>Document: PDF; JSON</td>
<td>5-10MB each</td>
<td>N</td>
<td>N/A</td>
<td>Reference for training courses</td>
<td>Partially EXT</td>
<td>N</td>
</tr>
<tr>
<td>05.05</td>
<td>Data about the participants in the exercises?</td>
<td>Profiles of participants, together with their consent for participation</td>
<td></td>
<td></td>
<td></td>
<td>Document: DOC; XLS;</td>
<td>5MB</td>
<td>N</td>
<td>N/A</td>
<td>Internal record</td>
<td>INT</td>
<td>Y</td>
</tr>
<tr>
<td>06.01</td>
<td>Workshops and Events' agendas</td>
<td>Internal documents collected by the document team</td>
<td>WP6</td>
<td>WP6</td>
<td>DBC</td>
<td>document: PDF; DOC; DOCX</td>
<td>1-2 MB each</td>
<td>N</td>
<td>N</td>
<td></td>
<td>INT</td>
<td>Y</td>
</tr>
<tr>
<td>07.01</td>
<td>Meeting Presentations, minutes and internal reports</td>
<td>Internal documents collected by the document team</td>
<td>WP7</td>
<td>WP7 (all WPs)</td>
<td>INLE (all)</td>
<td>document: PDF; DOC; DOCX</td>
<td>1-2 MB each</td>
<td>N</td>
<td>N</td>
<td></td>
<td>INT</td>
<td>Y</td>
</tr>
<tr>
<td>08.01</td>
<td>Deliverables</td>
<td>Documents that will be submitted to the EC</td>
<td>All WPs</td>
<td>Main authors</td>
<td>document: PDF; DOC; DOCX</td>
<td>1-20MB each</td>
<td>N</td>
<td>N</td>
<td></td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>08.02</td>
<td>Structured feedback, recomendations from workshops, AB etc.</td>
<td>Records or questionnaires</td>
<td>All WPs</td>
<td>Respective author</td>
<td>document: PDF; DOC; DOCX</td>
<td>1-20MB each</td>
<td>N</td>
<td>N</td>
<td></td>
<td>INT</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>
3.5 DMP in ERATOSTHENES WPs and WP leaders’ responsibilities and Allocation of resources

Data management in ERATOSTHENES belongs to Task 7.4 Quality Assurance, Data Management and Risks Management [M1-M42] (SINTEF). This includes data management life cycle monitoring for all datasets to be collected, processed or generated by the project. This task is responsible for the Data Management Plan (DMP, D7.3 and D7.9) including and respecting GDPR policies and procedures for personal/sensitive information protection. The plan will cover the rules of handling research data during and after the project, including the characterization of the data that will be collected, processed or generated. Special attention will be paid to Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and repealing Directive 95/46/EC (GDPR).

In order to ensure compliance with all the previously described data management decisions as they relate to the DMP, the following overall ERATOSTHENES measures will apply in ERATOSTHENES:

- SINTEF is the leader of the task on Quality Assurance, Data Management and Risks Management
- WP leaders will be responsible for adhering to the specifications above in their respective work packages.
- The project manager of each organization will be responsible for the DMP actions and will be accessible by the partner team in case of issues related to DMP.
- Data Owners (see section 4.2.2) have the ultimate responsibility of complying with the specifics of the ERATOSTHENES Data Management plan, as well as to the related GDPR policies.
- For the overall ERATOSTHENES project management activities, INLE will be responsible for complying with the data management plan.
- The project manager and main contact from each and every partner should ensure that personnel working on the project have read the data management plan and apply/exercise all the principles as described in the ERATOSTHENES DMP (this document).
4 General Data Protection Regulation (GDPR)

This chapter summarizes the GDPR compliancy of ERATOSTHENES, as the GDPR was formally introduced in May 2018, and has been applicable in all Member States in the European Union, as well as in the countries in the European Economic Area (EEA).

4.1 GDPR Compliancy

Data confidentiality is an overriding concern throughout the ERATOSTHENES project and beyond, as the solution to be developed in ERATOSTHENES will continue to be used afterwards, to this end ERATOSTHENES aims to be fully compliant with the GDPR. All data to be collected from stakeholders in the project will be done in accordance with applicable ethical standards and requirements in the respective countries of the data collection, as well will be processed and handled in a secure way and in line with applicable rules and regulations on privacy and data protection.

<table>
<thead>
<tr>
<th>Personal Data Description</th>
<th>Access</th>
<th>Storage</th>
<th>Purpose</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>XLS list of ERATOSTHENES contacts</td>
<td>Internal to ERATOSTHENES (project partners only)</td>
<td>Teamwork (folder: contacts)</td>
<td>ERATOSTHENES internal communications</td>
<td>31/03/2025</td>
</tr>
<tr>
<td>Meeting related material (agendas, presentations, signature lists, minutes)</td>
<td>Internal to ERATOSTHENES (project partners only)</td>
<td>Teamwork (folder: meetings)</td>
<td>ERATOSTHENES meetings’ related</td>
<td>31/03/2025</td>
</tr>
<tr>
<td>Workshops/Conferences and Training sessions</td>
<td>Internal and external to ERATOSTHENES</td>
<td>Teamwork (folder: meetings), LinkedIn, twitter, ERATOSTHENES website</td>
<td>Large event dissemination</td>
<td>31/03/2025</td>
</tr>
<tr>
<td>Reporting (C forms)</td>
<td>Internal to ERATOSTHENES (project partners only)</td>
<td>INLECOM server</td>
<td>ERATOSTHENES reporting and consolidation of financial reports</td>
<td>5 years after the project end (for the case of audit)</td>
</tr>
<tr>
<td>Deliverables, internal documents and other ERATOSTHENES reports</td>
<td>Depending on deliverable type could be public or consortium restricted</td>
<td>Teamwork (folder: deliverables)</td>
<td>ERATOSTHENES documents and deliverables</td>
<td>31/03/2025</td>
</tr>
<tr>
<td>Publications</td>
<td>Internal and external to ERATOSTHENES</td>
<td>Teamwork (folder: publications)</td>
<td>Dissemination and publication of research results</td>
<td>Internal: 31/03/2025 External: Depending on publisher</td>
</tr>
<tr>
<td>List of stakeholders (external to ERATOSTHENES)</td>
<td>Internal and external to ERATOSTHENES</td>
<td>Teamwork (folder: contacts)</td>
<td>ERATOSTHENES mass-dissemination, list of potential users, exploitation</td>
<td>31/03/2025</td>
</tr>
</tbody>
</table>

---

9 Overall data description.
10 Determines who has access to the particular data (internal, external to consortium).
11 Storage places of actual data.
12 Intended purpose of data and reasons for keeping.
13 Duration of stored data (until when they will be kept).
4.2 General Data Protection Policy

4.2.1 Introduction

This General Data Protection Policy (the “Policy”) is drafted by Inlecom Ltd. (the “Project Coordinator”) with regard to the EU H2020 Project ERATOSTHENES Contract Number 780075 (the “Project”) executed by the list of partners included therein (the “Project Partners”) in order to:

- Comply with the policy and legal requirements of the EU General Data Protection Regulation (Regulation EU 2016/679, the “GDPR”)
- Comply with all other applicable national and EU regulations and guidelines on personal data processing;
- Comply with applicable regulations and best practices with regard to research projects within the EU H2020 Research Programme;
- Raise awareness and improve knowledge among the Project Coordinator, the Project Partners, as well as their employees and/or agents and/or contractors (collectively, the “Policy Recipients”).

Because the field of data protection is a dynamic legal field of constant change, new developments, in the form of new regulations, official reports and/or guidelines, are issued by EU and national legislators, as well as, competent national authorities at a constant pace. In this context, this Policy may need to be periodically updated by the Project Coordinator, in order to remain relevant to legislative change. Accordingly, Policy Recipients will be duly informed, and will be asked to provide their renewed consent upon any such updates.

While every effort has been undertaken by the Company to compile a comprehensive, accurate, relevant and lawful Policy, it is expressly clarified that this Policy does not constitute legal advice neither does it warrant compliance to any applicable laws or regulations. This Policy makes no warranties, express or other, on lawfulness, completeness, fitness for a purpose, or merchantability. Recipients and addressees of this Policy are advised to engage legal counsel prior to applying this Policy for their own aims and purposes.

4.2.2 Definitions

For the purposes of this Policy the GDPR definitions, as set in Article 4, apply. In addition,

“Personal data” means any information relating to an identified or identifiable natural person that is processed by any Project Partner and Policy Recipient during execution of the Project.

“Controller” means the owner of the personal data (usually the creator of the data itself), unless otherwise expressly clarified in this Policy or elsewhere in Project deliverables and/or reports.

“Processor” means each Project Partner, unless otherwise expressly clarified in this Policy or elsewhere in Project deliverables and/or reports.

“Consent” of the data subject means any freely given, specific, informed, unambiguous and in writing indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.

“Supervisory authority” means the competent Data Protection Authorities within the Project Partners’ jurisdictions.

Aim of the above definitions is to particularise and complement the definitions of Article 4 of the GDPR. Policy Recipients are advised to consult both texts in order to formulate the applicable definitions each time.

4.2.3 Policy Scope

The Controller determines in advance what is the law applicable to the processing of personal data in a particular case, considering that according to EU law such determination comes from legal principles and cannot be derogated by the parties.

---

14 EU GDPR Website, GDPR portal: https://www.eugdpr.org/.
4.2.3.1 Establishment

Each Project Partner is established on the territory of EU Member States. In the event of any change in establishment, the respective Project Partner shall notify the Project Coordinator duly and in writing.

Unless otherwise expressly specified, each Project Partner is considered the controller in that Member State.

4.2.3.2 Processor outside the EU

In the event of any subcontracting to an organization not established on EU territory (such as subsidiaries pertaining to the same corporate group) that processes personal data of people staying on EU territory, on behalf of a Project Partner, that organization qualifies as Processor and ensures the fulfilment of the obligations imposed by the GDPR for that specific part of processing.

4.2.4 Personal data processing

4.2.4.1 Personal data

Personal data means any information relating to natural persons, that is or can be identified, even indirectly, by reference to any other information including a personal identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

➢ Special categories of data

Special categories of personal data include data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, data concerning health or data concerning a natural person's sex life or sexual orientation as well as the processing of genetic data and biometric data for the purpose of uniquely identifying an individual.

In the event of such processing the Controller and/or Processor respectively comply with specific rules related to the processing of such data of special categories, as collecting specific informed consent from data subject and applying stricter safeguards.

When the Controller and/or Processor relies on data subject's consent as a legal ground for processing special categories of data, it will meet all legal consent requirements; otherwise, they are only processed if and to the extent it is based on one of the legal grounds listed in the GDPR for the processing of such data.

➢ Data anonymisation

Whenever possible, including non-detrimental to Project execution purposes, Controller and Project Partners shall undertake efforts to keep personal data processed by them for Project purposes anonymous or pseudonymous.

According to the GDPR, "anonymous information" is information which does not relate to an identified or identifiable natural person, or personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. In this context, the GDPR does not apply to the processing of such anonymous information, including for statistical or research purposes.

Similarly, "pseudonymisation" means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

➢ Newsletters, social media and other dissemination material

Unless otherwise expressly specified in Project contract, Controller shall be responsible for the personal data processing carried out for Project dissemination purposes. To this end, Controller shall:

- Collect and keep all relevant personal data (including lists of contact details), or copies thereof;
- Monitor relevant communications;
• Address to Project Partners instructions and guidelines on Project dissemination activities (including any EU or other state guidelines, whenever available);
• Inform Project Partners of any policy or legal requirements reviews and changes.

4.2.4.2 Data subjects

➢ Minors
Processing of children’s personal data requires a special legitimate basis. In the event of such processing the Controller shall be informed in advance and in writing by Project Partners.

4.2.4.3 Data processing
Data processing means any operation, or set of operations, carried out with or without the help of electronic or automated means, concerning the collection, recording, organization, keeping, interrogation, elaboration, modification, selection, retrieval, comparison, utilization, interconnection, blocking, communication, dissemination, erasure and destruction of data whether the latter are contained or not in data bank.

4.2.4.4 Principles for legitimate processing
European Union data protection law set forth the following specific principles which have to be complied with for the processing to be legitimate.

Pertinence and necessity - The Controller should implement management practices to fulfil the obligation to collect only relevant and necessary data for a specified purpose.

Purpose limitation - Personal data is collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. The Controller has a clear overview of all purposes for which personal data is processed. Personal data is not processed for purposes besides the original purposes, unless the (secondary) use is compatible.

Data minimization - Personal data collected by the Controller must be adequate, relevant and limited to what is necessary in relation to the purposes for which they are collected and further processed; if the same purposes can be realized in a less data intensive way a preference is given to that method.

Data update - Personal data is accurate, and, where necessary, kept up to date. Every reasonable step is taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.

Data retention - Personal data is kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed. The Controller and/or Processing concerned should have processes and policies in place to:

1) determine what the applicable (minimum and maximum) retention periods are for the personal data that is being processed;
2) ensure that relevant retention periods are monitored.

4.2.5 Data protection legal roles

4.2.5.1 Controller
By determining the purposes and means of the processing of personal data, unless otherwise expressly specified in this Policy, the Controller is considered by law as the “Controller” and it is the primary target of the provisions of the law.

➢ Identification
The data controller previously identifies itself as such and ensures an effective implementation of data protection measures in order to comply with the principle that personal data are processed fairly and lawfully. The legal role of controller implies specific responsibilities because provisions setting conditions for lawful processing are essentially addressed to the controller.

➢ **Accountability**

The GDPR provides full accountability of the company/controller regarding the compliance of its processing of personal data with the law. To ensure the effectiveness of that obligation, it prompts the Controller to follow an overall approach, achieving a genuine system of control and management of its pertinent information. So, accountability and compliance system are elements of the framework for the protection of personal data, in the cause / effect relationship: to be compliant and able to prove it (accountability), the Controller needs to put in place a comprehensive compliance system.

➢ **Data protection by design**

The Controller considers data protection issues from the outset and from the design of the Project, within the whole lifecycle of processing, in order to manage the issues in a proactive way, to reduce costs and improve efficiency.

➢ **Data protection by default**

The Controller standardizes data protection principles in personal data processing, products and services. The measures adopted ensure that

- personal data is processed for purposes not different from the original purposes,
- only data necessary for these purposes are collected, and
- data are not disclosed without human intervention.

### 4.2.5.2 Joint controller

In the event that at any time during Project execution the Controller processes personal data in conjunction with a third party, by jointly determining the purposes and means of the processing, they both act as joint controller. Both joint controllers determine the mutual responsibilities with a specific arrangement.

### 4.2.5.3 Processor

Unless otherwise specified expressly in this Policy, all Project Partners act as Processors during Project execution.

A processor processes personal data on behalf of the Controller – that is, the Controller delegates all or part of the processing activities to them. In such event the Project contract assumes the role of the relevant required written agreement as per GDPR requirements.

The processor warrants that it shall provide sufficient guarantees to ensure compliance with the GDPR, has implemented appropriate controls to meet data protection requirements defined by the agreement, instructions and/or legal requirements and ensures the protection of the rights of data subjects.

➢ **Auditing**

The Controller ensures the commitment of the Processor(s) to enable and contribute to any review activities, including inspections, carried out by the Controller or other (EU authorities’) auditors and/or reviewers, as appropriate.

➢ **Security**

Each Project Partner undertakes that it adopts appropriate security measures to ensure the security, integrity and confidentiality of personal information and electronic communications at an adequate level with regard to Project purposes, and at any event at no lower lever than processing of similar data within its own organisation.
4.2.5.4 DPO

Whenever required, following applicable GDPR and Member State respective legal requirements, the Controller and each Processor, may designate a data protection officer ("DPO") for assistance in monitoring internal compliance with the GDPR.

➢ Identification

Each Processor appoints a DPO in accordance with the criteria and the requirements set forth in the GDPR, as applicable to it. In such event, it shall notify the Controller in writing accordingly.

➢ Designation compulsory vs. voluntary

Each Processor documents the reasons supporting the designation of the DPO or, rather, the reasons why such designation is deemed not necessary. This documentation forms part of the data protection documentation system of that Processor.

➢ Professional requirements

The DPO has sufficient authority, professional qualities and independence to ensure success in his role, according to the GDPR provisions.

➢ Tasks

The organization assigns to the DPO at least the tasks listed in the GDPR.

➢ Notification to Supervisory Authority

Whenever a DPO is appointed, the organization notifies the Supervisory Authority of such designation and publishes DPO's contact details.

4.2.5.5 People in charge of processing

Individuals who process personal data under the authority of the Controllers or Processor(s) must receive specific formal instructions. Hence, the Controller gives specific instructions, relating also to the implementation of security measures and safeguards, to all of its personnel in charge of processing personal data.

➢ Training and awareness

All Project Partners’ employees should be well informed and aware of data protection implications and be able to carry out their obligations in their work. A data protection education and communication program should be in place and supported by a monitoring system that confirms all employees and/or contractors are appropriately trained on their data protection responsibilities.

➢ Policies and procedures

Data protection policies and procedures exist, are documented in writing, are formally approved by management, implemented, reviewed, updated and approved when there are changes to applicable laws and regulations.

All Project Partners understand and the Controller may ask them to overview all their personal data processing, the data protection risks and the applicable rules and procedures. In such event, they shall provide it with all requested information to the best of their ability without undue delay.

4.2.6 Notice and consent

4.2.6.1 Notice

Each Controller and/or Processor, as appropriate, provides the information required by law to the data subject in a concise, transparent, intelligible and easily accessible form, using clear and plain language.
The data protection notice informs data subjects about the processing of personal data relating to them, even when the personal data is not collected from them as well as of their rights, in order to let them verify in particular the accuracy of the data and the lawfulness of the processing.

### 4.2.6.2 Free and informed consent

Personal data is processed if and to the extent that the data subject has given valid consent to the processing for one or more specific purposes, or another legal basis for processing exists.

Systems or applications are able to document the explicit consent of the data subject so that it can be evidenced at any time.

Other legal grounds for a legitimate personal data processing are the following:

1. performance of a contract;
2. legal obligation;
3. vital interest of data subject;
4. public interest;
5. legitimate interest of the controller or third party.

If "legitimate interest" is used as a basis, the interests that have preceded to the decision, need to be documented as well as any possible mitigating measures which will be taken to be able to proceed with personal data processing based on the defined interests.

### 4.2.6.3 Withdrawal of consent

Data subject’s consent can be withdrawn at any time; even though it will not affect the lawfulness of processing based on consent before its withdrawal.

### 4.2.7 Rights of data subjects

The individual whom the data refers to (data subject) is entitled with specific rights set forth by the law. The GDPR requires that each Controller and/or Processor, as appropriate, must facilitate the exercise of the data subject’s rights, take action on the request within a specific time frame and must communicate the information requested in an intelligible and easy to access form.

#### 4.2.7.1 Right of access

Any individual must be able to exercise the right of access to data relating to him which are being processed.

#### 4.2.7.2 Right to rectification

Each Controller and/or Processor, as appropriate, should have a procedure in place for data subjects to request rectification of their personal data. The procedure specifies in which cases rectification is legitimate.

If a data subject’s request for rectification is legitimate, this is executed across all relevant data storage facilities, including those managed by third parties.

#### 4.2.7.3 Right to erasure

Each Controller and/or Processor, as appropriate, should have a procedure in place for data subjects to request erasure of their personal data. The procedure specifies in which cases erasure is legitimate.

If a data subject’s request for erasure is legitimate, this is executed across all relevant data storage facilities, including those managed by third parties.

#### 4.2.7.4 Right to restriction of processing

Each Controller and/or Processor, as appropriate, should have a procedure in place for data subjects to request restriction of processing of their personal data. The procedure specifies in which cases restriction is legitimate.
If a data subject’s request for restriction of processing is legitimate, this is executed across all relevant data storage facilities, including those managed by third parties.

4.2.7.5 Right to data portability

Each Controller and/or Processor, as appropriate, determines which processes are subject to the right of data portability as well as when the requirements for such right are met.

Data subject can request the organization to receive a machine-readable copy of the personal data the organization holds about them and where possible, enable the transfer of this data to another data controller.

Portability right can be exercised when:

1. processing operations are based on data subject's consent or on contract
2. personal data concerns the data subject and are the same that the latter has provided to the organization
3. the right does not adversely affect rights and freedoms of others
4. the processing is carried out by automated means.

Each Controller and/or Processor, as appropriate, implements appropriate measures and procedures to provide data subject, who is entitled to, with a structured, commonly used and machine-readable copy of the personal data it holds about him and where possible, to enable the transfer of this data to another data controller indicated by data subject.

4.2.7.6 Right to object

Where personal data are processed for scientific or historical research purposes or statistical purposes, the data subjects have the right to object on grounds relating to their particular situation (unless the processing is necessary for the performance of a task carried out for reasons of public interest). The right to object is explicitly brought to the attention of the data subject at the latest at the time of the first communication with the data subject, presented clearly and separately from any other information. Measures should be in place to assess such objections and to ensure that such processing ceases when the request is legitimate and needs to be respected.

Data subjects have right to object, on request and free of charge, to the processing of personal data relating to them for purposes of direct marketing.

4.2.7.7 Automated decision making

Data subject has the right to object to any automatic decision-making (including profiling).

Each Controller and/or Processor, as appropriate, will have determined which processes entail automated decision-making (including profiling) and will have established measures to allow data subjects to object to such automated decision making and profiling. Suitable measures are in place to safeguard the data subject's rights and freedoms and legitimate interest, at least the right to obtain human intervention on the part of the Company/controller, to express his or her point of view and to contest the decision.

4.2.7.8 Timely response to exercise of rights

Each Controller and/or Processor, as appropriate, must confirm to data subjects without delay whether data relating to them are processed and communicate the data to them in an intelligible form. Each Controller and/or Processor, as appropriate, should implement internal procedures in order to be able to provide a timely response to the requests of data subject for the exercise of his rights.

Measures have to be implemented in a way that effectively allows an individual to exercise his or her right to personal data, and that enables Each Controller and/or Processor, as appropriate, to respond to such request appropriately within the required timeframes.

➢ Notification to recipients

In case of a legitimate exercise of rights to rectification, erasure or restriction of processing recipients of the personal data should be informed of the rectification, erasure of that data or of the restriction of processing.
Each Controller and/or Processor, as appropriate, should have a procedure in place for communicating any rectification or erasure of personal data or restriction of processing to the recipients to whom the personal data has been disclosed and for disclosing these recipients to the data subject, if so requested.

### 4.2.8 Data protection documentation system

#### 4.2.8.1 Register of processing

Each Controller and/or Processor, as appropriate, with regard to their processing activities must set up a relevant record, maintained in writing (including in electronic form) and made available easily and swiftly to the supervisory authority on request, as per applicable legal requirements within their respective Member States. The record of processing activities shall contain all the information required by GDPR.

Consequently, the Controller shall have an up-to-date overview of all personal data processing activities and shall maintain records within the Project, that meet the legal requirements posed by the GDPR. By so doing, the Controller will be able to demonstrate compliance to any Supervisory Authority or other state or EU authority concerned.

For the avoidance of doubt, each Project Partner carries the same responsibility above within its own respective organisation.

#### 4.2.8.2 Register of data breaches

A specific register where the breaches have to be recorded together with other information specified by the law, must be maintained by the Controller and shown to the Supervisory Authority upon request. This register is an important element of the data protection documentation system.

Project Partners need to notify immediately and in writing the Controller of any personal data breach within their respective organisations that affects execution of the Project in any way, and to cooperate with the Controller while applying relevant GDPR legal requirements.

### 4.2.9 Data protection assessment

#### 4.2.9.1 Assessment

In the event that a Data Protection Impact Assessment (“DPIA”) is carried out under the Project, the Controller shall ensure that personal data receives the appropriate level of protection in accordance with the assessed data protection risk.

The decision whether to carry out a DPIA under the Project, unless undertaken in respective Project contract, will be made by the Controller upon prior written consultation with the Project Partners.

- **Adequacy of protection**

  The Controller, assisted by Project Partners, should have a process in place in order to assess for all processing the risks of varying likelihood and severity for the rights and freedoms of natural persons, taking into account the nature, scope, context and purposes of personal data processing.

- **Impact assessment in case of high risk (DPIA)**

  When the preliminary assessment highlights that processing represents high risks, a formal and documented DPIA is carried out by ascertaining possible impact on data subject.

  DPIA is conducted in such a way to meet all the requirements set forth by the GDPR (art. 35) in order to confirm the quality and validity of the findings.

- **Prior consultation to Supervisory Authority**

  The Controller has a process in place and roles are assigned in order to ensure that when a DPIA determines that the processing represents high risks, the competent Supervisory Authority is consulted prior to the processing.
4.2.10 Technical and organizational measures

The Controller and each Project Partner, as appropriate, adopts appropriate technical and organisational measures with regard to Project execution (the “Measures”), and reviews and updates them where necessary, to ensure and to be able to demonstrate that processing is in compliance with GDPR.

Each Project Partner shall notify relevant Measures to the Controller in writing. In the event of any queries or further requests by the Controller, each Project Partner undertakes to address them duly and in writing.

In the event that any Project Partner has notified the Measures to its competent Supervisory Authority, it shall inform the Controller thereof, and shall provide respective copies thereof.

4.2.11 Data breach

According to GDPR, the Controller and/or Processor, as appropriate, has to implement adequate Measures in order to prevent personal data breaches.

In addition, the Measures should be able to minimize the adverse effects, in case a security breach to personal data relating in any manner to the Project occurs anyhow.

Should a data breach occur, GDPR sets forth that the Controller and/or Processor, as appropriate, has to notify it to the Supervisory Authority providing specific information, without undue delay and in any case no later than 72 hours from the time of knowledge.

When the breach leads to significant risk of serious adverse effects on the data subject(s) or serious adverse consequences for the protection of personal data, also the latter must be informed without undue delay.

4.2.12 Data transfers to third countries

No international transfers of personal data are expected to take place under the Project.

In the event that any Project Partner wishes to carry out such personal data processing in a third country, it shall notify the Controller in writing and in advance. Unless otherwise expressly specified, any international data transfers carried out by any Project Partner for any reason during Project execution take place at its own exclusive liability and responsibility; same Project Partner shall hold all other Project Partners (including the Controller) harmless from any legal or other claims arising for such personal data processing.

4.2.13 Sanctions and damages

In case of violation of data protection principles and rules, each Project Partner (including the Controller) is responsible for damages and is subject to sanctions. Possible violations may involve civil liability and sanctions in order to ensure that any relevant damage is compensated.

The Project Partner (including the Controller) that is liable for said damages and/or sanctions shall hold all other Project Partners harmless from any claims, costs, and expenses arising from relevant GDPR infringement.

4.2.14 ERATOSTHENES Web-server Personal Data Protection and Privacy policy

The following Personal Data Protection and Privacy Policy is uploaded onto the Project website and document server:

1. Introduction. This Personal Data Protection and Privacy Policy (the “Policy”) aims at providing details of the processing, and related methods of use, of personal data referred to users/visitors (the “User(s)”) of this website that can be reached at the address [teamwork.com] (the “Website”).

   This Policy refers to EU Project [ERATOSTHENES, 780075], (the “Project”).
Web users and visitors are recommended to read carefully this Privacy Policy before sending any personal information and/or filling in any electronic form posted on this website.

This information is given in accordance with applicable EU data protection law, in particular the EU General Data Protection Regulation, and EU applicable Privacy law.

2. Controller. The Controller is the actual data owner per data case i.e., it is expected to be an ERATOSTHENES partner that has full ownership or is the creator of the data set.

3. Scope. This Policy covers this web site only, and no other personal data processing under the Project or any other websites owned or run in any manner by the Controller or Project Partners.

4. Policy and information notice. This site has been designed with the main function of providing information on the activities of the Project. Therefore, in most cases, the collection of the user's personal data is not required.

In certain instances, such as the "newsletter" section and in order to allow the transmission of our newsletter, the interested user is required to fill out a data collection form. In these cases, the user is always free to provide his/her own data and consent to relevant processing. We recommend reading this Policy before providing the data.

In addition, should it be necessary in limited cases to collect personal information for other purposes, this will be clearly shown in the information privacy notices required by law, in order to enable transparency and user awareness. Consent forms and other documentation will be used each time, as appropriate.

The above information aims to define limits and methods of personal data processing of each service, according to which the visitor can freely express his consent and eventually allow the collection of data and its subsequent use.

5. Traffic data. The computer systems and software procedures used to operate this website acquire, during their normal operation, some personal data whose transmission is implicit in the use of Internet communication protocols.

This category of data includes: IP addresses, browser type, operating system, the domain name and website addresses from which you are logged in or out, the information on pages visited by users within the site, the time of access, time period of user's staying on a single page, the internal path analysis and other parameters regarding the user's operating system and computer environment.

This technical / IT data is collected and used only in an aggregated and not immediately identifiable manner; they could be used to ascertain responsibility in case of hypothetical crimes against the site or upon public authorities’ request.

6. Cookies. No cookies are used by this website.

7. Redirections to other websites. From this website, you can connect through special links to other websites of Project Partners within the Project, or of third parties as applicable each time. Controller hereby assumes no responsibility regarding the possible processing of personal data by third-party sites and in respect of the management of authentication credentials provided by third parties.

8. Purposes of processing and data retention. The processing of personal data is carried out mainly by using electronic procedures and media for the time strictly necessary to achieve the purposes for which the data were collected. The User, however, has the right to obtain the cancellation of his data for legitimate reasons.

9. Optional supply of personal information. The supply of personal data required from the User, unless otherwise noted, is optional, but in case of refusal it could be impossible to fulfill the request, or the related activity mentioned.

10. Place of personal data processing. Data processing related to web services of this website takes place, unless otherwise expressly stated, at Controller's establishment, which provides for the corresponding server management. Personal data are only handled by technical personnel of the Controller, specifically in charge of processing, or others charged with occasional maintenance operations. These persons have received specific instructions on the confidentiality of this data.

11. Scope of data flow and dissemination. The data may be used by Controller and/or Project Partners’ employees, as persons in charge of processing, to whom appropriate operating instructions have been given, as well as by third parties who perform operating activities on behalf of them and who act as data processors, in order to fulfill contractual obligations with regard to the Project. Personal data are not disseminated to unspecified recipients. Detailed information on the names of the data processors can be requested by writing to the project coordinator.
12. **Data protection rights.** With regard to the processing of personal data, User has the right, at any time, to obtain confirmation of whether or not the data exists and to have it communicated to him/her in an intelligible format. Users also have the right to know the content and the origin of the data, to check its accuracy or to ask that it be integrated, updated or adjusted. Finally, Users have the right to ask that the data be deleted or made anonymous or to request the blocking of data processed in violation of the law; moreover, they may oppose the processing of the data for legitimate reasons. Requests should be addressed to the project coordinator.

13. **Policy updating.** The possible entry into force of new laws, as well as the evolution and updating of User services or developments in the Project could make it necessary to vary the method of processing of personal data. It is therefore possible that our policy may be modified over time and therefore we invite the visitor to periodically visit this page. To this end, the policy document highlights the date of last version.

4.2.15 **ERATOSTHENES Day-to-Day Data Usage and Related Processes**

Despite the fact that ERATOSTHENES does not use any direct personal data (in the form of data coming out or processed during its research activities), it recognises the needs for creating some process related policies so that there is overall agreement of the usage/storage/retention/opt-out etc of data from every-day (day-to-day) project activities. A list of such matters is included below where the means that the consortium will tackle them reflects the whole consortium agreed approach.

4.2.15.1 **ERATOSTHENES list of contacts**

The ERATOSTHENES list of contacts relates to a single XLS file that includes the names of all the consortium partners and persons and their email address. It also indicates the purposes of contacting each person per organisation (admin, technical, legal etc) and the emailing lists that each belongs to. Only ERATOSTHENES consortium partners have access to this list of contacts. The purpose of this list is to keep a well organised list of contacts for the ERATOSTHENES communications. The data will be erased after the project end (31/12/2020) and not kept or maintained after the project end. This list is being stored at the ERATOSTHENES teamwork document server. Any person has the right to opt out of this list by direct email to the project coordinator.

4.2.15.2 **Meetings’ related material**

This relates to any document created and used for the purposes of project meetings. These may relate to agendas, presentations, minutes, signature lists or any other internal document created for the purposes of ERATOSTHENES meetings. All these documents will be created and maintained for internal purposes of ERATOSTHENES and only ERATOSTHENES partners will have access to them at the ERATOSTHENES teamwork under the meetings section. They will be kept for 5 years after the project end (for auditing reasons, ie 31/12/2025). Any person has the right to opt out of being mentioned in these by direct email to the project coordinator before or after the meeting.

4.2.15.3 **Workshops/Conferences and Training sessions**

These data relate to the creation of workshops, agendas, programmes, participants’ lists etc and in general dissemination material related to ERATOSTHENES organised workshops. Regarding the external publication of this material, we consider that this material can be fully anonymized so that it excludes personal information from the presenters/participants in the related programmes/agendas that will be shared publicly. For the parts of the related material that will be used for the workshop organisation internally to ERATOSTHENES, the related files will be stored in the ERATOSTHENES teamwork server under the section meetings. The data will be kept for 5 years after the project end for auditing reasons (i.e., 31/12/2025). Any person has the right to opt out of being mentioned in these by direct email to the project coordinator before or after the event.

4.2.15.4 **Reporting**

Reporting refers to internal and external (EC) documents including ERATOSTHENES progress of activities, technical overviews etc. Related files will be including documents (reports with no personal identifiable information) and financial data (C forms) sometimes including personal data. The purpose of these data is financial so that partners can claim budget requests for their related effort in ERATOSTHENES. C forms will be maintained by the project coordinator only and stored at internal and secure server. These (per partner) data are not to be shared with anyone internally or externally to ERATOSTHENES, will be kept for 5 years after the project end (for audit purposes, i.e.,
31/12/2025) and will be deleted after this date. Opting out of these data will be possible but will require an updated Form C to be submitted by the project partner.

4.2.15.5 Deliverables, internal documents and other ERATOSTHENES reports

During the ERATOSTHENES project run-time, a large series of documentation and reporting will be provided relating to the project deliverables and/or internal documents etc. These files will be used for the project contractual obligations and shared to: ERATOSTHENES partners, EC, everyone (depending on deliverable type). In these documents, the name or email of authors may be included. Following this, as far as the internal (to ERATOSTHENES) and EC distributed documents are related, they will be used only for the purposes of reporting and stored in the ERATOSTHENES teamwork under the deliverables section. Reports that will be shared publicly (public deliverables) will mention only the partner name and not any other personal information. All reports will be kept for 5 years after the project end for auditing reasons (i.e., 31/12/2025).

4.2.15.6 Source codes

As far as the inclusion of personal information inside source codes is concerned, ERATOSTHENES intends to not use any such information into actual source code files produced in the framework of ERATOSTHENES foreground. In case that any partner wishes to include any personal information, a related consent form will have to be created, used and signed by the data owner(s).

4.2.15.7 Usage of cookies (in ERATOSTHENES sites)

In the cases that in any ERATOSTHENES application (web) requires the usage of cookies, a related pop-up window informing the user must be present, prompting the user to accept (or not) the conditions under which her/his personal information are stored. ERATOSTHENES will maximize efforts to reduce the usage of cookies in its web developments.

4.2.15.8 Lists of stakeholders and ERATOSTHENES contacts

This list refers to internal to ERATOSTHENES lists of external stakeholders including potential technology/results up-takers, major links with end-users and other stakeholders. This list will be used for communication purposes of ERATOSTHENES, no external access will be allowed (restricted to ERATOSTHENES partners) and will be stored in the ERATOSTHENES teamwork (contacts section). When people are being registered to this list, a consent by email will have to be sent by the data owner. The data will be kept until the ERATOSTHENES end, i.e. 31/12/2020. Any person has the right to opt out of being mentioned in these by direct email to the project coordinator.

4.2.15.9 Project related research data (data from living labs)

Any data circulated internally to ERATOSTHENES for research purposes (i.e., data from living labs for source code analysis, sensor data for analytics etc.) must be fully anonymized by the data owner (in this case the data controller) and not relating in any case to personal information as stated in the chapters above.

4.2.15.10 Any other ERATOSTHENES related data

In case that personal information needs to be added in any other document in ERATOSTHENES, the controller (document creator) will have to notify the data owners of their personal details being included into the related document, purpose, retention, storage etc.
5 Making Data FAIR

The FAIR principles describe how research outputs should be organised so they can be more easily accessed, understood, exchanged and reused. Major funding bodies, including the European Commission, promote FAIR data to maximise the integrity and impact of their research investment. Fair data management relates to the EC guidelines on the Data being Findable, Accessible, Interoperable, Reused. The structure of this chapter complies with the FAIR data management template of the EC (DMP component 2).

5.1 Making data findable

To make data findable within the ERATOSTHENES Project Repository, naming conventions are laid out.

For facilitating common browsing and storage in different platforms, no spaces should be used in the document names, and instead the dash character “-” should be used. Project document names must start with the prefix “[ERATHOSTHENES]” in order to facilitate quick identification and indexing. In particular, the following conventions are mandatory for certain types of documents. Names of deliverable documents should follow the convention:

“[ERATHOSTHENES] Dw.n – Deliverable Name.vX.Y.ext”

here:

- "Deliverable Name" is the name of the deliverable as indicated in the Part A of the GA
- “Dw.n[m]” is the deliverable number
  - “w” is the WP number
  - “n” is the numbering within the specific WP
- “vX.Y” is the version number
  - “X” is the version
  - “Y” is the sub-version
- “ext” is the file extension pertaining to the format used. It is normally "doc" or "docx" during the preparation period, and "pdf" for the formal submitted version.

For instance, the name of (the final version of) deliverable D1.1 sent to the EC is “[ETAPAS] D1.1 – Title.v1.0.pdf”

The name of the ERATOSTHENE Technical Reports will follow the convention:


where “ddd” is a three-digit decimal number that will be assigned automatically to a new Technical Report

For the other type of datasets that are not in the form of only a single document, the folder, package, or online repository will follow the naming convention:

- “[ERATOSTHENES] Data – Data Asset Name.vX.Y.ext”

5.2 Making data openly accessible

5.2.1 Open Access to scientific publications

In accordance with the Grant Agreement, all peer-reviewed scientific publications related to the results of the ERATOSTHENES-Project will be published as open access.

This includes the obligation to
• deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications together with the research data needed to validate the results presented in the deposited scientific publication as soon as possible.

Open access to the deposited publication via the repository must be ensured at the latest

• on publication, if an electronic version is available for free via the publisher, or
• within six months of publication (twelve months for publications in the social sciences and humanities) in any other case.

Open access via the repository on the bibliographic metadata that identifies the deposited publication must be ensured. It must be provided in a standard format and must include

• the terms “European Union (EU)” and “Horizon 2020”;
• the name of the action, acronym and grant number;
• the publication date, and length of embargo period if applicable, and
• a persistent identifier.

5.2.2 Open Access to research data

In accordance with the Grant Agreement, research data will be made available to the highest possible extent. The research data will be made available for the project members via the ERATOSTHENES repository during the project. After the project, research data will be made available via Zenodo.

• Data from technical research. Data used by the research partners or generated from the research work related to the development and evaluation of the prototype tools will be opened as much as possible.

• Use Case Specific Data. As it is not yet exactly clear, which data will be generated throughout the use cases, at this point it cannot be stated which of the retrieved data will be made publicly available. Some use cases involve (sensitive) personal data which is protected by the GDPR and therefore will not be shared publicly or within the entire consortium. Some other data that are sensitive to the business of the use case providers will also be protected.

• Stakeholder-related data. Most of the stakeholder data contains personal information, which is why it will not be made openly available. However, anonymized results from workshops and other stakeholder engagement events will be made openly available through different deliverable.

5.2.2 How will the data be made available?

The data will be made available using Zenodo. Zenodo is an open-access repository developed under the European OpenAIRE program operated by CERN which provides researchers the sharing, curation and publication of data and software. The OpenAIRE project was commissioned by the EC to support their nascent Open Data policy by providing a catch-all repository for EC funded research.

Zenodo allows to create an own collection and accept or reject uploads submitted to it. It is possible to update all research outputs from all fields of science. In the upload form it is possible to choose between different types of files: publications, posters, presentations, datasets, images, software, videos/audio and interactive materials such as lessons.

Zenodo assigns all publicly available uploads a Digital Object Identifier (DOI) to make the upload easily and uniquely citeable.

5.3 Making data interoperable

All data will be store as standard formats (e.g., PDF, doc, JSON, blockchain record, etc.).

The documents and reports created by ERATOSTHENES will contain the executive summary in the beginning of the document, summarizing the contents, target readers, and the expected way to use the document. All source code
repository will contain a README file in under the root path with instructions about how to build, run and contribute to the code base.

For other types of data assets we will create the metadata, using as much as possible the format and vocabularies as defined in the DCAT standard\textsuperscript{15}.

5.4 Increase data re-use

- **How will data be licensed to permit the widest reuse possible?** We aim to enable open access to all research data via CC-BY licence. However, as it is yet to emerge which data exactly will be generated by the use cases, this might be adapted throughout the project.
- **When will the data be made available for re-use?** The data will be made available for re-use at the soonest moment possible, however it is not yet clear, when this will be, as it depends on which data will be generated throughout the use cases.
- **Which data quality assurance process will be in place?** Data related to deliverables will go through the same internal review processes, and the quality of the data will be part of the criteria for internal review.

5.5 Security Aspects of Repository Server (TEAMWORK)

For the internal purposes of ERATOSTHENES, INLECOM has configured and adapted the TEAMWORK™ server. Below we describe the server capabilities and services as far as physical, network and content security are concerned\textsuperscript{16}.

5.5.1 Physical Security

TEAMWORK’s servers are hosted within AMAZON’s Web Services (AWS) environments as world class servers and data centres making utilising AMAZON’s web/application servers, file servers and databases. TEAMWORK spans several layers of the AWS using the Elastic Computing (EC2) for application and web-servers. The file servers are connected to the AWS storage facility (S3) and the applications are connected to the Relational Database System (RDS).

Physical security is guaranteed here through increased physical security at the AWS servers supported by AMAZON’s access control and extensive anti-seismic bracing as well as fire detection/protection systems, monitored on a 24x7x365 basis.

Digital security is residing on the AWS firewall system, SPAM and DOS protections and visibility over only trusted IPs. Multiple zone handling is achieved through the Virginia AWS Data Center.

5.5.2 Network Security

Inside the network security aspects, we consider the server security and the data included into the server. A 256-bit secure socket layer (SSL) with the AES algorithm and 2048-bit key length are used to guarantee on a both server authentication and data encryption levels.

5.5.3 Content Security

For this aspect, full access control is managed by TEAMWORK ensuring the highest clearance to access the data-center data while there is strictly regulated access to data. User authentication is used to provide access to the server data based on unique accounts, password protected. The project administrator is also capable of providing or restricting access to particular folders/data by setting rights and permissions.

\textsuperscript{15} https://www.w3.org/TR/vocab-dcat/

\textsuperscript{16} TEAMWORK security overview report.
6 Conclusions

This document complements WP7 activities and contributes to the ERATOSTHENES project management processes and policies by providing a complete data management plan (as part of WP7). The report starts with an exhaustive description of the data that will be used in all technical, project management, administrative and the pilots of ERATOSTHENES. This provides the ERATOSTHENES data lifecycle overview and the types of data assets that range within IoT-related data, trust-related data, stakeholder-related and project management data. Then an analysis of the data in each WP and task is included together with extensive descriptions of their properties.

What follows is the ERATOSTHENES policy and overall overview for respecting the General Data Protection Regulation (GDPR). This includes also the ERATOSTHENES GDPR policy that is agreed with the consortium.

This document also includes the potential data that can be generated or collected through the ERATHOSTHENES project, how we plan to make the data FAIR, and how to assurance the security and privacy of personal and sensitive data.

This deliverable is the first version of the data management plan (DMP), and will be updated in the D7.9 (second and final version of the report). This will be updated towards the project end, including latest updates on data, actual data shared, metadata provided as well information on their public sharing and the related platforms.
### Annex I – Data Management Report

<table>
<thead>
<tr>
<th>Performance Indicator</th>
<th>Means of verification</th>
<th>Target Values</th>
<th>Compliance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Creation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format</td>
<td>Compliance with existing standards of data exchange</td>
<td>XLS, XML, etc.</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Availability and Readability</td>
<td>Whole package of data available, non-corruption, whole percentage collected</td>
<td>100% received 100% accessible</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Fit For Use</td>
<td>Data follow data compliancy for proper processing and review</td>
<td>100% usable by intended beneficiary/ies</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Consistency and Completeness</td>
<td>Data are consistent and complete for the intended purpose</td>
<td>Including 100% of information for the intended purpose</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Relation</td>
<td>Data follow a precise relation to their purpose</td>
<td>100% purpose precision</td>
<td>√ or ×</td>
</tr>
<tr>
<td><strong>Data Processing and Analysis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data logic</td>
<td>Data can be and are processed following a concise logic and approach</td>
<td>New and processed data follow precise data logic</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Organization and Utility</td>
<td>Suitable content organization of data under processing</td>
<td>100% organized data</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Validation</td>
<td>Ensuring that the data under processing are correct and relevant</td>
<td>100% validated and relevant data</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Aggregation</td>
<td>Whenever multiple data need to be aggregated ensure that this is done in a concise approach</td>
<td>100% aggregate-able data</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Transformation</td>
<td>Transformation of data to the proper format(s) for processing</td>
<td>Capability of data for transformation (if needed)</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Calibration</td>
<td>Calibration of data for their intended purpose</td>
<td>Data properly calibrated</td>
<td>√ or ×</td>
</tr>
<tr>
<td><strong>Data Publication and Utilization</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means-independent</td>
<td>Transferring of the data in a means-independent approach</td>
<td>100% means independent transferability</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Security (a)</td>
<td>Data stored in a secure enough server</td>
<td>At least access control provided over a TLS protocol</td>
<td>√ or ×</td>
</tr>
<tr>
<td><strong>Data Storage, Archiving and Re-Use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up to date</td>
<td>Ensuring that the stored data are up to date for the specific purpose and no later version exists</td>
<td>100% updated</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Meta Data</td>
<td>Existence of meta data in stored files</td>
<td>Relevant metadata have been included into the archive per data set</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Security (b)</td>
<td>Access control provided</td>
<td>Access control setup</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Security (c)</td>
<td>Server is considered as safe enough (TLS connection protocol)</td>
<td>At least TLS connection configuration</td>
<td>√ or ×</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------------------------------------------</td>
<td>--------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>Control of server bandwidth</td>
<td>Effective storage server bandwidth &gt; 2 MBPS</td>
<td>√ or ×</td>
</tr>
<tr>
<td>Expiration</td>
<td>Properly setting expiration dates for all data after which the data will be deleted</td>
<td>Expiration date noted</td>
<td>√ or ×</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GDPR Compliancy¹⁷</th>
<th>Data Subjects Details</th>
<th>Overall Compliancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Data Description¹⁸</td>
<td>Access¹⁹</td>
<td>Storage²⁰</td>
</tr>
<tr>
<td>eg conference programme</td>
<td>Internal and external</td>
<td>ERATOSTHENES document server, ERATOSTHENES website</td>
</tr>
</tbody>
</table>

¹⁷ To be completed for each type of data falls under GDPR or is connected to it in any way.
¹⁸ Overall data description.
¹⁹ Determines who has access to the particular data (internal, external to consortium).
²⁰ Storage places of actual data.
²¹ Intended purpose of data and reasons for keeping.
²² Duration of stored data (until when they will be kept).
Annex II - Non-Disclosure Agreement (DNA) for the Advisory Board (AB) members

Secure management of IoT devices lifecycle through identities, trust and distributed ledgers

Advisory Board – Non-Disclosure Agreement

This disclosure agreement has been prepared in support to the project activities and to restrict the release of confidential/restricted data of ERATOSTHENES. This NDA will be signed by all members of the ERATOSTHENES Advisory Board (AB).

NON-DISCLOSURE AGREEMENT

For the project ERATOSTHENES.

This Non-Disclosure Agreement (“NDA”), is made as of …………………….., hereinafter referred to as “Effective date”, and is entered into by and between the following parties (each a “Party and collectively the “Parties”):

(1) INLECOM INNOVATION ASTIKI MI KERDOSKOPIKI ETAIREIA.
(2) name of entity or member name.

Party #1 above are hereinafter referred to as a “Project Coordinator” and collectively as the “Project Coordinator”.

-AND-

The other Parties above are hereinafter each referred to as a “Advisory Board Party” and collectively as the “Advisory Board Parties”.

WHEREAS, the Project Coordinator of project “ERATOSTHENES” (Secure management of IoT devices lifecycle through identities, trust and distributed ledgers) funded by the European Union’s Horizon 2020 research and
innovation programme under the Grant Agreement No 101020416 has entered into a consortium agreement (the “Consortium Agreement”), with effective date October 2020, regarding ERATOSTHENES.

The Advisory Board Parties are members of an advisory committee (the “Advisory Board”), connected to ERATOSTHENES, but are not partners in ERATOSTHENES (in which case not parties to the Consortium Agreement).

WHEREAS, the Project Coordinator and the Ethics Committee Parties, for their mutual benefit, anticipate the need to disclose to and receive from each other information regarding ERATOSTHENES, which the disclosing Party considers to be proprietary and which relates to the Advisory Board Parties’ participation in the Advisory Board (hereinafter referred to as the “Project”).

The Parties are aware that the Consortium Parties’ undertakings towards each other regarding confidentiality and non-disclosure are governed by the Consortium Agreement and consequently not by this NDA.

NOW, THEREFORE, in consideration of the covenants and agreements herein contained, the sufficiency and adequacy of which are acknowledged, the Parties agree as follows:

1. “Information” as used in this Non-Disclosure Agreement (NDA) is any information disclosed by a Consortium Party to an Advisory Board Party related to the Project or by an Advisory Board Party to a Consortium Party related to the Project. A Consortium Party or an Advisory Board Party who discloses Information pursuant to the first sentence of this Article 1 is hereinafter referred to as a “Disclosing Party”. An Advisory Board Party or a Consortium Party who receives Information pursuant to the first sentence of this Article 1 is hereinafter referred to as a “Receiving Party”.

2. All information in a tangible form (including information transmitted electronically) which the Disclosing Party desires the Receiving Party to treat as “Confidential Information”, shall be marked by the Disclosing Party with the legend “CONFIDENTIAL” or “PROPRIETARY” in order to identify its confidential nature. If Confidential Information is disclosed orally, visually or in any other non-tangible manner of disclosure, such information shall also be entitled to protection as Confidential Information, if the information disclosed is identified as Confidential Information at the time of disclosure and is subsequently reduced to appropriate written form and furnished to the Receiving Party within one (1) month of the time of original disclosure.

3. Each Party hereto agrees to receive Information from the other Party for the sole purpose of evaluating the progress and research results of ERATOSTHENES.

4. For the term of the confidentiality obligation provided for in this NDA, the Receiving Party agrees to hold the Confidential Information received from the Disclosing Party in confidence and not to disclose such Confidential Information to any other Party or party and to use such Confidential Information only for specified purposes in this NDA. The Receiving Party agrees that it will use the same degree of precaution and safeguards as it uses to protect its own information of like importance, but in no case with any less than reasonable care. For the avoidance of doubt, the Receiving Party also agrees that it will not use Confidential Information to develop or produce any product or to develop or perform any service without an agreement in writing with the Disclosing Party authorizing such development, production or performance.

5. The obligation of confidentiality shall not apply to:
(i) Information, which is now or later becomes publicly available through no act or failure to act on the part of the Receiving Party in violation of this NDA; or

(ii) Information, which the Receiving Party can prove is rightfully acquired by the Receiving Party from a third party which is not under any obligation of confidentiality with respect to the Information; or

(iii) Information, which the Receiving Party can prove is developed by or for the Receiving Party independently of information which the Receiving Party is required to keep confidential under this NDA; or

(iv) Information, which is required to be disclosed by national law or any court having jurisdiction over the Receiving Party; or

(v) Information, which the Receiving Party can show it possessed before the Disclosing Party disclosed it to the Receiving Party.

6. The Receiving Party may disclose Confidential Information to its employees who need to know and use Confidential Information in furtherance of the purposes of this NDA and who are under an obligation to keep the Confidential Information confidential to the same extent as the Receiving Party.

7. The Receiving Party will return Confidential Information to the Disclosing Party upon written request by the Disclosing Party. However, the Receiving Party will be allowed to retain one (1) archival copy of any document if required by national law.

8. The Agreement shall cover Confidential Information disclosed by the Disclosing Party to the Receiving Party for the purpose of this NDA. Termination of obligations of confidentiality and non-use shall not be construed, however, as a grant of any license under patent rights or copyrights of the Disclosing Party, any other Party or party.

9. Nothing contained in this NDA shall obligate any Party to enter into any agreement with any other Party for the purpose of any development project or for any other purpose. Also, this NDA does not include, expressly or by implication, any representations or warranties as to the accuracy, efficacy, completeness, capabilities, safety or any other qualities whatsoever of any Information or materials provided under this NDA, nor does this NDA grant the Receiving Party any license on the Information of the Disclosing Party.

10. This NDA shall be construed in accordance with and governed by the laws of Belgium. All disputes arising out of or in connection with this NDA, which cannot be solved amicably, shall be finally settled under the Rules of Arbitration of the International Chamber of Commerce by one or more arbitrators appointed in accordance with said Rules. The place of arbitration shall be Brussels, if not otherwise agreed by the conflicting Parties. The arbitration will be final and binding upon the Parties.

11. Nothing in this NDA shall limit the Parties’ right to seek injunctive relief or to enforce an arbitration award in any applicable competent court of law.

12. This NDA can be amended or modified only by an amendment in writing signed by all Parties. Parties foresee to enlarge the number of Advisory Board Parties and agree to amend this NDA for accession upon request.

13. This NDA constitutes the entire agreement and sole understanding of the Parties with respect to the subject matter hereof, save for the Consortium Parties’ undertakings towards each other regarding confidentiality and nondisclosure which are, as stated above, governed by the Consortium Agreement and not by this NDA.
14. An entity becomes a Party to this NDA upon signature of this NDA. This NDA shall have effect from the Effective Date and will expire upon complete fulfilment of ERATOSTHENES. The secrecy obligations last until five (5) years after the expiration of this NDA. Provisions and/or obligations which naturally are intended to continue to exist after the expiration of this NDA, survive such expiration.

15. The Information is provided ‘as is’ without any kind of warranty, express, implied and/or statutory, including but not limited to the fact that the application and/or use of the Information does not infringe the intellectual property rights and/or other rights of a third party. The Parties are liable towards each other only for damages, which are the direct result of a culpable shortcoming namely a breach of contract, on the part of the breaching Party under this NDA. The Parties are not liable to each other for any kind of other damages, losses, expenses, indirect and/or consequential damages, which the Receiving Party suffers, arising out of and/or in connection with the accuracy, completeness and/or other quality issue with respect to, and/or the application and/or use by the Receiving Party of the Information.

16. The Receiving Party acknowledges and agrees that all property, including intellectual property, in the Information shall remain and be vested in the Disclosing Party.

17. By signing this NDA, the Receiving Party consents to the project coordinator keeping their personal data (name, email and affiliation) for the purposes of the ERATOSTHENES advisory board and only. Advisory Board personal data will be stored at ERATOSTHENES's web-space (TEAMWORK®), as a secure server that only ERATOSTHENES partners will be having access to. Any party reserves the right to be removed (opt-out) from this list at any point via direct contact to the project coordinator. These data will be kept for audit purposes up to 5 years after the project end date.

IN WITNESS WHEREOF, the Parties have accepted the terms and conditions of this NDA and caused it to be duly signed by the undersigned authorized representatives in separate signature pages.

Company name:

Name(s):

Title(s):

Signature(s):

Date: