IoT-NGIN

D7.1 – Data Management Plan (DMP)

WORKPACKAGE	WP7	**PROGRAMME IDENTIFIER**	H2020-ICT-2020-1
DOCUMENT	D7.1	**GRANT AGREEMENT ID**	957246
REVISION	V0.1	**START DATE OF THE PROJECT**	01/10/2020
DELIVERY DATE	31/03/2021	**DURATION**	3 YEARS

DISCLAIMER

ACKNOWLEDGEMENT

This document is a deliverable of IoT-NGIN project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 957246.

The opinions expressed in this document reflect only the author's view and in no way reflect the European Commission's opinions. The European Commission is not responsible for any use that may be made of the information it contains.

PROJECT ACRONYM	IoT-NGIN
PROJECT TITLE	Next Generation IoT as part of Next Generation Internet
CALL ID	H2020-ICT-2020-1
CALL NAME	Information and Communication Technologies
TOPIC	ICT-56-2020 - Next Generation Internet of Things
TYPE OF ACTION	Research and Innovation Action
COORDINATOR	Capgemini Technology Services (CAP)
PRINCIPAL CONTRACTORS	Atos Spain S.A. (ATOS), ERICSSON GmbH (EDD), ABB Oy (ABB), INTRASOFT International S.A. (INTRA), Engineering-Ingegneria Informatica SPA (ENG), Bosch Sistemas de Frenado S.L.U. (BOSCH), ASM Terni SpA (ASM), Forum Virium Helsinki (FVH), Optimum Technologies Pilroforikis S.A. (OPT), eBOS Technologies Ltd (EBOS), Privanova SAS (PRI), Synelixis Solutions S.A. (SYN), CUMUCORE Oy (CMC), Emotion s.r.l. (EMOT), AALTO-Korkeakoulusaatio (AALTO), i2CAT Foundation (I2CAT), Rheinisch-Westfälische Technische Hochschule Aachen (RWTH), Sorbonne Université (SU)
WORKPACKAGE	WP7
DELIVERABLE TYPE	PLAN
DISSEMINATION LEVEL	PUBLIC (Public, fully open, e.g. web)
DELIVERABLE STATE	FINAL
CONTRACTUAL DATE OF DELIVERY	31/03/2021
ACTUAL DATE OF DELIVERY	31/03/2021
DOCUMENT TITLE	Data Management Plan (DMP)
AUTHOR(S)	Rizkallah Touma (I2CAT)
REVIEWER(S)	Josep Escrig (I2CAT) Veli Airikkala (FVH) Antonis Gonos (OPT) Antonello Corsi (ENG)
ABSTRACT	SEE EXECUTIVE SUMMARY
HISTORY	SEE DOCUMENT HISTORY
KEYWORDS	IoT-NGIN, Data Management Plan, DMP, FAIR Data, GDPR, IPR

Document History

Version	Date	Contributor(s)	Description
V0.1	29/01/2021	I2CAT	Table of content
V0.2	16/03/2021	FVH, OPT, SYN, I2CAT, BOSCH, ASM, EMOT	Dataset questionnaire and details
V0.3	23/03/2021	I2CAT	First complete version with internal revision
V0.4	26/03/2021	FVH	Review / feedback from FVH
V0.5	30/03/2021	OPT	Review / feedback from OPT
V0.6	30/03/2021	ENG	Review / feedback from ENG
V1	30/03/2021	I2CAT	Final reviewed version

Table of Contents

List of Tables

List of Acronyms and Abbreviations

AGLV	Automated Guided Land Vehicles
AGV	Automated Guided Vehicles
API	Application Programming Interface
CC	Creative Commons
CKAN	Comprehensive Knowledge Archive Network
DCAT	Data CATalogue Vocabulary
DMP	Data Management Plan
DOI	Digital Object Identifier
FAIR	Findability, Accessibility, Interoperability, Reusability (of Data)
GDPR	General Data Protection Regulation
GPS	Global Positioning System
HTTPS	HyperText Transfer Protocol Secure
IoT	Internet of Things
IP	Internet Protocol
IPR	Intellectual Property Rights
LAN	Local Area Network
MaaS	Mobility as a Service
MQTT	Message Queuing Telemetry Transport
ODC	Open Data Commons
ORDP	Open Research Data Pilot
PMU	Phaser Measurement Units
PQA	Power Quality Analyzers
SMX	Smart Meter eXtension
SSL	Secure Sockets Layer
ToF	Time of Flight
URL	Uniform Resource Locator
UWB	Ultra-Wideband
VPN	Virtual Private Network
WP	Work Package

Executive Summary

This deliverable is the first version of the Data Management Plan (DMP) to be followed by the IoT-NGIN consortium. The main pillars of the DMP are the participation of IoT-NGIN in H2020's Open Research Data Pilot and the implementation of FAIR data principles, while at the same time respecting the Intellectual Property Rights of the consortium partners and the protection of personal data of the participants in the trials and living labs following the GDPR. Due to its importance, GDPR compliance will be further detailed in a separate deliverable, *D1.1 - Definition and analysis of use cases and GDPR compliance*.

A summary description of the datasets identified so far and a preliminary data inventory table are also provided in the DMP. Data management processes, such as data security, data sharing and data archiving, are also highlighted and discussed in varying degrees of detail depending on the information available at this early stage of the project. The DMP is expected to be updated multiple times throughout the project, in line with the guidelines on H2020 data management plans, and the information provided within will be modified according to the review process and timetable provided in this document.

1. Introduction

Throughout its living labs and trials, IoT-NGIN is expected to collect, generate and process a vast amount of data in the fields of Smart Mobility, Smart Agriculture, Industry 4.0 and Smart Energy. This data will stem from multiple types of sources, such as IoT sensors, social networks, drones and radars, in a multitude of formats and types. The management of this data is a complex task that requires coordinated efforts by the consortium partners to address its many challenges and guarantee that unified data management processes are followed throughout the project.

One of the most important aspects that must be addressed in the Data Management Plan (DMP) is the protection of Intellectual Property Rights (IPR) of the consortium partners. Moreover, some of the IoT-NGIN trials will involve human participants and will require gathering basic personal data (e.g. images, location in trial sites). The protection of this data is of utmost importance to the consortium and will be specifically addressed in an independent deliverable, *D1.1 - Definition and analysis of use cases and GDPR compliance,* in alignment with the General Data Protection Regulation 2016/679 (GDPR) and other relevant regulations at the national level.

Furthermore, given the cutting-edge technologies that will be used and researched in IoT-NGIN, this data is expected to be extremely valuable to practitioners in the abovementioned fields, both within the project consortium and beyond. Thus, the IoT-NGIN consortium, in its commitment to H2020's Open Research Data Pilot, will also put in place procedures to publish the generated data in a free accessible manner following FAIR data principles as part of the data management plan.

The rest of this deliverable is structured as follows. *Section 2* explains the main guiding principles of the DMP, most notably the participation in H2020's Open Research Data Pilot. *Section 3* summarizes of the data that has been identified so far in the project's trials and living labs. *Section 4* discusses the processes that the consortium will put in place to follow the FAIR data principles for the datasets that will be openly shared with third-parties. *Section 5* describes the DMP review process and timetable. Finally, *Section 6* concludes the deliverable.

The deliverable also includes two annexes. *Annex I* includes a data inventory table with a detailed description of the datasets identified so far, whereas *Annex II* summarizes all the FAIR data considerations that the consortium will put in place for the identified datasets.

2. Guiding Principles

2.1. H2020's Open Research Data Pilot

Grounded in its belief in the importance of Open Data and its benefits to the scientific community, the European Commission started the Open Research Data Pilot (ORDP)[1] as part of the European Union Framework Programme for Research and Innovation, Horizon 2020. As stated in the 2017 Work Programme, the pilot is being extended to cover all thematic areas for the remaining duration of H2020.

As such, IoT-NGIN and all other projects funded under the H2020 framework programme are by default part of the ORDP, unless they choose to opt-out of the pilot. The legal requirements for their participation are set out in Article 29.3 of the Grant Agreement[2]. The IoT-NGIN consortium believes in the power and necessity of Open Data that is freely accessible to the scientific and research communities and will participate in the pilot.

The strategy of the ORDP is very flexible and usually described as "*as open as possible, as closed as necessary*". Thus, taking part of the pilot does not entail that all research data and results collected, processed or produced during the project must be made openly and freely accessible. It is left at the discretion of the project consortium to decide which data is suitable for open access and which must be protected. The most common reasons for partial or complete opt-out of the ORDP are[3]:

- Participation is incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited,
- Participation is incompatible with the need for confidentiality in connection with security issues,
- Participation is incompatible with rules on protecting personal data,
- Participation would mean that the project's main aim might not be achieved,
- The project will not generate / collect any research data, or
- Any other legitimate reasons explained at the proposal stage or later on during the project.

[1] OpenAIRE - What is the EC Open Research Data Pilot: https://www.openaire.eu/what-is-the-open-research-data-pilot

[2] AGA - Annotated Model Grant Agreement, Article 29.3 - Open access to research data: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf#page=246

[3] Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, Opting out - partially or entirely: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf#page=8

The details of which datasets are to be included in the ORDP and which are to be left out in IoT-NGIN, along with other relevant aspects, are discussed in detail throughout the remaining sections of this Data Management Plan.

Consortium Statement. Open Research Data Pilot.

The IoT-NGIN consortium firmly believes in the importance of Open Data and the advantages that sharing data beyond the consortium brings to the scientific and research communities. As such, IoT-NGIN will participate in the Open Research Data Pilot (ORDP). The consortium partners will give their best efforts to allow open and free access to as much as possible of the data collected, processed and produced during the project lifetime following the principles of FAIR data.

Nonetheless, the consortium reserves the right to perform a partial opt-out of the ORDP for datasets whose nature in terms of security, privacy and / or Intellectual Property does not allow it to be freely and openly shared, in accordance with the ORDP strategy and the exceptions therein mentioned.

2.2. Intellectual Property Rights

Matters related to Intellectual Property Rights (IPR) handling were defined in the Consortium Agreement, where rules were set out to identify and protect the results of the IoT-NGIN work, especially innovations. The Consortium Agreement was aligned to the European Commission's policy on knowledge and intellectual property, and at the same time respected any restrictions from the project partners which may have IPR protecting some of their technologies or data. Aspects addressed in the Consortium Agreement include:

- Ownership and protection of foreground,
- Confidentiality,
- Publication of foreground,
- Access rights to foreground,
- Access rights to background.

Moreover, whenever the IPR agreement allows, the consortium commits itself to publish the overall project results on the project website and any other suitable portals without charging intellectual property rights. For scientific publications in particular, all consortium members will apply the H2020 Open Access Policy, as set out by Article 29.2 of the Grant Agreement[4], and deposit scientific peer-reviewed publications in an open and free centralized repository to be determined at the time of publication.

[4] AGA - Annotated Model Grant Agreement, Article 29.2 - Open access to scientific publications: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf#page=245

Consortium Statement. Intellectual Property Rights.

The IoT-NGIN consortium recognizes and respects the Intellectual Property Rights (IPR) of the consortium partners over the technologies and data produced during the project. Matters related to IRP were defined in the Consortium Agreement, including ownership and protection of foreground, confidentiality, publication of foreground, access rights to foreground and access rights to background, among others.

Furthermore, in alignment with the European Commission's policy on knowledge and intellectual property and with Article 29.2 of the Grant Agreement, all peer-reviewed scientific publications produced by the consortium will be deposited in an open and free centralized repository.

2.3. Data security

Throughout the living labs and trials that form part of IoT-NGIN, a vast amount of data is needed to successfully complete the research and innovation objectives of the project. This includes data collected from IoT sensors deployed by the partners involved in the different trials in fields ranging from Smart Mobility and Smart Agriculture to Industry 4.0 and Smart Energy Grids.

In all these trials, a holistic data security approach will be followed to protect and ensure respect to the main pillars of Information Security[5]:

- **Confidentiality:** this is the assurance that information is not disclosed to unauthorized individuals, groups, processes, or devices. Highly confidential data must be encrypted so third parties cannot easily decrypt it. Only those who are authorized to view the information are allowed access.
- **Integrity:** the accuracy and completeness of vital information must be safeguarded. Data should not be altered or destroyed during transmission and storage. This involves making sure that an information system is not tampered by any unauthorized entities. Policies should be in place so that users know how to properly utilize their system.
- **Availability:** this means that authorized users have timely and easy access to information services. IT resources and infrastructure should remain robust and fully-functional at all times even during adverse conditions, such as database conundrum or fall-overs. It involves protecting against malicious codes, hackers, and other threats that could block access to the information system.
- **Authenticity:** this security measure is designed to establish the validity of a transmission, message, or originator, or a means of verifying an individual's authorization to receive specific information. Authentication prevents impersonation and requires users to

[5] The 5 Pillars of Information Security and How to Manage Them: https://resourcecenter.infinit-o.com/blog/the-5-pillars-of-information-security-and-how-to-manage-them/

confirm their identities before being allowed access to systems and resources. This includes usernames, passwords, emails, biometrics, and others.

- **Non-Repudiation:** this assures that the sender of data is provided with proof of delivery and the recipient is provided with proof of the sender's identity, so neither party can deny sending, receiving, or accessing the data. Security principles should be used to prove identities and to validate the communication process.

These pillars will be fulfilled through a methodical security analysis that will be done on a case-by-case basis in all the trials of IoT-NGIN. The analysis will include a study on the nature of the data collected and identify any security risks related to the collection, processing and sharing of the data. Special provisions will be implemented in the case of personal data, as discussed further in *Section 2.4*.

Additionally, several consortium partners have their own policies on data security that will be followed and taken into consideration, such as FVH's Data Security Policy or SYN's Data Security Plan, which includes provisions on data control, data backup, communication security, software security, dynamicity management, robustness management, security monitoring and physical security measures, among others.

On a technical level, security measures will include:

- Promoting the use of secure network protocols, such as HTTPS or SSL, for the transmission of the data,
- User authentication and verification methods, such as login procedures with verified credentials or two-factor authentication, for sensitive data that requires an additional layer of protection,
- Techniques to protect the data from automated requests and malicious access, such as CAPTCHAs, and
- Data encryption, for both storage and transmission, when needed.

Consortium Statement. Data Security.

The IoT-NGIN consortium will ensure the data security of all data collected and produced during the trials of the project, and in particular of WP7. Thorough security measures will be put in place to respect the main pillars of Information Security: confidentiality, integrity, availability, authenticity and non-repudiation. These measures will include the use of secure communication protocols, user authentication and anti-malicious access techniques, among others.

2.4. Personal data protection and ethical issues

Part of the data collected during the activities of IoT-NGIN, particularly the trials of WP7, includes sensitive personal information that requires specific protection. This includes data such as names, video images, email addresses and mobility data of the participating subjects in some trials. The procedures to collect, store and process this data will follow

relevant regulations, such as the EU's General Data Protection Regulation (GRPD) 2016/679[6], and any additional regulations at the national level.

Furthermore, given the need for Machine Learning and Artificial Intelligence techniques to achieve the project objectives, the impact of the GDPR on the use of data in AI techniques will be specifically examined and analyzed, in accordance with a recent study commissioned by the Directorate-Generate for European Parliamentary Research Service[7].

In addition to this, and to ensure the fundamental right to data privacy of the data subjects participating in the project, the consortium will consider all relevant ethical issues that may arise from its data collection and processing procedures. These procedures will be thoroughly studied and optimized to include techniques such as explicit informed consent forms, data anonymization or pseudonymization whenever needed.

Understanding the sensitivity and complexity of data privacy, the consortium will dedicate a separate deliverable, *D1.1 - Definition and analysis of use cases and GDPR compliance*, to address these issues and provide a detailed analysis of the use cases, conditions, technical and privacy (GDPR) requirements that would govern the actual measurements in the trials and living labs.

Consortium Statement. Personal Data Protection.

The IoT-NGIN consortium holds in high regard the basic right to privacy of all data subjects involved in the project. As such, all personal data needed for the execution of some of the project's trials will be collected and processed in accordance with current regulations, such as the EU's General Data Protection Regulation (GDPR) 2016/679, as well as any additional regulations at the national level.

Given the importance of this aspect, an independent deliverable dedicated to GDPR compliance and data protection, *D1.1 - Definition and analysis of use cases and GDPR compliance*, will be produced by the consortium.

2.5. Data sharing

Two different contexts for data sharing will be considered by the IoT-NGIN consortium:

1. Data sharing between the **consortium partners** to ensure the successful execution of the different tasks of the project,
2. Data sharing with **third parties** to promote the reuse of the research data produced during the project.

Considering the first context, each partner of the consortium will be responsible for implementing mechanisms for the data they want to share. The use of Application

[6] General Data Protection Regulation (GDPR) 2016/679: https://eur-lex.europa.eu/eli/reg/2016/679/oj
[7] The impact of the General Data Protection Regulation (GDPR) on artificial intelligence: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf

Programming Interfaces (APIs) will be promoted and given priority when possible, as a means of providing fast, standardized and secure data sharing. The partners will also commit to respect all the data security and personal data protection mechanisms discussed in previous sections before sharing the data they collect and / or produce.

As for the second context, the consortium will provide all data possible for reuse to third parties in an open and free manner after performing the corresponding security and privacy clearance. This aspect of data sharing will form part of the ORDP discussed in *Section 2.1*, and will be implemented following the FAIR principles further discussed in *Section 4*. All of the data shared as part of the ORDP will be uploaded to a centralized Open Access repository, such as Zenodo[8], and the creation of a new repository specific to IoT-NGIN will be considered if no suitable options are found.

Consortium Statement. Data Sharing.

The IoT-NGIN consortium will implement secure data sharing mechanisms that respect intellectual property rights and data privacy and protection regulations, both in the context of the project and when sharing data with third parties. Secure APIs will be promoted by the consortium partners to share the data among the consortium, whereas the data sharing aspects with third parties beyond the consortium will be governed by the FAIR data principles further discussed in this Data Management Plan.

2.6. Long-term archiving and preservation

As specified by the "*Rules of Good Scientific Practice*" adopted by the Max Planck Society[9], primary research data should be reliably secured and stored for at least 10 years. During the lifetime of the project, this will be accomplished by the various storage solutions provisioned by the project partners. After the project's life, all of the data shared as part of the ORDP will be conserved in the chosen Open Access repository. Additionally, all of the archiving processes of the data will be documented to ensure the correct handling and reuse of the data after the end of the project.

Consortium Statement. Long-Term Archiving and Preservation.

The IoT-NGIN consortium will follow the "*Rules of Good Scientific Practice*" and ensure that the primary research data produced is preserved and shared beyond the lifetime of the project through the use of external, openly-accessible and centralized data repositories, such as Zenodo. Additionally, the creation of a project-specific repository will be considered if the need arises.

[8] Zenodo - Research. Shared.: https://zenodo.org/
[9] Rules of Good Scientific Practice: https://www.mpg.de/16404553/rules-scientific-practice.pdf

2.7. Allocation of resources

All costs stemming from the participation in the ORDP, including but not limited to data preparation for archiving, data security and sharing mechanisms and the creation of new open repositories, are eligible for reimbursement as covered by the conditions of Article 6, and in particular Article 6.2.D.3 of the Grant Agreement[10]. Consortium partners incurring such costs related to making data accessible, in particular to third parties beyond the consortium, will be responsible to apply for reimbursement of these costs.

Consortium Statement. Allocation of Resources.

The IoT-NGIN consortium recognizes that all costs stemming from the participation in the ORDP are eligible for reimbursement as covered by the conditions of Article 6 and Article 6.2.D.3 of the Grant Agreement. The consortium also plans to use existing, freely-accessible repositories for the long-term archiving of data beyond the project's lifetime with the goal of minimizing the costs of this archiving or eliminating them all together.

3. Data Summary

This section includes a brief description of the main datasets identified so far by the IoT-NGIN consortium partners as relevant to the tasks of the project. The datasets focus on the living labs and trials that will take part in the project, and in particular those that are part of WP7:

- **Trial 2 - Human-Centered Twin Smart Cities Living Lab (T7.2):** aims to adopt an innovative cross-border-by-default twin city context between the cities of Helsinki, Finland and Tallinn, Estonia, focusing on Mobility as a Service (Maas).
- **Trial 3 - Smart Agriculture IoT Living Lab (T7.3):** aims to combine multiple sources of data (ground micro-climate/soil/leaf information stations, drones, mobile robots and wearable devices) to enable the assertion of the evolution of the crop, detect diseases, optimize irrigation and fertilization, reduce spraying and support manual fruit harvesting.
- **Trial 4 and Trial 5 - Industry 4.0 Use Cases and Living Lab (T7.4):** aims to validate the IoT-NGIN federation framework in two real-world Industry 4.0 scenarios: (1) a BOSCH factory in Barcelona, Spain that manufactures car electromechanical brake boosters (Trial 4), and (2) an ABB factory in Helsinki, Finland that manufactures electrical accessories and electrical equipment (Trial 5).

[10] AGA - Annotated Model Grant Agreement, Article 6.2.D.3 - Costs of other goods and services: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf#page=87

- **Trial 6 - Smart Energy Grid Active Monitoring / Control Living Lab (T7.5):** aims to implement a smart energy grid pilot in Terni, Italy to demonstrate the capabilities of smart grid asset performance management and create human-centered micro-contracts and micro-payments in a fully distributed energy marketplace.

As for **Trial 1 - IoT-NGIN Integration Infrastructure Technology Lab**, and **Trial 7 - IoT-NGIN Technology and Living Labs Federation**, they are both focused on integrating the results and services developed in the other trials and are not expected to include any additional data collection. Thus, the datasets described in this section and throughout the deliverable do not include any that are collected or produced by these two trials.

3.1. Dataset description

We provide below a brief description of the datasets that have been identified by the consortium partners as relevant to the different trials mentioned above. For a detailed description of the datasets, including aspects such as access level, data origin, data purpose, among others, we refer the reader to *Annex I. Data Inventory Table*.

Note that this is a preliminary list of datasets that is likely to be extended during the project as the requirements of the trials are further defined and other relevant datasets are identified or generated. These updates will be included in future versions of the DMP or other corresponding deliverables as discussed in *Section 5*.

[T7.2] Trial 2 – Human-Centered Twin Smart Cities Living Lab

2.1 **On-location camera feed:** on-location cameras will be installed to monitor the movement of vehicles and pedestrians on the site of the living lab. The produced data will be aggregated to count the number of the different types of objects passing through the living lab location at a given time.

2.2 **Radar imagery:** radar imagery used to detect cars in a given area and count the totals of different types of vehicles passing through an area at a given time.

2.3 **Social network feeds:** the social network feeds of the volunteers in the living lab will be collected to retrieve routes, commute times and features of the provided commuting solution.

[T7.3] Trial 3 – Smart Agriculture IoT Living Lab

3.1 **AGLV sensor measurements:** measurements acquired from sensors installed and configured specifically for the trial on Automated Guided Land Vehicles (AGLVs) will be collected. The data will be used to allow the use of the AGLVs as carrier machines, and to enable them to locate and avoid workers and trees.

3.2 **Field sensor measurements:** data collected from SynField[11] IoT platform and integrated sensors, including micro-climate data (air temperature, air humidity, wind direction, wind speed, rain volume, rain intensity) and soil and crop data (leaf wetness, soil type, soil temperature, soil humidity, soil conductivity). This dataset will be used to calculate the crop growing degree days (ripening indicator).

3.3 **Drome camera images:** images collected from multi-spectral cameras on drones. The images will be associated with time information and geospatial/location information provided by GPS.

[T7.4] Trial 4 and Trial 5 – Industry 4.0 Use Cases Living Lab

4.1 **Indoor surveillance camera videos:** video images taken from surveillance cameras installed on the ceilings of the factories where the trials will take place. The cameras record one of the working areas and capture the movement of workers and Automated Guided Vehicles (AGVs) while working in the factory.

4.2 **Workers' indoor trajectories:** data derived from the dataset 4.1 about the location of workers at different time intervals inside the factories participating in the trials. The data will be in the form of positioning data with the X and Y coordinates of the factory workers.

4.3 **UWB raw data:** signals collected through Ultra-Wideband (UWB) units installed on AGVs at the factories participating in the trials. These signals will be sent to UWB beacons installed at the factories, which will in turn register the Time of Flight (ToF) of the received signals for further processing.

4.4 **Indoor AGV trajectories:** data derived from the dataset 4.3 about the location of AGVs at different intervals inside the factories participating in the trials. The data will be in the form of positioning data with the X and Y coordinates of the AGVs.

[T7.5] Trial 6 – Smart Energy Grid Active Monitoring / Control Living Lab

5.1 **Smart Meter eXtension (SMX) dataset:** measurements of voltage, currents, and power derived by ASM energy units, collected through an MQTT protocol via public IP and / or LAN/VPN connections. To ensure secure access to the dataset, credentials will be required.

5.2 **Power Quality Analyzers (PQA) dataset:** measurements of voltage, currents, and power derived by ASM energy units, collected through HTTP protocol only via LAN/VPN connections, which provide built-in security access features.

[11] SynField - The path toward Smart Agriculture: https://www.synfield.gr/

5.3 **Phaser Measurement Units (PMU) dataset:** measurements of voltage, currents, and power derived by ASM energy units, collected through HTTP protocol only via LAN/VPN connections, which provide built-in security access features.

5.4 **Charging station data:** real-Time and historical data collected from the charging stations deployed in Terni, Italy in preparation for the trial. The data will include information about the charging station and each charging session, without any identifying information about the use of the station.

5.5 **Electric vehicle data:** real-Time and historical data collected from electric vehicles deployed in the trial site (Terni, Italy). The data will include battery capacity, power and life, along with data about the movement and trajectories of the vehicles, such as latitude, longitude and speed.

4. FAIR Data Considerations

In 2016, the *"FAIR Guiding Principles for scientific data management and stewardship"*[12] were published in *Scientific Data*[13]. The authors intended to provide guidelines to improve the **F**indability, **A**ccessibility, **I**nteroperability, and **R**euse of digital assets. The principles emphasize machine-actionability, which is the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention).

As part of the Open Research Data Pilot, IoT-NGIN will adopt the FAIR principles for all datasets that will be made available to third parties beyond the consortium. The rest of this section discusses the processes and provisions that the consortium will put in place to ensure compliance with each of the FAIR principles. Moreover, *Annex II. FAIR Data Considerations* contains a detailed listing of all relevant aspects to FAIR data for each of the datasets described in *Section 3*.

4.1. Making IoT-NGIN data findable, including provisions for metadata

IoT-NGIN will deposit all datasets that can be shared with third parties in a centralized repository accessible on the web. The consortium will decide on which repository to use at a later stage in the project, and will utilize available directories, such as OpenDOAR[14], to choose the most adequate repository for the datasets that will be published.

[12] FAIR Principles: https://www.go-fair.org/fair-principles/
[13] Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al. The FAIR Guiding Principles for scientific data management and stewardship*. *Sci Data* 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
[14] OpenDOAR: https://v2.sherpa.ac.uk/opendoar/

The most prominent and consolidated option currently being studied is Zenodo, an Open Data portal maintained by OpenAIRE and CERN that allows researchers to upload any type of data (publications, datasets, videos, presentations, posters, etc.) and provides tools to link them together. It also follows the FAIR data principles by design, making it an excellent option for IoT-NGIN.

On one hand, Zenodo assigns a DOI to every published record including datasets, making all records referenceable and findable with a universal unique identifier. Furthermore, Zenodo provides a rich set of metadata to describe the datasets (e.g. contributor, owner, keywords, license, version, language, etc.), which increase the findability of the data. Finally, all metadata in Zenodo is indexed and searchable directly in Zenodo's search engine and through APIs immediately after publishing.

Another option currently considered by the consortium is to publish datasets on a custom-built CKAN repository[15], if no suitable portal is found. In this case, CKAN also assigns a unique URL identifier to each published dataset and has a rich metadata schema to describe the datasets (e.g. title, groups, tags, license, etc.)[16]. CKAN also provides a search engine that allows quick keyword search as well as faceting by tags and browsing between related datasets, which increases the findability of data when the unique identifier is unknown.

Consortium Statement. Making IoT-NGIN data Findable.

The IoT-NGIN consortium will publish all datasets that can be made accessible in Zenodo, another Open Data repository, or create its own CKAN repository if needed. Regardless of the chosen repository, the following guidelines will be implemented to ensure that IoT-NGIN data is findable:

- Each published dataset will be assigned a **unique identifier**. The identifier will either be a DOI or a unique URL where the dataset can be directly accessed.
- Each dataset will be described by a rich schema of **metadata**, that will at least include the following, along with any other domain-specific metadata that are deemed necessary:
  - Title
  - Description
  - Publisher / owner
  - License
  - Keywords / tags
  - Version
  - Language
  - Publication date
- To increase the findability of the data, the consortium will prioritize Open Data portals and repositories with built-in indexing and **search capabilities** that allow third parties to search for relevant datasets based on specific search criteria.

[15] CKAN - The Open Source Data Portal Software: https://ckan.org/
[16] Metadata - CKAN: https://ckan.org/portfolio/metadata/

4.2. Making IoT-NGIN data openly accessible

The IoT-NGIN consortium is committed to making the wide variety of data produced, collected and processed during the living labs and trials of the project freely accessible to third parties. As such, partners responsible for each trial will deposit the datasets in an online, freely-accessible repository, such as Zenodo. The repository of choice will be decided at a later stage of the project, as discussed in *Section 4.1*.

By default, all datasets generated in the trials and living labs will be made available for open access, with the following exceptions:

- Datasets for which one or more partners hold **Intellectual Property rights**, as outlined in the Consortium Agreement. This includes any pre-existing datasets that the partners have obtained prior to the project's start, and that might be used during the execution of the project.
- Datasets that contain **personal information protected by the GDPR** and other relevant regulations. The consortium will decide on a case-by-case basis whether these datasets are reusable by third parties, and will release processed versions of the same after going through proper security checks, aggregation and / or anonymization processes.
- Datasets for which a consortium partner decides to exercise their right to a **partial opt-out of the ORDP**, in line with the *Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020*, and as previously highlighted in *Section 2.1* of this deliverable.

Table 1 summarizes the expected levels of accessibility to the datasets described in *Section 3.1*. Note that any changes to these levels will be communicated in future versions of this DMP or other relevant project deliverables.

Table 1. Expected levels of accessibility of IoT-NGIN datasets.

Dataset Number	Dataset Name	Open / Restricted	Reasons for Restriction
[T7.2] Trial 2 - Human-Centered Twin Smart Cities Living Lab			
2.1	On-location camera feed	Restricted / Open	GDPR - Open after anonymization
2.2	Radar imagery	Restricted / Open	GDPR - Open after anonymization
2.3	Social network feeds	Restricted	GDPR

[T7.3] Trial 3 - Smart Agriculture IoT Living Lab			
3.1	AGLV sensor measurements	Open	
3.2	Field sensor measurements	Restricted / Open	IPR - Open after aggregation
3.3	Drone camera images	Restricted / Open	IPR / GDPR - Open after anonymization
[T7.4] Trial 4 and Trial 5 - Industry 4.0 Use Cases Living Lab			
4.1	Indoor surveillance camera videos	Restricted	GDPR
4.2	Workers' indoor trajectories	Restricted / Open	GDPR - Open after anonymization
4.3	UWB raw data	Open	
4.4	Indoor AGV trajectories	Open	
[T7.5] Trial 6 - Smart Energy Grid Monitoring / Control Living Lab			
5.1	Smart Meter eXtension (SMX) dataset	Restricted	IPR / data security and privacy
5.2	Power Quality Analyzers (PQA) dataset	Restricted	IPR / data security and privacy
5.3	Phaser Measurement Units (PMU) dataset	Restricted	IPR / data security and privacy
5.4	Charging station data	Restricted	IPR / data security and privacy
5.5	Electric vehicle data	Restricted	IPR / data security and privacy

Given that the consortium is planning on using a certified Open Access repository to publish these datasets, all documentation and tools necessary to access them will be available through the repository's website or data portal itself. In case a complementary project repository is implemented, relevant documentation will be produced by the consortium as well. Finally, the licenses under which these datasets will be made available are further discussed in *Section 4.4*.

Consortium Statement. Making IoT-NGIN data Accessible.

The IoT-NGIN consortium is committed to publishing the datasets collected, produced and processed during the project's trials and living labs in a consolidated Open Access repository. Notwithstanding, the consortium will take into consideration the following exceptions to exercise a partial opt-out on some of the datasets and withhold them from public access:

- **Intellectual Property Rights:** held by one or more of the consortium partners as per the Consortium Agreement.
- **Personal Data Protection:** pursuant to the GDPR 2016/679 and other relevant regulations.
- **Other opt-out reasons:** under the petition of a partner of the consortium, as highlighted in the *Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020*:
  - Participation is incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited,
  - Participation is incompatible with the need for confidentiality in connection with security issues,
  - Participation is incompatible with rules on protecting personal data,
  - Participation would mean that the project's main aim might not be achieved,
  - The project will not generate / collect any research data, or
  - Any other legitimate reasons explained at the proposal stage or later on during the project.

The repository of choice will be decided at a later stage during the project and discussed in future versions of the DMP. All documentation and tools relevant to accessing the data will either be provided by the chosen repository directly, or produced by the consortium in case a new repository is created.

4.3. Making IoT-NGIN data interoperable

When it comes to datasets, interoperability has a significant relation with which metadata standards and vocabularies are used to describe the data. Thus, and given that the data produced in IoT-NGIN will come from a variety of fields, we will focus our efforts on making that data interoperable at the level of metadata that will be provided. Details on which standards of formats, types and vocabularies will be used in the datasets themselves will be provided in future versions of this DMP or other trial-specific deliverables.

That being said, the used metadata standards and vocabularies depend on the repository chosen to publish IoT-NGIN data. In the case of Zenodo, the provided metadata is compliant with DataCite's Metadata Schema's minimum as well as recommended terms[17], and the metadata can be exported using other standard formats such as Dublin Core[18] or

[17] DataCite Schema: https://schema.datacite.org/
[18] Dublin Core Metadata Initiative: http://dublincore.org/

MARCXML[19]. On the other hand, metadata provided by a potential CKAN repository is mapped to the Data Catalogue Vocabulary (DCAT) ontology[20], a widely used standard for metadata terms of datasets.

Consortium Statement. Making IoT-NGIN data Interoperable.

The IoT-NGIN consortium will promote the use of standard vocabularies and terms for the metadata of all datasets made accessible, most prominently the Dublin Core Metadata Initiative, the DataCite Metadata Schema, and the Data Catalogue Vocabulary (DCAT) ontology. In cases where this is not possible, mappings between the used metadata terms and one or more of these standards will be provided.

4.4. Increasing IoT-NGIN data reuse (through clarifying licenses)

The data produced through the IoT-NGIN trials and living labs spans multiple state-of-the-art research fields (Industry 4.0, Smart Mobility, Smart Agriculture, etc.) and will be highly valuable to practitioners of these fields outside of the consortium. Thus, the datasets that will be published by the consortium, as summarized in *Table 1*, will be presented to the scientific and research communities with licenses as broad as possible in order to increase their reuse.

Given the different requirements and needs of each trial and consortium partner, it might not be possible for all datasets to be made available under the same license. *Table 2* lists some of the most common Open Access licenses that will be considered for the datasets published by the IoT-NGIN consortium. The decision on which license is used for each dataset corresponds to the partner that generates or collects the data, and will be defined in future versions of this DMP or other corresponding deliverables.

Table 2. Open Access licenses considered for IoT-NGIN datasets.

License	Publisher	Description
ODbL ODC Open Database License	Open Data Commons (ODC) https://opendatacommons.org/licenses/	A *share-alike* license agreement intended to allow users to freely share, modify, and use a dataset while maintaining this same freedom for others.
ODCaL ODC Attribution License		A less restrictive version of ODbL license that only requires users to attribute the owner of the data they plan to utilize, without forcing them to share it in the same free manner.

[19] MARC in XML: https://www.loc.gov/marc/marcxml.html
[20] Data Catalogue Vocabulary (DCAT) - Version 2: https://www.w3.org/TR/vocab-dcat-2/

PDDL ODC Public Domain Dedication and License		A public-domain-equivalent license that acts as a waiver and makes the dataset usable by anyone without any conditions whatsoever.
CC BY Attribution only	Creative Commons (CC) https://creativecommons.org/licenses/	Lets others distribute, remix, adapt, and build upon the licensed work, even commercially, as long as they credit the license holder for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.
CC BY-SA Attribution-ShareAlike		Lets others remix, adapt, and build upon the licensed work even for commercial purposes, as long as they credit the license holder and license their new creations under identical terms. Recommended for materials that would benefit from incorporating content from similarly licensed projects.
CC BY-ND Attribution-NoDerivs		This license lets others reuse the licensed work for any purpose, including commercially, as long as they credit the license holder. However, the work cannot be shared with others in any adapted, modified or built-upon form.
CC BY-NC Attribution-Non Commercial		Lets others remix, adapt, and build upon the licensed work only non-commercially. Furthermore, although their new works must also acknowledge the license holder and be non-commercial, they do not have to license their derivative works on the same terms as the original.
CC BY-NC-SA Attribution-Non Commercial-ShareAlike		Lets others remix, adapt, and build upon the licensed work non-commercially, as long as they credit the license holder and license their new creations under identical terms.
CC BY-NC-ND Attribution-Non Commercial-NoDerivs		This license is the most restrictive of the CC licenses, only allowing others to download the licensed work and share them with others as long as they credit the license holder, but they cannot change them in any way or use them commercially.

Although currently not foreseen for any of the identified datasets, partners reserve the right to impose an embargo period on a particular dataset during or after the project execution if it is deemed necessary to allow for time to publish or seek patents or other IP rights. Such embargo periods will also be identified and determined in future versions of this DMP or other corresponding deliverables.

Consortium Statement. Increasing IoT-NGIN data Reuse.

The IoT-NGIN consortium commits to licensing datasets made available to third parties under the most permissive licenses possible, taking into consideration the IPRs of consortium partners as defined in the Consortium Agreement. The consortium partners will also put in place rigorous quality assurance processes to guarantee that the published data is of high quality and maximize its reuse.

5. DMP review process and timetable

This deliverable provides an initial version of the Data Management Plan of IoT-NGIN and will operate as a manual on the data policy of the project. It contains preliminary information about the data that the project will collect, generate or process through its multiple living labs and trials, including whether and how it will be made accessible to third parties, as identified at this early stage of the project.

However, the requirements, needs and conditions to the DMP are expected to change during the lifetime of the project, in particular through events such as:

- Incorporation of new datasets,
- Changes to the description, details, or conditions of existing datasets,
- Changes in the IPRs or data access policies of the consortium or one of the partners,
- Unlikely changes in the composition of the consortium and other external factors.

To accommodate the changes stemming from these events or any other that might affect the information described in this deliverable, regular updates will be provided throughout the lifetime of the project. *Table 3* lists the expected versions of the DMP that will be provided throughout the lifetime of IoT-NGIN, in line with the H2020 guidelines on data management plants. In case it is deemed necessary, further intermediate updates to the DMP will be provided through the periodic technical reports of the project to the European Commission.

Table 3. Expected revisions to the IoT-NGIN Data Management Plan.

Version	Deliverable	Delivery date	Description
v1	D7.1 - Data Management Plan (DMP)	M06 - March 2021	Initial DMP version and guidelines on IoT-NGIN data policy

v2	D7.2 - Trial site set-up, initial results and DMP update	M18 - March 2022	Review and updates based on the requirements and initial results of the trials
v3	*D7.4 - IoT-NGIN Living Labs use cases Assessment and Replication guidelines*	*M36 - September 2023*	*Final review and updates based on final trial results (only if needed)*

6. Conclusions

This document provided the first version of the IoT-NGIN Data Management Plan. The DMP is anchored in participation in H2020's Open Research Data Pilot, the adoption of FAIR data principles and the adherence to Intellectual Property Rights and GDPR regulations. An initial list of datasets identified as relevant to the project's trials and living labs was also provided.

Next steps will focus on studying available Open Access repositories and choosing which will be used for the published data, and whether a new project repository is needed or not. Once a repository is chosen, the metadata provided with the published datasets will be specified and priority will be given to standard vocabularies such as DataCite, Dublin Core and DCAT. In parallel, the consortium partners will decide on the licenses under which the data will be published, taking into consideration their IPR and the policies of the chosen repository.

The next version of the DMP is expected by M18 of the project (March 2022). It will cover, along with the points mentioned above, any new identified datasets as well as any modifications to the details of the datasets already mentioned in this first version. Other aspects that might be covered include further details or modifications to the processes specified in this DMP, such as data security procedures, data sharing mechanisms between the consortium partners, and data anonymization and aggregation techniques for protected personal data.

Annex I. Data Inventory Table

Data Inventory Table								
Dataset no.	Dataset Name	Data Description	Data Controller	Data Format	Data Size	Data Origin[21]	Data Purpose	Ethical Issues[22]
[T7.2] Trial 2 - Human-Centered Twin Smart Cities Living Lab								
2.1	**On-location camera feed**	On-location cameras will be installed to monitor the movement of vehicles and pedestrians in the site of the living lab. The produced data will be aggregated to count the number of the different types of objects passing through the living lab location at a given time.	FVH	Video files, XLS, XML, JSON	≈ KBs	Observational: from on-location cameras.	On-location video analysis (edge computing) to aggregate the number of vehicles and pedestrians in the videos.	Yes
2.2	**Radar imagery**	Radar imagery will be used to detect cars in a given area and count the totals of different types of vehicles passing through an area at a given time.	FVH	Video files, XLS, XML, JSON	≈ KBs	Observational: from radars.	Detection of vehicles in a particular area in order to produce aggregates of the number of detected vehicles.	Yes

[21] **Data Origin** should be one of the following:

- **Observational:** Data captured in real time, often not reproducible (e.g. sensor readings, images, telemetries, sample data).
- **Experimental:** Data from lab equipment, often reproducible, but with high costs (e.g. chromatograms, magnetic fields readings).
- **Simulation:** Data generated by computational models (e.g. climate models, economic models, materials models).
- **Derived / Compiled:** Data coming from analysis or compilation. Reproducible but with high costs (e.g. results of text and data mining, compiled databases).
- **Reference / Canonical:** Collection or conglomeration of smaller datasets published and curated (e.g. chemical structures, gene sequence databanks, spatial data portals).

[22] **Ethical Issues?** a (Y/N) answer indicating whether the data has any potential ethical issues, particularly related to protected personal data.

Data Inventory Table								
Dataset no.	Dataset Name	Data Description	Data Controller	Data Format	Data Size	Data Origin[21]	Data Purpose	Ethical Issues[22]
2.3	**Social network feeds**	The social network feeds of the volunteers in the living lab will be collected to retrieve routes, commute times and features of the provided commuting solution.	FVH	Text (DOC, PDF), XLS	≈ KBs	Derived: from social network feeds of volunteer participants.	Analysis of routes, commute times and features of the provided commuting solution.	Yes
[T7.3] Trial 3 - Smart Agriculture IoT Living Lab								
3.1	**AGLV sensor measurements**	Measurements acquired from sensors installed and configured specifically for the trial on Automated Guided Land Vehicles (AGLVs) will be collected. The data will be used to allow the use of the AGLVs as carrier machines, and to enable them to locate and avoid workers and trees.	OPT	CSV, JSON	≈ MBs per month	Observational: from AGLV sensors.	Analysis of measurements from AGLV sensors to aid in the harvesting of crops within the trial site(s).	No
3.2	**Field sensor measurements**	Data collected from SynField IoT platform and integrated sensors, including micro-climate data (air temperature, air humidity, wind direction, wind speed, rain volume, rain intensity) and soil and crop data (leaf wetness, soil type, soil temperature, soil humidity, soil conductivity). This dataset will be used to calculate the crop growing degree days (ripening indicator).	SYN	CSV, JSON	≈ 300 KBs per day per SynField node	Observational: from field sensors.	Conduct tests related to smart irrigation and precision aerial spraying within the trial site(s).	No

Data Inventory Table								
Dataset no.	Dataset Name	Data Description	Data Controller	Data Format	Data Size	Data Origin[21]	Data Purpose	Ethical Issues[22]
3.3	**Drone camera images**	Images collected from multi-spectral cameras on drones. The images will be associated with time information and geospatial/location information provided by GPS.	SYN	Image (JPEG, TIFF)	≈ 500 MBs per day per drone	Observational: from multi-spectral drone cameras.	Collect images of crop leaves at the trial site(s) to test crop disease prediction.	Yes
[T7.4] Trial 4 and Trial 5 - Industry 4.0 Use Cases Living Lab								
4.1	**Indoor surveillance camera videos**	Video images taken from surveillance cameras installed on the ceilings of the factories where the trials will take place. The cameras record the working areas and capture the movement of workers and Automated Guided Vehicles (AGVs) while working in the factory.	BOSCH / I2CAT	Image files	≈ 10,000 HD images	Observational: from indoor surveillance cameras.	Train and validate computer vision algorithms in detecting workers throughout the shop floor of the trial site based on the collected images.	Yes
4.2	**Workers' indoor trajectories**	Data derived from dataset 4.1 about the location of workers at different time intervals inside the trial factories. The data will be in the form of positioning data with the X and Y coordinates of the factory workers.	BOSCH / I2CAT	CSV	≈ 1 GB (17M position points)	Derived: from analyzing images of dataset 4.1.	Analyze the trajectories of workers within the shop floor of the trial site to predict and avoid AGV-human collisions.	Yes
4.3	**UWB raw data**	Signals collected through Ultra-Wideband (UWB) units installed on AGVs at the factories participating in the trials. These signals will be sent to UWB beacons installed at the factories, which will in turn register the Time of Flight (ToF) of the received signals for further processing.	BOSCH / I2CAT	CSV	≈ 2 GB (100M UWB signals)	Observational: from UWB units installed on AGVs.	Collect positions of AGVs at different points in time throughout the shop floor / working area of the trial sites.	No

Data Inventory Table								
Dataset no.	Dataset Name	Data Description	Data Controller	Data Format	Data Size	Data Origin[21]	Data Purpose	Ethical Issues[22]
4.4	**Indoor AGV trajectories**	Data derived from the dataset 4.3 about the location of AGVs at different intervals inside the factories participating in the trials. The data will be in the form of positioning data with the X and Y coordinates of the factory AGVs.	BOSCH / I2CAT	CSV	≈ 2 GB (100M UWB signals)	Derived: from analyzing UWB signals in dataset 4.3.	Analyze the trajectories of AGVs within the shop floor of the trial site to predict and avoid AGV-AGV and AGV-human collisions.	No
[T7.5] Trial 6 - Smart Energy Grid Monitoring / Control Living Lab								
5.1	**Smart Meter eXtension (SMX) dataset**	Measurements of voltage, currents, and power derived by ASM energy units, collected through an MQTT protocol via public IP and / or LAN/VPN connections. To ensure secure access to the dataset, credentials will be required.	ASM	JSON	≈ 1 MB per day	Observational: from ASM energy units.	Provide timely alarms when the system approaches unstable operational boundaries by collecting measurements that indicate the health of the grid.	No
5.2	**Power Quality Analyzers (PQA) dataset**	Measurements of voltage, currents, and power derived by ASM energy units, collected through HTTP protocol only via LAN/VPN connections, which provide built-in security access features.	ASM	CSV / JSON	≈ 1 KB per day	Observational: from ASM energy units.	Provide timely alarms when the system approaches unstable operational boundaries by collecting measurements that indicate the health of the grid.	No
5.3	**Phaser Measurement Units (PMU) dataset**	Measurements of voltage, currents, and power derived by ASM energy units, collected through HTTP protocol only via LAN/VPN connections, which provide built-in security access features.	ASM	JSON	≈ 1 MB per day	Observational: from ASM energy units.	Provide timely alarms when the system approaches unstable operational boundaries by collecting measurements that indicate the health of the grid.	No

Data Inventory Table								
Dataset no.	Dataset Name	Data Description	Data Controller	Data Format	Data Size	Data Origin[21]	Data Purpose	Ethical Issues[22]
5.4	**Charging station data**	Real-time and historical data collected from the charging stations deployed in the trial site (Terni, Italy). The data will include information about the charging station and each charging session, without any identifying information about the use of the station.	EMOT	JSON	≈ KBs per day	Observational: from EMOT charging stations.	Conducting driver-friendly, dispatchable charging of EVs based on energy demand-response with human-centered micro-contracts and micro-payments.	No
5.5	**Electric vehicle data**	Real-Time and historical data collected from electric vehicles deployed in the trial site (Terni, Italy). The data will include battery capacity, power and life, along with data about the movement and trajectories of the vehicles, such as latitude, longitude and speed.	EMOT	JSON	≈ KBs per day	Observational: from EMOT electric vehicles.	Conducting driver-friendly, dispatchable charging of EVs based on energy demand-response with human-centered micro-contracts and micro-payments.	No

Annex II. FAIR Data Considerations

FAIR Data Considerations									
Dataset no.	Dataset Name	Findability		Accessibility			Interoperability	Reusability	
		Metadata Description	Permanent Identifiers	Storage Medium	Access Level[23]	Access Methods	Standard Vocabularies	IP / License	Access / Reuse Restrictions[24]
[T7.2] Trial 2 - Human-Centered Twin Smart Cities Living Lab									
2.1	**On-location camera feed**	*TBD*	-	FVH / AALTO secured servers (internal) Open Access repository (external)	Restricted / Open	API and / or file sharing (internal) Open Access repository (external)	-	CC	GDPR - Open after anonymization and after end of project
2.2	**Radar imagery**	*TBD*	-	FVH / AALTO secured servers (internal) Open Access repository (external)	Restricted / Open	API and / or file sharing (internal) Open Access repository (external)	-	CC	GDPR - Open after anonymization and after end of project
2.3	**Social network feeds**	*TBD*	-	FVH / AALTO secured servers (internal) Open Access repository (external)	Restricted	API and / or file sharing (internal) Open Access repository (external)	-	CC	GDPR - sensitive personal information
[T7.3] Trial 3 - Smart Agriculture IoT Living Lab									

[23] **Access Level** should be one of the following:

- **Open:** the data is made publicly available to third parties, with or without restrictions (as indicated in the **Access / Reuse Restrictions** column).
- **Restricted:** the data is only available to members of the IoT-NGIN consortium, and the EU Commission for revision and evaluation purposes if needed.

[24] **Access / Reuse Restrictions** in the case of data with Restricted access, indicates one or more of the following: (1) limitations and restrictions to **access** the data, and if they are linked to a specific timeframe. (2) Restrictions on the **reuse** of the data by third-parties after the end of the project (e.g. confidentiality agreements) and if any embargo period is required.

FAIR Data Considerations									
Dataset no.	Dataset Name	Findability		Accessibility			Interoperability	Reusability	
		Metadata Description	Permanent Identifiers	Storage Medium	Access Level[23]	Access Methods	Standard Vocabularies	IP / License	Access / Reuse Restrictions[24]
3.1	**AGLV sensor measurements**	Location, timestamp	-	OPT IoT platform database (internal) Open Access repository (external)	Open	API (internal) Open Access repository (external)	-	*TBD*	Open after end of project
3.2	**Field sensor measurements**	Location, timestamp, SynField node	*TBD*	SynField IoT platform database (internal) Open Access repository (external)	Restricted / Open	API and / or file sharing (internal) Open Access repository (external)	*TBD*	ODC	IPR - Open after aggregation and after end of project
3.3	**Drone camera images**	Location, timestamp	*TBD*	SYN cloud services (internal) Open Access repository (external)	Restricted / Open	API and / or file sharing (internal) Open Access repository (external)	*TBD*	ODC	IPR / GDPR - Open after anonymization and after end of project
[T7.4] Trial 4 and Trial 5 - Industry 4.0 Use Cases Living Lab									
4.1	**Indoor surveillance camera videos**	Timestamp, number of AGVs, number of workers	-	Secure cloud file server (internal)	Restricted	File sharing (internal)	*TBD*	ODC	GDPR - sensitive personal information
4.2	**Workers' indoor trajectories**	*TBD*	-	Secure cloud file server (internal) Open Access repository (external)	Restricted / Open	File sharing (internal) Open Access repository (external)	*TBD*	ODC	GDPR - Open after anonymization and after end of project

FAIR Data Considerations									
Dataset no.	Dataset Name	Findability		Accessibility			Interoperability	Reusability	
		Metadata Description	Permanent Identifiers	Storage Medium	Access Level[23]	Access Methods	Standard Vocabularies	IP / License	Access / Reuse Restrictions[24]
4.3	**UWB raw data**	Timestamp, UWD beacon position, AGV position	-	Secure cloud file server (internal) Open Access repository (external)	Open	File sharing (internal) Open Access repository (external)	*TBD*	ODC	Open after end of project
4.4	**Indoor AGV trajectories**	*TBD*	-	Secure cloud file server (internal) Open Access repository (external)	Open	File sharing (internal) Open Access repository (external)	*TBD*	ODC	Open after end of project
[T7.5] Trial 6 - Smart Energy Grid Monitoring / Control Living Lab									
5.1	**Smart Meter eXtension (SMX) dataset**	*TBD*	-	ASM servers (internal)	Restricted	Authorized credentials through VPN (internal)	-	-	IPR / data security and privacy
5.2	**Power Quality Analyzers (PQA) dataset**	*TBD*	-	ASM servers (internal)	Restricted	Authorized credentials through VPN (internal)	-	-	IPR / data security and privacy
5.3	**Phaser Measurement Units (PMU) dataset**	*TBD*	-	ASM servers (internal)	Restricted	Authorized credentials through VPN (internal)	-	-	IPR / data security and privacy

FAIR Data Considerations									
Dataset no.	Dataset Name	Findability		Accessibility			Interoperability	Reusability	
		Metadata Description	Permanent Identifiers	Storage Medium	Access Level[23]	Access Methods	Standard Vocabularies	IP / License	Access / Reuse Restrictions[24]
5.4	**Charging station data**	*TBD*	-	EMOT servers (internal)	Restricted	API and / or file sharing (internal)	-	-	IPR / data security and privacy
5.5	**Electric vehicle data**	*TBD*	-	EMOT servers (internal)	Restricted	API and / or file sharing (internal)	-	-	IPR / data security and privacy