Horizon 2020

Societal Challenge: Improving the air quality and reducing the carbon footprint of European cities

Project: 690105 – ICARUS

Full project title:

Integrated Climate forcing and Air pollution Reduction in Urban Systems

D9.1 Data Management Plan

WP 9 Management

Lead beneficiary: AUTH

Date: January 2021

Nature: Report

Dissemination level: Public

TABLE OF CONTENTS

Document Information

Grant Agreement Number	690105	**Acronym**	ICARUS
Full title	Integrated Climate forcing and Air pollution Reduction in Urban Systems		
Project URL			
Project Officer	Richard Tavares		

Delivery date	**Contractual**	October 2016	**Actual**	January 2021
Status	Draft □		Final fifth version √	
Nature	Demonstrator □	Report √	Prototype □	Other □
Dissemination level	Confidential □	Public √		

Responsible Author (Partners)	**Aristotle University of Thessaloniki (AUTH)**			
Responsible Author	Denis Sarigiannis		**Email**	sarigiannis@auth.gr
	Partner	AUTH	**Phone**	
Other partners (Institution)	ALL			

Document History

Name (Institution)	Date	Version
AUTH	October 2016	First draft
AUTH	October 2016	Final
AUTH	February 2018	Final second revised
ALL	December 2018	Draft third version
AUTH	December 2018	Final third version
AUTH, ALL	June 2020	Final fourth version
AUTH, ALL	January 2021	Final (fifth) version

1 Executive Summary

This Data Management Plan (DMP) is a continuously updated document that describes already existing and new data generated in the ICARUS project, its type, format and structure, the arrangements for its storage and security, and its potential for being used by others outside of the ICARUS Consortium. The structure of this DMP is based on the EC's *Guidelines on Data Management in Horizon 2020*[1] following the FAIR principles.

The ICARUS project participates in the Pilot on Open Research Data launched by the European Commission along with the H2020 programme. The use of a Data Management Plan is required for all participating projects and the development of this DMP has been done to facilitate the release of the data generated within the project through storage in research data repositories.

The ICARUS Data Management plan is a document that is submitted to the EU as project deliverable D9.1. It is important to note however that the document evolved and further developed during the project's life cycle. It can be identified by a version number and a date. Updated versions have been uploaded by Aristotle University of Thessaloniki (AUTH), which is the primary responsible for data management. ICARUS partners can forward questions and suggestions, as to (additions to) the contents and use of the data management plan to AUTH, and will be informed when a new version will be uploaded in the ICARUS web site (http://icarus2020.eu).

This document is the fifth and final version of the DMP. It includes updated and final information about the datasets produced by the project according to the FAIR principles to make the data findable, accessible, interoperable and reusable.

[1] EC. (2016). *Guidelines on Data Management in Horizon 2020*. Version 2.1. Brussels: European Commission. Available online at: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

2 Research data types in ICARUS

For this final release of the DMP, the data types produced during the project are focused on the Description of the Action (DoA) and on the results obtained during the project activities.

According to this consideration, the table below reports a list of types of research data that ICARUS produced. A detailed description of each dataset is given in the following sections of this document.

#	Dataset	Lead partner	Related WP(s)
1	Future activity-emission factor matrices for the participating cities	USTUTT	WP2
2	Future activity-emission factor matrices for the whole of Europe	USTUTT	WP2
3	Emission inventory for the participating cities	USTUTT	WP2
4	Emission inventory for the whole of Europe	USTUTT	WP2
5	Life Cycle Data for relevant processes and activities	USTUTT	WP2
6	Atmospheric dispersion modelling results for the participating cities	AUTH	WP3
7	Atmospheric dispersion modelling results for the whole of Europe	AUTH	WP3
8	Environmental measurements dataset	NCSRD	WP3
9	Light Manned Aircraft (LMA) dataset	ARTEMIS	WP3
10	Climate data and climate indicators	CMCC	WP3
11	Multi-sensor data for personal exposure monitoring	AUTH	WP4
12	ICARUS survey data	AUTH	WP4
13	ICARUS TAD data	AUTH	WP4
14	Health effects at the community level	AUTH	WP4/WP5
15	Identification of feasible mitigation and abatement options	USTUTT	WP5
16	Monetary valuation of impacts and cost-benefit analysis of policy options	EXETER	WP5
17	Visions for smart, green and healthy cities	ADDMA/EXETER	WP6
18	Stakeholders List	AUTH	WP8

19	Scientific publications	AUTH, ALL	WP8

Research data linked to exploitable results has not been put into the open domain if they compromise its commercialization prospects or have inadequate protection, which is a H2020 obligation. The rest of research data have been deposited in an Open Access repository as described in Chapter 3 except the Scientific publications which have been stored in the Zenodo repository.

3 ICARUS global data repository

To share data between ICARUS partners, two secure platforms has been established: the ICARUS global open-data repository accessible at ftp://155.207.30.29 where all the project data are stored, and the Sensor Data Platform available at http://icarusdss.upcom.eu:80820/web where only the data collected in the frame of the sensor campaigns are stored as backup copy.

Creating and maintaining a global data repository is very important for the normal evolution of the project, as well as for fulfilling the commitments made in the proposal for FAIR data. To this end, building the repository aims to cover the need for data sharing and data archiving after the project's lifetime and making data available to the public.

The global open-data repository contains all the project data (including the sensor data) and free access is provided (no need for a username/password); the sensor data is also available at the sensor data platform upon free user registration.

3.1.1 Data sharing

During the project's lifetime, a significant amount of collected and generated data needs to be exchanged among the partners of the consortium. Additionally, the repository served data sharing needs with Organizations outside the ICARUS consortium. The repository offers a virtually unique place for data storage and retrieval. Moreover, it offers homogeneity in data formats by applying the standards described in the present document.

3.1.2 Data archiving after project's lifetime

Most of the data generated or collected during the project needs to be persisted and archived even after its completion. This is to enable consortium members, stakeholders and third-party researchers to retrieve and use the data in the future. The repository covers this need by providing an organized structure for data storage. Both the ICARUS global open-data repository and the sensor data platform will remain operational for at least **3 years** after the project's end. The current structure and access rights will remain the same after the project end for at least 3 years unless new data are generated. In that case AUTH and UPCOM will decide on the access rights in collaboration with the Data Owner/Data Provider.

3.1.3 Public availability

It is a commitment of the consortium to make any data collected or generated via the use of public funding as available as possible, unless this is contradictory to personal or corporate privacy requirements or could be harmful to the commercialization of the project. These data must be available to the public during and after the project's lifetime. Another aim is to ease the procedure of data search and retrieval (i.e. making them Findable and Accessible) by archiving them all in the same place and under a clear and concrete structure.

3.2 Architecture

The global open-data repository is used only for storing and archiving data, as well as for making them available to the public. It will not be used as an online database to retrieve data for processing or for storing generated data directly. Consequently, there is no need for it to contain any database management systems.

The global open-data repository has been built using the infrastructure of Aristotle University of Thessaloniki (AUTH). It offers management and data retrieval functionalities and supports data exchange through an FTP server. The sensor data platform has been designed and will be maintained by UPCOM Ltd. A table with the full components of the two platforms of the Global Data Repository can be found below.

Table 1: The ICARUS Global Data Repository

Name	ICARUS Global open-data repository	ICARUS sensor data platform
Protocol Type	FTP	HTTP
URL	ftp://155.207.30.29	http://icarusdss.upcom.eu:8080/web
Accessible via	Any FTP client (Filezilla, WinSCP, etc.)	Any Web Browser (Google Chrome, Firefox, etc.)
Username	None (free access)	Decided upon user registration
Password	None (free access)	Decided upon user registration
Web Interface	Not Available	Available
List of Datasets	All project data	Sensor Campaigns Data
Responsible	AUTH	AUTH, UPCOM

3.2.1 ICARUS Sensor Data Platform - Web interface

The web interface of the ICARUS Sensor Data Platform supports accounts with different access levels. The supported account types are:

- Administrative: Reserved for the project coordinator. Administrators can create, manage and delete user accounts, configure the FTP server and modify the web interface. They are able to access all three components of the web interface, i.e. Administration, Management and Data retrieval. The administrative accounts also handle third parties' requests for data provisioning with or without a simple user account creation.
- Data Management: All project partners that collect, generate or otherwise consume project data must have a data management account. This account type is able to upload data, browse and download any data uploaded by other users and delete data they uploaded themselves. They can also choose whether the data uploaded by them will be publicly or internally available. Data management accounts will have access to Management and Data retrieval components.
- Simple user account: In order to allow third parties to access public project data, the requirement of creating an account has been considered.

3.2.2 Administration

The administration component provides functionality for configuring the FTP server, managing user accounts, granting access to data sets and recovering deleted data.

3.2.3 Management

The management component, accessible to administrative and data management accounts, provides an appropriate interface for uploading data and for choosing whether they are public or internal. It also enables the administrators to rename, move, modify or delete data. It also allows data management account users to rename, modify or delete the data they have uploaded.

3.2.4 Data retrieval

The data retrieval component, accessible to all account types as well as to unsubscribed users, provides functionality for searching data according to various criteria, depending on their type, e.g., for geospatial air concentration data: location, date/time, substance etc.

3.2.5 Access rights control

This component decides whether data can be browsed, uploaded or downloaded in respect to user privileges. It can be configured only by the administrative accounts and it controls all communications between the web interface and the storage component.

3.2.6 Storage

Storage component ensures data is regularly and efficiently backed-up in order to minimize the possibility of data loss due to a hardware failure. Moreover, a data recovery mechanism has been implemented to enable the administrator and/or the data management users retrieve accidentally modified or deleted data. An alternative to the recovery mechanism is the obvious requirement that all partners keep copies of all data they upload to the data repository.

3.2.7 ICARUS Sensor Data Platform - Sensor Campaigns

The ICARUS Sensor Data Platform has been set up to store, process and share the data of the ICARUS Sensor Campaigns among the members of the Consortium. To this end a Sensor Data Platform has been developed, containing all the anonymized data collected by the sensors. Access to this platform is restricted to the Consortium and access is protected by a username and a system-enforced strong password. Authentication and authorization are enforced by Spring Security. The Sensor Data Platform is accessible from the DSS Platform website http://icarusdss.upcom.eu:8080/web/login by a separate role created for this functionality.

The Sensor Data Platform has been developed using the Spring Framework (a framework for building web applications on top of the Java EE (Enterprise Edition) platform. Its architecture is based on the Spring Web MVC (Model-View-Controller), ideal for handling the business and processing logic of the sensor data entities. Data acquisition is done through REST clients that receive HTTP requests from the backend environments of the sensors' producers. The acquired data is stored in Couchbase, a NoSQL database.

There are functionalities allowing the field workers to declare the assignments of devices to anonymized individuals and households for a specific period, as well as to define the correspondence of households to individuals. Researchers can monitor the status of the devices that are transmitting data live (uHoo and IoTech Portable PM sensors) and can export data in CSV/Excel format either on a per-device basis or combined data of all devices.

A more detailed description of the requirements, the architecture and the functionalities of the ICARUS Sensor Data Platform can be found in Deliverable D4.4 *Multi-sensor data collection IT platform for personal exposure monitoring data fusion*.

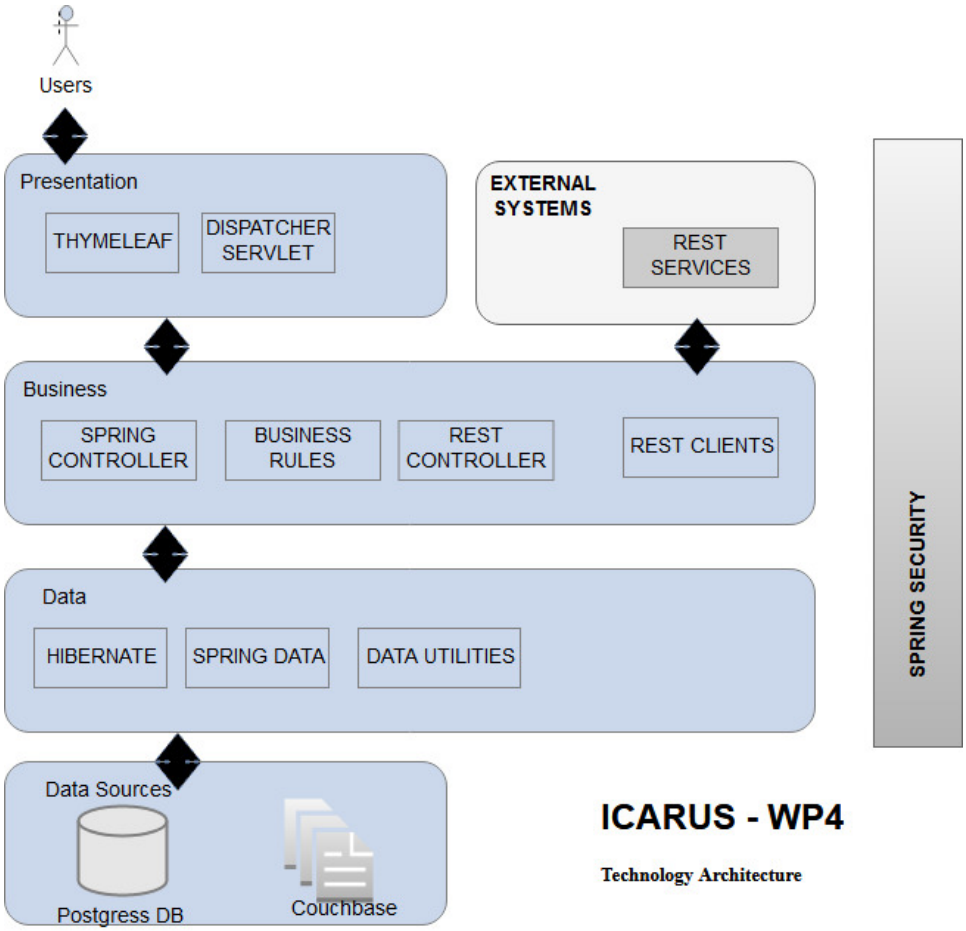


Figure 1: architectural components of the ICARUS Sensor Data Platform

4 Methodology

The methodological approach that has been used for the compilation of D9.1 follows the updated version of the "Guidelines on Data Management in Horizon 2020" released by the European Commission Directorate - General for Research & Innovation which states that research data should be 'FAIR' that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard or implementation-solution.

The DMP covers the complete research data life cycle. It describes the types of research data that has been generated or collected during the project, the standards used, how the research data will be preserved and what parts of the datasets will be shared for verification or reuse (Figure 2).

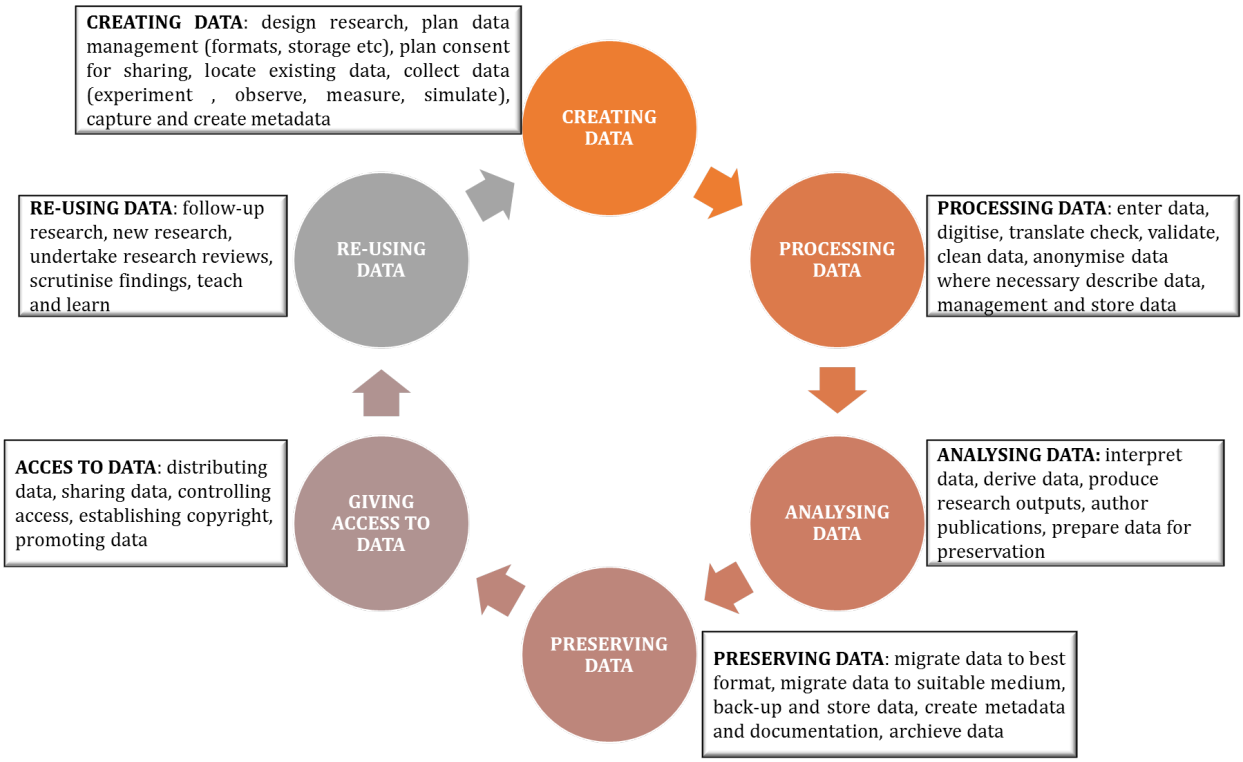


Figure 2: Research data life cycle (adapted from UK data archive https://www.ukdataservice.ac.uk/manage-data/lifecycle)

5 DMP components

Taking into account the proposed methodology and according to the FAIR principles, this chapter describes the following DMP components.

5.1 Data summary

Points to be addressed:

Provide a summary of the data, addressing the following issues

- *State the purpose of the data collection/generation*
- *Explain the relation to the objectives of the project*
- *Specify the types and formats of data generated/collected*
- *Specify if existing data is being re-used (if any)*
- *Specify the origin of the data*
- *State the expected size of the data (if known)*
- *Outline the data utility: to whom will it be useful*

5.1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

The purpose was to collect and produce different datasets across Europe to quantitatively assess the impact of current and alternative national and local policies on reducing greenhouse gas (GHG) emissions and improving air quality leading to the design and implementation of appropriate abatement strategies in European cities. Access to use these data is needed to address the different ICARUS objectives as specified in the grant agreement and summarized below.

1. Quantitatively assess the impact of current and alternative national and local policies on reducing greenhouse gas (GHG) emissions and improving air quality through a full chain approach and evaluate the future public health and well-being impacts of these policies in European cities.
2. Evaluate (using source apportionment and atmospheric modelling) the current contributions of the different pollution sources linked to urban activities including heat and power use in the urban building stock, urban traffic and transportation needs, energy production, industrial activities including energy production, agriculture and trans-boundary pollution with respect to GHG-emissions, air quality loading, public health and well-being of the population.
3. Propose measures of a technological (i.e. measures that will lead to a reduction of emissions at the source) and non-technological (i.e. measures that induce behavioral changes) nature to reduce both the carbon footprint and the air quality burden (win-win solutions). Techno-economic analysis of possible scenarios for the introduction of such measures will result in the definition of cost-effective environmental and climate protection and air quality management plans adapted to the specific needs of different EU cities and regions. The effect of these measures will be evaluated jointly taking into account the socioeconomic drivers related to the existing and projected scenarios.
4. Develop visions of green cities with clean air, close to zero or negative carbon footprint and maximal wellbeing, develop a pathway for the realization of these visions in the next 50 years and propose the first steps down that road in the form of a concrete plan towards achieving these visions in the participating cities.

5. Raise awareness of the citizens about the impacts on public health and the climate causes by their activities or with changes in their activities.

5.1.2 What types and formats of data will the project generate/collect?

5.1.2.1 Types of the data

1. Future activity-emission factor matrices for the participating cities
2. Future activity-emission factor matrices for the whole of Europe
3. Emission inventory for the participating cities
4. Emission inventory for the whole of Europe
5. Life Cycle Data for relevant processes and activities
6. Atmospheric dispersion modelling results for the participating cities
7. Atmospheric dispersion modelling results for the whole of Europe
8. Environmental measurements dataset
9. Light Manned Aircraft (LMA) dataset
10. Climate data and climate indicators
11. Multi-sensor data for personal exposure monitoring
12. ICARUS survey data
13. ICARUS TAD data
14. Health effects at the community level
15. Identification of feasible mitigation and abatement options
16. Monetary valuation of impacts and cost-benefit analysis of policy options
17. Visions for smart, green and healthy cities
18. Stakeholders List
19. Scientific publications

5.1.2.2 Formats of the data:

Several formats have been used on the basis of the data type. Hereinafter are summarized some of the most common formats used. More details are provided in Chapter 6 “Dataset of ICARUS”

- MS Excel compatible files including comma separated, txt and xls(x).
- Standard GIS format (e.g. raster, shapefile)
- Data base format (e.g. accdb)
- Network Common Data Form (NetCDF)
- MS Word (doc) and/or Adobe Portable Document Format (pdf) format

5.1.3 Will you re-use any existing data and how?

Yes, it is encouraged to make existing data available for research. These include for example emission inventory in the ICARUS cities or environmental data measured in the frame of the ICARUS campaigns.

5.1.4 What is the origin of the data?

Dataset generated by the ICARUS project and to lesser extent data from previous research projects in which, for example, climate data or emission data have been collected across Europe.

5.1.5 What is the expected size of the data?

The expected size depends on the extent and the nature of the data that is made available. The estimated size of the ICARUS Global Data Repository is ca. 2.5 TB. The non-relational databases that store the Sensors Campaign Data size is ca. 15GB.

5.1.6 To whom might it be useful ('data utility')?

- ICARUS consortium;
- Public Authorities at National, Regional and Local level responsible for Air quality and Public Health
- City planners, environmental and health professionals
- European Commission services and European Agencies;
- EU National Bodies;
- The general public including the broader scientific community

5.2 FAIR data

Points to be addressed:

In general terms, your research data should be 'FAIR' that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard or implementation-solution.

5.2.1 Making data findable, including provisions for metadata

Points to be addressed:

- *Outline the discoverability of data (metadata provision)*
- *Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?*
- *Outline the approach for clear versioning*

5.2.1.1 Metadata provision

ICARUS data always includes metadata. To ensure the project data is made available using the same standards towards consistency and usability all the datasets provide as far as possible the following metadata according to the DataCite Metadata Schema[2].

- Digital Object Identifier (DOI)

[2] DataCite Website http://schema.datacite.org/

- Publication date
- Title
- Description

5.2.1.2 Standards for metadata creation

See point above.

Spatial data follows the INSPIRE metadata elements for spatial data sets and services[3]

5.2.1.3 Naming conventions used

The ICARUS dataset identification follows the naming: Data_<WPn°>_<serial number of dataset>_<dataset title>. Example: Data_WP2_1_ future activity-emission factor matrices for the participating cities data set.

5.2.1.4 Clear versioning

The versioning management of the data, metadata and in general the files stored into the Repository are applied via the naming convention and the use of the date as suffix, indicating the last version of the file uploaded into the Repository;

5.2.2 Making data openly accessible

Points to be addressed:

- *Specify which data will be made openly available? If some data is kept closed provide rationale for doing so*
- *Specify how the data will be made available*
- *Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?*
- *Specify where the data and associated metadata, documentation and code are deposited*
- *Specify how access will be provided in case there are any restrictions*

5.2.2.1 Data are accessible via ICARUS repository to ICARUS consortium partners

Most data generated in ICARUS has been made openly available. Some exceptions are relevant to the personal data collected in the frame of the sensor's campaigns and to the activity-emission factor matrices which cannot be made available to the public due to privacy and confidentiality issues. Full details are provided in Chapter 6 for each dataset identified.

The ICARUS global open-data repository has been set up for this project as it:

- Is a platform that facilitates sharing of data, intermediate results, and results
- Is needed to enable the analysis of the impact of current and alternative national and local policies on reducing greenhouse gas (GHG) emissions and improving air quality policies, but also of accessory external exposure data and health data to meet the goals of ICARUS.
- Enables data users to work with selected quality-controlled data sets and versions approved by the Data Owners/Data Providers.
- Aims to reach the highest level of GDPR compliancy, amongst others by:

[3] INSPIRE Metadata Regulation: http://data.europa.eu/eli/reg/2008/1205/oj#d1e600-14-1

- Relying on the EU authentication platform and security protocols for data sharing.
- Logging of user identity during data access, download, and upload, including version control. This enables to restore the availability and access to the data in a timely manner in the event of a physical or technical incident**.**

In an effort to develop a truly open-access data repository, access to the ICARUS global open-data repository does not need credentials. Access to the sensor data platform is granted only to authenticated users. Users can be freely authenticated by their username and password. Authentication credentials must be exchanged over SSL to ensure they are kept private.

5.2.3 Making data interoperable

Points to be addressed:

- *Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.*
- *Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?*

Strong focus has been given to interoperability of the data produced. In ICARUS we use the established European and international standards for the storage, exchange and dissemination of project data. INSPIRE (the European Directive on Infrastructure for Spatial Information) compliance has been used wherever possible (spatial dataset). Maps are made available for use by others as standard GIS files (raster and vector) usable by GIS systems including OpenGIS.
Other dataset produced are in common format (e.g. xls(x), csv, txt, AccDb, netCDF) which assures full interoperability allowing easy parsing and information exchange.

5.2.4 Allocation of resources

5.2.4.1 Estimation of costs

Costs for establishing and maintaining the ICARUS global open-data repository are covered by the AUTH financial budget of ICARUS.

5.2.4.2 Responsibilities

Each ICARUS partner has to respect the policies set out in this DMP. Datasets have to be created, managed and stored appropriately and in line with applicable legislation.

The Project Coordinator has a particular responsibility to ensure that data shared are easily available, but also that backups are performed and that proprietary data are secured.

AUTH, as WP9 leader, will ensure dataset integrity and compatibility for its use during the project lifetime and beyond by different users.

Validation and registration of datasets is the responsibility of the partner that generates the data in the WP.

Backing up data for sharing through open access repositories is the responsibility of the partner possessing the data. Quality control of this data is the responsibility of the relevant WP leader, supported by the Project Coordinator.

If datasets are updated, the partner that possesses the data has the responsibility to manage the different versions and to make sure that the latest version is available in the case of publicly available data. WP9 will provide naming and version conventions.

Last but not least, all partners must consult the concerned partner(s) before publishing data in the open domain that can be associated to an exploitable result.

5.2.4.3 Responsibilities for data management

AUTH and UPCOM are co-managers of the two data repositories, with these specific responsibilities

UPCOM is responsible for:

- Initial set-up of the hardware and software components of the data repositories
- Maintenance of the hardware and software components of the data repositories
- Carrying out the initial security assessment of the repositories
- Perform Security Assessment on a regular basis (e.g., one year) in order to guarantee the agreed security level
- Reporting and blocking any possible security threat, taking appropriate measures accordingly.

UPCOM responsibilities include:

- setting up and upgrading, when needed, of the hardware and software components of the repositories
- creation, maintenance and upgrading of the User Group Account database
- co-creation, under specific instructions provided by AUTH, of the data repository folders/sub-folders for each user group and document types
- Capacity management of hardware and software components.

AUTH is responsible for:

- Supervise the users request for access to and download of data
- The content (of data and documents) reported into the data repositories
- Definition, creation, updating of the data repository structures, i.e.: structure of folders and subfolders, names, contents and access, upload, download rights
- Co-creation with the UPCOM of the data repository folders/sub-folders for and document types
- Providing instructions to the UPCOM about the data repository structure

Each member of the ICARUS consortium is responsible for transferring and updating their data on the global open-data repository. The data must conform to the standards defined by the present document including but not limited to naming, structure and anonymization. When storing publicly available data, the publisher is responsible to ensure that these do not infringe personal or corporate privacy and that they don't jeopardize the commercialization of the project's products.

5.2.5 Data security

Points to be addressed:

- *Address data recovery as well as secure storage and transfer of sensitive data*

With respect to Privacy and Data Protection, the so-called "GDPR" EU-legislation (i.e. the General Data Protection Regulation) imposes several new obligations upon the consortium partners being data

processors. Moreover, several new rights are granted to data subjects and significant fines are introduced in case of a data breach.

Apart from this legislation, the consortium partners regard privacy and data protection as a fundamental principle and hence apply a strict policy on this matter.

5.2.5.1 Data confidentiality and integrity

Data stored at the ICARUS sensor data platform is protected against unauthorised access by means of standard login procedure. Access to data is granted only to authenticated users. Users can be authenticated by their username and password. Authentication credentials must be exchanged over SSL to ensure they are kept private. Appropriate access levels may be granted by the creation of groups

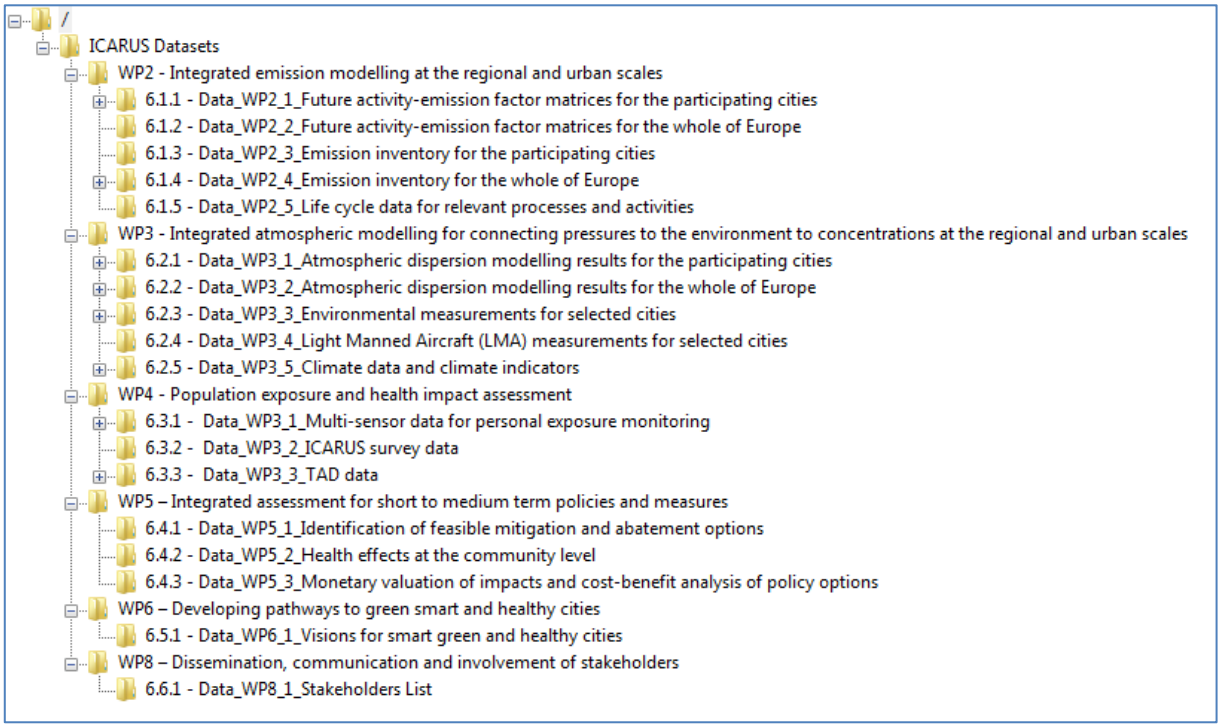
5.2.6 Ethical aspects

The transfer of data on human subjects to the ICARUS repository is only considered when: informed consents, ethics approval and – when applicable - approval by local data protection authorities cover the purpose that the data are envisaged to be used within ICARUS and allow transfer of individual or aggregated data to the ICARUS repository.

All data that have been transferred and stored into the ICARUS global open-data repository are either pseudonymised or completely anonymized. The Data Owner/Data Provider was responsible for the anonymization or pseudonymity process and for ensuring that identifiable variables are not transferred to the ICARUS repository. Directly identifiable variables include - but are not limited to - national ID number, name, phone number, ZIP-code, e-mail address, address, geographical coordinates (at a resolution that risks identification). One shall also be aware that a combination of just of few indirect identifying variables (such as birth data, gender, and zip-code) can be used to identify a large portion of individuals on any dataset. In this context, the Data Owner/Data Provider shall only provide such variables at the lowest possible resolution that is necessary to for analysis, e.g. district instead of zip-code; year of birth or age instead of birth date.

6 Datasets in ICARUS

Datasets in the ICARUS global open-data repository are stored according to the Work-Package from where they have been generated. To this end a tree-structure labelled according to the WP names has been created in the root folder of the repository (see figure below).



6.1 Datasets in WP2 - Integrated emission modelling at the regional and urban scales

6.1.1 Future activity-emission factor matrices for the participating cities

Data set reference and name	Data_WP2_1_Future activity-emission factor matrices for the participating cities
Data summary	The dataset includes data on future activity-emission factor matrices for all the ICARUS participating cities and for the classical air pollutants (NH_3, NMVOC, NO_X, PM_{10}, $PM_{2.5}$, CO, SO_2, BC and OC), GHGs (CO_2, CO, N_2O, CH_4), heavy metals, PAHs and dioxins, relevant to the years 2015, 2020 and 2030. A detailed description of the structure of the DB can be found in deliverable D2.2 freely downloadable from the ICARUS website. Additionally, bottom-up activity data and emission factors are combined to sectoral emission grids (1x1 km) for each city. The sites targeted cover all the ICARUS pilot cities.
Standards and metadata	The data are available in MS-Access format (accdb) to maximise

	the interoperability for use with a variety of analytical software. The sectoral emission grids are provided in standard GIS vector format (shp).
Data sharing	Aggregated bottom-up data in form of sectoral emission are publicly freely available to anyone with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. Raw data with activity-emission factor matrices are freely available for project internal use only as they cannot be made available to the public due to privacy and confidentiality issues. However, raw data might be shared with interested users through individual requests made to the Project Coordinator.
Archiving and preservation	Bottom-up emissions aggregated on sectoral level are stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end. Due to their sensitive nature, raw data are stored by the responsible (city) partners.

6.1.2 Future activity-emission factor matrices for the whole of Europe

Data set reference and name	Data_WP2_2_Future activity-emission factor matrices for the whole of Europe
Data set description	The dataset includes data on future activity-emission factor matrices for the whole of Europe for the classical air pollutants (NH_3, NMVOC, NO_x, PM_{10}, $PM_{2.5}$, CO, SO_2, BC and OC), GHGs (CO_2, CO, N_2O, CH_4), heavy metals, PAHs and dioxins, relevant to the years 2015, 2020 and 2030. A detailed description of the structure of the DB can be found in deliverable D2.2 freely downloadable from the ICARUS website.
Standards and metadata	The data are available in MS-Access format (accdb) to maximise the interoperability for use with a variety of analytical software.
Data sharing	Sectoral emission data are publicly freely available to anyone with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. Raw datasets relevant to activity-emission factor matrices cannot be made available to the public but they may be shared with interested users by means of individual requests made to the Project Coordinator.
Archiving and preservation	The data are stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end.

6.1.3 Emission inventory for the participating cities

Data set reference and name	Data_WP2_3_Emission inventory for the participating cities data set
Data set description	The dataset includes model-ready gridded sector-specific emissions at a high spatial resolution for six air pollutants (NH_3, NMVOC, NO_X, PM_{10}, $PM_{2.5}$, SO_2) for 2015, 2020 and 2030 using a bottom-up approach. The sites targeted cover the pilot cities addressed in ICARUS (i.e., Thessaloniki, Athens, Basel, Brno, Milan, Madrid, Ljubljana and Stuttgart). These emission inventories have been used as input for the atmospheric pollution models applied in WP3.
Standards and metadata	For each city the dataset is available in csv format as it is best suited for project internal use (especially as input for the atmospheric pollution models applied in WP3). Requirements of the INSPIRE standards have been taken into account.
Data sharing	The dataset with emission inventory is publicly freely available with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The data are stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project. The same datasets are stored in the ICARUS Decision Support System developed in WP7.

6.1.4 Emission inventory for the whole of Europe

Data set reference and name	Data_WP2_4_Emission inventory for the whole of Europe data set
Data set description	This dataset includes model-ready gridded sector-specific emissions at a high spatial resolution for six standard air pollutants (NH_3, NMVOC, NO_X, PM_{10}, $PM_{2.5}$, SO_2) and GHGs (CO_2, CO, N_2O, CH_4) for 2015, 2020 and 2030 using a top-down approach for the spatial distribution. Gridded sector-specific emissions for the standard air pollutants are stored in the sub-folder named "*Air_pollutants*" while gridded emissions for GHHs are stored in the sub-folder called "*GHG*". These emission inventories were used as input for the atmospheric pollution models applied in WP3.
Standards and metadata	The dataset is available in txt format (semicolon separated) as it is best suited for project internal data exchange (especially as

	input for the atmospheric pollution models applied in WP3). Requirements of the INSPIRE standards have been taken into account.
Data sharing	These datasets are publicly freely available with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The data are stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end as well as in the ICARUS Decision Support System.

6.1.5 Life cycle data for relevant processes and activities

Data set reference and name	Data_WP2_5_Life_cycle_data for relevant processes and activities data set
Data set description	The dataset includes life cycle emission data relevant to the activities and processes identified in WP5. This includes life cycle emission factors for the road transport, heating sector as well as electricity consumption. Lifecycle emission factors for electricity consumption (filename: *LCDB_Electricity_FINAL.xlsx*) comprise several possible future developments of the European electricity market, including higher penetration of renewable energy. For the buildings, separate emission factors are generated for insulation (filename: *LCDB_Buildings_insulation_FINAL.xlsx*) and heating technologies (filename: *LCDB_Buildings_heating_FINAL.xlsx*). Lifecycle emissions comprise emissions from production, transport (of goods/fuels) and disposal. For transport (filename: *LCDB_Transport_Final.xlsx*), separate lifecycle emission factors for different vehicle types are collected, split into production, operation, disposal, infrastructure and fuel cycle. All data are documented in deliverable D2.3 freely downloadable from the ICARUS website.
Standards and metadata	These datasets are available in MS Excel (xlsx) format.
Data sharing	These datasets are publicly freely available to anyone with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and	The data are stored in the ICARUS Global Data Repository which

preservation (including storage and backup)	will remain operational for 3 years after the project end.

6.2 Datasets in WP3 - Integrated atmospheric modelling for connecting pressures to the environment to concentrations at the regional and urban scales

6.2.1 Atmospheric dispersion modelling results for the participating cities

Data set reference and name	Data_WP3_1_Atmospheric dispersion modelling results for the participating cities
Data set description	This dataset is operationally produced by the atmospheric dispersion models applied in ICARUS and contains data stored in three main subfolders as follows: i) Folder "EURO-CORDEX data for Weather Clustering" which includes weather data from three Regional Climate Models (RCMs) in three separate further sub-folders named according to the RCM model used, namely, SMHI-RCA4; KNMI-RACMO22E; and IPSL-INERIS-WRF331F. Each subfolder includes climate daily data of the European domain for the years2001 to 2050 under RCP4.5 extracted from the EUROCORDEX database. These data were used for the ICARUS developed weather clustering methodology. The filename conventions is the following: *variable/domain/institute/drivingmodel/experiment/RCM model/timefrequency/timeperiod* (e.g. hurs_EUR-11_IPSL-IPSL-CM5A-MR_rcp45_r1i1p1_IPSL-INERIS-WRF331F_v1_day_20060101-20101231). More details on the naming convention can be found in https://www.euro-cordex.net/060378/index.php.en ii) Folder "Air quality data of Representative days" which includes air quality data for the Representative days (RD) which were estimated by the EUROCORDEX climate data (see above). This data has been derived from the application of a nesting approach at the urban scale of the Eulerian WRF/WRF-Chem model in the 9 pilot cities addressed in ICARUS i.e. Thessaloniki, Athens, Roskilde/Copenhagen, Basel, Brno, Milan, Madrid, Ljubljana and Stuttgart. On the urban scale, nine (9) nests with dimensions 2x2km have been used consisting of 42x42 grid cells each. Daily concentrations were produced and stored for NO_2, O_3, PM_{10}, $PM_{2.5}$. Data are stored in various subfolders named according to the city to which they are referring to and in further subfolders named according to the policy/measure

	assessed. Within these subfolders the filename convention is the following: *Domain/date of the RD*. iii) The "GHG" folder which includes GHGs (CO_2 and CH_4) concentration data calculated for each ICARUS city under the BaU emission scenario. The filename conventions is the following: *Cityname.csv*
Standards and metadata	The "EURO-CORDEX data for Weather Clustering" dataset is available in net-CDF file format. The "Air quality data of Representative days" dataset is available in standard WRF output net-CDF format. The GHGs concentration dataset for each city is available in csv format.
Data sharing	All the various datasets stored in the "Data_WP3_1_Atmospheric dispersion modelling results for the participating cities" folder are publicly freely available and can be downloaded and used with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The files are stored in the ICARUS Global Data Repository which will remain operational for 3 years after project end.

6.2.2 Atmospheric dispersion modelling results for the whole of Europe

Data set reference and name	Data_WP3_2_Atmospheric dispersion modelling results for the whole of Europe
Data set description	This dataset is stored in various subfolders named based on the simulation day to which the modeling results are referring to (i.e. *yyyymmdd-hh*). The dataset has been operationally produced by the Atmospheric dispersion models applied in ICARUS. Air quality data (NO_2, O_3, PM_{10} and $PM_{2.5}$) and meteorological data (precipitation, relative humidity, temperature, wind speed and direction) are available at European level using an outer grid with dimensions 12x12km consisting of 303 x 303 cells. The dataset has been used within WP3 to assess air quality levels at European scale and to produce concentration maps.
Standards and metadata	The dataset is available in the form of concentration maps stored in png format.
Data sharing	The data is stored in the form of concentration maps (png format) freely available to anyone. Row data in net-CDF format

	were not stored due to their huge size (i.e. we produced ca. 5,000 files each one of about 55GB). However, raw data in net-CDF format may be shared with interested users by means of individual requests made to the Project Coordinator with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics in doing so.
Archiving and preservation	The files are stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end.

6.2.3 Environmental measurements dataset

Data set reference and name	Data_WP3_3_Environmental measurements for selected cities
Data set description	The dataset contains files with environmental measurements that were obtained from ad-hoc targeted field campaigns in six ICARUS cities (i.e. Athens, Thessaloniki, Madrid, Stuttgart, Ljubljana, Brno). The dataset is organized in a series of subfolders each one named according to the city name and which includes all the data measured in that specific ICARUS city. Each of these subfolders contains several files which include: the location where the measurements were carried out, the time stamp, the air pollutant concentrations measured (NOx, SO_2, CO, O_3) as well as the results of the chemical analysis of the PM2.5 samples collected (PAHs, ions, heavy metals, OC/EC). A further file named "GHGs Ground sampling.xlsx" reports the GHGs (CO_2, N_2O, CH_4, SF_6) concentration levels measured in the different cities.
Standards and metadata	The data are in MS-Excel xls(x) format to maximise the potential for use with a variety of analytical software.
Data sharing	The datasets are publicly freely available and can be downloaded and used by other researchers with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. Raw data (primary data from filters weighting and chemical analysis procedures) have not been stored. These datasets may however be shared with interested users through individual requests made to the Project Coordinator with the expectation

	that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation (including storage and backup)	This dataset is stored in the ICARUS Platform Global Data Repository, which will remain operational for 3 years after the project end.

6.2.4 Light Manned Aircraft (LMA) dataset

Data set reference and name	Data_WP3_4_Light Manned Aircraft (LMA) measurements for selected cities
Data set description	The dataset contains air quality (VOCs) and GHGs (CO_2, N_2O, CH_4, SF_6) data at different height profiles collected in the frame of the ICARUS aerial campaigns. The aircraft operated at different height profiles above three European cities (Athens, Thessaloniki and Ljubljana) over sites with different characteristics (urban, suburban, rural) providing measurements of air pollutants and GHGs concentrations using sensor technologies as well as special sampling instrumentation. The data are stored in a single MS-Excel file named "*Aerial_campaign_data.xlsx*" which encompasses two worksheets: one named "*Airplane_GHGs"* which includes GHGs concentration levels measured over the 3 cities and the second named *"Airplane_VOC's_Conc"* which includes VOCs (benzene, toluene, o – Xylene, and m,p – Xylenes) concentrations measured over the 3 cities. Data are reported as average concentrations respectively above and below the atmospheric mixing layer.
Standards and metadata	The data is in MS-Excel xls(x) format to maximise the potential for use with a variety of analytical software.
Data sharing	This dataset is publicly freely available and can be downloaded and used with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation (including storage and backup)	These data will be stored in the ICARUS Platform Global Data Repository, which will remain operational for 3 years after the project end.

6.2.5 Climate data and climate indicators

Data set reference and name	Data_WP3_5_Climate data and climate indicators data set
Data set description	This dataset is stored in one subfolder named "EUR-11". The latter includes three subfolders named: "historical", "rcp45" and "rcp85 " which contain high-resolution climate simulation daily results for the whole Europe, respectively covering the historical period (from 1950 to 2005) and the future scenarios (RCP4.5 and RCP8.5) from 2006 to 2065. Dataset are derived from a bunch of four Regional Climate Models (RCMs) participating to EURO-CORDEX EUR-11 with a horizontal resolution of about 12 km and includes the following variables: 2-meters surface relative humidity (hurs), Mean precipitation flux (pr), surface solar radiation downwards (rds), 10-meters wind speed (sfcWind), 10-meters maximum wind speed (sfcWindmax), 2-meters average temperature (tas), 2-meters maximum temperature (tasmax), 2-meters minimum temperature (tasmin), 10m u-component of wind (uas) and 10m v-component of wind (vas). Boundary conditions for downscaling purposes, over ICAURS domains, were provided by the General Circulation Model GFDL-GCM. The filename convention is the following: *variable/domain/institute/drivingmodel/experiment/RCM model/timefrequency/timeperiod* (e.g. tas_EUR-11_MPI-M-MPI-ESM-LR_rcp85_r1i1p1_CLMcom-CCLM4-8-17_v1_day_20960101-21001231.nc)
Standards and metadata	The aforementioned datasets are provided in the standard format (net-CDF CF compliant) for climate data sharing. Deliverable D3.1 freely accessible on the ICARUS web site defines the exact data set specifications.
Data sharing	These data sets are publicly freely available and may be used by other interested parties with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics in doing so.
Archiving and preservation	Climate data are stored in the ICARUS Platform Global Data Repository, which will remain operational for 3 years after project end.

6.3 Datasets in WP4 - Population exposure and health impact assessment

6.3.1 Multi-sensor data for personal exposure monitoring

Data set reference and name	Data_WP4_1_Multi-sensor data for personal exposure monitoring in the participating cities data set
Data set description	This dataset includes data collected by personal sensors worn by each participant, and static sensors located in each household of the city campaigns in Athens, Thessaloniki, Madrid, Milan, Brno, Basel and Ljubljana. There are three types of sensors: the custom made by IoTech, the one provided by Garmin (Vivosmart 3) and the static sensors provided by uHoo. The custom-made sensors measured air pollution by capturing three types of particulate matter, PM_1, $PM_{2.5}$ and PM_{10}, temperature, relative humidity, altitude and speed. The Garmin sensor provided measurements of personal movement/activity and intensity of activity, amount of burned calories, sleeping time, stress levels and heart rate in 15-second time intervals. The static sensors provided by uHoo took measurements of temperature, relative humidity, and air pollution levels (particulate matter – $PM_{2.5}$, ozone, nitrogen dioxide, carbon dioxide, carbon monoxide and total volatile organic compounds. All the data measured by these three sensor types are stored in the open repository in different subfolders named according to the city where the data were collected and in further subfolders named according to device used (i.e., “GARMIN”; “PPM SENSORS”; and “uHoo”) and finally in two subfolders named on the basis of the season (“WINTER” and “SUMMER”). Single files encompassing the data measured are named according to the following scheme: City_Season_ParticipantID.xlsx (e.g. ATH_S_P039.xlsx).
Standards and metadata	The data are stored in MS-Excel xls and xlsx formats to maximise the potential for use with a variety of analytical software. The same data are also stored as backup copy in the ICARUS Sensor Data Platform developed in the frame of WP4, using the Couchbase NoSQL database.
Data sharing	The multi-sensor data for personal exposure monitoring is freely available to anyone as Open Data in anonymous format. Data can be freely downloaded and used with the expectation

	that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of individual participants and its tracking (GPS coordinates) has been removed.
Archiving and preservation	Data are stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end. A backup copy of the same data is stored in the Sensor Campaigns Data Platform developed in the frame of WP4 and accessible to anyone upon free registration.

6.3.2 ICARUS survey data

Data set reference and name	Data_WP4_2_ ICARUS survey data set
Data set description	The dataset consists of quantitative and qualitative information collected from each participant during the personal exposure campaigns in the case-study cities (Athens, Thessaloniki, Madrid, Milan, Brno, Basel and Ljubljana). The survey data captured information related to participants' socioeconomic status, age, gender, nationality, place of birth, health status, occupation, education level, marital status, children, annual income, daily schedule regarding work and exercise and time spent indoors and outdoors, details on their working conditions, preferred means of transportation and time spent weekly, medical treatments, illnesses and smoking. In addition, data regarding each participants' household has also captured, such as the number of people (minor and adults) living in it, the house structure and surface, isolation, heating (type and autonomy of the thermostat), proximity to pollutants, cooking, cleaning, smoking and ventilation habits, renovations, income and existence of pets. This data has been used to evaluate how SES and individual lifestyle should be taken into account in exposure modelling. There are two types of files available containing respectively raw data relevant to individuals and households and aggregated data on all the individuals and all households for each question in the survey.
Standards and metadata	The data are stored in MS-Excel xls(x) format to maximise the potential for use with a variety of analytical software.
Data sharing	The ICARUS survey data set is freely available to anyone as Open Data in anonymous format. Data can be freely downloaded and

	used with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of individual participants has been removed from the "individual" datasets.
Archiving and preservation	Data is stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end.

6.3.3 ICARUS TAD data

Data set reference and name	Data_WP4_3_ ICARUS Time Activity Data
Data set description	The datasets include data collected by time activity diaries (TAD) in the form of questionnaires from each participant of the city campaigns in Athens, Thessaloniki, Madrid, Milan, Brno, Basel and Ljubljana. The questionnaires were digitized in excel format to match the sensors data. The information the time activity diaries provided relates to the type and the place of an activity the participant performs in a specific time frame. The activity (resting, playing, sleeping, sports, cooking, smoking, cleaning, running) can be indoor or outdoor, the participant can be in transit using public or personal means of transportation (bus, car, motorbike, bicycle, foot, other). Also, information regarding the house conditions (candles on/off, air conditioning – fan on/off, fireplace, windows open/closed) is also provided. All the data collected through the TAD surveys are stored in the open repository in different subfolders named according to the city where the data were collected and in further two subfolders named according to the season when the data was collected ("WINTER" and "SUMMER"). Single files encompassing the data measured are named according to the following scheme: City_Season_ParticipantID.xlsx (e.g. ATH_S_P032.xlsx).
Standards and metadata	The anonymized data is stored in MS Excel xls and xlsx formats to maximise the potential for use with a variety of analytical software.
Data sharing	The ICARUS TAD dataset is freely available to anyone as Open Data in anonymous format. Data can be freely downloaded and used with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program

	are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of individual participants has been removed.
Archiving and preservation	Data is stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end.

6.4 Datasets in WP5 – Integrated assessment for short to medium term policies and measures

6.4.1 Identification of feasible mitigation and abatement options

Data set reference and name	Data_WP5_1_Feasible mitigation and abatement options data set
Data set description	This dataset consists of databases where all the policies and measures options initially identified in all the cities are reported. Based on the discussions with the city stakeholders this list was then narrowed down to 5 policies per city which have been then integrally assessed in Task 5.4 according to the ICARUS paradigm. Both technical (e.g., changing emission factors due to the penetration of more environment-friendly technologies) and non-technical measures (that influence the citizens behavior) are covered and both short term measures that can be implemented in the immediate future and medium-term measures which entail for instance changes of infrastructure or a certain time for market penetration and will be thus fully effective only after 2020 will be considered.
Standards and metadata	Data is stored as MS Excel xlsx format. In the MS Word file named "*Overview of policies and measures_documentation.docs*" the description of the content of MS excel file is reported
Data sharing	The dataset containing the identified policies and measures options has been shared with interested stakeholders (e.g., policymakers and city authorities). The dataset is publicly freely available to anyone with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation (including storage and backup)	The files are stored in the ICARUS Global Data Repository, which will remain operational for 3 years after the project end.

6.4.2 Health effects at the community level

Data set reference and name	Data_WP5_2_Health effects at the community level data set.
Data set description	This dataset contains the quantitative assessment of health impact at the population level as well as on vulnerable population subgroups (e.g. children) for several health endpoints. The latter includes the followings: $PM_{2.5}$ mortality (all causes) >30 years old; PM_{10} infant mortality (0-1 year); PM_{10} Incidence of Chronic bronchitis (adult); PM_{10} Cardiac Hospital admissions, PM_{10} Respiratory Hospital admissions; PM_{10} Prevalence of Chronic bronchitis (children aged between 6 and 12); NO_2 mortality (all causes) > 30 years old; and NO_2 Prevalence of bronchitic symptoms in asthmatic children (between 5 and 14). Data are reported as central estimates as well as lower and upper bound to reflect the associated uncertainty. The impact on human health has been estimated up to the year 2040 as the total number of additional health outcomes for each health end point above mentioned in each ICARUS city for each policy/measure options identified in WP5. The filename convention is the following: *City_ Health Impact_Pollutant LT.xls* (e.g. Milan Health Impact_PM LT.xlsx). All the files include two worksheets where the total number of additional health outcomes are reported respectively for the whole computational domain and for the municipality only. The only exceptions are Athens and Thessaloniki where results are reported only for the whole computational domain as all the policies analyzed are addressing the greater city areas.
Standards and metadata	The datasets are available in MS–Excel xls(x) format to ensure full interoperability allowing easy parsing and information exchange.
Data sharing	The datasets are publicly freely available and can be downloaded and used by other researches with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The files are stored in the ICARUS Global Data Repository which will remain operational for 3 years after the project end.

6.4.3 Monetary valuation of impacts and cost-benefit analysis of policy options

Data set reference and name	Data_WP5_3_Monetary valuation of impacts and cost-benefit analysis data set
Data set description	This dataset includes the results of the cost-effectiveness and cost-benefit analysis performed for each of the policy/measure options defined in Task 5.1. As such it complements the dataset "Data_WP5_2_Health effects at the community level data set" in providing the assessment of the economic dimension for the same policies/measures. Cost-effectiveness analysis examined the costs of these options and calculated for example the cost per ton of CO2eq. For the cost-benefit analysis the dataset includes all benefits, damages and costs and the non-monetary (intangible) items which were transformed into monetary values (where possible) including social costs, monetized health impacts, monetized contributions to climate change, utility and gain losses. The full dataset is organized in three different files according to the sector addressed by the policy/measure options analyzed. In this light, the file named "CBA active transport" includes the full results for the active transport policies; the file named "CBA alternative fuel vehicles" results for all the alternative fuel vehicles policies; and the file "CBA energy efficiency" for the energy efficiency policies. Every single file includes multiple worksheets which respectively encompasses a summary of all the CBA results for the policy sector addressed, as well as other worksheets including the detailed results for each specific policy up to the year 2040.
Standards and metadata	The data are available either in MS–Excel xls(x) format to ensure full interoperability allowing easy parsing and information exchange.
Data sharing	The dataset containing the results of the monetary evaluation of the chosen policies and measures options is publicly freely available to anyone with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics in doing so.
Archiving and preservation	This dataset is stored in the ICARUS Global Data Repository, which will remain operational for 3 years after the project end.

6.5 Datasets in WP6 – Developing pathways to green smart and healthy cities

6.5.1 Visions for smart green and healthy cities

Data set reference and name	Data_WP6_1_Visions for smart green and healthy cities data set
Data set description	This dataset consists of narrative visions of ICARUS smart, green and healthy city developed in the frame of WP6 for all the ICARUS cities. The narrative-based visions have been updated through participatory expert and stakeholder workshops, literature review and horizon scanning. Supporting data in the development of visions include videos made at workshops as part of the ICARUS project. The dataset includes the steps towards the development of a transition pathway and how the transition from the current state to different city visions can be managed. The file includes data and information for each ICARUS city resulting from the interactions with local stakeholders.
Standards and metadata	The data set is available pdf format. Supporting document (i.e. three videos) are available in mp4 format.
Data sharing	This dataset is publicly freely available to anyone. The data and information included may be used by other researchers with the expectation that both the ICARUS project and the European Commission's Horizon 2020 as the funding program are properly cited.
Archiving and preservation	This dataset has been stored in the ICARUS Global Data Repository, which will remain operational for 3 years after the project end. The three videos recorded during the Madrid workshop are also available in the ICARUS website.

6.6 Datasets in WP8 – Dissemination, communication and involvement of stakeholders

6.6.1 Stakeholders List

Data set reference and name	Data_WP8_1_Stakeholders list
Data set description	The dataset consists of the identified ICARUS stakeholders.
Standards and metadata	The dataset is available in xls(x) format. For each stakeholder identified the following metadata information is provided in the concise dataset: • Organization name

	• Field of activity • Type of stakeholder Full details including: • Name and Surname • Role in the Organization • Job title • Contact details (i.e. email address, phone number, etc.) cannot be shared for privacy issues and are included in the extended version.
Data sharing	The concise dataset is publicly freely available to anyone. The extended dataset cannot be shared for privacy issues. Nonetheless, all the details are stored in an excel file for internal use only and may be shared upon written request to the Project Coordinator.
Archiving and preservation	The concise dataset has been stored in the ICARUS Global Data Repository, which will remain operational for 3 years after the project end.

6.6.3 Scientific publications

Data set reference and name	Data_WP8_2_Scientific Publication data set
Data set description	The dataset consists of all the ICARUS scientific publications produced during the project life.

Standards and metadata	Metadata maximises the discoverability of publications and ensure the acknowledgment of EU funding. The inclusion of metadata is necessary for adequate monitoring, production of statistics, and assessment of the impact of H2020. In addition to basic bibliographic information about deposited publications the following metadata information are provided:
	• EU funding acknowledgement:
	  ○ Contributor: "European Union (EU)" & "Horizon 2020"
	• Project Information:
	  ○ Grant number: “690105”
	  ○ Project Acronym: “ICARUS”
	  ○ Project Name: “Integrated Climate forcing and Air pollution Reduction in Urban Systems”
	• Publication Date.
	• Journal name
	• DOI (if any)
	• Keywords
	• Authors and Contributors.

Data sharing	ICARUS scientists committed themselves to follow the ‘Open Access Publishing Model’, meaning that the scientific articles

	are provided in open access mode by the scientific publisher.
Archiving and preservation	Publications are freely available through the ZENODO repository at https://zenodo.org/communities/icarus-2020/?page=1&size=20.