



Data Management Plan

D3.2 (5.0 - Public)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689954.

Project Acronym and Name	iSCAPE - Improving the Smart Control of Air Pollution in Europe	
Grant Agreement Number	689954	
Document Type	Report	
Document version & WP No.	V.5.0	WP3
Document Title	Data Management Plan	
Main authors	Guillem Camprodon, Óscar González, Santa Stibe	
Partner in charge	IAAC	
Contributing partners	-	
Release date	11/2019	

The publication reflects the author's views. The European Commission is not liable for any use that may be made of the information contained therein.

Document Control Page			
Short Description	<i>one paragraph summary of the report</i>		
Review status	Action	Person	Date
	Quality Check	<i>Coordination Team</i>	08/10/2019
	Internal Review	Santa Stibe (UCD)	15/11/2019
Francesco Pilla (UCD)		15/11/2019	
Distribution	Public		

Revision history			
Version	Date	Modified by	Comments
v1.0	12/2017	Guillem Camprodon	First internal version
V2.0	01/2018	Guillem Camprodon	Second internal version
V3.0	02/2018	Guillem Camprodon	First public release
V4.1	10/2018	Guillem Camprodon, Santa Stibe, Óscar González	Report refactoring after EU review feedback
V4.2	02/2019	Guillem Camprodon, Óscar González Santa Stibe and Francesco Pilla	Preliminary review
V4.3	11/2019	Santa Stibe and Francesco Pilla	Final internal review
V4.3.1	11/2019	Guillem Camprodon, Óscar González	Final public version
v5.0	05/2020	Guillem Camprodon	Enhanced 2.4. Increase data re-use section with more information on the public release and archiving process Added Consolidated Environmental Datasets in Annex with OpenAIRE compatible repositories URL and DOI. Mapping the original datasets Trace Codes with the consolidated datasets publicly released. Added the same for Software and Hardware Public Releases

Table of Contents

1. Data summary	10
1.1 Research data	11
1.2.1 Observational and Experimental data	11
1.2.3 Models and simulations	15
1.2 Operational data	22
1.1.1 Project management and monitoring	22
1.1.2 Research activities involving human subjects	24
1.1.3 Public events organized by the project	25
1.1.4 Project online platforms	27
2. FAIR Data	29
2.1 Making data findable, including provisions for metadata	30
2.1.1 The iSCAPE Datasets Tool	30
2.1.2 Provisions for metadata	34
2.2 Making data openly accessible	35
2.2.1 Observational and experimental environmental data	35
2.3. Making data interoperable	37
2.4. Increase data re-use	38
2.5 Managing personal data	41
3. Allocation of resources	43
4. Data security	45
5. Ethical aspects	46
6. References	47
Annex 1: Datasets	48
Environmental research datasets	48
Consolidated Environmental Datasets	48
Traced Environmental Datasets	51
Software and Hardware	59
Personal data	60

List of abbreviations and symbols

ADMS: Advanced Disaster Management Simulator

API: Application Programming Interface are a set of subroutine definitions, and tools for building application software.

CERN: European Organization for Nuclear Research,

CFD: Computational Fluid Dynamics

CSV: Comma Separated Values

DMP: Data Management Plan

DOI: Document Object Identifier

EDT: ENVI-met data file

ENVI-met: A three-dimensional microclimate model designed to simulate the surface-plant-air interactions in urban environment

EPA: Environmental Protection Agency

ESRI: An international supplier of geographic information system software

FCC: Future Cities Catapult

FMI: Finnish Meteorological Institute

GCM: Global Climate Model

GDPR: EU General Data Protection Regulation

GIS: Geographic Information System

GPL: The GNU General Public License is a widely-used free software license, which guarantees end users the freedom to run, study, share and modify the software.

GPS: Global Position System

GTFS: General Transit Feed Specification

IAAC: Institute for Advanced Architecture of Catalonia

IMOB: Hasselt University Mobility Research Institute

ISPRA: Italy Higher Institute for Environmental Protection and Research

JSON: JavaScript Object Notation is an open-standard file format that uses human-readable text to transmit data objects consisting of attribute–value pairs and array data types

LBW: Low Boundary Wall

ORCID: The ORCID is a nonproprietary alphanumeric code to uniquely identify scientific and other academic authors and contributors

OSM: Open Street Map

OpenAIRE: A network of Open Access repositories, archives and journals that support Open Access policies

PCS: Passive Control Solution

REST: Representational state transfer is an architectural standard for web services

SAF: Sensor Analysis Framework is a Python based environmental data sensors post processing tool built for the iSCAPE project

SC Platform: Smart Citizen Platform, the software platform ingesting, processing, archiving and presenting the iSCAPE sensors data

TUDO: Technical University Dortmund

UCD: University College Dublin

URL: Uniform Resource Locator, is a reference to a web resource on the Internet

VLL: Virtual Living Lab is an online platform developed and used in the iSCAPE project

List of Tables

TABLE 1 FAIR DATA PRINCIPLES	31
TABLE 2 PROJECT DATASETS TRACE CODES	32
TABLE 2 PROJECT DATASET TYPOLOGIES	33
TABLE 3 PROJECT DATASETS SOURCE TYPES	33
TABLE 4 PROJECT DATASETS STORAGE PLATFORM	34
TABLE 5 PROJECT DATASETS ACCESSIBILITY LEVELS	34
TABLE 6 GUIDANCE ON METADATA TOPICS TO ADDRESS	36
TABLE 7 DATASET INFORMATION IN ZENODO	41
TABLE 8 REVIEW DATES	45

List of Figures

FIGURE 1 SC PLATFORM SOFTWARE ARCHITECTURE	17
FIGURE 2 ENVIRONMENTAL SENSORS DATA LIFECYCLE	37

Executive Summary

iSCAPE is an EU H2020 project working on integrating and advancing the control of air pollution and carbon emissions in European cities in the context of climate change through the development of sustainable and passive air pollution remediation strategies, policy interventions and behavioural change initiatives.

The following report describes the Data Management Plan for the data collected, processed or generated by the project. The design and implementation of the Data Management Plan are led by the Data Manager (IAAC) with support from all the project partners. The purpose of the document is to provide a set of standard guidelines and detailed protocols to ensure project data is managed, following FAIR Data principles and Open Access practice while being compliant with the EU General Data Protection Regulation (GDPR) and other relevant national and international laws safeguarding individual rights to privacy and personal data protection.

This report includes provisions for both research data and operational data. Research data includes observational air pollution and climate data from sensors, intervention analysis and participant surveys conducted to assess the socio-economic impact of the project interventions, as well as models derived from these datasets. Operational data includes the practices carried out to successfully implement the project and oversee its operation, as well as events, publications and platform usage data.

Since the beginning of the project, the Data Management Plan has been a living document. To support the data management lifecycle for all data that is collected, processed or generated by the project, this document has been regularly updated. The current document is the fourth public release and the last iteration. Compared to the previous versions, this version also includes a detailed list of datasets generated by the project, also providing guidelines for data management. In order to provide a comprehensive overview of the datasets collected, the document was in continuous editing until the end of the project, when most of the data collection activities were completed.

1. Data summary

The overall objective of the iSCAPE project is to develop an integrated strategy for air pollution control in European cities that is grounded on evidence-based analysis. To achieve the project's objectives, a multi-stakeholder and multi-disciplinary research approach was implemented, which resulted in a great diversity of research data to manage.

Data collected, processed or generated during the project include:

- **Observational and experimental data.** These are, for example, environmental metrics gathered through the measurement campaigns, data collected using remote sensing stations, as well as participant surveys conducted to assess the socio-economic impact of the project interventions. In addition, multiple air pollution control strategies were deployed during the course of the project, which produced various datasets of observational and experimental data.
- **Models and simulations.** These include environmental, air pollution control and climate change simulation data, as well as transport mobility models.
- **Project operations data.** These include, for example, datasets generated to successfully implement the project and oversee its operations as part of project management and monitoring activities. This also includes data produced to document research activities, data collected to organize public events, as well as data collected through the usage of the project online platforms and their statistical tools. In addition, project operations data includes personal data which was collected in compliance with GDPR for a specific purpose and to an extent only necessary for that purpose.

The following chapter provides a comprehensive overview of both research and operations data. As described below, the iSCAPE project involves a wide typology of data. To describe research and operation data thoroughly, a table for each data category was created and it covers the following aspects:

- Data collection purpose and its relation to project objectives (also includes the description of data origin);
- Description of the data types
- Data archival and utility

1.1 Research data

This section provides an overview of the research data collected, processed, and generated by the project, which includes observational and experimental data, and data produced as part of the modelling and simulation efforts. The iSCAPE project has collected a wide variety of datasets and used them to generate various models and perform simulations from them. Input data ranges from sensors data, socio-economic surveys, geospatial datasets, among others. Each of them was used accordingly to generate models or perform simulations that are described in the following subsections. Finally, further details can be found in the Annex 1.

1.2.1 Observational and Experimental data

DATA COLLECTION PURPOSE AND RELATION TO THE PROJECT OBJECTIVES

Observational air pollution and climate data were critical for the development of the project, and its usage extended from WP2 to WP6. Observational data was primarily collected to monitor the performance of the project activities from the deployment of passive control strategies to citizen science initiatives. This data collection process was not exclusively limited to the observational tasks themselves and extended to other tasks to complement and validate modelling and simulation activities.

Observational and environmental data was collected as part of the project sensor deployments. Living Labs interventions were monitored with both, low cost sensors and high-end scientific instrumentation (WP3, WP5), which then served to help improve sensor readings from the former. As well, the citizen science activities were run using the Citizen Kits (WP4), which are detailed below.

These datasets also include information on the internal validation of the developed sensors. (WP3 and WP4). In addition, observational data includes data collected to survey people's perception of air pollution (WP2) and assess the socio-economic impacts of the project interventions (WP5).

Moreover, observational data was used to evaluate the interventions (WP5) and validate simulations (WP6), see Models and simulations below.

DESCRIPTION OF THE DATA TYPES

- **Citizen Kits**, formerly known as the Low-Cost Sensors, are aimed at providing the Living Labs with a tool for Citizen Science and Citizen Awareness activities in T5.0 and beyond the project. During the project, the

kits were deployed at the six Living Lab sites across Europe. Data collected by these kits contains environmental metrics such as light, temperature, humidity and atmospheric pressure, as well as air pollution indicative data from indoor tVOC sensors (Volatile Organic Compounds) and more accurate PM2.5 readings. This data was uploaded to the SC Platform managed by IAAC in real-time through a Wi-Fi connection or manually uploaded by their users afterwards. Data is currently available on the platform and downloadable by the users and project stakeholders. This data is also available on the Sensors Analysis Framework¹ used to assess the performance of the sensors as well as to generate data visualisations for the citizen science activities. All those activities are documented in D4.7 Citizen Science Communities Report².

- **Living Lab Stations**, formerly known as the High-End Sensors, are aimed at providing the Living Labs with a system for monitoring the performance of their interventions. The stations were deployed at the six Living Lab sites across Europe in T3.4. Data was collected with two purposes, assessing the sensors performance, and monitoring of the PCS interventions. The first purpose was achieved by performing multiple collocations of the sensors near regional official reference stations across the totality of the project Living Lab cities. The collected data included environmental data such as the one mentioned above for the Citizen Kit, as well as more complex readings from atmospheric chemistry such as CO, NO₂ and O₃ and particle composition (PM₁, PM_{2.5} and PM₁₀). This data was then compared with the reference stations data obtained mostly from the European Environmental Agency or provided by partners individually, containing air pollution data from available sensors on the site. This data generally contained low-mid resolution (~15'-1 day) datasets with some, but not all, of the targeted metrics (CO, NO₂, O₃, PM₁, PM_{2.5} and PM₁₀). The second goal focused on supporting the monitoring of the different interventions implemented in T3.5. Most of the interventions were monitored using high-end air pollution and weather instruments as part of T5.2. Since high-end instrumentation has generally tight usage timeframes, the usage of the stations with high end instrumentation allowed for more extended periods of measurement, once the high end instrumentation was needed for another experiments. Finally, all the sensor data was managed by the SC Platform operated by IAAC.
- **Heterogeneous air pollution and weather instrumentation** time series sensors data coming from the field observations conducted by the

¹ <https://github.com/fablabbcn/smartcitizen-iscape-data>

² https://www.iscapeproject.eu/wp-content/uploads/2019/07/iSCAPE-Report-D4.7_2019-05-24_final.pdf

partners in WP3. In Task 5.2 several monitoring campaigns were performed in multiple sites including Bologna and Lazzareto in Italy, Dublin in Ireland, Guildford in the United Kingdom and Vantaa in Finland. The campaigns had three main objectives:

- Provide experimental data to be used for the calibration and validation of the simulations conducted at larger spatial and temporal scales as part of WP4 and WP6, particularly for T6.2 Simulating the effect of infrastructural solutions on air quality at the neighbourhood level.
- Validate sensor solutions developed within WP3 by confronting the sensor against referenced scientific instrumentation in real contexts.
- Evaluate the potential of different PCSs, including low boundary walls (Dublin), photocatalytic coatings (Lazzareto site on the outskirts of Bologna) and various forms of green infrastructure (Bologna, Guildford and Vantaa).

Partners collected and stored the data internally for their usage in WP3 - WP6. When available, data was shared with IAAC to be used for the validation of the internal sensor using the Sensors Analysis Framework in WP3. All the monitoring campaign details including the purpose, instrumentation used and data collection procedures was documented part of D3.8 Report on deployment of neighborhood level interventions

- **An anonymous socio-economic survey** was used to collect information from the participants before and after the Citizen Science activities in the 6 Living Labs in T5.0, and was approved by a Research Ethics Committee. In addition, similar data was gathered from citizens and local stakeholders involved in the Living Labs activities (WP2) and in the deployment of the interventions (WP3). In the surveys, all external actors were asked to provide some general information about themselves which was necessary for the socio-economic assessment. This included :
 - Demographic variables: age, gender, income, cultural background, family composition, occupation.
 - Psychographic variables: soft variables investigating actors' attitudes, interests and behaviours.

The primary outcome of the data collected was to inform T4.1 about participants' opinion about environmental issues and their pro-environmental behaviour. The results of the research was made

available on D5.3 Report on interventions³ as well as on D4.1 Report on environmental effects of behavioural actions as well .

- **Individual activity-travel** data from small groups of citizens from Hasselt (Belgium), Bologna (Italy), and Guildford (UK) was collected using a GPS based smartphone application developed by the Transportation Research Institute (IMOB) at the University of Hasselt. This data was then used in an algorithm that aimed to identify relatively soft aspects of travel behavior with a potential to reduce the negative impacts of mobility on the environment and individual health⁴. Once the app called “SPARROWS/HTB” is installed on the participants phones, it collects individual location tracking as geographic coordinates and a timestamp and automatically reports the data to a database at the IMOB. This study and was approved by a Research Ethics Committee and was posteriorly published⁵. The database is password-protected, and it is exclusively used for research purposes. Trips and activities of a participant are recorded through the mobile application, and the purpose of the data collected was to provide feedback to the individuals about their exposure to air pollutants, vehicle use and CO2 emission and physical activity involvement as part of T4.1 Managing city need to adequate solutions at an urban level.

DATA ARCHIVAL AND FURTHER USE

All the data collected by the Citizen Kits and the Living Lab Stations will remain available on the SC Platform⁶. It will also be archived on the Zenodo platform⁷ following the process described in this report. Coupled with heterogeneous instrumentation data, this data is useful for assessing the performance of the low-cost sensors. IAAC, for instance, is using this data as part of the SAF (see Models and Simulations). This framework and its corresponding datasets will also be archived in Zenodo upon the completion of the project.

³ <https://www.iscapeproject.eu/scientific-reports/#environmental-effects>

⁴ Ahmed, S., Adnan, M., Janssens, D. et al. Pers Ubiquit Comput (2019) 23: 653. <https://doi.org/10.1007/s00779-018-1187-5>

⁵ <https://doi.org/10.1007/s00779-018-1187-5>

⁶ <https://smartcitizen.me/kits>

⁷ <https://zenodo.org/>

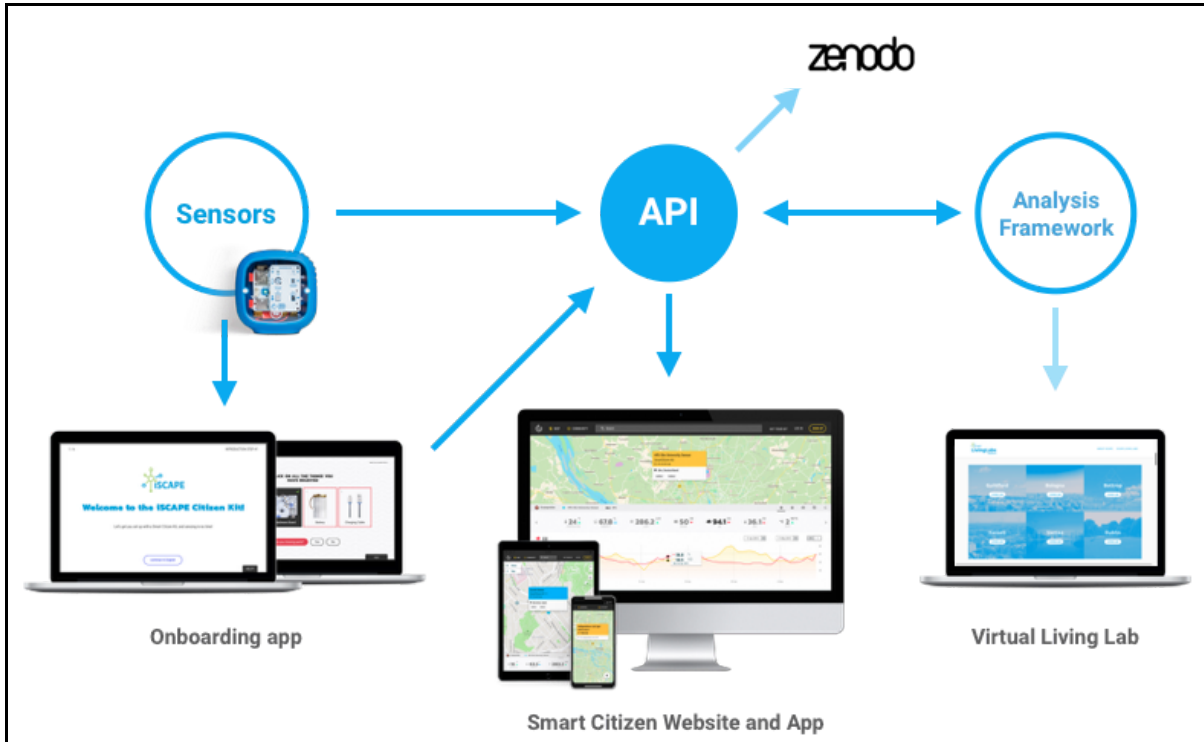


Figure 1. SC Platform Software architecture

The Zenodo repository is used due to its principles for open access and open data, being in Open Data policy by providing a catch-all repository for EC funded research. It provides integration of the FAIR data principles⁸, OpenAIRE⁹ support and tools for scholarly processes to handle data and software.

Regarding the socio-economic survey and the individual activity-travel, both contain extremely sensitive data and their public disclosure it's not possible, as is protected in accordance to GDPR. However, their aggregated results are available on their corresponding deliverables and are publicly accessible on the project website.

1.2.3 Models and simulations

DATA COLLECTION PURPOSE AND RELATION TO THE PROJECT OBJECTIVES

⁸ <https://doi.org/10.1038/sdata.2016.18>

⁹ <https://www.OpenAIRE.eu/>

Current, state-of-the-art simulation methods were at the core of the project research activities. To emphasize the role of such tools, the project has an entire work-package dedicated to simulations (WP6) with the primary goal to evaluate the efficacy of passive control systems (PCS) in the mid/long term and quantify their effects in terms of air pollutants levels and climate change. Those simulations used the most up-to-date numerical approach and primarily focused on Computational Fluid Dynamics (CFD) simulations using atmospheric pollution dispersion models such as the Atmospheric Dispersion Modelling System (ADMS-3). Even more, simulation tools were used in other work packages such as WP4 to assess the behavioural activity-based models in the field of transportation simulation models. Finally, modelling tools were also used in WP3 to post-processes the data obtained by some of the project sensors designed in T3.1.

As for any simulation process, input information determines the quality of outputs and final results. For this reason, most of the simulations in the project were run using a combination of both existing and publicly available datasets, and data collected by the project as their input data. Existing data mostly came from open data sources published by governments and scientific organizations, containing, for instance, GIS data of urban infrastructure, urban transportation and meteorological data. Data collected during the project and used in the simulation tasks as their input data are described in the Observational and Experimental data section above. Such data includes environmental data collected in WP5 as well as activity data collected in WP4.

DESCRIPTION OF THE DATA TYPES

- **Environmental simulations** were run based on the environmental data collected during the monitoring campaigns in Task 5.2 (see 1.2.1 Observational data) and existing meteorological, geographical and other relevant environmental datasets. All the simulations details including the purpose, procedure and results are documented in D6.3 Detailed report based on numerical simulations of the effect of PCSs at the urban level and D6.2 Microscale CFD Evaluation of PCS Impacts on Air Quality. Below, the purpose, data usage and results for each of the simulations are described:
 - **Gridded product of the city of Bottrop.** Based on a data set of the city of Bottrop (Germany) with GIS data of buildings,

roads and urban green infrastructure provided by the City of Bottrop alongside with other meteorological data available from official sources. The result is a gridded dataset for all baseline and scenario models (2 x 2 meters) with air temperatures at different heights, Physiologically Equivalent Temperature (PET) at 1.50m height, wind information at different heights and other meteorological and chemical parameters. Format: EDT (ENVImet-Fomat), ESRI shapefile (point information – not all information), Excel (point information – not all information). All the data is stored on-premises of the Technische Universität Dortmund (TUDO).

- **Climate scenarios with their changes per decade for multiple iSCAPE cities** to obtain a model using ADMS that show different climate scenarios with their changes per decade for multiple iSCAPE cities. These scenarios are based on several datasets from different sources: Global Climate Model (GCM) simulations, part of the Coupled Model Intercomparison Project, under the RCP8.5 and RCP4.5 greenhouse gas scenarios; CMIP5 GCM datasets downloaded from the EARTH System Grid Federation data archive¹⁰; atmospheric forcing datasets obtained from a numerical weather prediction system (SURFEX) and datasets on climatological baseline observations from the simulated region (ENVIMet). The result are multiple climate scenarios (changes per decade) for multiple variables for each the iSCAPE cities. These variables include daily means, minimum and maximum temperatures, diurnal temperature range, precipitation, solar radiation, air pressure, wind speed and wind direction. The output is a gridded (500 x 500 meters) weather dataset describing the state and processes of the surface-atmosphere system (SURFEX) and a gridded data (4 x 4 meters) meteorological and chemical parameters describing the state and processes of the surface-atmosphere system (ENVIMet). All the data is

¹⁰ <http://pcmdi9.llnl.gov>

available from FMI at request and will be made public by the end of the project in the Zenodo platform.

- **Climate scenarios with their changes per decade in Lazzaretto street canyons** obtained using the Atmospheric Dispersion Modelling System (ADMS) that shows different climate scenarios with their changes in the Lazzaretto street canyon. Data used included: multiple datasets of meteorological and air quality data gathered during the measurement campaign in Bologna (one summer 2018 in the Lazzaretto site) and from the Bologna airport synoptic meteorological station. A dataset of NO_x emissions estimated from ISPRA emissions database, as the source was a Euro 2 car placed in the middle of the street canyons. The simulation results include climate scenarios (changes per decade) for multiple variables (daily mean, minimum and maximum temperature, the diurnal temperature range precipitation, solar radiation and air pressure, wind speed and wind directions). Gridded data (500 x 500 meters) weather describing the state and processes of the surface-atmosphere system (SURFEX). Gridded data (4 x 4 meters) meteorological and chemical parameters are describing the state and processes of the surface-atmosphere system (ENVIMet). The collected data will be released after the completion of the project but not the results.
- **Climate scenarios with their changes per decade in Bologna street canyons** obtained using the Atmospheric Dispersion Modelling System (ADMS). Multiple datasets of meteorological and air quality data gathered during the two measurement campaigns in Bologna (one during summer 2017 and one during winter 2018). Meteorological and air quality data gathered at the Bologna airport synoptic meteorological station was collected and posteriorly simulated results are temperature maps and temperature at the receptor sites. The collected data will be released after the completion of the project and the results will be published in an academic journal.

- **Model for the Dublin low boundary wall** aiming to obtain a Computational Fluid Dynamics (CFD) model for the low boundary wall (LBW) intervention in Dublin, and to assess its effectiveness as an air pollution passive control strategy (PCS). The input data was total NO_x and NO concentration data, wind direction, and traffic volume. NO_x and NO data was collected for two weeks in December 2018 by using Chemiluminescent NO/NO₂/NO_x Analyser (Model 200EU). Wind velocity and direction was collected with a wind anemometer (Model 05103L) at an elevation of 19.3m above street level, located on the roof of a building on Pearse street. Finally, traffic volume data collected manually during the same period. The result is a CFD simulation with pollutant concentration profiles in the street canyon. Data is stored on the University College of Dublin (UCD) premises and it will be released after the completion of the project in the Zenodo platform.
- **Integrated transport behavioural simulations by the University of Hasselt (UH)**, using multiple available datasets in order to build an estimation of traffic volume and network travel time. All the simulation details including the purpose, procedure used and results have been documented in D5.0 Prototype of a fully Integrated behavioural (data-driven) simulator. A variety of datasets has been gathered to estimate sub-models within the activity-based model and also to calibrate the outputs derived from the FEATHERS-MATSIM Integration. The simulations primarily used as an input the Onderzoek Verplaatsings Gedrag (OVG) dataset, although additional datasets were also gathered and listed below, all of which are publicly available:
 - GPS (Global Positioning System) based activity-travel routine data was used and was collected as part of Task 4.1
 - General Transit Feed Specification (GTFS) data with information about public transport routes, stops and their schedule in a standard format

- The Centraal Referentie Adressen Bestand (CRAB) with detailed land use data in terms of residential units and firms and shop addresses within Flandes.

The OVG dataset was generated with a survey that contains details of the personalized mobility routine of individuals in the Flemish region of Belgium, as well as information about household characteristics of a sampled individual, their personal characteristics and at least one-day trip diary information. This dataset was originally owned by the Department of Mobility and Public Works of Flandes, being the sole role of UH to post-process it. For this reason, this dataset cannot be published fully, although some anonymized aggregated statistics will be made publicly available.

Finally, the above listed geospatial datasets were complemented by Open Street Map (OSM) data, which is open access data publicly available.

- **Sensor Analysis Framework** gathers all the observational and experimental data from the iSCAPE sensors stored on the SC Platform as well as reference data from high-end sensors co-located with them. It's purpose is to compare and contextualize the sensors data. The framework has three primary functions: evaluate the performance of the sensors by comparing the data with collocated reference equipment, obtain insights on the collected data by presenting the results in a meaningful way, and to improve the performance of the sensors by using calibration and correction models.

This dataset contains sensor data from the Living Lab Stations deployed in different iSCAPE cities in co-location with high-end equipment in order to improve the sensor understanding and calibration. The stations data is available in the SC Platform and will be released as part of a publication which aims to make the calibration methods available. The readings contain all the raw data from the available sensors in the stations, referring to CO, NO₂ and O₃ readings for the electrochemical units, and PM/PN metrics for the particle matter ones. Reference equipment data is, in some cases, provided by local authorities and they will be released according to the public authority data strategy. In other cases,

reference equipment data is managed by the project stakeholders, and will conform part of the dataset made publicly available in the above mentioned publication.

Additionally, the resulting models conform a dataset that can be used to improve the sensor readings. Furthermore, based on this data, predictive models are used to anticipate the behaviour of sensors and detect possible outliers. All these models are currently stored on-premises at IAAC and will be publicly released by the end of the project.

The framework is open-source software released under a GPL v3.0 license and was built mainly for the project. The results are already being exploited at IAAC for commercial projects. The software is built as a collection of Python tools that can be interfaced using Jupyter Notebooks and shell commands. The tools are built using state of the art data science methods to foster researchers adoption and contributions.

DATA ARCHIVAL AND UTILITY

During the project, the collected data and its models was managed by the project partners on-premises. As stated in the subsection above, due to the wide variety of datasets and sources, careful archival and publication has to be considered to comply with GDPR rules and data ownership. This is detailed per-dataset in the previous subsection. All publically available data will be published on the Zenodo platform or in OpenAIRE compliant databases upon the completion of the project.

1.2 Operational data

Operational data includes the practices carried out to successfully implement the project and oversee its operation, as well as events, publications and platform usage data.

1.1.1 Project management and monitoring

DATA COLLECTION PURPOSE AND RELATION TO THE PROJECT OBJECTIVES

To implement the project as outlined in the Grant Agreement, several data lists and documents containing institutional information and contact information of the individuals contributing to the project have been created. They include names, affiliations, emails and phone numbers of those involved in the project. Their contact details are shared with the project partners and the funding agency in cases specified in the Grant Agreement only. Personal data collected from the iSCAPE project partner employees are processed safeguarding the privacy and rights of individuals.

DESCRIPTION OF DATA TYPES

- **Project contact list:** The project contact list, which includes a name, affiliation, email address and/ or phone number of the individuals contributing to the project, has been created and stored in the project Google Drive folder. Only those individuals who are involved in the implementation, management or monitoring of the project have access to the folder. It is the responsibility of each iSCAPE partner to ensure that the project contact list is accurate and up to date. The project contact list is not publicly available, and after the project, the project contact list will be destroyed.
- **Internal mailing lists:** A number of internal mailing lists have been created to facilitate internal communication which plays a crucial role in ensuring project success. These mailing lists were created for 1) project-wide communication; 2) communication of the iSCAPE working groups (WP mailing lists); and 3) communication of the iSCAPE management bodies based on the informed decision of the project partners. The mailing lists

are managed by UCD using webarch.net service¹¹. It is the responsibility of each iSCAPE partner to inform UCD about changes in their personnel to keep the mailing lists up to date. The internal project mailing lists are not publicly available, and after the project, these mailing lists will be destroyed.

- **Internal documents:** Several internal documents which contain personal data, for example, names and email addresses, have been created during the project. Such documents include meeting minutes, questionnaires and registration forms developed to gather information and feedback from the project partners, as well as to organize internal project meetings. A project partner, coordinating such activity, may grant their personnel and other project partners access to such data only when it is necessary for implementing, managing or monitoring the project. The internal documents are not publicly available and only those internal documents which are necessary to fulfil the obligation to keep records and other supporting documentation as stated in Article 18 of the Grant Agreement, will be kept for a period of five years after the payment of balance to prove the proper implementation of the project and costs the project partners declare are eligible. It is the responsibility of each project partner to ensure appropriate data storage security for such documents.
- **Photos and videos:** During the project, a significant number of photos and videos have been taken to document project activities, including internal project meetings, city visits and field campaigns. With the consent of the project personnel, some of the photos and videos include identifiable people in action. Such photos and videos are stored in the project Google Drive folder in addition to the data storage systems used by the project partners. Only those individuals who are involved in the implementation, management or monitoring of the project have access to the project Google Drive folder. However, with the consent of individuals, some of the project photos or videos might be publicly disseminated using the project website and social media.

DATA ARCHIVAL AND UTILITY

¹¹ for more details see <https://www.webarch.net/security>

Data lists and documents containing institutional information and contact information of the individuals contributing to the project will not be made publicly available, given that such lists and documents were created only for the project implementation, management and monitoring.

1.1.2 Research activities involving human subjects

DATA COLLECTION PURPOSE AND RELATION TO THE PROJECT OBJECTIVES

iSCAPE is fully committed to the advancement of ethical research, ensuring that all research activities undertaken by the iSCAPE project partners which involve human subjects and personal data are undertaken in a way that safeguards the dignity, rights, safety, and privacy of those involved. As research activities conducted by the interdisciplinary team play a crucial role in achieving the project's objectives, the Research Ethics Framework (D10.1 - D10.3, available for internal use only) was developed to provide an overarching framework and establish a unified approach for iSCAPE research ethics. As specified in the document, the project partners undertake to obtain formal ethical approval from their organization's Research Ethics Committee or an equivalent body responsible for research ethics, following their internal procedures as long as their practice is in compliance with applicable international, EU and national law, and follow the standards and best practices of ethical research. This includes meeting requirements of informed consent. Examples of the information sheets and consent forms used for the research activities are provided in the Research Ethics Framework.

DESCRIPTION OF THE DATA TYPES

During the course of the iSCAPE project, several research activities involving human participants and collecting and processing of personal data have been planned. These activities have been carried out in the following WPs:

- WP2 - while engaging stakeholders and citizens in iSCAPE Living Lab activities (for example, the Citizen stories study run by FCC);

- WP4 - while implementing the behavioural interventions and solutions (for example, the Behavioural studies run by UH);
- WP5 - while assessing socio-economic impacts and benefits of the proposed passive air pollution remediation strategies and behavioural change initiatives (for example, the Socio-economic impact research run by T6).

Personal data in research studies have been collected following the personal data protection principles, including the principles of purpose limitation and data minimisation. Depending on the purpose and the nature of the planned research activities, personal data collected might include a name, contact information such as email and GPS-based travel data. Such personal data have been collected based on informed consent. Informed consent requires that participants have a genuine understanding of the research, which involves full disclosure of information about the research to potential subjects including an adequate understanding of the research procedures, the risks and benefits of the research, rights of the subjects, and the voluntary nature of participation. More details on the process of seeking informed consent are provided in the Research Ethics Framework (D10.1 - 10.3).

DATA ARCHIVAL AND UTILITY

Personal data provided by research participants has been kept strictly confidential and only the researchers involved in the studies had access to it. Research data has been processed as specified in the information sheets and consent forms developed for each research activity which involves human subjects and collection of their personal data. It might include data sharing with other organizations (for example, other iSCAPE project partners), however, this requires data anonymization prior to proceeding with any further data processing.

1.1.3 Public events organized by the project

DATA COLLECTION PURPOSE AND RELATION TO THE PROJECT OBJECTIVES

To implement the project as outlined in the Grant Agreement, a number of public events have been organized. These include the events organized by the iSCAPE

Living Labs such as co-creation and citizen science workshops, as well as public events organized by the project, for example, the project mid-term and final events. Personal data provided by participants voluntarily are not used by the project partners for any other purposes than the organization and administration of these events.

DESCRIPTION OF THE DATA TYPES

Event registration forms: Participant contact information (name, email address and/or phone number) is typically collected through event registration forms to organize and administer public events organized by project partners. Online registration forms are created using services of well-established and secure platforms such as Eventbrite (for more details see [Eventbrite Security and Safety Guide](#)) and Google Forms ([for more details see Google Privacy Policy](#)). A project partner, organizing a public event, may grant their personnel access to registration lists only when it is strictly necessary for organizing and administering public events.

Photos and videos: During the public events, some photos or videos which include identifiable people in action might be taken. With the consent of individuals, some of the public event photos or videos might be publicly disseminated using the project website and social media.

DATA ARCHIVAL AND UTILITY

Registration lists will not be made publicly available, however, they might be kept for a certain period of time by a project partner if required so by their accounting and bookkeeping requirements to comply with relevant national and local laws or to fulfil the obligation to keep records and other supporting documentation as stated in Article 18 of the Grant Agreement. It is the responsibility of a partner, organizing a public event, to ensure that registration information available on the online platforms is destroyed after the event and that only exported files are securely stored by the partner to fulfil the accounting and bookkeeping requirements.

1.1.4 Project online platforms

DATA COLLECTION PURPOSE AND RELATION TO THE PROJECT OBJECTIVES

Communication and dissemination of the project-related information, including news, project results and deliverables are of significant importance in achieving the desired project impacts. Therefore, a number of online communication channels and platforms were established to reach a great variety of project target audiences.

DESCRIPTION OF THE DATA TYPES

Project website and Virtual Living Lab platform: The website is hosted by UCD, the project coordinator, while the website development and content is managed by T6. The disclaimer and privacy policy¹² provides detailed information on what type of data, for what purpose and how this data is being used when people are visiting the project website and the Virtual Living Lab platform. In addition, it provides detailed information on the use of external services such as MailChimp¹³ and Google Analytics¹⁴.

- Contact form: As stated in the privacy policy, if an email is sent using the online contact form available at the project website, the sender's email address and any other personal information were only used for the individual correspondence with the sender. Data entered by the sender has been used only by UCD and has been treated confidentially according to the prevailing regulations.
- Newsletter mailing list: As stated in the privacy policy, the website uses MailChimp to send the iSCAPE newsletter. The privacy policy further specifies that data entered in the registration form has not been used for any other purpose than sending of the newsletter. Once a person unsubscribes from the newsletter (link provided at the end of each newsletter), data has been deleted permanently.

¹² <https://www.iscapeproject.eu/disclaimer-privacy-policy/>

¹³ <https://mailchimp.com/>

¹⁴ <https://analytics.google.com/analytics/web/>

- Google Analytics: As stated in the privacy policy, the website uses Google Analytics, a web analytics service provided by Google, Inc. More detailed information on how Google Analytics uses cookies to help the website analyse how users use the site and how such information generated by the cookie may be used by Google is provided in the privacy policy.

SC Platform: The SC Platform manages all the project sensors data. The platform is a front and backend solution for ingesting, storing and interacting with sensor data with a particular focus on crowd sensing applications. The platform is used by the Citizen Kits and Living Lab Stations across the different iSCAPE Living Lab and manages the end-to-end flow from users registration to data ingestion, visualisation and storage. The main components are the Smart Citizen Website, the Smart Citizen Engine API and the Onboarding app. The platform has been upgraded to fulfil all the project requirements with a particular emphasis on the GDPR compliance. IAAC is the platform data controller and is responsible for any issues related to personal data management. The policy section¹⁵ provides detailed information on what type of data, for what purpose and how this data is being used when people are visiting and contributing data to the platform.

DATA ARCHIVAL AND UTILITY

Personal data provided through online communication channels and platforms will not be made publically available, given that this information was collected only for a limited and specific purpose.

¹⁵ <https://smartcitizen.me/policy>

2. FAIR Data

iSCAPE advocates for openness and effectiveness in sharing research data. Hence, the project strives to make data collected, processed or generated during the project as available as possible, following the Fair Data Principles while safeguarding individual rights to personal privacy and personal data protection. These principles guide scientific data management and stewardship and are relevant to all stakeholders in the current digital ecosystem.

The project has followed the new European Commission Guidelines on FAIR Data Management in Horizon 2020¹⁶, whose primary goal is to improve the findability, accessibility, interoperability, and reuse of digital assets. This is implemented by agreeing upon standard and measurable set of principles that are referred as the FAIR Data Principles. Wilkinson et al. (2016)¹⁷ describe the FAIR Data Principles as follows:

FINDABLE	INTEROPERABLE
<ul style="list-style-type: none"> ● Metadata are assigned a globally unique and eternally persistent identifier. ● Data are described with rich metadata. ● Data (and metadata) are registered or indexed in a searchable resource. ● Metadata specify the data identifier. 	<ul style="list-style-type: none"> ● Data (or metadata) use a formal, accessible, shared and broadly applicable language for knowledge representation. ● Data (or metadata) use vocabulary that follows the FAIR principles. ● Data (or metadata) include qualified references to other (meta)data.
ACCESSIBLE	RE-USABLE
<ul style="list-style-type: none"> ● Data (and metadata) are retrievable by their identifier using a 	<ul style="list-style-type: none"> ● Data (or metadata) have a plurality of accurate and relevant attributes. ● Data (or metadata) are released

¹⁶

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

¹⁷ <https://paperpile.com/c/o1bTXg/nVV0>

<p>standardized communications protocol.</p> <ul style="list-style-type: none"> • The protocol is open, free, and universally implementable. • The protocol allows for an authentication and authorization procedure, where necessary. • Metadata are accessible, even when the data are no longer available. 	<p>with clear and accessible data usage licenses.</p> <ul style="list-style-type: none"> • Data (or metadata) are associated with their provenance. • Data (or metadata) meet domain-relevant community standards.
--	--

Table 1 Fair Data Principles

2.1 Making data findable, including provisions for metadata

Provisions for findable (meta)data were developed during the iSCAPE project. This was put in practice by the creation of the iSCAPE Datasets Tool, which aimed to create metadata and data trace codes to ensure traceability, which are described in the subsection below. Furthermore, provisions for metadata were developed and described in the following subsections, with the addition of a set of metadata keywords were created to classify and find data for later re-use.

2.1.1 The iSCAPE Datasets Tool

At the beginning of the project, the iSCAPE Datasets Tool (iSCAPE DT) was developed to collect information and trace the project datasets. Accessible by the consortium members only, this tool was used throughout the course of the project to identify project datasets and partners responsible for each dataset, including datasets containing personal data to guarantee the traceability of all the project data. In accordance to the Data Management Plan, the following information was collected using this tool:

- Dataset
- Short Description
- Trace Code (see. Table 1)

- Type (see. Table 2)
- Source (see. Table 3)
- Storage Platform (see. Table 4)
- Current Accessibility (see. Table 5)
- Archiving Access Policy
- Data Controller Responsible Partner
- Data Controller Contact Information
- Data Collection Status

The database was designed such that it automatically provides a unique trace code (see Table 2) and registry date for each dataset. The database was updated by project partners regularly. iSCAPE DT was also used to update the Data Management Plan.

TRACE CODE	DESCRIPTION
PD	Personal data
SV	Single value
TS	Time series
MV	Multiple Values
SC	Source Code
MA	Maps

Table 2 Project datasets trace codes

TYPE
Observational
Experimental
Models or simulation
Compiled data

Table 2 Project dataset typologies

SOURCE
Living Lab Stations
Citizen Kits
Field Measuring Equipment
EPA Monitoring Station
GPS Tracker
Online Usage Metrics
Interviews
Surveys
Mixed Environmental and Spatial Data
Online Usage Metrics
Interviews
Surveys
Mixed Environmental and Spatial Data
Participant Registration Platforms
Project management & monitoring
Project events
Project online platforms
Others

Table 3 Project datasets source types

STORAGE PLATFORM
On-premises
Consortium Google Drive
SC Platform

Sensor Analysis Framework
Github
Google Analytics
Other platforms

Table 4 Project datasets storage platform

ACCESSIBILITY
Publicly Accessible Anytime
Publicly Accessible After Pending Publication
Publicly Accessible After Anonymization and Aggregation
Publicly Accessible Once the Project Finishes
Non-public disclosure (Personal Data)
Non-public disclosure (Patent Pending)
Under discussion (for planned activities)

Table 5 Project datasets accessibility levels

2.1.2 Provisions for metadata

Considering best practice, iSCAPE adopted the guidelines of Natural Environment Research Council (NERC)¹⁸ for metadata provisions to describe its project datasets. In Table 6, the different metadata topics are addressed for each activity, providing guidelines for each specific case.

METADATA TOPIC TO ADDRESS	DETAILS
Experimental Design / Sampling Regime	Metadata should provide details of the experimental design and/or sampling regime, where applicable
Collection / Generation / Transformation Methods	Metadata should provide covering methods used for collection of samples/observations. Alternatively, where data values are derived/generated/ transformed, then the details of how this was achieved should be provided.
Fieldwork and / or Laboratory Instrumentation	Information should be supplied on instruments/machines used for collection/analysis of samples/observations where relevant.
Calibration Steps and Values	Details of the steps taken to calibrate any instruments/machines used, including the use of blanks, and the values used for calibration should be provided.
Nature and Units of Recorded Values	Metadata should describe the nature of the recorded values and used units to unambiguously define what has been measured and recorded in the dataset.
Analytical Methods	Full descriptions of any analytical methods used to generate the data values included in the dataset should be included.
Quality Control	Any quality control measures to ensure the quality of the data values

¹⁸ Natural Environment Research Council <http://www.nerc.ac.uk/>

	included in the dataset should be detailed.
Format of Stored Data	The format which was used to store the dataset during the lifetime of the project, and the format in which the dataset is made publicly available, if different, should be named in the contextual metadata.

Table 6 Guidance on metadata topics to address

2.2 Making data openly accessible

Following the Fair Data Principles recommended by the EC H2020 guidelines, many importance was put on adequate data archiving and public accessibility, in particular for observational and experimental environmental data as they are in the core of the project.

2.2.1 Observational and experimental environmental data

Data sharing, accessibility and archiving are essential topics on any research project with multiple stakeholders, being environmental sensor data one of the primary data types we have designed an end to end data flow for it.

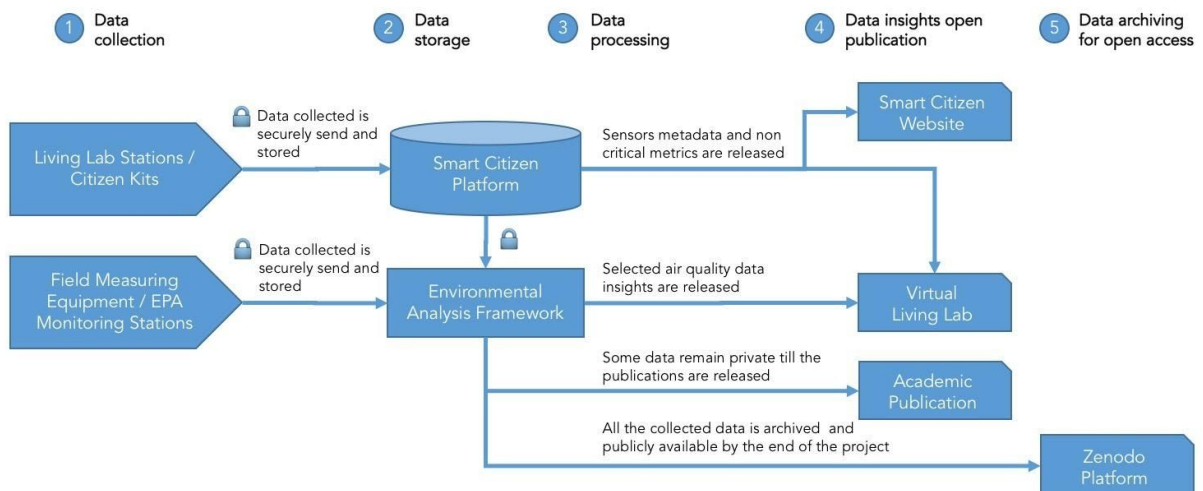


Figure 2 Environmental Sensors Data Lifecycle

Figure 1 represents the data flow for all the environmental sensor data collected on the project, considering all the steps it follows until its public release. Firstly, a layer of security was added to ensure all the data collected is securely sent and stored

privately. Secondly, the first version of the document thought the SC Platform as the single data repository. On the present version, we introduce a platform submodule, the Sensors Analysis Framework (SAF). The module contains all the algorithms that were developed to process the sensor environmental data, and it was stored on-premises during development. Resulting code and models will be stored in the Zenodo platform and, once the core functionalities of the module are safely deployable, it will be integrated into the core of the SC Platform.

The data public release process involves different steps. First, to reach citizens and local communities, especially from the six iSCAPE cities, the Virtual Living Lab (VLL)¹⁹ was developed in Task 8.1. That is especially important since having a sense of ownership over a technology intervention has been associated with sustained community engagement (Balestrini et al. 2014). The VLL works together with the Smart Citizen Website where participants can browse and explore their sensor data. On the Smart Citizen Website, the metadata of the sensor such as location, sensor types and some sensor metrics is available in near real-time and facilitate the exploration of data with other contextual data (maps, keywords).

Upon the project completion, all the data from the SAF will be made publicly available on the Zenodo digital repository²⁰. However, that method does not provide real-time neither interactive access to the environmental sensor, for that reason, some of the data is also available using the Smart Citizen API.

¹⁹ <https://livinglabs.iscapeproject.eu/>

²⁰ <https://zenodo.org/record/2566518>

2.3. Making data interoperable

All the data aimed to be publicly released within the project is managed using well-known formats or documented accordingly, including extensive metadata.

Most of the data standard regarding environmental sensor data is accessible at the Smart Citizen API, a JSON REST API web service fully documented on the platform webpage. The service also supports batch data downloads as CSV. Metadata is accessible via the same interface. Data related to sensor location, sensor type or manufacturer is available within the same API. Connectors for other data standards can be built on top of the existing API. On top of that, all the data used to assess the performance of the sensor and develop the calibration models are managed on the Sensor Analysis Framework. Data standards related to other project activities might vary depending on partners' internal tools and methodologies. As part of the projects, open data access and sustainability strategy once the project finishes, environmental sensor data stored in the SC Platform and the SAF will also be programmatically downloaded and archived. To ensure the project data is made available using the same standards towards consistency and usability all the datasets should provide the following metadata (see 2.1.2 Provisions for metadata).

2.4. Increase data re-use

Following the Fair Data Principles, all the collected data which can be publicly disclosed will be released by the end of the project and made available to the Zenodo digital repository²¹ due to its principles for open access and open data, providing a catch-all repository for EC funded research. It provides integration of the FAIR data principles, OpenAIRE support and tools for scholarly processes to handle data and software. We also consider the possibility that partners can use existing data repositories already provided by their universities when they are compliant with OpenAIRE policies. It is important to notice that excludes personal data only disclosable when anonymization and aggregation are possible. Furthermore, most of the environmental sensor data will be preserved on the SC Platform as long as the Smart Citizen project by IAAC exists. That guarantees that any applications developed on top of the SC Platform will remain accessible. Regarding data archiving and safety, the SC Platform relies on a distributed cluster of Cassandra databases²² split across different datacenters in Europe.

As already described in the previous section 2.1.1 the iSCAPE DT has been implemented where each partner must register all the datasets they are currently collecting (see Annex 1 for the current database snapshot). Information to be filled includes the release timeframe and the disclosure process.

The usage of an external platform as Zenodo, allows for an increase and reuse protocol and to guarantee consistency on the use, reference, and preservation of the sources once they are archived by the end of the project. Having followed that protocol, all the project partners have uploaded all the datasets registered on the iSCAPE Datasets Tool and indicated as “Publicly Accessible” on the Zenodo platform²³. Table 7 aims to assess partners to map the existing datasets information with the Zenodo requirements.

As a default standard, the project uses Creative Commons Licenses²⁴ for public data sharing. In particular, an *Attribution 4.0 International*²⁵ License was selected in order to provide third-party researchers the right to redistribute the material in any medium or format. This license also gives the possibility to build upon the material for any purpose, even commercially. The license however, requests to attribute the

²¹ As a European Commission supported the initiative and technically supported by CERN, we trust this as the best way to ensure access to the generated data remains long after the project ends.

²² <https://cassandra.apache.org>

²³ Zenodo Platform upload URL <https://zenodo.org/deposit>

²⁴ <https://creativecommons.org/>

²⁵ <https://creativecommons.org/licenses/by/4.0/>

original source, giving appropriate credit, provide a link to the license, and indicate if changes were made to the original publication. Whenever other copyleft (or copyleft compatible) licenses exist more specific to the content to be publicly released, we suggest the use of them. For example, most of the project software²⁶ has been released under GNU GPL licenses, while the hardware design files using the CERN Open License.

Annex 1 contains a list of all the publicly released datasets with their openAIRE compatible repositories URLs. In the release process, multiple datasets have been consolidated into collections to facilitate the reusability of them. For that reason, each consolidated dataset includes the compilation of numerous ISCAPE trace codes and are mapped on the Annex 1 Table. Also, by promoting the archival of all the datasets under the ISCAPE Zenodo community, we increase the accessibility by making it available under a single URL²⁷.

²⁶ See Annex 1 - Software and Hardware table

²⁷ iSCAPE Collection on the Zenodo platform <https://zenodo.org/communities/iscape/>

ZENODO FIELDS	INFORMATION SOURCE
Digital Object Identifier (DOI)	Already provided by Zenodo
Publication date	Already provided by Zenodo
Title	Description from the iSCAPE DT
Authors	Partner name of Authors, we recommend ORCID ²⁸
Description	Extended description based on the iSCAPE DT. Use Table 1 recommendations.
Keywords	Include the Trace Code Label and the Label from the iSCAPE DT and any other relevant keywords
Additional notes	Data Set field from iSCAPE DT
Grant	Add the iSCAPE EU Project Grant 689954
Access right	Set to Open Access
License	Creative Commons Attribution 4.0 is the default iSCAPE license
Community	Search for the iSCAPE community https://zenodo.org/communities/iscape

Table 7 Dataset information in Zenodo

²⁸ ORCID iD (Open Researcher and Contributor ID) is a nonproprietary alphanumeric code to uniquely identify scientific and other academic authors and contributors.

2.5 Managing personal data

The project collects and uses personal data to implement the activities as outlined in the Grant Agreement Nr. 689954. The purposes of processing personal data include project management and coordination, project outreach, communication and dissemination activities, research activities, administration and compliance with the Grant Agreement and statutory obligations of the project partners, etc. Personal data collected by the iSCAPE project partners is processed safeguarding the privacy and rights of individuals following the applicable national laws and regulations in compliance with the EU General Data Protection Regulation (GDPR).

Article 5 of GDPR sets out seven key principles that iSCAPE adheres to as follows

Lawfulness, fairness and transparency

iSCAPE obtains and processes personal data fairly and in compliance with the Grant Agreement and statutory obligations of the project partners.

Purpose limitation

iSCAPE collects personal data for specified, explicit and legitimate purposes. Personal data will only be processed in a manner compatible with these purposes.

Data minimisation

Data collected by iSCAPE is adequate, relevant and limited to what is necessary in relation to the data collection purposes.

Accuracy

iSCAPE takes reasonable steps to ensure the accuracy of any personal data, as well as to periodically update the information.

Storage limitation

iSCAPE retains personal data no longer than for the purposes for which the personal data are processed.

Integrity and confidentiality

iSCAPE collects personal data securely and takes appropriate measures against unauthorised access to, or alteration, disclosure or destruction of data and against accidental loss or destruction.

Accountability

iSCAPE puts in place appropriate technical and organizational measures to demonstrate compliance with these principles, including the implementation of data protection policies.

3. Allocation of resources

As specified in the Grant Agreement, data management, including the development of the Data Management Plan as part of Task 3.1 with a dedicated project budget led by IAAC. The Data Management Plan has been a living document, and it was updated regularly by the Data Manager (IAAC) and reviewed by WP-leads to support the data management lifecycle for all data that will be collected, processed or generated by the project.

The responsibilities of the project Data Manager (IAAC) include:

- Oversight and coordination of data management in iSCAPE;
- Development of the Data Management Plan;
- Management of internal tools developed to ensure efficient and effective data management;
- Guidance on best practices in data management.

The responsibilities of the project partners include:

- Implementation of the Data Management Plan;
- Ensuring compliance with applicable international, EU and national law, including GDPR when collecting and processing personal data;
- Bearing the costs of the data management to implement the Data Management Plan.

With the EU GDPR having entered enforcement in May 2018, special care was put on the management of personal data. IAAC assisted the partners in collecting and storing personal data during the project.

To map out all data that will be collected, processed or generated by the project, the iSCAPE Datasets Tool has been developed as described in section 2.3.1. Information provided in there has been used to update the Data Management Plan, following the timeline specified below:

Review Dates	Project Month	DMP Version
December 2017	M16	V1.0 (Internal)

January 2018	M17	V2.0 (Internal)
February 2018	M18	V3.0
October 2018	M26	V4.1 (Internal)
February 2019	M30	V4.2 (Internal)
November 2019	M36	V4.3 (final)
May 2020	-	V5.0 (final)

Table 8 Review dates

To increase the value of long-term preservation in addition to reducing the costs associated with data management, free Open Science platforms such as Zenodo will be used. Developed based on best practice and following the [FAIR data principles](#), [Zenodo](#) provides an infrastructure for research projects to join in Open Science.

4. Data security

All personal data collected and processed during the project will happen in platform conforming with the EU General Data Protection Regulation (GDPR). Also, external platforms operated by US companies (Google Drive, Mailchimp and Google Analytics) are all ensured to comply with the US-EU Privacy Shield. A particular emphasis was set to ensure the SC Platform, complies with GDPR. All the other experimental data, as well as models and simulations, was managed by partners individually and stored safely on-premises. We have implemented a policy around sensitive data, in particular, personal data, to minimize transfer and ensure storage is on partners premises. Now, as the end of the project arrives, all the publicly available data will be made available on open access certified repositories as explained in 2.4. Increase data re-use.

5. Ethical aspects

As previously stated, research activities conducted by iSCAPE's multidisciplinary team play a crucial role in achieving the project's objectives. Therefore, an overarching framework to establish a unified approach for iSCAPE research ethics was developed. In addition, with the Research Ethics Framework (D10.1 - D10.3), iSCAPE aims to demonstrate that the project partners are fully committed to the advancement of ethical research, ensuring that all research activities undertaken by the iSCAPE project partners which involve human subjects and personal data are undertaken in a way that safeguards the dignity, rights, safety, and privacy of those involved. As specified in the Research Ethics Framework, the project partners conducting such research activities will obtain formal ethical approval from their organization's Research Ethics Committee or an equivalent body responsible for research ethics, following their internal procedures as long as their practice is in compliance with applicable international, EU and national law, and follow the standards and best practices of ethical research. That includes meeting requirements of informed consent. Examples of the information sheets and consent forms used for the research activities are provided in the Research Ethics Framework (D10.1 - D10.3).

6. References

Commission, E, 2018. *Open access & Data management - H2020 Online Manual*. [online] Ec.europa.eu. Available at: http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm [Accessed 16 Feb. 2018].

Commission, E, 2018. *EU General Data Protection Regulation (GDPR)* [online] eugdpr. Available at: <https://www.eugdpr.org/> [Accessed 16 Feb. 2018].

IT Governance, 2017. *EU General Data Protection Regulation Infographic* [online] eugdpr. Available at: <https://itgovernance.eu/eu-gdpr-infographic> [Accessed 16 Feb. 2018].

Wilkinson, M.D. et al., 2016. *The FAIR Guiding Principles for scientific data management and stewardship*. *Scientific data*, 3, p.160018.

Balestrini, M. et al., 2014. *Understanding sustained community engagement*. In Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14. Available at: <http://dx.doi.org/10.1145/2556288.2557323>.

6. Annex 1: Datasets

The following tables contain a summarized version of the iSCAPE Datasets Tool described in section 2.1.1 The iSCAPE Datasets Tool. More importantly, it also includes the archival URL on OpenAIRE compatible repositories of all the publicly released content.

Environmental research datasets

The following section covers the observational, experimental, models and simulation data generated within the project related to environmental research. All the datasets will be made public as part of the FAIR Data Policy.

Consolidated Environmental Datasets

The following table contains all the publicly available datasets consolidated and stored within an OpenAIRE Compatible Repository, primarily Zenodo²⁹, and all of them are citable using a DOI. As most of the datasets are combined with other related datasets, we provide the original ISCAPE Trace Code to locate them on the original table also available after. The full DOI, Publication date, Title, Authors, Description, Keywords, Access right and Licensing is available in the datasets metadata on the linked OpenAIRE repositories.

Title	iSCAPE Trace Code	Description	OpenAIRE Compatible Repository
iSCAPE Low Cost Sensor Development Data	DS_TS_058, DS_TS_059, DS_TS_060, DS_TS_061, DS_TS_062, DS_TS_063, DS_TS_064, DS_TS_065, DS_TS_066, DS_TS_067, DS_TS_068, DS_TS_069, DS_TS_070, DS_TS_071, DS_TS_072, DS_TS_073, DS_TS_074, DS_TS_075, DS_TS_076, DS_TS_077, DS_TS_078, DS_TS_079, DS_TS_080, DS_TS_081, DS_TS_082, DS_TS_083, DS_TS_084, DS_TS_085, DS_TS_086, DS_TS_087, DS_TS_088, DS_TS_089, DS_TS_090, DS_TS_091,	Sensors test data used for the low-cost sensor development during the iSCAPE project. The dataset is divided in a series of tests, each of them described on a yaml file with the test name. Each csv file contains time series data of each experiment, and the yaml	http://doi.org/10.5281/zenodo.3570688

²⁹ iSCAPE Collection on the Zenodo platform <https://zenodo.org/communities/iscape/>

	DS_TS_092		
iSCAPE Outdoor Sensor Deployment Data	DS_TS_049, DS_TS_050, DS_TS_051, DS_TS_052, DS_TS_053, DS_TS_055, DS_TS_056, DS_TS_057, DS_TS_048	Sensors deployment data using low cost sensors during the iSCAPE project. The dataset is divided in a series of deployments, each of them described on a yaml file with the test name. Each csv file contains time series data of each experiment, and the yaml f	http://doi.org/10.5281/zenodo.3570700
iSCAPE Citizen Science Workshops Data	DS_TS_093, DS_TS_094, DS_TS_095, DS_TS_096, DS_TS_097, DS_TS_098, DS_TS_047	Sensor data recorded during the Citizen Science Workshops during the iSCAPE project . Several workshops were held as part of the Citizen Science activities during the project in the cities of Vantaa, Dublin, Bologna, Bottrop, Hasselt and Guildford.	http://doi.org/10.5281/zenodo.3570680
Matsim Simulation Aggregated Results Bologna	DS_MV_014	Contains aggregated results for the following scenarios based on Matsim simulations.	http://doi.org/10.5281/zenodo.3604314
Matsim Simulation Results (disaggregated)	DS_MV_017	Matsim Simulation results at disaggregate level are based on the following 2 datasets files for each city	http://doi.org/10.5281/zenodo.3604336
Matsim Simulation Aggregated Results Vantaa	DS_MV_015	Aggregated results for the following scenarios based on Matsim simulations	http://doi.org/10.5281/zenodo.3604320
Information Based Behavioural Intervention Study (Hasselt case)	DS_PD_010, DS_MA_002, DS_PD_015	Datasets acquired and generated during the implementation of behavioural intervention study in Hasselt	http://doi.org/10.5281/zenodo.3565621
Information Based Behavioural Intervention Study (Guildford case)	DS_PD_014, DS_PD_019	Datasets acquired/generated during the implementation of behavioural intervention study in Guildford	http://doi.org/10.5281/zenodo.3565741
Behavioural Simulator Matsim Input Data	DS_MV_016	Behavioural simulator input data files	http://doi.org/10.5281/zenodo.3604333
Information Based Behavioural Intervention Study (Bologna case)	DS_PD_011, DS_MA_003, DS_PD_016	Datasets are acquired and generated during the implementation of behavioural intervention study in Bologna	http://doi.org/10.5281/zenodo.3565699
Living Lab Station data Jun-Oct 2019, Sutherland Memorial Park, Guildford	DS_TS_099	Data collected before and after hedge in Sutherland Memorial Park, Guildford using Living lab station during Jun-Oct 2019.	http://doi.org/10.5281/zenodo.3609187

All site data for Citizen Science (Guildford case study)	DS_TS_100	Data collected during citizen science activities in Guildford using low-cost sensors.	http://doi.org/10.5281/zenodo.3609191
Air quality Modelling data for Guildford City	DS_TS_101	Simulation data for Guildford City using ADMS-Urban at 1 km spatial resolution.	http://doi.org/10.5281/zenodo.3609197
Field campaign data for evaluating green infrastructure in Guildford	DS_TS_102	Field campaign data collected using reference instruments before and after hedges (Green infrastructure) in Guildford by University of Surrey.	http://doi.org/10.5281/zenodo.3609193
Living Lab Station data Feb-Sept 2019, Stoke Park, Guildford	DS_TS_103	Data collected before and after hedges in Stoke Park, Guildford using Living lab stations.	http://doi.org/10.5281/zenodo.3609183
Repository for data used in STOTEN article "Impacts of town characteristics on the changing urban climate in Vantaa"	DS_MV_017	Dataset contains data used in the Science of The Total Environment journal article "Impacts of town characteristics on the changing urban climate in Vantaa" (https://doi.org/10.1016/j.scitotenv.2020.138471).	http://doi.org/10.5281/zenodo.3753897
Matsim Simulation Aggregated Results Hasselt	DS_MV_013	Dataset contains aggregated results for multiple scenarios based on Matsim simulations.	http://doi.org/10.5281/zenodo.3604310
A route to school informational intervention for air pollution exposure reduction	DS_PD_020	Datasets are gathered during the implementation of route to school intervention study in Antwerp (Belgium).	http://doi.org/10.5281/zenodo.3565751
Field environmental data used to study the effects of Photocatalytic Coatings in reducing NOx pollutant concentrations	DS_TS_038, DS_TS_039, DS_TS_040, DS_TS_041, DS_TS_042, DS_TS_043, DS_TS_046,	Observations of air quality collected within one experimental campaign conducted in August 2018 at Lazzaretto campus of the University of Bologna to verify the effectiveness of photocatalytic coatings in reducing pollutant concentrations. a, which can be done by including a proper citation in any documents or publications using these data.	http://doi.org/10.6092/unibo/amsacta/6281
Field environmental data used to study the effects of Vegetation Urban Street Canyons	DS_TS_026, DS_TS_027, DS_TS_028, DS_TS_029, DS_TS_030, DS_TS_031, DS_TS_032, DS_TS_033, DS_TS_034, DS_TS_035, DS_TS_036, DS_TS_037, DS_TS_044, DS_TS_045	Observations of air quality collected within two experimental campaigns conducted in summer 2017 (07 August 2017-26 September 2017) and winter 2018 (15 January 2018-15 February 2018) in two street canyons in the city of Bologna, to verify the	http://doi.org/10.6092/unibo/amsacta/6317

		effectiveness of vegetation in altering ventilation levels and pollutant concentration in real-world street canyons.	
--	--	--	--

Traced Environmental Datasets

The following table contains the original environmental datasets traced during the project now consolidated and publicly disclosed on the table above. The table is kept for legacy purposes to follow the previous versions of the document.

Data Set	Description	Trace Code	Type	Source
DS_MA_002	Behavioural Intervention study - pollutant concentration maps - Dublin	Maps	Environmental	Mixed Environmental and Spatial Data
DS_MA_003	Behavioural Intervention study - pollutant concentration maps - Dublin	Maps	Environmental	Mixed Environmental and Spatial Data
DS_MA_004	Behavioural Intervention study - pollutant concentration maps - Dublin	Maps	Environmental	Mixed Environmental and Spatial Data
DS_MA_005	Behavioural Intervention study - pollutant concentration maps - Dublin	Maps	Environmental	Mixed Environmental and Spatial Data
DS_MA_006	Behavioural Intervention study - pollutant concentration maps - Guildford	Maps	Environmental	Mixed Environmental and Spatial Data
DS_TS_018	Projected meteorological variable data for Dublin by considering effect of climate change	Time Series	Environmental	Mixed Environmental and Spatial Data
DS_MV_007	CFD simulations of the low boundary wall intervention in Dublin	Multiple Values	Environmental	Mixed Environmental and

				Spatial Data
DS_TS_019	Dublin Field Experiment - NO2 data	Time Series	Environmental	Field Measuring Equipment
DS_TS_020	Dublin Field Experiment - NO data	Time Series	Environmental	Field Measuring Equipment
DS_TS_021	Dublin Field Experiment - NOx data	Time Series	Environmental	Field Measuring Equipment
DS_TS_022	Dublin Field Experiment - PM10 data	Time Series	Environmental	Field Measuring Equipment
DS_TS_023	Dublin Field Experiment - PM2.5 data	Time Series	Environmental	Field Measuring Equipment
DS_TS_024	Dublin Field Experiment - Wind speed data	Time Series	Environmental	Field Measuring Equipment
DS_TS_025	Dublin Field Experiment - Wind Direction	Time Series	Environmental	Field Measuring Equipment
DS_TS_026	Bologna Field Experiment - NO data (summer 2017 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_027	Bologna Field Experiment - NO2 data (summer 2017 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_028	Bologna Field Experiment - CO data (summer 2017 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_029	Bologna Field Experiment - O3 data (summer 2017 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_030	Bologna Field Experiment - PM10 data (summer 2017 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_031	Bologna Field Experiment - PM2.5 data (summer 2017 campaign)	Time Series	Environmental	Field Measuring Equipment

DS_TS_032	Bologna Field Experiment - NO data (winter 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_033	Bologna Field Experiment - NO ₂ data (winter 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_034	Bologna Field Experiment - CO data (winter 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_035	Bologna Field Experiment - O ₃ data (winter 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_036	Bologna Field Experiment - PM ₁₀ data (winter 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_037	Bologna Field Experiment - PM _{2.5} data (winter 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_038	Lazzaretto Field Experiment - NO data (summer 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_039	Lazzaretto Field Experiment - NO ₂ data (summer 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_040	Lazzaretto Field Experiment - CO data (summer 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_041	Lazzaretto Field Experiment - O ₃ data (summer 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_042	Lazzaretto Field Experiment - PM ₁₀ data (summer 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_043	Lazzaretto Field Experiment - PM _{2.5} data (summer 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_044	Bologna field campaign - meteorological data (summer 2017 campaign)	Time Series	Environmental	Field Measuring Equipment

DS_TS_045	Bologna field campaign - meteorological data (winter 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_046	Lazzaretto field campaign - meteorological data (summer 2018 campaign)	Time Series	Environmental	Field Measuring Equipment
DS_TS_047	Environmental data collected as part of the Citizen science workshops in Bologna	Time Series	Environmental	Citizen Kits
DS_TS_048	Living Lab Stations Lazzaretto	Time Series	Environmental	Field Measuring Equipment
DS_MV_008	CFD simulations in Bologna	Multiple Values	Environmental	Mixed Environmental and Spatial Data
DS_MV_009	CFD simulations in Lazzaretto	Multiple Values	Environmental	Mixed Environmental and Spatial Data
DS_MV_010	ADMS urban simulations in Bologna	Multiple Values	Environmental	Mixed Environmental and Spatial Data
DS_MV_011	ADMS urban simulations in Lazzaretto	Multiple Values	Environmental	Mixed Environmental and Spatial Data
DS_TS_049	2018-02_EXT_BOLOGNA_TESTS_LB_MA (First iteration of iScape Station in Feb-2018 in Bologna)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_050	2018-03_EXT_DUBLIN_TESTS (Dublin tests with first sensors iteration)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_051	2018-09_EXT_BOLOGNA_TEST_WALL_MO (Bologna Lazzaretto site. Monitoring of Pureti photocatalytic coating)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_052	2018-09_EXT_BOLOGNA_TEST_WALL_RE (Bologna Lazzaretto site. Monitoring of Pureti photocatalytic coating)	Time Series	Environmental	Sensor Analysis Framework

DS_TS_053	2019-03_EXT_UCD_URBAN_BACKGROUND_API (Urban Background data at Dublin City Council.)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_054	2019-06_INT_TEST_STATION_21_CSIC (Palau Reial Barcelona. Urban Background monitoring station.)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_055	2019-03_EXT_UOS_URBAN_BACKGROUND_API (UOS urban background data - first hedges iteration)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_056	2019-05_EXT_UCD_TRAFFIC (Urban traffic data at M50 in Dublin)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_057	2019-06_EXT_HASSELT_SCHOOL (Hasselt test in school)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_058	2019-06_EXT_UOS_HEDGES_API_SECOND_BATCH (Hedges evaluation in second location)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_059	2017-06_INT_FIRMWARE_VALIDATION (Validation of 1.5 Firmware changes on kits)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_060	2018-01_INT_ALPHASENSE_VALID (Hardware validation with two set of sensors)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_061	2018-01_INT_BOLOGNA_RELEASE (Pre Bologna release validation with 6 sensors (4 kits and 2 alphasenses))		Time Series	Environmental	Sensor Analysis Framework
DS_TS_062	2018-04_INT_MICS_POWER_STEPS (Stabilised Gas (no specific composition) with variation of heater current with steps of 2h)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_063	2018-04_INT_MICS_STABLE (Stabilised over 6 days in controlled atmosphere for the SCK1.5)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_064	2018-06_INT_ALPHASENSE_STABILISED_TEST (Test out the W A currents using prototype board)		Time Series	Environmental	Sensor Analysis Framework
DS_TS_065	2018-06_INT_OUTDOOR_TESTS_GRAN_VIA (Collection of data for alphasense and mics calibration in the lab and gran via.)		Time Series	Environmental	Sensor Analysis Framework

DS_TS_066	2018-06_INT_OUTDOOR_TESTS_LAB (Collection of data for alphasense and mics calibration in the lab and gran via)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_067	2018-07_INT_MICS_VERSION_COMPARISON (Test comparing the mics on SCK V2.0 and SCK V2.0Rev1 to validate the hardware.)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_068	2018-07_INT_PMS5003_ON_OFF_PARTICLE (The purpose of these test is to determine the time needed to obtain representative readings in a one-shot mode)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_069	2018-07_INT_TEMP_CALIB (Temperature calibration test at different temperatures and enclosures situations. With and without (Time Series	Environmental	Sensor Analysis Framework
DS_TS_070	2018-07_INT_TEMPERATURE_TEST_STATION) (Temperature test in house with station)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_071	2018-08_INT_STATION_TEMPERATURE_COMPARISON (Weekend exposure of station + External dallas temperature and kit with external SHT31)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_072	2018-08_INT_STATION_TEST_SUMMER_HOLIDAYS (Long term exposure for late-stage development station.)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_073	2018-10_INT_PM_TEST_BATT (Test for checking battery influence on PM)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_074	2018-10_INT_TEST_TEMP_CHARGE (Test for checking temperature of external and internal SHT31 with the Charger at 98 to 100% charging rate)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_075	2018-10_INT_TEST_TEMP_CHARGE_CORRECTION_LEARNING (Test for checking temperature of external and internal temp...)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_076	2018-10_INT_TEST_TEMP_DALLAS_SHT31 (Test for checking the temperature correlation between Dallas and SHT31)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_077	2018-10_INT_TEST_TEMP_FAN (Test for checking temperature of SHT31 Dallas and Kit with and without fan pointing at it)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_078	2018-10_INT_TEST_TEMP_PM (Test for checking temperature of external and internal SHT31 with the PM ON. This is known to have an effect due to the step up. Different recordings with PM ON OFF charging and not.)	Time Series	Environmental	Sensor Analysis Framework

DS_TS_079	2018-12_INT_BENCHMARKING (TEST FOR BME680 SHT31 DALLAS AMS CCS811)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_080	2018-12_INT_CCS811 (Test for CCS811)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_081	2018-12_INT_NOISE_SAMPLING (Noise sampling Test)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_082	2019-01_INT_CCS811 (CCS811 sensor assessment)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_083	2019-01_INT_PM_PN_ANALYSIS (Calculate PM-PN size distribution)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_084	2019-01_INT_TEMP_EFFECT_CCS811 (Verify temperature effect of CCS811. Indoor conditions CCS811 facing the SHT31 directly)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_085	2019-03_INT_PROTO_V21-VALIDATION (First validation of V2.1 proto)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_086	2019-03_INT_STATION_V21-VALIDATION (First validation of V2.0 STATION)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_087	2019-04_INT_PROTO_V21-TEMP_TEST (Temperature comparison of V2.1 and external sensors)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_088	2019-04_INT_V20_T_H_TEST_MODES (Temperature comparison of V2.0 and external sensors - All Modes)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_089	2019-06_INT_TEST_STATION_21_CSIC (Co-location of the Station V2.1 in CSIC EPA)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_090	2019-08_INT_TEMPERATURE_DELTA (Evaluation of temperature delta for SCK 2.0)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_091	2019-11_INT_CHECK_PM_ISCAPE (4 Stations for iSCAPE with PM problems)	Time Series	Environmental	Sensor Analysis Framework

DS_TS_092	2019-12_INT_WATER_TESTS (4 kits with different protection levels to check water resistance)	Time Series	Environmental	Sensor Analysis Framework
DS_TS_093	Citizen Kits Citizen Science Campaign - Vantaa	Time Series	Environmental	Citizen Kits
DS_TS_094	Citizen Kits Citizen Science Campaign - Bottrop	Time Series	Environmental	Citizen Kits
DS_TS_095	Citizen Kits Citizen Science Campaign - Dublin	Time Series	Environmental	Citizen Kits
DS_TS_096	Citizen Kits Citizen Science Campaign - Bologna	Time Series	Environmental	Citizen Kits
DS_TS_097	Citizen Kits Citizen Science Campaign - Surrey	Time Series	Environmental	Citizen Kits
DS_TS_098	Citizen Kits Citizen Science Campaign - Hasselt	Time Series	Environmental	Citizen Kits
DS_MV_013	Matsim Simulation Aggregated Results Hasselt	Multiple Values	User Activity	Others
DS_MV_014	Matsim Simulation Aggregated Results Bologna	Multiple Values	User Opinions	Others
DS_MV_015	Matsim Simulation Aggregated Results Vantaa	Multiple Values	User Activity	Others
DS_MV_016	Behavioural Simulator Matsim Input Data	Multiple Values	User Activity	Others
DS_TS_099	Living Lab Station data Jun-Oct 2019, Sutherland Memorial Park, Guildford	Time Series	Environmental	Others
DS_TS_100	All site data for Citizen Science (Guildford case study)	Time Series	Environmental	Others
DS_TS_101	Air quality Modelling data for Guildford City	Time Series	Environmental	Others
DS_TS_102	Field campaign data for evaluating green infrastructure in Guildford	Time Series	Environmental	Others
DS_TS_103	Living Lab Station data Feb-Sept 2019, Stoke Park, Guildford	Time Series	Environmental	Others
DS_MV_017	Repository for data used in STOTEN article "Impacts of town characteristics on the changing urban climate in Vantaa"	Multiple Values	Environmental	Others

Software and Hardware

For archiving purposes all the software developed for the project as well as the hardware design files and related documentation previously-stored on multiple Github repositories has been archived on the Zenodo platform to comply with the project FAIR data principles. Most of the datasets are part of the Low-Cost and High-End sensor tools developed within T3.1. The full DOI, Publication date, Title, Authors, Description, Keywords, Access right and Licensing is available in the datasets metadata on the linked OpenAIRE repositories.

Title	Description	Datasets Type	OpenAIRE Compatible Repository
Smart Citizen Web Front End	Sensors Web Platform Front-end	Software Open Access	http://doi.org/10.5281/zenodo.2566531
Smart Citizen Sensor Enclosures	Sensors Enclosures Design Files	Hardware Open Access	http://doi.org/10.5281/zenodo.3822376
Smart Citizen Sensor Analysis Framework	Python Framework for Sensor Analysis	Software Open Access	http://doi.org/10.5281/zenodo.2566518
Smart Citizen PM Board	Sensors - Station PM Component	Hardware Open Access	http://doi.org/10.5281/zenodo.2566524
Smart Citizen Platform	Sensors Platform Backend	Software Open Access	http://doi.org/10.5281/zenodo.2566516
Smart Citizen Hardware Platform documentation	Sensors Technical Documentation	Software documentation Open Access	http://doi.org/10.5281/zenodo.2555030
Smart Citizen Hardware and Firmware (2.0)	Sensors Hardware Design and Embedded Software	Hardware Open Access	http://doi.org/10.5281/zenodo.3403063
Smart Citizen Gases Pro Board	Sensors - Station Gases Pro Board	Hardware Open Access	http://doi.org/10.5281/zenodo.2566522
Sense Your Data: Sensor Toolbox	R Tools for Sensors Data Processing	Software Open Access	http://doi.org/10.5281/zenodo.3618523
Air Pollution Quiz	Air Pollution Quiz Web Application	Software Open Access	http://doi.org/10.5281/zenodo.2555939

Personal data

The following table includes most of the datasets used internally for project development purposes that potentially contain Personal Data. The resulting datasets will not be made public.

Data Set	Description	Trace Code	Type	Source
DS_PD_001	Users transport data collected over GPS	Personal Data	User Activity	GPS Tracker
DS_PD_002	Users surveys during WP2 activities	Personal Data	User Opinions	Survey
DS_PD_003	VLL visitors survey	Personal Data	User Opinions	Survey
DS_PD_004	Participant lists: 2 Citizen science workshops in Dublin	Personal Data	Other	Others
DS_PD_005	Dublin Living Lab - a new LBW co-creation activities (1)	Personal Data	User Opinions	Interviews
DS_PD_006	Dublin Living Lab - a new LBW co-creation activities (2)	Personal Data	User Opinions	Survey
DS_PD_007	Socio-economic impact assessment "before" survey Bottrop	Personal Data	User Opinions	Survey
DS_PD_008	Socio-economic impact assessment "after" survey Bottrop	Personal Data	User Opinions	Survey
DS_PD_009	Users transport data collected over GPS in Bottrop	Personal Data	User Activity	GPS Tracker

DS_PD_010	GPS_based_Activity-travel diary data (Hasselt citizens)	Personal Data	User Activity	Survey
DS_PD_011	GPS_based_Activity-travel diary data (Bologna citizens)	Personal Data	User Activity	Survey
DS_PD_012	GPS_based_Activity-travel diary data (Dublin citizens)	Personal Data	User Activity	Survey
DS_PD_013	GPS_based_Activity-travel diary data (Bottrop citizens)	Personal Data	User Activity	Survey
DS_PD_014	GPS_based_Activity-travel diary data (Guildford citizens)	Personal Data	User Activity	Survey
DS_PD_015	Behavioural Intervention Study Questionnaires from Hasselt	Personal Data	User Opinions	Survey
DS_PD_016	Behavioural Intervention study Questionnaires from Bologna	Personal Data	User Opinions	Survey
DS_PD_017	Behavioural Intervention study Questionnaires from Dublin	Personal Data	User Opinions	Survey
DS_PD_018	Behavioural Intervention study Questionnaires from Bottrop	Personal Data	User Opinions	Survey
DS_PD_019	Behavioural Intervention study Questionnaires from Guildford	Personal Data	User Opinions	Survey
DS_MV_002	Social Impact assessment study - Bottrop	Personal Data	User Opinions	Survey
DS_MV_003	Social Impact assessment study - Dublin	Personal Data	User Opinions	Survey

DS_MV_004	Social Impact assessment study - Giulford	Personal Data	User Opinions	Survey
DS_MV_005	Social Impact assessment study - Vantaa	Personal Data	User Opinions	Survey
DS_MV_006	Social Impact assessment study - Bologna	Personal Data	User Opinions	Survey
DS_PD_020	Participant list: Final project event in Dublin	Personal Data	Other	Other
DS_PD_021	iSCAPE Newsletter subscribers (on Mailchimp)	Personal Data	Other	Others
DS_PD_022	Participant registrations for mid-term event Bologna (on Eventbrite)	Personal Data	Other	Participants Registration Platforms
DS_PD_023	Participant registrations for EU Green week 2018 event in Bologna (on Eventbrite)	Personal Data	Other	Participants Registration Platforms
DS_PD_024	Participant registrations for EU Green week 2018 event in Bruxelles (on Eventbrite)	Personal Data	Other	Participants Registration Platforms
DS_PD_025	Participant list: 2 citizens workshops in Bologna	Personal Data	Other	Others
DS_PD_026	Participant list: co-creation workshop in Bologna	Personal Data	Other	Others
DS_MV_012	Data from co-creation and citizen science workshops delivered by the LLs	Multiple Values	Other	Others

