Grant Agreement: 945391

URBANOME

Urban Observatory for Multi-participatory Enhancement of Health and Wellbeing

URBANOME
Urban health, wellbeing, liveability

WP10 Project and Consortium management

D10.3: Data Management Plan – second version

Partner responsible: AUTH

Authors: AUTH, USTUTT, AMBIT, CIEMAT, RGU, VIL, UPCOM, JSI, ENoLL, ADDMA, AU, IUSS, ENVE.X, INSERM, CdM, ISCIII

Date: February 2023

Table of contents

List of Acronyms

accdb	MS Access database format
AUTH	Aristotle University of Thessaloniki
CERN	Conseil Européen pour la Recherche Nucléaire
COHb	Carboxyhemoglobin
CSV	Comma separated Value
DMP	Data Management Plan
DoA	Description of the Action
DOI	Digital Object Identifier
ECG	Electrocardiography
FAIR	Findability, Accessibility, Interoperability, and Reusability
FeNO	Exhaled Nitric Oxide Fraction
FEV1	Forced Expiratory Volume in the 1st second
FEF 25-75	Forced Expiratory Flow 25–75%
FTP	File Transfer Protocol
FVC	Forced Vital Capacity
GA	Grant Agreement
GDPR	General Data Protection Regulation
GIS	Geographic Information System
H2020	Horizon 2020
HTTP	HyperText Transfer Protocol
MD5	Message Digest Method 5
MRIO	Multi-Regional Input Output modelling
NCD	Non-Communicable Diseases
netCDF	Network Common Data Form
OA	Open Access
OSA	Obstructive Sleep Apnea
PDF	Adobe Portable Document Format
PSG	Polysomnographic
RMD	R Markdown file
RPM	Revolutions Per Minute
SDG	Sustainable Development Goals
SSL	Secure Sockets Layer
TB	Terabyte
TTL	Time-To-Live
ULL	Urban Living Lab
UN	United Nations
WP	Work Package
XLS	MS Excel format

Index of Figures

1. Executive summary

The Data Management Plan (DMP) is a continuously updated document that describes data generated in the URBANOME project, its type, format and structure, the arrangements for its storage and security, and its potential for being used by others outside of the URBANOME Consortium. The structure of this DMP is based on the EC's *Guidelines on Data Management in Horizon 2020 (EC, 2016)* following the FAIR principles.

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

The URBANOME Data Management plan is a document that is submitted to the EU as project deliverables in the frame of WP10 "Project and Consortium management". It is important to note however that the document will evolve and further develop during the project's life cycle. It can be identified by a version number and a date. Updated versions will be uploaded by Aristotle University of Thessaloniki (AUTH), which is the primary responsible for data management. URBANOME partners can forward questions and suggestions, as to (additions to) the contents and use of the data management plan to AUTH, and will be informed when a new version will be uploaded in the URBANOME web site (http://URBANOME.eu).

This document is the second version of the DMP (deliverable D10.3), delivered in Month 24 of the project. It includes updated information about the datasets produced by the project according to the FAIR principles to make the data findable, accessible, interoperable and reusable.

A final version of the DMP will be released (deliverable D10.5) at the end of the project on February 2025 (Month 48), in order to finally update it in accordance with the final data generated from the studies and potential changes to the initially expected data production and use and to reflect the evolving needs of URBANOME.

2. Research data types in URBANOME

For this second release of DMP, the data types that are going and/or will be produced during the project are focused on the Description of the Action (DoA) and on the results obtained in the first 24 months of the project.

According to the above, the table below reports the list of types of research data that currently URBANOME is expected to generate. This list may be adapted with the addition or removal of datasets in the next and final versions of the DMP to take into consideration the project developments. A detailed description of each dataset is given in the following sections of this document.

#	Dataset	Lead partner	Related WP(s)
1	Pre-existing experiences dataset with Urban Living Labs	ENoLL	WP2
2	Intervention catalogue	CIEMAT	WP2
3	Emission inventory for the URBANOME ULLs	USTUTT	WP4
4	Atmospheric dispersion modelling results for the ULLs	AUTH	WP4
5	Noise levels in the ULLs	AUTH	WP4
6	Indoor air quality levels in the ULLs	JSI	WP4
7	Multi-sensor data for personal exposure monitoring in the ULLs	ENVE.X	WP4
8	Exposure modelling results in the ULLs	AUTH	WP4
9	Physical Health Assessment	INSERM	WP5
10	Mental Wellbeing Assessment	AUTH	WP5
11	Sleep quality	AUTH	WP5
12	Medical examinations	AUTH/INSERM	WP5
13	Cardiorespiratory clinical assessment	AUTH/INSERM	WP5
14	Integrated assessment of personal benefits	ISCIII	WP6
15	Integrated health and wellbeing assessment	USTUTT	WP6
16	Environmental Footprint Assessment	CIEMAT	WP6
17	Policy cost-benefit and cost-effectiveness assessment	CIEMAT	WP6
18	Interviews and surveys results on policy learning in the ULLs	AU	WP7

19	Stakeholders Database	CdM	WP9
20	Scientific publications	AUTH	WP9
21	Dissemination and Communication Activities Monitoring Database	VIL	WP9
22	Database of Standards applicable to URBANOME research and interventions	VIL	WP9

In addition, specific datasets may be associated to scientific publications (i.e., underlying data), public project reports and other raw data or curated data not directly attributable to a publication.

Research data linked to exploitable results will not put into the open domain if they compromise its commercialization prospects or have inadequate protection, which is a H2020 obligation. The rest of research data have been deposited in an Open Access repository as described in this document.

3. Methodology

The methodological approach that has been used for the compilation of D10.3 follows the updated version of the "Guidelines on Data Management in Horizon 2020" released by the European Commission Directorate - General for Research & Innovation. Taking into account the proposed methodology, this second version of the URBANOME DMP addresses and updates the following points below on a dataset-by-dataset basis:

- Data set reference and name
- Data set description
- Standards and metadata
- Data sharing
- Archiving and preservation (including storage and backup).

The DMP covers the complete research data life cycle. It describes the types of research data that has been generated or collected during the project, the standards used, how the research data will be preserved and what parts of the datasets will be shared for verification or reuse (Figure 1).

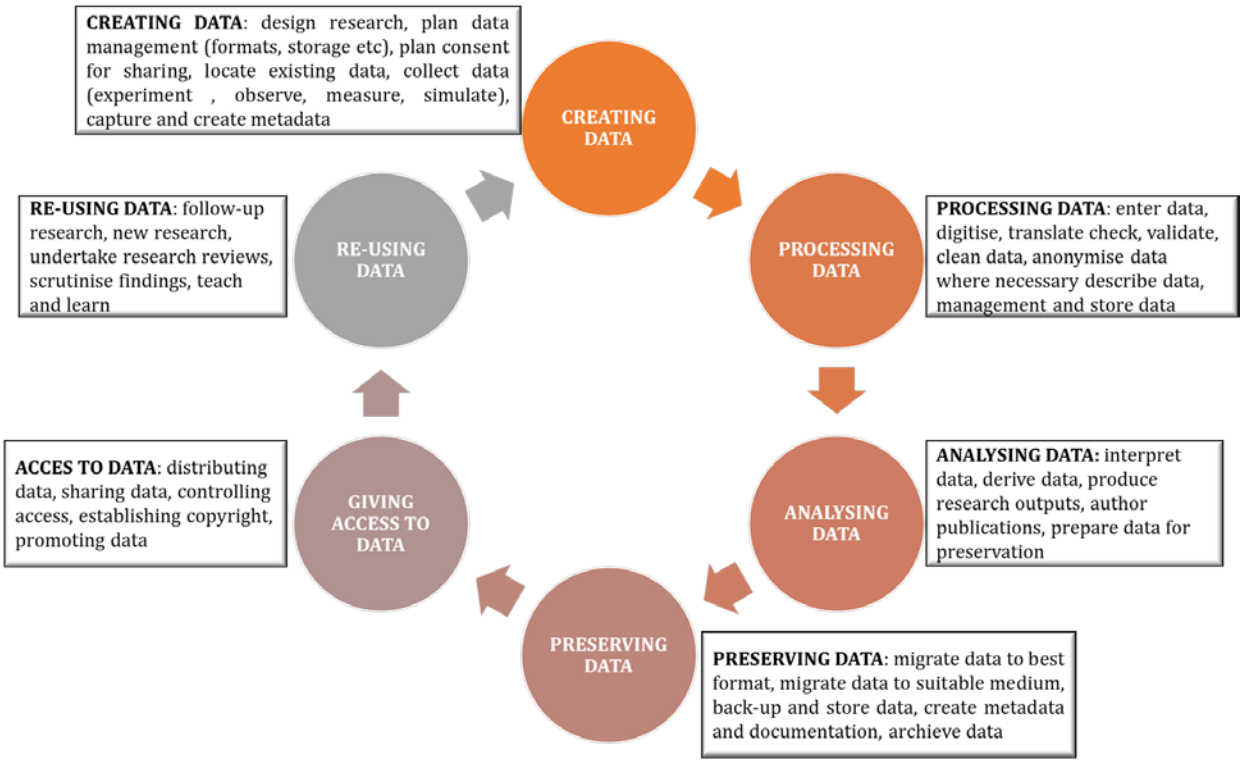

Figure 1 : Research data life cycle (adapted from UK data archive
Source: https://www.ukdataservice.ac.uk/manage-data/lifecycle

Data scientists from each URBANOME partner formed a committee that is responsible for the proper data management throughout the aforementioned research data lifecycle. More specifically, its main duties have been defined as follows: a) data archiving and preservation, b) data security, c) data protection and d) defining a policy setting standard retention period to comply with documentation requirements. These will be analyzed in more detail in the following chapters.

3.1. Data set reference and name

This point is the identifier for the dataset to be produced. The URBANOME dataset identification follows the naming:
<WPn°>_<serial number of dataset>_<dataset_title>.

Example: WP2_1_ *dataset_title*.

3.2. Data set description

In this point the data that are and/or will be generated is described, including references on their origin (in case it is collected), nature and scale and to whom it could be useful. Where applicable, information on the existence (or not) of similar data and the possibilities for integration and reuse, are mentioned.

3.3. Standards and metadata

This point refers to existing suitable standards of the discipline, as well as an outline on how and what metadata are created. Therefore, at this stage, the available data standards (if any) accompany the description of the data that are and/or will be generated, including the description on how the data are organised during the project, mentioning for example naming conventions and version control.

3.4. Data sharing

This point describes how data are and/or will be shared, including access procedures, embargo periods (if any) and defines whether access will be widely open or restricted to specific groups. Identification of the repository where data are going to be stored, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.). In case the dataset cannot be shared, the reasons for this is mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).

This section also contains information on the license under which the data can be used by third parties. In compliance with the commitments of the consortium, this license should allow them to use the data free of charge provided they make any derivative work is available under the same license and mentions the data origin. The exact terms of the one or more licenses for data sharing will be defined in the final version of the DMP.

Within URBANOME all personal data used in the project will be protected, while the informed consent of all participants will be a prerequisite for their participation in the project activities. Finally, it is expected that the URBANOME project results in a number of publications in scientific, peer-reviewed journals. Project partners are committed to collaborate with each other and jointly prepare publications relevant to the URBANOME project. Scientific journals that provide Open Access (OA) to all their publications will be strongly preferred, as it is required by the European Commission.

3.5. Archiving and preservation (including storage and backup)

According to the GDPR Principle "e", the consortium defines the maximum period for data archiving and preservation. In this point the procedures that will be put in place for long-term preservation of the data is described, along with the indication of how long the data should be preserved. This point emphasizes in the long-term preservation and curation of data, beyond the lifetime of the URBANOME project. According to the decision taken among members of the URBANOME consortium, this period is estimated at least five (5) years beyond the project's end (see Section 4.1). The committee may also decide whether and how the retention period will be extended for specific purposes. Apart from the retention period, the committee is also responsible of periodically reviewing the data, securing their anonymization and proper protection and also erasing it when the retention period expires.

In particular, storage and backup activities have been carefully designed so as to address the following questions:

1. What features should the URBANOME storage system have in order to be GDPR-compliant?
2. How is GDPR-compliance expected to affect the performance of the various components of the storage system?
3. Which are the technical challenges need to be addressed in order to achieve strict GDPR-compliance without performance degradation?

This second version of the DMP has confirmed a series of GDPR articles that have been preliminary identified in the initial version of the DMP and that describe how a storage system addresses the aforementioned questions. Each article is associated with a specific key requirement, which should be implemented by a storage feature. However, there are two articles (5.2 & 5.13) that involve the entire URBANOME storage system.

Table 1 Description of the GDPR articles which have a significant impact on the design, interfacing and performance of the URBANOME storage system. The article key requirements and their mapping to the URBANOME storage system features are also reported. Adjusted version by Aashaka Shah et al (2019).

Article No.	***Article Title***	***Key Requirement***	***URBANOME Storage Feature***
5.1	Purpose limitation	Data must be collected and used for specific purposes	Metadata indexing
5.1	Storage limitation	Data should not be stored beyond its purpose	Timely deletion
5.2	Accountability	Controller must be able to demonstrate compliance	All
13	Data collection conditions	Get user consent on how their data would be managed	All
15	Right of access by users	Provide users a timely access to all their data	Metadata indexing
17	Right to be forgotten	Find and delete groups of data	Timely deletion
20	Right to data portability	Transfer data to other controllers upon request	Metadata indexing
21	Right to object	Data should not be used for any objected reasons	Metadata indexing
25	Protection by design and by default	Safeguard and restrict data access	Access control & Encryption

32	Data security	Implement appropriate data security measures	Access control & Encryption
33 & 34	Data breaches notification	Share insights and audit trails from concerned systems	Monitoring
46	Transfers subject to safeguards	Control where the data resides	Manage data location

The current DMP version confirmed the following key features that the URBANOME storage system must support in order to be GDPR-compliant:

1) ***Timely deletion:*** Personal data cannot be retained for an indefinite time period. The URBANOME storage system should define a personal data retention period and develop the appropriate mechanism which associate time-to-live (TTL) counters for personal data. These TTL counters will automatically erase personal data from all internal URBANOME subsystems in a timely manner. The counters could be either a static time or a policy criterion which would be objectively evaluated by the URBANOME data management committee.
2) ***Monitoring & logging:*** The URBANOME storage system should be able to demonstrate its GDPR compliance by developing an audit trail of both its internal actions and their interaction with the other components of the URBANOME ecosystem. This will facilitate the monitoring and logging of every operation either inside the data path (read/write actions) or inside the control path (metadata alterations, access control).
3) ***Access control:*** The URBANOME storage system should support data access roles in order to support fine-grained and dynamic access control to only permitted entities, for established purposes and for predefined time duration.
4) ***Encryption:*** The URBANOME storage system needs to develop encryption mechanisms to the personal data it manages. Data encryption needs to be implemented beyond data pseudonymization which will take place at a previous step.

4. URBANOME data repository

To share data between URBANOME partners and make them publicly available, two secure platforms have been established: the URBANOME global open-data repository and the Zenodo repository. Regarding the global open-data repository, it is hosted in a secure and maintained FTP server at the Aristotle University of Thessaloniki (ftp://155.207.30.156) and it is openly accessible to the public with a login requirement. With regard to the Zenodo repository, a project page (community) has been setup for easy upload and download of project datasets https://zenodo.org/communities/urbanome/?page=1&size=20. Both the repositories will include the totality of the project generated data that will be made finally public.

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

The global open-data repository has been built using the infrastructure of Aristotle University of Thessaloniki (AUTH). It offers management and data retrieval functionalities and supports data exchange

through an FTP server.

Zenodo is a research output publication and archival service built and developed by researchers from the CERN IT department, to ensure that everyone can join in Open Science. Based on the open source Invenio digital library framework, Zenodo was created through the OpenAIRE project and is hosted at CERN. Zenodo is open, free, searchable and structured with flexible licensing allowing for storing all types of data: datasets, images, presentations, publications and software. In addition, Zenodo allows researchers to deposit both publications and data, while providing tools to link them.

The use of Zenodo ensures data management procedures are unified across the project. Moreover, as a European Commission supported 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.

A DOI is assigned to datasets for effective and persistent citation when it is uploaded to the Zenodo repository. This DOI can be used in any relevant publications to direct readers to the underlying dataset

Search keywords will be provided when the dataset is uploaded to Zenodo which will optimise possibilities for re-use.

All open data files will be stored along with a MD5 checksum of the file content. Regular checks of files against their checksums will be made.

The datasets in Zenodo will be preserved in line with the European Commission Data Deposit Policy. The data will be preserved indefinitely (minimum of 5 years) and there are currently no costs for archiving data in this repository.

If datasets are updated, the partner that possesses the data has the responsibility to manage the different versions and to make sure that the latest version is available in the case of publicly available data. Quality control of the data is the responsibility of the relevant responsible partner generating the data.

4.1. Data archiving after project's lifetime

Most of the data generated during the project needs to be persisted and archived even after its completion. This is to enable consortium members, stakeholders and third-party researchers to retrieve and use the data in the future. The repositories cover this need by providing an organized structure for data storage. Both the URBANOME global open-data repository and the Zenodo community will remain operational for at least 5 years after the project's end.

4.2. Public availability

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

4.3. Architecture

The URBANOME global open-data repository is used only for storing and archiving data, as well as for making them available to the public. It will not be used as an online database to retrieve data for processing or for storing generated data directly. Consequently, there is no need for it to contain any database management systems. To simplify the process and make it easily accessible to every user, a very simple architecture has been followed (Figure 2).

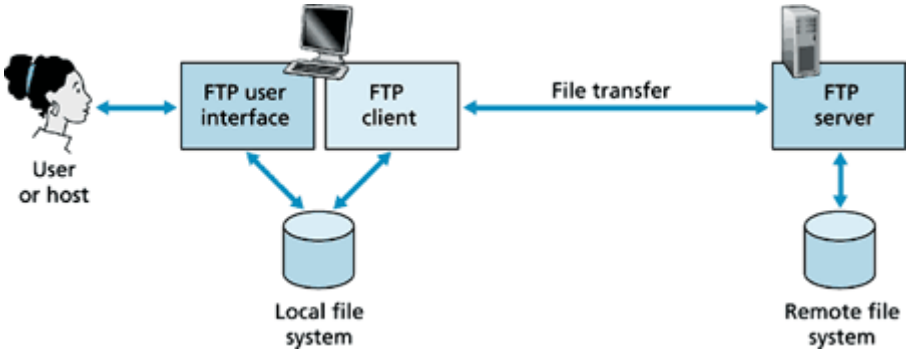

Figure 2 : URBANOME global open-data repository architecture

Users can access the project data via a standard FTP client or via a web browser interface. Since no UAC (User Access Control) method has been implemented, users that access the FTP server (ftp://155.207.30.156) are granted with Read rights only.

5. DMP components

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

5.1. Data summary

Points to be addressed:

Provide a summary of the data, addressing the following issues

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

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

The purpose is to produce different datasets across Europe to quantitatively and qualitatively assess the impact of "interventions" at city level to improve air quality, human health and wellbeing eventually leading to the design and implementation of appropriate environmental and health strategies in European cities. Access to use these data is needed to address the different URBANOME objectives as specified in the grant agreement and summarized below.

- ✓ Develop an integrated methodological framework and a big data computational ecosystem and decision support system accounting for the physical, mental, socio-economic, operational, and structural characteristics of a city which affect human health.
- ✓ Provide an accurate spatial and temporal resolution of environmental exposures in a personalized manner and identify the patterns of socio-spatial environmental inequalities finally leading to social cohesion improvement.
- ✓ Investigate the impact of environmental stressors on mental, cognitive, social, and physical wellbeing particularly on vulnerable population groups, including hard-to reach citizens, and considering gender differences.
- ✓ Raise public awareness regarding urban health and wellbeing and engage urban citizens through Urban Living Labs (ULLs), including big data collection and citizen science-led participatory governance.
- ✓ Alleviate environmental health injustice in cities.
- ✓ Build solid partnerships with other relevant parties in urban health, including policy makers, experts, local authorities, business and local communities to better connect research to practical needs and user demands.
- ✓ Develop and promote effective, innovative, and inclusive governance and planning adapted to local urban contexts which foster long-term health and wellbeing in cities.
- ✓ Provide prognostic markers regarding the effect of environmental stressors in the main non-communicable diseases (NCDs).
- ✓ Establish an evaluation framework to assess cost-effectiveness as well as barriers and facilitators to the implementation of identified policy actions aiming to support urban development and city planning towards improved urban health and wellbeing as well as social cohesion.
- ✓ Creation of new business opportunities (in mobility, entertainment, culture, etc.) aimed at enhancing urban health and wellbeing through co-creation and co-maintenance of sustainable, egalitarian, and healthy cities.

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

5.1.2.1 Types of the data

At the time of drafting this second version of the Data Management Plan the following data types have been already identified to be generated during the project life. The following list may be updated in the next and

final version of the DMP in accordance with the data generated from the studies and potential changes to the currently expected data production and use and to reflect the evolving needs of URBANOME.

1. Pre-existing experiences dataset with Urban Living Labs
2. Intervention catalogue
3. Emission inventory for the URBANOME ULLs
4. Atmospheric dispersion modelling results for the ULLs
5. Noise levels in the ULLs
6. Indoor air quality levels in the ULLs
7. Multi-sensor data for personal exposure monitoring in the ULLs
8. Exposure modelling results in the ULLs
9. Physical health assessment
10. Mental Wellbeing assessment
11. Sleep quality
12. Medical examinations
13. Cardiorespiratory clinical assessment
14. Integrated assessment of personal benefits
15. Integrated health and wellbeing assessment
16. Environmental Footprint Assessment
17. Policy cost-benefit and cost-effectiveness assessment
18. Interviews and surveys results on policy learning in the ULLs
19. Stakeholders Database
20. Scientific publications
21. Dissemination and Communication Activities Monitoring Database
22. Database of Standards applicable to URBANOME research and interventions

5.1.2.1 Formats of the data:

Several formats have been already identified based on the data type. Hereinafter are summarized some of the most common formats used. More details are provided in Chapter 6 “Datasets of URBANOME”

- MS Excel compatible files including comma separated CSV, txt and xls(x).
- Standard GIS format (e.g., raster, shapefile)
- Database format (e.g., accdb)
- MS Word (doc/docx) and/or Adobe Portable Document Format (PDF) format
- R data format (e.g., RDM/RMS file, R file)
- Image: JPEG, PNG, TIFF
- European Data Format (EDF files) for storing raw polysomnographic (PSG) files for sleep related data

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

Yes, it is encouraged to make existing data available for research. These may include for example environmental and health data collected in the frame of the other past and/or ongoing research projects.

5.1.4. What is the origin of the data?

Dataset are mainly generated by the URBANOME project and to lesser extent are gathered from previous research projects in which, for example, emission data have been collected across Europe.

5.1.5. What is the expected size of the data?

The expected size depends on the extent and the nature of the data that is made available. Currently, the estimated size of the URBANOME Global Data Repository is 1 TB.

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

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

5.2. FAIR data

Points to be addressed:

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

5.2.1. Making data findable, including provisions for metadata

Points to be addressed:

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

5.2.1.1 Metadata provision

URBANOME data will always include metadata. To ensure the project data is made available using the same standards towards consistency and usability all the datasets provide at least the following metadata according to the DataCite Metadata Schema[1].

- Digital Object Identifier (DOI)
- Publication date
- Title
- Description

5.2.1.2 Standards for metadata creation

See point above.

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

Spatial data follows the INSPIRE metadata elements for spatial data sets and services[2].

5.2.1.3 Naming conventions used

The URBANOME dataset identification follows the naming:
<WPn°>_<serial number of dataset>_<dataset_title>.
Example: WP2_1_ dataset_title.

5.2.2. Making data openly accessible

Points to be addressed:

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

5.2.2.1 Data are accessible via URBANOME repository to URBANOME consortium partners

As detailed in Chapter 6 several datasets will be collected by various partners as part of their project activities. The URBANOME Consortium considers data sharing is necessary as it will enable the Parties to understand how environmental health determinants, the spatial distribution of these in the cities, and the social distribution of their impact among different population groups, accounting for different life styles and behaviours eventually affect human health and wellbeing. This will allow us to draw conclusions regarding the determinants of urban health and wellbeing that will be translated into evidence–based policy recommendations considering socio-economic and environmental factors leading to urban health inequalities.

To this end an International Data Sharing Agreement is currently being drafted among the dataset suppliers and the final dataset collector (i.e., Aristotle University of Thessaloniki). With this agreement the URBANOME consortium commit itself to made most of the data generated in URBANOME openly available. Some exceptions are relevant to the personal data collected in the frame of the exposure and health campaigns which cannot be made available to the public due to privacy and confidentiality issues. Full details are provided in Chapter 6 for each dataset identified.

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

- Is a platform that facilitates sharing of data, intermediate results, and results
- Is needed to enable the analysis of the impact of “interventions” at city level to improve air quality, human health, and well-being policies, but also of accessory external exposure data and health data to meet the goals of URBANOME.
- Enables data users to work with selected quality-controlled data sets and versions approved by the Data Owners/Data Providers.
- Aims to reach the highest level of GDPR compliancy, amongst others by:

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

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

5.2.3. Making data interoperable

Points to be addressed:

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

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

5.2.4. Allocation of resources

5.2.4.1 Estimation of costs

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

5.2.4.2 Responsibilities

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

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

AUTH, as WP10 leader, will ensure dataset integrity and compatibility for its use during the project lifetime and beyond by different users.

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

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

If datasets are updated, the partner that possesses the data has the responsibility to manage the different versions and to make sure that the latest version is available in the case of publicly available data. WP10 will provide naming and version conventions.

Finally, all partners must consult the concerned partner(s) before publishing data in the open domain that can be associated to an exploitable result.

5.2.4.3 Responsibilities for data management

AUTH will be the data repository manager. It will be assisted by UPCOM for the implementation and maintenance of the data repository, with these specific responsibilities:

UPCOM is responsible for:

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

UPCOM responsibilities include:

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

AUTH is responsible for:

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

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

5.2.5. Data security

Points to be addressed:

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

With respect to Privacy and Data Protection, the so-called "GDPR" EU-legislation (i.e. the General Data Protection Regulation) imposes several new obligations upon the consortium partners being data processors. Moreover, several new rights are granted to data subjects and significant fines are introduced in case of a data breach.

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

5.2.5.1 Data confidentiality and integrity

Data collected in the frame of the URBANOME exposure and health campaigns will be protected against unauthorised access by means of standard login procedure. Access to data will be granted only to authenticated users. Users can be authenticated by their username and password. Authentication credentials must be exchanged over SSL to ensure they are kept private. Appropriate access levels may be granted by the creation of groups.

5.2.6. Ethical aspects

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

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

The GDPR is an expansive set of regulations which cover the entire lifecycle of personal data. It consists of 99 articles that describe its legal requirements and 173 recitals that provide additional context and clarifications to the articles. More specifically, there are four (4) articles that codify the URBANOME participant rights in terms of their personal data and how these are related with the data storage. These articles have been already taken into consideration by the URBANOME community and are highlighted within the current DMP version:

1. ***Access rights by URBANOME participants*** (Article 15): URBANOME participants are allowed to obtain detailed information about their data usage in terms of a) processing purposes, b) any disclosed data recipients, c) storage period duration and d) how their data will be used within automated decision-making approaches.
2. ***Right to be forgotten*** (Article 17): URBANOME participants are constantly granted with the right to require the deletion of their personal data without any undue delay. Therefore, the URBANOME repository has been designed so as to be able to delete the requested data in a timely manner (e.g. within 72 hours). This obligation is also valid for any backup files or replicas.
3. ***Data portability right*** (Article 20): URBANOME participants have the right to ask to obtain in a commonly used format all their personal information being stored by the repository. They can also have the right to request the direct transmission/dissemination of their data to another health expert

or a third party. So, the URBANOME repository should be capable to access and transmit all their personal data in a timely manner.

4. Objection rights ***by URBANOME participants*** (Article 21): URBANOME participants have the right to decline at any time the use of their personal data for any research, publication, historical archiving or profiling purposes. So, the URBANOME DMP and repository system must be able to know both white/black-listed purposed associated with the URBANOME datasets at all times and being able to dynamically control their access.

6. Description of the URBANOME Datasets

Datasets in the URBANOME global open-data repository will be stored according to the Work-Package from where they have been generated.

6.1. Datasets in WP2 - Urban Living Labs for participatory planning and pilot interventions

6.1.1. Pre-existing experiences dataset with Urban Living Labs

Data set reference and name	WP2_1_Pre-existing_experiences_dataset_with_Urban_Living_Labs
Data set description	This dataset includes previously gathered information and research on Urban Living Labs compiled by ENoLL. The dataset will be used for planning, training and handbook purposes.
Standards and metadata	The dataset will be stored in docx and pdf format.
Data sharing	The data used for the training purpose will be shared internally with all project partners. Excerpts of the data used for training in addition to other data on Urban Living Labs components and management for the "Good Practice guidelines for setting up and managing Urban Living Labs" will be shared publicly without restrictions.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.1.2. Intervention catalogue

Data set reference and name	WP2_2_Intervention_catalogue
Data set description	Interventions related to individual lifestyle and on the urban scale will be collected in each city and compiled in a catalogue which will be shared with all project partners. The catalogue may be updated during the project course due to account for any potential modification which may arise during the project activities.
Standards and metadata	The catalogue will be stored in docx and pdf format
Data sharing	The catalogue will be made freely available to all project participants, partners and stakeholders in each city.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.2. Datasets in WP3 - Development of Big Data ecosystem

No datasets are expected to be generated in WP3.

6.3. Datasets in WP4 - Exposure analysis focusing on patterns of socio-spatial environmental inequality

6.3.1. Emission inventory for the URBANOME ULLs

Data set reference and name	WP4_1_Emission_inventory_for_the_URBANOME_ULLs
Data set description	The dataset will include model-ready gridded sector-specific emissions at a high spatial resolution (at least 1 km x 1 km) for six air pollutants (NH_3, NMVOC, NO_X, PM_{10}, $PM_{2.5}$, SO_2). The sites targeted will cover the pilot cities addressed in URBANOME (i.e., Aarhus, Aberdeen, Athens, Thessaloniki, Milan, Madrid, Ljubljana, Montpellier and Stuttgart). These emission inventories will be used as input for the atmospheric pollution models applied in WP4.
Standards and metadata	For each city the dataset will be available in csv format as it is best suited for project internal use (especially as input for the atmospheric pollution models applied in WP4). The datasets are georeferenced and ready to be used in a GIS environment following the INSPIRE standards.
Data sharing	The dataset with emission inventory will be publicly freely available with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.3.2. Atmospheric dispersion modelling results for the ULLs

Data set reference and name	WP4_2_Atmospheric_dispersion_modelling_results_for_the_ULLs
Data set description	This dataset is operationally produced by the atmospheric dispersion models applied in URBANOME. More specifically they will be derived from the application of Eulerian models in the 9 ULLs addressed in URBANOME i.e. Aarhus, Aberdeen, Athens, Ljubljana, Madrid, Milan, Montpellier, Stuttgart and Thessaloniki. The dataset will be used within WP4 to support the data fusion approach.
Standards and metadata	The dataset will be available in the standard GIS raster grid format

	following the INSPIRE standards.
Data sharing	The data will be in the form of GIS grid maps shared with interested users by means of individual requests made to the Project Coordinator with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics in doing so.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.3.3. Noise levels in the ULLs

Data set reference and name	WP4_3_Noise_levels_in_the_ULLs
Data set description	Available data sets of noise level modelled and /or measured in those cities with network stations managed by local authorities will be collected. Additionally, data obtained in those cities whose ULLs propose activities that include in situ measurements of noise levels at sites of interest will be included.
Standards and metadata	For each city the dataset will be available either in standard GIS format (e.g. raster or shapefile) or in MS-Excel xls and xlsx as well as csv formats to maximise the potential for use with a variety of analytical software. Requirements of the INSPIRE standards are considered.
Data sharing	Datasets with noise levels registered at monitoring stations managed by local authorities will be publicly freely available. Datasets obtained from ULL activities (measured or modelled) will be publicly freely available with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.3.4. Indoor air quality levels in the ULLs

Data set reference and name	WP4_4_Indoor_air_quality_levels_in_the_ULLs
Data set description	Data obtained in those cities whose ULLs propose activities that include in situ measurements of indoor air quality levels at sites of interest will be included.
Standards and metadata	Data sets obtained from ULLs activities will be available in MS-Excel xls and xlsx as well as csv formats.
Data sharing	Depending on the location of the ULLs activities, such as public

	schools or private houses, there could some restrictions regarding data protection or ethics, in relation with the diffusion of the data sets obtained. In this case, any information that might permit identification of the individuals will be removed or obfuscated. Only fully anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects and households will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses. Data will be stored in the URBANOME open repository as well as in the Zenodo repository in different subfolders named according to the city where the data were collected. A coding mechanism for naming the single files has been developed.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.3.5. Multi-sensor data for personal exposure monitoring in the ULLs

Data set reference and name	WP4_5_Multi_sensor_data_for_personal_exposure_monitoring_in_ the_ULLs
Data set description	This dataset will include data collected by personal sensors worn by each participant in the cities performing the sensor campaigns. Data includes individual levels of exposure to pollutants, noise, physical activity and geo-referenced position. All the data measured by these sensors will be stored in the URBANOME open repository as well as in the ZENODO repository in different subfolders named according to the city where the data were collected. A coding mechanism for naming the single files has been developed.
Standards and metadata	The data will be stored in MS-Excel xls and xlsx formats to maximise the potential for use with a variety of analytical software. Storing data as RMD files will also be considered to preserve more meta data (e.g. data type of variable) about datasets.
Data sharing	The multi-sensor data for personal exposure monitoring will be freely available to anyone as Open Data in full anonymous format. Only anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses.

	Fully anonymized data will be freely downloaded and used with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of individual participants will be removed or obfuscated.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end

6.3.6. Exposure modelling results in the ULLs

Data set reference and name	WP4_6_Exposure_modelling_results_in_the_ULLs
Data set description	The dataset will contain time dynamic personal exposure levels derived from the application of ABM model informed by personal sensors data and information from surveys. Personal exposure levels will be calculated as inhaled adjusted exposure to air pollutants by assigning pollutant concentrations to a person depending on different coordinates, different activities, the level of intensity and the corresponding inhalation rate.
Standards and metadata	The data will be available either in csv or xls to maximise potential for use with a variety of analytical software. Storing data as RMD files will also be considered to preserve more meta data (e.g. data type of variable) about datasets
Data sharing	The personal exposure profiles will be made available as Open Data in a full anonymous format. Only anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses. Fully anonymized data will be freely downloaded and used with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of the individuals will be removed or obfuscated.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years

	after the project end.

6.4. Datasets in WP5 - Assessment of physical and mental health and sleep quality

6.4.1. Physical Health Assessment

Data set reference and name	Data_WP5_1_Physical_Health_Assessment
Data set description	This dataset aims at estimating the physical health status of the volunteers participant in the URBANOME health study. The assessment makes use of the following surveys: 1) Large Analysis and Review of European housing and health Status (LARES) study (Ormandy, 2009) 2) ELSA-EPAR questionnaire on asthma and allergies
Standards and metadata	The data will be stored in MS-Excel xls and xlsx as well as csv formats to maximise the potential for use with a variety of analytical software. Storing data as RMD files will also be considered to preserve more meta data (e.g. data type of variable) about datasets.
Data sharing	The physical condition features will be freely available to anyone as Open Data in a full anonymous format. Only anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects and households will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses. Data will be freely downloaded and used with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of individual participants will be removed or obfuscated.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.4.2. Mental Wellbeing Assessment

Data set reference and name	WP5_2_Mental_Health_Assessment
Data set description	This dataset aims at estimating mental wellbeing, anxiety, depressive or mood disorders even at a preclinical stage. Therefore, it targets the general population focusing on the identification of vulnerable groups. It consists of the following tests: 1) State-Trait Anxiety Inventory (STAI): It is a commonly used measure of trait and state anxiety. It can be used in clinical settings to diagnose anxiety and to distinguish it from depressive syndromes. It also is often used in research as an indicator of caregiver distress. 2) The Perceived Stress Scale" (PSS) (Cohen et al., 1983) which will allow us to estimate the level of stress of the participants included in our study 3) the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) test (Tenant et al., 2007). The WEMWBS test was developed to enable the measuring of mental wellbeing in the general population and the evaluation of projects, programmes and policies which aim to improve mental wellbeing. WEMWBS has 2 scales: the original 14-item scale and the short 7-item scale. The 14-item scale WEMWBS has 5 response categories, summed to provide a single score.
Standards and metadata	The data will be stored in MS-Excel xls and xlsx as well as csv formats to maximise the potential for use with a variety of analytical software. Storing data as RMD files will also be considered to preserve more meta data (e.g. data type of variable) about datasets.
Data sharing	The mental health features will be freely available to anyone as Open Data in a full anonymous format. Only anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects and households will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses. Data will be freely downloaded and used with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might

	permit identification of individual participants will be removed or obfuscated.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.4.3. Sleep quality

Data set reference and name	WP5_3_Sleep_Quality
Data set description	This dataset aims at providing quantifiable indices of sleep quality and how they are related to participants' health. Therefore, it targets the general population aiming to assess sleep macro-architecture, sleep disorders (insomnia, restless leg syndrome, micro-arousals, wakes after sleep onset), sleep-related breathing disorders (hypopnea, apnea, oxygen desaturation). It consists of the following tests: 1) Pittsburgh Sleep Quality Index (PSQI): Self-rated, assesses sleep quality and disturbances over the last month. 2) Epworth Sleepiness Scale (ESS): Self-administered questionnaire with 8 questions. The total ESS score (the sum of 8 item-scores) gives an estimate of a more general characteristic, the person's 'average sleep propensity' or ASP, across a wide range of activities in their daily lives. The reference range of 'normal' ESS scores is zero to 10. Sleep disordered breathing and depression contribute to the total ESS score. 3) Berlin Questionnaire (Sleep Apnea): It reliably identifies middle-aged and older persons in the community who are at high-risk for Obstructive Sleep Apnea (OSA) In Addition, in some ULLs we will also carry out Polysomnographic (PSG) studies during entire sleep (four (4) per participant on two (2) different season intervals)
Standards and metadata	The data will be stored in MS-Excel xls and xlsx as well as csv formats to maximise the potential for use with a variety of analytical software. Raw sleep data from Polysomnographic studies will be stored in EDF format. Sleep macro-architecture features will be stored in csv and xls/xlsx format so as to maximise the potential for use with a variety of analytical software. Storing data as RMD files will also be considered to preserve more meta data (e.g. data type of variable) about datasets.
Data sharing	The sleep quality dataset will be freely available to anyone as Open Data in a full anonymous format. Only anonymized data can be downloaded from the URBANOME repositories for

	subsequent tabulation and statistical analysis. Subjects and households will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses. Data will be freely downloaded and used with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of individual participants will be removed or obfuscated.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.4.4. Medical examinations

Data set reference and name	WP5_4_Medical_examinations
Data set description	This dataset aims at providing some general medical information derived from some basic medical examinations of the population recruited. The datasets include the collection general personal information, as well as the measurement of the Body Mass Index (BMI) and blood pressure.
Standards and metadata	The data will be stored in MS-Excel xls and xlsx as well as csv formats to maximise the potential for use with a variety of analytical software.
Data sharing	The medical examinations dataset will be freely available to anyone as Open Data in a full anonymous format. Only anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects and households will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses. Data will be freely downloaded and used with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might

	permit identification of individual participants will be removed or obfuscated.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.4.5. Cardiorespiratory Clinical Assessment

Data set reference and name	WP5_5_Cardiorespiratory_clinical_assessment
Data set description	This dataset aims at providing quantifiable clinical outcomes of participants cardiorespiratory health. Therefore, it targets the URBANOME participants aiming to identify risk factors that may degrade their cardiorespiratory health. It consists of the following examination and tests: 1) Exhaled nitric oxide level measured through non-invasive breath testing with a portable device. It is an established marker of airway inflammation. It relies on the fact that NO is a gaseous molecule produced in a type of inflammatory response. The exhaled nitric oxide fraction (FeNO) has been associated with both short-term and long-term exposure to air pollution. The sampling will be performed in morning hours with the NIOX device. 2) Spirometric evaluation involving the measurement of lung function (FEV1, FVC, FEF 25-75 as well as vital capacity). The recordings will be performed measured on a monthly basis with a spirometer Spirobank Smart by MIR capable to provide remote assessments of lung function. 3) Quantification of CO carbon monoxide levels after binding with hemoglobin in the lungs and forming carboxyhemoglobin (COHb), which impairs the transport of oxygen. 4) Cardiometabolic assessment will be performed through an incremental maximal cycle ergometer test started with two (2) minutes of seated rest followed by three (3) minutes of unloaded cycling. Then, the workload protocol would be increased with incremental steps of 10, 15 or 20 Watts each minute according to the participant's height, weight and age. They will be instructed to cycle at a speed of 60-70 rpm and would be encouraged to continue cycling until exhaustion. During the test participants will breathe through a facemask connected to a calibrated metabolic cart, while the expired gases pass through a flow metre, oxygen (O2) and carbon dioxide (CO2) analyzers. Both the flow metre and the analyzers would be connected to a computer for estimating breath-by-breath minute

	ventilation (VE), oxygen uptake (VO2), carbon dioxide output (VCO2) and the respiratory exchange ratio calculated as: RER=VCO2/VO2. Peak oxygen consumption (VO2peak) will be calculated as the average value over the last 30 seconds before subjective exhaustion. 5) The heart rate would be continuously measured during the exercise through 2 leads electrocardiography (ECG). R-R intervals and spectral heart rate variability (HRV) features would be extracted through the Lomb-Scargle periodogram analysis as in Chriskos et al., 2018. The HRV analysis assesses the energy expenditure and the interplay among the sympathetic and the parasympathetic branch of the autonomic nervous system
Standards and metadata	Cardiorespiratory features will be stored in csv and xls/xlsx format so as to maximise the potential for use with a variety of analytical software. Storing data as RMD files will also be considered to preserve more metadata (e.g. data type of variable) about datasets.
Data sharing	The cardiorespiratory features dataset will be freely available to anyone as Open Data in a full anonymous format. Only anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects and households will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses. Data will be freely downloaded and used with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. Due to the sensitive nature of these data which implies data protection and ethical concerns, any information that might permit identification of individual participants will be removed or obfuscated.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.5. Datasets in WP6 - Assessment of personalised interventions and proposed policies

6.5.1. Integrated assessment of personal benefits

Data set reference and name	WP6_1_ Integrated_assessment_of_personal_benefits
Data set description	The dataset contains the final outcomes of the integrated assessment of personal benefits of interventions at the personal level of the individuals who participate in the URBANOME ULLs The dataset will include qualitative data from questionnaires and interviews with participants. All data at the level of the individual will be fully anonymized at a local level in each ULL. Only anonymized data can be downloaded from the URBANOME repositories for subsequent tabulation and statistical analysis. Subjects will be assigned a unique identification code that does not include any identifying information. Information such as address and phone number will be kept in a separate, encrypted database and will not be linked to subject data in analyses.
Standards and metadata	The dataset will be made available in MS-Excel xls(x) format to ensure interoperability allowing easy parsing and information exchange. There will be one MS-Excel file for each pilot city.
Data sharing	Both qualitative and quantitative data at individual level will not be shared due to the sensitive nature of these data which implies data protection and ethical concerns. Aggregated data will be made public once the data has been subject to appropriate analysis and with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited.
Archiving and preservation	The aggregated data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.5.2. Integrated health and wellbeing assessment

Data set reference and name	WP6_2_Integrated_health_and_wellbeing_assessment
Data set description	The dataset contains the final outcomes of the integrated health and wellbeing assessment of the selected interventions at urban level in the pilot cities where interventions at urban level will be undertaken. The dataset will include the quantitative assessment of health impacts for several health end-points at the population level as well as on vulnerable population subgroups (e.g. children and

	elderly). The dataset will also include qualitative assessment and indicators regarding perceived livability and well-being, and the effects of the urban level interventions undertaken.
Standards and metadata	The dataset will be made available in MS-Excel xls(x) format to ensure interoperability allowing easy parsing and information exchange. There will be one MS-Excel file for each pilot city.
Data sharing	Both qualitative and quantitative data at individual level will not be shared due to the sensitive nature of these data which implies data protection and ethical concerns. Dataset with results on the integrated health and wellbeing assessment aggregated for population groups will be publicly freely available with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these aggregated data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.5.3. Environmental Footprint Assessment

Data set reference and name	WP6_3_Integrated_health_and_wellbeing_assessment
Data set description	The dataset contains the results of the Environmental Footprint Assessment of the selected interventions in the form of a set of environmental impact results. Additionally, the dataset will also include the input cost vector for the MRIO assessment of the investments made in a selected set of relevant interventions and the obtained results in the form of impacts on employment and value-added generation.
Standards and metadata	The dataset will be made available in MS-Excel xls(x) format to ensure interoperability allowing easy parsing and information exchange. There will be one MS-Excel file for each pilot city.
Data sharing	The dataset with results on environmental footprint assessment will be publicly freely available with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.5.4. Policy cost-benefit and cost-effectiveness assessment

Data set reference and name	WP6_4_Policy_cost-benefit_and_cost-effectiveness_assessment
Data set description	The dataset contains the final outcomes of the assessment of cost-benefit and cost-effectiveness of the policies developed in WP7. These will relate the pilot cities addressed in URBANOME (i.e., Aarhus, Aberdeen, Athens, Thessaloniki, Milan, Madrid, Ljubljana, Montpellier and Stuttgart). The dataset will include both data concerning the tangible and intangible factors which are judged to affect the policies. These will include the monetary values of the assessed health effects and the monetary values of the environmental impacts quantified in Task 6.2. Additionally, the dataset will include the relevant costs associated to the interventions. The data will also include the outcomes of the benefit-transfer modelling in relation to policy options. The dataset will include a cost-benefit analysis for each of the interventions applied in the ULLs, identifying predicted and modelled benefits, costs, risks and uncertainties. Additionally, the dataset will include, to the extent possible, an assessment of the impact of the policy options in relation to the UN SDGs (including qualitative assessment of impact on health, energy, emissions and climate at aggregated level).
Standards and metadata	The dataset will be made available in MS-Excel xls(x) format to ensure interoperability allowing easy parsing and information exchange. There will be one MS-Excel file for each pilot city.
Data sharing	The dataset with results on the integrated health and wellbeing assessment will be publicly freely available with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited. The non-sensitive nature of these data means that there are no concerns regarding data protection or ethics.
Archiving and preservation	The data will be stored in Zenodo as well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.6. Datasets in WP7 - Innovative and co-creation governance: Targeted policy development and implementation

6.6.1. Interviews and surveys results on policy learning in the ULLs

Data set reference and name	WP7_1_Interviews_and_surveys_results_on policy_learning_in_the_ ULLs
Data set description	Two types of data are included in this dataset. Firstly, data that

	generated through qualitative interviews which is recorded and transcribed, in total or parts thereof, and stored safely and in anonymous forms. This data informs the analysis of how different idealized engagements can lead to policy learning in the URBNANOME ULLs and prompt novel forms and practices of policy making. Secondly, data is generated through a survey among cities. The aggregated dataset covers different types of cities, based on the findings from the interviews and engagement with the URBANOME ULLs.
Standards and metadata	The data will be only shared in aggregated form and paying due respect to GDPR regulation. These datasets will be made available in .docx and/or in .pdf format. This sustains transparency of the analysis and enables reproducible use of data, thus enhancing the validity of the URBANOME findings.
Data sharing	The data from the interviews will be kept confidential, due to confidentiality and GDPR. Furthermore, the data from the qualitative interviews and the survey will, as meta-data and as input to the analyses, be integrated in deliverables, scientific articles, and public dissemination. Aggregated data to preserve anonymities may be shared publicly with the expectation that both the URBANOME project and the European Commission's Horizon 2020 as the funding program are properly cited.
Archiving and preservation	The aggregated data will, when confidentiality and GDPR allows, be stored in Zenodo a s well as in the URBANOME Global Data Repository which will remain operational for 5 years after the project.

6.7. Datasets in WP8 - Evidence-based feedback to the policy makers

No datasets are expected to be generated in WP8.

6.8. Datasets in WP9 - Dissemination, communication, training and exploitation

6.8.1. Stakeholders Database

Data set reference and name	WP9_1_Stakeholders_Database
Data set description	The dataset consists of the identified URBANOME stakeholders
Standards and metadata	The extended dataset will be available as an SQL database on the URBANOME intranet webpage with restricted access, and therefore downloadable in xls(x), csv, pdf format accordingly.

	For each stakeholder identified the following metadata information will be provided in the concise dataset: • Organization name and address • Field of activity • Type of stakeholder • Website link and Social media • Stakeholder Logo • Description of expertise • Services/products deliverable for URBANOME • Geographical reach • Description of interest regarding the project Full details including: • Reference Person for the Stakeholder: Name and Surname; E-mail address; Phone number; Role in the Organization; Job title. Contact details (i.e. email address, phone number, etc.) will be not shared for privacy issues and will be included only in the extended version kept by the project only for internal use. They are collected with the specific aim of actively involving stakeholders in the project and to proceed with targeted communications.
Data sharing	The concise list and description of Stakeholders will be publicly freely available to anyone on URBANOME public webpage. The extended dataset cannot be shared for privacy issues. Nonetheless, all the details are stored in private intranet URBANOME webpage for internal use only and may be shared upon written request to the Project Coordinator. The goal is to build an active community with the creation of an internal and public database where stakeholders can interact and find new opportunities for collaboration.
Archiving and preservation	The concise dataset has been stored in URBANOME public webpage, and in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.8.2. Scientific publications

Data set reference and name	WP9_2_Scientific_Publications
Data set description	The dataset consists of all the URBANOME scientific publications produced during the project life.
Standards and metadata	Metadata maximise the discoverability of publications and ensure the acknowledgment of EU funding. The inclusion of metadata is necessary for adequate monitoring, production of statistics, and assessment of the impact of H2020. In addition to

	basic bibliographic information about deposited publications the following metadata information are provided:
	• EU funding acknowledgement:
	  • Contributor: "European Union (EU)" & "Horizon 2020"
	• Project Information:
	  • Grant number: "945391"
	  • Project Acronym: "URBANOME"
	  • Project Name: "Urban Observatory for Multi-participatory Enhancement of Health and Wellbeing"
	  • Publication Date.
	  • Journal name
	  • DOI (if any)
	  • Keywords
	  • Authors and Contributors

Data sharing	URBANOME scientists committed themselves to follow the 'Open Access Publishing Model', meaning that the scientific articles are provided in open access (either Gold or Green) mode by the scientific publisher.
Archiving and preservation	Publications are freely available through the ZENODO repository as well as in the URBANOME website

6.8.3. Dissemination and Communication Activities Monitoring

Data set reference and name	WP9_3_Dissemination_and_Communication_Activities
Data set description	The dataset aggregates all the actions performed by URBANOME partners to disseminate and/or communicate the project towards the target stakeholder groups. The dataset will draw input from an online reporting form created for this purpose (https://ec.europa.eu/eusurvey/runner/urbanomedissrep)
Standards and metadata	The dataset will be available in xls(x) format. Basic information regarding the performed activity will be featured in respective fields: • Name of Organisation • Type of activity • Title of activity / slogan • Title of the event (if applicable) • Venue (if applicable) • Date • Title of the presentation (if applicable)

	• URL of the activity (if applicable) • Type and number of audiences reached
Data sharing	The dataset has a unified form and it will be accessible by URBANOME partners
Archiving and preservation	The dataset will be stored in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

6.8.4. Database of Standards applicable to URBANOME research and interventions

Data set reference and name	WP9_4_URBAMOME_Standards_Database
Data set description	This dataset entails all the identified national and global standards that are relevant to URBANOME research activities throughout WP2 to WP6. It also, lists the aspects developed with the context of URBANOME research and interventions that present strong relation to existing standards. This inventory will be assessed periodically based on partners feedback, so that it stays updated in terms of the standards state of the art and corresponding to WPs and Tasks progress. In a more mature project stage, the dataset will aggregate all the standardization activities performed by partners in order to engage with relevant standardization bodies/authorities.
Standards and metadata	The dataset is available in xls(x) format. Basic information regarding the reported information will entail the following fields: • Standard's Number/reference • Title • Date of publication • Author(s)/Publisher(s) • Standard definition • URL
Data sharing	The dataset has a unified form and it will be accessible by URBANOME partners
Archiving and preservation	The dataset will be stored in the URBANOME Global Data Repository which will remain operational for 5 years after the project end.

References

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

Aashaka Shah, Vinay Banakar, Supreeth Shastri, Melissa Wasserman, and Vijay Chidambaram. "Analyzing the Impact of GDPR on Storage Systems." In: Proceedings of the 11th USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage). July 2019

Chriskos, P., Frantzidis, C. A., Gkivogkli, P. T., Bamidis, P. D., & Kourtidou-Papadeli, C. (2018). Achieving accurate automatic sleep staging on manually pre-processed EEG data through synchronization feature extraction and graph metrics. Frontiers in human neuroscience, 12, 110.

Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983;24:385–396. Cited: in: : PMID: 6668417

Ormandy D. Housing and Health in Europe: The WHO LARES project [Internet]. Taylor & Francis; 2009. Available from: https://books.google.fr/books?id=BXx9AgAAQBAJ

Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, Parkinson J, Secker J, Stewart-Brown S. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health and Quality of Life Outcomes. 2007;5:63. doi: 10.1186/1477-7525-5-63

Acknowledgments

The URBANOME project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 945391.