

Data Management Plan

Project acronym:

URBAG

Project Title:

**Integrated System Analysis of
Urban Vegetation and Agriculture (URBAG)**

Grant Agreement No: **818002**

The project has received funding from the European Research Council

Duration of the project:

1st September 2019 – 31st August 2024

Version 1



European
Research
Council

UAB
Universitat Autònoma
de Barcelona

ICTA



Institut de Ciència
i Tecnologia Ambientals - UAB

Table of Contents

1. Introduction.....	3
2. Data Summary	4
2.1. Primary Data	4
2.2. Data from subcontracts funded by URGAG	5
2.3. Third Party Data.....	5
3. FAIR data principles	6
3.1. Making data findable	6
3.2. Metadata.....	6
3.3. Making data openly accessible	7
3.4. Making data interoperable	7
3.5. Making data re-usable.....	7
4. Data storage and security	8
4.1. Long-term data preservation	8
4.2. Data security	8
4.3. Ethical aspects.....	9
5. References.....	10
6. Appendix.....	11
6.1. Appendix 2: Port d'Informació Científica (PIC) - service specifications.....	11
6.2. Appendix 1: informed consent form to participate in interview	12

1. Introduction

The URBAG project aims to find out how urban green infrastructures can be most efficient in contributing to urban sustainability. Different combinations of urban, peri-urban agriculture and green spaces are evaluated to determine the best performance in terms of local and global environmental impact. For this purpose, URBAG is developing a novel and comprehensive analysis integrating life cycle impacts of the resources required for green infrastructures with the understanding of how green infrastructures impact the urban atmosphere interaction (see Figure 1). For a full description of the tasks set out to accomplish this goal, please visit www.urbag.eu or contact Gara Villalba (gara.villalba@uab.cat).

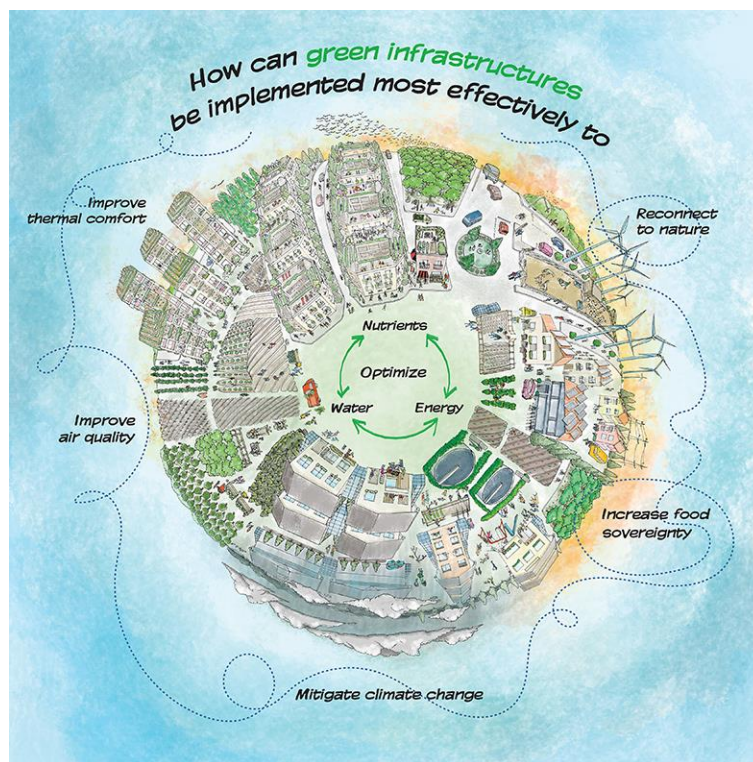


Fig 1. Graphic representation of the urban interactions that will be studied by URBAG to understand how green infrastructures can be most efficient in contributing to sustainability.

URBAG will be focusing on two case study cities and their metropolitan area to develop its research: Barcelona (Spain) and Oslo (Norway). For each of them, data related to climate change, urban metabolism, ecosystem services and social vulnerabilities will be collected and analysed using GIS, life-cycle assessment, modelling and stakeholder workshops discussion in order to provide concrete results on the performance of the green infrastructures.

As an ERC-funded project, URBAG is committed to the Open Research Data (ORD) policy. For this purpose, the FAIR data principles are followed to make sure data is Findable, Accessible, Interoperable and Re-usable.

This document constitutes the first of the Data Management Plan (DMP) of the URBAG project and includes a description of the management life cycle for all research data belonging to the URBAG project.

2. Data Summary

The data collected and produced in URBAG can be divided into three main groups: Primary data, data from subcontracts funded by URBAG and third-party data (see Figure 2). These three types of datasets have their own characteristics and each of them will be treated using the FAIR principles

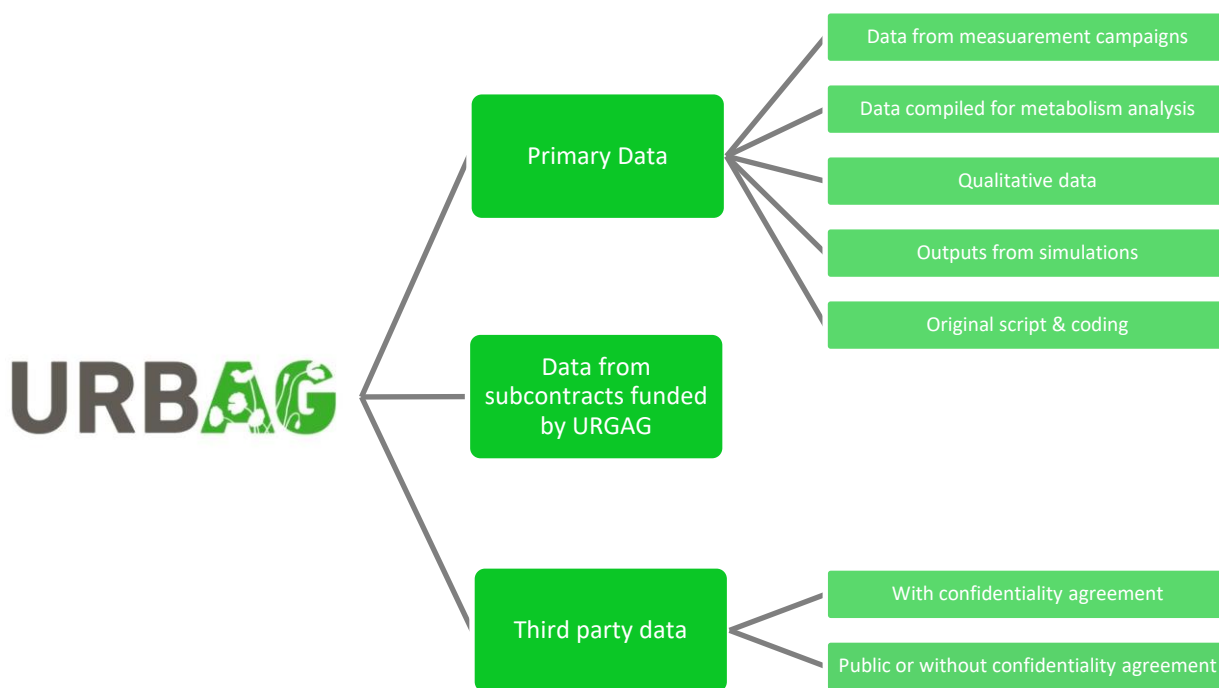


Fig 2. Datasets collected/produced in URBAG divided by its origins.

2.1. Primary Data

Primary data will be that produced or obtained directly by URBAG and is divided in five different groups:

- Data from measurement campaigns: URBAG project is generating primary data with sampling campaigns to study different characteristics of the atmosphere in the regions of study. Air samples

are taken to measure concentration of different gases; associated to these air samples, atmospheric conditions such as temperature, pressure, wind velocity or wind direction are also being recorded.

- Data compiled for metabolism analysis: URBAG project will be analyzing the food-energy-water linkages in order better how to better understand how to optimize the production of urban food, considering urban morphology and life cycle impacts. In order to do this, URBAG will be collecting data on water use for crop irrigation, as well as fertilizer use per crop.
- Qualitative data: URBAG will be holding workshops with different stakeholders from each one of the cities. There, discussions about the relationship between nature-based solutions and city resilience will be held. These exercises will provide useful information on the current state and perception of the urban environmental challenges and how should these be prioritized. Similarly, surveys and questionnaires will be conducted in these events. It is also possible that interviews to stakeholders take place during the project.
- Outputs from simulations: URBAG project is generating simulated data by executing computer models. The objective of these models is to increase the granularity in which the atmosphere conditions are known for a specific region. The inputs for these models include detailed morphological description of the region of study as well as initial conditions of the atmosphere. After executing these models in a supercomputer, we obtain large datasets describing the state of the atmosphere with large level of resolution in both the temporal and special dimensions.
- Original script & coding: URBAG team will develop its own coding and scripts to generate models that simulate how the atmospheric conditions can be impacted by green infrastructures. Moreover, GIS-Python based scripts will be created to display land uses in the cities, including green and constructed areas, and waste treatment plants.

2.2. Data from subcontracts funded by URGAG

Through agreements with third party organizations, URBAG will also acquire some datasets, in particular those related with anthropogenic emissions in the metropolitan area of Barcelona. In these situations, the data will be submitted to the same policies as the primary data.

2.3. Third Party Data

Data collected by third parties will also be used during URBAG's research. In this case there are two types of datasets:

- Third Party data with confidential agreements: through agreements with third party organizations, the URBAG project has access to datasets of water at energy consumption at a low level. More information about its treatment can be found in section *4.3 Ethical considerations*.

- Public data or third-party data without confidential agreements: this data will come from public entities or third parties that explicitly express their willingness to share and offer the data for public use.

3. FAIR data principles

URBAG follows the ‘FAIR data principles’ by making the project’s research data findable, accessible, interoperable and reusable. The following sections describe how the data is organized to fulfil each requirement in an appropriate manner.

3.1. Making data findable

Each one of the datasets of the project will be accessible when published in the repositories Zenodo and GitHub, which were chosen because allow open access, long-term storage, and downloads for all kind of data types. Within the Zenodo platform, an URBAG community will be created, this way the datasets will not only be linked to their authors, but also grouped together with other datasets from URBAG.

When it comes to larger datasets that cannot be stored in these repositories due to their size (such as simulations), we will be offering a description of these on their respective research publications and, upon request, we will provide a summarized or full version of them.

3.2. Metadata

Each file included in our inventory of datasets has a data description sheet with the necessary information to describe the underlying data in sufficient detail so others can use it. URBAG metadata will include:

- title
- creator(s) and contact person(s): names, first names, email
- date,
- version,
- location
- contributor: e.g., funding body, including the grant agreement number,
- data format(s),
- keywords,
- identifiers: DOI and url
- access rights: license (s)
- suggestion for citation
- a description of the methodology that has produced the data
- brief description of the URBAG project.

URBAG will also provide a README.txt file with a description of the structure of the data sets and subsets, on how the data is organized, how data sets are related to each other and on how to use the data. The metadata will follow the Dublin Core Standard (<http://dublincore.org/>). Furthermore, and regarding dataset version, we'll rely on the version tracking feature of Zenodo.

3.3. Making data openly accessible

All URBAG's journal publications will be open access and so the datasets associated with them. As previously mentioned, datasets will be stored in the Open Science platform Zenodo. Similarly, any code developed will be made openly accessible on GitHub.

It should be noted that there is a particular situation to consider for large datasets, often generated by simulations, that cannot be available on the web due to its large size. For these cases, we will be offering either a summarized version with the most relevant data or the full version, all upon request.

Finally, an exception needs to be made when it comes to Third Party datasets containing information about water or energy consumption at a low level. In these cases, URBAG will not be publishing them because, although they do not contain personal information, they could be used to infer it.

3.4. Making data interoperable

To guaranty the interoperability of the data, whenever possible, the URBAG project will use open file formats to store the data. To avoid that researchers from outside the URBAG network must purchase software licenses or learn how to function with nonstandard file types, the following file types will be used:

- Text documents: docx, txt, pdf
- Tabular data: csv, xlsx
- Meteorological data: NETCDF
- Georeferenced data and maps: QGIS (.shp, .kml, .dbf, .qgz, tif)
- Simulation outputs: NETCDF4

3.5. Making data re-usable

The URBAG data will be public on the Open Science platform ZENODO and GitHub. All published URBAG data will be reusable and it is expected to be useful to both academic and public sector groups linked to urban development, urban food production, climate resilience and urban metabolism. Currently, there is no estimate of how long the data will remain re-usable, but it is expected to be kept available for as long as the aforementioned repositories are active.

When making our data publicly available we will attach a license, this is a way to make it safe for others to use it. The chosen license is a general Creative Commons license, and particularly to the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA license). This license allows re-distribution and re-use of the licensed work and data on the condition that the creator is appropriately credited, the data are not commercially used and that modified data will be distributed under the same license.

4. Data storage and security

The following sections will detail the practices related to data security and long-term data preservation.

4.1. Long-term data preservation

The data collected and produced by URBAG will be available in the repositories (Zenodo, GITHUB, UAB) as long as these remain active over time. In the case of larger datasets that cannot be stored in these repositories, it is planned to keep them in the PIC (Port d'Informació Científica), a computer and digital storage center located at the Autonomous University. For detail information about the PIC services, please see Annex 6.1.

URBAG does not contemplate the destruction of data during or after the project, except in two cases:

- When project participants or researchers specifically request it.
- For digital copies of consent forms obtained from participants and interviewees, these will be destroyed two years after the end of the project.

Table 1 includes the long-term data preservation policy for each one of the repositories and dataset types.

4.2. Data security

To ensure the maintenance of the data, URBAG will be employing different digital storages and backup spaces. In a first instance, the datasets (or their access paths, in the case that their size is very large) will be concentrated in OneDrive, a Microsoft storage and synchronization service contracted by the Autonomous University of Barcelona. The information stored in OneDrive will be backed up weekly, using an external hard drive to ensure that there is no risk of data loss. Finally, and as the data become public, it will also be accessible through the Zenodo and GitHub repositories. Table 1 presents the compatibility between the datasets and the digital storages and backup systems.

	Data from measurement campaigns	Data compiled for metabolism analysis	Qualitative data	Outputs from simulations	Original script & coding	Storage capacity	Open Access	Long-term storage
Zenodo	x	x	x			50 Gb per dataset. No data set limits	YES	YES
GitHub					x	5 Gb per dataset. No data set limits	YES	YES
PIC				x		10 TB	NO	YES
OneDrive	x	x	x			1 TB per user	NO	YES
External hard drive	x	x	x			1.8 TB	NO	YES

Table 1. URBAG digital storages and backup spaces, its compatibilities with the different datasets, open access and longevity

4.3. Ethical aspects

Studies involving the collection of personal information will be reviewed by the “Comissió d'Ètica en Experimentació Animal i Humana (CEEAH)” the Ethical board of the Autonomous University of Barcelona. Any future survey conducted either in person or online will also go through ethical approval by the same committee.

In the case of interviews, participants will be asked to fill a consent form and a questionnaire related to the research purposes of the interview. A template of the document can be found on Annex 6.2.

Moreover, any dataset containing personal information from participants will have its access strictly controlled and will go through anonymization process (e.g. by using pseudonyms during the analysis and presentation of qualitative data), hence confidentiality is ensured during the entire processes. The published outcomes of the analyses should report only aggregated summaries of the data and by no way will it be possible from the results to personally identify any of the participants.

URBAG is also maintaining stakeholder's contact information with the only purpose of sending information about the development of the URBAG project to any person that has shown interest on it.

Finally, personal data will be kept in a form which permits identification of individuals for no longer than is necessary for the purposes of the URBAG project. Within two years after the end of the project all personal data will have been destroyed.

5. References

EC, 2017. H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Version 3.2.

ERC, 2017. Guidelines on Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020. Version 1.1.

ERC, 2016. Mono-Beneficiary Model Grant Agreement, ERC Starting Grants, Consolidator Grants and Advanced Grants (H2020 ERC MGA— Mono). Version 3.0, p. 55-56.

6. Appendix

6.1. Appendix 2: Port d'Informació Científica (PIC) - service specifications

- Access
 - POSIX
 - Internal access from PIC, UIs (Interactive servers) and WNs (servers) bat batch)
- Structure and Features
 - /data/co2flux/common – NFS
 - An external copy and 7 snapshots local daily
 - Write and read permission for group
 - Quote max group 10 TB
 - Data usage: data to access all project users
 - NFS server based on ZFS file system
 - HW Equipment: Disk server with SAS disks and level redundancy ZFS
 - CentOS7 operating system configured with most up-to-date versions from security packages
 - /software/co2flow – NFS
 - Local Snapshots: 3 monthly, 5 weekly, 8 daily, 2 hours
 - External Backups: 1 snapshot external
 - Write permission for an admin (gara/co2flow) and read by group
 - max Max 100 GB quota
 - High IOP performance
 - Usage: common software for the project
 - Appliance NFS
 - HW Equipment: NetApp Fas2552 cockpit and redundancy at hardware level
 - Operating system: Data ONTAP
 - /home/user – NFS
 - Local Snapshots: 3 monthly, 4 weekly, 16 daily, 3 hourly
 - External Backups: 1 snapshot external
 - Quote max 10GB
 - High IOPS Performance
 - Usage: local user data
 - Appliance NFS
 - HW Equipment: NetApp Fas2552 cockpit and redundancy at hardware level
 - Operating system: Data ONTAP

6.2. Appendix 1: informed consent form to participate in interview

Information about the study

You are asked to participate in a research study coordinated by _____ from the Autonomous University of Barcelona (UAB) This study will assess _____.

Please read the information below and ask questions about anything you do not understand before deciding whether or not to participate in the interview.

- This interview is voluntary. You have the right to refuse to answer any question and to stop the interview at any time.
- There are no notably foreseeable consequences for you or your organization from engaging in this interview, i.e. from answering the questions.
- You will not be compensated for your participation.
- We would like your permission to record this interview. If you agree to have this interview recorded, you may revoke recording permission at any time. If you agree to be recorded:
 - We would like permission to indicate your name and position title in any publications where we use direct quotations or refer to information you provide during this interview (without your name). If you do not grant permission, only the name of your organization and a generic position title will be used. Your name and actual position title will remain confidential.
 - If you do not want your name and position title linked to this interview, we will ensure your confidentiality by not putting your name or any other information that can identify you on the tape or digital recording file or on the transcript we make of the interview. Instead, we will assign a numeric code to your materials. Only the primary investigators named above will have access to information that links the numeric code to your identity.
 - You have the right to ask that all data in link with you be removed from the project.
 - The recording will be stored in a secure workspace for 10 years after finishing the project.
 - The data that you share with us will be used for academic publication purposes and reports.
 - The analysis of the data that you share will not involve automatized decisions, nor the elaboration of profiles predicting your personal preferences, behaviors, and attitudes.

Study contact details for further information:

This research is coordinated by _____. You can contact at _____ at any time with any concerns or questions about your rights regarding data access, corrections, rectification, suppression, opposition, transfer, and limitation of analysis.

CONSENT FORM

Please tick the appropriate boxes (you will be given a copy of this informed consent form)

Yes No

Taking part in the study:

I have read and understood the study information, or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.

I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.

I understand that taking part in the study involves an audio-recorded interview which will be transcribed as text and be stored in a secure workspace for 10 years after finishing the project.

Use of the information in the study:

I understand that information I provide will be used for academic publication purposes and reports.

I understand that personal information collected about me that can identify me, such as my name, will not be shared beyond the study team.

I agree that my information and position title can be quoted in research outputs.

I agree to be audio/video recorded.

Future use and reuse of the information by others:

I give permission for the information that I provide to be archived in a safe data repository so it can be used for future research and learning.

Signatures:

Participant Name _____

Participant Signature _____ Date _____

I have accurately read out the information sheet to the above-named research participant in her/his native language, and to the best of my ability, ensured that the participant understand to what she/he is freely consenting.

Researcher Name _____

Researcher Signature _____ Date _____