

Data Management Plan GARLOCK

Project Name GARLOCK: Unlocking the climate change adaptation and mitigation potential of the garden complex - Data Management Plan GARLOCK

Project Identifier 3E210512

Principal Investigator / Researcher Ben Somers, Trui Steen, Pieter Raymaekers, Florian Lange, Stien Heremans, Kelly Wittemans, Valerie Dewaelheyns, Janne Teerlinck, Marc Olijslagers

Project Data Contact Valerie Dewaelheyns, +32 16 32 84 97, valerie.dewaelheyns@kuleuven.be

Description Gardens are an integral part of the urban green and blue network but their potential for climate change mitigation and adaptation remains underutilized. In our GARLOCK project, we consider gardens as social-ecological systems and hypothesize that a full realization of this potential thus requires a better insight in and integration of their ecological and social dimensions. In this project, we therefore focus on: (i) obtaining timely and reliable data on garden composition, configuration, use and management through an integrated use of novel data acquisition and analysis approaches like remote sensing, deep learning, and citizen science; (ii) strengthening the empirical knowledge base on the link between these garden characteristics and urban climate, by combining remote sensing and autonomous sensor networks; (iii) gathering insight into the decision-making patterns of garden owners and identifying policy pathways for mobilizing them towards climate resilient gardening. Finally, we anticipate to integrate our new insights into the development of policy support tools for local and regional governments; allowing them to design and implement effective climate mitigation and adaptation policies on domestic green space and to monitor the effects of these policies.

Institution KU Leuven

1. General Information

Name of the project lead (PI)

Ben Somers

Internal Funds Project number & title

KU Leuven bijzonder onderzoeksfonds C2-project 4 jaar SAP code: 3E210512

2. Data description

2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data
- reuse existing data 

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

WP1 - Combining remote sensing and citizen science for inventorying the composition, configuration and management of the garden complex

Input data

	Dataset Name	Origin of data	Type of data	File format	Volume (estimated)
1.1	Scientific literature	<i>Secondary</i> : Online repositories of scientific manuscripts	Textual data	.pdf	500 MB
1.2	Tuinenkaart	<i>Secondary</i> : Internal dataset originating from the GARMON project by the Flemish government	Geospatial data, Vector Geodata	.shp, .dbf, .shx	2.21 GB
1.3	Grootschalig referentiebestand (GRB)	<i>Secondary</i> : Flanders Information Agency, available on Geopunt.be	Geospatial data, Vector Geodata	.shp, .dbf, .shx, .sld, .prj, .lyr	4.32 GB
1.4	Landgebruikskaart-2019	<i>Secondary</i> : Department of environment and spatial development, available on Geopunt.be	Geospatial data, Raster Geodata	.cpg, .dbf, .lyr, .ovr, .pdf, .qml, .sld, .tif, .xml, .lyr, .qlr	68.6 MB
1.5	NARA landgebruikskaart (level 1, 2 and 3)	<i>Secondary</i> : INBO and VITO, available on geo.inbo.be/ecosysteemdiensten	Geospatial data, Raster Geodata	.tfw, .tif, .tif.vat.dbf, .xml, .txt, .lyr	244 MB
1.6	Groenkaart Vlaanderen 2009, 2012, 2015 and 2018	<i>Secondary</i> : Agency for nature and forest, available on Geopunt.be	Geospatial data, Vector Geodata	.tfw, .tif, .pdf, .xml	1.14 GB + 1.04 GB + 1.16 GB + 1.05 GB
1.7	Landbouwgebruikspercelen	<i>Secondary</i> : Agency for agriculture and Fisheries, available on Geopunt.be	Geospatial data, Vector Geodata	.pdf, .xml, .dbf, .lyr, .prj, .shp, .shx, .sld	411 MB
1.8	Gewestplan	<i>Secondary</i> : Department of environment and spatial development, available on Geopunt.be	Geospatial data, Vector Geodata	.pdf, .lyr, .dbf, .prj, .qlr, .shp, .shx, .cpg	171 MB

1.9	Statistical sectors Belgium	<i>Secondary:</i> Statbel, available on statbel.fgov.be	Geospatial data, Vector Geodata	.pdf, .dbf, .prj, .shp, .shx, .cpg, .doc	250 MB
1.10	Vlaams Hydrografische Atlas - Waterlopen	<i>Secondary:</i> VMM, available on klimaat.vmm.be	Geospatial data, Vector Geodata	.pdf, .xml, .dbf, .prj, .shp, .shx, .cpg, .WOR	89.4 MB
1.11	Climate data VMM	<i>Secondary:</i> Open data on climate impact, available on klimaat.vmm.be	Geospatial data, Vector Geodata	.pdf, .xml, .dbf, .prj, .shp, .shx, .cpg, .WOR	1 GB
1.12	Verstedelijkt, randstedelijk en landelijk Vlaanderen, indeling op basis van statistische sectoren	<i>Secondary:</i> Department of environment and spatial development, available on Geopunt.be	Geospatial data, Vector Geodata	.cpg, .dbf, .lyr, .pdf, .prj, .qlr, .shp, .shx, .sld	25.6 MB
1.13	Biologische Waarderingskaart en Natura 2000 Habitatkaart	<i>Secondary:</i> INBO, available on Geopunt.be	Geospatial data, Vector Geodata	.pdf, .xml, .lyr, .dbf, .prj, .shp, .shx, .sld	552 MB
1.14	AVIRIS-NG dataset (hyperspectrale beelden)	<i>Secondary:</i> Internal dataset, originating from EU's CHIME hyperspectral mission APEX sensor, 2 m pixels, spectral range: 350-2500 nm, spectral resolution ~10 nm	Geospatial data, Raster Geodata	.enp, .hdr	124 GB
1.15	LiDAR Digitaal Hoogtemodel Vlaanderen II - ruwe remote sensing data	<i>Secondary:</i> Flanders Information Agency, available on www.remotesensing.vlaanderen.be (Average resolution 15pts/m2)	Geospatial data, Point cloud	.laz	101,3 GB per map sheet
1.16	Sentinel data	<i>Secondary:</i> Sentinel images from Copernicus (www.copernicus.eu/en)	Geospatial data, Raster geodata	.jp2, .xml, .gml	1 GB per image
1.17	Orthophotos	<i>Secondary:</i> annual orthophotos (RGB, 45 cm pixel size) provided by the Flemish government (https://overheid.vlaanderen.be/en/producten-diensten/aerial-photographs).	Geospatial data, Vector geodata	.pdf, .xml, .cpg, .dbf, .prj, .sbn, .sbx, .shp, .shx, .j2w, .jp2, .j2w	27,9 GB
1.18	Data from MijnTuinlab	<i>Secondary:</i> Data of garden composition and configuration collected from citizen science platform MijnTuinlab - reference data for calibration and validation of models	Geospatial data, Vector geodata	.shp, .shx, .prj, .dbf	< 100 MB
1.19	Field data of gardens	<i>Primary and secondary:</i> Data collected of garden composition and configuration in > 150 gardens - reference data for calibration and validation of models	Geospatial data, Vector geodata, observational	.shp, .shx, .prj, .dbf	< 100 MB
1.20	Survey responses on garden management	<i>Primary:</i> Online survey responses by Flemish citizens regarding their garden design, landscaping and management	Textual/numerical data (survey), observational	.csv; .xlsx	< 10 MB
1.21	Analytical scripts and coding	<i>Primary and secondary:</i> R-scripts, SQL coding, FME Workflows and GIS Workspaces using packages from online repositories such as CRAN	Code or workspace	.R, .sql, .fmw, .qgz	< 100 MB

The geospatial input data will be processed using R scripts, SQL coding, FME workflows and GIS tools (QGIS and ArcGIS). It has not yet been decided how the survey data will be further processed. Once a decision is made, this will be added to the DMP.

Output data

	Dataset Name	Origin of data	Type of data	File format	Volume (estimated)
--	--------------	----------------	--------------	-------------	--------------------

2.1	Annotations	<i>Primary:</i> Desk research: notes, summaries and other data collected as part of the review of academic and practitioner literature during the project	Textual data	.url (to KU Leuven OneDrive/ OneNote), .docx, .pdf	500 MB
2.2	Tuinenkaart 2.0	<i>Primary:</i> Processed data based on the garden map	Geospatial data, derived	.shp, .shx, .prj, .dbf	2.21 GB
2.3	Characterization of the garden complex	<i>Primary:</i> Processed data based on the garden map and other geodatasets	Geospatial data, derived	.shp, .shx, .prj, .dbf	> 2 GB
2.4	Statistical analysis of garden characteristics	<i>Primary:</i> Processed data based on the garden map and other geodatasets	Numerical, derived	.xlsx, .csv	< 100 MB
2.5	Connectivity Indices/ Metrics	Processed data based on the garden map, land use maps and green maps	Geospatial data, derived	.shp, .shx, .prj, .dbf	> 2 GB
2.6	High resolution garden composition and configuration of Leuven	<i>Primary:</i> Derived data based on hyperspectral images and lidar data – detailed surface characterization map for all gardens in Leuven	Geospatial data, derived	.shp, .shx, .prj, .dbf	> 2 GB
2.7	Garden composition and configuration of Flanders	<i>Primary:</i> Derived data based on satellite data and orthophotos – Land cover map for all gardens in Flanders	Geospatial data, derived	.shp, .shx, .prj, .dbf	> 2 GB
2.8	Quantitative analysis of the garden complex	<i>Primary:</i> Quantitative assessment of the spatial configuration of the garden complex for Flanders (general classes) and Leuven (detailed classes)	Numerical, derived	.xlsx, .csv	< 100 MB

2.9	Processed survey data on garden composition and management	<i>Primary:</i> Processed survey data for analysis – geospatial statistics on garden management and use	Numerical, derived	.xlsx, .csv	< 10 MB
2.10	Paper manuscripts	<i>Primary:</i> Desk research using all above data	Textual data	.docx	< 10 MB

WP2 - Modelling the climate change alleviation potential of the garden complex

Input data

	Dataset Name	Origin of data	Type of data	File format	Volume (estimated)
3.1	Scientific literature	<i>Secondary:</i> Online repositories of scientific manuscripts	Textual data	.pdf or .txt	500 MB
3.2	Meteorological data: Leuven	<i>Primary and secondary:</i> Raw 16 seconds data collected by the Leuven.cool weather station Network (Leuven.cool)	Numerical, observational	.csv	6 GB per quarter
3.3	Meteorological data: Reference	<i>Secondary:</i> 10 minute aggregated data collected by the official equipment of the RMI (Royal meteorological institute of Belgium) in Uccle, Diepenbeek and Humain	Numerical, derived	.csv	20 MB
3.4	Thermal remote sensing data	<i>Secondary:</i> Thermal infrared (TIR) band of a Landsat 5 thematic mapper (TM) image (Earth Explorer - USGS)	Geospatial data, Raster geodata	.tiff	1 GB per image
3.5	Soil moisture and temperature data	<i>Primary:</i> Raw Measurements each 15 minutes collected by sensors, previously used in curieuzeneuzen network (TOMST TMS4 sensor)	Numerical, observational	.csv	6 GB per quarter
3.6	Air quality data	<i>Primary:</i> Raw measurements collected by sensors	Numerical, observational	.csv	6 GB per quarter
3.7	LiDAR Digitaal Hoogtemodel Vlaanderen II - ruwe remote sensing data	<i>Secondary:</i> Flanders Information Agency, available on www.remotesensing.vlaanderen.be Average resolution 15pts/m2)	Geospatial data, Point cloud	.laz	101,3 GB per map sheet
3.8	Data Bodemkundige dienst België	<i>Secondary:</i> Soil data collected by the Bodemkundige dienst van België, vzw	Numerical, observational	csv	< 100 MB
3.9	Field data of gardens	<i>Primary and secondary:</i> Data collected of garden in > 150 gardens - in situ measurements of above (via allometric measurements of the vegetation) and below ground (via destructive chemical analysis of soil samples)	Geospatial data, Vector geodata, observational	.shp, .shx, .prj, .dbf	< 100 MB
3.10	Natuurpunt waarnemingen.be	<i>Secondary:</i> Open data on plant and animal observations in Belgium	observations	csv	< 100 MB
3.11	Analytical scripts and coding	<i>Primary and secondary:</i> R-scripts and SQL coding using packages from online repositories such as CRAN	code	.R, .sql	< 100 MB

The input data will be processed using R scripts and SQL coding.

Output data

	Dataset Name	Origin of data	Type of data	File format	Volume (estimated)
4.1	Annotations	<i>Primary:</i> Desk research: notes, summaries and other data collected as part of the review of academic and practitioner literature during the project	Textual data	.url (to KU Leuven OneDrive/OneNote), .docx, .pdf	500 MB
4.2	Corrected Air temperature data + other atmospheric meteorological variables	<i>Primary and secondary:</i> Meteorological Data (Leuven.cool + RMI)	numerical, derived	csv and shp	40MB per quarter
4.3	Corrected soil moisture and soil temperature data	<i>Primary:</i> Soil moisture and temperature data (sensors TOMST TMS4 sensor)	numerical, derived	Csv and shp	40MB per quarter
4.4	Corrected air quality data	<i>Primary:</i> Air quality sensor data	Numerical, derived	Csv and shp	40MB per quarter
4.5	Carbon stock	<i>Primary:</i> Quantitative dataset on below and aboveground carbon stocks in gardens	Numerical, derived	Csv and shp	1 GB
4.6	Spatial assessment of the current and potential climate mitigation and adaptation effect of the Flemish garden complex	<i>Primary:</i> Relate garden and urban land cover composition and configuration to the potential of gardens to mitigate climate pressures	Numerical, derived	Csv and shp	1 GB
4.7	Paper manuscripts for scientific and popular science and policy supporting communication	<i>Primary:</i> Desk research using all above data	Textual data	.docx	< 10 MB

WP3 - Designing a toolset for engaging citizens as stewards of climate proof gardening and inducing sustainable behavioral change

Input data

	Dataset name	Origin of data	Type of data	File format	Volume (estimated)
5.1	Scientific literature	<i>Secondary:</i> Online repositories of scientific manuscripts	Textual data	.pdf	500 MB
5.2	Policy	<i>Secondary:</i> Publicly available policy papers, reports, fact sheets, parliamentary questions	Textual data	.pdf, .html	500 MB
5.3	Survey on behavioral antecedents, acceptance of policy instruments and socio-economic context	<i>Primary:</i> Survey aiming at Flemish citizens; validated survey scales; adaptations of validated survey scales; unvalidated survey scales, data collected during May-July 2022	Textual data	.docx	< 500 kb
5.4	Survey responses on behavioral antecedents, acceptance of policy instruments and socio-economic context	<i>Primary:</i> Survey responses by Flemish citizens	Textual data, Numerical data, derived	.csv; .xlsx	1-5 MB
5.5	Interview questionnaire	<i>Primary:</i> Interview questionnaire, content to be determined	Textual data	.docx	< 500 kb
5.6	Interview transcripts (audio and text)	<i>Primary:</i> Audio files and transcripts from interviews with relevant stakeholders, to be determined who will be interviewed	Textual data, audio data	To be determined which program will be used, probably Nvivo; .docx, .mp3	< 1 GB
5.7	Observational data from participation processes	<i>Primary:</i> Available meeting reports, interview results, observational notes etc.	Textual data	To be determined which program will be used, probably Nvivo; .docx	< 500 kb
5.8	Survey experiment on behavioral interventions and policy communication strategies	<i>Primary:</i> Survey responses by Flemish citizens	Textual data; Numerical data; derived	.csv; .xlsx	1-5 MB

It has not yet been decided how the data will be further processed. Once a decision is made, this will be added to the DMP.

Output data

	Dataset name	Origin of data	Type of data	File format	Volume (estimated)
6.1	Annotations	<i>Primary:</i> Desk research: notes, summaries and other data collected as part of the review of academic and practitioner literature during the project	Textual data	.url (to KU Leuven OneDrive/ OneNote), .docx, .pdf	500 MB
6.2	Processed policy data	<i>Primary:</i> Results from the analysis of (policy) documents	Textual data	To be determined which program will be used, probably Nvivo; .docx	< 500 kb
6.3	Processed survey data on behavioral antecedents, acceptance of policy instruments and socio-economic context	<i>Primary:</i> Processed survey data for analysis	Numerical data; textual data	.xlsx, .cs; To be determined which program will be used	1-5 MB
6.4	Coding results from interview analyses	<i>Primary:</i> Processed interview data	Textual data	To be determined which program will be used, probably Nvivo.	< 500 kb
6.5	Results from the observation of participation processes	<i>Primary:</i> Processed observational data	Textual data	To be determined which program will be used, probably Nvivo; .docx	< 500 kb
6.6	Processed survey experiment on behavioral interventions and policy communication strategies	<i>Primary:</i> Processed survey experiment data for analysis	Numerical data; textual data	.xlsx, .cs; To be determined which program will be used	1-5 MB

3. Ethical and legal issues

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

Yes.

Short description of the kind of personal data that will be used:

WP2

- Sensor networks (data 3.2, 3.4, 3.5)
 - contact information (name, address, telephone number) to contact participants of the sensor networks
- Survey on garden management (data 1.2)
 - Part A - garden profile (garden context, design...)
 - Part B - Context (spatial context, housing type...)
 - Part C - garden management (what, when, frequency, why) focussed on soil, water and vegetation; gardening knowledge; sources of gardening inspiration; awareness of the importance of the garden
 - Part F - socio economic context (age, sex, income level, educational level, number of residents, number of children per age category, importance of climate compared to other aspects of environmental quality of life)
 - Part G - experiences with a specific citizen science initiative - optional part for specific evaluation of citizen science initiatives

WP3

- Survey on behavioral antecedents, acceptance of policy instruments and socio-economic context (data 5.4)
 - Part A – garden profile (garden context, design...)
 - Part B - Context (spatial context, housing type...)
 - Part D: behavioral antecedents: values (16 items, validated scale Schwartz, 1992); perception of climate risk (3 to 6 items on drought, flooding and urban heat island effect); personal norms; social norms; relatedness with nature (6 items validated NR6 schaal, Nisbet et al., 2009)
 - Part E – acceptance of policy instruments (trust; acceptance of 6 types of instruments regulations, subsidies, taxes, communication, nudging, co-creation/participation)
 - Part F - socio economic context (age, sex, income level, educational level, number of residents, number of children per age category, importance of climate compared to other aspects of environmental quality of life)
- Interview (data 5.6): to be determined
- Survey experiment on behavioral interventions and policy communication strategies (data 5.8): to be determined

3.2. Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s).

Yes.

Reference to ethical committee approval (PRET application):

- autonomous sensor networks: G-2022-4847 (submitted)
- online survey: G-2022-4860 (submitted)
- evaluation of citizen science initiatives: G-2022-4974 (submitted)
- online experiment: G-2022-4861 (concept)

3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

No.

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

Governmental input data (data 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.10, 1.11, 1.12, 1.13, 1.15, 1.17, 3.7, 5.2) has the following restrictions: <https://overheid.vlaanderen.be/modelllicentie-gratis-hergebruik>. Any commercial or non-commercial reuse for an indefinite period of time is permitted at no cost. The sole condition of use is an obligation to acknowledge the source.

Copernicus satellite data (data 1.16) has the following license: <https://cds.climate.copernicus.eu/cdsapp#!/terms/licence-to-use-copernicus-products>. The USGS satellite data (data 3.4) has also no restrictive license: <https://www.usgs.gov/faqs/what-are-terms-uselicensing-map-services-and-data-national-map>.

Regarding the official meteorological data provided by the RMI (data 3.3), we have the right to use but not to share the data. Some other 3rd party agreements still need to be drafted (e.g. data 3.8 and 3.10). The DMP will be updated every time a 3rd party agreement will be signed.

As co-developer and/or co-owner of some datasets such as Tuinenkaart (data 1.2), AVIRIS-NG (data 1.14) and Mijntuinlab (data 1.18), we will respect the applicable agreements.

For the personal data (1.2, 3.2, 3.4, 3.5, 5.4, 5.5, 5.6, 5.8), an informed consent will be signed by all participants/respondents.

4. Documentation and metadata

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

For each work package and subtask, output data and supported documentation will be created in specified folders:

- A separate README text file with a fixed template that will be used by all researchers involved in the project, containing project and contact information, a description of the type and format of all input data (+ when applicable the data collection method), the processing methods or scripts/workflows used to generate the output data and a clear description of the output data will be kept in the same folder.
- The scripts and workflows will be also have “in-file” documentation in order to better understand data preparations and analysis. Moreover, a virtual environment will be created in which the code runs. The versions of the programming languages and packages used in it, will be documented. This environment can then be recreated for reusability/verifiability.

When creating files, a clear file organization system will be used:

1. File naming: File names include the creation date, the project name, a clear (short) description of the data and the version

E.g. 20220325_GARLOCK_DMP_v3

2. Versioning: creation dates and version names (v1, v2.2, ...) are used to keep track of versions. Obsolete versions will be deleted. Changes will also be documented.
3. File structure: subfolders for different projects and different work packages have been created.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

For most input data, metadata files exist and will be stored together with the data

For each dataset that will be published or archived, a metadata file will be foreseen. The metadata file will follow the prescriptions defined by the Open data repository and/or will be based on the metafile structure of the input data. A data standard will be used when applicable.

For satellite derived products metadata are included in GeoJSON file format. The GeoJSON implementation is based on the

following standards: <https://www.ogc.org/standards/eo-geojson>.

- OGC-17-003r1 – OGC EO Dataset
- Metadata GeoJSON(-LD) Encoding Standard OGC-17-084-v0.9.0-D2 - GeoJSON EO
- Collection Metadata Encoding Standard

5. Data storage and backup during the project

5.1. Where will the data be stored?

- **Non-personal data** will be stored on the KU Leuven **secured network discs** and the **OneDrive accounts** of the involved researchers (backed up in the cloud) and on the Microsoft Teams pages of this project (backed up in the cloud). A time-stamped master copy will always be kept on the KU Leuven's cloud storage systems. Copies can be made and kept on personal devices and/or external hard-drives.

Depending on the processing power and storage needed during the research project, large geospatial data could also be stored on the HPC VSC - High Performance Computing from the Vlaams Supercomputer Centrum – large volume storage (automatic backup).

- **Pseudonymized data:**
 - The raw weather station data of the Leuven.cool network (data 3.2) is currently stored in a secured phpMyAdmin MySQL database (leuven.cool / localhost | phpMyAdmin 5.1.0). For the acquisition of the weather station data, we make use of a secured MySQL database that is hosted on a virtual private server (VPS) hired for the project. We use a cloud VPS from a private company based in The Netherlands. This VPS is used for both the reception and the storage of the weather stations data. There are no personal data of the participants stored on this VPS.
 - Secure storage of the sensor data of soil moisture/temperature (data 3.5) will technically be supported by Curieuzeneuzen in de tuin and the partnership of Orange.
 - Data storage of other sensor data, i.e. air quality (data 3.6), has not yet been defined, but will follow a parallel trajectory as the weather station data. This will technically be supported by sensor and communication network experts at ESAT, KU Leuven.
 - Survey data (data 1.20, 5.3, 5.4, 5.8) will be stored in a secured PostgreSQL database, hosted on the Spatial Application Division Leuven (SADL) server (ICTS-S-DI-SADL1.lunah.kuleuven.be).
 - Interview transcripts (5.6) will be stored at a secured network drive at KU Leuven.
- **Personal data** will be stored at a secured network drive at KU Leuven (data 1.20) and a secured PostgreSQL (data 5.3, 5.4, 5.8). Personal data of the sensor networks (data 3.2, 3.5 and 3.6) are stored on a drive where only the main researchers have access to.

5.2. How will the data be backed up?

KU Leuven Onedrive and the secured network drives have an integrated back-up option.

An offline backup of the weather data stored on the VPS is currently performed on a regular basis.

5.3. Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

Yes.

- Capacity for non-personal data:
 - KU Leuven OneDrive and KU Leuven Teams: 2 TB/person (this can if necessary be raised to 5 TB without additional costs)
- Capacity for personal data
 - KU Leuven secured network disc: 1 TB (FNL) + 1 TB (IO)

5.4. What are the expected costs for data storage and backup during the project? How will these costs be covered?

KU Leuven provides every student and employee a standard OneDrive storage capacity of 2 TB, this can if necessary be extended to 5 TB without any additional costs.

The cost of data storage on the servers of the division is estimated at €2000. This cost is foreseen in the project budget.

The rental costs of the VPS (and other sensor network services) are covered by the operating costs of the department and the project.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Non-personal data: OneDrive and Teams are an enterprise cloud storage service with centralized security controls and multi-layered encryption, authorized by KU Leuven (KU Leuven Authenticator tool).

Personal data: Personal information is stored separately from the survey and sensor data. An pseudonymized ID is used in the research data. At the database level, the researcher will not be able to identify the respondent.

Where possible, access to part of the survey data (data 5.3, 5.4, 5.8) is further refined and secured by an intermediate layer that further shields the actual access to the database. Except for the necessary database administration, the data is only available via this intermediate layer. Researchers of the project do have access to personal data via a specific account where this is necessary for the research, in addition to the database administration already mentioned (for example if someone needs to be contacted in the context of the research).

Other personal data (data 1.20, 3.2, 3.5, 3.6) are located on secured drive where only the researchers as defined in the SMEC application have access to.

6. Data preservation after the end of the project

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

All relevant data for the project will be retained for 10 years after the end of the project, in line with KU Leuven RDM-policy. This includes data 1.19-1.21; 2.2-2.10; 3.2, 3.5, 3.6, 3.9, 3.11; 4.2-4.7; 6.2-6.6; 1.19-1.21; 5.3-5.8. Necessary documentation and metadata files that will help others understand the data and make it fully reusable: interview guidelines, processing methods, etc. will also be retained.

Secondary data sources will not be retained. The services of these data sources are considered responsible for their own data archive. Furthermore, depending on the 3rd party agreements we are also not permitted to retain some datasets (cf. 3.3). In order to reuse the project outputs, we will make a step-by-step plan which includes information about the data provider, the general form of the data (i.e. column names and data type) of the secondary input data. Potential other users can consequently contact the responsible data provider.

After the project, personal data will be retained in accordance with the informed consent of the respondents.

6.2. Where will these data be archived (= stored for the long term)?

Data will be archived on the KU Leuven network drives.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

The expected cost for preserving this data is estimated at €2000. This cost will be covered by the working budget of PI Ben Somers.

7. Data sharing and re-use

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

Yes.

The personal data will not be shared in accordance with the informed consent of the respondents.

The official meteorological data provided by the RMI (data 3.3) will not be retained after the end of the project as we only have the right to use it and not share it.

Depending on the 3rd party data agreements (*in progress*) for datasets 3.8 and 3.10, restrictions on data sharing will be taken into account.

7.2. Which data will be made available after the end of the project?

All created output on the climate mitigation and adaptation potential of the garden complex (as defined in the GARLOCK project – data 2.2-2.10; 4.2-4.7; 6.2-6.6) will be made available. In addition all primary material (including anonymized personal data) to support the outputs (data 1.19-1.21; 3.2, 3.5, 3.6, 3.9, 3.11; 5.3-5.8) will be made available. Lastly, we will also include all necessary documentation and metadata files that will help others understand the data and make it fully reusable: interview guidelines, processing methods, etc.

7.3. Where/how will the data be made available for reuse?

- In an Open Access repository
- Upon request by mail

The datasets, scripts and workflows collected by the project (not including personal or pseudonymized data) will be uploaded in open-access repositories such as KU Leuven Research Data Repository (RDR) ([Research Data Repository \(RDR\) – RDR - Research Data Repository \(kuleuven.be\)](https://kuleuven.be/rdr)) under a CC-BY license.

7.4. When will the data be made available?

- Upon publication of the research results

7.5. Who will be able to access the data and under what conditions?

Personal data: only the researchers mentioned in the approved SMEC file will be able to access the personal data. The pseudonymized data will be stored in a secure drive and PostgreSQL database at the server of the division, where also only a limited group of people have access too.

Other data: these data will be made available to anyone for any purpose, provided that they give appropriate credit to the creators (CC BY license).

7.6. What are the expected costs for data sharing? How will these costs be covered?

There are no expected costs related to data sharing as RDR is free and open.

8. Responsibilities

8.1. Who will be responsible for the data documentation & metadata?

WP1: Kelly Wittemans & Janne Teerlinck

WP2: Kelly Wittemans

WP3: Valerie Dewaelheyns & Janne Teerlinck

8.2. Who will be responsible for data storage & back up during the project?

WP1: Kelly Wittemans & Janne Teerlinck

WP2: Kelly Wittemans

WP3: Valerie Dewaelheyns & Janne Teerlinck

8.3. Who will be responsible for ensuring data preservation and sharing?

WP1: Kelly Wittemans & Janne Teerlinck

WP2: Kelly Wittemans

WP3: Valerie Dewaelheyns & Janne Teerlinck

8.4. Who bears the end responsibility for updating & implementing this DMP?

The PI bears the end responsibility of updating and implementing this DMP. We plan a recurrent update of the DMP.

