

Solutions for outdoor climate adaptation - SOLOCLIM

Open Science Strategy / Data Management Plan

DATE 31 March 2020

GRANT AGREEMENT NUMBER	861119
WORK PACKAGE NUMBER	WP5 - Dissemination and communication; Month 1-48
DELIVERABLE NUMBER	D5.2
DUE DATE	Month 6 - 31 March 2020
LEAD BENEFICIARY	Project Coordinator - Wageningen University, the Netherlands
CONTACT PERSONS	Sanda Lenzholzer; Ivana Vujic

Objective

Open communication and discussion sit at the heart of scientific practice. As most of the growth of scientific understanding in recent centuries is due to open practices, this document tends to elaborate on open science (open access) strategy of SOLOCLIM project in order to contribute to global, transparent and reproducible scientific progress and innovation. Furthermore, this document will outline the data management plan of SOLOCLIM project answering how (research) data will be collected, processed or generated within the project, what methodology and standards will be adopted in this process, how this data will be shared and/or made open and how it will be curated and preserved during and after the project's life time.

This document may be updated as the policy evolves.

Project summary

Solutions for Outdoor Climate Adaptation (SOLOCLIM) is a European Industrial Doctorate project in the programme Innovative Training Networks and part of Marie Skłodowska-Curie Actions funded by the European Commission within the Horizon 2020 programme. The aim of SOLOCLIM is to develop a doctoral training programme that enables young researchers to generate solutions for urban outdoor environments.

Some solutions are known such as the use of vegetation, but there are still many remaining questions about the impact of vertical green as well as the proper distribution of green in cities to have an optimal effect. Moreover, upcoming systems using water vapour as a coolant as well as flexible systems that respond to microclimate have not been studied yet.

SOLOCLIM will develop solutions on different scale levels from small scale around buildings to a larger neighbourhood/city scale and test their effects. The industry in architecture, urban and landscape design is in need of these innovative solutions as climate adaptation in cities is one of the largest challenges for the future. All plans and designs that the industry develops have to respond to the urban (micro)climate challenges. Apart from the availability of solutions the industry also needs the expertise. This expertise involves design and research skills related to urban (micro)climate.

SOLOCLIM will train PhD students as well as some parts to a broader community beyond the duration of the project.

Official website with more information: www.soloclim.eu.

Contents

Objective	1
Project summary	1
1. Open Access Strategy	3
2. Data Summary	3
2.1 Purpose of the data collection/generation	3
2.2 Types of data generated/collected	4
2.3 File formats of data generated/collected	4
2.4 Origin of the data, size of data and data utility	4
2.5 Making data findable, including provisions for metadata	5
<i>Online dataset</i>	5
<i>Naming convention</i>	5
<i>Search keywords, data version numbers and metadata created</i>	5
2.6 Making data openly accessible	5
<i>Type of data made openly available as the default</i>	5
<i>Data accessibility and restrictions</i>	6
<i>Software</i>	6
<i>Certified repositories</i>	6
2.7 Making data interoperable	7
<i>Interoperable data, data exchange and re-use</i>	7
<i>Metadata vocabularies, standards or methodologies</i>	7
<i>Inter-disciplinary interoperability</i>	7
2.8 Increase data re-use	7
<i>Data licence for wider re-use and data re-use time period</i>	7
3. Allocation of resources	7
3.1 Costs and costs coverage	7
3.2 Data management responsibilities	8
3.3 Resources for long-term preservation	8
4. Data security	8
4.1 Data security provisions	8
4.5 Data safety	9
5. Ethical aspects	9
Annex I	10
Abbreviations list	11

1. Open Access Strategy

All SOLOCLIM peer-reviewed scientific publications will be deposited in repositories that are registered in the directory of authoritative academic Open Access repositories (Open DOAR), will be used to deposit publications as a machine-readable electronic copy at the latest upon publication by the publisher. To comply with the H2020 open access mandate, the articles published both via ‘gold’ and ‘green’ open access, will be deposited in institutional repositories (e.g. Research@WUR or KDR) and become open access not later than 6 months after publication by the publisher (12 months for publications in Social Sciences). Where possible, the institutional repositories will be providing information to continuous reporting of the project in OpenAIRE (the authoritative source to the European Commission for H2020 projects; once material is deposited, OpenAIRE facilitates the posting of publications and which automatically appear in the [EC Research Participant Portal](#) - SygMa). Those project partners that do not have an Open DOAR (Polimi, UKent) institutional repository, will instead make use of Zenodo for datasets and RE.PUBLIC for publications (Polimi) and, KDR and KAR (UKent) for Open Access publishing. Beside peer-reviewed publications, Open Access will be provided to other types of scientific output of the project results such as monographs, books, conference proceedings and reports. Open Access will also be provided to the bibliographic metadata that identify the deposited publication, according to the standard format from the H2020 Guidelines to participants. If the underlying data to a publication are available, they will be deposited in a specific data repository. The publication will include a DOI link to the location of these data.

Horizon 2020 steps to Open Access:

- *Submit papers to a journal of choice, there is no restriction provided you can comply with the policy. If a researcher chooses to publish in an Open Access journal, APC costs should be claimed within the project period and budget.*
- *Deposit the peer-reviewed author accepted manuscript or publisher’s PDF in a repository, compliant to the grant regulations and take care of the embargo period (6 months or 12 months for social sciences).*
- *Acknowledge project funding both in the publication and in the metadata in the format: “This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 86119.”*
- *Make sure you add the publication to your final report to the EC.*

2. Data Summary

2.1 Purpose of the data collection/generation

Given SOLOCLIM’s integrated design with industry actors and possibly confidential, commercially sensitive data, this document pays attention to managing data consistent with the project’s Grant Agreement and Consortium Agreement.

The document is developed translating the generic H2020 requirements and recommendations into specific guidelines and advice that can be applied in the SOLOCLIM project. The research data will be stored in three central repositories (ESR1 and ESR2: WUR, ESR3 and ESR4: Polimi, ESR5 and ESR6: UKent) according to the FAIR principles: Findable, Accessible, Interoperable and Reliable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or implementation-solution.

More information about FAIR: [Guidelines on FAIR Data Management in Horizon 2020](#), [FAIR data principles \(FORCE11 discussion forum\)](#), [FAIR principles \(article in Nature\)](#).

2.2 Types of data generated/collected

SOLOCLIM anticipates generating a range of data:

- textual,
- numerical,
- visual, and
- design prototypes.

Examples of data generated include:

- microclimate measurement data,
- energy data,
- simulation result data,
- questionnaire results, and/or
- interview transcripts.

2.3 File formats of data generated/collected

File formats to be used for textual data are:

- DOC,
- PDF, and
- TXT.

File formats to be used for visual data are:

- TIFF, and
- PNG.

File formats to be used for tabular data are CSV.

And finally, file formats to be used for audio data and video data are:

- MP3, and
- MP4.

2.4 Origin of the data, size of data and data utility

The origin of the data will differ. Some of the data comes from existing data sets (e.g. geo-data from Dutch 'Kadaster', Centraal Bureau voor de Statistiek - statistical data, such as publicly available demographic data, openly accessible data like AutoCAD). On the other hand, some data will be generated by ESRs and project-related researchers (such as microclimate measurements, etc.).

During the project, the existing and generated data will be useful for the ESRs and their research projects and the SOLOCLIM beneficiaries, i.e. universities and industry firms. After completion of the project, potential re-users of the data will be other researchers while the results may also be useful to companies.

The research 'know-how' and development of design guidelines will also be exploited by the companies, which will help them to strengthen their market position developing further expertise and provide new market niches in the world-wide knowledge economy. The guidelines will also benefit the wider professions, through active engagement with the relevant professional associations to disseminate the application to their members. Representatives from municipalities and local government will benefit from this knowledge, essential contextualization for climate change adaptation and mitigation measures and evaluating proposals for new developments.

The knowledge developed across the different disciplines will greatly impact educational provision at all cycles. Initially, it will be through participating academic staff making proposals for new modules, courses and programmes within their own institutions. The material will also provide CPD opportunities for professionals

across Europe, informing them of the latest developments in the related disciplines with potential implementation paths in their work across different sectors.

2.5 Making data findable, including provisions for metadata

Upon submitting to repositories, a dataset will receive a DOI (persistent identifier) as well as title, author, data, associated files, keywords, etc. The project will use DANS-EASY (WUR), Zenodo for datasets and RE.PUBLIC for publications (Polimi) and KDR and KAR (UKent) as repositories. The default metadata scheme used will be [Dublin Core](#).

Online dataset

SOLOCLIM will establish an online database not only with the results of the research, but also the dataset(s) collected from the various measurements campaigns, which will be useful to researchers involved in wider research beyond the duration of the project stored in the certified repositories of the three universities (WUR: DANS-EASY, Polimi: Zenodo for datasets and RE.PUBLIC for publications, UKent: KDR and KAR). In this way we ensure that the long-term data management can be guaranteed. Of course, not all datasets can be placed in a repository. If there are contractual restrictions, commercial interests, personal data, etc., there are all reasons to withhold the sharing of the datasets.

Naming convention

For the project's needs, a file naming and versioning protocol will be set up to ensure that files are named and versioned following the same standard by all participating members as follows:

YEAR_MONTH_DAY_MAIN TOPIC_SUBTOPIC_VERSION NUMBER

This protocol will be described in a readme.txt document that will be placed in dataset folder.

Search keywords, data version numbers and metadata created

By placing (some of) data in a repository we will be required to add keywords, similar to a publication. In fact, we will use the same keywords as the publication. Metadata is added during submission in a repository but also added within a dataset by creating a readme.txt file.

2.6 Making data openly accessible

Type of data made openly available as the default

In particular, SOLOCLIM aims to publish at least 36 ESR-authored papers for peer-reviewed journals and scientific conferences on specific aspects of climate responsive design, and 6 PhD theses. Access to the full thesis will be public by default, except if ESRs wish to restrict access in order to protect copyright. The data raised by the ESRs (e.g. measurement data related to prototypes) is not data that should be publicly available due to IPR in some cases. All other SOLOCLIM data that are not subject to IPR will be openly accessible for use in other research projects or for meta-analysis but even these would sometimes not be of a broad interest for other researchers. Thus, specific microclimatic measurement results from development of new prototypes (e.g. green walls) can only be made public partly, depending on IPR.

TYPE OF DATA / OUTPUTS	OPEN?	COMMENT
Textual (e.g. papers, thesis, interview transcripts, questionnaire results, etc.)	YES	except in the case if ESRs wish to restrict access in order to protect copyright and/or to comply with Creative Common licence

Numerical (e.g. microclimate measurement data, energy data, simulation result data, etc.)	YES / NO (depends)	data generated by the ESRs is not data that should be publicly available due to IPR in some cases. specific microclimatic measurement results from development of new prototypes can only be made public partly, depending on IPR.
Visual (e.g. graphs, plans, maps, renders, posters, infographics, etc.)	YES / NO (depends)	depends on the IPR situation.
Design prototype (e.g. models, simulations, etc.)	YES / NO (depends)	depends on the IPR situation. except in the case if ESRs wish to restrict access in order to protect copyright and/or to comply with Creative Common licence

Data accessibility and restrictions

Datasets produced during this project will be deposited in DANS-EASY (WUR), Zenodo for datasets and RE.PUBLIC for publications (Polimi) and KDR and KAR (UKent) according to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. As an exception, the beneficiaries do not have to ensure Open Access to specific parts of their research data if the achievement of the action's main objective (as described in Annex 1 of the Grant Agreement) would be jeopardised by making those specific parts of the research data openly accessible. An embargo period of max two years will be applied to dataset to first allow publishing of the results.

Software

The SOLOCLIM project does not anticipate the need for any type of certain software tool or method to access the data. The data can be accessed via the online interface of the chosen data repository.

Certified repositories

The dataset(s) will be shared through a data repository that has been certified and that guarantees data preservation for a minimum period of 10 years. The repository should furthermore provide a persistent identifier such as a DOI to the dataset, for the dataset to be easily found and cited.

The library at WUR is front office for the Dutch national archive, the CoreTrustSeal certified DANS-EASY (see <https://easy.dans.knaw.nl/ui/home>). As such, the library will submit the datasets and cover the costs of submission. Datasets underlying a manuscript that is not yet published will be placed under embargo until the associated manuscript is published. There is no need for a data access committee. The datasets in DANs-EASY will be assigned the licence CC-BY, which is a machine-readable licence. Data in DANs-EASY can be made available only for registered users. It can also be made 'restricted access' which requires permission from the data depositor before access is granted.

The library at Polimi are not involved at any level in the management of research data which remains entirely in the hands of the project group and researchers. In terms of costs, these fall within the project budget and are not centrally supported by the university. For the license associated with the dataset, Zenodo and RE.PUBLIC accepts all types. The datasets will also be assigned the licence CC-BY.

UKent uses KDR, that is the University's institutional data repository. It serves as catalogues for data created there but archived elsewhere. It's there for researchers if a specialist archive is not available. This repository accepts all types of research data, defined as "any recorded information that supports or validates research observations,

findings or outputs.” Storage of research data on KDR is part of the University of Kent standard service and is free. Each data record has citation information and unique identifiers, validated by a central body (DOI).

2.7 Making data interoperable

Interoperable data, data exchange and re-use

As much as possible, standardized methodologies and data formats will be used to facilitate data transfer and reuse among project partners. Ultimately the data that will be made openly available will be in open formats.

Metadata vocabularies, standards or methodologies

A discipline-specific metadata standard is not commonly used in this field of research (at the judgement of the researcher). We are not aware of standard metadata vocabularies within the field. Standard methodologies will be used for data collection and analysis.

Inter-disciplinary interoperability

We are not aware of standard metadata vocabularies between disciplines for the data types that we will produce. In case it is unavoidable that we use uncommon or generate project specific ontologies or vocabularies, we will discuss among project partners the possibility of providing mappings to more commonly used ontologies.

2.8 Increase data re-use

The main areas for the exploitation of the SOLOCLIM research results are commercial application, professional guidelines and educational provision.

All parties involved in the consortium have a business model strongly based on research and innovation. microclimatic adaptation at different scales is a key topic for all the parties and SOLOCLIM offers a valuable opportunity to strengthen and expand this research with particular application in urban environments.

Some of the research outputs have great potential to be fully commercialized, such as green facade systems, new water and novel responsive systems. Several WPs will thus result in product development and after the first prototypes and tests have been completed, the relevant partners will be able to explore options for filing for patent. Along with the technology transfer departments of the universities leading the respective WPs the beneficiaries will assess further steps in product exploitation.

Data licence for wider re-use and data re-use time period

Data will be licensed under CC-BY, which allows free reuse and adaptation of data on condition of attribution. Data will be available for reuse for a minimum of 10 years.

3. Allocation of resources

3.1 Costs and costs coverage

At the WUR and Polimi, one can quantify the costs of storing data on the network drive as follows: For 2020 it costs 0.80 euro per GB per year. The costs for submitting data to DANS-EASY are covered by the WUR library. The costs fall within those of the project budget. The costs for data management will be estimate considering backup costs, data preparation, repository costs, legal advice, etc. At UKent, storage of research data on KDR is part of the University of Kent standard service and is free.

3.2 Data management responsibilities

The document is based in the local regulations as well as the supra-national guidelines established by the European Commission regarding data protection, but also to facilitate open data. Data management activities concern the whole project and will be coordinated and monitored both at project and work package level. Data management is also linked to publication of project results and thus dissemination activities.

The responsibility for drafting and maintaining the Data Management Plan lies with the WP6 lead and each beneficiary will confirm that they have appointed a Data Protection Officer and the contact details will be made available to all data subjects involved in the research.

The Work Package leaders are responsible for:

- the implementation of the data management policy in their respective WPs;
- monitoring data management activities and deadlines and sending reminders to partners;
- offering customized help and further guidance for filling out the WP data surveys;
- asking partners for missing information or clarifications;
- providing input to the data management plan by analysing and summarising the WP-specific data surveys;
- offering customized help and further guidance for publishing open data and documents;
- monitoring that open results (data and software) are deposited in the default repository or a complementary OpenAIRE-compliant repository and sending reminders to partners;
- contacting the project management in case of questions and ethical and privacy issues that may forbid a publication of the data.

3.3 Resources for long-term preservation

Long term data preservation will be carried out by chosen repositories. The project coordinators decide on which data will be preserved. The general research data policy requires that all data underlying a publication should be archived for a minimum of 10 years.

4. Data security

4.1 Data security provisions

SOLOCLIM aims to provide:

- confidentiality that will assure ethical and legal treatment of data. All researchers must obtain from the subjects of study informed consent prior performing the research itself, being it an interview, questionnaire, focal group, etc. (model in Annex I). Once the research is performed, researchers must anonymize the outputs, i.e. in interview transcripts should never be mentioned any personal data that might led to the identification of the interviewee.
- archiving that will assure that the data, once confidential, can be obtained by anyone willing to use them for secondary or contrasting research.

With regards to managing personal data, SOLOCLIM will treat the data as confidential and will take every precaution to guarantee the privacy to participants, i.e., ensuring that personal data will be appropriately anonymised and be made inaccessible to third parties.

Any direct or indirect personal data that is collected will be anonymized as soon as possible after collection. Prior to anonymization, data will only be stored on an institutional network drive with restricted access and will comply with institutional policies on working with personal data. Along with these data, documentation on the data will be kept in README-files. All data will be stored on the institutional network drives with access restrictions and daily back-ups, and cross-institutional sharing will only be done using institutionally

approved cloud storage. To minimise the risk of data loss or corruption, external devices such as hard disks or USB drives will not be used. All ESRs will use the same methods for collecting and storing data to allow easy retrieval, sharing and further processing, compliant with the GDPR.

4.5 Data safety

The repositories of choice, DANS-EASY (WUR), Zenodo for datasets and RE.PUBLIC for publications (Polimi) and KDR and KAR (UKent) are certified repositories that carry out long-term preservation and curation.

5. Ethical aspects

SOLOCLIM will perform research in four different countries (but not limited to): the Netherlands, the United Kingdom, Italy and Germany. The Consortium is aware of the national and EU legislation, guides and codes that regulate management of data.

SOLOCLIM project uses textual, numerical, visual, and design prototypes data from microclimate measurements, energy, simulations, questionnaires and interviews, thus, the risks of ethical violations are low.

Participants of the project will conduct all work respecting the principles of the Charter of Fundamental Rights of the European Union, which covers issues of dignity, freedom, equality, solidarity, citizens' rights and justice.

In particular, researchers will aim to consider the sensitive implications of their proposals in terms of respecting privacy, inclusiveness and autonomy. The project and the proposed work will ensure that all material and data will be used with the utmost confidentiality and dignity, and thus no risk, nor harm to third parties is caused. Researchers will aim to comply with national and European Union legislation, respect international conventions and declarations and take into account the opinions of the European Group on ethics.

Informed consent will be sought whenever ICT research involves volunteers in interviews and the project will ensure that subjects have the information they need, specifying the alternatives, risks, and benefits for those involved in ways that they understand. None of the methodologies and technologies intended to be used are known to inflict any psychological damage on participants.

Additionally, through the outreach activities to secondary education SOLOCLIM can showcase new exciting cross-disciplinary pathways and combat negative stereotypes for women. As the leading investigators at the three universities are all women, the project can provide a role model to inspire young girls.

Annex I

Informed Consent Form¹

SOLOCLIM state of the art analysis will rely on interviews with people involved in the project. Summaries of the interviews will be made available in an online database with open access. For more information, see: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf

I, (first name + surname), hereby certify that I accept to participate to the SOLOCLIM project through one (or several) interview(s) with one (or several) researcher(s) from partner organisation.....

1. I accept that my name and my position inside (or outside) the organization under study does not appear in the SOLOCLIM deliverables and to have my personal data referred to as 'anonymous'.
2. I am aware of the nature and objectives of the SOLOCLIM project and have been informed of what is expected from me. I was given a copy of this consent form, preceded by a summary of the research objectives.
3. I am aware that the content of the interview(s) is subject to the researcher's analysis and interpretations; and that the data files contain the name and affiliation of the researcher(s) who interviewed me.
4. The interview(s) will be summarized, I accept that some parts will be made public through quotations.
5. I acknowledge that I approve the summaries based on the one (or several) interview(s) before they are made public in the open access database. I was made aware of the content of the summaries.
6. I accept that the data might be further processed for scientific purposes, in line with the (Name of Country) law of the Date on the protection of privacy regarding the processing of personal data.
7. I accept that the results of the study will be disseminated for scientific purposes, in compliance with the ethical standards of the scientific community.

Location:

Date:

Signature:

¹ The provision of the Form may vary in certain details in different beneficiaries' countries, but most preferential arrangements have common privacy statements concerning how participants' data is going to be anonymised, used, where it will be stored, for how long, etc.

Abbreviations list

APC - Article processing charge

CC-BY - Creative Commons by attribution

CDP - Continuing Professional Development

CSV - comma-separated values

DOI - Digital Object Identifier

EC - European Commission

ESR - Early Stage Researcher

EU - European Union

FAIR - Findable, Accessible, Interoperable, Reliable

GB - Giga bytes

GDPR - General Data Protection Regulation

H2020 - Horizon 2020

ICT - Information and communications technology

IPR - Intellectual property rights

KAR - Kent Academic Repository

KDR - Kent Data Repository

PDF - Portable document format

Polimi - Politecnico di Milano

SOLOCLIM - Solutions for outdoor climate adaptation

UKent - University of Kent

USB - Universal Serial Bus

WP - Work Package

WUR - Wageningen University and Research

All rights reserved

This document may not be copied, reproduced or modified in whole or in part for any purpose without the written permission from the SOLOCIM Consortium. In addition to such written permission to copy, reproduce or modify this document in whole or part, an acknowledgement of the authors of the document and all applicable portions of the copyright must be clearly referenced.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 86119.