SOPHIA

Sustainable Off-grid solutions for Pharmacies and Hospitals In Africa

DATA MANAGEMENT PLAN

DELIVERABLE D9.3

VERSION 1.0

PROJECT ACRONYM	SOPHIA
PROJECT TITLE	SUSTAINABLE OFF-GRID SOLUTIONS FOR HOSPITALS AND PHARMACIES IN AFRICA
FUNDED UNDER	EU H2020 – grant agreement No. 101036836
PROJECT URL	http://www.sophia4africa.eu/
COORDINATOR	UNIVERSITY OF APPLIED SCIENCES KARLSRUHE
PROJECT DURATION	01/10/2021 – 30/09/2025 (48 MONTHS)

Funded by the European Union's Horizon 2020 research and innovation programme under grant agreement N. 101036836. The sole responsibility for the content of this paper lies with the authors. It does not necessarily reflect the opinion of the European Commission (EC). The EC is not responsible for any use that may be made of the information it contains.

SOPHIA CONSORTIUM

NO.	BENEFICIARY	ADDRESS	MAIN CONTACT
1.		**HOCHSCHULE KARLSRUHE (HKA)** Moltkestr. 30, DE-76133 Karlsruhe, Germany	**Coordinator:** Prof. Dr.-Ing. habil. Michael Kauffeld michael.kauffeld@h-ka.de
2.		**OST - OSTSCHWEIZER FACHHOCHSCHULE (SPF-OST)** Oberseestrasse 10, CH-8640 Rapperswil, Switzerland	Dr. Mihaela Dudita mihaela.dudita@ost.ch
3.		**MAKERERE UNIVERSITY (MAK)** Main Campus, PO Box 7072, 256 Kampala, Uganda	Prof. Dr. Nicholas Kiggundu kiggundu@caes.mak.ac.ug
4.		**FONDATION 2IE ASSOCIATION (2IE)** Rue de la Science, PO Box 01BP594, 01 Ouagadougou, Burkina Faso	Dr.-Ing. habil. Kokouvi Edem N'Tsoukpoe edem.ntsoukpoe@2ie-edu.org
5.		**STEINBEIS INNOVATION GGMBH (SEZ)** Steinhäuserstr. 12, DE-76135 Karlsruhe, Germany	Dr. Anthony Salingre anthony.salingre@steinbeis-europa.de
6.		**MINISTERE DE LA SANTE PUBLIQUE (MPHC)** Boulevard Rudolph Manga Bell Yaounde, Cameroon	Ms. Jeanne Aurélie Abomo Ayinda aureliejeannea@yahoo.fr
7.		**INSTITUT INTERNATIONAL DU FROID (IIR)** Boulevard Malesherbes 177, FR-75017 Paris, France	Ms. Ina Colombo i.colombo@iifiir.org
8.		**OPERIEREN IN AFRIKA EV (OiA)** Wirthstrasse 11, DE-79110 Freiburg, Germany	Prof. Bernhard Rumstadt b.rumstadt@gmail.com
9.		**EVERFLO(PTY)LTD (EVERFLO)** 10 Mclaren Way Racing Park, 7441 Milnerton, South Africa	Mr. Evert Potgieter evert@everflo.co.za
10.		**KOVCO PTY LTD (KOVCO)** Century City, PO Box 134, 7446 Century City, South Africa	Mr. Brendan Prestage bprestage@kovco.co.za
11.		**MARTIN SYSTEMS GMBH (MS)** Address: Friedrichstrasse 95, DE-10117 Berlin, Germany	Dr.-Ing. Jose Ordonez jose-ordonez@martin-systems.com
12.		**SIMPLY SOLAR SCHEFFLER TECHNOLOGIES AND CONSULTING (SISO)** Graf von Werdenbergstrasse 6, DE-89344 Aislingen, Germany	Ms. Heike Hoedt h.hoedt@simply-solar.de
13.		**RAACH SOLAR GMBH (RS)** Stellwinkel 1, DE-88451 Dettlingen an der Iller, Germany	Mr. Jürgen Raach juergen.raach@raachsolar.com

DOCUMENT INFORMATION

RESPONSIBLE PARTNER	HKA
AUTHOR (S)	Irene Robles García (SEZ) Sophie von Stralendorff (SEZ) Oliver Schmid (HKA)
REVIEW	Mihaela Dudita (OST-SPF)
APPROVAL	Oliver Schmid (HKA)
KEYWORDS	Data management, FAIR data, open science
DISSEMINATION LEVEL [1]	PU
DUE DATE OF DELIVERABLE	30/04/2022
ACTUAL SUBMISSION DATE	03/05/2022

PU = Public
PP = Restricted to other program participants (including the Commission Services)
RE = Restricted to a group specified by the consortium (including the Commission Services)
CO = Confidential, only for members of the consortium (including the Commission Services)*

DOCUMENT HISTORY

DATE	VERSION	CHANGE HISTORY	AUTHOR
03/05/2022	1.0	First version submitted to EC	Oliver Schmid (HKA)

TABLE OF CONTENTS

ABBREVIATIONS

DMP	Data Management Plan
DOI	Digital Object Indentifier
EC	European Commission
EMCA	Environmental Management and Coordination Act
FAIR	Findable, accessible, interoperable, re-usable
GB	Gigabyte
GLP	Good Laboratory Practice
GMP	Good Manufacturing Practice
H2020	Horizon 2020
IP	Intellectual Property
IPR	Intellectual Property Rights
ISO	International Organization for Standardization
LCA	Life Cycle Assessment
NEA	National Environment Act
NEMA	National Environment Management Authority
PV	Photovoltaic
PVGIS	Photovoltaic Graphical Information System
WP	Work Package

EXECUTIVE SUMMARY

This report was made in the frame of Work Package 9 “Project Management”, led by HKA and presents the Data Management Plan (DMP), relating to the management of data generated and used during the SophiA project.

As part of making research data findable, accessible, interoperable and re-usable (FAIR), this DMP includes information on:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access and
- how data will be curated and preserved (including after the end of the project).

This first version of the DMP will serve as a basis to define the processes in place by each partner in respect to data management. The strategy of data management at project level is also described. Two updates of this deliverable are expected throughout the project, in M24 and M36.

A DMP questionnaire was sent to all project partners in order to collect all necessary information about their individual data management strategy. A summary of the results received is presented in this report, collected and presented at project level. The questionnaire template is included in Annex 1. The individually filled questionnaires have been stored in the project’s repository and can be made available to the European Commission upon request.

1. GENERAL INFORMATION REGARDING DATA MANAGEMENT IN SOPHIA

Among other, the SophiA project will generate the following data:

- Chemical-physical characterization, safety and environmental data for the cooling materials;
- Cost data for the production of the container;
- Market research data about the needs for SophiA services;
- Data on treatment efficiency and further water quality parameters;
- Energy consumption data of SophiA solution and possible up-scaling;
- Weather data of the locations for the SophiA demo-containers;
- Quantitative data on hospitals and remote medical centers;
- SophiA systems performance in trial tests;
- Non-personalized socio-economic data on gender/women regarding involvement in academia and economics.

In order to manage these data, a secured access knowledge database will be established and updated with the knowledge generated, used and disseminated in the course of the project.

From the consortium's knowledge, the following standards will be applied:

- GLP and GMP standard
- ISO 9001: Quality management systems – Requirements (2008)
- Environmental Management and Coordination Act (EMCA) (Section 58 to 69)
- Bureau of Standards (e.g KEBS) focusing on ensuring quality standards
- Ethical bylaws within the 4 countries for the trial tests of SophiA systems
- The National Environment Management Authority (NEMA) Uganda rules and regulations
- National Environment Act (NEA), Uganda, 1995
- The Uganda National Wetlands Policy, 1995
- EN 378 Refrigerating systems and heat pumps - Safety and environmental requirements and national standards

The measurement raw-data, calculations and evaluations based on these standards as well as project's research will be documented in reports and appropriate electronic files to be stored on the project's and on partners' knowledge bases. These will be shared with the consortium partners for verification and used in the scientific and technical publications (except for partners-owned intellectual property, which must be protected).

2. DATA SUMMARY

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

The purpose of the data collection differs among the partners depending on their role and specific work in the project. Data is mainly needed for the:

- SophiA systems design
- selection of the demonstration sites
- social and environmental impact assessment
- communication and dissemination of the project results
- development of the exploitation plan for the key exploitable project results
- capitalisation of experiences for future research actions
- management of IPR inside the consortium
- creation of a stakeholder community
- general project management

The collected data will be used for research, design and improvement of the systems installed in the containers, as well as technical and manufacturing record-keeping to provide an effective after sale service. Data on hospital facilities serve as important criteria to be able to select suitable sites for the demonstration of SophiA technologies and to adapt/optimize the SophiA technologies to the actual requirements at the chosen demo sites. Feedback from stakeholders, the general public and public organisations including ministries will be gathered during meetings, workshops and dedicated interviews to be able to assess the social acceptance and environmental impact of the SophiA systems.

Data collection is furthermore necessary for the successful communication and dissemination of the project results, the development of an exploitation plan and IPR management inside the consortium, as well as the creation of a stakeholder community and the capitalisation of experiences for future research actions. Finally, data collection is also required for general project management and reporting activities.

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

According to the different data's purpose the data types range from technical parameters (e.g. pressure, temperature, velocity, water flowrates, climate), monitoring data of pilot sites, technical drawings and simulation results to more general data types like name and contact data, publications/reports, presentations, photos/videos.

The main formats used are Excel, Word, PowerPoint, pdf, dwg, step, csv, image and video formats.

Will you re-use any existing data and how?

Already existing data is intended to be reused to facilitate the project activities and disseminated further via publications, outreach activities or teaching material where applicable.

On the technical side, this is mostly the case for previous research results in the field of cascade refrigeration cycles, thermal energy storage and solar/climatic data collected from weather stations, databases and open sources.

For the general project management, proven structures and templates of previous projects will be adapted to the SophiA project. This is especially the case of for instance IPR workshops, that will be modified to fulfil the partners needs in terms of IP management and exploitations in SophiA. Internal templates from other projects for the collection of data regarding project management, dissemination and communication activities will be adapted and used in the project.

What is the origin of the data?

The data is on the one hand directly produced and gathered via system tests on laboratory scale, questionnaires, assessment tools, the LCA data collection tool, partner's own drawings/designs as well as external climatic data. On the other hand, data originates from previous research projects, where partners have been involved with similar responsibilities and services like Metonorm, PV-GIS and POWER NASA.

What is the expected size of the data?

The expected size of the collected data will represent at the end of the project several gigabits, probably ranging around 1-3GB for most of the partners.

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

The raw data generated will first and foremost be useful and important to the SophiA consortium to process and assure the smooth running of project activities. Once published in scientific articles, the data will also be useful for the broader scientific community, governmental organisations and wider public, as well as local stakeholders.

3. FAIR DATA

3.1 Making data findable, including provisions for metadata

The consortium uses a data repository in Microsoft Office Teams to which all partners have access to and have the opportunity to read, edit, create, share, or delete documents and other files. The details of the data repository, its structure and best practices for its use have been collected in the Deliverable D9.1 Project Management Plan, and discussed with the full consortium in project meetings. Within the consortium's data repository, the project data is easily findable by being organised in clearly labelled folders. Files saved include clear version numbers. Optionally, some files may include information on the date of creation or last edit, and the initials of the editors of the document.

To make project results and associated metadata easily findable, SophiA will make use of the open science data platform Zenodo.org for the dissemination and sharing of the peer-reviewed academic publications of the project. The Zenodo platform makes all related metadata publicly available, allows the content to be found, accessed and cited easily and it is free of charge.

The platform, developed by CERN, the OpenAIRE initiative and the European Commission, uses open APIs and is a valued resource for sharing research results from H2020 projects in a FAIR manner. All academic peer-reviewed publications will have a Digital Object Identificator (DOI).

The metadata fields collected and made accessible through Zenodo (https://zenodo.org/deposit/new) include:

- Communities: One publication can be assigned to several communities
- Type of publication: e.g. Publication, Poster, Presentation, Dataset, Image, Video/Audio, Software, Lesson, Other. Further subcategories can also be defined.
- Mandatory basic information: publication date, title, authors, description (abstract according to the agreed internal guideline).
- Optional basic information: DOI, version, language, keywords, additional notes.
- License information: access rights (open, embargoed, restricted, closed access), license name (Creative Commons versions).
    - All academic publications of the SophiA project are planned to be submitted open access.
- Funding: Grant programme name and number
- Related/alternative identifiers
- Moreover, to target the refrigeration industry, FRIDOC – the leading platform dedicated to refrigeration information will be used

In addition to Zenodo, some partners may make use of their own research publication repositories. This is the case of IIR and the Tridoc platform, where scientific publications, conference proceedings and articles will be added. Nevertheless, all the entries included in separate repositories such as Tridoc will be also added to the Zenodo platform.

3.2 Making data openly accessible

The research results and data will be shared within the consortium as necessary for the successful implementation of the project. Background information or IP by partners that is necessary for the implementation of the project will be shared with other consortium partners during the duration of the project, as stipulated in the Consortium Agreement.

The accessibility of the data generated by the SophiA project to external persons will depend on the type of data, its purpose and commercial use of it for the consortium partners. The project results will generally be made available through presentations, publications, newsletters, reports and articles, which will be easily available and accessible to all through the project website. The extent to which the full data is made accessible will be decided taking into account data protection, confidentiality and possible IP concerns. The different data categories depending on their accessibility can be summarised as:

1. **General project information, public reports and educational materials**

Public information regarding the project results, project activities, findings, best practices, demonstrations and reports will be made available and openly accessible to everyone through the project's website. The SophiA website will have a Downloads section where public reports, public deliverables, press releases, brochures, articles, newsletters, presentations and

published scientific articles will be made accessible. In addition, educational materials, such as training materials, SophiA handbook, workshops or webinars will also be made accessible through the project's website. Links to this information, as well as direct links to relevant project reports, publication or newsletters will be shared through the SophiA social media channels.

2. **Scientific publications**

Scientific publications will be published in academic journals and will be made accessible open sources. Publications will be included in Zenodo repository as explained above, and will be made accessible to download on the SophiA project website.

3. **Data retrieved from questionnaires**

Within WP1, partners have developed questionnaires in order to identify the needs of the local hospital facilities where the SophiA containers will be demonstrated. In WP7, a social acceptance study will take place, involving interviews and questionnaires directed at local stakeholders and local public of the regions where SophiA will be demonstrated and where it may be of interest in the future. In WP8, the dissemination activities such as workshops, trainings and meetings may collect feedback from stakeholders. Any personal data collected through these questionnaires will be fully anonymised and will not be shared, according to data protection regulations. Data will from questionnaires will not be made accessible to protect the individual and business interests of the respondents. Only the results of the analysis of the data collected through questionnaires will be included as part of reports which may be made publicly available and accessible, if they do not interfere with commercial or IP interest of project partners or other stakeholders. In all cases, only the analysed data, or the conclusions derived from the analyses, will be shared, in a way that no participant to the questionnaires is identifiable.

4. **Software (open source code)**

Relevant software will not be made accessible as software developments by the partners themselves form part of the organisation's IP and is therefore protected. In other instances, such as drafting software, it cannot be made accessible as it is used under license.

Software needed to access data

For all the above data, the software tools needed to access the data will be commercially available software, including open source alternatives. No documentation about the software will be necessary to access the published data.

Restrictions of use

If necessary, in Zenodo, files can be added under restricted access with the possibility of sharing only if certain requirements are met, or after a certain time. These files would not be made publicly available immediately.
Files may be added under closed, open or embargoed access. Files deposited under "closed" access are protected against unauthorised access at all levels. For any files deposited under "embargo", the repository will restrict access to the data until the end of the embargo period, at which time the content will become automatically publicly available.

3.3 Making data interoperable

The data produced in the project will allow exchange and be re-usable between researchers, institutions, organisations and countries since it will use common formats, will be accessible and shareable using commercial software and will be compatible with open software applications. Standard vocabularies will be used for all data types present in the SophiA datasets in order to allow for inter-disciplinary interoperability.

All data and project documents will be provided in English, as the common language for the consortium partners. The project website will be available in English, as well as in French and German to facilitate communication efforts in the local settings of project partners. Certain communication, dissemination and capacity building materials, as well as questionnaires, will also be available in French and German, and always in English. The Communication and Dissemination WP leader, IIR, together with local project partners, will be responsible for the translation of communication materials to French and German.

3.4 Increase data re-use (through clarifying licenses)

Scientific publications, conference proceedings and public reports, newsletters, press releases and general articles will be widely shared through social media and in the project website and available for re-sharing and re-use. Data and summaries of the project results will also be presented at events in forms of talks, workshops, posters or training sessions.

In the case of research data, the data will be made available for re-use once the article has been published in a peer-reviewed journal.

Project data, which has been published through open access journals, made available in the Zenodo platform or shared in the project's website or social media will be accessible and useable by third parties immediately after publishing.

Certain data including software, technical drawings, technical designs of the container's components, manufacturing processes or specific methodologies developed by project partners that have a commercial interest to at least one partner will not be made available for re-use to third parties. This decision is made on the grounds of protecting the project partners' IP.

The project website and the data collected and downloadable from it will remain online and available for access and re-use until 5 years after the project end.

4. ALLOCATION OF RESOURCES

All academic and research partners of the project have foreseen costs for the publication of scientific peer-reviewed articles in open-access journals. In addition, industrial and research partners may use project costs for patent filling applications in order to protect and further exploit the exploitable results and IP generated from the project. The costs for publications and patent fillings will be covered by the SophiA project as these have been foreseen in the project budget as described in the Grant Agreement, Description of Action Annex 2, and they are eligible costs under H2020 programme.

The Zenodo.org platform is free of charge and therefore no additional costs are expected for the inclusion of project results in this open platform.

Regarding the internal project data repository, SEZ set up the repository using Microsoft Office Teams as part of its internal data system and using its own license. As such, the creation and maintenance of this repository is at no extra cost for the project. All project partners have access to the MS Teams or SharePoint files through their own organisation licenses, or may create a new account free of charge.

All partners have appointed a person to be responsible for the data management and data assurance of their organisation for the SophiA project. This is, in general, a project manager or project engineer strongly involved in the project.

The resources for long term preservation of the data will be discussed with the consortium and updated throughout the project where necessary. The website will be maintained for 5 years after the project end and the open-source publications will remain available long term. Further than this, long-term preservation or sharing of other data and project results may involve additional and bilateral agreements between project partners which will be agreed upon during the project duration.

5. DATA SECURITY

All partners are individually responsible for their data security and preservation. At partner level, most partners rely on local servers or external hard-disk drives, local servers with an additional back-up service at a different location, cloud storage protected by a firewall, all of which are regularly updated, backed up and are protected by a password.

The common project results and final data are saved during the project lifetime on the Microsoft Office Teams platform, managed by SEZ and regularly backed-up. All data available through the project's website are subject to the data security and recovery mechanisms of the host.

6. ETHICAL ASPECTS

Ethical issues of the data generated by the project have been discussed in Ethics deliverables under WP10 (D10.1-D10.4).

Informed consent for data sharing and long-term preservation is included in questionnaires dealing with personal data.

7. OTHER ISSUES

Currently, there is no need to make use of other national/funder/sectorial/departmental procedures for data management.

8. DEVIATIONS FROM THE WORKPLAN

No deviations from the work plan have occurred in the preparation of this deliverable.

9. CONCLUSIONS

This deliverable constitutes the first version of the SophiA Data Management Plan and lays out the strategy regarding data management for the full project and for partners individually. It provides information on the type of data collected and the measures that will be taken in order to ensure a FAIR data framework. This first version is seen as a benchmark to report the actions that have been implemented and to identify any which are still to be performed in order to fulfil the EC requirements in terms of data management and accessibility.

Improvements of DMP will be discussed during internal project meetings with all partners. Two updates are planned, for M24 and M36 of the project.

10. ANNEXES

10.1 Annex 1

The following questionnaire was completed by all partners in order to gather the information about the data requirements and systems in place for following a FAIR data treatment. The completed questionnaires are saved in file in the project repository and can be made available to the European Commission upon request.

10.1.1 Data summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?
What types and formats of data will the project generate/collect?
Will you re-use any existing data and how?
What is the origin of the data?
What is the expected size of the data?
To whom might it be useful ('data utility')?

10.1.2 FAIR data

10.1.2.1 Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata?
Are the data produced identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?
What naming conventions do you follow?
Will search keywords be provided that optimize possibilities for re-use?
Do you provide clear version numbers?
What metadata will be created? **In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.**

10.1.2.2 Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default?

If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.
How will the data be made accessible (e.g. by deposition in a repository)?
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)?
Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.
Have you explored appropriate arrangements with the identified repository?
If there are restrictions on use, how will access be provided?
Is there a need for a data access committee?
Are there well described conditions for access (i.e. a machine readable license)?
Where will the data and associated metadata, documentation and code be deposited? How will the identity of the person accessing the data be ascertained?

10.1.2.3 Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?
What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?
Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?
In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

10.1.2.4 Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?
When will the data be made available for re-use?

If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? **If the re-use of some data is restricted, explain why.**
How long is it intended that the data remains re-usable?
Are data quality assurance processes described?

10.1.3 Allocation of resources

What are the costs for making data FAIR in your project?
How will these be covered? **Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).**
Who will be responsible for data management in your project?
Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

10.1.4 Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?
Is the data safely stored in certified repositories for long term preservation and curation?

10.1.5 Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? **These can also be discussed in the context of the ethics review.** **If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).**
Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

10.1.6 Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

sophia4africa.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 101036836.