H2020-LC-SC3-2018-2019-2020 / H2020-LC-SC3-2020-RES RIA

ENERGY HARVESTING IN CITIES WITH TRANSPARENT AND HIGHLY EFFICIENT WINDOW-INTEGRATED MULTI-JUNCTION SOLAR CELLS

CITYSOLAR
(No 101007084)

D8.3 – Data Management Plan (DMP)

WP8 – **WP Management** (CNR)

Lead beneficiary of this deliverable:	<CNR>	Authors:
Due date of deliverable: 31. 05. 2021	Actual submission date: 10/06/2021	- Maria Rita Bruni (CNR) - Franca Rossi (CNR) - Andrea Laruffa (CNR) - Carlo Ottaviani (CNR) - Alessandro Ippoliti (CNR) Reviewer: - Aldo Di Carlo (CNR) - Steering Committee

Dissemination Level		
PU	Public	X
PP	Restricted to other programme 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)	

REVISION HISTORY

Rev.	Date	Changes

Disclaimer: the information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

Table of Content

Executive Summary

CITYSOLAR is a research and innovation project in which Transparent PhotoVoltaics (TPV) plays a fundamental role for a pervasive application of solar energy. CITYSOLAR project propose a new breakthrough concept for TPV development, by exploiting the combined use of emerging technologies based on a multi-junction combination of a near-ultraviolet absorbing cell and a near-infrared absorbing cell together with advanced concepts of light management such as photonic crystals, nanophotonics and photon recycling and an advanced integration scheme among modules.

During the course of the project, the data will be generate in a wide range of R&D activities from material synthesis to prototype device testing and validation going through modeling, characterizations and functionalization parameters tests.

As a project participating in the Open Research Data Pilot (ORDP) in Horizon 2020, CITYSOLAR will make its research data findable, accessible, interoperable and reusable (FAIR). Nevertheless, data sharing in the open domain may be restricted, taking in account "the need to balance openness and protection of scientific information, commercialization and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions" as stated in Guidelines on FAIR Data Management in Horizon 2020 published by the European Commission.

The DMP's purpose is, therefore, to provide the main elements of the data management policy to be used by the Consortium, with regard to its complete research data cycle. It describes: types and formats of data to be generated or collected, and how; the standards to be applied; the data-reservation methods; the data-sharing policies for re-use. The DMP traces the exploitation and IPR requirements, as defined in the Consortium agreement.

The CITYSOLAR Data Management Plan (DMP) follows the Horizon 2020 DMP template that was conceived to be applied to any Horizon 2020 project that produces, collects or process research data. This first Data Management Plan describes the CITYSOLAR data management principles and strategies, tools and data. The consortium will aim at open access when publishing papers and articles.

The DMP is a living document and, therefore, it will be updated as the implementation of the project progresses and whenever significant changes occur.

Acronym and Abbreviations

Acronym	Description
BIPV	BIPV prototype
CNR	National Research Council of Italy
DBR	Distributed Bragg Reflector
DMP	Data Management plan
FAIR	Findable, Accessible, Interoperable and Reusable
IPR	Intellectual Property Rights
LM	Light Management
NIR	Near-Infra Red
NUV	Near-Ultra Violet
OpenAIRE	Open Access Infrastructure for Research in Europe
ORDP	Open Research Data Pilot
UC	Up-Converter
TPV	Transparent PhotoVoltaics

Table 1 – Acronyms

1. Introduction

All European Union funded projects must try to disseminate as much information as possible, and on top of that the CITYSOLAR project was signed up to the "Open Research Data Pilot" (ORDP), hence committed to give open access to data generated during the project, unless it goes against its legitimate interests. The main purpose of the DMP, consequently, is to ensure the accessibility and intelligibility of the data generated during the CITYSOLAR project in order to comply with the Guidelines of the ORDP. Each data set created during the project will be assessed, and categorized, as open, embargo or restricted by the owners of the content of the data set.

CITYSOLAR, as said, will produce data in a wide range of R&D activities which, once generated (or collected), will be stored in several formats, which are: Documents, Images, Data etc.

The CITYSOLAR Data Management Plan aims at providing a strategy for managing key data generated and collected during the project and optimize access to and re-use of research data. The DMP is intended to be a 'living' document that will outline how the CITYSOLAR research data will be handled during and after the project, and so it will be reviewed and updated as the project implementation progresses and/or significant changes occur. Aa a minimum, it will be updated contextually to the periodic reporting/evaluation of the project.

2. Data Management Plan principles

All the data sets, regardless of their categorization, will be stored in each of the participant entities databases and in the cloud infrastructure at CNR-ARTOV, in a restricted access area specifically created as an internal database of the partners (as described in D8.1). They will be referenced within the CITYSOLAR website (www.citysolar-h2020.eu) in a restricted access page. In addition, both open or embargo data sets, will be publicly shared (in the case of embargo, after the embargo period is over) through the public section of the project website and **ZENODO (**https://zenodo.org/**).** In particular CITYSOLAR consortium defined the following community on Zenodo

ZENODO is an open access repository for all fields of science that allows uploading any kind of data file formats, which is recommended by the Open Access Infrastructure for Research in Europe (OpenAIRE).

For all data types, the Consortium will consider the aspects of potential conflicts against commercialization, and the IPR protection issues of the knowledge generated, before deciding which information needs to be made public, and when. The decision process, summarized in the figure below, follows the principle of Open Access according to the Horizon 2020 guideline.
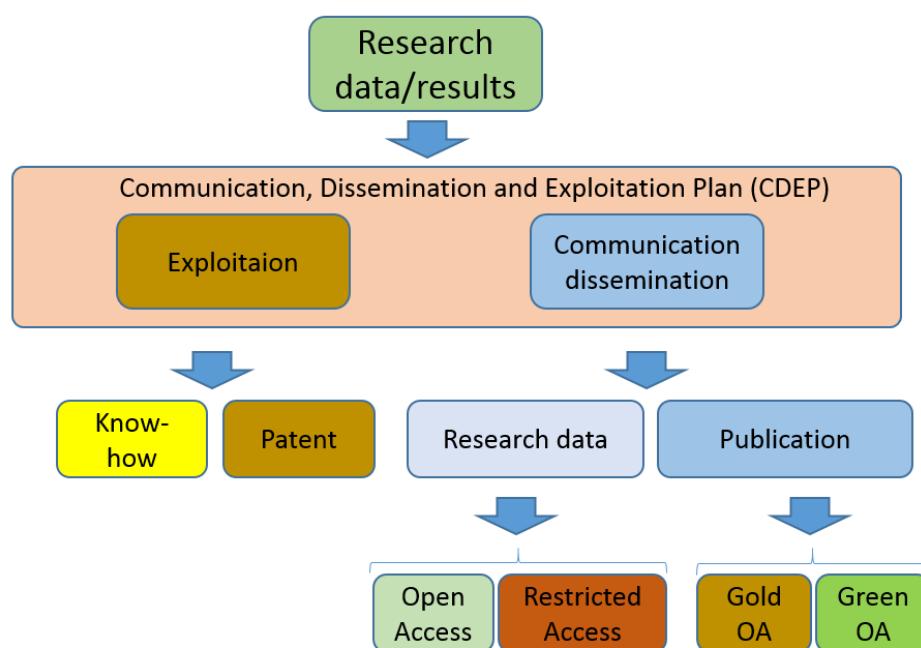

Fig.1 Open access to research data and publication decision diagram *(from Guidelines to the Rules on Open Access to Scientific publications and Open Access to Research Data in Horizon 2020)*

The other main issues for the CITYSOLAR project DMP are the following:

1. This Data Management Plan (DMP) has been prepared by taking into account the template of the "Guidelines on Data Management in Horizon 2020" on site: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

2. The DMP is an official project Deliverable (D8.3) due in Month 6 (end of MAY 2021), to be updated throughout the project. The first initial version will evolve depending on significant changes arising and periodic reviews at relevant reporting stages of the project.

3. The consortium complies with the requirements of Regulation (EU) 2016/679 and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Guidance on how these regulations interact with open-access data policy can be found here: https://www.openaire.eu/ordp/

4. The type of data, storage, confidentiality, ownership, management of intellectual property and access: procedures that will be implemented for data collection, storage, access, sharing policies, protection, retention and destruction will be in line with EU standards as described in the Grant Agreement and the Consortium Agreement.

As stated in the Grant Agreement (Article 29.3) "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 would be jeopardized by making those specific parts of the research data openly accessible."

Such an exception applies to CITYSOLAR whenever the project findings provide high innovation level (possibility of commercialization, etc.). In this case, the consortium will consider two forms of protection:

1) to withhold the data for internal use;

2) to apply for a patent in order to commercially exploit the invention and have in return financial gain.

In the former case, appropriate IPR protection measures (e.g., NDA), must be taken for data sharing outside the consortium, meanwhile, in the latter, publications will be delayed until the patent filing is completed. Otherwise, the results will be made "Open Access" by depositing the research data into an online repository service (ZENODO) or by publishing in journals (document, reports, articles, etc.) adhering to suitable "Open Access" ('green' or 'gold').

In parallel, public deliverables will be stored on one (or both) of the following locations: The CITYSOLAR website (https://www.citysolar-h2020.eu) after approval by the consortium, and the CITYSOLAR page on https://cordis.europa.eu/project/id/101007084 website where all public deliverables submitted to the European Commission are hosted.

3. FAIR data

As a project participating in the Open Research Data Pilot (ORDP) in Horizon 2020, the DMP's Data Management strategy of CITYSOLAR project is focused on the observation of FAIR Data Management Protocols.

To this end, each data set collected, processed and/or generated in the project is identified as follows:

Data set reference and name: internal project Identifier for the data set to be produced. This will follow the format:

CITYSOLAR_DS_Name-of-set_Partner-name_Subset-name_Dissemination-level_Version_Date-of-storage

where partner name represents the name of the data custodian (WP Lead/ Task Leader/activity leader) and subset name a kind of keyword defining the data type as: material, synthesis, characterization, structure etc.

Purpose and relation to the objectives of the project: description, type of research and aim of the data generated or collected, including its origin (in cases where data is collected), nature and scale and to whom it could be useful.

Data type: such as document, raw data, elaborated data etc.

File format: all common electronic formats and/or encrypted data for specific treatment software.

Reuse of existing data: condition in which and by the existing data will be reused.

Data production methods: by who and by which method the dataset will be generated.

Archiving and preservation (including storage and backup): description of the procedures to be put in place for long-term preservation of the data, including an indication of how long the data should be preserved, the approximate end volume, associated costs, and how these are planned to be covered.

Data utility: for which scope the dataset will be used.

Potential for reuse: to who, in addition of project itself, the dataset will be useful.

Diffusion principles: description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling reuse, and definition of whether access will be open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating the type of repository (institutional, standard repository for the discipline, etc.). In cases where the dataset cannot be shared, the reasons for this will be stated (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).

3.1. CITYSOLAR datasets

To focalise the Data Management strategy of CITYSOLAR project on the FAIR principles five datasets were identified:

3.1.1. Dataset on top Near-UV Perovskite Solar Cell (DS_NUVcell)

Data set reference and name	DS_NUVcell
Purpose and relation to the objectives of the Project	The purpose of the data is to define a set of materials, device architectures, fabrication protocols and characterizations used by the CITYSOLAR consortium to identify the Near-UV perovskite top cell with the efficiency, average visual transparence, and dimension identify by th eWP3 of the project.
Data types	Documents, data, images
File format	Document and images: All common electronic document formats (.docx, .pdf, .tex, etc.). Data: text format tables that are readable by common data analysis software, or encrypted for specific data treatment software (to be defined). For what concern the material, the format of the file is (M2.1): **MID:** The material ID is represented as a 6-digit number which consist of a first digit identifying the institute, two digits for the year and three digits for an increasing number to identify individual material batches **type** the type of material (absorber, HTL, ETL, electrode etc.) to improved searchability and to suggest relevant materials for a specific layer **Name** the name of the material **Supplier** the name of the supplier of the material **AdditionalFields** Additional fields like batch number, purity, amount, main user, and a picture can optionally be assigned to each material. For devices the following applies: **DID:** The device ID is represented as a 6-digit number which consist of a first digit identifying the institute, two digits for the year and three digits for an increasing number to identify individual device batches

- **Name** the name of the device
- **Stack** the stack sequence starting from the bottom to the top. This include also the thickness of the layers
- **AdditionalFields** Additional fields that could also link to attachments and related for example to AVT, PV characteristics, morphology study and other characterization data.

Reuse of existing data	Processed and aggregated data will be shared by partners not collecting data for the advancements of the Project, adhering to the access rights conditions to results and background as described in CA - Section 9
Data production methods	The dataset will be generated by partners laboratories through experimental trials, measurements, and numerical simulation. The dataset will also include summaries of Project meetings and discussions between partners, and relevant publications in scientific journals.
Archiving and preservation (including storage and backup)	See section 2
Data utility	These data are provided and use to all the partner of the project involved in WP1,2,3 an in addition will be used by partners working on the Organic Solar cell (WP4) and on the integration between organic and Perovskite solar cell (WP6). In addition WP5 will use these data to design the light management part of the CITYSOLAR device.
Potential for reuse	It will be useful for other research groups working especially in the field of perovskite-based photovoltaics, but also for other domains of (academic and industrial) PV research and development.
Diffusion principles	Currently data are shared via scientific communication, either in the form of scientific papers or at conferences via oral or visual presentations. While the peer-reviewed publications will be made available as open access, it is the intention to also provide the underlying data itself by storing it in an open repository (see Section 2)

3.1.2. Dataset on bottom Near-IR Organic Solar Cell (DS_NIRcell)

Data set reference and name	DS_NIRcell
Purpose and relation to the objectives of the Project	The purpose of the data is to define a set of materials, device architectures, fabrication protocols and characterizations used by the CITYSOLAR consortium to identify the Near-IR perovskite bottom cell with the efficiency, average visual transparence, and dimension identify by the eWP4 of the project.
Data types	Documents, data, images
File format	Document and images: All common electronic document formats (.docx, .pdf, .tex, etc.). Data: text format tables that are readable by common data analysis software, or encrypted for specific data treatment software (to be defined). For what concern the material, the format of the file is (M2.1): **MID:** The material ID is represented as a 6-digit number which consist of a first digit identifying the institute, two digits for the year and three digits for an increasing number to identify individual material batches **type** the type of material (acceptor, donor, HTL, ETL etc.) to improved searchability and to suggest relevant materials for a specific layer **Name** the name of the material **Supplier** the name of the supplier of the material **AdditionalFields** Additional fields like batch number, purity, amount, main user, and a picture can optionally be assigned to each material. For devices the following applies: **DID:** The device ID is represented as a 6-digit number which consist of a first digit identifying the institute, two digits for the year and three digits for an increasing number to identify individual device batches **Name** the name of the device **Stack** the stack sequence starting from the bottom to the top. This include also the thickness of the layers

	• **AdditionalFields** Additional fields that could also link to attachments and related for example to AVT, PV characteristics, morphology study and other characterization data.
Reuse of existing data	Processed and aggregated data will be shared by partners not collecting data for the advancements of the Project, adhering to the access rights conditions to results and background as described in CA - Section 9
Data production methods	The dataset will be generated by partners laboratories through experimental trials, measurements, and numerical simulation. The dataset will also include summaries of Project meetings and discussions between partners, and relevant publications in scientific journals.
Archiving and preservation (including storage and backup)	See section 2
Data utility	These data are provided and use to all the partner of the project involved in WP1,2,4 an in addition will be used by partners working on the Perovskite Solar cell (WP3) to tune the cell for tandem objectives of the project and on the integration between organic and Perovskite solar cell. (WP6). In addition WP5 will use these data to design the light management part of the CITYSOLAR device.
Potential for reuse	It will be useful for other research groups working especially in the field of organic-based photovoltaics, but also for other domains of (academic and industrial) PV research and development. Results on material could also be used in the organic electronic context as well as for Organic Light Emitting Diodes (OLED)
Diffusion principles	Currently data are shared via scientific communication, either in the form of scientific papers or at conferences via oral or visual presentations. While the peer-reviewed publications will be made available as open access, it is the intention to also provide the underlying data itself by storing it in an open repository (see Section 2)

3.1.3. Dataset on Light Management (DS_LM)

Data set reference and name	DS_LM
Purpose and relation to the objectives of the Project	The purpose of the data is to define a suitable set of materials, architectures, fabrication protocols and characterizations for the Light Management of the top NUV perovskite cell, bottom NIR organic solar cell and the tandem structure of CITYSOLAR. This include data for both Distributed Bragg Reflectors and Up-Converter
Data types	Documents, data, images
File format	Document and images: All common electronic document formats (.docx, .pdf, .tex, etc.). Data: text format tables that are readable by common data analysis software, or encrypted for specific data treatment software (to be defined). For LM architecture the following applies: • **AID:** The architecture ID is represented as a 6-digit number which consist of a first digit identifying the institute, two digits for the year and three digits for an increasing number to identify individual device batches • **Name** the name of the architecture • **Stack** the stack sequence starting from the bottom to the top. This include also the thickness of the layers • **AdditionalFields** Additional fields that could also link to attachments and related for example to AVT, PV characteristics, morphology study and other characterization data.
Reuse of existing data	Processed and aggregated data will be shared by partners not collecting data for the advancements of the Project, adhering to the access rights conditions to results and background as described in CA - Section 9
Data production methods	The dataset will be generated by partners laboratories through experimental trials, measurements, and numerical simulation.

	The dataset will also include summaries of Project meetings and discussions between partners, and relevant publications in scientific journals.
Archiving and preservation (including storage and backup)	See section 2
Data utility	These data are provided and use to all the partner of the project involved in WP5 an in addition will be used by partners working on the Perovskite Solar cell (WP3) and on Organic Solar Cell (WP4) to define the efficiency and AVT of the cell. Data will be used by WP6 to design the final CITYSOLAR tandem cell and module
Potential for reuse	It will be useful for other research groups working especially in the general field of photovoltaics, but also for other domains of (academic and industrial) related to optoelectronics devices.
Diffusion principles	Currently data are shared via scientific communication, either in the form of scientific papers or at conferences via oral or visual presentations. While the peer-reviewed publications will be made available as open access, it is the intention to also provide the underlying data itself by storing it in an open repository (see Section 2)

3.1.4. Dataset on Tandem cells and modules (DS_TANDEM)

Data set reference and name	DS_TANDEM
Purpose and relation to the objectives of the Project	The purpose of the data is to define a suitable set of lamination materials, architectures, fabrication protocols and characterizations for the CITYSOLAR tandem cell and module composed by the top NUV perovskite cell/module and the bottom NIR organic solar cell/module
Data types	Documents, data, images
File format	Document and images: All common electronic document formats (.docx, .pdf, .tex, etc.).

	Data: text format tables that are readable by common data analysis software, or encrypted for specific data treatment software (to be defined). For TANDEM architecture the following applies: • **AID:** The architecture ID is represented as a 6-digit number which consist of a first digit identifying the institute, two digits for the year and three digits for an increasing number to identify individual device batches • **Type** if tandem cell or tandem module • **Name** the name of the architecture • **Stack** the stack sequence starting from the bottom to the top. This include also the thickness of the layers • **Area** aperture area of the tandem cell/module • **AdditionalFields** Additional fields that could also link to attachments and related for example to AVT, PV characteristics, morphology study and other characterization data.
Reuse of existing data	Processed and aggregated data will be shared by partners not collecting data for the advancements of the Project, adhering to the access rights conditions to results and background as described in CA - Section 9
Data production methods	The dataset will be generated by partners laboratories through experimental trials, measurements, and numerical simulation. The dataset will also include summaries of Project meetings and discussions between partners, and relevant publications in scientific journals.
Archiving and preservation (including storage and backup)	See section 2
Data utility	These data are provided and use to all the partner of the project involved in WP6 and in addition will be used by partners working on the Perovskite Solar cell (WP3) and on Organic Solar Cell (WP4) to tune the efficiency and AVT of the cell in order to reach the final objectives of the project. Data will be used by WP6 to design the final CITYSOLAR BIPV demonstrator

Potential for reuse	It will be useful for other research groups working especially in the general field of photovoltaics, in particular in tandem applications, but also for other domains of (academic and industrial) related to optoelectronics devices.
Diffusion principles	Currently data are shared via scientific communication, either in the form of scientific papers or at conferences via oral or visual presentations. While the peer-reviewed publications will be made available as open access, it is the intention to also provide the underlying data itself by storing it in an open repository (see Section 2)

3.1.5.   Dataset on Building Integrated PhotoVoltaic prototype (DS_BIPV)

Data set reference and name	DS_BIPV
Purpose and relation to the objectives of the Project	The purpose of the data is to define a suitable set of integration architecture, procedure and characterizations for the CITYSOLAR BIPV prototype integrating the tandem cell/module composed by the top NUV perovskite cell/module and the bottom NIR organic solar cell/module
Data types	Documents, data, images
File format	Document and images: All common electronic document formats (.docx, .pdf, .tex, etc.). Data: text format tables that are readable by common data analysis software, or encrypted for specific data treatment software (to be defined).
Reuse of existing data	Processed and aggregated data will be shared by partners not collecting data for the advancements of the Project, adhering to the access rights conditions to results and background as described in CA - Section 9
Data production methods	The dataset will be generated by partners laboratories through experimental trials, measurements, and numerical simulation.

	The dataset will also include summaries of Project meetings and discussions between partners, and relevant publications in scientific journals.
Archiving and preservation (including storage and backup)	See section 2
Data utility	These data are provided and use to all the partner of the project involved in WP6 and in addition will be used by partners working on the Perovskite Solar cell (WP3) and on Organic Solar Cell (WP4) to tune the efficiency and AVT of the cell in order to reach the final objectives of the project.
Potential for reuse	It will be useful for other research groups working especially in the general field of BIPV or other photovoltaic application where integration of PV in other objects is mandatory (including for example agrivoltaic, Vehicle integrated PV etc.).
Diffusion principles	Currently data are shared via scientific communication, either in the form of scientific papers or at conferences via oral or visual presentations. While the peer-reviewed publications will be made available as open access, it is the intention to also provide the underlying data itself by storing it in an open repository (see Section 2). The final BIPV prototype report including data will be public available (D6.7)

3.2.   Making data findable, including provisions for metadata

Metadata

Metadata is data on the research data themselves. It enables other researchers to find data in an online repository and is, as such, essential for the reusability of the dataset. Abundant and detailed metadata enable other researchers, to better determine whether the dataset is relevant and useful for their own research. Metadata (type of data, location, etc.) will be uploaded in a standardized form. This metadata will be kept separate from the original raw research data.

As described in the project Grant Agreement (Article 29.2), the bibliographic metadata include all of the following:

-   the terms "European Union (EU)" and "Horizon 2020";
-   the name of the action, acronym and grant number;
-   the publication date, and length of embargo period if applicable;
-   a persistent identifier.

Persistent and unique identifier

Digital Object Identifier (DOI) and Creative Common's license numbers will be used as persistent identifiers on open data repositories. If the object does not have a DOI (as in the case, for example, of publications), ZENODO will assign a proper DOI.

Naming conventions

Files and folders in data repositories will be versioned and structured using a name convention consisting of project name datasetname and ID as described before:

CITYSOLAR_DS_Name-of-set_Partner-name_Subset-name_Diss-level_Version_Date-of-storage<.fileformat>

Search keywords

Example Keywords (will be modified with the project advancement):

>   DS_NUVcell: perovskite solar cell, transparent solar cell, synthesis, structure, simulation
>
>   DS_NIRcell: organic solar cell, transparent solar cell, synthesis, structure, simulation
>
>   DS_LM: Distributed Bragg reflector, up-conversion, light management
>
>   DS_TANDEM: lamination, tandem cell, tandem module, encapsulation
>
>   DS_BIPV: Building Integrated photovoltaics, encapsulation, Photovoltaic windows

Version numbers

Individual file name will contain version numbers that will be incremented at each version.

3.3.   Making data openly accessible

In order to maximise the impact of CITYSOLAR research data, the results are shared within and beyond the Consortium. Selected data and results will be shared among the scientific community, and other stakeholders, through publications in scientific journals and presentations at conferences, as well as through open access data repositories.

CITYSOLAR open data will be collected in an open online research data repository: ZENODO**,** whose structure, facilities and management are in compliance with FAIR data principles.

ZENODO is an Open Access Infrastructure for Research in Europe (OpenAIRE) that allows researchers to deposit both publications and data, providing tools to linking them to these through persistent identifiers and data citations, it is set up to facilitate the finding, accessing, re-using and interoperating of data sets, which are the basic principles that ORD projects must comply with.

ZENODO repository is provided by OpenAIRE and hosted by CERN and is a catch-all repository that enables researchers, scientists, EU projects and institutions to:

- Share research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science;

- Display their research results and get credited by making the research results citable and integrating them into existing reporting lines to funding agencies like the European Commission;

- Easily access and reuse shared research results;

- Integrate their research outputs with the OpenAIRE portal.

CITYSOLAR consortium has defined a "Community" on Zenodo named: **CITYSOLAR H2020 Project**

The CITYSOLAR project datasets are first stored and organized in a database by the data owners (personal computer, or on the institutional secure server) and on the project database (project website's private section). All data are made available for verification and re-use, unless the task leader can justify why data cannot be made openly accessible.

To protect the copyright of the project knowledge, Creative Commons license will be used in some cases. The CITYSOLAR dataset deliverables are both public (data access policy unrestricted) and they will be accessible by:

- CITYSOLAR project web site (www.citysolar-h2020.eu);

- Partners database;

- OpenAIRE;

- ZENODO (https://zenodo.org) for ORDP data and datasets;

- Open access journals.

All data deposited on ZENODO are accessible without restriction for public. For other data, potential users must contact the IPR team or the data owner in order to gain access. If necessary, appropriate IPR procedure (such as non- disclosure agreement) will be used.

Most data are produced in common electronic document/data/image format that do not require specific software (.docx, .pdf, .tex, .jpg, etc), the data that requiring specific compilers will be specified.

Zenodo FAIR priciples are available at https://about.zenodo.org/principles/

3.4. Making data interoperable

OpenAIRE guidelines for online interoperability, including OpenAIRE Guidelines for Literature Repositories, wil be observed by the CITYSOLAT partners.

 OpenAIRE Guidelines for Data Archives, these guidelines can be found at: https://guidelines.openaire.eu/en/latest/.
Partners will also ensure that CITYSOLAR data observes FAIR data principles under H2020 open-access policy:

- Controlled vocabularies will be used in descriptive metadata field to support consistent, accurate and quick indexing and retrieval of relevant data;

- Keywords and their synonyms will be used for indexing and subject headings of the data e metadata;

- Keywords, as controlled vocabularies change within different disciplines of Science, will be updated during the course of the project to increase the interoperability.

In order to ensure the interoperability, all datasets will use the same standards for data and metadata capture/creation.

As the project progresses and data are identified and collected, further information on making data interoperable will be outlined in subsequent versions of the DMP.

3.5.   Increase data re-use (through clarifying licences)

It is the intention of the CITYSOLAR consortium to allow the widest possible re-use of data. Creative Commons licensing is considered for some datasets for which the Consortium has decided to make that part public. Both Share-Alike and NonCommercial-ShareAlike licenses will be considered.

The date of data release will be immediately after the Consortium decision to make data Open Access. However, an embargo period may be applied if the data (or parts of data) are used in published articles in "Green" open access journals. The recommended maximum embargo period length by European Commission is 6 months. For datasets deposited on a public data repository (ZENODO) the access is unlimited.

Restrictions on re-use policy are applied for all protected data whose re-use will be limited within the project partners.

CITYSOLAR data will remain re-usable for at least 1 year.

4. Allocation of resources

Costs related to open-access to research data in Horizon 2020 are eligible for reimbursement under the conditions defined in the H2020 Grant Agreement. Project beneficiaries will be responsible for applying for reimbursement for costs related to making data accessible to others beyond the consortium.

The costs for making data FAIR include:

- Fees associated with the publication of scientific articles containing project's research data in "Gold" Open access journals. The cost sharing, in case of multiple authors, shall be decided among the authors on a case-by-case basis;

- Project Website and cloud storage operation: *free of charge*;

- Data archiving at ZENODO: *free of charge*;

- Copyright licensing with Creative Commons: *free of charge*.

During the project, data will be updated regularly as new results are submitted by partners.

Each partner is responsible for the data they produce. Any fee incurred for Open Access through scientific publication of the data will be responsibility of the data owner (authors) partner(s) in compliance with the CA, Art. 8.4.2.1.

The project members of General Assembly are also responsible for the Data Management of CITYSOLAR dataset and research data in accordance with each organization internal Data Protection Officer.

Data preservation of at least 1 year after the project is required by GA (Art.31.3). The associated cost for dataset preparation for archiving will be covered by project itself. Long-term preservation will be provided and associated cost covered by selected disciplinary repository.

5. Data security

CITYSOLAR will use methods that emphasize easy access and extended contact, and trust building among participants.

The following guidelines will be followed in order to ensure the security and integrity of the data:

- Access to data will be granted via an HTTPS connection with ECDHE-RSA-AES256-GCM-SHA512:DHE-RSA-AES256-GCM-SHA512:ECDHE-RSA-AES256-GCM-SHA384:DHE-RSA-AES256-GCM-SHA384 ciphering

- Data backup resides on an independent server array separated from the access server to avoid loss of data;

- Encrypt data is not presently implemented; this feature will be added if it is deemed necessary by the participating researchers;

- Limit the use of USB flash drives;

- Label files in a systematically structured way in order to ensure the coherence of the final dataset.

All project deliverables and data will be stored and shared on a cloud facility managed by CNR-ARTOV and will be referenced inside a restricted area within the CITYSOLAR website (www.citysolar-h2020.eu) As an initial step, only the Consortium Partners will have access to the server storage area where dataset and metadata are filed.

Moreover, scientific publications and articles, the dataset deliverables, and the final demonstrator research results will be shared through ZENODO and other database to promote FAIR data making.

6. Ethical aspect

Ethical requirements have been defined in WP9 and in particular in deliverables

D.9.1 POPD - Requirement No. 1

D.9.2 NEC - Requirement No. 2

All the CITYSOLAR research, including data management aspects, in compliance with H2020 ethical standards, in compliance with the ethical principles (see Article 34) and confidentiality (see Article 36 as set out in the Grant Agreement).

7. Conclusion

This documents describes the man principles and guidelines of Data Management for the CITYSOLAR project. Being a living document, it will be updated throughout the project lifetime.

Further revisions of the Data Management Plan will include, as the case, updating of online research data repository where data are collected and shared, and the data, the description of dataset, and research data gradually generated and collected.

END OF DOCUMENT