D6.7 Data Management Plan

WP6

Lead Partner: TNO

Partner Contributors: EURAC, IMEC, SUPSI

Dissemination Level: Public

Deliverable due date: M6 Actual submission date: M6

Deliverable Version: V1.0

Funded by the European Union

Project funded by
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

Project Acronym	MC2.0
Project Title	Mass customization 2.0 for Integrated PV
Grant Agreement n°	101096139
Call	HORIZON-CL5-2022-D3-01
Topic	HORIZON-CL5-2022-D3-01-03 Advanced manufacturing of Integrated PV
Starting Date	1 January 2023
Duration	38 months

Document history

Version	Date	Comments
V0.1	2023-05-10	Start of first draft DMP based on template "EU Grants: Data Management Template (HE):V1.0 – 05.05.2021", composed by TNO
V0.2	2023-05-25	Version based on feedback from WP6 (EURAC, IMEC, SUPSI) and presented for review from SC and reviewer according to table from PQP.
V1.0	2023-06-23	Final version ready for upload
V1.1	2023-06-30	Typo at page 6 correct. This final version is uploaded in the HE-portal

Summary

This data management plan (DMP) ensures that the data associated with the MC2.0-project will be stored, accessible and distributed in accordance with the FAIR principles and HE guidelines. This is the first version of the plan that will be updated over the course of the project. This document will serve as a guide for the partners for the implementation of data storage and handling. This DMP will always be kept up-to-date with detailed descriptions of the data handling and storage details over the course of the project and therefor will be a live document maintained on the MC2.0 sharepoint.

Contents

Abbreviations and Acronyms

[BOM] Bill of Materials
[CA] Consortium Agreement
[DMP] Data Management Plan
[DoA] Description of Action
[DOI] Digital Object Identifier
[DSSA] Data and Sample Sharing Agreement
[EC] European Commission
[FAIR] Findable Accessible Interoperable Reusable
[GA] Grant Agreement
[Gen.As] General Assembly
[IB] Industrial Board
[IPV] Integrated Photovoltaic
[MC] Mass Customization
[PC] Project Coordinator
[PQP] Project Quality Plan
[SC] Steering Committee
[WP] Work Package

1 Introduction

The MC2.0-project will demonstrate how automated manufacturing of Integrated PV (IPV) elements using the Mass Customization (MC) approach can deliver a cost breakthrough in IPV facades, IPV windows and IPV roofs. To realize this ambition, the MC2.0 consortium brings together experts and companies on materials for PV laminates (including PV cells), on manufacturing of PV laminates, on manufacturing of IPV products and on market and application of IPV products. In the Project Quality Plan (PQP) and on the website www.mc2dot0.eu many details can be found on the precise results and deliverables of the project. Various types of data will be generated by many project partners. This Data Management Plan (DMP) explains how the project deals with that data.

2 Data Summary

Project partners might consider re-use any existing data of their own for the purpose of the project goals. In case of newly produced semi-fabricates or IPV-prototypes with e.g. a new BOM or a new manufacturing process as developed within MC2.0, it is likely that existing data on older semi-fabricates or IPV-prototypes is of limited value to the project. In that case new measurements will be performed on the new semi-fabricates and new IPV-prototypes.

Many different types and formats of data will be generated by the project. The quality and reliability of MC2.0 semi-fabricates or IPV-prototypes will be determined by the many different testing procedures as described in the requirements document (D1.1). On top of that the Environmental impact of the semi-fabricates is quantified. This impact is quantified through a.o. Life Cycle Assessment (LCA) that will generate data. Also the economic aspects of the newly developed MC2.0-process is analysed based on input data and models.

The purpose of the generated data is directly related to the objective of the project to demonstrate functional and reliable mass-customized semi-fabricates. And integrate those in functional and reliable IPV-prototypes. With a minimum impact on the Environment. And with a solid business case in the market of the building industry.

The expected size of the data that will be generated by the project can vary from small single measurements on small batches, to larger sets of multiple test results (from the standards described in D1.1). However, huge data sets with millions of data points are not expected because within MC2.0 there is no extensive outdoor testing of e.g. a full year. In conclusion, it is foreseen that the amount of data within MC2.0 stays manageable.

The origin/provenance of the data will be produced by measurement equipment used by the project partners to quantify functionality and reliability as described above. Moreover, it could be that some project partners would build a new measurement device that would produce data in a format that is not known yet. The data related to economical and LCA-aspects will not come from measurements, but from datasheets, calculation methods and models. As they offer detailed insight into the bills of material and manufacturing processes of the commercial IPV products developed in MC2.0, they cannot be published, considering the high value of this data.

The data of MC2.0 might be useful ('data utility') outside our project by other IPV-integrator companies. These companies might like to have measurement reports with data sets. However, due to confidentiality not all data can be shared public. In those cases, the knowledge from the analysis of that data can be shared public, without sharing the underlying data, contributing in such a way to the public report.

3 FAIR data

More in detail the so-called FAIR-principle will be used, which is an acronym for the following four principles Findable, Accessible, Interoperable and Reusable. Whenever data is shared, Project Coordinator (PC) TNO will remind the project partners to respect the FAIR-principle. The General Assemblies (Gen.As.) that are planned each half year with the full consortium are a perfect moment to check whether the partners use this FAIR-principle.

At the moment of writing this DMP, it is not clear into detail, which data will be produced by which partner. However, trying to use the FAIR-principle already now, the following remarks can be made.

3.1 Making data findable, including provisions for metadata

Data will have a persistent identifier.

Rich metadata will be provided to allow discovery. E.g. in case a newly produced laminate is produced at the pilot line of TNO, the actual settings and the used BOM is mentioned in the metadata, as long as this is not conflicting the commercial interest of the involved partners.

It is no problem to provide this rich metadata, as it is normal practice of the researchers and research assistants and technician to note down in their logbook which experiments and test runs were performed on which samples (BOM and identifier). Without this practice it would not be possible at all to analyse data.

Search keywords are present in the metadata to optimize the possibility for discovery. E.g. project coordinator TNO is using an internal database to search efficiently through measurement data. Due to confidentiality, not all content of this database can be made public.

Data can be searched, sorted or retrieved via the metadata.

3.2 Making data accessible

3.2.1 Repository

Public data will be deposited on Zenodo (www.zenodo.org). Zenodo is an online digital academic repository for scientific publications and data, sponsored by the European Commission, and found to be a trusted repository by the international community. The website is hosted by co-founder CERN. The European Commission is involved through the OpenAIRE project. Zenodo offers the possibility to upload, store and make available datasets free of charge. Each dataset is provided with a Digital Object Identifier (DOI) so that it is easy to link to the data.

The dataset(s) must be classified as open or restricted. At this stage of the project, it is not clear yet which parts of the data will be open and which parts will be restricted. In the license field of Zenodo one can choose for:

- CC-BY-ND
  This license lets others reuse the work for any purpose, including commercially; however, it cannot be shared with others in adapted form, and credit must be provided to you.
- CC-BY-NC-ND
  This license lets others reuse the work for any purpose, however excluding commercial use of the work; and it cannot be shared with others in adapted form, and credit must be provided to you.
- Restricted Access;
  Users may deposit restricted files with the ability to share access with others if certain requirements are met. These files will not be made publicly available and sharing will be made possible only by the approval of depositor of the original file.

3.2.2 Data

Data will be made available in the Zenodo-repository as mentioned in the previous paragraph. However, one should understand that not all data can be made openly available. Certain datasets cannot be shared (or need to be shared under restricted access conditions) because it would reveal the precise materials or precise processes used by one (or multiple) project partners. If such precise and detailed information is available in public, competitors could start producing exactly the same materials without having the costs of all the research and development that the MC2.0-partner(s) put into it.

The project partners will make sure that the data that is accessible does not need a special access protocol.

In case of public data, the identity of the person accessing the data does not need to be ascertained.

There is no need for a data access committee. The judgement whether data is sensitive or public can be made by the General Assembly. If appropriate, the steering committee will prepare all context to smoothen the process of this judgement by the Gen.As.

3.2.3 Metadata

Metadata of deposited data will be open under Creative Commons conditions (CC BY). By using Zenodo this is automatically respected.

The metadata will be composed with enough information to enable the user to find and access the data efficiently.

The data will remain available and findable for at least 5 years after the project MC2.0 has been officially closed.

There will no special software needed to access or read the public data provided by MC2.0. To be more explicit regarding foreseen data types:

- Reports: .PDF-files readable in any (free) PDF-reader
- Data: .CSV-files readable in any (free) text reader. Advised can be the very robust and free software Notepad++ (https://notepad-plus-plus.org/downloads/)
- Images: .JPG-files or .PNG-files readable in any (free) picture program. Advised can be the established and free software IrfanView (https://www.irfanview.com/)
- Calculation methods/models: .ODF-files , which is the OpenDocument version of Excel; can be read and written by the commonly used commercial Microsoft Excel, but also by the freeware OpenOffice (https://www.openoffice.org/)

3.3 Making data interoperable

Due to the nature of the data in MC2.0, we will follow the vocabularies, standards, formats or methodologies of the following international standards:

3.3.1 International BIPV-standards

EN 50583-1:2016	Photovoltaics in buildings - Part 1: BIPV modules
EN 50583-2:2016	Photovoltaics in buildings - Part 2: BIPV systems
IEC 63092-1:2020	Photovoltaics in buildings - Part 1: Requirements for building-integrated photovoltaic modules
IEC 63092-2:2020	Photovoltaics in buildings - Part 1: Requirements for building-integrated photovoltaic systems

3.3.2 International PV-standards

IEC 61730-1	Photovoltaic (PV) module safety qualification - Part 1: Requirements for construction
IEC 61730-2	Photovoltaic (PV) module safety qualification - Part 2: Requirements for testing
EN 61215	Terrestrial photovoltaic (PV) modules - Design qualification and type approval - Part 1: Test requirements
EN 61215-2	Terrestrial photovoltaic (PV) modules - Design qualification and type approval - Part 2: Test procedures
EN 61853 –X series	Photovoltaic (PV) module performance testing and energy rating

3.3.3 Other relevant standards

3.3.3.1 Construction Products Regulation (CPR)

Many standards with respect to 'glass in buildings' which is relevant for two IPV-products in the MC2.0-project, which are 'Venetian blinds in Insulation Glazing Unit (IGU)' and 'Transparent Glass'. If appropriate, we will refer to specific standards when using wording/vocabulary, definitions, formats, etc… that is used by the relevant standard.

With respect to the IPV-product 'Façade', European Assessment Document (EAD) 090062-00-0404 "Kits for External Wall Claddings Mechanically Fixed"[1] is taken as a good reference for using wording/vocabulary, definitions and formats.

[1] https://www.eota.eu/download?file=/2016/16-09-0062/ead%20for%20ojeu/ead%20090062-00-0404_ojeu2018.pdf

3.3.3.2 Low Voltage Directive (LVD)

Standards that are falling within the scope of Directive 2014/35/EU (Low Voltage Directive - LVD), of which most important for our project are:

IEC 60335-1	Household and similar electrical appliances - Safety - Part 1: General requirements
IEC 60335-2-97	Household and similar electrical appliances - Safety - Part 2-97: Particular requirements for drives for rolling shutters, awnings, blinds and similar equipment

Many standards with respect to 'glass in buildings' which is relevant for 2 IPV-products in the MC2.0-project, which are 'Venetian blinds in Insulation Glazing Unit (IGU)' and 'Transparent Glass'.

3.3.3.3 Electromagnetic Compatibility (EMC)

Standards that are falling within the scope of Directive 2014/30/EU (Electromagnetic Compatibility—EMC), of which most important for our project are:

EN IEC 55014-1	Electromagnetic compatibility — Requirements for household appliances, electric tools and similar apparatus — Part 1: Emission
EN IEC 55014-2	Electromagnetic compatibility — Requirements for household appliances, electric tools and similar apparatus — Part 2: Immunity — Product family standard

3.3.4 Other documents relevant to the project

To be foreseen relevant at the moment are the following documents:

- The IEA PVPS Task15 'Enabling Framework for the Development of BIPV' is already active for about 10 years. In this period, many public available reports have been published; see: https://iea-pvps.org/research-tasks/enabling-framework-for-the-development-of-bipv/

  Thanks to this work, many definitions, wording, methodologies have now an established (and international recognized) status. It is obvious that project MC2.0 will comply as much as possible to this context.

Please note that this list of relevant documents could be updated during the project.

The used vocabulary (and ontologies if appropriate) will be explained in detail in case it is not evidently clear to a more general public, and if all mentioned documents in this paragraph do not provide enough background explanation to the matter.

In case reference to other data (outside MC2.0-project) is meaningful, this will be done in such a way that also that data is as easy accessible as possible. Providing enough detailed reference and metadata to ease the process of finding and downloading that external, public, data.

3.4 Increase data re-use

In case of e.g. data that is provided as Column Separated Value (CSV)-files, the header in the first row normally contains very condensed (sensor)names as column label. In such a case an additional readme-file will be added to the same folder as the data (or included in the ZIP-file) for the purpose of easy data re-use.

In case of e.g. and Excel workbook that contains multiple tabs/sheets, a first explanatory tab/sheet is added as kind of manual for using the full Excel workbook.

The provenance of the data will be documented using the appropriate standards, as mentioned explicitly in the paragraph 3.3 ‘Making data interoperable’

In case of straightforward data files, the Project Coordinator (TNO) will review the quality of the data before uploading/publishing in the repository Zenodo. In case of more complex data, this review will be done in close collaboration with the relevant WP-leader.

4 Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

In T3.1 of the WP3, we will analyse data collected within the consortium regarding the cost of the IPV semi-finished and end products developed by the partners of MC2.0. These data will be used as key inputs for the studies related to the validation of a cost calculation tool for the IPV cost based on MC2.0 and a sensitivity analysis showing the sensitivity of cost to the most determining parameters. The use of the modelling tool will be made publicly available, not the source code nor the input data, as also mentioned in the Grant Agreement under section 1.2.11. The research outputs of T3.1 thus refer also to the results, outputs, findings and guidelines published in the two deliverables produced: D3.1 (public) and D3.2 (sensitive). The level of detail in these reports will aim to avoid any disclosure of commercially sensitive information from MC2.0 partners. At the same time, intermediate results will be presented, and some data will be shared only in highly aggregated form, to allow for the results in these deliverables to be reproduced.

In T3.2 and T3.3 of the WP3, we will analyse data collected within the consortium regarding the IPV semi-finished and end products developed by the partners of MC2.0. These data will be use as key inputs for the studies related to the market entry roadmap, the definition and the optimization of the value chain, the identification of the strategies to engage the stakeholders and the analysis of new drivers for business cases, which will be made publicly available. The research outputs of T3.2 and T3.3 thus refer to the results, outputs, findings and guidelines published in the three public deliverables produced (D3.3, D3.4 and D3.5). The level of detail in these reports will aim to avoid any disclosure of commercially sensitive information from MC2.0 partners. At the same time, intermediate results will be presented, and some data will be shared only in highly aggregated form, to allow for the results in these deliverables to be reproduced.

In WP5, we will analyse data collected within the consortium regarding the manufacturing of the commercial IPV products developed by the partners of MC2.0. These data will be used as key inputs for the studies related to LCA, supply-chain optimization, R-ladder sustainability considerations, and the product-level circularity metric analyses. They offer very detailed insight into e.g., the BOM of the IPV products and their manufacturing processes, and as such, these input data cannot be shared publicly, as also mentioned in the Grant Agreement under section 1.2.11. The research outputs of WP5 thus refer to the results, outputs, findings and guidelines published in the four public deliverables produced in WP5. The level of detail in these reports will aim to avoid any disclosure of commercially sensitive information from MC2.0 partners. At the same time, intermediate results will be presented, and some data will be shared only in highly aggregated form, to allow for the results in these deliverables to be reproduced.

5 Allocation of resources

All partners will be responsible themselves for the data management of their own work (Task, Deliverable, co-authoring, etc…) in the MC2.0-project. Therefore, potential costs of that partner specific data management will be brought in by the partner self.

As mentioned, there is no big data foreseen in our project. Therefore the costs of making data available (in FAIR-principle) are rather small. If during the project, this changes into the direction of larger amounts of data, then we will consider declaring the costs (or parts that are eligible) as further specified in our Grant Agreement.

The preservation is ensured for at least 5 years after the project closure. It is not needed to reserve resources from within MC2.0-budget to accomplish this, because it is part of normal practice in case TNO is participating into a H2020 or Horizon Europe project.

6 Data security

For all internally used data a very professional (Microsoft) Sharepoint environment is in place already, made available by TNO from start of the project. A two-step verification process makes this a very secure platform for sharing data. This allows partners to work and share sensitive data within the project without having to worry about security or data recovery in case of accidental deleting data. Moreover, the version management of all files (reports, data, metadata, logbooks, pictures, etc…) is superb.

As mentioned above, data to be shared public will be stored in Zenodo. Because Zenodo is supported by the EC, we believe this is one of the most trusted repositories available on the internet.

7 Ethics

It is unlikely that in the MC2.0-project ethics will have a negative impact on data sharing (maybe a positive impact, see first paragraph below), given the nature of the project in which we develop a new manufacturing process for semi-fabricates with solar energy (PV) functionality to be integrated in building products. Nevertheless, the Grant Agreement does contain an Ethics Self-Assessment which has two positive answers regarding ethical aspects, related specific to Environment, Health and Safety. Addressing each topic in the following paragraphs:

7.1.1 Potentially causing harm to environment

The full self-assessment question is: 'Does this activity involve the use of substances or processes that may cause harm to the environment, to animals or plants (during the implementation of the activity or further to the use of the results, as a possible impact) ?'

The effect of the semi-fabricate on the environment (which includes animals and plants indirectly) is well described by Life Cycle Assessment (LCA). This is the core of WP5. By quantifying the LCA of BOM and processes, potential users of semi-fabricates (as developed within our project) can make a much better choice with respect to lower impact on the environment! It is our expectation that these potential users (integrator companies) and end-users of the IPV-product are willing to pay more for products that cause less harm to the environment (with a lower LCA). Therefore, we can conclude that this ethical aspect is an important part of the project. WP5 will produce several public deliverables. The detailed data underpinning the LCA analyses, and the other analyses in WP5 based on LCA results, are too commercially sensitive to be shared. We will strive to publish intermediate LCA results (highly) aggregated and generic input data to offer more insights in the results obtained. As an example, we can show how switching from one polymer to another, reducing material consumption of essential and critical PV materials will affect the overall environmental impact of IPV products.

7.1.2 Potentially causing harm to humans

The full self-assessment question is: 'Does this activity involve the use of substances or processes that may cause harm to humans, including those performing the activity (during the implementation of the activity or further to the use of the results, as a possible impact) ?'

During the manufacturing ('processes') of the semi-fabricates, various materials ('substances') are used. If not handled with care they could harm the persons working ('during the implementation of the activity') in the project. In WP2, work will be done at the premises of many MC2.0-partners that may involve the use of substances that require controlled disposal, and equipment that require adequate safety measurements. These partners have the necessary safety procedures in place and a prevention officer appointed, in order to be compliant with the legal requirements on environment, health and safety.

There is no foreseen negative aspect of data sharing coming out of this ethical aspect.

Finally for completeness, it is noted that there is no personal data at all involved in MC2.0 . Hence also no data sharing or long term preservation of such data.

8 Data and Sample Sharing Agreement (DSSA)

On top of the data and metadata for sharing outside the project, there is also data and samples that will be shared between the MC2.0 project partners within the project internally. A template for this so-called Data and Sample Sharing Agreement (DSSA) is available at the project SharePoint for project partners only.

The DSSA contains at first the names of the partners involved in the sharing, and the date on which the agreement will be signed. Contact details of the Data (or Sample) provider are given in the next section. Also the contact details of the Data (or Sample) receiver are specified; this can be more persons from multiple project partners. Thereafter, the WPx (or specific Task x.y) is ticked, multiple ticks are possible.

Thereafter, in case of data, the type of this data is specified. It contains metadata and more elaborate descriptive texts if necessary to understand the data. Moreover, the confidentiality level of the data is provided. Data format and location of the data (URL or server name) is specified.

In case of samples, the sample is described. The metadata and more elaborate descriptive text about the sample is provided.

Regardless whether data or samples are shared, the final chapter of the template specifies the objective of the sharing from both the view point of the receiver(s) and the providing party. The (or each) receiver writes down what he/she likes to achieve with the data or sample(s). The provider writes down what he/she likes to achieve in line to which objective of the DoA (e.g. mentioning/referring once more to specific WPx and/or Task Tx.y). In case the data or sample(s) needs to be disseminated outside the project by the receiver, this has to be mentioned by the provider.

The Data and Sample Sharing Agreement (DSSA) is finally signed by provider and receiver(s).