



# POCITYF

## Data Management Plan

D11.8: Data Management Plan

WP11, T11.4

Authors: Paul Kessler (EDPL), José Miguel Costa (EDPL)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 864400.

## Technical references

Project Acronym	POCITYF
Project Title	A POSitive Energy CITY Transformation Framework
Project Coordinator	João Gonçalo Maciel EDPL JoaoGoncalo.Maciel@edp.com
Project Duration	60 months (from October 2019 - to September 2024)

Deliverable No.	D11.8
Dissemination level*	PU
Work Package	WP11 - Project Management
Task	T11.4 - Data and Ethics Management
Lead beneficiary	1 (EDPL)
Contributing beneficiary/ies	1 (EDPL), 3 (UNINOVA), 15 (GA), 30 (E@W), 38 (CERTH), 39 (VTT)
Due date of deliverable	31 March 2020
Actual submission date	30 March 2020

\* PU = Public

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)

v	Date	Beneficiary	Author
0.1	04/03/2020	EDPL	Paul Kessler
0.2	12/03/2020	EDPL	José Miguel Costa
0.3	17/03/2020	EDPL	Paul Kessler, José Miguel Costa
1.0	30/03/2020	UNINOVA, VTT	EDPL, including comments by two reviewers



## Executive Summary

The present deliverable D11.8 - *Data Management Plan* - establishes the framework under which POCITYF consortium will monitor/generate, process and collect data during project's demonstration activities and citizen engagement actions. D11.8 tackles as well how the data will be exploited or made accessible for verification and re-use and how the data will be curated and preserved after the project is completed. Current document corresponds to the first version of POCITYF Data Management Plan (DMP), therefore composed by preliminary information and frameworks that will be followed, in the near future, to collect and characterise project's data sets (which will be presented in next iterations of the current deliverable, to be submitted by M24 - September 2021 - and M48 - September 2023).

Moreover, the DMP details the standards that will be used, how the research data will be preserved and what data sets will be published for verification or reuse as open access. The DMP covers the entire research data life cycle and must be consistent with exploitation and Intellectual Property Rights (IPR) requirements. Hence, research data linked to exploitable results will not be put into the open domain if they compromise the consortium's commercialisation prospects or have inadequate protection, which is also a H2020 obligation. Particularly, sensitive data provided by consortium partners for the demonstration scenarios and personal data of the users will be kept strictly confidential and anonymized, i.e., secured in a way to maintain compliance with General Data Protection Regulation (GDPR).

Information herein presented goes beyond POCITYF's Grant and Consortium Agreement, stating the progress that was achieved within the first six months of the project in what concerns Data Management methodologies.

This document has been built based on the Horizon 2020 FAIR DMP template (Version: 26 July 2016), which promotes the principle "as open as possible, as closed as necessary" together with the project ambition of "Open by Default". POCITYF is, indeed, part of the European Commission (EC) Open Research Data Pilot (ORDP), adhering to all of its conditions<sup>1</sup>.

The first section of the document presents the deliverable and contextualises it within POCITYF project. The DMP contains, in the second section a description for the characterisation of the data sets to be generated by the project and how the data will be generated/acquired, processed and archived. The third section of this deliverable is devoted to ethics, privacy and security considerations to ensure compliance with European and national legislation as well as directives relevant to the country where the data collection is/will taking place. This includes that the collection, processing and transmission of personal data will be analysed under principles of: (a) the recently published GDPR (Regulation (EU) 2016/679)<sup>90</sup>; (b) the Universal Declaration of Human Rights and the Convention 108 for the Protection of Individuals with regard to Automatic Processing of Personal Data; and (c) the national laws applying its provisions. All beneficiaries are informed about consenting, data processing, data security, and the pertinent regulations such as GDPR which are addressed in the third section. The DMP is concluded, in section 4, by outlining the guidelines how the data is archived and preserved for long-term beyond the project's conclusion.

---

<sup>1</sup> <https://www.openaire.eu/what-is-the-open-research-data-pilot>



## Table of contents

1	Introduction .....	6
1.1	Objectives and Scope .....	6
1.2	Relation to other activities .....	6
1.3	Structure of the deliverable.....	6
2	Data Set Description Methodology .....	7
3	Ethics, Privacy & Security Considerations .....	10
3.1	Ethics Requirements .....	10
3.2	Data Privacy and Personal Data.....	11
3.3	Data Security .....	12
4	Data Management, Sharing & Open Access .....	14
4.1	Data Handling and Management .....	15
4.2	Sharing and Open Access.....	17
4.3	Document Management .....	17
5	Data Curation & Preservation .....	18
6	Conclusions .....	19

### List of Tables

Table 1	Template for characterisation and description of the data sets .....	9
Table 2	Template to specify the physical instances of the data set .....	9
Table 3	- Template for data handling and management summary.....	15
Table 4	- Tool to characterise data curation and preservation .....	18

### List of Figures

Figure 1	Template for Responsibility Assignment Matrix to track responsibilities within POCITYF Project .....	16
Figure 2	- Data Management and IPR Protection .....	17



## Acronyms

Acronym	Meaning
DMP	Data Management Plan
DPO	Data Protection Officer
EC	European Commission
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
ORDP	Open Research Data Pilot
PEB	Positive Energy Block



# 1 Introduction

---

## 1.1 Objectives and Scope

The current deliverable D11.8 - Data Management Plan - constitutes a guideline complying to European and national legislation on the acquiring, handling, processing and archiving of data generated during POCITYF project and beyond the project's end, being updated by Month 24 and 48.

D11.8 allows for a consortium-wide alignment around the project's data management processes. The objective is to promote the principles of: (a) the recently published GDPR (Regulation (EU) 2016/679)90; (b) the Universal Declaration of Human Rights and the Convention 108 for the Protection of Individuals according to Automatic Processing of Personal Data; and (c) the national laws applying its provisions.

## 1.2 Relation to other activities

T11.4, and, subsequently, its respective deliverables, have a transversal actuation throughout POCITYF, providing a guideline on how partners should generate, acquire, handle, share and curate research data during all activities carried out within and beyond POCITYF, especially important for the demonstration- and replication-related WPs (WP6, WP7 and WP8), and for the one dealing with citizen engagement activities (WP4). The DMP will allow that all consortium is aware of the practices that shall be followed within the conduction of those activities.

## 1.3 Structure of the deliverable

Deliverable 11.8 is structures as follows:

- I. Section 2 - Data Set Description Methodology: presentation of the methodology that will be used to identify and characterise the different data sets that will be generated or acquired, their origin, purpose and its relation to the objectives of the project, as well as the type and format of the data. Furthermore, this methodology will describe the data sets in terms of communication protocols and involved hardware and software;
- II. Section 3 - Ethics, Privacy & Security Considerations: constituting ethical, privacy or legal issues that can have an impact on data sharing and handling. The processes for recruitment of participants are outlined and data privacy and issues regarding personal data are presented. Furthermore, the role of a Data Protection Officer (DPO), the one who ensures adherence to GDPR, is explained. Complementing these general efforts, city- and partner-specific processes on privacy and security will be elaborated in the future versions;
- III. Section 4 - Data Management, Sharing & Open Access: contextualising the data handling descriptions, procedures and access rights to publications and data sets, the principle of Open Research Data and Open City Data as well as the document management system;
- IV. Section 5 - Data Archiving & Preservation: presenting the procedures and processes on the curation and preservation of the generated data sets within and beyond the project's life cycle.
- V. Section 6 - Conclusion: addressing the main highlights of the deliverable.



## 2 Data Set Description Methodology

Within this chapter, the methodology to collect the data to be generated or acquired, the origin, nature and scale and to whom it could be useful, and whether it underpins a scientific publication or not, is described. In addition, information and guidelines on the reuse of the data are stated.

As the project is still in its infancy, it is still not possible to list the various and final data that will be generated/acquired during the project, since relevant activities towards data sets' identification are still being carried out in *WP1 - POCITYF Smart City Framework Towards an Integrated Development*, *WP2 - Setting up, Planning and Execution of Performance Monitoring Activities*, *WP4 - Citizens Engagement and Open Innovation Activities*, *WP6 - Evora Lighthouse City demonstration activities* and *WP7 - Alkmaar Lighthouse City demonstration activities*. A first comprehensive identification of the data sets to be generated/acquired, with all evaluation metrics and formulas derived, will be available in *D2.3 - Monitoring Data & Infrastructure Definition, expanding to Cultural Heritage cases* (M12, September 2020), and included in the updated versions of the DMP in month 24 and 48. The next versions of the DMP will, as well, characterise the identified data sets as defined in the chapter on "Data Management, Sharing & Open Access".

Table 1 presents one of the tools that will be used to characterise and describe the data sets that are expected to be generated/acquired by the project. The includes standardised items under which every data set will be classified. That table will be stored, as an excel file, in POCITYF's collaborative online tool.

The **1<sup>st</sup> column** identifies the ID of the data set. The **2<sup>nd</sup> column** indicates the 'Title of the data set'. The title of the data set shall be self-explanatory regarding the nature and purpose of the data set.

In the **3<sup>rd</sup> column**, the 'Filename' under which the data set will be stored/archived/curated is specified.

The **4<sup>th</sup> column** specifies the WPs for which the data set is generated or acquired.

In the **5<sup>th</sup> column**, the city corresponding to the data set is specified. Here, it shall be distinguished between *Lighthouse* and *Fellow* cities.

The **6<sup>th</sup> column** indicates the consortium partner(s) and/or contact person(s) responsible for the data set.

In the **7<sup>th</sup> column**, the 'Relation to the project objective' is detailed. The relation adheres to the subordinate 10 specific objectives of POCITYF Project:

- **Objective 1:** Demonstrate solutions at building and district level that enable the increase of energy self-consumption, energy savings and high share of locally produced renewable energy - leading to energy positive districts, located in mixed use urban districts including that of cultural heritage ones;
- **Objective 2:** Demonstrate P2P energy management and storage solutions supporting grid flexibility and curtailment reduction;
- **Objective 3:** Demonstrate the integration of electro-mobility solutions as an enabler to grid flexibility;
- **Objective 4:** Demonstrate the integration of the latest generation of ICT solutions within existing city platforms over open and standardized interfaces enabling data exchange/monitoring for the development of new innovative services;



- **Objective 5:** Demonstrate active citizen engagement services and solutions providing an open innovation ecosystem for citizens to participate in co-creation, decision making, planning and problem solving within the Smart Cities;
- **Objective 6:** Design bankable business models and robust investment concepts that consider the whole PED lifecycle, and test them to reduce technical and financial risks for investors guaranteeing replicability at EU scale;
- **Objective 7:** Strengthening the links and cooperative innovation with other SCC Projects in a large number of Member States with a wide coverage of cities ranging in size, geography, climatic zones and economic situations;
- **Objective 8:** Measure, validate and evaluate the demonstration results after a 2-year large-scale demonstration at district level within 2 highly innovative EU Cities;
- **Objective 9:** Demonstrate solutions that facilitate the energy transition of historical and cultural heritage city areas;
- **Objective 10:** Identify related regulatory barriers, legal aspects and data security/protection and propose practical recommendations on how to overcome them.

In the **8<sup>th</sup> column**, the ‘Data type’ of the data set is indicated. The expected data types include but are not limited to:

- Integers
- Booleans
- Characters
- Floating-point numbers
- Alphanumeric strings
- Other (to be specified)

The **9<sup>th</sup> column** specifies the ‘Data format’ of the data set. The expected data formats include but are not limited to:

- ASCII text-formatted data (TXT)
- CAD data (DWG)
- Comma-separated values (CSV)
- dBase (DBF)
- eXtensible Mark-up Language (XML)
- Tab-delimited file (TAB)
- Geospatial open data based upon JavaScript Object Notation (GeoJSON)
- Geo-referenced TIFF (TIF, TFW)
- Hypertext Markup Language (HTML)
- Keyhole Markup Language (KML)
- MS Word (DOC/DOCX)
- MS Excel (XLS/XLSX)
- MS Access (MDB/ACCDB)
- OpenDocument Spreadsheet (ODS)
- Open Document Text (ODT)
- Rich Text Format (RTF)
- SPSS portable format (POR)
- Other (to be specified)





The **10<sup>th</sup> column** details the ‘Description of the data set’<sup>2</sup>. The description of the data set shall be clear and precise so that each data set can be distinguished from the others by the information enclosed.

The **11<sup>th</sup> column** indicates the ‘Origin of the data’. The origin of the data set shall indicate how the data was collected and what kind of equipment was used to obtain the data set. In the **12<sup>th</sup> column**, it is specified whether the data is personal (traced back to an individual entity) or not. This characteristic shall help to identify sensitive data in terms of GDPR as well as requirements emanating from ethical, data privacy and security considerations at the point where it is generated/acquired.

**Table 1 Template for characterisation and description of the data sets**

ID	Title of data set	File Name	WPs included	City	Responsible Partner	Relation to project objective	Data type	Data format	Description of data set	Origin of the data	Personal data

Complementing the sheer characterisation of the data sets, a second tool is used in form of a second table (Table 2) to specify the physical instances of the data set (building, origin equipment/tool, repository equipment/tool as well as communication between origin and repository and access rights).

**Table 2 Template to specify the physical instances of the data set**

ID	Title of data set	PEB	Building	Origin equipment/tool	Repository equipment/tool	Communication between origin and repository	Access (confidential or public)

This table will further contextualise data sets within POCITYF universe, directly linking them with demo sites and hardware and software.

The **3<sup>rd</sup> column** indicates the Positive Energy Block (PEB) where the data sets will be generated and acquired, whereas the **4<sup>th</sup> column** the specific building within that PEB/s. The **5<sup>th</sup> column** identifies the hardware (equipment) and/or software that generates that data set. The **6<sup>th</sup> column**, in its turn, points to the system that will store the data set. Following this characterisation, **7<sup>th</sup> column** renders information on the communication protocol that will be used between the source and the repository of the data set. The last one - **8<sup>th</sup> column** - characterises the data set in terms of confidentiality.

<sup>2</sup> It may be sufficient to introduce keywords in context with the data set description. The keywords used shall still allow to identify each data set individually. The characterisation of the data sets based on keywords may allow the use of database tools, such as SQL to organise, maintain and access the data sets during POCITYF project.



## 3 Ethics, Privacy & Security Considerations

The project involves carrying out data collection (in the context of the piloting and validation phase) and a set of large-scale validation tests to assess the technology and effectiveness of the proposed framework in real life conditions. Hereby, the DMP specifies ethics, privacy and security considerations to comply with all European and national legislation and directives relevant to the country where the data collections are taking place and, moreover, where the data set is curated and preserved. This work is complemented by deliverables D12.1, D12.2 and D12.3 - submitted in November 2019 - and D11.11, on “Ethical Monitoring and GDPR Conformation Plans”.

As POCITYF project includes multiple cities with diverse characteristics and specific and local objectives, the nature of the collected and processed data is heterogenous. To cope with the different nature of the data sets, methodologies to address ethical requirements as well as data privacy and security are of utmost importance to establish a general framework in this heterogenous environment. The following three sections will detail the different methodologies and principles used to address a number of privacy and data protection issues raised by the project’s activities.

### 3.1 Ethics Requirements

The internal ethical agenda of POCITYF follows the guidelines of various expert communities in the field of data ethics (e.g. the European group on ethics in science and new technologies to the European Commission). The ethical agenda includes a code of conduct to specify correct behaviour and corresponding principles in relation to data collection and processing. The following six principles are contemplated and addressed within POCITYF ethical code of conduct:

- I. **Ownership** - Who is the owner of the data?
- II. **Transaction transparency** - What access is given to the owner and how transparent is the access (a transparent access for the use of the individual’s personal data has to be established, the individuals should have full and transparent access to the algorithms used to and generate and aggregate the data sets)?
- III. **Consent** - Which individuals or other entities need to give consent to use the data (individuals or other entities shall be explicitly informed of what personal data moves to whom, when and for what purpose from the owner of the data)?
- IV. **Privacy** - What efforts and measures are in place to ensure data privacy (in terms of data processing, the project partners are obliged to invest all reasonable effort to preserve the privacy of the individual)?
- V. **Currency** - If applicable, what is the financial value of the personal data and how is that communicated with the owner of the data (individuals shall be explicitly informed of any financial transactions resulting from there data)?
- VI. **Openness** - how much of the aggregate data sets are freely available (if applicable, under adherence to point 1. - 5. aggregate data should be freely available)?

Following these principles shall maintain the responsible and sustainable use of the data generated and processed in POCITYF. The ethical code of conduct shall further serve as a complement to the mere compliance with data protection laws and current regulations (specified in the next section). In fact, the code of conduct of POCITYF shall reflect a principle that promotes honesty and genuine transparency in data management.



Besides the afore-mentioned 6 principles, any data generation/acquiring activity involving humans will be strictly held confidential at any time of the research. This entails the following practical actions each project partner shall obey interacting with volunteers/individual stakeholders:

- I. Explicitly inform all participants about the first 6 principles of the code of conduct and give them the opportunity to provide their consent to the data management process (Why the data is being collected? How it is going to be used? How long it will be stored? How it can be amended by the individual concerned?);
- II. Ensure the volunteers that no personal or sensitive data will be centrally stored. In addition, the volunteers shall be informed about the data security measures in place (see subsection 2.3 Data Security) to avoid potential identification of individuals.

The briefing and information provided shall be conducted in the volunteers' native language. Additional information provided shall include a (i) written description of the project and its goals, (ii) the project's progress and the related testing, evaluation and validation procedures (respecting POCITYF confidential information), and (iii) information on unrestricted disclaimer rights on their agreement.

To oversee the adherence of the POCITYF consortium to the ethical code of conduct and to establish an entity communicating with the European Commission, an Ethics Board was appointed, as stated in the Grant Agreement. The Ethics Board, chaired by POCITYF Ethics Manager, includes one responsible person for each partner, whom is responsible that the activities follow the ethics, privacy and security considerations detailed in this chapter. Partners with reduced effort and limited scope had the possibility to assign a proxy to represent them in this board, i.e., someone deeply involved, for instance, in the demonstration activities and that has a comprehensive view on the conducted and planned work.

### 3.2 Data Privacy and Personal Data

POCITYF project will engage multiple cities collecting multiple sets of data and carrying out various measurements of data per city. Hence, the project stakeholders (internal and external) as well as the citizens affected by the data collection will vary from task to task. The collection, processing and transmission of personal data will be analysed under principles of:

- a) The recently published GDPR (Regulation (EU) 2016/679)90;
- b) The Universal Declaration of Human Rights and the Convention 108 for the Protection of Individuals with regard to Automatic Processing of Personal Data; and
- c) the national laws applying their provisions.

Any additional regulations at national level that do not fall under the GDPR and apply to data protection or any other sensitive information will also be taken into account.

The project will respect the privacy of all stakeholders and citizens and follow a stakeholder management procedure assuring that stakeholders and citizens are fully informed about their rights, the objective and handling of their data. The stakeholder management procedure obliges the project partners to obtain consents where personally identifiable data is collected and processed as described above, implementing suitable data handling procedures and protocols to avoid potential identification of individuals. This will include participants' data sets in activities that use techniques such as



interviews, questionnaires, workshops, or mailing lists as well as building, energy, and mobility data collection.

Data managed during the project will be processed only under the following preconditions which need to be met:

- (i) When the data subject has given her/his consent;
- (ii) When the processing is necessary for the performance of or the entering into a contract;
- (iii) When processing is necessary for compliance with a legal obligation; and
- (iv) When processing is necessary in order to protect the vital interests of the data subject.

To this end, personal data managed within POCITYF will be anonymised and stored in a form which does not permit identification of users. POCITYF will establish a data management framework that guarantees security of collected personal data from potential abuse, theft, or loss.

This framework will establish the guidelines that the project partners shall obey when data is generated/acquired, transferred and stored, as well as preserved and curated.

POCITYF Ethics Manager, already appointed, will be responsible to ensure the consortium's compliance with the laid-out data protection and privacy measures. The Ethics Manager will be assisted by POCITYF Advisor on EU Data Protection Law, also already appointed. The Ethics Manager will further support the consortium partners in the process of generating/acquiring data and consult both the internal and external stakeholders of POCITYF during any activity related to the generation/acquisition, handling, sharing and curation of data.

### 3.3 Data Security

A fundamental task is to manage the data in a secure way. Besides using data anonymisation techniques (such as data masking, pseudonymisation or data swapping), POCITYF will promote data encryption and backup distribution dealing with sensitive data of individual stakeholders. Moreover, the goal of these measures will be to ensure that data remains consistent over the lifetime of the project and there exist alternatives to the main files, in case they disappear or get corrupted. The encryption component adds an extra layer of security to the data files and information.

For each data set (which will be classified with Table 1 and Table 2, at least), the partners will state the provisions and measures to be implemented to ensure data security, privacy and ethical requirements.

The secure management of information will adhere to the guidelines of relevant standards (e.g. ISO/IEC 27001 and 27002; Code of practice for information security management) to ensure the triad of cyber security:

- I. **Confidentiality** - Preventing unauthorised disclosure of information;
- II. **Integrity** - Assuring that data cannot be modified in an unauthorised manner;
- III. **Availability** - Making information available for authorised users.

The information security management will further contain the Directive on security of network and information systems ('Cybersecurity directive', NIS-Directive 2016/1148) on the security of critical infrastructures and the ePrivacy Directive 2002/58, as well as European Union Agency for Network and Information Security (ENISA) guidance. Storage of information will fully comply with the national and EU legal and regulatory requirements.



The city processes on privacy and security already established as well as data security measures and engineering best practices considered will be encapsulated in the first update of the DMP by month 24.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 864400.



## 4 Data Management, Sharing & Open Access

---

POCITYF will implement a comprehensive data management system complying to the ethics, privacy and security considerations and facilitating a swift processing between the project participants. After the data sets have been generated/acquired, the procedures of how the data is managed shall ensure trackability, transparency and usability among the consortium partners.

In general, POCITYF DMP details 4 different categories of data generated or acquired in the project:

- I. **Research data** - all the data necessary to evaluate the quantitative KPIs of the project and data necessary to validate the results presented in public deliverables or scientific publications. The consortium strongly believes in and applies the concepts of open science, and benefits arising from the European innovation ecosystem and economy by facilitating the reuse of data at a larger scale. The DMP covers the entire research data life cycle and must be consistent with exploitation and IPR requirements. Hence, research data linked to exploitable results will not be put into the open domain if they compromise its commercialisation prospects or have inadequate protection, which is also a H2020 obligation. The Project Coordinator will be responsible to ensure that provisions on Scientific publications and guidelines on Data Management in H2020 are adhered to. As indicated, scientific research data should be findable, accessible interoperable and reusable (FAIR) to ensure it is soundly managed beyond the original purpose for which it was collected.
- II. **Operational and observational data** - all the data, raw data generated/acquired as well as curated data during the implementation, testing and operation of the demonstration activities (*operational data*), and data from qualitative activities including surveys, interviews, fieldwork data or engagement activities (*observational data*, such as the one that will be collected under WP4 activities on citizen engagement). Particularly, sensitive data provided by consortium partners for the demonstration scenarios and personal data of individual stakeholders will be kept strictly confidential to protect their competitive advantage and in terms of personal data anonymised and secured to maintain compliance to GDPR.
- III. **Monitoring and evaluation data** - all the data related to the monitoring of project specific KPIs (WP2) to track the performance of the developed LH energy solutions. This data will be regularly reported and published in relevant repositories (e.g. SCIS) with respective access rights set there.
- IV. **Documentation, instruments and reusable knowledge** - all the data and documentation produced by POCITYF consortia, including specific documentation of the project and demonstration and implementation activities, such as tools, equipment, instruments, software, and underlying source code. The DMP covers the entire research data life cycle and must be consistent with exploitation and IPR requirements. In terms of public results, sufficient and consistent documentation



and publication will support the project's dissemination activities. All public Deliverables will be published on the project website in Open Access.

## 4.1 Data Handling and Management

POCITYF will ensure that all the research data generated is *findable, accessible, interoperable and reusable (FAIR)* complying with the H2020 Guidelines on FAIR Data Management. To implement FAIR Data Management, POCITYF will facilitate the use of a data handling system following the template presented in Table 3.

Table 3 - Template for data handling and management summary

Task/Demo/Activity	Task Name
Description	
Purpose and relevance of data collection and relation to objectives	
Methodology	
Data source, data ownership	
Standards, data formats, vocabularies	
Storage	
Security & Privacy considerations	
Exploitation/Dissemination	
Dissemination Level, Limitations, Approach, Justification	
Stakeholders	

The template will serve as a guideline to detail the description, purpose and relevance of the data sets as well as the methodology and collection procedure used to obtain them. Furthermore, the data management system herein presented allows to complement the general information enclosed in Table 1 and Table 2, by addressing the question of ownership, which standards are applicable, how the information is stored and what data security and privacy considerations shall be taken into account/apply.

For data sets that are further used in publications and/or deliverables as a result of POCITYF, the deliverable, corresponding dissemination level and the stakeholders involved shall be detailed. In case the prediction of whether a data set will be published as part of a deliverable or not, the responsible consortium partner shall update the status of each data set as soon this can be foreseen.

To assure the trackability and transparency of responsibilities for each data set, the concept of a Responsibility Assignment Matrix is incorporated, in an adapted way towards the data management topic. In Figure 1 the basic concept of this approach is illustrated.



	Partner 1	Partner 2	...	Partner n-1	Partner n
Data set 1		R,C,G			
Data set 2	R	C		G	
...				C	
Data set n-1	C	G		R	
Data set n	R				G

R	Responsible
G	Gatherer
C	Curator

Figure 1 Template for Responsibility Assignment Matrix to track responsibilities within POCITYF Project

In the rows of the Responsibility Assignment Matrix all the data sets generated/acquired are listed (making link with the data set ID of Table 1 and Table 2), whilst the columns are assigned to individual project partners. The matrix itself abstracts three different forms of data responsibility: the Responsible (R), the Gatherer (G) and the Curator (C). One partner could potentially cover all the responsibility layers vertically (see Partner 2 for Data set 1 in Figure 2).

The specific tasks corresponding to each responsibility level include but are not limited to:

- i. **Responsible** - responsible to track the data set through its entire life cycle during POCITYF Project and beyond. The data set **Responsible** shall further liaise with the **Gatherer** and **Curator** to ensure transparency and compliance of all activities concerning the data set.
- ii. **Gatherer** - including all the tasks related to the collection or generation/acquisition of the data set. The **Gatherer** shall assure that the data is collected in an appropriate way and complying to the general procedures stated in the DMP.
- iii. **Curator** - responsible to archive and, if applicable, preserve the data set during the project and 5 years after the project’s completion. The **Curator** is further responsible to ensure that the guidelines on data curation and preservation detailed in the following section are implemented (e.g. appropriate storage of data set in repository, costs in relation to data curation, etc.).

Both guidelines, the Data Handling System and the Responsibility Assignment Matrix, will ensure that the data sets are clearly documented and responsibilities are well defined within the life cycle of POCITYF Project.

All information and data gathered and elaborated will be appropriately described in the respective deliverables. All public deliverables will be made available and archived on the project website and through the EU Community Research and Development Information Service (CORDIS) for the project. The project aims to make research data and publications





that comply with the ethics, privacy and security considerations freely available through *Open Access* and suitable repositories.

## 4.2 Sharing and Open Access

The activities within POCITYF will include a number of public deliverables as well as scientific and other publications. The consortium will provide timely open access to research data in project-independent repositories and provide the link to the respective publications, to allow the scientific community to examine and validate the results based on the underlying data.

The policy for open access to research data and publications follows the H2020 Guidelines to Open Access. It is the will and commitment of the partners to share non-commercially sensitive knowledge and experience with the city socioeconomic sector to ensure learning is transferred and errors are not repeated. The consortium members have committed to allocate specific budget to ensure and encourage the academic partners to publish research results as *Gold Open Access*. The policy for open access to research data (technical platform, energy services, non-personal data, knowledge, etc.) arising from this project is illustrated in the following picture.

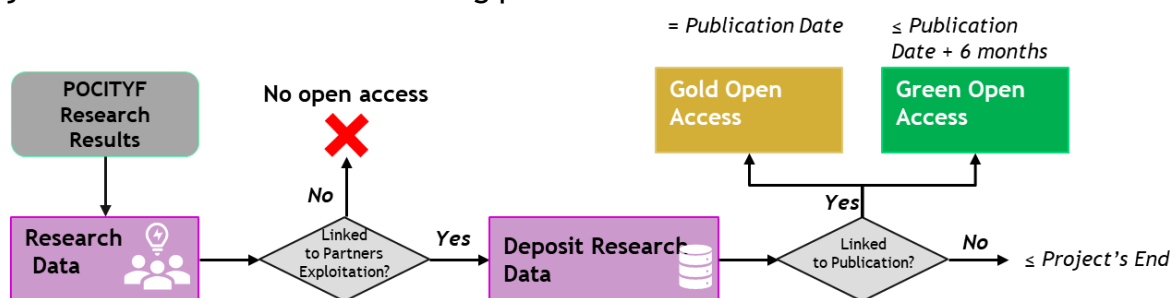


Figure 2 - Data Management and IPR Protection

In case *Gold Open Access* in peer-reviewed repositories cannot be achieved, the consortium will resort to *Green Open Access* strategy.

## 4.3 Document Management

All documents and files generated within the POCITYF are handled in one platform (Microsoft Teams) to ease the collaboration and increase transparency among the various consortium partners to locate and access the project documentation. The naming and terminology of the files and documents follows a rigid structure agreed on by all consortium partners, addressed in D11.1 - *Project Management Roadmap*.

The partners are responsible that the documents and files they upload are complying to the overall guidelines laid-out in this DMP. The partners are advised to coordinate the sharing of sensitive documents with the Ethics Board, Ethics Manager and POCITYF Advisor on EU Data Protection Law.



## 5 Data Curation & Preservation

This chapter describes the procedures for the long-term curation and preservation of the data. The procedures included will indicate how long the data shall be stored, where it will be stored and what cost are associated with it.

Public deliverables will be published and curated on the project website while internal data sets will be backed up to allow recovery for re-use and/or verification. Primary data will be archived for a minimum of years, to be determined still, by the data responsible/partner generating the data.

The classification system presented in Table 4 shall facilitate appropriate measure for the data curation and preservation procedures.

Table 4 - Tool to characterise data curation and preservation

ID	Title of data set	Repository			Long-term preservation plan			
		Storage location	Type of repository	Reasons to use this repository	Time period	Approximated end volume	Storage repository	Associated costs for preservation



## 6 Conclusions

---

The present deliverable constitutes the first version of the DMP for POCITYF at the time of delivery: March 2020 (M6). POCITYF Project Management Team will regularly reflect with the consortium members to refine and update the DMP.

More detailed procedures, descriptions, forms, etc. will be added as they become available through the ongoing work in the respective Work Packages whilst existing templates will be verified, complemented and maintained. The next update will include detailed data set descriptions for the activities within that period.

This first version of the DMP presents an innovative way to deal with the Data Handling and Management, presenting, in that respective section, an adapted Responsibility Assignment Matrix, identifying the roles of Responsible, Gatherer and Curator, for each POCITYF data set.

The DMP will be updated in month 24 and 48 of the project schedule. These updates will detail the defined data management procedures and contain more detailed data set descriptions.

