

Grant Agreement no. 829010



Call: H2020-FETOPEN-2018-2020 **Topic:** FETOPEN-01-2018-2019-2020

Type of Action: RIA (Research and Innovation action)

Name of Lead Beneficiary: CSIC, Spain

Project Start Date: 1st May 2019 Project Duration: 48-Months

DELIVERABLE 4.3: DATA MANAGEMENT PLAN

Due date of Deliverable: 31st October 2019 **Actual Submission Date:** 30th October 2019

Responsible partner: CSIC

Report Author(s): Carlos Sánchez Somolinos

Type11: ORDP

Dissemination Level²: PU

¹ **Type**: Use one of the following codes (in consistence with the Description of the Action):

R: Document, report (excluding the periodic and final reports)

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

ORDP: Open Research Data Pilot OTHER: Software, technical diagram, etc.

² Dissemination level: Use one of the following codes (in consistence with the Description of the Action)

PU: Public, fully open, e.g. web

CO: Confidential, restricted under conditions set out in the Model Grant Agreement CI: Classified, information as referred to in Commission Decision 2001/844/EC















DELIVERABLE D4.3: Data Management Plan

Table of Contents

1. Document History	3
2. List of Abbreviations	4
3. Deliverable description	5
4. Data Summary	5
5. Fair Data	6
5.1 Making data findable, including provisions for metadata	7
5.2 Making data openly accessible	10
5.3 Making data interoperable	10
6. Allocation of Resources	10
7. Data Security	11
8. Ethical Aspects	11
9. Other Issues	11
10. Conclusions	11
11. Bibliography and Sources	11



1. DOCUMENT HISTORY

Version	Date	Authors/ who took action	Comment	Modifications made by
0.1	26.09.2019	Carlos Sánchez Somolinos (CSIC)	First draft sent to PIs	
0.2	03.10.2019	Falk, Alfaro, Ladenhauf (BNN)	Draft review by BNN, TU/e and UZ	
0.3	24.10.2019	Carlos Sánchez Somolinos (CSIC)	Draft sent to BNN for comments	
0.4	24.10.2019	Alfaro (BNN)	Draft review by BNN	
1.0	30.10.2019	Carlos Sánchez Somolinos (CSIC)	Submitted to Commission	



2. LIST OF ABBREVIATIONS

DMP: Data Management Plan

DOI: Digital Object Identifier

ORD: Open Research Data

GA: Grant Agreement

EC: European Commission

CERN: Organisation Européenne pour la Recherche Nucléaire

FAIR: Findable, Accessible, Interoperable and Re-usable

OA: Open Access



3. DELIVERABLE DESCRIPTION

This DMP describes the data management life cycle for the data arising from the research carried out within the framework of the PRIME project. This DMP aims to establish a common consortium strategy to identify the data to be collected, processed or generated. We will also define how to preserve data, which part of the data will be shared and how this will be done.

To elaborate this DMP we have taken into account the Horizon 2020 Guidelines on FAIR Data Management in Horizon 2020^[1] and the Rules on Open Access to Scientific Publications and Open Access to Research Data.^[2] We have adopted to a large extent the H2020 Data Management Plan template (Version: 13 October 2016).^[3] Attending to the general description of DMP, participation in the ORD Pilot does not imply that all project research data become public. In this respect, this DMP follows the EC recommendation "as open as possible, as closed as necessary". Knowledge generated within the framework of the project can be protected, for example through patenting, or disseminated as scientific publications. To avoid issues related to Intellectual property (IP) rights and their access, the data publicly released by PRIME will be only that related to scientific publications.

This document follows the article 29.3 of the GA about "Open Access to research data". All consortium partners will bind to this document and follow the guidelines and procedures hereby described. This DMP is a first release and it could be updated if new needs arise during the implementation of the project.

4. DATA SUMMARY

The PRIME project will develop a platform to create a new generation of active, tubeless and contactless microfluidic chips effectively changing the currently established paradigm of large-scale, complex-to-use, expensive off-chip equipment. To implement this platform, PRIME will develop and integrate new responsive materials and elements in the chip, effectively providing it with all the fluidic and sensing functions. PRIME research will allow the implementation with a large freedom of design of new smart-integrated and easy-to-operate microfluidic devices. The project will generate data that will be crucial for the development of the technology. Besides, the data could be useful for other researchers and scientists in the field of responsive materials, soft actuators, (bio-)sensors, microfluidics, in vitro diagnostic and organ-on-chip devices.

The data collected, in the laboratory notebook or in a file provided by a characterization equipment, or that generated in a computer simulation, within the framework of the PRIME project will be processed and analysed in useful forms leading to reports, publications, etc. Data will be preserved using appropriate rules, regarding names and metadata schemes by each individual partner. Identification of the datasets will be systematically done by each partner to ensure coherence along the datasets. Quality of datasets will be ensured by all the researchers working on the project and in all stages, including data collection, generation, handling and preservation. Each partner will be responsible of their own datasets.

PRIME DMP establishes how to provide the underlying data of scientific publications related to the project with the appropriate, names, metadata and standards ending in publicly accessible datasets in the long term.

Widely used, preferentially non-proprietary, file formats, will be used. Typically, numerical data collected and processed, for example through measurements, or generated in a computer





simulation will be incorporated in a spreadsheet that will be saved as a comma separated value (.csv) file. Optical or electron microscopy images from samples will be provided as image files in tagged image file format (.tiff). Computer aided design (CAD) drawings for active element or microfluidics chip designs will be saved in drawing exchange format (.dxf). Table 1 presents an overview of example dataset types and formats related to the project tasks that could be generated. Depending on the research needs arising in the course of the project, this list might be expanded with new types of datasets and formats. The size of each dataset cannot be estimated in all the cases but it is expected that the datasets generated within the project will never exceed 50 GB, the limit at the repository chosen by the consortium, ZENODO.^[4]

Table 1 Task and related datasets together with their format

Task	Dataset	Format
Polymer synthesis	Fourier transform infrared (FTIR) data	
	Nuclear magnetic Resonance (NMR) data	
	Differential scanning calorimetry (DSC) data	.csv
	Polarization Optical microscope (POM) images	.tiff
Polymer processing	UV-Vis Spectroscopic data	.csv
	Reflectivity data (UV-Vis)	.csv
	Differential scanning calorimetry (DSC) data	.csv
	POM images	.tiff
	Scanning electron microscopy (SEM) images	.tiff
Actuator preparation and characterization	Strain and stress vs light and/or temperature data	.csv
Analytical and numerical modelling of actuators	Strain and stress vs light and/or temperature data	.csv
Nanoparticle synthesis and functionalization	SEM and Transmission electron microscopy (TEM) images	.tiff
	Dynamic light scattering (DLS) data	.csv
Biosensor characterization	Sensor sensitivity and selectivity. UV-Vis vs analyte concentration. Spectroscopic data	.CSV
Valve and pump design and preparation	Computer aided design (CAD) drawings	.dxf
	Strain and stress vs light and/or temperature data	.csv
	Fluidic performance data	.csv
Numerical modelling of fluidic systems	Fluidic performance data	.csv
In vitro diagnostic (IVD) device and Organ on Chip (OoC) design and preparation	Computer aided design (CAD) drawings	.dxf
	Fluidic performance data	.csv
	IVD performance data	.csv
	OoC performance data	.csv





5. FAIR DATA

5.1 Making data findable, including provisions for metadata

As mentioned, PRIME datasets corresponding to the underlying data of PRIME scientific publications will be available at the Zenodo (www.zenodo.org) repository. Zenodo allows harvesting of the entire repository via the Open Archives Initiative Protocol for Metadata Harvesting (OAI- PMH), a broadly used protocol to harvest metadata also, by most repository software.

PRIME project has created its own public PROJECT-PRIME community.

https://zenodo.org/communities/project-prime/

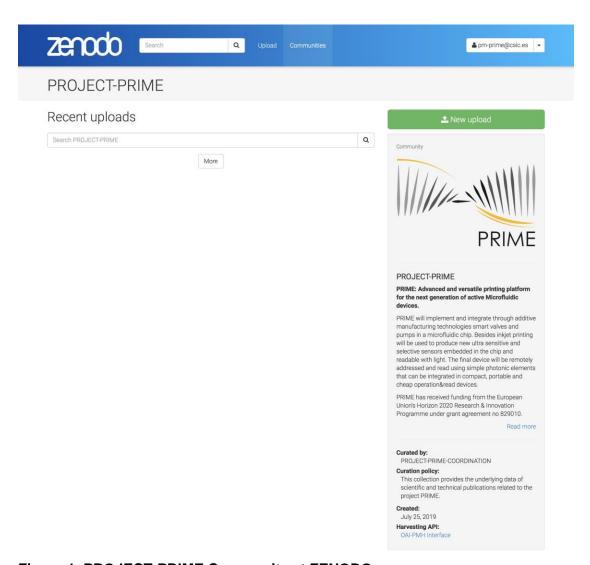


Figure 1. PROJECT-PRIME Community at ZENODO





Each partner, after reporting results in a scientific publication, will be responsible to make a "new upload" in this community for each of the datasets that it is responsible in the publication.

The dataset file format is established in section 4 according to the dataset type. The name of the dataset related to a research publication figure will be assigned as follows: "Standard journal Abbreviation+Year_Issue Nr._Page_First author surname_Figure Nr.". For example: "AdvFunctMater2020_31_12345_Reynolds_Fig3". For figures involving several datasets files, each of them will be treated separately clearly identifying them (e.g.: _Fig3a, _Fig3b, _Fig3c...). If the dataset relates to a table, the name will set as follows: "Standard journal Abbreviation+Year Issue Nr. Page First author surname Table Nr."

Together with the dataset file and in the same upload, a second file in .txt format, will be also incorporated including dataset key information shown in Table 2, complying in this way with the FAIR data principles.

Table 2 Metadata incorporated in the associated .txt file associated to each dataset.

Standard journal Abbreviation+Year_Issue Nr._Page_First author surname_Figure Nr_METADATA.txt (example: AdvFunctMater2020_31_12345_Reynolds_Fig3_METADATA.txt)

PRIME. Advanced and versatile PRInting platform for the next generation of active Microfluidic dEvices)

- Publication title
- Publication reference (in standard format: Standard journal Abbreviation, Year, Issue, Page.)
- Publication DOI (as provided by the publisher)
- Researcher(s) (including all the authors and affiliations as shown in the publication)
- Publication keywords (besides those employed in the publication, these will be complemented with specific keywords related to each precise dataset if needed, e.g.: Fluidic performance data)
- Publication date (in YYYY-MM-DD format)
- Figure Nr. and caption or Table Nr. and caption (as appearing in the publication)
- DOI of the dataset (ZENODO will register a new DOI before upload. This will allow others to easily and unambiguously cite the uploaded dataset)
- Version (using the format v1.0, v1.1, v1. 2...)
- Sample name and description
- Equipment employed
- Experimental conditions
- Sampling procedure
- Variables and units
- Filetype (In case non-proprietary file formats, as those described in section 4 cannot be employed, a reference to the used file format and the software supporting it will be outlined here.
- Access rights: Data files may be either open access and subject to a license described below or closed access and not available for download.
- License: Creative Commons Licensing





Associated to these two files (dataset file and Metadata .txt file), key information corresponding to the fields of the ZENODO upload form will be incorporated as shown in table 3.

Table 3 Dataset metadata incorporated in ZENODO upload form

- Community: PROJECT-PRIME

Upload type: DatasetDOI of the dataset

- Publication date
- Title (as described above using "Standard journal Abbreviation+Year_Issue Nr._Page_First author surname_Figure Nr.")
- Authors and affiliation as shown in the publication
- Description:
 - o Project
 - o Publication title
 - Publication reference
 - o Publication DOI
 - Figure footnote
- Version: (using the format v1.0, v1.1, v1. 2...)
- Language: (e.g.: English)
- Keywords (those employed in the publication will be complemented for each specific dataset as described)
- Additional notes:
 - o Project acronym and name: PRIME. Advanced and versatile PRInting platform for the next generation of active Microfluidic dEvices).
 - Project funding: PRIME has received Funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 829010.
- Access rights:
- License:
- Funding: European Commission (EU); Grant Agreement No. 829010

After a partner publishes a dataset and the associated key information, the coordinator will be responsible to accept the dataset in the PROJECT-PRIME community so the data is both public and linked to this community.





5.2 Making data openly accessible

Project results will be made available Open Access in making use of online repository services (e.g.: DIGITAL.CSIC) or by publishing in journals adhering to Open Access policies (green or gold). Dissemination through the project's website (https://www.project-prime.eu/) will also be carried out.

As already stated, only the underlying data of scientific publications from the project will be made openly available in the research data and publication repository Zenodo. This is a repository for EC funded research provided by CERN, an OpenAIRE partner. The website Zenodo.org does not require an account, login or password giving free access to the deposited data.

5.3 Making data interoperable

To ensure interoperability, allowing re-use and exchange of data between researchers from different institutions, organizations or countries, the datasets will make use of the standards for data and metadata capture and elaboration described above.

Whenever possible the data will be provided in non-proprietary file formats as those described in section 4 (e.g.: .txt, .csv, .tiff, .dxf...). In case this is not possible, a reference to the file format and the software supporting it will be outlined in the metadata file.

A controlled vocabulary will be employed in the metadata fields to provide an accurate and consistent indexing and retrieving of research data. Keywords employed in the publications will be complemented with specific keywords related to each precise dataset and with synonyms to facilitate interdisciplinary interoperability of the project's data and metadata.

5.4 Increase data re-use (through clarifying licences)

It is very likely that datasets from the PRIME project can be re-used in the future to benchmark results of other projects or to study the suitability of PRIME materials and devices in the proposed application or in others. Each partner principal investigator (PI) and later the coordinator will check the uploaded datasets to ensure that they are re-usable. To ensure re-usability during and after the project, the datasets will be deposited in Zenodo and a creative commons license will be used to protect our ownership.

The underlying datasets of scientific publications will be made available immediately after publication release for "gold access" publications. For "green access" publications an embargo period may apply.

6. ALLOCATION OF RESOURCES

Compliance with the FAIR data principles implies activities that have associated costs:





- Open-access publications: Costs associated to open-access publications are eligible for reimbursement in H2020 under the conditions defined in Article 6 and Article 6.2.D.3. of the GA. These OA costs were already considered in the project budget.
- Project website maintenance: by the project partner BioNanoNet.
- ZENODO repository: Free of charge.
- Data preservation and storage: Data preservation for at least 1 year after the project end is required (GA 31.3). Each partner is responsible for the data that produces. Any other fee will be responsibility of the data owner. The underlying data of scientific publications related to the project will be transferred to the ZENODO repository, which ensures sustainable archiving.
- In case needed, additional data storage will be ensured by each partner institutions' data repositories.

7. DATA SECURITY

Each partner is responsible of the security of their own data. The storage of, at least, two backup copies at different locations will be carried out to avoid data lost.

The underlying data of scientific publications will be safely stored for long-term preservation in the ZENODO repository, the same cloud infrastructure as CERN's own LHC research data. According to Zenodo the data and metadata will be stored in CERN Data Center with multiple independent replicas. [4] In the unlikely event that ZENODO stops this service, PRIME will migrate all the project content (datasets and associated metadata) to other equivalent repositories. On the long-term storage, the institutional repositories of each partner should provide a satisfactory level of security.

No issues regarding personal data security are foreseen.

8. ETHICAL ASPECTS

As PRIME DMP only affects to the underlying data of scientific publications no ethical or legal issues that could have impact on data sharing are foreseen.

9. OTHER ISSUES

No other issues to report at this moment.

10. Conclusions

This DMP defines the data management life cycle for the data generated within the framework of the PRIME project. This document that has been done following the FAIR data principles, is the first release of the PRIME DMP and thus it will be updated if new needs arise during the implementation of the project.

11. BIBLIOGRAPHY AND SOURCES

[1] European Commission, "Guidelines on FAIR Data Management in Horizon2020 (v3.0)," European Commission, 2016.





- [2] European Commission, "Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research data in Horizon 2020," European Commission Directorate For Research and Innovation, Brussels, 2017.
- [3] European Commission, "H2020 templates: Data Management Plan v1.0," 2016.
- [4] https://help.zenodo.org/features/