



Innovative Training Networks (ITN)  
H2020-MSCA-ITN-2019  
Grant Agreement No: 860497

# GMOS-Train

## Global Mercury Observation Training Network in Support of the Minamata Convention

### Deliverable D8.13 "Data Management Plan (DMP)"



Start date of project: 01.01.2020

Duration: 48 months

## Bibliographical Information

Grant Agreement Number	860497
Project Acronym	GMOS-Train
Project Full title	Global Mercury Observation Training Network in Support of the Minamata Convention
Funding Scheme	H2020-MSCA-ITN-2019
Project URL	<a href="https://www.gmos-train.eu/">https://www.gmos-train.eu/</a>
Project Coordinator	Prof Milena Horvat, Jožef Stefan Institute, Slovenia
Contact info	Jožef Stefan Institute Jamova 39 SI - 1000 Ljubljana Slovenia Phone: +386 1 588 54 15 e-mail: <a href="mailto:info@gmos-train.eu">info@gmos-train.eu</a>

WP Number	8
WP Title	Dissemination, Communication, Outreach
Deliverable Number	8.13
Deliverable Title	Data Management Plan (DMP)
Lead Beneficiary	CNR-IIA
Author(s) (Partner)	Ian M. Hedgecock (CNR-IIA), Nicola Pirrone (CNR-IIA), Jeroen Sonke (CNRS), Aurelien Dommergue (UGA), Milena Horvat (JSI), Vanja Usenik (JSI)
Nature	<input type="checkbox"/> Report
	<input type="checkbox"/> Other
	<input type="checkbox"/> Websites, patents filing, etc.
	<input checked="" type="checkbox"/> ORDP: Open Research Data Pilot
	<input type="checkbox"/> Ethics
Dissemination Level	<input type="checkbox"/> CO - Confidential, only for members of the consortium (including the Commission Services)
	<input checked="" type="checkbox"/> PU - Public
Est. Del. Date (annex I)	M6 (30.6.2020)
Actual Del. Date	25.9.2020
Version	1.0
Key Words	GMOS-Train, Early Stage Researcher, ESR, Data Management Plan, DMP, metadata, open access, Open Research Data Pilot
Copyright	© 2020 by GMOS-Train This project has been funded with support from the European Commission. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use, which may be made of the information contained therein. We certify that all of the facts submitted in document are true to the best of our knowledge and belief.



## Acronyms and Abbreviations

### PROJECT BENEFICIARIES:

AMU	Université d'Aix-Marseille – Mediterranean Institute of Oceanography, France
CNR IIA	Institute of Atmospheric Pollution Research of the Italian National Research Council, Italy
CNRS	Centre National de Recherche Scientifique, France
HZG	Helmholtz-Zentrum Geesthacht Zentrum für Material- und Küstenforschung GmbH, Germany
IFREMER	French Research Institute for Exploitation of the Sea, France
IOS	Institute for Environmental Protection and Sensors, Slovenia
JSI	Jožef Stefan Institute, Slovenia
PSA	PS Analytical, United Kingdom – project exit date 1.7.2020
UGA	Université Grenoble Alpes, France
UPPA	Université de Pau et des Pays de l'Adour, France
SU	Stockholm University, Sweden

### PROJECT PARTNER ORGANISATIONS:

AMAP	Arctic Monitoring and Assessment Programme, Norway
AUTH	Aristotle University of Thessaloniki, Greece
EEB	European Environmental Bureau, Belgium
Harvard	Harvard University, USA
IPSJS	International Postgraduate School Jožef Stefan, Slovenia
IRD	Institut de Recherche pour le Développement, France
Lumex	Lumex, Germany/Russia
MIT	Massachusetts Institute of Technology, USA
MSC-E	Meteorological Synthesizing Centre – East of EMEP, Russia
PSA	PS Analytical, United Kingdom
SPRS	Swedish Polar Research Secretariat, Sweden
Tekran	Tekran, Canada
UBL	Université Bretagne Loire, France
UNEP	United Nations Environmental Programme, Switzerland
UPS	Université Paul Sabatier, France
VSL	Dutch National Standard Laboratory, The Netherlands
API	Application Programming Interface
DMP	Data Management Plan
EC	European Commission
ESR	Early Stage Researcher
FAIR	Findable, Accessible, Interoperable, Reusable
IPR	Intellectual property rights
SBA	Societal Benefit Areas



## Executive Summary

The GMOS-Train consortium will draw on the expertise of its participants to draw up tailor-made Data Management Plans for each of the ESRs according to their specific research areas, the type of data they will produce, the international initiatives and programs to which their research is relevant and can contribute, and ensure that the results of their work adhere to the principles of FAIR data management. The GMOS-Train Data Management Plan will be a living document, and as such, it will be reviewed and updated regularly during the project implementation to reflect new challenges and needs of the scientific community and all stakeholders as well as technological advances in the field. The final DMP version will be presented in Deliverable D8.11 “Final exploitation plan” (M46).



## Table of Contents

Bibliographical Information.....	2
Acronyms and Abbreviations.....	3
Executive Summary.....	4
1. Introduction.....	6
2. Principles.....	6
3. DMP Questionnaire.....	8
4. Ethical aspects.....	11



## 1. Introduction

The GMOS.Train project will benefit directly from the years of work which have gone into the development of the Global Mercury Observation System (GMOS), which began with the eponymous project in the Seventh Framework Programme (<https://cordis.europa.eu/project/id/265113>), starting in 2010. One of the primary objectives of the original project was “to develop interoperable tools to allow the sharing of observational and model output data produced by GMOS. The overarching goal of GMOS was to support the achievement of goals set by the GEO / GEOSS, and specifically of the GEO Task HE-09-02d and contribute to the advancement of our scientific understanding in the nine Societal Benefit Areas (SBA) established in GEOSS”. The Project evolved into the GMOS program (<http://www.gmos.eu/>) with the intention of supporting the Minamata Convention (specifically articles 19 and 22, <http://www.mercuryconvention.org/>), and the GEO Task evolved into GOS4M; “GOS4M is a GEO Flagship aimed to support the Minamata Convention on Mercury Secretariat, the UN Environment Mercury Fate & Transport Partnership and all Nations in the follow up of the Conferences of Parties (COP) related to the Effectiveness Evaluation and Global Monitoring framework” (<http://www.gos4m.org/>), one of whose objectives is to “develop IT applications which will be based on observations, model outputs and web tools to exploit the wealth of data made available through platforms like, for example Copernicus DIAS and GEO Knowledge Hub”.

The close relationship with GEO (<https://www.earthobservations.org/index.php>) and hence with GEOSS (<https://www.earthobservations.org/geoss.php>) and the fact that recently GEO has become an Observer at COP-3 of the Minamata Convention has forged strong cooperation between the mercury science community and GEO, committed to “Full and open access to Earth observation data, information and knowledge is crucial for humanity as it faces unprecedented social, economic and environmental challenges. monitoring and earth observation Effectiveness Evaluation under a larger framework”.

Within this context GMOS-Train will draw on the expertise of its participants to draw up tailor-made Data Management Plans for each of the ESRs according to their specific research areas, the type of data they will produce, the international initiatives and programs to which their research is relevant and can contribute, and ensure that the results of their work adhere to the principles of FAIR data management. The GMOS-Train Data Management Plan will be a living document, and as such, it will be reviewed and updated regularly during the project implementation to reflect new challenges and needs of the ESRs, scientific community and all stakeholders as well as technological advances in the field. The final DMP version will be presented in Deliverable D8.11 “Final exploitation plan” (M46).

## 2. Principles

The Data Management Plan of the individual ESRs will be based on the H2020 DMP, which has been adapted into a Questionnaire (see section 3) which the ESR supervisors will tailor to their specific fields of research, particularly with respect to the questions regarding the Initiatives and Programs to which the ESR’s project contributes. This will also be important for choice of data format, documentation and where the data are preserved. It should be borne in mind that the ESR projects cover a very wide range of fields even if they have



a unifying theme. The data produced during the projects will therefore be a heterogeneous mix of types, formats and quantities, and thus this DMP is general and it will be the responsibility of the supervisors and ESRs to ensure that their individual plans are appropriate to the project's research field.

In order to help both the ESRs and their supervisors in the development of their own DMPs, the Principles of Data Management will be addressed at the Kick-off meeting. The approach from GEO to the implementation of Data Management Principles, as laid out below under five headings: discoverability, accessibility, usability, preservation, and curation, provides a broad overview.

### *Discoverability*

1. **Data and all associated metadata will be discoverable through catalogues and search engines**, and data access and use conditions, including licenses, will be clearly indicated.

### *Accessibility*

2. **Data will be accessible via Open Access publications, or equivalent online publication repositories.**

### *Usability*

3. **Data will be structured using encodings that are widely accepted in the target user community** and aligned with organizational needs and observing methods, with preference given to non-proprietary international standards.
4. **Data will be comprehensively documented**, including all elements necessary to access, use, understand, and to process it.
5. **Data will include descriptions of the origin** and processing history of raw observations and derived products, to ensure full traceability of the product chain.
6. **Data will be quality-controlled.**

### *Preservation*

7. **Data will be protected from loss and preserved for future use.**

### *Curation*

8. **Data will be assigned appropriate persistent, resolvable identifiers** to enable documents to cite the data on which they are based and to enable data providers to receive acknowledgement of use of their data.

The specific requirements within H2020 and therefore GMOS-Train will then be addressed using the DMP Templates (section 3).



### 3. DMP Questionnaire

The Questionnaire has been designed using the Horizon 2020 DMP - Initial DMP.

Questions	Answers
Name of person/organization responding to the survey	
What is the purpose of the data collection / generation and its relation to the objectives of the project?	
<b>Data Summary</b>	
To whom might it be useful ('data utility')?	
<b>Making data findable, including provisions for metadata</b>	
Are the data produced and/or used in the project discoverable with metadata? What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.	
Are data identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?	
<b>Making data openly accessible</b>	
Are datasets openly accessible? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions. Also explain where and how the conditions for access are described (i.e. a machine readable license)?	
Does dataset access require specific software tools? If yes, are they easily available (e.g. executable or source code)?	
Where will the data and associated metadata, documentation and code be deposited? Preference should	





<p>be given to certified repositories which support open access where possible.</p> <p>Will data and all associated metadata be discoverable through catalogues and search engines?</p>	
<b>Making data interoperable</b>	
<p>Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?</p>	
<p>Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?</p>	
<b>Increase data re-use</b>	
<p>Is the data safely stored in certified repositories for long term preservation and curation?</p>	
<p>How will the data be licensed to permit the widest re-use possible?</p> <p>Does the data require attribution?</p> <p>If significant in a scientific paper, does it require citation or authorship?</p>	
<p>When will the data be made available for re-use?</p> <p>If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.</p>	
<p>Will data include provenance metadata indicating the origin and processing history of raw observations and derived products, to ensure full traceability of the product chain?</p>	
<p>Are data quality assurance processes described?</p>	
<p>Will data be fully documented including all elements necessary to access, use, understand, and process,</p>	



preferably via formal structured metadata? Will data also be described in peer-reviewed publications referenced in the metadata record?	
Will data be accessible via online services, including user-customizable services for visualization and computation?	
Is the complete dataset available for download? For datasets that are composed of several elements (e.g. a long time series), is there a protocol or a API for automatic downloading?	
<b>Allocation of resources</b>	
What are the costs for making data FAIR in your project? Who will be responsible for data management and preservation? Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?	
Who will be responsible for data management and preservation?	
Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?	

### Compliance with H2020 Open Research Data Pilot

1. Data generated in the project are available in a repository with proper metadata:

- Yes  
 Partially: \_\_% of datasets are available in a repository with proper metadata  
 No

Comment/Justification:



2. Data generated in the project are available free-of-charge:

- Yes
- Partially: \_\_% of datasets are available in a repository with proper metadata
- No

Comment/Justification:

3. Scientific results (e.g. publications) of the project are reproducible (data are available free-of-charge, information on required tools and instruments are provided):

- Yes
- Partially: \_\_% of datasets are available in a repository with proper metadata
- No

Comment/Justification:

4. Do the Project results contribute to National or International Initiatives or Programmes?

- Yes
- Partially: \_\_% of datasets are available in a repository with proper metadata
- No

Comment/Justification:

## 4. Ethical aspects

No personal data will be exchanged. The ethical standards and guidelines of Horizon 2020 will be rigorously applied regardless of the country in which the research and training activities are carried out.

