CrossGov
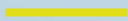
Coherent & Cross-compliant Ocean Governance for Delivering the EU Green Deal for European Seas

Deliverable 6.1

Plan for Data Management

Grant Agreement number	101060958
Project title	CrossGov - Coherent and Cross-compliant Ocean Governance for Delivering the EU Green Deal for European Seas
Deliverable title	Plan for Data Management
Deliverable number	6.1
Deliverable version	Version 1
Contractual date of delivery	28.02.2023
Actual date of delivery	28.02.2023
Document status	Final
Document version	1.0
Online access	Public
Diffusion	Public document
Nature of deliverable	Public
Work Package	6
Partner responsible	s.Pro – Sustainable Projects GmbH
Contributing Partners	all
Author(s)	Pedro Brosei
Editor	Paulina Ramirez-Monsalve
Reviewed by	Frode Sundnes
Approved by	Froukje Maria Platjouw
Project Officer	Gonzalo Garcia de Arboleya
Abstract	This document presents an Initial data management plan (DMP) for the CrossGov project, as defined in the Grant Agreement with respect to the EU Horizon Europe guidelines/template. This DMP provides guidelines for the partners to follow and should be updated with specific information on data sets throughout the project. The DMP includes a summary of data that is to be collected, under different Work Packages and how the CrossGov project will adopt FAIR data policy.
Keywords	CrossGov, Data management, Fair Principles

Table of contents

D 5.5

List of Tables

Acronyms

DMP - Data Management Plan

EC – European Commission

EU – European Union

WP – Work Package

IPR - Intellectual Property Rights

Executive Summary

This document presents a data management plan (DMP) of the CROSSGOV project, as defined in the Grant Agreement and with respect to the EU Horizon Europe guidelines/template. The plan includes the definition of the data and the steps to make all project outputs FAIR (Findable, Accessible, Interoperable, Reusable): its first version (dated from Month 6) was developed on the best knowledge possible at this point. It will be regularly updated throughout the project's lifetime. The Data and Knowledge Management task leader will look at the DMP every 6 months starting from the submission of this initial version and discuss with the project coordinator if and if yes, which updates are necessary and proceed with the update in the respective case.

The DMP provides guidelines for the partners to follow and should be updated with specific information on CROSSGOV data sets throughout the project. The DMP includes a summary of data foreseen to be collected, primarily through interviews and workshops. It also describes how the project will adopt FAIR data policy through several technical guidelines and documents by widely accepted frameworks and guidelines. Critical data management issues will be discussed, including but not limited to common (meta)data standards and formats, data quality control, (meta)data identification mechanisms and vocabulary.

How the project deals with Intellectual Property Rights (IPR) is mainly dealt with in the separate Strategy for Ethics & IP (D6.2), which was developed alongside this DMP. The DMP only addresses the protection of IPR in relation to the findability, accessibility, interoperability and reusability of data produced under CROSSGOV.

Introduction

The Data Management Plan (DMP) outlines how data is collected, generated and/or processed throughout the lifespan of a research project. As part of the European Commission's goal to advance Open Science policy and practices – it is a mandatory component in Horizon Europe that all projects involving data include a DMP.

Having this in mind, the DMP of CROSSGOV defines the data collected, produced, handled and stored in the project in the following way: *Information that has been collected through questionnaires, interviews, focus groups, workshops or observation.*

The DMP further describes the steps to make all project outputs FAIR (Findable, Accessible, Interoperable, Reusable), ensuring the long-term storage of project outputs where relevant and feasible through self-archiving.

The DMP will outline how the consortium deals with the sourcing, processing, analysis and storage of data generated in the project (e.g. in relation to surveys, interviews and co-creation events), the standards, guidelines and methodologies it adopts, as well as how data will be made available and to whom.

It will be implemented in accordance with the Strategy for Addressing the Ethics and Intellectual Property Rights (D6.2), facilitating potential reuse of the data collected and processed and outputs produced within the project in accordance with the principle of 'as open as possible, as closed as necessary' (HE Programme Guide, 2022) and also ensuring that the project is entirely in line with the GDPR.

All project partners will ensure that the data they collect, produce, handle, and store is in agreement with the DMP and with any national requirements. Project partners will appoint a dedicated contact person responsible for this within their institution and will inform the task leader and the project coordinator in writing that they are that they have to comply with any national requirements.

The first version of the DMP was developed with the involvement of all PPs filling in the table that was created based on the HE DMP template (see Annex I). In this table, they specified, to the extent possible at this stage, the datasets which will be produced in relation to each work package/task and their most important characteristics in line with FAIR principles. The table and the DMP will be regularly updated throughout the project's lifespan.

Data summary

In the research context of CROSSGOV and in line with the definition given in the Introduction above, examples of data include questionnaire results and key points extracted from interviews, focus groups, workshops, and observation processes. The focus is on primary research data that is available in digital form. (European Commission, 2016). Data is collected, named, stored, analysed, compiled, and managed and may also be compiled into digital outputs.

The data and digital outputs outlined above are expected to have long-term value for a broad research community and various stakeholder groups and organisations. A detailed list of datasets per Task per Work Package identified in the initial stage of the project, with the most essential information based on the HE DMP template, can be found in the ANNEX.

Foreseen outputs

Datasets

- **Description:** Data will be collected for most project deliverables (see Tables 1 & 2). More information will be provided in the DMP updates.
- Moreover, it is foreseen that one Excel database will be created for the Stakeholder Forum in WP5, which will be used and updated throughout the project. The database will be kept internally. Relevant data and information might be compiled and summarised into respective deliverables, publicly available.
- **The data and information will be collected:** through questionnaires, interviews, focus groups, workshops, and observations with stakeholders. One-pagers with key information collected from these events will be made available following the principles of Open Science (OS).

Title of the output	Dataset	WP/Tasks	Responsible partner
Scenarios	Transposed discussions; One-pagers with key information collected from interaction with stakeholders	WP1/T1.1	UU/RIFS
Methodological guidelines for Coherence and Cross-Compliance	One-pagers with key information collected from interaction with stakeholders	WP1/T1.2	NIVA
Science-Policy-Society interfaces in CrossGov: Methodological guidelines	One-pagers with key information collected from interaction with stakeholders, literature and information sources accessed	WP1/1.3	CNR-ISMAR
Mapping of Coherence & Cross-compliance (Report)	Mapping documents	WP2/2.1	UEF
Report on Horizontal coherence	One-pagers with key information collected from interaction with stakeholders	WP2/2.3	UEF
Report on Vertical coherence	One-pagers with key information collected from interaction with stakeholders	WP2/2.3	UEF

Guiding the case study process (Report)	One-pagers with key information collected from interaction with stakeholders	WP3/3.1	ACTeon
Strengthening of integrative planning - Case studies (Report; online presentation)	One-pagers with key information collected from interaction with stakeholders	WP3/3.2	NIVA
Internalising resilience and marine biodiversity into sectoral decisions (Report, Policy briefs)	One-pagers with key information collected from interaction with stakeholders	WP3/3.3	ACTeon
Web-based interactive roadmaps (recommendations & Guidance)	Based on outputs from WP 1-3	WP4/4.1	s.Pro
Toolbox for assessment of coherence and cross-compliance	Based on outputs from WP 1.2 and WP 2-3	WP4/4.2	NIVA
Blueprint for SPS interface	Based on outputs from WP 1.3 and WP 2-3	WP4/4.3	CNR-ISMAR
Stakeholder Mobilisation Charter	n.a.	WP5/5.1	ACTeon
Events	Contact data etc., in line with the GDPR	WP5/5.2	s.Pro
PEDC, Dissemination material	Based on outputs from WP 1-4	WP5/5.3	s.Pro
Development and testing of training modules	Based on outputs WP 1-4	WP5/5.4	ACTeon

Table 1 List of the main foreseen outputs per each task/Work Package in the CROSSGOV project (more details available in the Annex)

Data Collection Objectives

Data collection under the CROSSGOV project will enable the development of a better methodological and conceptual framework for assessing the related concepts of Coherence and Cross-Compliance for analysing the challenges of implementing the Green Deal with the current legal and policy instruments.

Existing data will be collected/reused, as well as information from stakeholders involved in the Stakeholder Forum (Tasks 5.1, 5.2, 5.4). The ultimate goal is to produce co-created roadmaps towards enhanced Cross-compliance in the European Seas (T4.1) and a toolbox for assessing Coherence and Cross-compliance (T4.2).

Principles for data treatment and storage

The informants (research participants, e.g. stakeholders) who provide qualitative data from questionnaires, interviews, focus groups, workshops and observations (will be informed in writing about the purpose of the research, their role in the project and how collected data will be used (and fill in the **Informed consent statement**, Annex to D.6,2 Strategy for Ethics & IP). Informants will also be informed that they can withdraw their consent to participate by sending a written request at any time. Anonymity will also be guaranteed to research participants. The data-generating project partner organisation is responsible for securing that personal information is handled and stored safely and in line with legal requirements and ethical research rules. Further details are elaborated in the Ethics & IP Strategy (D.6.2).

Internal sharing of data within the partnership, such as interview summaries/analyses, internal outputs and drafts, etc., will occur via the project's MS Teams channel, which is hosted by the project coordinator (NIVA). In the CROSSGOV MS Teams channel, each WP and task has a folder to which WP teams can upload data and outputs to be shared. All project team members have access to the CrossGov MS Teams channel. If considered useful, private groups can be created in Teams to limit access to a smaller group of persons. As already specified, no personal or sensitive data should be uploaded on the CROSSGOV MS Teams channel. Once the material has been uploaded, other project members should be informed appropriately (e.g., at a meeting and/or via email or similar). The data and knowledge management task leader (s.Pro) will ensure that these principles are followed as part of the overall project coordination (WP6).

Certain data generated in the project via, for example, interviews and participant observations will be **sensitive** (e.g., controversial opinions, political views etc.). Therefore, there will be a need for data-generating partners to protect and safely store sensitive data. While the DMP does not rule out that primary data from interviews etc. (like transcripts and recordings) are made accessible to other project partners on the shared repository (MS Teams channel) by the data generating partners, the latter should carefully assess if the raw data might be sensitive and should be securely stored and only shared internally among project researchers in an aggregated and anonymous form (e.g. as interview summaries, aggregated analysis documents of a set of interviews etc.).

However, one-pagers with key information collected from interaction with stakeholders should be made openly accessible via the storage repository (Zenodo) when the dataset is finalised.

Types and formats of data

For the types of data generated and used in CROSSGOV, please see "Data Summary". A preliminary list of data to be collected/compiled in the project for each deliverable/output produced per each task, covering type/format, is given above in Table 2. This list was based on the current inventory summarised in the Annex and will be updated throughout the project.

The best practice in long-term archiving and sharing data and outputs is using standard file formats that most software can interpret. This DMP recommends[1] that CROSSGOV

---

[1] based on the list prepared by Utah State University / USA (following Library of US Congress' recommendations): https://library.usu.edu/data-management/archiving-data/file_formats

researchers use the following file formats for long-term storage (and sharing in a public file repository, i.e. Zenodo):

Data Type	Preferred Formats	Acceptable Formats
Text Documents	Open Document Format (.odt or .ods)	Portable Document Format (.pdf); XML with Standard DTD or Link to Schema (.xml); HTML (.htm or .html); Plain Text - encoding: UTF-8 or ASCII (.txt); PDF - other subtypes (.pdf); Open Office XML (.docx or .xlsx); Rich Text Format (.rtf); EPUB (.epub)
Datasets	Comma Separated Variables (.csv) Office Open XML (.xlsx)	Tab Delimited (.txt); Open Office Spreadsheet (.ods); SQL DDS; dBASE (.dbf)
Raster Image	TIFF - Uncompressed (.tif or .tiff)	JPEG (.jpeg or .jpg); JPEG2000 - Uncompressed or Lossless Compressed (.jp2 or .j2k); PDF/A or PDF/X - Graphic Exchange Format (.pdf); PNG (.png); TIFF - Lossless Compressed (.tif); GIF (.gif); RAW Image Format (.raw); Photoshop Files (.psd); BMP (.bmp)
Audio Files	Broadcast Wave (.wav) Audio Interchange File Format (.aif, .aiff)	MP3 (.mp3); Advanced Audio Coding (.aac, .m4p, .m4a); MIDI (.mid, .midi); Ogg Vorbis (.ogg)

Table 2 List of the main foreseen communication & dissemination digital outputs in the project

Reuse of existing data

The project will reuse existing data in WP1 (related to the assessment frameworks used by the EU and the abstract policy targets of the Green Deal) to develop methodological guidelines that will serve as input for the other work packages (WP2 + 3, and at a later stage, WP4).

Otherwise, no already generated data are expected to be reused at this stage. Information on the reused data that can be found will be updated in the Annex later in case these are still to be defined as the project progresses.

Origin of data

Data origin is dataset specific. It is primarily data generated from interviews, questionnaires, interviews, focus groups, workshops, and observations.

Data utility

CROSSGOV data will be useful to other projects working on similar topics. We aim to guarantee this by means of the One-pagers with key information collected from the interaction with stakeholders. But first, it is expected that data and outputs generated during the project implementation will be used by key users and beneficiaries, such as policymakers, researchers, and governance authorities.

Fair data

Under Horizon Europe, the EU insists all projects adhere to the FAIR data policy. "The FAIR Data Principles are a set of guiding principles to make data findable, accessible, interoperable and reusable" (Wilkinson et al., 2016). Here we refer to the primary data we collect from questionnaires, interviews, focus groups, workshops, observation. This chapter explains how the CROSSGOV project will adhere to all components of the FAIR data policy. As the project progresses, dataset elements will be added to the DMP and checked with the FAIR data principles.

Findability of data

Data must be findable and retrievable over prolonged periods. Therefore, metadata files and standardised identification mechanisms are required where metadata are assigned a globally unique and persistent identifier; data are described with rich metadata, and metadata are registered or indexed in a searchable resource (Wilkinson et al., 2016).

The collected data will be described with standard metadata, and DOIs will be assigned. Links and metadata exchange with the existing major thematic repositories will be established where possible to maximise the findability and impact of the project outputs. Where relevant and feasible, CROSSGOV will ensure the data storage and archival of project outputs in open-access targeted e-infrastructures such as Zenodo.

Accessibility of data

To make data accessible means that corresponding metadata are retrievable by their identifier using a standardised communications protocol; the protocol is open, free, and universally implementable; the protocol allows for an authentication and authorisation procedure where necessary; and metadata are accessible, even when data are no longer available (Wilkinson et al., 2016).

The CROSSGOV project follows the Horizon Europe guidelines to the rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020 (EC Directorate-General for Research & Innovation, 2017). Using and linking with existing repositories (e.g. Zenodo), the project output metadata will be accessible through standard protocols.

Interoperability of data

Interoperable data means that (meta)data use is formal, accessible, shared, and broadly applicable language for knowledge representation; metadata are released with a clear and accessible data usage license; (meta)data use vocabularies that follow FAIR principles; (meta)data include qualified references to other (meta)data (Wilkinson et al., 2016). Making data interoperable is linked to making data reusable and accessible.

CROSSGOV will use existing standards, formats and vocabularies for data and metadata (e.g. INSPIRE). Regarding data management, if there is a need for a specific open format or data requirement from the users, the standards can be adjusted to particular needs enhancing impact.

Reusability of data

The data will be made available to the public, and its reuse will be encouraged.

While the aim is to have all the data and project results publicly available in line with the Strategy for Addressing the Ethics and Intellectual Property Rights (D6.2), the IPR agreement will be established in the project to ensure innovation and competitiveness are maximised. Where relevant and feasible, CROSSGOV will ensure the data storage and archival of project outputs in open-access targeted e-infrastructures such as Zenodo.

Allocation of resources

Curation and storage/preservation costs: Where relevant and feasible, CROSSGOV will ensure collected data storage and archival in open-access targeted e-infrastructures such as Zenodo.

The data and knowledge management task leader (s.Pro) is responsible for developing and updating the Data Management Plan. The CROSSGOV consortium has assigned Pedro Brosei (s.Pro) as Data Management Officer responsible for data management and quality assurance. He will be supported by the project coordinators and managers within the project coordination WP (WP6). He will also be supported by data management contact persons from all project partners specified in the ANNEX table, especially in quality assurance. Paulina Ramirez-Monsalve (NIVA) acts as Deputy Data Management Officer.

Privacy

Personal data resulting from research participants and data security are being handled in line with D6.2 Strategy for Addressing the Ethics and Intellectual Property Rights, which means having data protection measures including pseudonymisation or anonymisation of personal data, data minimisation, applied cryptography, using data-protection focussed service providers and storage platforms and informed consent when working with research participants. This means that research participants will always be asked before participation to sign an informed consent document, available in the Annex to D6.2.

Data Security

For data security and storage, different options are being explored in the CROSSGOV project. Currently, data and information are kept in the MS Teams created for the project, to which the partners have access.

Ethical Aspects

Deliverable 6.2, the Strategy for Addressing the Ethics and Intellectual Property Rights, covers the ethical principles: the principle of proportionality, the right to privacy, the right to protection of personal data, the right to physical and mental integrity of a person, the right to non-discrimination and the need to ensure high levels of human health protection. Based on these principles, an informed consent procedure was set up (Form in Annex to D6.2), and data protection measures comply with the GDPR. The CROSSGOV project envisages several questionnaires, interviews, focus groups, workshops and observations that will be conducted in line with the mentioned informed consent procedure.

Update of the Data Management Plan

The first version of the Data Management Plan is based on the HE Template and MS Excel tables filled in by CROSSGOV partners in the first six months of the project and per each WP, which is annexed to this DMP. The table and the DMP will be updated regularly during project

implementation (every six months) until the Final Data Management Plan is completed at the end of the project's lifetime.

References

Horizon Europe Data Management Plan Template, Version 1.0, 05 May 2021

Data management plan, OCEANS PACT project, December 2020

Annex

The excel table containing questions from HE DMP template filled in by project partners per WP – is to be continuously updated throughout the project.