



NATURANCE

Nature for insurance,
insurance for nature

(Grant Agreement 101060464)

Deliverable D6.1 – Data Management Plan

WP6 – Coordination

Version final | March 2023

HORIZON-CL6-2021-BIODIV-01-06 - Nature-based solutions, prevention and reduction of risks and the insurance sector



Deliverable Title	Data Management Plan
Brief Description	A plan for the dissemination of data within and outside the project
WP number	6
Lead Beneficiary	VU
Author(s)	Max Tesselaar (Review: JoAnne Linnerooth-Bayer)
Deliverable Due Date	31/03/2023
Actual Delivery Date	31/03/2023
Nature of the Deliverable	Other
Dissemination Level	PU - Public

Document History

Date	Ver.	Contributors	Comment
21/12/2022	0.1	Max Tesselaar	Outline of the DMP
22/02/2023	0.2	Max Tesselaar; JoAnne Linnerooth-Bayer	Draft
28/03/2023	1.0	Max Tesselaar	

Table of Contents

Executive Summary.....	1
1. Introduction	2
2. Data summary	2
3. FAIR data.....	4
3.1 Making data findable & metadata.....	4
3.2 Making data openly accessible.....	4
3.3 Making data interoperable.....	5
3.4 Making data interoperable.....	6
4. Data ownership and allocation of resources	6
5. Data security	6
6. Ethical aspects	8



Executive Summary

A Data Management Plan (DMP) is especially helpful in a large group of independent institutes like the NATURANCE consortium because it makes it possible to synchronize data streams. It guides the group from an early date to think about sharing and storing data and about making it findable and interoperable for (in the first place) the group members of the consortium institutes themselves.

The data gathered and generated by the NATURANCE project will be primarily stored in Zenodo, a general-purpose open-access repository developed under the European OpenAIRE program and operated by CERN.

Once all reports have been written, the full datasets gathered and analysis generated within the scope of the project, data will be opened for further use by any interested party. This allows for further re-use of datasets and knowledge. It also allows for verification of the NATURANCE projects results.

Some of the data gathered in the NATURANCE project will never be directly opened due to security and privacy reasons. As of writing this document, we expect this to only apply to stakeholder contact information and interview transcript data.

This is the first version of the DMP for the NATURANCE project. It is intended to be a living document, which will be updated regularly when new insights are gained.



1. Introduction

The NATURANCE project has received funding from the European Union's Horizon 2020 research and innovation program call HORIZON-CL6-2021-BIODIV-01-06 under grant agreement number 101060464.

During the lifecycle of a project the generated data may evolve in terms of volume, type, speed, sources, analysis processes and so on. It is therefore important to consider a DMP as a dynamic framework that might change during the project, thus new data, new types of data or changes in consortium policies, processes or composition will occur in the future. Every 6 months this DMP will be revisited to ensure that it is up to date. The open software standards also often differ between different partners and countries. All final datasets and collection activities will be based on the FAIR data management (Findable, Accessible, Interoperable and Reusable) and the following documentation will provide the framework for the datasets and processes created by NATURANCE partners over the next 4 years.

This DMP presents a summary of the key data that will be used and generated by the NATURANCE project, and how this data will be managed according to FAIR-principles. For the composition of this deliverable, the DMPs of well-established HORIZON2020-projects were consulted, including COACCH, RECEIPT, and MYRIAD-EU. The deliverable is organized as follows: Section 2 summarizes the data that is collected and applied in each of the work packages of the project; Section 3 describes how the data management is done in accordance with FAIR-principles; Section 4 describes the allocation of resources for the data management; and Section 5 presents data security.

2. Data summary

The NATURANCE Project involves both primary and secondary data collection, and the generation of new data. Data collection is planned for activities in work packages 2, 3, and 4. Data collection under these tasks will mainly comprise stated stakeholder interviews, as well as preference surveys and/or choice experiments. Specific examples of activities and secondary data sources employed include:

- Work package 2 is concerned with obtaining data and insights from experts and stakeholders. The primary method of data collection are innovation labs, where information is exchanged between experts and researchers, which is then used for a variety of purposes, such as the generation of new scientific output, or the validation of research generated by the NATURANCE project. The generation of data, therefore, includes collaborative dialogue activities, such as online citizen forums, capability-building forums, innovation labs, and expert working groups. But also outreach activities may function as a method to obtain relevant information, including festivals and “webstivals”, webinars and/or podcasts, and contributions to major conferences.



- Stock-taking exercises in Tasks 3.1 and 4.1 will consist of reviews of established academic literature. The collection of literature will be done using search engines including Google Scholar, Web of Science, and Scopus. Resulting collections of literature will become available as bibliographies in subsequent deliverables. Data used for the meta-analysis planned for Task 4.1 will comprise of findings from published academic literature from databases including ESVD, and will become openly available on Zenodo.
- Data collection in WP2 and 3 will largely entail conducting and analyzing stakeholder interviews and surveys. The methods of collecting data in this work package are manifold, including direct research activities using human subjects, such as interviews and surveys, but also collaborative dialogue activities, such as online citizen forums, capability-building forums, innovation labs, and expert working groups. Moreover, interviews will be recorded and transcribed for accurate analysis at later stage. This is done in full compliance with ethical and privacy policies as described in the NATURANCE ethics plan.
- Task 4.2 promises innovations in the assessment of (co-)benefits of NBS. This task will involve both primary data collection and assessments using secondary data. Novel survey collection and assessments, such as contingent valuation or choice experiments, will be conducted to estimate benefits generated by NBSs, such as aesthetic or recreational values. Collected data will be anonymous and untraceable to individual respondents. Further details on data security and ethical aspects of the data collection is provided in Sections 5 and 6 in this document. Data collections will become publicly available on Zenodo after publication.
- Besides primary data collection using surveys or experiments, Task 4.2 will proceed in the development of novel methods to assess the risk-reduction potential of NBS for climatic risks, such as flood risk. This task requires assessment methods that estimate such climatic risks under current and future climate- and socio-economic conditions. Data underlying these assessments is acquired from databases, such as IIASA's SSP database (<https://doi.org/10.1016/j.gloenvcha.2016.05.009>). Assessments of flood risk, which will be used to estimate the risk-reduction potential of specific NBSs, are done using the global flood risk model GLOFRIS, output of which is already publicly available at: <https://www.wri.org/aqueduct/data>. Exact data used for the assessments done in Task 4.2 will be provided on Zenodo.



WP5 is concerned with the dissemination of research findings. This includes:

- mobilizing and fostering engagement of key knowledge networks and policy & action partners (in WP1), as well as other policy and societal actors, citizens and EU Climate Pact Ambassadors,
- designing high-quality and tailored dissemination, exploitation and communication activities, e.g. events & campaigns, instruments (e.g. social media,) and products (e.g. innovation briefs, compendia of good practices), monitor and periodically assessing the impact of the project activities and develop conditions for sustained exploitation beyond the life-span of the project.

3. FAIR data

3.1 Making data findable & metadata

Data collections used in the NATURANCE project, that are suitable to be made openly accessible will be uploaded to Zenodo (<https://zenodo.org/communities/naturance/>), where they will receive a unique DOI-number. The data will be accompanied by rich metadata and registered/indexed in a searchable database. The presence and location of data collections will be thoroughly advertised by NATURANCE Deliverables, publications and various dissemination channels (including the project-website and the knowledge networks).

Naming conventions for data files and folders are yet to be established. Consistent keywords for searching will also be defined later.

3.2 Making data openly accessible

All primary and rehabilitated (i.e. restored to good conditions) secondary data, except for data held sensitive under the EU's data protection law, will be stored in Zenodo, providing free, immediate and permanent access. To make them findable, the data will be equipped with Digital Object Identifiers (DOI) and shared through knowledge portals such as **Climate ADAPT**.

Table 1, below, describes the possibilities for open access provision per type of data used and produced in NATURANCE.

Table 1: Data types and open access availability

#	Data type	Data openly available?	Justification and alternative solution
1	Stakeholder contacts collection	No	Although the contacts of the collection are professionals' contacts that may be publicly available, the consortium cannot publish them due to GDPR considerations.



			Furthermore, the contact information will only be used in the context of the NATURANCE project and will be destroyed at most 1 year after the project ends. The statistical information on the stakeholder data (such as how many, from which countries, which professions etc.) will be integrated in the public report.
2	Expert interview data	No	The data from expert interviews (recordings, protocols and transcriptions) will not be published as primary data due to privacy and security concerns and will be destroyed at most 1 year after the project ends. Anonymization is not considered as an alternative; the sample size allows drawing conclusions on the respondents.
3	Database of NBS valuation	Yes	Will become available on the NATURANCE Zenodo community
4	Geospatial data (Used for D4.2)	Yes	Will become available on the NATURANCE Zenodo community

3.3 Making data interoperable

To ensure interoperability, all data will be stored in formal, accessible, shared, and broadly applicable formats, and described using vocabularies that follow FAIR principles.

Table 2, below, provides an overview of data types that may be used and produced by NATURANCE. Each data type is accompanied with one or more recommended data formats.

Table 2: Frequently used data types

Type of data	Recommended formats
Tabular data	CSV, XLSX
Textual data	Rich Text Format (.rtf). Plain text, ASCII (.txt). Hypertext Mark-up Language (.html).
Geospatial data	NetCDF; TIFF
Image data	PNG; SVG
Audio data	Audio Codec (FLAC) or MPEG-1 Audio Layer 3 (.mp3) if originally created in this format.
Video data	MPEG-4
Documentation	MS-Word; PDF
Papers & Articles	LaTeX; MS-Word



When existing datasets are used or new datasets are generated in the context of the NATURANCE project, they will be added to the NATURANCE Zenodo community (<https://zenodo.org/communities/naturance>) if the license allows it.

If or when following these recommendations is not possible, the party responsible for the addition of the dataset will provide a motivation. Possible motivations may include being unable to modify the data from its initial format for licensing reasons or being forced to use specific software that cannot handle the recommended format. Having to spend additional effort to change the format is not a valid reason for deviating from the recommendations.

3.4 Making data interoperable

In order to foster data re-use, all open datasets will be described with accurate and relevant attributes and released with a clear and accessible data usage license.

In addition, concerning the meta-analysis planned for Task 4.1, and survey analysis for Task 4.2, datasets and results will be uploaded to open-source databases that specialize in ecosystem valuation, including ESVD. Such databases are often consulted to gather information of values of particular ecosystem services. This way, the project also contributes to the knowledge base of ecosystem services more broadly.

4. Data ownership and allocation of resources

Uploading the data and ensuring it complies with FAIR-principles will be the responsibility of the project-partner in charge of the deliverable the data is used for. The specific project-partner remains responsible for the quality for the duration of the NATURANCE project. Coordination of data provision on Zenodo will be done by VU.

The costs for making the data generated by the project 'FAIR' will be minor, and are included as part of the budget assigned to the project partners responsible for collecting the data (VU, KIT, IIASA, CMCC).

5. Data security

The Zenodo NATURANCE community will be used for the collection and sharing of all (finalized) datasets and documents in the NATURANCE project. Zenodo offers a data retention for the lifetime of the repository, which is stated to be guaranteed for at least 20 years.

Zenodo's policies will be adhered to by all members of the NATURANCE project. For clarity and reference, a copy of version 1.0 is provided here:

Content



- **Scope:** All fields of research. All types of research artifacts. Content must not violate privacy or copyright, or breach confidentiality or non-disclosure agreements for data collected from human subjects.
- **Status of research data:** Any status is accepted, from any stage of the research lifecycle.
- **Eligible depositors:** Anyone may register as user of Zenodo. All users are allowed to deposit content for which they possess the appropriate rights.
- **Ownership:** By uploading content, no change of ownership is implied and no property rights are transferred. All uploaded content remains the property of the parties prior to submission.
- **Data file formats:** All formats are allowed - even preservation unfriendly.
- **Volume and size limitations:** Total files size limit per record is 50GB. Higher quotas can be requested and granted on a case-by-case basis.
- **Data quality:** All information is provided “as-is”, and the user shall hold Zenodo and information providers supplying data to Zenodo free and harmless in connection with the use of such information.
- **Metadata types and sources:** All metadata is stored internally in JSON-format according to a defined JSON schema. Metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema (according to the OpenAIRE Guidelines).
- **Language:** For textual items, English is preferred but all languages are accepted.
- **Licenses:** Users must specify a license for all publicly available files. Licenses for closed access files may be specified in the description field.

Access and Reuse

- **Access to data objects:** Files may be deposited under closed, open, or embargoed access. Files deposited under closed access are protected against unauthorized access at all levels. Access to metadata and data files is provided over standard protocols such as HTTP and OAI-PMH.
- **Use and re-use of data objects:** Use and re-use is subject to the license under which the data objects were deposited.
- **Embargo status:** Users may deposit content under an embargo status and provide an end date for the embargo. The repository will restrict access to the data until the end of the embargo period; at which time, the content will become publicly available automatically.

Removal

- **Revocation:** Content not considered to fall under the scope of the repository will be removed and associated DOIs issued by Zenodo revoked. Please signal promptly,



ideally no later than 24 hours from upload, any suspected policy violation. Alternatively, content found to already have an external DOI will have the Zenodo DOI invalidated and the record updated to indicate the original external DOI. User access may be revoked on violation of Terms of Use.

- **Withdrawal:** If the uploaded research object must later be withdrawn, the reason for the withdrawal will be indicated on a tombstone page, which will henceforth be served in its place. Withdrawal is considered an exceptional action, which normally should be requested and fully justified by the original uploader. In any other circumstance reasonable attempts will be made to contact the original uploader to obtain consent. The DOI and the URL of the original object are retained.

Longevity

- **Versions:** Data files are versioned. Records are not versioned. The uploaded data is archived as a Submission Information Package. Derivatives of data files are generated, but original content is never modified. Records can be retracted from public view; however, the data files and record are preserved.
- **Replicas:** All data files are stored in CERN Data Centers, primarily Geneva, with replicas in Budapest. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape on a nightly basis.
- **Retention period:** Items will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental program defined for the next 20 years at least.
- **Functional preservation:** Zenodo makes no promises of usability and understandability of deposited objects over time.
- **File preservation:** Data files and metadata are backed up nightly and replicated into multiple copies in the online system.
- **Fixity and authenticity:** All data files are stored along with a MD5 checksum of the file content. Files are regularly checked against their checksums to assure that file content remains constant.
- **Succession plans:** In case of closure of the repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories.

6. Ethical aspects

Good research practices are based on fundamental principles of research integrity. They guide researchers in their work as well as in their engagement with the practical, ethical and intellectual challenges inherent in research. These principles will be adhered to by all NATURANCE project partners who are expected to be familiar with the European Code of



Conduct (European Science Foundation, & All European Academies, 2011). These principles are:

- Reliability in ensuring the quality of research, reflected in the design, the methodology, the analysis and the use of resources.
- Honesty in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair, full and unbiased way.
- Respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment.
- Accountability for the research from idea to publication, for its management and organization, for training, supervision and mentoring, and for its wider impacts.

For the ethical aspects of stakeholder interactions, we refer to the NATURANCE deliverable D6.2 Ethics Management Plan. This document includes (among others) informed consent form.

Concerning primary data collection under Task 4.2, in particular the survey/experiment that is planned there, ethical guidelines of VU-Amsterdam will be followed strictly (<https://vu.nl/nl/over-de-vu/faculteiten/faculteit-der-betawetenschappen/meer-over/research-ethics-review-committee-beta>). These guidelines provide clear instructions to meet ethical standards regarding data collection using human participants.