



# D19.5

## 1st UPDATED DATA MANAGEMENT PLAN

### WORK PACKAGE 19 – PROJECT MANAGEMENT

LEADING BENEFICIARY: UNIVERSITY OF HELSINKI

Author(s):	Beneficiary/Institution
Ari Asmi	University of Helsinki
Jean-Daniel Paris	CEA
Zhiming Zhao	University of Amsterdam
Anders Tjulin	EISCAT association
Helen Glaves	NERC
Antti Pursula	CSC

**Accepted by:** Werner Kutsch

**Deliverable type:** [REPORT]

**Dissemination level:** PUBLIC

**Deliverable due date:** 31.5.2016/M12

**Actual Date of Submission:** 31.5.2016/M6



Project internal reviewer(s):

<b>Project internal reviewer(s):</b>	<b>Beneficiary/Institution</b>
Werner Kutsch	ICOS ERIC

Document history:

<b>Date</b>	<b>Version</b>
30.5.2016	Draft for comments
31.5.2016	Accepted by WK



# UPDATED DATA MANAGEMENT PLAN

## ABSTRACT

This document presents the updated ENVRIplus data management plan (DMP), developed by the project Data Management Team in their workshops and email conversation. This is the second version of the DMP and will be updated annually. The document describes shortly the project, scope of DMP, default policies and then explains the current plan on data set collection level.

## DATA MANAGEMENT TEAM

The DMP team met first time in during ENVRIplus kickoff meeting in Helsinki, Finland in May 2015. The initial meeting concluded with discussion on the overall procedure towards proper data management within the ENVRIplus cluster. The draft version of the document was then developed during the first 6 months of the project, using emails and teleconferences within the DMT, and with the relevant project participants.

The overall structure of this document is based on the DMP templates available in the UK Digital Curation Center (<http://www.dcc.ac.uk/resources/data-management-plans>).

The DMP has been updated by email conversation in the May 2016. Updated parts (only titles in APPENDICES) in 2016 are marked in PURPLE.

## ENVRI PLUS AS A PROJECT

The objective of ENVRI<sup>PLUS</sup> is to provide common solutions to shared challenges for European Environmental and Earth System Research Infrastructures (RIs) in their efforts to deliver new services for science and society. The main actions of this work are (with their relevant data types and sources):

1. Improve the RI's abilities to observe the Earth System, particularly in developing and testing new sensor technologies, harmonizing observation methodologies and developing methods to overcome common problems associated with distributed remote observation networks;
  - a. measurements from new sensors under tests or intercomparison of sensors and sampling systems – these are rarely needed for long term preservation, but important during project period;
  - b. Data collection of field measurements from RI scientists;
2. Generate common solutions for shared information technology and data related challenges of the environmental RIs in data and service discovery and use, workflow management and documentation, data identification and citations methodologies, data provenance, and user characterization and interaction;
  - a. Software products for data management and applications;
  - b. RI questionnaires and interview results on existing and future needs for IT solutions;
3. Develop harmonized policies for access (physical and virtual) for the environmental RIs, including access services for the multidisciplinary users;
  - a. RI questionnaires and interview results on existing and future needs for access, including experiences on new TNA structures, details of the applications, etc.
  - b. (potentially) data sets provided by TNA activities
4. Investigate the interactions between RIs and society: Find common approaches and methodologies how to assess the RIs' ability to answer the economical and societal challenges, develop ethics guidelines for RIs and investigate the possibility to enhance the use Citizen Science approaches in RI products and services;



- a. Questionnaires and interviews of RIs
  - b. Analysis results of societal interaction methodologies
  - c. Citizen science participant data, and produced data sets
5. Ensure the cross-fertilisation and knowledge transfer of new technologies, best practices, approaches and policies of the RIs by generating training material for RI personnel to use the new observational, technological and computational tools and facilitate inter-RI knowledge transfer via a staff exchange program;
    - a. Experience (questionnaires) from personnel exchange
    - b. Training materials
    - c. Personnel exchange personal information
  6. Create RI communication and cooperation framework to coordinate activities of the environmental RIs towards common strategic development, improved user interaction and interdisciplinary cross-RI products and services.
    - a. Reports of meetings, their discussions and participants
    - b. Dissemination material

Thus the overall data products generated are heterogeneous, and mostly for internal consortium use, and expected storage needs are minor in volume. Relatively lot of “soft” research information is collected in contrast to direct observational “hard” research data.

### ***Related policies***

The RIs have their own associated data management policies, either existing or in development. Similarly, the Beneficiaries might have their own overarching policies, although the effect of these will be controlled by the relationship defined in the Consortium Agreement document of the ENVRIplus. This DMP is descriptive to only the data directly generated within the project ('foreground' in the CA).

### **Scope of the DMP**

The DMT had several discussions on the scope of the DMP. After these discussion it was decided the that the ENVRIplus project DMT will have following scope definitions:

- It covers all research data sets produced directly in ENVRIplus project activities mentioned in the Description of Action. It specifically does not cover the auxiliary data sets produced by the Research Infrastructures using the ENVRIplus products (e.g. using common harmonized metadata system, etc.) as these are subject to the data management plans of the individual Research Infrastructures.
- Even though the project partners or non-partner personnel taking a part of the Trans National Activities (TNA) are strongly suggested to use the overall DMP structure presented here, the data sets produced by non-ENVRIplus partners in their TNA activities are outside the scope of this DMP and the Consortium Agreement. However, clauses regarding DMP used can be included in the TNA agreements.
- This DMP covers observational data, results from questionnaires, interviews and developed software. For the software, only relevant parts of the DMP template need to be covered.
- The overall level of the DMP is more on the “collection” level than individual data sets. Similar data sets from similar source are considered a collection, if the policies attached to them are similar or the data sets can be considered to form a single whole.



## DEFAULT ENVRIPLUS DATA POLICIES

This section presents the default data policies, which can be then referred in individual data collection policy results later in the document.

### Metadata and documentation

ENVRIplus contains a work package on harmonization of metadata for Earth System observation RIs. This work is naturally intended for long-term harmonization, and thus most of the data produced during the project will not immediately benefit for this work. However, the RIs taking part of the project have their own metadata schemes, and the overall **ENVRIplus data policy is to use these schemes and documentation methods as much as possible**. Use of suitable international standards (e.g. INSPIRE directive) are strongly recommended.

For non-observational data or data produced by non-ENVRIplus partners, other methods are needed. Major part of the project is concerned on human or policy data collected from the participating RIs. These data sets are not part of the typical RI collected data sets and thus require more attention on metadata and documentation, which need to be defined case-by-case basis, with overall requirement of ensuring data usability for re-use.

Similarly, the citizen science data products and non-ENVRIplus partners data from Pilot Transnational Access are not necessarily covered by typical RI metadata standard procedures. The overall ENVRIplus data policy in these cases requires inclusions of sufficient and suitable metadata for these datasets. The WP leaders of relevant WPs are required to provide information on these datasets on the next iteration of the DMP.

### Ethical and privacy issues

Standard (RI) observational data sets in ENVRIplus are not expected to have ethical or privacy concerns. This should be evaluated on case-to-case basis by the WP leader in question.

However, the questionnaire data and Citizen Science collected datasets can have significant privacy concerns. For data sets which have these concerns, the ENVRIplus general policy is to follow the procedure:

1. Collection of any private data (including e.g. opinions, names, positions, etc) is to be avoided if not useful for the purpose of the study.
2. If such information is needed, the target of the study must be informed before the data collection about the
  - a. Collected information,
  - b. Why it is needed,
  - c. How the targets were selected,
  - d. Who has access to the data,
  - e. Of any anonymization scheme (if any) is used,
  - f. How the data will be analyzed,
  - g. How the target and either agree on disagree on the terms,
  - h. How the access to the data will be controlled and,
  - i. How long and where the data will be stored.
3. The data retention must be carefully considered, and if the raw data sets do not need to be stored, they must be destroyed efficiently after use.
4. Data storage must be adequate to the level of sensitivity of the data.



5. Access control of the data must be adequate and clearly defined, including access policies in the long-term storage.

All of these procedures of course need a formal ethical review for doing the study. The template and procedure above should be used as the basis of the ethical approval requests.

**UPDATE 2016:** The procedure for questionnaires is currently (as of May 2016) submitted for review in the University of Helsinki Ethical review board. This is to make sure that each questionnaire in the project will not need to be separately evaluated in participating institutions. Each ENVRIplus beneficiary making a questionnaire inside the project will instead follow the accepted procedure (basically outlined above, with minor changes.). The additional questionnaires are currently being launched.

### **IPR issues**

Ownership of the data sets produced within the project partners are defined in the Consortium Agreement. Ownership of non-partner data sets produced within the project envelope (e.g. Citizen Science and TNA partners) must be clearly defined in the Terms of Collaboration to be defined with these partners.

**UPDATE 2016:** Ownership of the TNA partners data was in the end left to the TNA partners. This also means that the datasets produced in their activities are outside of the ENVRIplus data products, and of this document. No citizen science activities have yet been started.

### **Storage and backup during project time**

The individual partners (and RIs) are responsible for the project time storage and backup of the collected data sets. However, the ENVRIplus has collaboration with the EUDAT2020 H2020 project, which provides B2DROP service for data set storage (max size 2Gb per file), sharing and backup. The DMT recommends partners to use this facility.

### **Access management**

For many datasets produced, the storage and access management can be done using the RIs own repositories. Access to other ENVRIplus partners to these data sets (not others) must be provided as defined in the Consortium Agreement. For the data stored in EUDAT2020 services, the access management will be provided by their Access Management modules. Access must be provided to the Commission officials and their appointed reviewers. Access to sensitive data must be adequately controlled.

**UPDATE 2016:** Ownership of the TNA partners' data was in the end left to the TNA partners as they are not directly considered to be primary results of the Project and would have otherwise made the TNA participation too low for the main use cases. This also means that the datasets produced in their activities are outside of the ENVRIplus data products, and of the scope of this document. No citizen science activities have yet been started.

### **Retention and preservation**

Decision on retention of the data sets should be done carefully. In the case of sensitive data, the retention must consider the absolute need of storing such information, and effective anonymization procedures must be followed. If sensitive data is to be destroyed, such actions must be done effectively and with clear responsible person determined before data collection.



The retention of the general data sets done on the decision of the WP leader in question, however this decision can be changed if the data originator or other project participant complains the situation to the ENVRIplus Executive Board. Project Executive Board has final say on the retention issues. In the data retention decision the following aspects are to be considered: 1) Re-usability of the data (including metadata), 2) needed resources for long term storage (size, access), 3) expected storage period, 4) possibility of external data storage using non-project related repositories.

### **Long term preservation plan**

Data selected to preservation must include long-term preservation plan. Data can be stored in generally accepted long-term preservation system, with minimum of 10 years of guaranteed storage. Such preservation systems must also provide acceptable sustainability plan for migration to other storage systems. The DMT suggests use of EUDAT services (B2SHARE, B2SAFE) for such storage.

### **Sharing policy**

Unless required by Consortium Agreement or sensitive data policies, all ENVRIplus data products must be made openly accessible and findable (i.e. via metadata harvesting access) in reasonable time after data collection. Default time for this is 6 (six) months data end of data collection. This is controlled by the relevant WP leaders. Exceptions for this rule can be made on for good reasons by Project Executive Board on request.

### **Restrictions on sharing**

Only restrictions on sharing data sets are connected to sensitive information, or temporary grace period right after data collection. Otherwise all ENVRIplus results are open to use for any purpose, although all data sets must be appropriately licensed using very open licensing policy (e.g. Creative Commons 4.0BY or similar).

**UPDATE 2016:** WP5 raw interview data is not shared outside of the participating institution. Processed (not personal) datasets from them are documented in Deliverable D5.1. All deliverables are available in the project website and their publication in Zenodo is under consideration.

### **Responsible person**

Each data set produced must have an appointed responsible person. The data providing institution must provide a way to maintain the responsibility even in the case of personnel change – i.e. a clear procedure must be detailed on how the responsibility can be transferred and how this can be controlled.

### **Resources used**

Data set storage, curation and maintenance costs are valid ENVRIplus costs. The long term resources needed for storage must be considered.



## DATA SETS DESCRIPTION AND PLANS

Note: This is done in the initial DMP in data collection level, i.e. we do not follow each separate data set, but instead work on identified major types of data produced in the project. The DMT together with WP leaders found 7 data collection types, which will be described in this section:

1. Project documentation
2. Questionnaire and interview data
3. Software products
4. Testing data
5. Citizen Scientist collected data
6. Data collected by Transnational Access participants

As of 2016, only data types 1 and 2 have been this far materialized.

### **1. Project documentation**

#### **Data collected**

Deliverables, meeting notes, project personnel lists, email lists, email messages, other project documents (agreements, MoUs, etc.) collected by the Project Management Team (PMT). The work related to these datasets is done in the Work Packages 18 and 19.

#### **Collection method**

Deliverable collection is defined in the ENVRIplus deliverable guidelines. Meeting notes are collected by the meeting rapporteurs in the ENVRIplus ActiveCollab site. Project personnel lists, including email addresses, are collected on ActiveCollab site by the persons in question. Email messages of the project management team are collected on personal email archives.

**UPDATE 2016:** Special attention is now made to ensure that the project personnel email addresses are NOT available in any publicly available documents in easily harvestable form.

#### **Metadata and documentation**

Documents will be included with necessary documentation to understand them later on. Email archives have their internal metadata schema attached to messages. The formal metadata will be included to the documents when they are inserted to the project long term storage by the PMT.

#### **Ethical and privacy issues**

Deliverables have no privacy issues as they are intended and clearly labeled to be open knowledge. Meeting notes are collected to avoid any personal information. Personnel lists, including email addresses, are collected only for project purposes, and their subjects are clearly informed on their collection – only name, institution, email address and project tasks are stored. The information is done using the standard Finnish Personal Data Act (523/1999) section 10 declaration sheet (**ANNEX I**). Email messages have potential privacy issues, although only directly project-related emails are stored, and the control of the store is by the recipient (not overall project management).





## **IPR issues**

ENVRIplus default.

## **Storage and backup during project time**

Documents are stored in the ENVRIplus ActiveCollab site. The contract includes back up services. Additional backups are made by the project staff. Email archives are regularly backed up by the PMT on personal level.

## **Access management**

ENVRIplus ActiveCollab has an Access management mechanism, controlled by the Project Management Team. Only Beneficiaries have access to the collected email addresses or personnel list. Deliverables are open to all and are actively published in the project public website. Other project documents (i.e. agreements, MoUs) access is controlled by the Project Management Team and handled in the ActiveCollab site. Access is generally open for all Beneficiaries as defined by the Consortium Agreement.

## **Retention and preservation**

Deliverables will be retained. The email lists and personal information will be deleted latest 4 years after project ending to maintain necessary communications for reporting and financial auditing purposes. For email archives and other documents, the decision on retention will be done before project end by the PMT, after consulting the project Executive Board.

## **Long term preservation plan**

Retained documents will be stored with the minimum of 10 years. The long term preservation is done by using one of the following facilities (to be decided at Project finalization): OpenAire for documents suitable for that service, EUDAT B2SHARE for deliverables, etc. For documents not suitable for either, the Finnish national TTA ETSIN service is used.

## **Sharing policy**

Deliverables are openly shared, unless specifically determined to be confidential (in Description of Action, none are). Email addresses are only available to project participants. Stored documents (e.g. MoUs) are only shared to project beneficiaries. Email archives are not shared on default. [As of 2016, special attention to potentially publishing Deliverables in the ZENODO service is under investigation \(see ANNEX V for the current questions.\) This has not yet been accepted by the ENVRIplus Executive Board.](#)

## **Responsible person**

Deliverable, email list, personal information and general document management is by Magdalena Brus from DMT. Both Ari Asmi and Magdalena Brus are responsible for their own email archive. University of Helsinki will organize replacements for this staff if needed (personnel changes, long absences etc.).

## **Resources used**

ActiveCollab is funded by the ENVRIplus management budget. The work for data curation and storage is part of the Management PMs. Long term storage resources are institutional or handled by other EU funded activities (e.g. EUDAT2020).



## **2. Questionnaire/interview data**

### **Data collected**

Initially the questionnaires themselves, or audio recordings of the interview. For longer term storage, the reports collecting the findings of the questionnaires/interviews.

### **Collection method**

Individuals specifically selected for collecting such information collect the information. They are detailed in the Ethical Review of the questionnaire/interview done before the collection. The collection includes the initial information retrieval, analysis, consultation with the people interviewed (for correct interpretation) and synthesis. All participants are informed (via [header information](#) or consent forms) of the data collection, processing and access before data collection.

### **Metadata and documentation**

The collected data sets will have machine created metadata. The end result (reports) will have the standard ENVRIplus documentation format (i.e. deliverable template, [see attachment II](#)) with included metadata.

### **Ethical and privacy issues**

All such activities go through ethical review, as they potentially include personal data and opinions. [Procedures are summarized for project internal use in ANNEX IV](#). Each questionnaire needs a responsible person who is responsible on the data security, access and destruction.

### **IPR issues**

ENVRIplus default.

### **Storage and backup during project time**

Raw documents are securely stored by the interviewing or collecting [clearly identified responsible](#) person. Responsibility of this storage and related backups is with this person. The interpreted reports are stored in the Project internal document management system, which have automatic back up systems.

### **Access management**

Access to the raw documents are only accessible to the [clearly identified responsible](#) person doing the interviews or data collection. He or she can provide access to the raw information to the participant during the after-interview discussion on conclusion for the final documentation. The final results in the form of the report are public, and meant for general use in the form of a deliverable. These documented final results will also be shared to collaborating H2020 projects, currently there is a data sharing MoU with EUDAT2020 project.

### **Retention and preservation**

The raw data (questionnaires and audio recordings) will be destroyed immediately after processing to the final report format. The resulting documents are stored as Deliverable reports (see earlier point). [Clearly identified responsible person is responsible on data destruction and potential preservation](#).



**Long term preservation plan**

For the deliverables, the long term preservation is explained in the previous section.

**Sharing policy**

Deliverables are publicly available. None of the raw data will be shared in or out to the Project.

**Responsible person**

Each questionnaire/interview will have a **clearly identified** responsible person, defined when doing the Ethical Review. The responsibility of final reports are with the Project Management Team.

**Resources used**

Local resources per partner with the raw results. Deliverables are using Project Management Team resources.



### **3. Software products**

#### **Data collected**

Developed software packages or modules. These are specifically those packages developed in the ENVRIplus project, i.e. these do not include existing software.

#### **Collection method**

Done during the project period, mostly in Theme II work packages. The data (software codes) are collected during work using suitable versioning systems, which can be controlled via web interfaces.

#### **Metadata and documentation**

Documented by inline comments and documentation principles of each IT partner in question.

#### **Ethical and privacy issues**

None

#### **IPR issues**

All produced software will stay as the ownership of the original authors. However, all will need to be licensed using an open software license, as described in the Grant Agreement.

#### **Storage and backup during project time**

IT providers own service for version control will be used, with their internal backup systems.

#### **Access management**

Provided modules will have free access after project period. During the project work, the produced draft modules will be available to RI participants, and the author can also share it more widely. The access practicalities depend on the software platform, but these can be also shared via the project document management system.

#### **Retention and preservation**

Final versions of the program codes are stored and preserved. Draft versions can also be stored if the author or RI representative require this.

#### **Long term preservation plan**

The long term preservation of the software uses ENVRIplus standard methodologies (actual method will be chosen latest on project end).

#### **Sharing policy**

All final software products are to be shared publicly.

#### **Responsible person**

Individual researchers/programmers are responsible on version control, documentation and initial storage. WP leader is responsible to identify all software produced in his/her WP and to control that the final versions are made available.

#### **Resources used**

Partner individual version control systems. Long term storage: as ENVRIplus standard.



#### **4. Testing data**

##### **Data collected**

Data sets collected during testing of e.g. new equipment, or endurance tests. These are typically of only engineering interest, and do not have much additional value.

##### **Collection method**

Depends on each test.

##### **Metadata and documentation**

Only local metadata – the results of the instrument tests are reported in Deliverables, but actual testing data sets are not specifically documented.

##### **Ethical and privacy issues**

None

##### **IPR issues**

ENVRIplus default.

##### **Storage and backup during project time**

During testing with the RI individual storage options, or ENVRIplus default methods.

##### **Access management**

Only to WP in question.

##### **Retention and preservation**

The WP leader will make the final decision on storing such test data. If the data is determined to be stored, it will then be included in the documentation process and will be stored according to ENVRIplus standard methodology.

##### **Long term preservation plan**

Usually not retained. If retained as ENVRIplus standard.

##### **Sharing policy**

Only within the WP. Retained data is openly available, and will be shared according to ENVRIplus standard methods.

##### **Responsible person**

Persons doing the testing, WP leader in question (for retention decision)

##### **Resources used**

Only local resources. If data is to be retained, then ENVRIplus standard.



## **5. Citizen Scientist collected data**

### **Data collected**

Data collected by Citizen Scientists during the (potential) Citizen Science pilot

### **Collection method**

Unknown yet – depends on the method chosen

### **Metadata and documentation**

This will be decided when the CS WP decides on the potential test case.

### **Ethical and privacy issues**

Most likely these will be issues. These will be investigated later in the project.

### **IPR issues**

ENVRiplus default.

### **Storage and backup during project time**

To be decided

### **Access management**

To be decided

### **Retention and preservation**

To be decided

### **Long term preservation plan**

If retained as ENVRiplus standard.

### **Sharing policy**

To be decided

### **Responsible person**

WP leader of Citizen Science.

### **Resources used**

To be decided



## **6. Data collected by Transnational Access participants**

### **Data collected**

Data sets collected Transdisciplinary Transnational Access Pilot (TTAP). The content will depend on selected TTAP projects.

### **Collection method**

Depends on each TTAP project.

### **Metadata and documentation**

We will require clear documentation and (discipline specific) typical and complete metadata records.

### **Ethical and privacy issues**

Depends on TTAP, but unlikely.

### **IPR issues**

The data ownership must remain with the TTAP participant, but we require these data to be openly available (as in the ENVIplus Grant Agreement). [This requirement still holds by May 2016.](#)

### **Storage and backup during project time**

Up to the TTAP participant. However, the Project will provide support for this and will require TTAP proposals to explain how this is done.

### **Access management**

To be decided, but should follow ENVIplus normal procedures.

### **Retention and preservation**

To be decided, in collaboration with the TTAP participant.

### **Long term preservation plan**

Up to the TTAP participant. However, the Project will provide support for this and will require TTAP proposals to explain how this is done.

### **Sharing policy**

Data must be made open and accessible. All produced data sets must follow the ENVIplus general directions, although exceptions can be made if necessary, on decision of ENVIplus Executive Board.

### **Responsible person**

TTAP participant and the TTAP WP leader.

### **Resources used**

TTAP participants' own data services, General ENVIplus resources.



## CONCLUSIONS

### ***IMPACT ON PROJECT***

This is the initial version of the Data Management Plan. It is clear that this document needs to be further developed, detailed and corrected during project period. However, it presents the overall DM principles in the ENVRIplus and most likely data types collected.

### ***IMPACT ON STAKEHOLDERS***

Many of the Data Management Plan actions are crucial for RIs, as this document will clearly present that the produced documents, software and key datasets are available to all personnel.







A document of ENVRI<sup>plus</sup> project - [www.envri.eu/envriplus](http://www.envri.eu/envriplus)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654182



## APPENDICES

### ANNEX I DESCRIPTION OF PROJECT INTERNAL PERSONNEL INFORMATION STORAGE

#### DESCRIPTION OF THE FILE

Personal Data Act (523/1999) section 10

Date of drafting: 29.10.2015

Use an enclosure if necessary.

1. Controller	Name  Ari Asmi
	Contact information (address, tel. etc...)  P.O.Box 48, 00014 UHEL, Finland, +358407709729, ari.asmi@helsinki.fi
2. The person in charge and/or contact person	Name Magdalena Brus
	Contact information (address, tel. etc...) P.O.Box 48, 00014 UHEL, Finland, +358504154762, magdalena.brus@helsinki.fi
3. Name of the register  (should describe the content)	Contact information of ENVRIplus project participants
4. The purpose for processing the personal data / the purpose for the use of a register  (If the processing of personal data is outsourced [external service providers are used for the processing], a mention about it can be included to this point)	Maintaining communication and organisational information for the ENVRIplus H2020 project. This is meant as a register of the project participants for the internal use of the project. The Grant Agreement of the project (between the University of Helsinki and European Commission) specifically mentions "In particular, the Coordinator shall be responsible for: .... keeping the address list of Members and other contact persons updated and available"

OFFICE OF THE DATA PROTECTION OMBUDSMAN      MODEL FORM      25 November 2004  
Tel: +358 10 36 66700  
Tel: +358 10 36 16670 (information service 9 a.m. to 3 p.m.) Fax: +358 10 36 66735 [www.tietosuojafi.fi](http://www.tietosuojafi.fi)

[Unofficial translation]



<p>5. Content of the register</p> <p>(For instance: name, address and telephone number of the data subject)</p>	<p>Name, professional title, address, telephone number, email address, institution they work in, (potentially) associated Research Infrastructure, involvement on specific tasks in the project.</p>
<p>6. Regular sources of information *</p> <p>(Which data is received, by whom and on what ground. For instance: consent or provision of a law)</p>	<p>The information is collected from the project participants, specifically from the Primary Investigators of project Beneficiaries.</p>
<p>7. Regular destinations of disclosed data and whether the data is transferred to countries outside the the European Union or the European Economic Area</p>	<p>Data is available to the Project Beneficiaries, and people working directly (project participants) within the project. Data is also available for the participants outside of the European Union (Switzerland, Norway), but the participating organizations have declared that they will follow the ethical regulations of the Horizon 2020.</p> <p>Similarly, the data can be shared to the European Commission services, if needed in their reporting or auditing purposes.</p>
<p>8. The principles how the data file/register is secured.</p>	<p>A. Manual register (place of storage and the methods of protection)</p> <p>Copy of the register is stored in encrypted drive of the personal work computer of M. Brus.</p> <p>B. Data register/ADP register (principles for the right to use a register, monitoring of the use and actual protection of hardware)</p> <p>Data is stored in the Activecollab project collaboration site, which is password protected. The project partners have a right to use the register. The site access is controlled by the Controller (Ari Asmi) and the Contact Person (Magdalena Brus). They also monitor the site access.</p>

\* Section 10 of the Personal Data Act does not require to make a note of this information, but it's recommendable to do so, regarding the right of access of the data subjects. Furthermore, it also describes the content of the register.

OFFICE OF THE DATA PROTECTION OMBUDSMAN      MODEL FORM      25 November 2004  
Tel: +358 10 36 66700  
Tel: +358 10 36 16670 (information service 9 a.m. to 3 p.m.) Fax: +358 10 36 66735 [www.tietosuoja.fi](http://www.tietosuoja.fi)

[Unofficial translation]



## ANNEX II

Removed in this version, as ANNEX IV is a newer version



## ANNEX III

### QUESTIONNAIRE ETHICAL REVIEW TEMPLATE

This is the Ethical review document prepared by Malcolm Atkinson and Cristina Alexandru (U. Edinburgh) to their ethical board. This can be used as basis for similar documents in other parts of the project.



## ANNEX IV

### Guidelines for ENVRIplus Questionnaires

Before making any questionnaires, consider the following requirements:

1. Collection of any private data (including e.g. opinions, names, positions, etc) **is to be avoided** if not useful for the purpose of the study. Make sure that all information you collect are **strictly required** for the actions described in the Description of Action document of the Grant Agreement with the European Commission.
2. If such information is needed, the target of the study must be informed before the data collection about the
  - a. Collected information,
  - b. Why it is needed,
  - c. How the targets were selected,
  - d. Who has access to the data,
  - e. Of any anonymization scheme (if any) is used,
  - f. How the data will be analyzed,
  - g. How the target and either agree on disagree on the terms,
  - h. How the access to the data will be controlled and,
  - i. How long and where the data will be stored.

**This is best done by including a header information on your questionnaire.**

**COPY/PASTE the text under the line on the next page your questionnaire and fill in the information in the BOLDED parts.**

3. The data retention must be carefully considered, and if the raw data sets (answers) do not need to be stored, they must be destroyed after use. We require that all personal information is destroyed latest on the project end, preferably immediately after the conclusions of the study are finalized.
4. You must indicate a responsible person for the questionnaire who is responsible that the material is adequately stored and handled. He or she is also responsible on the access control, data storage and on destruction of personal information.
5. Data storage must be adequate to the level of sensitivity of the data. Use of separate protected areas in the common ActiveCollab virtual platform can be arranged - please contact the project office if needed. Also protected services such as EUDAT B2DROP can be used in this context, if the access is carefully considered.
6. Access control of the data must be adequate and clearly defined, including access policies in the long-term storage (if needed).
7. You **MUST** fill any registration information legally required by your country of operations regarding the storage of personal information. This is **YOUR** responsibility.

**IMPORTANT: Including the header information is only valid if you are NOT collecting any sensitive information. For sensitive information, you MUST also include the consent form agreement.**

**YOU MUST INFORM THE ENVRIPLUS PROJECT OFFICE ([envriplus-coordination@helsinki.fi](mailto:envriplus-coordination@helsinki.fi)) ON ANY QUESTIONNAIRES YOU ARE CONDUCTING AND PROVIDE INFORMATION ON HOW THIS TEMPLATE IS USED.**



---

## INFORMATION ON THIS QUESTIONNAIRE

This questionnaire is meant to collect your professional knowledge related to the ENVRIplus research infrastructures which you are connected to. As we can also collect your name, position and professional position and potentially other personal information, it is important that you understand the reason and procedure of this questionnaire.

Information on ENVRIplus project: (<http://www.envriplus.eu>)

*ENVRIPLUS is a cluster project, funded by the European Commission Horizon 2020 programme, on collaboration of research infrastructures (RIs) in Environmental and Earth System sciences, built around ESFRI roadmap and associating leading e-infrastructures and Integrating Activities together with technical specialist partners. ENVRIPLUS is driven by 3 overarching goals: 1) favouring cross-fertilization between infrastructures, 2) implementing innovative concepts and devices across RIs, and 3) facilitating research and innovation in the field of environment to an increasing number of users outside the RIs. ENVRIPLUS organizes its activities along a main strategic plan where sharing multidisciplinary expertise will be most effective. It aims to improve Earth observation monitoring systems and strategies, including actions towards harmonization and innovation, to generate common solutions to many shared information technology and data related challenges, to harmonize policies for access and provide strategies for knowledge transfer amongst RIs. ENVRIPLUS develops guidelines to enhance transdisciplinary use of data and data products supported by applied use cases involving RIs from different domains. ENVRIPLUS coordinates actions to improve communication and cooperation, addressing Environmental RIs at all levels, from management to end-users, implementing RI staff exchange programs, generating material for RI personnel, and proposing common strategic developments and actions for enhancing services to users and evaluating the socioeconomic impacts. ENVRIPLUS is expected to facilitate structuration and improve quality of services offered both within single RIs and at pan RI level. It promotes efficient and multidisciplinary research offering new opportunities to users, new tools to RI managers and new communication strategies for environmental RI communities. The produced solutions, services and other project results are made available to all environmental RI initiatives, thus contributing to the development of a consistent European RI ecosystem*

Responsible person for this questionnaire: **[Name]**, **[email]**, **[Institution]**

You can always ask for further information from the responsible person above, or from the ENVRIplus project office: [envriplus-coordination@helsinki.fi](mailto:envriplus-coordination@helsinki.fi)

The questionnaire aims at providing **[SPECIFY INFORMATION TYPE, e.g. technical details on energy requirements of observations]** information on **[SPECIFY REASON]** regarding European public Research Infrastructures (RI). Answering the questionnaire is voluntary and you can stop answering at any moment. The questionnaire will be done using a **[SPECIFY, e.g. web-based form]**, and will take approximately **[XXX]** minutes to answer. You have been selected to answer the questionnaire as your professional capacity as the representative of the RI you are working with.

**[SELECT ONE:**



No personal data is requested, although you have an opportunity to optionally leave your contact information for further information, if needed.

**OR**

Your personal information: name, contact information, organization, and position in your organization are stored for analysis purposes.]

All data will be stored securely on [**FILL IN STORAGE LOCATION, e.g. Secure Servers in University of Helsinki, Department of Physics**] and will only be used within the framework of the ENVRIplus project. Access to the answers is restricted to the responsible person and the data analyzers selected by him/her. The answers will be analyzed offline. The questionnaire technical results and conclusions deduced from the results can be published within the ENVRIplus project deliverables, reports and documentation, however no personal information will be published in any form. All questionnaire answers will be deleted latest at the end of the ENVRIplus project. If your contact information is stored with your answers, you can also request to be informed on the reports and documents generated from the information collected in this questionnaire.





## CONSENT FORM

in addition to the information form above if sensitive information, e.g. age, sex, opinions, political or religious opinions, medical information or similar, is strictly needed for the questionnaire. Please contact the project office if you are unsure of this. This can be done as a virtual form. If some points are not filled YES, you must make sure that there is no possibility to continue.

---

## CONSENT FORM

**Title of Project:** ENVRlplus

Questionnaire title: **[title]**

Responsible person for this questionnaire: **[Name]**, **[email]**, **[Institution]**

**Date:** **[fill or auto-fill]**

**This questionnaire can include potentially sensitive information on you. Please confirm the following**

1. I confirm that I have read and understand the information for the above project and have had the opportunity to ask questions.

[yes]

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without my legal rights being affected.

[yes]

3. I agree for the answers to be recorded, and for each recording to be kept until associated report is completed. No identification information will published in any form and the personal information will be destroyed immediately after the questionnaire results are analyzed.

[yes]



## ANNEX V

### Consideration on publishing the Deliverables in ZENODO service.

*Some of the ENVRIplus partners asked if the project public deliverables could be published in Zenodo (or similar) service, giving them DOIs. The general agreement to publish (public) deliverables in ZENODO or a similar service was approved by the Commission officer. However, this document needs still to be discussed and approved in the ENVRIplus Executive Board. At this stage, this represents the issues to be discussed on this stage.*

1) Who does the submission? The submission process is a little involved, and has also a responsibility issue to take the relevant contributors (see below) into account. Even references can be included (but are not absolutely necessary). If coordination does this, some might be missed, and is an extra requirement, which would in some cases better fit to the main author (see 2).

DMT recommendation: The submission is done primarily by the main author. However, if needed the Project Office can do this as well. A mention and the DOI of the submission must then be submitted to the Project Office and to the Project internal Activecollab site.

2) Who are the authors? Even though we do have author list, the Zenodo metadata also includes term “contributors”. We need a policy on difference between these.

DMT recommendation: The authors of each deliverable should be all people who have significantly contributed to the creation of the document. For minor contributions, the metadata field "contributors" can be used. Responsibility of naming these persons in with the main author of the deliverable, although the deliverable (internal) review is responsible to check these issues.

3) Even though the GA and CA strictly require us to make the deliverables openly available, it is not so clear will this also include publishing (as much as Zenodo is publishing, see below). This is not so much a legal issue than a common decision in the consortium. Of course if all involved are clearly asked their permission, this is fine. We should aim for consortium-level agreement, but...

DMT recommendation: The Commission officer already agreed on publication also in Zenodo, so we suggest that we will have a General Assembly agreement on our internal policies.

4).. IF Zenodo is publication some might not be willing to do this. This is because at least traditionally parts of Deliverables have been used as parts of peer reviewed journal publications. This has been quite common practice in some fields, where the first report in the Deliverable is used as the basis of a more mature report in a journal. If Zenodo is publication, this might lead to issues when the journal asks if it is published earlier in some form. Even though our deliverables ARE peer reviewed (in a fashion), the Web of Science unfortunately does not see it so (a fact which I do not like, but what can one do).

DMT recommendation: For each individual case, we need to make sure we have agreement of all partners. This requires still further discussion within the Consortium.

5) We need a formal decision on using Zenodo, and not some other service (e.g. Figshare). We should have all the deliverables in a same place. It could be good to have the DOA in the same place as well, so we can use it as a “related publication”.



DMT recommendation: Single service (e.g. Zenodo) should be chosen. Key part is that it should be free, as reliable as possible, and preferably located within the European Union. The decision on which service (e.g. Zenodo, Figshare, B2SHARE) should be well explained.

6) In DOA we specify CC4.0BY, so this must be used in publishing

DMT recommendation: General Assembly should be informed also on license when making the decision.

7) The ENVRIPLUS project must be properly included in the metadata (field Funding)

DMT recommendation: Always include the Project number and funder in the metadata.

8) Keywords should be at least somewhat similar in similar deliverables

DMT recommendation: Project office should prepare list of potential keywords, which can then be supplemented by partners. ActiveCollab site is a good tool for this.

9) Which version is included? Now we have already several versions of the D5.1... DOIs currently do not properly handle versioning and child/parent/sibling relations, thus the old version must then always be remembered to be include in the “related publication” fields.

DMT recommendation: At least the version sent to the Commission must be then published. Further developments MUST use the Related Publication Field.

10) Zenodo offers possibility to “reserve” a DOI. This way it could be used to form a workflow a) upload a document b) include metadata c) get a DOI d) put proper citation information INTO the deliverable (e.g. first page) including the DOI e) change the file on the bottom of the Zenodo metadata input form. A little involved system, but proper citation method should be somehow included.

DMT recommendation: This procedure should be made a part of the General Assembly decision.

