

PERIVALLON

Protecting the **Eu**Ropean territory from organised **enV**ironmental crime through **intell**Ligent threat **detecti**ON tools

D1.2 - Initial Data Management Plan

WP1 - Project Management and Coordination



Project ref.	PERIVALLON HE - 101073952
Deliverable version date	24/05/2023
Deliverable version nr.	1.0
Lead Partner Organisation	UNIVIE
Dissemination Level	PU – Public

Document Information

D1.2 Initial Data Management Plan		
Deliverable number:	D1.2	
Version:	1.0	
Contractual delivery date:	31/05/2023	
Type*:	DMP — Data Management Plan	Dissemination Level**: P-Public
Contributing WP:	WP1	
Lead author(s):	Nikolaus Forgó Mariana A. Risetto Eva Korenjak Lalovič	UNIVIE UNIVIE UNIVIE
Contributor(s):	Jordan Thompson Despoina Chatzakou, Konstantinos Gkountakos Thanasis Kapoutsis, Anastasia Moumtzidou, Athanasios Petsanis, Klearchos Stavrothanasopoulos Luca Di Nuovo Eduardo Villamor Medina Dafni Delioglani Dries Borloo Eleni Gkozgkou Ana Ghiumiuşliu Vagia Pelekanou, Zeta Aggariti Vasiliki Efstathiou Piero Fraternali Ioana Paunescu Alessandro Castagnetti David Fortune Jose Santos, Juan Romero Luigi Căldăraru, Sever Calit Christoffer Hagberg, Jimmy Berggren Nir Haimov	CENTRIC CERTH DYLOG ETRA DRAXIS DWG HP IGP KEMEA MT POLIMI RBP SAFE SAH SATCEN SMOB SPA TMR
Reviewer(s):	Kieran Dennis (internal review) Luigi Căldăraru (internal review) Dimitris Myttas (security assessment control)	CENTRIC SMOB KEMEA

*Type. R – Document, report; DMP – Data Management Plan; OTHER; DEM – Demonstrator, pilot, prototype; ETHICS

**Dissemination Level. PU – Public, fully open, e.g., web (Deliverables flagged as public will be automatically published in CORDIS project’s page); SEN - Sensitive (limited under the conditions of the Grant Agreement); Classified R-UE/EU-R – EU RESTRICTED under the Commission Decision No2015/444.

Document History

Version	Issue Date	Author(s)	Content and changes
0.0	31/01/2023	Nikolaus Forgó Eva Korenjak Lalovič Mariana A. Risetto	Design of table of contents
0.1	22/02/2023	Nikolaus Forgó Eva Korenjak Lalovič Mariana A. Risetto	Final and accepted TOC
0.2	24/04/2023	Nikolaus Forgó Eva Korenjak Lalovič Mariana A. Risetto	Final draft version of D1.2
0.3	08/05/2023	Nikolaus Forgó Eva Korenjak Lalovič Mariana A. Risetto	Final version submitted to internal reviewers
0.4	10/05/2023	Kieran Dennis	Internal review
0.5	11/05/2023	Luigi Căldăraru	Second internal review
0.6	15/05/2023	Eva Korenjak Lalovič Mariana A. Risetto	Version ready for security assessment
0.7	22/05/2023	Eduardo Villamor	Security assessment control passed and minor updates and comments
0.8	23/05/2023	Mariana A. Risetto	Version ready for submission to Project Coordinator
1.0	24/05/2023	Eduardo Villamor	Version ready for submission to EC

Disclaimer

This document contains material, which is the copyright of certain PERIVALLON beneficiaries, and may not be reproduced or copied without permission.

The information appearing in this document has been prepared in good faith and represents the views of the authors. Every effort has been made to ensure that all statements and information contained herein are accurate; however, the authors accept no statutory, contractual or other legal liability for any error or omission to the fullest extent that liability can be limited in law.

This document reflects only the view of its authors. Neither the authors nor the Research Executive Agency nor European Commission are responsible for any use that may be made of the information it contains. The use of

the content provided is at the sole risk of the user. The reader is encouraged to investigate whether professional advice is necessary in all situations.

No part of this document may be copied, reproduced, disclosed, or distributed by any means whatsoever, including electronic without the express permission of the PERIVALLON project partners. The same applies for translation, adaptation or transformation, arrangement or reproduction by any method or procedure whatsoever.

Copyright message

© **PERIVALLON Consortium, 2023**. This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgment of previously published material and of the work of others has been made through appropriate citation, quotation or both. Reproduction is authorised provided the source is acknowledged.

PERIVALLON

PERIVALLON aims to provide an improved and comprehensive intelligence picture of organised environmental crime and develop effective and efficient tools and solutions for detecting and preventing such types of criminal activities and for assessing their environmental impact based on geospatial intelligence, remote sensing, scanning, online monitoring, analysis, correlation, risk assessment, and predictive analytics technologies, by leveraging the latest advancements in Artificial Intelligence (AI) in the fields of computer vision and multimodal analytics. As a result, enhanced investigation processes and methodologies will be derived through the capabilities provided by the developed tools and solutions, and the insights obtained through the proposed Environmental Crime Observatory.

The capacity of end users (including Police Authorities and Border Guards) will also be improved and will enable them to tackle such criminal activities in an effective manner based on advanced tools and solutions and also on the innovative training curricula developed using physical and/or digital twins of relevant environmental crime scenarios. Moreover, improved international cooperation will be facilitated through improved data sharing enabled by blockchain technologies, while improved regulation shaping and tuning will be supported through relevant policy recommendations.

PERIVALLON will be validated in field tests and demonstrations in four operational use cases. Extensive training, hands-on experience, joint exercises, and training material will boost the uptake of PERIVALLON tools and technologies. With a Consortium 5 Police and Border Guard Authorities, 3 authorities related to environmental protection, 6 Research/Academic institutions, 8 industry partners (including seven SMEs), one EU Agency, and one Foundation, PERIVALLON delivers a strong representation of the challenges, requirements and tools to meet its objectives.

Table of Contents

Executive summary	9
Acronyms.....	10
1. Introduction	12
1.1 Purpose of the document	12
1.2 Scope of the document.....	12
1.3 Structure of the document	13
1.4 Methodology of the document.....	14
1.5 Definitions and main legal implications	14
2. Handling of PERIVALLON data according to FAIR principles	16
2.1 Administrative data	22
2.2 Research data	27
2.3 Allocation of Resources for FAIR data	60
2.4 Data Security for FAIR data	61
3. Legal Aspects - Data Protection.....	68
3.1 General aspects	68
3.2 Legal responsibilities.....	71
3.3 Data sharing within the consortium.....	72
3.4 Data Processing Agreements	73
3.5 International data transfers	74
3.6 Data Protection Impact Assessment	74
4. Ethical Aspects.....	75
4.1 General.....	75
4.2 Ethical responsibilities.....	76
5. Open Data Initiatives	77
6. Conclusions	79
6.1 Summary.....	79
6.2 Future work and recommendations	79
6.3 Timetable for DMP Updates	80
Annexes	81
Annex 1 - DMP Questionnaire	81

Annex 2 – Data flows 88

Annex 3 – Informed consent form template 91

Annex 4 – Information sheet template 92

References..... 96

Index of Figures

Figure 1 - PERIVALLON Data sets 16

Figure 2 - PERIVALLON Research Data Types..... 17

Figure 3 - The Fair Principles 19

Figure 4 - Data Management Plan Questionnaire Part 1 85

Figure 5 - Data Management Plan Questionnaire Part 2 87

Figure 6 - Overview of tasks and their respective contributing Research data types..... 89

Figure 7 - PERIVALLON Research data types 90

Index of Tables

Table 1 - Partners’ data management policy..... 19

Table 2 - Responsible person in charge of Administrative data per partner 23

Table 3 - Data Summary Administrative data 25

Table 4 - FAIR principles Administrative data 27

Table 5 - Responsible person in charge of Research data per partner 29

Table 6 - DT1 Data Summary 33

Table 7 - DT2 Data Summary 37

Table 8 - DT3 Data Summary 39

Table 9 - DT4 Data Summary 40

Table 10 - DT5 Data Summary 41

Table 11 - DT1 FAIR Approach 46

Table 12 - DT2 FAIR Approach 52

Table 13 - DT3 FAIR Approach 55

Table 14 - DT4 FAIR Approach 57

Table 15 - DT5 FAIR Approach 59

Table 16 - Allocation of Resources per partner 61

Table 17 - Data Security Measures per partner..... 67

Table 18 - Data Protection Contact Person per partner 71

Executive summary

PERIVALLON will ensure that all data sets that will be created and processed will be handled in relation to guidelines for FAIR data management, and personal data processing will be conducted in accordance with Regulation (EU) 2016/679 ('GDPR'). Within this aim, this Initial Data Management Plan ('DMP') guides all aspects of data management by

- outlining the types of data that will be collected and generated during the course of the project, including data from the pilots related to tackling environmental crime produced by the technical Work Packages (WPs);
- describing the methodology and standards required, and also identifying whether and how data will be collected, shared, exploited, re-used, or made accessible for verification, and how they will be curated and preserved;
- specifying the degree of privacy and confidentiality of the data that will be collected and generated;
- outlining the measures foreseen for adequately managing data from the legal, ethical, and security viewpoints;
- outlining the main elements of the data management policy that will be followed by the PERIVALLON Consortium to handle collected and generated data with respect to their sensitiveness during and after the project; and
- outlining the guidelines to be followed by the PERIVALLON Consortium with respect to Open Data initiatives.

For this purpose, DMP includes an overview of the data collected and processed during the project and focusing on the Research and Administrative data sets, it describes how such data sets are planned to be actionable following the FAIR approach.

Created under the scope of T1.4 ('Data and IPR management' (M1-M36)), this DMP addresses data that are/will be processed during PERIVALLON project only.

Finally yet importantly, this DMP is an initial iteration and a living document that will be updated through the lifetime of the project.

Acronyms

Acronym	Full name
ARPA	Agenzia regionale per la protezione dell'ambiente (arpa) della Lombardia
BayHFOD	Hochschule Fur Den Offentlichen Dienst In Bayern
CA	PERIVALLON Consortium Agreement. Version 1.1 – 22.08.2022
CENTRIC	Sheffield Hallam University
CERTH	Ethniko Kentro Erevnas Kai Technologikis Anaptyxis
D	Deliverable
DMP	Data Management Plan
DoA	Description of Action
DOI	Digital Object Identifier
DPO	Data Protection Officer
DRAXIS	DRAXIS ENVIRONMENTAL SA
DYLOG	DYLOG HITECH S.R.L.
DWG	Vlaamse maatschappij voor watervoorziening
ETRA	ETRA INVESTIGACION Y DESARROLLO SA
EU	European Union
FAIR	Findable, Accessible, Interoperable and Reusable
GA	Grant Agreement number: 101073952
GDPR	Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data
HP	Hellenic Police
IGP	INSPECTORATUL GENERAL AL POLITIEI
IPR	Intellectual Property Rights
KEMEA	Kentro Meleton Asfaleias

KR	PERIVALLON Key Results
LEAs	Law Enforcement Authorities
MT	Marintrafik opereisons anonymsi etaireia pliroforikis
MS	Milestone
M	Month
N/A	No applicable
PMB	Project Management Board
POLIMI	Politecnico di Milano
PUC	Pilot use case
RBP	Inspectoratul general al politiei de frontiera
SAFE	Fondazione safe
SATCEN	European Union Satellite Centre
SMOB	Set mobile srl
SPA	Polismyndigheten swedish police authority
T	Task
TMR	Tamar Israeli advanced quarrying CO Ltd
WP	Work package

1. Introduction

1.1 Purpose of the document

This DMP aims to outline the data collected and processed during the lifespan of the PERIVALLON project in order to secure a ‘good data management’, which encompasses various benefits. Namely, it gains visibility, secures that research data are stored in a sustainable and well-documented way and facilitates that data is rendered ‘FAIR’¹, among others. It addresses legal and ethical aspects on the handling of personal and non-personal data as well as details Open Data initiatives, to which also PERIVALLON data could contribute.

In this regard, this initial DMP pursues to:

- outline the types of data that will be collected and generated during the course of the project (see [Subsection 1.2 - ‘Scope of the document’](#)) regarding the analysed data sets in this initial DMP);
- address per data set:
 - the methodology and standards required;
 - whether and how data will be collected, shared, exploited, re-used, or made accessible for verification, and how they will be curated and preserved;
 - the degree of privacy and confidentiality of the data that will be collected and processed;
- outline the measures foreseen for adequately managing data from the legal, ethical, and security viewpoints;
- outline the main elements of the data management policy that will be followed by the PERIVALLON Consortium to handle collected and generated data with respect to their sensitiveness during the project; and
- outline the guidelines to be followed by the PERIVALLON Consortium with respect to Open Data initiatives.

1.2 Scope of the document

Following the Horizon 2020 Guidelines on Data Management², a DMP should include information on what kind of data will be collected and generated within and after the end of the project, the purpose of that processing and its relation with the objectives of the project as well as its origins and the types of data. Additionally, a DMP is required to set out the standards and methodologies that are applied during the data processing. Further, it should set out how data is shared among the project partners and whether it is shared outside the project/ consortium.

The PERIVALLON DMP follows such guidelines taking into consideration that the scope of this DMP and its initial character covers data processed during the course of the project only.

Additionally, given the initial nature of this document, the consortium decided to present a detailed and complete methodology on the handling of two data types identified: Administrative data and Research data. The remaining data sets (Technical data, Platform data and PUC data) were identified, however, the collection of information and the results of such collection will be delivered in the following Update#2 as per in the ‘Timetable for DMP Updates’ (see [Subsection 6.3](#)) since decisions on these data sets need still

to be taken. This analysis will be reflected in the Deliverables to be submitted after this DMP submission deadline:

- D2.1 ‘Co-creation of use case scenarios, specification of user and security requirements’ – M9
- D3.1 ‘Initial version of AI-based geospatial intelligence, remote sensing and scanning tools’ – M16
- D4.1 ‘Initial version of online monitoring and analysis tools’ – M17
- D5.1 ‘First version of PERIVALLON’ platform, secure data management and decision support system’ – M18

Given the living aspect of a DMP, updates over the course of the project are foreseen in the ‘Timetable for DMP Updates’ (see [Subsection 6.3](#)), which will supplement and/or complement this initial DMP in accordance with the respective deliverables and will contribute to the collection and processing of data within PERIVALLON. According to the GA, an update of the DMP is foreseen in M36 planned within the scope of D1.3. The consortium may evaluate ad-hoc reviews of the DMP.

It needs to be noted that the data sets identified at this stage of the project may be altered, modified, or may not be used upon the decision of the partners/consortium at a later stage of the project.

1.3 Structure of the document

This document is organised as follows:

- ❖ Section 1 provides a description of the purpose, scope of the document and its structure and definitions and main legal implications.
- ❖ Section 2 provides an overview of the PERIVALLON data sets identified at this initial stage of the project, and addresses the handling of two data sets according to the FAIR principles per data set. A general analysis on the allocation of resources and data security is provided in this section, as well.
- ❖ Section 3 identifies the main categories of personal data and cases where personal data will be processed. It defines the data protection measures to be put in place within the consortium according mainly to the Regulation 679/2018 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and to the repealing Directive 95/46/EC (‘GDPR’).³
- ❖ Section 4 delineates the most relevant ethical risks from the data protection perspective, in particular regarding the generation, use, storage and sharing of data in the project, as well as a description of ethical responsibilities relating to these risks.
- ❖ Section 5 enumerates the relevant legal framework and other open data initiatives at the European level as well as the main guidelines arising thereof and analyses the implications of such guidelines for the project.
- ❖ Section 6 concludes with the main findings and formulates the steps taken and covered concerning the handling of PERIVALLON data sets.
- ❖ Annex 1 includes the PERIVALLON - Data management plan questionnaire - Part 1 on Research data set, circulated to the technical partners, including KEMEA and BAYHFOD and end users, and PERIVALLON - Data management plan questionnaire - Part 2 on Administrative data set, circulated to all partners.

- ❖ Annex 2 shows the Research data types flow within the tasks.
- ❖ Annex 3 provides an Informed consent form template.
- ❖ Annex 4 provides an Information sheet template.

1.4 Methodology of the document

Additionally, the methodology followed by the authors in order to draft this DMP is described below:

1) Collection of specific information on processed data sets for PERIVALLON purposes.

Contribution from the consortium was key to the drafting of this Deliverable, and therefore information was collected through

- a questionnaire circulated by UNIVIE (Annex 1 – DMP Questionnaire) divided in two parts using the online tool: LimeSurvey⁴. Part 1 was circulated to technical partners, including KEMEA and BAYHFOD and end users, about Research data set, and Part 2 to all consortium partners about Administrative data set.
- the aggregated inputs provided by partners either in written form or as an oral contribution during bilateral consultations with UNIVIE.

For the purposes of information collection and materialization in the tables under the subsections ‘Administrative data’ and ‘Research data’, the wording of partners’ replies was slightly modified in order to provide for a harmonized and easy-to-read approach. The meaning of the replies was not changed in any manner whatsoever. Lastly, the order in which partners appear in the tables follows an alphabetical order.

2) Consultation and confirmation of PERIVALLON data sets and Research data types with consortium.

PERIVALLON data sets have been consulted with the PMB members and the Research data types have been confirmed by the consortium. The DMP has undergone the formal review under the Quality Assurance Procedure set forth in D1.1.

3) Analysis of the objectives of the Grant Agreement and Consortium Agreement.

This Deliverable has been drafted in view of the objectives described in the PERIVALLON GA⁵ and PERIVALLON CA⁶.

4) Research on DMP for research purposes and Guidelines on FAIR Data Management in Horizon 2020.

A comparative online research was performed by collecting methodologies on DMP in research activities⁷, having as the main source the Guidelines on FAIR Data Management in Horizon 2020.

1.5 Definitions and main legal implications

The main aim of this subsection is to provide a common understanding to the project partners of the PERIVALLON consortium of certain terms which are used alongside the deliverable regarding data

management, ‘publicly available data’, ‘open data’. These terms are often misinterpreted, which leads to misconceptions on the legal terms applicable to them. The simplistic approach to look for a common understanding is chosen as methodology for this subsection, avoiding entering into a scientific discussion on each of the definitions of these terms. Therefore, the legal instruments regulating ‘Data’ at the European Union level are resorted to for the purpose of this subsection⁸.

- **Data:** the European legal data initiatives define data in different ways, which depends on the material scope of such instrument will enter into play.
 - Data Governance Regulation⁹ and Data Act (Proposal)¹⁰: Article 2(1) ‘Data’ means any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording;’
 - Open Data Directive: Article 2(6) ‘document’ means: (a) any content whatever its medium (paper or electronic form or as a sound, visual or audiovisual recording); or (b) any part of such content;
 - Regulation on a framework for the free flow of non-personal data in the European Union¹¹: Art. 3(1) ‘data’ means data other personal data as defined in point (1) of Article 4 of Regulation (EU) 2016/679.
- **Non-personal data:** As stated above, the Regulation on a framework for the free flow of non-personal data in the European Union¹² defines ‘data’ as ‘other than personal data as defined in point (1) of Article 4 of Regulation (EU) 2016/679’, and it needs to be clarified that ‘[t]he category of “non-personal data” is generally defined in a negative way, as including all data which are not “personal”’.¹³
- **Open data:** The so called ‘Open Data Directive’¹⁴ in its Recital 16 defines open data as ‘[...]data in an open format that can be freely used, re-used and shared by anyone for any purpose’.
- **Personal Data:** The well-known GDPR defines personal data in its Article 4(1) as ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;
- **Publicly available data:** there is no definition provided in the instruments analyzed herein of the term ‘publicly available data’. One way to understand it is as any information or data that is made generally available to the public and is available on a non-exclusive basis under standard terms and conditions¹⁵. It must be highlighted that even when data is accessible to the public, these data might be subject to terms and conditions on its use, and even fall under another category of data, as e.g., personal data. In the later example, the GDPR applies irrespective of if such data are or were publicly available¹⁶.

Depending on the category of data, data is protected on grounds of intellectual property rights of third parties or under commercial confidentiality. This protection is granted either by law or on contractual basis and therefore attention should be given to the terms of use when the partners using such data set do not own a data set.

2. Handling of PERIVALLON data according to FAIR principles

This section details the data sets being collected and processed or expected to be collected and/or processed during the lifespan of the PERIVALLON project.

An array of data will be collected and processed and the handling of PERIVALLON data reflects the project’s initial status. Therefore, this first iteration of the DMP addresses the analysis of two data sets, Research data and Administrative data and their FAIR handling. It should be noted that some aspects of data processing are yet to be discussed, agreed, and decided upon at a later stage of the project.

As of the time of writing this DMP, taking into consideration the currently available and known information, the PERIVALLON project is expected to

- collect and process the following data sets:
 - Administrative data
 - PUC data
 - Research data
 - Technical data (Sensor data)
- process the following data set:
 - Platform data

The PERIVALLON data sets are visually represented below:

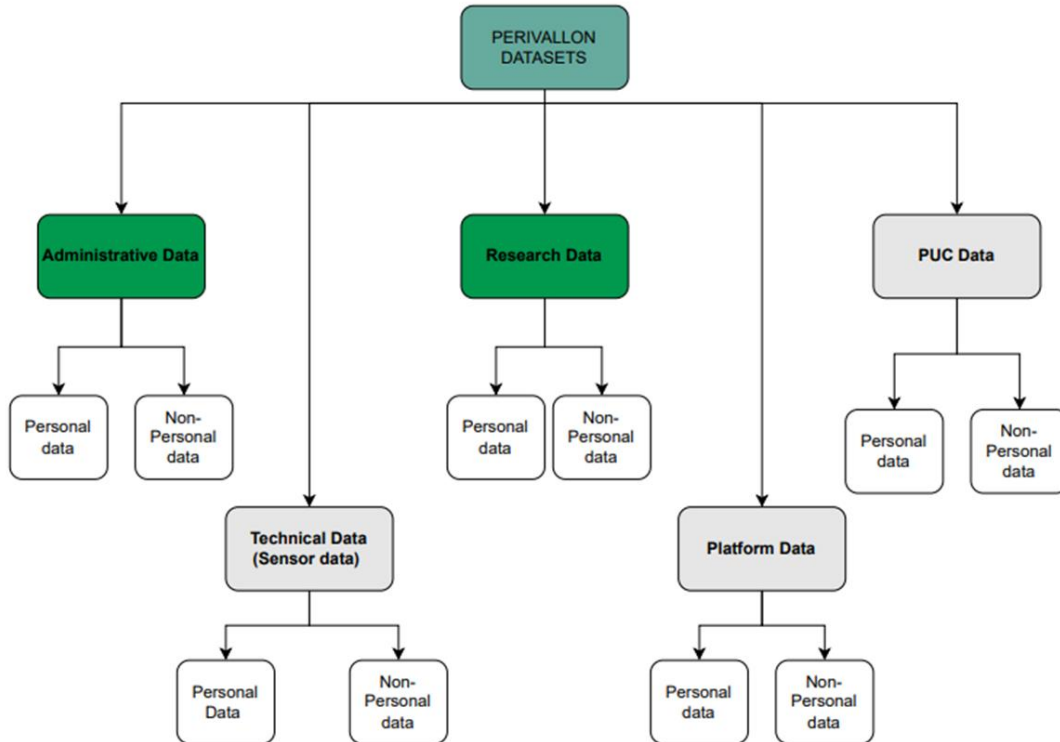


Figure 1 - PERIVALLON Data sets

Each data set depicts the type of data collected or produced under the umbrella of such data set. Its analysis, included in the [Subsection 2.1 ‘Administrative data’](#) and [Subsection 2.2 ‘Research data’](#), consists of

- the Data Summary, and
- the FAIR Approach.

PERIVALLON Research data set is subdivided in five different taxonomies, which are depicted below:

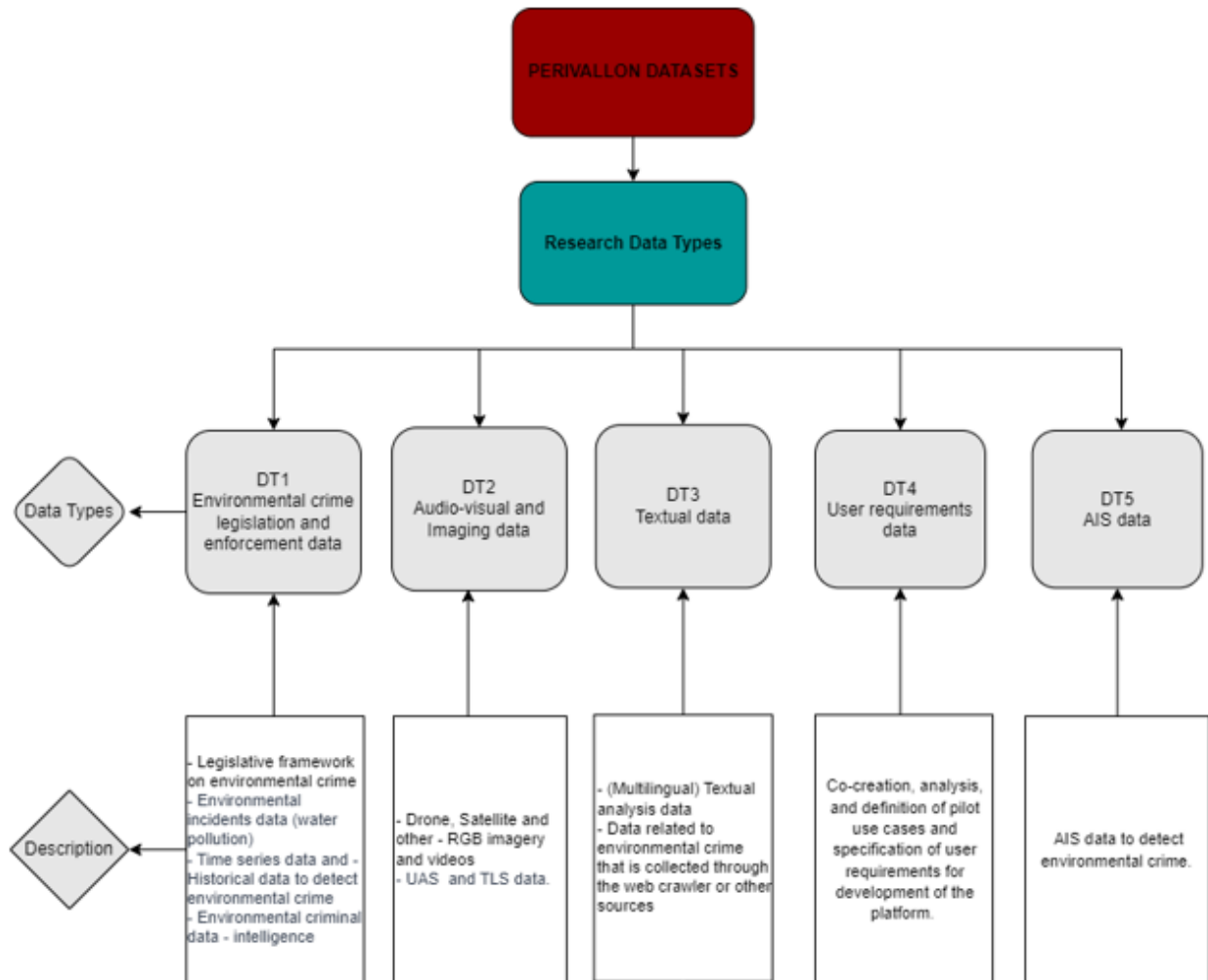


Figure 2 - PERIVALLON Research Data Types

DATA SUMMARY

The analysis of each data set addresses the horizontal approach to the FAIR principles that partners are expected to provide to the data sets they declared to process. A general summary to the data set and the general approach to FAIR data management of such data set is showed in a form of tables, which provides a clear visualization of the gathered information (See [‘Table 3 - Data Summary Administrative data’](#), [‘Table](#)

4 - FAIR principles Administrative data’, ‘Table 5 - Responsible person in charge of Research data per partner’, ‘Table 6 - DT1 Data Summary’, ‘Table 7 - DT2 Data Summary’, ‘Table 8 - DT3 Data Summary’, ‘Table 9 - DT4 Data Summary’, ‘Table 10 - DT5 Data Summary’, ‘Table 11 - DT1 FAIR Approach’, ‘Table 12 - DT2 FAIR Approach’, ‘Table 13 - DT3 FAIR Approach’, ‘Table 14 - DT4 FAIR Approach’, ‘Table 15 - DT5 FAIR Approach’, ‘Table 16 - Allocation of Resources per partner’, ‘Table 17 - Data Security Measures per partner’, ‘Table 18 - Data Protection Contact Person per partner’). Additionally, there is ongoing work on a repository of pre-existing and publicly available data, as declared by partners, which will provide an overview and necessary description of such data.

The allocation of resources for turning data FAIR is dealt with in a separate section (See [Subsection 2.3 ‘Allocation of Resources for FAIR Data’](#)) as it corresponds to a general appeal to the PERIVALLON project’s budget. Data security is also handled in a separate section (See [Subsection 2.4 ‘Data Security for FAIR Data’](#)) given the individual measures that each partner has and the common approach that PERIVALLON should look upon.

It is important to note that the information contained in the description of data sets analyzed in this DMP corresponds to the stage of the project at M6; some project tasks have not started yet and open issues (e.g. naming convention, interoperability of data) need to be agreed upon. Changes and updates to the data sets and their management are expected and will be reflected in the next informal update of the DMP planned for M15 (see [Subsection 6.3 ‘Timetable for DMP Updates’](#)).

Additionally, to the consortium approach established in this document, the following partners confirmed to follow their own data management guidelines:

Partners’ data management policy		
	Partner	Confirmation (link to it)
1.	ARPA	Yes.
2.	MT	Yes, Compliance Officer.
3.	CENTRIC	Yes, https://www.shu.ac.uk/research/excellence/ethics-and-integrity/data-management .
4.	CERTH	Yes, CERTH has clear internal security policy as defined and reviewed by the legal department and CERTH’s DPO.
5.	DWG	Yes, Data protection officer/ Legal department/ Security department.
6.	ETRA	Yes, Data Protection Officer and Legal Department.
7.	HP	Yes, European Union regulation of Data protection, incorporated in our legislation. All Hellenic Police personnel are required to comply with requirements of their Information Systems Security Policy of the Hellenic Police and to contribute to strengthening security practices, realizing the obligations they have in context of the existing legislation for security of both personal data and information systems. Guidelines for the protection of personal data: https://www.astynomia.gr/odigos-tou-politi/prostasia-dedomenon-prosopikou-%20charaktira/ (in Greek)

8.	IGP	Yes, Legal department.
9.	KEMEA	Yes, Data protection officer, Legal department and Security department.
10.	MT	Yes / Compliance Officer.
11.	POLIMI	Yes, DPO as required by GDPR.
12.	RBP	Yes, According to article 37 of Regulation (EU) no. 679/2016 , a personal data protection officer was appointed at the General Inspectorate of the Romanian Border Police level.
13.	SAFE	Yes, Data Protection Officer.
14.	SATCEN	Yes, EU legislation on data protection.
15.	SMOB	SMOB contract legal services from law companies project based. It has application-oriented security policy. The applications dedicated to end-users are accompanied by <i>Terms and conditions</i> and <i>Confidentiality policy</i> documents, explaining how the user data are gathered.
16.	SPA	Yes, internal policies.

Table 1 - Partners' data management policy

FAIR APPROACH

The 'Guidelines on FAIR Data Management in Horizon 2020'¹⁷ foresee four principles that govern the management of data in order to make them more easily understood, exchanged and open for re-use, and which is materialized by being findable, accessible, interoperable and reusable. In this regard, the 'PERIVALLON will ensure that all data sets that will be created and processed will be handled in relation to guidelines for FAIR data management'.¹⁸

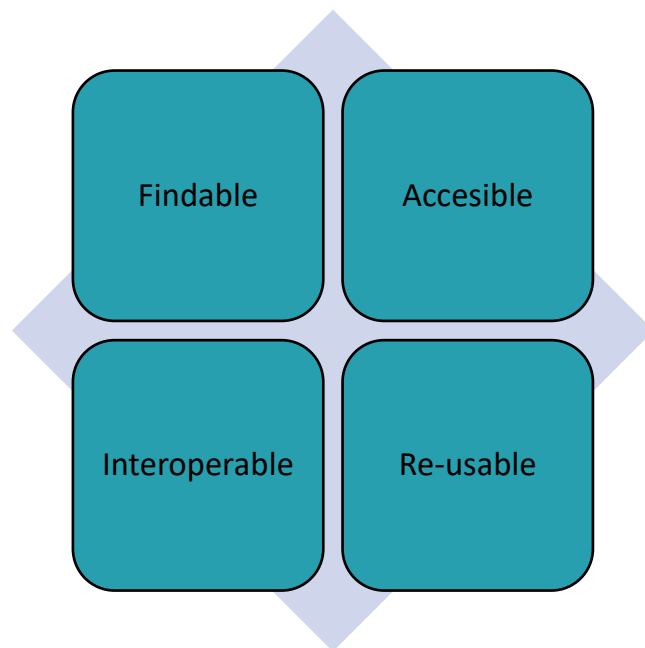


Figure 3 - The Fair Principles¹⁹

The PERIVALLON FAIR Approach is outlined in the GA, where the project partners decided the following:

Findable: Public data sets will be available through existing open APIs (e.g. Zenodo API, <https://developers.zenodo.org/>) and will be published under a Digital Object Identifier (DOI) thus making them easily findable. Data will be fully described using metadata, e.g. data description, creators, encoding formats, etc.

Accessible: Public data will be stored in trusted data repositories (e.g. Zenodo) for an indefinite time, allowing the scientific, and not only, community to improve PERIVALLON. Part of the data generated from PERIVALLON pilots may be confidential and protected against unauthorised access since they may relate to potential sensitivity issues. Anonymization techniques would be considered to allow the public access to this data.

Interoperable: Open data formats (such as CSV and JSON) will be used for meta(data). Meta(data) will make use of existing vocabulary (e.g. Schema.org) as internal representations of metadata. Finally, (meta)data could be exported to popular formats including referenced external pieces of URI-qualified metadata.

Reusable: License terms (e.g. GPL) will be included making the meta(data) subject to the specified terms by the uploader. Authors and meta(data) versions will be included.

Additionally, the project's FAIR approach is enriched by the results of the analysis of partners' replies.

Findable

For the PERIVALLON project, original data collection takes place at partners' site, stored at originating partner local storage and shared with partners, which are dependent on that data according to their task. The storage should be organised in a way that the data provenance of the data can be traced back.

Each set of data produced (data set, deliverables, etc.) will be named in a uniform way and will include a table with a version control.

For documents: The approach to naming documents of the project is detailed in D1.1 - *Project Management and Quality Assurance Handbook* and is summarized as follows:

The deliverables are categorised according to the following types:

- R: Document, Report
- DMP: Data Management Plan
- DEM: Demonstrator, pilot, prototype
- ETHICS: Ethics requirement
- OTHER

With respect to the dissemination level of deliverables and other documents, including presentations, the following levels of dissemination are considered in PERIVALLON:

- PU: Public, fully open, e.g., web (Deliverables flagged as public will be automatically published in CORDIS project's page)
- SEN: Sensitive (limited under the conditions of the Grant Agreement)

The documents will be **named and numbered** according to the following rules, in order to facilitate the quick identification and indexing:

PERIVALLON_<dtype>_<dnum>_<sec>_v<ver>.pdf

All the documents' names start with the word "PERIVALLON" in order to facilitate the identification, and to raise awareness about the project within a number of people that will download the documents from the public website.

With regard to project-related contractual documents, copies are kept on the Alfresco repository.

For data sets: Where possible, data sets should be named as follow

PERIVALLON WP [Work Package number] T [Task number] - [description of the activity] i.e.:
PERIVALLON WP1 T1.2 Results of DMP.

Partners maintain their collected data and metadata about the collected data in different versions with backup versions. No periodic strategy is defined at partners' side at the time of writing this DMP. The PERIVALLON procedure of project data sets in Zenodo are detailed in Deliverable D7.1.

The project repository provided through Alfresco serves as a collective working space where the partners can exchange documents and collaborate in the development of deliverables and data sets.

Metadata standards will be used for Research data types such as e.g. Github, Mendeley, Zenodo. Potential users can access the data sets through the Alfresco repository.

Accessible

For the PERIVALLON project,

Administrative data, which are marked as public, will be published on the project website. Otherwise, it will be accessible for the project partners via the Alfresco repository.

For **some Research data types**, data will be available on the PERIVALLON platform or, if agreed among the consortium, in another platform like Zenodo. These data types are accessible via internet and log in details are needed in the case the chosen platform requires it. Regarding the methods for accessing the data, these vary from project's internal platform to publicly available tools, like AerialWaste toolkit (<https://github.com/nahitorres/aerialwaste>).

Information which is explicitly marked as 'confidential' is 'Sensitive Information'. The PERIVALLON CA provides for a series of obligations on the handling of such data in its Art. 10.2 CA.

For research metadata, accessibility will be provided through Zenodo, Github for code and other public websites, like aerialwaste.org.

Interoperable

For the PERIVALLON project,

Administrative data will follow the naming convention as defined in D1.1, as well as methodologies for sharing such data. Common vocabularies have not been defined yet.

Research data will be interoperable to the extent that the licenses pertaining to data types allow so. Where depending on the data type, this is foreseen, will e.g. mark columns (CSV) and objects (JSON) with clearly descriptive names and standard formats (e.g. integer, strings, floats). Additionally, AIS approved recommendation standards: <https://www.itu.int/rec/R-REC-M.1371-5-201402-1/en> will be provided. Geospatial data sets that will be used extensively in the project will follow OGC standards <https://www.ogc.org/standards>.

Reusable

For the PERIVALLON project,

Administrative data will be available for reuse to the extent that are marked with the level ‘Public’.

Research data will be, in principle, available for reuse, in many cases upon agreement with the consortium. The license terms applicable to the different types of data vary and indicate the terms applicable to reuse of data. The audience for reuse is mainly Police Authorities, Border Guards and National/Regional Authorities and researchers/developers, which would have/ required access to the PERIVALLON data sets.

2.1 Administrative data

All data related to the management and administration of the project; namely, templates, deliverables, audio and visual recordings, presentations, financial reports, among others are hereby encompassed under ‘Administrative data’.

Personal data, like contact details of partners, are also considered to fall under Administrative data but for methodological reasons are analyzed under [Subsection 4 ‘Legal Aspects – Data Protection’](#).

These data are planned to be collected in the performance of T1.1, T1.2, T1.4, T3.1- T3.4, T4.1-T4.5, T5.2, T5.4, T7.1, T7.2, T7.3, T7.5.

This data set contains public data. Public data are collected in the PERIVALLON sharing platform, Alfresco, hosted by ETRA, as leader of the WP1 on Project Management. The platform is used to exchange, collect and store documents relevant to the project, as detailed in Section 4.1 of D1.1 ‘Project management and Quality Assurance Handbook’.

Responsible person in charge of Administrative data per partner*			
	Partner	Full Name	Contact details
1.	CENTRIC	Jordan Thompson	jordan.thompson@shu.ac.uk
2.	CERTH	Theodora Tsikrika Thanasis Kapoutsis Kostas Gkountakos	theodora.tsikrika@iti.gr athakapo@iti.gr gountakos@iti.gr
3	DRAXIS	Anastasios Karakostas Dafni Deligioglani	akarakos@draxis.gr ddelioglani@draxis.gr

4	DWG	Han Vervaeren	han.vervaeren@dewatergroep.be
5	DYLOG	Francesco Trincherò Luca Di Nuovo	trincherò@dylog.it dinuovo@dylog.it
6	ETRA	Data Protection Officer	dpo@grupoetra.com
7	IGP	Radu Bors	radu.bors@igp.gov.md
8	KEMEA	Skarlatos Efstathios	e.skarlatos@kemea-research.gr
9.	POLIMI	Piero Fraternali	piero.fraternali@polimi.it verdiana.bredice@polimi.it
10.	RBP	Ovidiu-Mircea Manolache Sergiu-Razvan Malita (data protection officer)	ovidiu.manolache@igpf.ro sergiu.malita@igpf.ro
11.	SAFE	Alessandro Castagnetti (project management related data), Davide Gallo (financial data), Silvia D’Adda (DPO)	Alessandro.castagnetti@safe-europe.eu Davide.gallo@safe-europe.eu Silvia@safe-europe.eu
12.	SATCEN	Patricia Romeyro	Patricia.Romeyro@satcen.europa.eu
13.	SAHER	David Fortune Andrew Staniforth	dave@saher-eu.com
14.	SPA	Åsa Granström	asa.granstrom@polisen.se
15.	TMR	Nir Haimov Lea Haim	nir@tamar-group.com lea@tamar-group.com

Table 2 - Responsible person in charge of Administrative data per partner

*This table shows the responsible persons designated by the partners in charge of Administrative data. In any case, partners are reachable through their assigned contact person(s)' contact details available in the project's contact list.

The data summary of Administrative data is presented below based on the methodology presented in [Subsection 1.3 \(Structure of the document\)](#):

DATA SUMMARY ADMINISTRATIVE DATA	
Type of data collected/generated and processed	<ul style="list-style-type: none"> • project planning, meeting presentations, • templates, • email addresses, • names, • minutes,

	<ul style="list-style-type: none"> • interviews, • questionnaires, • surveys, • deliverables, • reports, • WPs and tasks presentations, record documentation of communications among members of the projects.
Use of pre-existing data	<p>No for partners: CENTRIC, ETRA, DRAXIS, DYLOG, DWG, KEMEA, IGP, POLIMI, RBP SACTEN, SAHER, SPA, TMR.</p> <p>Yes for partner: CERTH</p>
Origin of data	Books, journals, research publications, provided by project partners, online sources provided by PERIVALLON end users.
Purpose of the data collection	<ul style="list-style-type: none"> • organisation, monitoring, coordination and management of tasks. • obtain input from partners regarding the needs/ requirements of tasks of interest. • creation of the scenarios and the requirements that came up after discussion with the end users of the project, will give feedback to the technical partners. • administration and report related to demonstration event in Sweden. Fulfilling obligations as partners in the project (SPA). • make an analysis of the requirements for platform's development, to monitor and report the evolution of the project (IGP). • visualization. • record of all the activities of the people involved, • keep track of the progress and the resources used.
Way of data collection	Collected from all partners / Outlook/Teams.
Formats	Storage: Locally in a dedicated server or stored in Alfresco.
Size	Size: GB, minimal.
	Format: MsOffice files, PDF, MP4, emails .eml/.html.
Data traceability	Via Alfresco site, and using the established naming conventions in deliverable D1.1.
Use of procedures for data management	N/A
Tools or software needed to	Office suite (i.e., MS Office), PFD reader (i.e., Adobe Reader), Email client (i.e., Microsoft Outlook, Mozilla Thunderbird).

create/process/visualize the data	
IPR Protection	No.
Storage and backup strategy for the collected data	<ul style="list-style-type: none"> Mainly data will be stored on the Alfresco site, but also downloaded and stored locally. Local servers perform daily backups, e.g. backup strategy of De Watergroep (data is stored in Microsoft Teams), DRAXIS servers. <p><u>Particular to partners:</u></p> <ul style="list-style-type: none"> Data will be stored locally in a secure server located in the premises of CERTH and remotely on the project’s repository (CERTH). Data is stored in the computer of project's responsible people. In case if the file needs to be shared and the size is too big data is stored remotely, most frequently is used Google Drive (IGP). DropBox (SAHER). Any data locally stored at the SatCen is automatically backup (SATCEN). stored in POLIMI email system and Microsoft OneNote repository. Drafts of planning docs temporarily stored in Google doc and Google sheet documents, consolidated in the Alfresco project repository (POLIMI).
Tasks	T1.1, T1.2, T1.4, T3.1- T3.4, T4.1-T4.5, T5.2, T5.4, T7.1-T7.3, T7.5.
Partner in charge	CENTRIC, CERTH, ETRA, DRAXIS, DYLOG, DWG, KEMEA, IGP, POLIMI, RBP, SACTEN, SAFE, SAHER, SPA, TMR, UNIVIE.

Table 3 - Data Summary Administrative data

FAIR PRINCIPLES ADMINISTRATIVE DATA	
Making data findable	
Data identifiers	N/A
Directory and file naming convention	Various directories are available in Alfresco for storing Administrative data sets. The naming conventions defined in D1.1 will be used.
Place to find data	Alfresco repository. Organisation stricted links.
Clear version numbers	Yes.
Search keywords	This is not foreseen/not defined yet.
Metadata creation	Metadata such as date of creation, last save, editing time, word count, etc... are created automatically by MS Office. Date, ownership.

Documentation and Metadata standards	Not defined yet. Built in metadata of teams.
Potential users	<ul style="list-style-type: none"> • By accessing the Alfresco site. • Dissemination, scientific papers and project’s web site.
Making data openly accessible	
Data identifiers	-N/A-
Directory and file naming convention	Various directories are available in Alfresco for storing Administrative data sets. The naming conventions defined in D1.1 will be used.
Place to find data	Alfresco repository. Organisation stricted links.
Clear version numbers	Yes. Versioning bulit in.
Search keywords	This is not foreseen/not defined yet.
Metadata creation	<ul style="list-style-type: none"> • Metadata such as date of creation, last save, editing time, word count, etc... are created automatically by MS Office. • Date, ownership.
Documenation and Metadata standards	Not defined yet.
Potential users	<ul style="list-style-type: none"> • By accessing the Alfresco site. • Dissemination, scientific papers and project’s web site.
Making Data interoperable	
Interoperable data	Naming convention was defined for the Administrative data sets in D1.1, as well as methodologies for sharing such data. Common vocabularies have not been defined yet. Partners will inform UNIVIE if certain data cannot be made interoperable. Records thereof will be maintained by UNIVIE.
Data and metadata vocabularies, standards or methodologies to facilitate interoperability	<ul style="list-style-type: none"> • Not defined yet.
Making Data re-usable	
Data sharing and re-use (licence)	<ul style="list-style-type: none"> • The following ones have been identified so far: - Deliverables with dissemination level "Public" - Presentations with dissemination level "Public". Where certain data cannot be made interoperable, this must be informed by the partners to UNIVIE, who will maintain record.

	Partners will inform UNIVIE if certain data cannot be shared or re-used. Records thereof will be maintained by UNIVIE.
Audience for reuse	<ul style="list-style-type: none"> • Any stakeholder interested in the project. • Consortium partners.
Period of re-usability	Not defined yet.
Restrictions on data re-use	Not defined yet.

Table 4 - FAIR principles Administrative data

2.2 Research data

Research data is herein referred to data collected and processed or expected to be collected and processed for the development phase of the components, services, and tools under the PERIVALLON platform. It excludes the further processing of data for the integration as a platform (Technical and Platform data sets) and the processing of data for demonstration of Pilot Use Case scenarios (PUC data sets).

These data will be collected in the performance of different tasks according to the data type collected and processed.

Personal data processed for the purposes of Research data will not be analyzed herein, but addressed under [Subsection 4 ‘Legal Aspects – Data Protection’](#) of this DMP.

For a detailed understanding of the description and FAIR approach to Research data, partners provided the following statements in their reply to the DMP Questionnaire Part 1.

- ARPA: ARPA could participate in the process of collecting and analyzing data, particularly, for DT2 (thinking at satellite data, and data/surveys from drone), and also in DT1 (georeferenced list of already mapped/known critical site for waste management, illegal landfills, etc..). This data will be exploited, as expected, particularly in the workflows in WP3 and WP6.
- CERTH: CERTH should/could use audiovisual and online data for analysis. Specifically, satellite data will be used for research development work in T3.1, aiming to detect water pollutants. In addition, drone imagery data will be used in tasks T3.2, T3.3, and T3.4 to address project objectives such as detecting water pollutant events, 3D terrain mapping, swarm intelligence and optimization for mission planning and surveillance. Moreover, visual data will be used in the context of T4.4 for detecting objects related to HFCs and refrigerants. Finally, online data will be analyzed for text analysis and concept extraction in the scope of T4.3. In all cases mentioned above, visual data in the form of still images or videos can be used. However, the audio will not be processed and immediately removed -if any- from the videos.
- ETRA: Since most of our research development work (T5.1, T5.3, T5.5) focus on integrating and processing various data sources produced by other partners in the project (mainly from WP3 and WP4), and while at this moment the different data sources to be exploited and their integration are still under discussion, our answer is based on the Description of Action. A definitive answer is expected to be provided by M12, when the platform architecture definition is finalised.

Considering the above, [ETRA] expect[s] to collect and process research data involving’ Data types 1 to 3 and Data type 5.

As ETRA is not the main collector of the data types, the partner leaves to the main collectors the possibility to address the details of the Data summary and FAIR approach. Where the collection of ETRA materializes, ETRA plans to FAIR approach as follows:

- On tools or software needed to visualize data: For T5.1, ETRA expects to use the following technologies: NodeJS, Meteor, React, semanticUI, JSON, mongoDB, aldehyd:tabular, NATS, MQTT/AMQP, DDP. For T5.3, ETRA expects to use the following technologies: NodeJS, Meteor, React, semanticUI, JSON, mongoDB, aldehyd:tabular, NATS, MQTT/AMQP, DDP. For T5.5, ETRA expects to use the following technologies: BigchainDB, IPFS.
- On metadata creation: ETRA expects at least to create/capture timestamps. MIME types are planned to be used.
- DRAXIS: DRAXIS is responsible for the development of PERIVALLON's "Waste crime data monitoring tool" (KER8). This repository will contain any data partners send to DRAXIS.
- DWG: DWG collects data from environmental incidents that might impact drinking water production quality. These data are collected. Different kinds of information is collected in this database.
- RBP: RBP participates as an end-user partner in EURMARS project and as a public body, and do not engage in research data activities.
- SAFE: will gather and elaborate data relative to the assessment of the PUCs and tests as well as during the monitoring of the related activities.
- SATCEN: SATCEN will not collect research data, technical partners will be the ones collecting and processing research data. SATCEN will provide expertise in relation to the type of data to be collected but will not be directly involved in that data collection.
- SMOB: SMOB will only transfer data between Platform components via integration tools. Within this process, SMOB transfers output from one or more components to be input for one or more components. The specific integration data will be defined during the System Architecture design. SMOB will use testing data of each type that is produced by some Platform components and that are needed by other Platform components. No FAIR analysis will be accomplished, since no data type is produced by the integration process. The collection of data within PERIVALLON project is accomplished by software components, which will be then collected under the Technical dataset.

Responsible person in charge of Research data per partner*			
	Partner	Full Name	Contact details
1.	ARPA	To be discussed	
2.	BAYHFOD	Stefan Heisenhuber	stefan.heisenhuber@pol.hfoed.bayern.de
3.	CENTRIC	Helen Williamson Helen Gibson Jordan Thompson	dpo@shu.ac.uk (DPO) h.gibson@shu.ac.uk jordan.thompson@shu.ac.uk

4.	CERTH	Theodora Tsikrika Thanasis Kapoutsis Kostas Gkountakos	theodora.tsikrika@iti.gr athakapo@iti.gr gountakos@iti.gr
5.	DRAXIS	Anastasios Karakostas Dafni Deliolgani	akarakos@draxis.gr ddeliolgani@draxis.gr
6.	DWG	Han Vervaeren	han.vervaeren@dewatergroep.be
7.	ETRA	Data Protection Officer	dpo@grupoetra.com
8.	HP	Ioannis Petropoulos	i.petropoulos@astynomia.gr
9.	IGP	Radu Borş Sergiu Sîrbu Nicoleta Coşcodan	radu.bors@igp.gov.md sergiu.victor.sirbu@igp.gov.md nicoleta.coscodan@igp.gov.md
10.	KEMEA	Skarlatos Efstathios Vagia Pelekanou Zeta Aggariti	e.skarlatos@kemea-research.gr v.pelekanou@kemea-research.gr z.aggariti@kemea-research.gr
11.	MT	Vasiliki Efstathiou	vasiliki.efstathiou@marinetraffic.com
12.	POLIMI	Piero Fraternali	piro.fraternali@polimi.it giacomo.boracchi@polimi.it
13.	RBP	Malita Sergiu-Razvan	sergiu.malita@igpf.ro
14.	SAFE	Alessandro Castagnetti	Alessandro.castagnetti@safe-europe.eu
15.	SATCEN	Patricia Romeyro	Patricia.Romeyro@satcen.europa.eu
16.	SMOB	Sever Calit	sever.calit@setmobile.ro
17.	SPA	Christoffer Hagberg	christoffer.hagberg@polisen.se

Table 5 - Responsible person in charge of Research data per partner

*This table shows the responsible persons designated by the partners in charge of Research data. In any case, partners are reachable through their assigned contact person(s)' contact details available in the project's contact list.

SUMMARY OF RESEARCH DATA TYPES

Research data set is subdivided in five different taxonomies according to the information provided by partners to the DMP questionnaire PART 1.

- DT 1. Environmental crime legislation and enforcement data.
- DT 2. Audio-visual and Imaging data.
- DT 3. Textual data.
- DT 4. User requirements data.
- DT 5. AIS data.

This section provides a data summary per data type as well as its FAIR principles.

DATA SUMMARY PER RESEARCH DATA TYPE

Data Type 1: Environmental crime legislation and enforcement data				
Contributing partners: CERTH, DWG, HP, IGP, UNIVIE				
Data Summary				
Partner in charge	CERTH	DWG	HP	IGP
Type of data collected and processed	<ul style="list-style-type: none"> Time series data and associated metadata for the detection of trends, patterns and prediction of potential threats. Historical data, including information about past environmental crimes, statistics, socio-economic factors, etc. <p>CERTH plans to use openly available historical data. However, it is not certain at this stage of the project if such data exists or if the Police Authorities, Border Guards and National/Regional Authorities will be able to provide such data - upon to their decision.</p>	Data from environmental incidents that might impact the drinking water production quality. Different kind of information is collected in this database.	<ul style="list-style-type: none"> Criminal data – intelligence (probably) statistics of criminal records, related only to Greece, collected from HP annual activity reports. 	Legislative framework, type and effectiveness of enforcement action and legal data
Use of pre-existing data	Yes.		No.	Yes.
Origin of data	<ul style="list-style-type: none"> Publicly available data for model development/evaluation EUROSTAT databases 	Data collected over the years by ourselves and the VMM.	N/A	Open source online data sources, online marketplaces, journals,

	<ul style="list-style-type: none"> Publicly accessible relevant databases with recorded environmental crimes. 			provided by project partners.
Purpose of the data collection	Development of methods for the detection of trends, patterns and prediction of potential criminal activities.	Delivery of data that is useful for PUC2. Risk assessment regarding impact on the drinking water production.	Evaluation.	For the creation of needs for the PERIVALLON platform.
Way of data collection Formats Storage Size	<ul style="list-style-type: none"> Already existing data sets publicly available for research purposes. Provided by Police Authorities, Border Guards and National/Regional Authorities partners (if possible). <p>Storage: CERTH secured servers.</p> <p>Size: not known yet.</p> <p>Format:</p> <ul style="list-style-type: none"> Time series data and associated metadata; Criminal databases, including information about past activities, modus operandi and metadata (e.g. location, timestamp) 	<p>Format: Excel, shape files.</p>	<p>Collected by corresponding authorities.</p> <p>Storage: Platform/ALFRESCO</p> <p>Format: MP4, MPEG, MsOffice files.</p>	<p>Collected by Police Authorities, Border Guards and National/Regional Authorities.</p> <p>Storage: Locally or remotely stored.</p> <p>Size: not known yet.</p> <p>Format: MsOffice files (Excel, Word, PowerPoint) PDF files .jpg/PNG.</p>
Data traceability	Data will be marked with unique identifiers and meaningful names.	Ourselves and received from VMM.	Using suitable tools.	Not known yet.
Use of data management procedures	No.			

Tools or software needed for data creation/ processing/ visualization	Python-based libraries for data processing: will be specifically determined during the course of the project.	Microsoft Office (for Excel), GIS-viewer (for Shape files).	Unknown.	No.
IPR Protection	No.			
Annotation	Not known yet.	No.	Unknown.	No.
Storage and backup strategy for the collected data	Data will be stored at CERTH secured servers. Incremental backup strategy will be followed using a separate hard drive disk.	Yes.	No.	Not yet.
Tasks	T1.4, T2.1, T2.2, T3.1, T4.1, T5.2, T5.3, T5.4, T6.1, T6.2, T6.3, T7.3			
KR tentatively benefitting from RD	All KRs available, in particular depending on the partner: KR1-6KR13, KR14, KR19, KR20, KR21.			

Table 6 - DT1 Data Summary

Data Type 2: Audio-visual and Imaging data
Contributing partners: CERTH, CENTRIC*, POLIMI, SPA

Data Summary

Partner in charge	CERTH	POLIMI	SPA
Type of data collected and processed	Drone, Satellite and other RGB imagery and videos.	Satellite images.	UAS and TLS data. RGB images from environmental crimes.
Use of pre-existing data	<ul style="list-style-type: none"> • Drone, Satellite data. Source: Copernicus database • RGB data. Source: open available databases. (i.e., https://cocodataset.org) • Existing academic data set. Data from past missions that could likely be used. 	Yes.	No.
Origin of data	<ul style="list-style-type: none"> • Test trials • Provided by partners, potentially the same ones who will provide the equipment • Online, publicly available / open source academic research data, e.g. Copernicus data. 	<ul style="list-style-type: none"> • Orthophotos generated by an airborne campaign in 2018, executed by AGEA. • WorldView-3 (WV3): high-resolution pan-sharpened RGB images acquired by a commercial satellite sensor (no pan-sharpened near infrared images were used), from the campaign in 2021. • Google Earth (GE): images downloaded using the Google API 3, 	From SPA

		free to the public and can be used for remote sensing studies.	
Purpose of the data collection	Analysis as well as testing, training and evaluation of models.	Analysis and territory monitoring.	Analysis and evaluation.
Way of data collection Formats Size	<ul style="list-style-type: none"> Collected by CERTH through the Copernicus Open Access Hub API/ other open-source repos. Size: GB Format: mp4, WebM, mpeg, png, jpg, json, xml, tif, laz <ul style="list-style-type: none"> Initial data: SAFE format, JP2 Produced data: GeoJSON, geoTIFF 	Collected by ARPA Lombardia, Format: PNG image.	Collected by SPA and the consortium partners Size: not known yet. Format: MP4, MOV, JPG, RAW
Data traceability	Uploaded to the Perivallon website or repository (Alfresco) with a meaningful name (e.g. Data set, Version_Month_Year).	Trace system of the ZENODO repository.	N/A
Use of data management procedures	No.		
Tools or software needed for data creation/ processing/ visualization	Data visualisation: WebODM, QGIS (outside of platform - no Web application), python or web-based languages Data processing: Python pandas, pillow and OpenCV libraries, eodag, scikit-lear, keras/Tensorflow, eo-learn, scikit-image, geopandas, joblib - other software: GNU Parallel, ESA Sentinel	Python code is provided to visualize and process the data https://github.com/nahitorres/demo-inspection-tool	Initially, Agisoft Metashape, Cloud Compare, Leica software (connected to the use of TLS). Might be subject to change in the process.

	Application Platform (SNAP) with the relevant Sentinel toolboxes (via GPT and snappy)		
IPR Protection	<p>Drone imagery: No.</p> <p>Satellite data:</p> <ul style="list-style-type: none"> Original data: open and free Produced data: licensed under CC BY-NC-SA (https://creativecommons.org/licenses/by-nc-sa/2.0/) <p>RGB data: Some yes: https://cocodata.set.org/#termsofuse, Creative Commons Attribution 4.0 License.</p>	<p>Yes, Creative Commons CC BY licensing scheme.</p> <p>The usage of Google Imagery must respect the Google Earth terms and conditions.</p>	No.
Annotation	<p>Drone imagery: no</p> <p>Satellite data:</p> <ul style="list-style-type: none"> Original data set: retrieved by Copernicus and it isn't annotated. Produced data set: expected to contain name of the labels that corresponds to a specific water pollutant, and the label for each pixel of the geoTiff/ geoJSON image. <p>RGB data: Some yes: The MSCOCO data set, which is publicly available and already annotated, contains more than 200,000 images that depict 91</p>	<ul style="list-style-type: none"> Binary labels (waste, no waste) Waste types (15): Rubble/excavated earth and rocks, Bulky items, Fire Wood, Scrap, Plastic, Vehicles, Tires, Domestic appliances, Paper, Asphalt milling, Corrugated sheets (presumed asbestos-cement) etc. Storage modes (7): Heaps not delimited, Container, Big bags, Pallets, Delimited heaps (by barriers/walls/etc), Cisterns etc. Most positive instances are characterized by metadata about the 	No.

	categories (i.e., refrigerator, chair, tv, car etc.) that occur in both indoor and outdoor scenes.	evidence, severity, and area type of the site.	
Storage and backup strategy for the collected data	Storage: Locally and possibly cloud storage. Some could be open-access (e.g. Github). Given that collected data (i.e., original data) are open, and can be retrieved again if it is deemed necessary, no specific backup strategy is used.	Storage: locally and in ZENODO repository. https://zenodo.org/record/7034382#.Y9fueHbMJnI Periodic backup.	Storage: Locally, on SPA controlled servers. Backup strategies.
Tasks	Drone data: T3.3. Possibly also T3.4, T3.2 Satellite data and Waste discovery in aerial images: T3.1, possibly also T5.1, T5.3 RGB data: T4.4 UAS and TLS data: T2.2, T2.3, T3.2, T3.3.		
KR tentatively benefitting from RD	Drone data: KR4, possibly also KR5, KR3 Satellite data: KR2 RGB data: KR10	KR1	KR3

Table 7 - DT2 Data Summary

*CENTRIC: The partner will collect visual data (images, videos) obtained from online sources.

Data Type 3: Textual data Contributing partners: CENTRIC, CERTH		
Data Summary		
Partner in charge	CENTRIC	CERTH
Type of data collected and processed	Text data related to environmental crime that is collected through the web crawler, types of open-source data integrated from datasets of T4.2 with any other platform data that will be stored in the T5.4 SDMA. This could include web page text data.	Textual analysis data with the objective determining whether there are indications (e.g., concepts of interest) of a potential threat.
Use of pre-existing data	No.	Yes.
Origin of data	N/A.	Publicly available data for model development/evaluation.
Purpose of the data collection	Monitoring of dark web sources for environmental crime indicators. Analyzes the collected data to determine its relevance and potential value in identifying environmental crimes.	Development of methods for supporting multilingual content analysis, such as machine translation, name entity recognition and concept extraction aiming at detecting indications of potential threat.
Way of data collection	N/A.	Already existing data sets publicly available for research purposes
Formats		Size: not known yet
Size		Format: Text-based data.
Data traceability	The URL of the data is captured during dark web crawling, All data stored in the secure data store will be audited by the auditing module.	Data will be marked with unique identifiers and meaningful names.
Use of data management procedures	No.	

Tools or software needed to create/process/visualize the data	The tools needed to create, process, and visualize data for the web crawler and link evaluator modules include Selenium for crawling, Java Spring for processing and MongoDB for storage. Selenium is useful for automating web browser interactions to extract data, Java Spring provides features for data processing and manipulation, and MongoDB is a NoSQL database for storing unstructured data.	Python-based libraries for data processing: will be specifically determined during the project.
IPR Protection	Yes. A crawled web data set is generally considered to be publicly available information and not subject to IPR protection. However, some websites may have terms of use or copyright restrictions that prohibit web crawling and data extraction without permission.	No.
Annotation	No.	Not known yet.
Storage and backup strategy for the collected data	No strategy is in place.	Stored at CERTH secured servers. An incremental backup strategy will be followed using a separate hard drive disk.
Tasks	T4.3, T4.4, T5.1, T5.3, T5.5 and T5.6	
KR tentatively benefitting from RD	KR7-8, KR11, KR16, KR17	

Table 8 - DT3 Data Summary

Data Type 4: User requirements data Contributing partner: KEMEA	
Data Summary	
Partner in charge	KEMEA
Type of data collected/generated and processed	Co-creation, analysis, and definition of pilot use cases and specification of user requirements
Use of pre-existing data	Yes.
Origin of data	Through interviews with the technological partners and the end users for each PUC, workshops, online research and maybe some statistic data.
Purpose of the data collection	Analysis and definition of PUCs and specification of the requirements.
Way of data collection, Formats, Size	Storage: Platform. Size: not known yet. Format: MS Office files, shape files
Data traceability	Some data will be traced after online research, while some other will be retrieved from technical partners and end users, through online meetings that take place regularly.
Use of procedures for data management	No.
Tools or software needed to create/process/visualize the data	This question refers to the technological partners of the project.
IPR Protection	No.
Annotation	We are working with it.
Storage and backup strategy for the collected data	This question will need to be dealt with technological partners of the project.
Tasks	T2.1-T2.4, T3.2-T3.4, T5.2, T5.3, T6.1-T6.4, T7.1, T7.3, T7.4.
KR tentatively benefitting from RD	KR1- KR6, KR8, KR17, KR19.

Table 9 - DT4 Data Summary

Data Type 5: AIS data Contributing partner: MT	
Data Summary	
Partner in charge	MT
Type of data collected/generated and processed	AIS data
Use of pre-existing data	Yes.
Origin of data	Proprietary data owned by MT
Purpose of the data collection	Analysis.
Way of data collection	Proprietary AIS data
Formats	Locally stored
Size	Format: .csv / in the form of MS SQL tables in databases
Data traceability	Size: unknown
Data traceability	Via SQL queries.
Use of data management procedures	No.
Tools or software needed to create/process/visualize the data	Python geopandas, folium QGIS H3 hexagonal layers
IPR Protection	Proprietary commercial data will be used for the analysis. Aggregated data in the form of results will be used for the purposes of the use case.
Annotation	No.
Storage and backup strategy for the collected data	Daily backups at company's premises.
Tasks	T4.5
KR tentatively benefitting from RD	KR10

Table 10 - DT5 Data Summary

FAIR PRINCIPLES DIVIDED PER DATA TYPE

Data Type 1: Environmental crime legislation and enforcement data				
Contributing partners: CERTH, DWG, HP, IGP				
Making Data Findable				
Partner in charge	CERTH	DWG	HP	IGP
Data identifiers	<ul style="list-style-type: none"> • Already publicly available data set. • For data that may be provided by end users (historical data), unique identifiers will be used. 	No.	Unknown/ Not known yet.	
Place to find data	Not available yet.	The data set that we will deliver (data set of incidents from last years) will be send by mail to the relevant partners.	Unknown.	Document name, IGP and PERIVALLON. Also, the relevant data, needed for developing the platform will be send via email to the partners.
Directory and file naming convention	To be defined in due course by the project consortium.	None.	Unknown.	link to public website/repository/Alfresco directory.
Clear version numbers	Yes.	No.		Yes.
Search keywords	No.	To be agreed within the consortium.		Yes.
Metadata creation	To be defined in due course.	No.		N/A
Metadata standards	To be defined in due course.	N/A	N/A	Not known yet.
Potential users find data method	By accessing the project’s platform.			

Making data openly accessible				
Openly available data	<ul style="list-style-type: none"> • Already existing data sets publicly available will be used. • Provided by Police Authorities, Border Guards and National/Regional Authorities partners: it is at the discretion of the partners who will provide the data. 	Data is confidential, only available for consortium partners. This data is not to be used outside this project or distributed in any way or form.	To be agreed within the consortium.	Through project website and social media channels, as official web site, Facebook and Telegram. Accessible will be made the data which will be agreed within the consortium.
Means of making data openly available	To be discussed for the second bullet point.	The data will be sent via mail or uploaded on the project platform to the partners that are involved in that task.	Idem.	Locally stored in partners' servers located, shared via email.
Methods or software tools needed to access data; Software documentation required to access data	TBD for the second bullet point	Microsoft Office, GIS viewer.	Via project's platform.	Will need to have Internet access, login details to the project's platform.
Place of metadata deposition	To be discussed for the second bullet point	Located on local servers. We can provide data (e.g. dataset of incidents) to PERIVALLON platform. Data provided by DWG can only be used within the consortium and needs to be handled	To be agreed within the consortium.	PERIVALLON Platform.

		securely so that unauthorised people cannot access it.		
Restrictions on use	To be discussed for the second bullet point	No requirement for login, considering the fact that the data will only be used inside the project consortium for the duration of the project. We assume that confidential data will be handled with care. Provided data needs to be deleted after the project.	Idem.	If the data will be shared on social media, there are no restrictions on use.
Correct execution of the access process	To be discussed for the second bullet point	The data is provided to a certain partner, this partner is responsible for the access and respect its confidentiality. The consortium has to manage how will be the identity of the person accessing the data.	To be agreed within the consortium	<ul style="list-style-type: none"> • The Protection of State Secrets Section, the Personal Data Protection Section, the Information Technologies and Communications Directorate. • Throw user account log-in verification.
Making Data interoperable				
Interoperable data	Yes. Mark columns (CSV) and objects (JSON) with clearly descriptive names and standard formats (e.g. integer, strings, floats).	No.	Yes.	No.

<p>Data and metadata vocabularies, standards or methodologies to facilitate interoperability</p>	<p>No.</p>	<p>To be agreed within the consortium</p>	<p>Yes.</p>
<p>Making Data re-usable</p>			
<p>Data sharing and re-use (license)</p>	<p>In case metadata is created, it will potentially be publicly available and will likely be re-shared with the same license as the original.</p> <ul style="list-style-type: none"> • Provided by Police Authorities, Border Guards and National/Regional Authorities partners: it is at the discretion of the partners. 	<p>No.</p>	<p>Yes. To be agreed within the consortium, whereas not personal data and not classified data.</p>
<p>Audience for reuse</p>	<ul style="list-style-type: none"> • Police Authorities, Border Guards and National/Regional Authorities Researchers/developers who would like to evaluate or develop modules of a similar nature. 	<p>No reuse.</p>	<p>Police Authorities, Border Guards and National/Regional Authorities.</p>
<p>Period of re-usability</p>	<ul style="list-style-type: none"> • Already publicly available data 	<p>Never.</p>	<p>To be agreed within the consortium.</p>

	<ul style="list-style-type: none"> • Provided by Police Authorities, Border Guards and National/Regional Authorities partners: it is at the discretion of the partners. 			
Restrictions on reuse	N/A.	Not to be reused in any shape or form.	To be agreed within the consortium	The data will be available for Police Authorities, Border Guards and National/Regional Authorities.

Table 11 - DT1 FAIR Approach

Data Type 2: Audio-visual and Imaging data			
Contributing partners: CERTH, POLIMI, SPA			
Making Data Findable			
Partner in charge	CERTH 1. Drone imagery 2. Satellite imagery and relevant insights from image classification algorithms	POLIMI	SPA
Data identifiers	1. Location and UAV (or other vehicle) that gathered the data and it will receive a respective denomination 2. Pollutant type	AerialWaste	Each data set will be associated with a unique identifier specific for each data set.
Place to find data	1. Not currently. 2. Original data set: - Can be found from Copernicus Open Access Hub (https://scihub.copernicus.eu/dhus/#/home) which requires credentials) Produced data set: - Can be found in Public website (e.g., M4D site), PERIVALLON repository.(for consortium partners)	aerialwaste.org	Local storage. Available by request.
Directory and file naming convention	1. See Data Summary. 2. To be decided within the consortium, but the project Acronym, the KR, and the pollutant type will be probably part of the file naming convention and the directory.	Consortium naming conventions.	
Clear version numbers	Yes.		
Search keywords	1. Yes.	Yes.	No.

	2. To be confirmed depending on the end application that will be provided by MT, and no automatic creation.		
Metadata creation	1. To be discussed initially and then expanded on that according to arising needs. Original images will be processed in order to identify pollutants on water surface. The metadata from the analysis include: - geoTIFF/ geoJSON file that indicates the pixels where the pollutant is found - accompanying JSON file the provides information on the original product used for the analysis and the outcome product.	See description of the annotations in Data Summary.	To be agreed within the consortium.
Metadata and documentation standards	1. Github, Mendeley, Zenodo 2. Produced data set: If published - DOIs, Github repos	DOI	Not decided.
Potential users find data method	Link to the respective repository provided in the project's platform, through the project's website, and through M4D team website. Project's platform, link that directs them to a repository or cloud storage.	Publication in Nature scientific data, Message to scientific mail lists	By request.
Making Data Openly Accessible			
Openly available data	1. Our initial assessment is to develop an openly available data set. 2. Original data:	AerialWaste is made public	No data will be made openly available.

	<ul style="list-style-type: none"> publicly available in Copernicus Open Access Hub <p>Produced data:</p> <ul style="list-style-type: none"> data will be openly available, under the license mentioned and can be published in project website, etc. 		
<p>Means of making data openly available</p>	<ol style="list-style-type: none"> in a repository, locally stored in partners' servers. <p>Original data:</p> <ul style="list-style-type: none"> accessible through Copernicus open Access Hub API <p>Produced data:</p> <ul style="list-style-type: none"> locally stored in partner's servers located in CERTH premises in Greece. 	<p>Zenodo for the data Github for the code public web site (aerialwaste.org)</p>	<p>Locally stored and accessible at request.</p>
<p>Methods or software tools needed to access data; Software documentation required to access data</p>	<ol style="list-style-type: none"> Depending on the partners' needs, some could be displayed in a user-friendly manner in the project's platform without many technical tools, simply internet access and clicking. The rest could be private, with logic access or they could be openly available. In either case, viewing this data would require a 3Dmap visualization tool such as WebODM. <p>Original data:</p> <ul style="list-style-type: none"> partners will need to have Internet access. credentials for accessing Copernicus API 	<p>The data, as well as the tools employed, are publicly released:</p> <ul style="list-style-type: none"> AerialWaste toolkit: Instructions and scripts to visualize it and draw basic statistics are published in the https://github.com/nahitorres/aerialwaste. AerialWaste model: the models along and the code to execute them are released on https://github.com/nahitorres/aerialwaste-model. Inspection Tool: the Jupyter notebook-based tool is also publicly released on 	<p>Partners will need to have Internet access or login details to the project's platform, as well as access to 3D software.</p>

	<p>Produced data:</p> <ul style="list-style-type: none"> • Internet access 	<p>https://github.com/nahitorres / demo-inspection-tool. After the execution of the models in new areas, the predictions can be visualized with the inspection tool.</p> <ul style="list-style-type: none"> • ODIN: the tool used for the data set annotation process and for the model evaluation phase is a previous work, available at https://nahitorres.github.io/odin-docs/. 	
Place of metadata deposition	<p>See Data Summary.</p> <ul style="list-style-type: none"> • Produced data/ code/ documentation: • Locally stored in partner's services located in CERTH premises in Greece. 	<p>Zenodo for the data github for the code public web site as an entry point (aerialwaste.org)</p>	<p>Locally stored on servers.</p>
Restrictions on use	<ol style="list-style-type: none"> 1. if stored in PERIVALLON platform only via login details. 2. The data set is expected to be open under the aforementioned license and thus no restrictions will be on use. Data retention: JSON metadata should accompany the geoTiff/ geoJSON that will allow the interpretation of the data. 	<p>Open access, see license info</p>	<p>By request, via PERIVALLON platform.</p>
Correct execution of the access process	<ol style="list-style-type: none"> 1. Funder 2. PI 	<p>Data are open access in ZENODO. Via ZENODO authentication.</p>	<p>N/A. User account log-in verification on the PERIVALLON platform.</p>
Making Data interoperable			

Interoperable data	<ol style="list-style-type: none"> No. STAC specification (https://stacspec.org/en) that is used for describing geospatial information will be considered. 	Yes. European standard waste code	Not yet decided.
Data and metadata vocabularies, standards or methodologies to facilitate interoperability	<ol style="list-style-type: none"> Yes. Not defined yet. 	No.	
Making Data re-usable			
Data sharing and re-use (licence)	<ol style="list-style-type: none"> Data all are re-usable, but not necessarily openly accessible. No license. Yes. Processed GeoTIFF/ geoJSON files that are the outcome of the analysis Metadata accompanying the image files. Under CC BY-NC-SA. 	Yes. See license info already provided DATA SUMMARY	No.
Audience for reuse	<ol style="list-style-type: none"> We will use this data to train our decision-making algorithms. In the future, this data set could be utilized by several robotic teams that are interested in either expanding this data set with their own data or to develop improved control algorithms. In the future, several academic / research groups could be interested in using the data set for evaluating their methods and results against the methods used by CERTH. 	Police Authorities, Border Guards and National/Regional Authorities and researchers.	

<p>Period of re-usability</p>	<p>Regarding the publicly available data: 1. As long as it is decided so and the data are produced, immediately. No embargo periods 2. For as long as there exists ample storage space (on local or cloud storage) and the server provider is not shutdown. N/A.</p>	<p>Minimum 3 years after project termination.</p>	<p>Not known yet. Discussions on reusability of data collected by SPA must be held within the organisation and then within the consortium. SPA needs information (e.g., purpose of reuse, addressees) not yet known that determines the reusability of the dataset.</p>
<p>Restrictions on reuse</p>	<p>No restrictions.</p>	<p>See license information, no restriction on types of users or duration of use.</p>	<p>Yes. Accessible at request.</p>

Table 12 - DT2 FAIR Approach

Data Type 3: Textual data Contributing partners: CENTRIC*, CERTH *this Data type will contribute to T4.3 (CENTRIC)		
Making data findable		
Partner in charge	CENTRIC	CERTH
Data identifiers	For the web crawling and URL scoring module, project and data identifiers could include a unique project name or ID, a timestamp, and a source URL. The data set denomination could be "Web Crawler and URL Scoring Data".	Already publicly available data set
Place of deposition	The data set is yet to be collected.	To be defined in due course by the project consortium.
Directory and file naming convention	Data will be in the dedicated data store.	
Clear version numbers	Timestamp, Source	Yes.
Search keywords	-II-	To be defined in due course.
Metadata creation	Timestamps, Source URL, Item Identifiers.	
Documentation & Metadata standards	For the documentation and metadata standards of the two software modules, we will use our own internal domain ID system which includes the source and a unique identifier for each data item. In addition to this, we will also include a timestamp as metadata to provide information about when the data was collected or stored. This approach will enable us to maintain consistency across all data items and allow for easy identification and tracking of individual pieces of data.	By accessing the project's platform.
Potential users find data method	Accessing the project's platform and viewing the items in the data store.	Link to the respective repo provided in the project's platform.

Making data openly accessible		
Openly available data	No the crawled web-data set.	Already existing data sets publicly available will be used.
Means of making data openly available	In the secure data store accessed through the platform.	Already publicly available.
Methods or software tools needed to access data; Software documentation required to access data	Internet access, log in details, and the right access permissions in the system.	
Place of documentation & metadata deposition	Platform: data and associated metadata, documentation.	Already publicly available. A copy of the data (and possible metadata) will be stored at CERTH secured servers.
Restrictions on use	PERIVALLON platform only via login details. Access will be audited.	Already publicly available. In case of usage restrictions (e.g. due to the generated metadata), access will be provided with credentials.
Correct execution of the access process	Access to the data store is managed by CENTRIC, the data audit module may also provide automated checks and alerts to ensure that access to the data store is in compliance with the established policies and procedures.	Funder. Already publicly available. In case of usage restrictions, user-account login verification will be required.
Making Data interoperable		
Interoperable data	No.	
Data and metadata vocabularies, standards or methodologies to facilitate interoperability	No.	Mark columns (CSV) and objects (JSON) with clearly descriptive names and standard formats (e.g. integer, strings, floats). Standard vocabularies yes. No mapping of ontologies.
Making Data re-usable		

<p>Data sharing and re-use (license)</p>	<p>Data will be shared. License: Data will be licensed.</p>	<p>Yes. Already publicly available data will be used. In case metadata is created, it will potentially be publicly available; will be determined in due course. In case metadata is created, the data set will likely be re-shared with the same license as the original.</p>
<p>Audience for reuse</p>	<p>Police Authorities, Border Guards and National/Regional Authorities.</p>	<p>Police Authorities, Border Guards and National/Regional Authorities. Researchers/developers who would like to evaluate or develop modules of a similar nature.</p>
<p>Period of re-usability</p>	<p>The data collected by the web crawler and link evaluator modules should only be made available for re-use after any necessary redaction or anonymization has been completed to remove any identifiable personal information. Additionally, there may be legal or ethical considerations that dictate any embargo periods for making the data available for reuse. For example, if the data is part of an ongoing investigation, it may not be appropriate to release the data until the investigation has concluded.</p>	<p>Already publicly available data will be used.</p>
<p>Restrictions on reuse</p>	<p>Police Authorities, Border Guards and National/Regional Authorities</p>	<p>Already publicly available data will be used.</p>

Table 13 - DT3 FAIR Approach

Data Type 4: User Requirements Data	
Contributing partners: KEMEA	
Making data findable	
Partner in charge	KEMEA
Data identifiers	In the related tasks leaded by KEMEA, no data identifiers will be used.
Directory and file naming convention	Not applicable yet.
Place to find data	-II-
Clear version numbers	Not known.
Search keywords	Currently not known.
Metadata creation	Currently not known.
Metadata standards	A standard template for the use case scenarios.
Potential users find data method	Currently not known.
Making data openly accessible	
All elements on making data openly accessible	Currently not known.
Making Data interoperable	
Interoperable data	No.
Data and metadata vocabularies, standards or methodologies to	Not known yet.

facilitate interoperability	
Making Data re-usable	
Data sharing and re-use (license)	No.
Audience for reuse	Currently not known.
Period of re-usability	
Restrictions on data re-use	

Table 14 - DT4 FAIR Approach

Data Type 5: AIS data	
Contributing partners: MT	
Making data findable	
Partner in charge	MT
Data identifiers	Unique identifiers for individual data sets. No link currently available
Directory and file naming convention	Project acronym and some geo-reference ID to the area of interest
Clear version numbers	Yes.
Search keywords	-II-
Metadata creation	To be confirmed depending on the end application that will be provided by MT, and no automatic creation.
Metadata standards	To be confirmed
Potential users find data method	Aggregated data resulting from the analysis of the described AIS data should be accessible by the project's platform.
Making data openly accessible	
Openly available data	Results (insights/conclusions) from the analysis of AIS data corresponding to maritime routes that pertain to e-waste trafficking may be made public through academic publications, online demos. At this stage of the Project, the availability of the results is under discussion.
Means of making data openly available	Identifiable by a DOI linked with a repository in a platform such as Zenodo.
Methods or software tools needed to access	Internet access.

data; Software documentation required to access data	
Place of metadata deposition	linked with a repository in a platform such as Zenodo.
Restrictions on use	TBC.
Correct execution of the access process	PI
Making Data interoperable	
Interoperable data	Yes.
Data and metadata vocabularies, standards or methodologies to facilitate interoperability	AIS approved recommendation standards: https://www.itu.int/rec/R-REC-M.1371-5-201402-I/en Geospatial data sets that will be used extensively in the project will follow OGC standards1 https://www.ogc.org/standards
Making Data re-usable	
Data sharing and re-use (licence)	Yes, especially aggregated data resulting from analysis of AIS data.
Audience for reuse	To be used by end users of the maritime use case.
Period of re-usability	TBC.
Restrictions on data re-use	For non-commercial purposes, period TBC.

Table 15 - DT5 FAIR Approach

2.3 Allocation of Resources for FAIR data

Research data types produced by PERIVALLON are aimed to be FAIR, taking into consideration the confidentiality and sensitiveness of data being handled within the project and its objectives.

While costs for making the project’s data FAIR should not amount to additional costs than the already budgeted, partners were asked in the DMP Questionnaire if they foresee any cost that might be encountered for making the project FAIR. Additionally, they were asked whether they would incur costs in order to prepare the data for sharing/preservation, whether there are any fees associated with storing data or if there is a need for payment to any specialist to assist or maintain or access data.

Whereas the majority of the partners stated that no expected or additional costs are to be incurred for the purposes of making data FAIR, some partners declared that either the costs at the time of writing this deliverable are unknown/undefined or possible costs need to be discussed and agreed within the consortium.

An update of these statements will be required in the next update of this deliverable (see [Subsection 2.3 ‘Timetable for DMP Updates’](#)).

The detailed reply by partners is shown in the table below.

Allocation of Resources per Partner	
Partner	Statement
CENTRIC	It would require the data to be transformed to not include personal data through anonymization.
CERTH	To be discussed.
DWG	None.
ETRA	Costs that might be encountered for making the project FAIR are currently unknown.
HP	No expected costs. To be agreed within the consortium
IGP	No.
KEMEA	No, this question should be answered by technical partners.
MT	No extra costs are foreseen.
POLIMI	Personnel cost for annotation and data set extension (e.g., 0.5 FTE per year). Anticipated costs sharing/preservation 0.5 FTE of early stage researcher.
RBP	Some costs are anticipated, but not (yet) defined.
SAFE	No.
SATCEN	No.

SMOB	SMOB will incur in fee(s) associated with storing data in repositories.
SPA	All costs will fall within the announced and allocated budget.

Table 16 - Allocation of Resources per partner

2.4 Data Security for FAIR data

KPI17 of the PERIVALLON GA foresees data security as a parameter where all data breaches are prevented.²⁰ It should be noted that this section does not include the PERIVALLON security approach given to the PERIVALLON platform.

Therefore, it is key to first identify which are the major risk to data security and to which tasks can they be related in order to secure a proper data security approach.

Major security risks to data were identified by partners as follows:

- Unauthorised access to the data;
- Tracing of personal data;
- Data breach;
- Security breach;
- Tracing personal data, such as names and addresses in crawled content, could lead to identifying individuals, posing a risk to data security;
- Unintentional gathering of personal data from drones’ activities (e.g., photos).

Despite the stated, it should be noted that no formal assessment to address major risks to data security has yet been performed by the consulted partners.²¹

Having identified possible risks to data, data sets must be securely managed in order not only to guarantee a FAIR management of PERIVALLON data sets but also to secure the project’s success.

Therefore, technical partners’ and end users’ data security measures will be detailed in the table below which will be applicable to those data sets and data types which they collect/process within the project’s context.²²

Security measures, designed for the protection of personal data, are identified below as the ‘Steps to protect privacy and confidentiality and additional measures’. Where personal data is processed, Art. 5(1)(f) GDPR becomes particularly relevant as it requires data controllers and data processors to process personal data ‘[...] in a manner that ensures appropriate security of personal data’.

To achieve security of personal data processing, the GDPR states that ‘the appropriate technical and organisational measures to ensure a level of security appropriate to the risk’.²³

Particular technical and organizational measures for the processing of personal data are addressed below in [Subsection 3.1](#).

Data Security Measures per Partner	
Partners	Data Security measures
CENTRIC	<p>Steps to protect privacy/confidentiality: Access control will be implemented to restrict who can view or modify data stored in the secure data store. The platform will be protected with permission-based access to prevent unauthorized access. The data audit module will keep track of any access or modification to the data and will be monitored to ensure that only authorized personnel have access. URL scoring using possible pseudonymised and anonymized data sets to reduce the risk of re-identification.</p> <p>Encryption: SSL if accessed over the internet.</p> <p>Encryption: Masking features and data anonymisation.</p> <p>Measures for data storage/management: Access Controls: This includes various security measures such as authentication, authorization, and encryption. Authentication ensures that only authorized users are allowed access to the data. The authorization ensures that each user has access only to the data that is necessary for their job responsibilities using permissions, and encryption ensures that the data is protected from unauthorized access by third parties.</p> <p>Data Retention Policies: This involves determining how long certain types of data should be kept and then disposing of it securely when it is no longer needed. This helps to reduce the risk of data breaches and protect sensitive information using settings within the secure data store.</p> <p>Data Integrity Checks: Data integrity checks are also crucial for safe data storage and management. These checks ensure that the data has not been tampered with or corrupted in any way. The audit modules perform these checks, including data hashing and block chaining Alerts and Monitoring: Finally, the use of alerts and monitoring can also help to ensure safe data storage and management. This involves monitoring the data store and audit module for any unusual activity, such as attempts to access data without proper authorization or changes made to the data that are not expected. If any such activity is detected, alerts are generated to notify the appropriate personnel, who can then take appropriate action to prevent any further unauthorized access or data breaches.</p> <p>By implementing these security measures, the audit module and secure data store can ensure that the data is kept safe and secure and that any attempts to access it are monitored and controlled.</p> <p>Data recovery protocol: None.</p> <p>Data breaches protocols: None.</p> <p>Data Preservation: No.</p>

	<p>Data Maintenance: The PERIVALLON platform secure data store. The data store settings, decided by the PERIVALLON consortium.</p> <p>Personal Data deletion: Secure data store will use an automated deletion process that can delete personal data that is no longer required to be stored. These processes can be triggered by predefined retention periods in the data store settings.</p>
<p>CERTH</p>	<p>Steps to protect privacy/confidentiality: All data will be collected in accordance with the licences and terms & conditions of the data providers and will be publicly available, or properly pseudo/anonymised if any personal information will unintentionally exist. All the security measures and access controls described in “Measures for data storage/management” will be applied.</p> <p>Encryption: blurring, mosaic, black bars etc, data pseudonymisation and/or anonymization.</p> <p>Measures for data storage/management: CERTH has already in place certain security measures such as: The server hosting the databases will be accessible only by authorised users through authentication (using passwords of high complexity). A firewall will also be in place to allow only specific (whitelisted) IPs to access the server and to restrict the access of each whitelisted IP only to specific ports/services. Different access privileges to the database will be available to ensure that the authorised users will only have access to the stored data on a need-to-know basis, i.e., that the authorised users will have access only to the stored data needed to fulfil their tasks. The server will be located inside a locked room, accessible only by authorised personnel. - Devices that will store a backup of the data will follow the same security procedures as the main server. For any remote interactions with the server (e.g., remote control or data transfer), secure protocols such as ssh/stfp are used. Any processing of the data will be performed on that server. In case processing will be needed on other machines, the same security measures of the server will be applied to the respective machine. - If none of the collected personal data are deemed useful for the project, they will be immediately deleted or anonymized. CERTH/ITI has also an internal security policy.</p> <p>Data breaches protocols: In case of data breach Art. 33 and 34 of GDPR are applicable. CERTH’s procedure enshrines the following steps: 1. All the necessary actions to stop the breach 2. Filing of all the information regarding the data breach 3. Analysis of the causes and the consequences of the incident 4. Legal analysis of the incident 5. Organization of the responding system to requests for information by the data subjects.6. Notification of the data breach to the Hellenic Data Protection Authority and to the data subjects, when this is considered as necessary, according to GDPR. 7. Organization of a recovery plan.</p> <p>Data Preservation: Yes. The storage of collected data will be realised according to the EC’s guidelines and will be up to 5 years after the completion of the project for auditing purposes. These data will be required for the duration of the project: (i) for scientific research purposes, (ii) to facilitate the functionality</p>

	<p>of other modules of the project, and (iii) for demonstration purposes. CERTH/ITI will be the only one responsible for the information collected.</p> <p>Data Maintenance: Locally in a secure server, located in CERTH premises. When data sets will be stored into the PERIVALLON repository these provisions will be addressed by the hosting partner.</p> <p>Personal Data deletion: As aforementioned, if none of the collected personal data are deemed useful for the project, they will be immediately deleted or anonymized..</p>
<p>DWG</p>	<p>Steps to protect privacy/confidentiality: We label data as confidential and provide it only under that label. Access control should be restricted, and it can not be used outside the purpose of the project without explicit consent of us. Regular privacy and security are applicable.</p> <p>Encryption: No.</p> <p>Measures for data storage/management: Security policy is applicable and removal of the provided data after the project is required.</p> <p>Data recovery protocol: none.</p> <p>Data breaches protocols: Yes. Data breach protocol is provided within De Watergroep.</p> <p>Data Preservation: No. The data is provided for the purpose of the project and the development of the tool. If the data is to be used after project life, we will need a separate agreement on that. In principal, data is not to be used after the project.</p> <p>Data Maintenance: N/A.</p> <p>Personal Data deletion: Privacy policy of De Watergroep is applicable.</p>
<p>ETRA</p>	<p>Steps to protect privacy/confidentiality: Not applicable as no raw data sharing is envisaged. ETRA will provide VPN connection and ensure only users with valid certificates will be able to access the data we produce, on a need-to-know basis. User and password authentication will be required</p> <p>Encryption: Not yet defined. Potentially by means of the secure version of data communication protocols (MQTTS, HTTPS,...).</p> <p>Encryption: Masking features and data anonymisation.</p> <p>Measures for data storage/management: ETRA expects to use access credentials and grant access only to specific IP ranges.</p> <p>Data recovery protocol: none</p> <p>Data breaches protocols: in place. 1) Identify the breach, 2) Contain and notify the breach, 3) Assess the risks and make a record, 4) Risk mitigation and implementation of improvements.</p> <p>Data Preservation: No.</p>

	<p>Data Maintenance: We do not foresee the long-term preservation of data.</p> <p>Personal Data deletion: Not applicable since we do not foresee the processing of raw data.</p>
HP	All details on data security are to be agreed within the consortium
IGP	No specific description. Other security measures are security at premises, deletion of data after use and regular testing, recording of data processing. Making records when the data are viewed or printed. Also, will be determined where the printed data will be stored and who will have the access.
KEMEA	On steps to protect privacy/confidentiality, encryption, measures for data storage/management, data recovery protocol, data breaches protocols, data preservation and maintenance, and personal data deletion information: This question refers to technological partners of the project.
MT	<p>Steps to protect privacy/confidentiality: End-to-end security mechanism to for authentication when accessing the data.</p> <p>Encryption: Masking features and data anonymization.</p> <p>Measures for data storage/management: Input sanitization and protection against logging breaches.</p> <p>Data recovery protocol: None</p> <p>Data breaches protocols: None.</p> <p>Data Preservation: no intention for long term preservation.</p> <p>Data Maintenance: Institutional, aggregated data should be accessible via API calls through the PERIVALLON platform. PI on behalf of Marine Traffic will maintain the data long term.</p> <p>Personal Data deletion: N/A.</p>
POLIMI	<p>No specific steps to protect privacy/confidentiality: no encryption, and standard data storage protection practices.</p> <p>Data Maintenance: ZENODO, and decision taken by POLIMI on data, which will be kept, and the period of storage.</p>
RBP	<p>Steps to protect privacy/confidentiality: Use of pseudonymised and anonymized data sets, access control in the storage service, Staff training in the field of personal data protection, staff training on granting access accounts to internal/external databases, data transmission through secure channels, physical access to the rooms where there is equipment that processes personal data is strictly limited to users designated by the controller, data backups, deletion of data after use and regular testing, recording of data processing.</p> <p>Measures for data storage/management: access to all stored data is limited to authorized persons, Establishing storage terms depending on the nature of the data and the purpose of the processing.</p>

	<p>General Inspectorate of the Romanian Border Police implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with GDPR.</p> <p>Those measures are:</p> <ul style="list-style-type: none"> • Staff training in the field of personal data protection, • Staff training on granting access accounts to internal/external databases, • Data transmission through secure channels, and access to all stored data is limited to authorized persons, • Establishing storage terms depending on the nature of the data and the purpose of the processing, • Physical access to the rooms where there is equipment that processes personal data is strictly limited to users designated by the controller, • There is a plan of measures regarding the minimum security requirements for the processing of personal data.
<p>SAFE</p>	<p>Steps to protect privacy/confidentiality: Data will be correctly categorized according to their level of confidentiality. Based on this categorization, access to data will be granted only to personnel with a defined “need-to-know”.</p> <p>Encryption: Data anonymization or pseudonymisation whenever necessary.</p> <p>Measures for data storage/management: Data will be accessed and managed by authorised personnel only. Access controls, including personal authentication, are active at SAFE premises according to different level of confidentiality and eventual classification.</p> <p>Data recovery protocol: N/A</p> <p>Data breaches protocols: In case of any data breach, SAFE qualified personnel will immediately inform the National Data Protection Authority and/or the National Security Authority depending on the type of breach.</p> <p>Data Preservation: Data collected for PERIVALLON will be preserved the shortest time as possible to ensure the current implementation and audit of the project.</p> <p>Data Maintenance: Eventual classified and/or confidential data will be maintained in physical copy only and stored in a dedicated room with adequate security protection and access restriction. All the other data will be maintained, for the concerned period of time, on the Data Management Platform used at institutional level.</p> <p>Personal Data Deletion: According to SAFE Privacy policy.</p>
<p>SATCEN</p>	<p>Questions are not applicable to the partners as far as we are not collecting, processing and/or storing research data.</p>

<p>SMOB</p>	<p>Steps to protect privacy/confidentiality: Use of pseudonymised and anonymized data sets Access control in the storage service Access control to PERIVALLON platform.</p> <p>Encryption: Not decided yet if test data needed for integration will be encrypted at storage location. Not decided yet, but during transfers between components, data most probable be encrypted using SSL. Will be decided during the design phase.</p> <p>Measures for data storage/management: Deletion of any local data after use and regular testing.</p> <p>Data recovery protocol: None.</p> <p>Data breaches protocols: None.</p> <p>Data Preservation: No.</p> <p>Data Maintenance: Not applicable.</p> <p>Personal Data deletion: Internal data deletion procedures.</p>
<p>SPA</p>	<p>Steps to protect privacy/confidentiality: SPA follow internal safety regulation routines and GDPR, including the national adaptations.</p> <p>Encryption: Unknown.</p> <p>Measures for data storage/management: Recording of data processing. Data backups.</p> <p>Data recovery protocol: None.</p> <p>Data breaches protocols: Yes. Own security/privacy policies.</p> <p>Data Preservation: No.</p> <p>Data Maintenance: Institutional.</p> <p>Personal Data deletion: SPA will aim to adapt to a joint view on the matter, as long as it is coherent with internal policies on the matter.</p>

Table 17 - Data Security Measures per partner

3. Legal Aspects - Data Protection

3.1 General aspects

Processing personal data is foreseen for the two data sets addressed in this first iteration of the DMP. Therefore, special attention to data protection aspects is initially stated in this DMP. A detailed analysis of applicable data protection legislation, rules and regulations applicable to the project will be dealt with in MS3 'Initial report on the legal and ethical framework' under UNIVIE responsibility due in M12.

A category of personal data is designed as different data protection implications arise out of such category.

First and foremost, the legal applicable framework on the protection of personal data constitutes, in particular:

- The Charter of Fundamental Rights of the European Union (the 'Charter'),²⁴
- The Treaty on the Functioning of the European Union ('TFEU'),²⁵
- The Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data ('GDPR'),²⁶
- The Directive (EU) 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data ('LED')²⁷ and its national implementations.

The applicability of the GDPR to the partner's data processing is not analyzed herein as well as the EU Directive 2016/680 in a law enforcement context²⁸. Such analysis will be one of the results materialized in the upcoming MS3 'Internal report on legal and ethical issues'. Despite the stated, two main concepts defined by the GDPR need to be highlighted: 'personal data' and 'anonymisation/ pseudonymization' of personal data. These two concepts are relevant to the extent of data herein processed can be categorized under personal data and therefore partners should be aware of the possible applicability of the GDPR or LED to their processing.

Personal data is defined in the GDPR as:

'any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'²⁹.

Making a distinction of personal and non-personal data (anonymous data) is crucial for all activities of the project. Therefore, when referring to the anonymisation of data, this term refers to the processing of personal data so as to irreversibly prevent the identification of the natural person to whom the personal data is linked³⁰.

Last but not least, as stated in PERIVALLON GA, PART B, p. 45, personal data will be processed and further processing of previously collected personal data is involved. In this regard, the partners have to ensure that personal data is collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes³¹. For the further processing of previously collected

personal data, Deliverable D8.2 – POPD – Requirement No. 2 will provide, where needed, an explicit confirmation that the beneficiary has lawful basis for the data processing and that the appropriate technical and organisational measures are in place to safeguard the rights of the data subjects must be submitted as a deliverable before the start of the relevant activities.

General data protection approach for the PERIVALLON project

The 'PERIVALLON consortium is committed to perform the project in compliance with the GDPR and implementing local legislation.'³² In this regard, the processing of personal data will follow data protection legislation, including the GDPR, the LED as well as partners' national data protection laws and their internal policies. It should be noted that three partners³³ are not under the territorial scope of the GDPR, and therefore, the confirmation of their applicable legal framework on data protection will be reflected in Deliverable D8.2 – POPD – Requirement No. 2.

Personal data processing will be safeguarded within the PERIVALLON project, in particular by³⁴:

- identifying the kind of personal data collected, the partners involved in the processing of personal data, and the legal basis declared by partners (see 'Personal data categories' below),
- implementing technical and organizational measures (see 'Technical and organizational measures' below), including those declared by partners in [Subsection 2.4 on Steps to protect privacy/confidentiality 'Data Security'](#), and
- concluding data processing agreements where needed (see [Subsection 3.4](#)).

Personal data categories

Personal data are defined in Article 4(1) GDPR, and therefore personal data is understood herein under such definition. In PERIVALLON, the following categories of personal data and special categories of data are/may be, in particular, processed:

- **Category 1:** Personal data of project partners.
- **Category 2:** Personal data of third parties to the project.
- **Category 3:** Personal data obtained from WP3, WP4 and WP5 activities.

The categories of personal data, the purposes of processing, the legal basis for processing and the partners involved are in detail described herein:

Category 1:

- Content: Personal data e.g., name, contact details, address, portrait images of project partners.
- Purpose(s): This category is collected for the purposes of performing the contractual obligations arising out the GA.
- Legal basis: Art. 6(1)b GDPR. Art. 11(5) paragraph 5 CA.
- Partners involved: partners to the PERIVALLON consortium.

Category 2:

- Content: Personal data of third parties, as external contacts for the performance of objectives of the project, e.g., name and contact details of external ethics advisors.
- Purpose(s): The category is collected for the purposes of performing the contractual obligations to the partners arising out the GA, like the establishment of the Independent Ethics Board, and promotion activities of PERIVALLON KRs.
- Legal basis: Art. 6(1)(a) or (f) GDPR, to be determined. Where Art. 6(1)(a) is the chosen appropriate legal basis for personal data processing, draft informed consent template is available to the consortium partners in this DMP (see [Annexes](#)).
- Partners involved: partners involved in the project management of the project, IGP.

Category 3:

- Content: Personal data may be collected in connection with the conduct of tasks WP3, WP4 and WP5, in particular in connection with the web scrapping and drone image collection. Data may include usernames, comments, posts in which personal identifiers can be detected.
- Purpose(s): For the purpose of performing tasks WP4 and WP5, within no purpose of identifying personal data but could be included in the data set that is collected for the respective tasks.
- Legal basis: Art. 6(1)f GDPR, Art 9(2)(j) GDPR, Art. 89 GDPR.
- Partners involved: mainly, CENTRIC, CERTH, RGP, SMOB, partners of WP3, WP4 and WP5.

Personal data declared by partners to be processed during the research:³⁵ CENTRIC, CERTH, IGP, SMOB.

Partners which do not expect to process personal data: DWG, HP, KEMEA, MT, POLIMI.

Technical and organizational measures

The PERIVALLON consortium agreed on the following technical and organizational measures:³⁶

- render anonymous personal data prior to granting access to/ contributing/ or otherwise sharing it with the other parties, in a manner that the data subject is not or no longer identifiable and that the data cannot anymore be considered Personal data pursuant to the GDPR, agreed upon with the Project Coordinator, Project DPO and accompanied by appropriate safeguards and agreements.
 - The common denominator method declared by partners to render personal data anonymous is encryption.
- pseudonymise online data where anonymization is considered impossible. Therefore, the following subsequent measures will apply:
 - Perform the research, namely collection and use of online data in a minimum necessary manner to perform the research (Data Minimisation Principle).

- Secure that online data will come from public sources.
- Ensure that all legal requirements are implemented, including requirements stemming from the GDPR and from national legislation, e.g., a legal basis for processing, and that before any personal data is shared, appropriate data sharing agreements (such as Joint Controllership Agreements or Processing Agreements) must be signed.
- assess the data protection roles in the consortium by UNIVIE and the coordinator (T2.4) and the relevant agreements drafted and signed.
- discuss other operational, managerial, and technical measures to be implemented.
- monitor on a continuous manner the compliance in the use of personal data, within T1.5. This includes a careful assessment of all data sets, to assess the risk of re-identification and whether they can be deemed anonymized based on legislation and state-of-the-art literature.
- destroy collected data within 12 months of the completion of the project to allow for the end of project dissemination and follow-up, unless specifically requested and agreed.

Finally yet importantly, regarding online data collected through web crawlers: for the lack of an API and clear research-use policy, the conditions for further use will be assessed on a case-by-case basis. It will nevertheless be ensured that data subject’s expectations of privacy will be considered, and if in doubt, the owners / administrators of websites will be contacted. Information about the users will be anonymised at the stage of data collection.

3.2 Legal responsibilities

The role of the project guardian of the application of the GDPR is mainly done by UNIVIE when performing the analysis of the legal and ethical framework compatible with the PERIVALLON project. Additionally, ‘while each partner having a Data Protection Officer (DPO) will have their DPO assist with collecting and processing personal data within the project’.³⁷

Despite the stated, partners will support the compliance with data protection at their premises by following data protection or security policies applicable to them, and/or engage, where necessary, their in-house DPO/ Legal department in charge of privacy and/or data protection issues. Below a list of those technical partners’ and end users’ partners whose data protection policies/ contact persons are available:

Partner	Data Protection Policy/ Data Protection Contact Person
ETRA	dpo@grupoetra.com
CERTH	A DPO has been appointed (dpo@certh.gr)
CENTRIC	Data protection officer - Helen Williamson dpo@shu.ac.uk . Contact points - h.gibson@shu.ac.uk jordan.thompson@shu.ac.uk

Table 18 - Data Protection Contact Person per partner

A major task to be performed by the consortium, when the processing of personal data is confirmed, is to define the roles of partners within the personal data processing activities, mainly defined in the GDPR as data controller³⁸ or data processor³⁹, which result will be reflected in the next update of this DMP.

3.3 Data sharing within the consortium

PERIVALLON partners have agreed⁴⁰ to comply fully with the GDPR and ‘in the event that it is necessary to transfer personal data outside the EEA, the party transferring the personal data submitted by another Party shall notify the latter before any transfer so that it can fulfil its obligations to inform the persons concerned’. It should be noted that any reference to the processing of personal data in the CA should prevail over the analysis performed in this DMP.

Nevertheless, personal data sharing is not prohibited among research partners provided legal requirements are complied, in particular⁴¹:

- Take appropriate measures to preserve personal data safety are taken,
- Use personal data only if the purposes specified and make no other use of them,
- Only transfer all or part of personal data thus transmitted outside the EU or any country ensuring an adequate level of protection within the meaning of the GDPR,
- Notify the other parties as soon as possible of any security breach concerning personal data transmitted by other parties,
- Assist each other in responding to any request from the natural persons concerned.

Where there is a need for personal data sharing within the EU, partners should enter into Data Processing Agreements before sharing such personal data within the respective partners. Where such a need arises, it is recommended to consult with ETRA and UNIVIE before any sharing takes place and UNIVIE will provide a Data Processing Agreement template where necessary.

Categories 1 to 3 have a potential sharing aspect where such agreements are required to be in place.

Special attention should be given to transfer of personal data with non-EU countries and countries apart from the additional EEA countries (Norway, Iceland, and Liechtenstein), as well (defined as ‘third countries’), to which the GDPR applies.

In this regard, personal data sharing with third countries is regulated under Chapter V of the GDPR (Article 44 to 50) and the conditions therein must be complied with.

In PERIVALLON, consortium partners CENTRIC (UK), IGP (Moldova) and TAMAR (Israel) are located in countries beyond the territorial scope of GDPR.

If a transfer is envisaged to happen within the research activity in PERIVALLON, the GDPR foresees that a data transfer to third country may take place where there is an *adequacy decision*⁴² in place.

- For **Israel**: In the case of Israel, such adequacy decision is in force, and therefore, no further measure is required.⁴³
- For **United Kingdom**: An adequacy decision applicable to the UK entered into force in June 2021, and therefore no further measures need to be in place.⁴⁴

As stated in Art. 46 GDPR, in the absence of an adequacy decision, a controller or processor may transfer personal data to a third country only if the controller or processor has provided appropriate safeguards without requiring any specific authorisation from a supervisory authority and on the condition that enforceable data subject rights and effective legal remedies for data subjects are available, by:

- a legally binding and enforceable instrument between public authorities or bodies;
- binding corporate rules in accordance with Article 47;
- standard data protection clauses adopted by the Commission in accordance with the examination procedure referred to in Article 93(2);
- standard data protection clauses adopted by a supervisory authority and approved by the Commission pursuant to the examination procedure referred to in Article 93(2);
- an approved code of conduct pursuant to Article 40 together with binding and enforceable commitments of the controller or processor in the third country to apply the appropriate safeguards, including as regards data subjects' rights; or
- an approved certification mechanism pursuant to Article 42 together with binding and enforceable commitments of the controller or processor in the third country to apply the appropriate safeguards, including as regards data subjects' rights.

As there is no adequacy decision in place for **Moldova**, the appropriate safeguards as detailed above will be evaluated in Deliverable D8.2 – POPD- Requirements No.2 in cooperation with UNIVIE and the respective partner's legal department/DPO, such personal data transfer would be required for the purposes of the project.

3.4 Data Processing Agreements

If data is expected/planned to be shared within the consortium, data processing agreements should be concluded before any personal data sharing takes place.

The first step to take in view of this purpose is to map out the processing of personal data within the project, and to identify which type of personal data intends to be shared with whom and for which purposes.

Partners declaring the need for personal data sharing with certain consortium partners within their research activities stated that the sharing will take place within the identified Category 3.

The second step is to identify the roles of partners in personal data processing activities. When considering the establishment of data processing agreements, one should identify the roles of the partners involved in the data processing, which can be defined according to the GDPR as data controller or data processor.

Once the roles are identified, the controllership models/structures normally identified are:

- Controller – Controller data controllership.
- Controller – Processor data controllership.
- Processor – Sub processor data controllership.
- Joint data controllership (Art. 26(1) GDPR).

For the processing of personal data that occurs in relation with the PERIVALLON, partners seem *prima facie* to act as individual controllers, and therefore attend individually to the duties and responsibilities arising from data protection legislation.

A dedicated session within the project partners involved will take place to find out the details on personal data sharing for Category 3, and will be organized by UNIVIE.

This is an initial analysis based on the information provided by project partners at this stage of the project. This analysis will be further developed for the updated DMP at M15 by UNIVIE, in particular once the remaining data sets are evaluated.

3.5 International data transfers

International data transfers during the course of the project have been addressed above under ‘Data sharing within the consortium’. Special attention should be given if any project task, resource, tool, service is outsourced to stakeholder is located outside the scope of the GDPR. This aspect will be reviewed in the update of this DMP in M15.

3.6 Data Protection Impact Assessment

The DPIA is a data protection organizational tool, which enables the identification and minimization of data protection risks. Art. 35 GDPR requires controllers to carry out a Data Protection Impact Assessment (‘DPIA’) on the processing activities, which are likely to result in a high risk for individuals, such as large data processing and processing of special category of personal data.⁴⁵

As a first step, the consortium has to determine whether a DPIA is needed for the project. The result of this assessment is planned to be performed by MS3 and complemented for deliverable D8.2 – POPD – Requirement No. 2.

4. Ethical Aspects

In addition to respecting legal obligations, all EU funded projects are guided by ethical considerations and the values and principles on which the EU is founded⁴⁶. The processing of data must therefore adhere to the highest ethical standards and the applicable EU, international and national law on ethical principles.⁴⁷

The following section will give a short overview of the most important ethical risks from the data protection perspective, regarding the personal data generation, use, storage, and sharing, that are or will be applied in the project. Additionally, ethical responsibilities relating to these risks will be outlined.

A further and detailed analysis of ethical risks and mitigation measures beyond the ethical aspects on data protection will be provided in D8.2 – POPD – Requirement No. 2. Additionally, ethical requirements that will shape the development of technology in PERIVALLON will be analysed in D2.4.

4.1 General

According to the European Commission's guidelines⁴⁸, projects that involve data processing operations that may entail higher ethics risks, require a detailed analysis of such risks.

In order to identify ethics risks in PERIVALLON, an overview of planned data collection and processing operations has been made, where information has been collected through the questionnaires, filled in by the respective partners. At this initial stage of the project the overview, presented in the chapter "Handling of PERIVALLON data according to FAIR principles", covers Research and Administrative data sets and will be updated throughout the project.

Based on the collected information the following activities have been identified as possibly posing higher ethical risks. Firstly, complex processing operations and/or the processing of personal data on a large scale and/or systematic monitoring of a publicly accessible area on a large scale might be conducted. Secondly, collecting data outside the EU or transferring personal data collected in the EU to entities in non-EU countries might be performed.

Such activities require particular attention since ethical risks, such as risk of discrimination, stigmatisation, data breaches, threats to the safety or security of participants and the potential for misuse of the research methodology or findings are higher based on the type of data, data subjects, processing techniques or different data protection standards⁴⁹.

4.1.1. Data collection and processing techniques

Complex processing operations and/or the processing of personal data on a large scale and/or systematic monitoring of a publicly accessible area on a large scale is likely to raise higher ethics risks.⁵⁰ Additionally, data mining (including data collected from social media networks), 'web-crawling' or social network analysis and using artificial intelligence to analyze personal data could also pose a higher ethical risks.⁵¹

In case when large amounts of personal data are processed, there is a higher risk of harm if that data is mishandled or misused. Additionally, such processing of data can amount to lack of control of data subjects. When data is collected through systematic monitoring on a large scale, the risk of potential abuse of power increases as well.

In PERIVALLON use of satellite data, use of data collected via UAV, web-crawling, web-scraping and large-scale or big data processing related to environmental crime is being foreseen. Therefore, special attention to the risks mentioned above is necessary.

4.1.2 Involvement of non-EU countries

Three partners from the PERIVALLON consortium are located in the country outside of the EU, namely in the UK, Israel and Moldova. Since transfer of data, including personal data, from the EU to these non-EU countries as well as import from them to another non-EU country and EU is planned, certain conditions must be fulfilled in order for such data exchange to be legally and ethically complaint.

The legal aspects of personal data transfer to non-EU-countries are dealt with in [Subsection 3.5](#) above, and mainly address the requirements set forth in Chapter V of the GDPR. Besides the risk of inadequate data protection, some ethical risk, often as a consequence to the latter, might also arise. Some examples are risks of exploitation of data for commercial purposes, lack of transparency of handling with data and risks related to the cultural differences, which may lead to different expectations and standards in regard to data protection.

4.2 Ethical responsibilities

In order to counter and minimise the ethical risks mentioned above partners must fulfil certain legal and ethical responsibilities. These include providing adequate legal basis for collection and processing of personal data, or carrying out data collection and processing operations that pose higher ethical risks under the control of the relevant national authorities. In such cases and if necessary, DPIA will be conducted to ensure an appropriate level of data protection and minimise risk to the data subjects' rights.

Due to the severity and number of the ethics issues raised by the project, an External Independent Ethics Advisor will be appointed to monitor the ethics issues.⁵² Additionally, UNIVIE will act as an Ethics Manager.

Beside appointed authorities who are responsible for the monitoring of ethical issues, researchers within a project must comply with ethical codes of conduct, such as the European Code of Conduct for research integrity (ALLEA)⁵³. This implies the compliance with⁵⁴.

- 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 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 organisation, for training, supervision and mentoring, and for its wider impacts.

5. Open Data Initiatives

Sharing and providing open access to data without any restrictions on its use or distribution can increase transparency, collaboration, and innovation. This is possible by making data available to a wider audience, such as researchers, policymakers, and the public.

The European Commission has several open data initiatives aimed at promoting the availability and accessibility of data,⁵⁵ including in the form of legal framework:

- **2003** -- Directive 2003/98/EC on the re-use of the public sector information (PSI directive)
- **2007** - Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)
- **2011** -- Commission Decision (2011/833/EU) on the reuse of Commission documents, established the basis for the former EU Open Data Portal, operated by the Publications Office of the EU (2012-2021).
- **2013** -- Directive 2013/37/EU amending Directive 2003/98/EC on the re-use of public sector information.
- **2019** -- Directive 2019/1024 on open data and the re-use of public sector information recasts Directives 2003/98/EC (PSI directive) and 2013/37/EU setting up a legal framework of minimum requirements for Member States regarding the accessibility of public data resources.

As well as in the form of projects, such as:

- **2012** -- launch of the EU Open Data Portal, central point of access to open data from EU institutions, agencies, and bodies.
- **2014** - launch of the Open Research Data Pilot, which requires certain Horizon 2020 projects to make their research data openly available for reuse
- **2015** -- launch of the European Data Portal publishing data from national, regional, and local open data portals.
- **2016** - launch of the European Open Science Cloud, virtual environment for researchers to store, share and reuse research data.
- **2020** -- Communication on a European strategy for data (COM/2020/66) aiming at creating a single market for data to ensure Europe's global competitiveness.
- **2021** -- The European Data Portal (data from the EU Member States and other European countries) and the EU Open Data Portal (data from the EU institutions, agencies, and bodies) were consolidated into the data.europa.eu portal.

In addition, the European Commission has set guidelines and principles that promote availability and reuse of data generated by the EU-funded research projects.⁵⁶ According to the policy, researchers are, in particular:

- required to create a data management plan that outlines how research data will be collected, processed, and shared during and after the project.
- required to deposit their data in a trusted repository or data center, and provide information about the data, such as metadata and documentation, to make the data findable and accessible.

- encouraged to make their data available for reuse, and to use open standards and formats to make the data more accessible and reusable.
- supposed to be aware of legal and ethical considerations related to their research data, such as data protection regulations or intellectual property rights.

Open research data refers to the data underpinning scientific research results that has no restrictions on its access, enabling anyone to access it.⁵⁷ The indicators of such data are availability of data repositories, policies of research funders and journals and researchers' attitude towards data sharing.

In PERIVALLON, 'all scientific publications related to project results will be available in open access. In addition, all of the research data will be publicly available, which will be done by fully respecting the privacy and fundamental rights. To this end, the data will be as open as possible (to ensure reproducibility and reuse) and as closed as necessary (to allow for protection of results).'⁵⁸

As to the repository, a general-purpose open access repository of research data and journal publications, Zenodo⁵⁹ is used. Additionally, partners have indicated to possibly use other open repositories, such as GitHub, where open source models, algorithms, designs, and software that will be developed by PERIVALLON will be published with open licenses (e.g. Apache).⁶⁰

6. Conclusions

6.1 Summary

This Deliverable D1.2 constitutes the Initial Data Management Plan for the PERIVALLON project. It guides all aspects of data management by outlining the five data sets (Administrative data, Research data, PUC data, Platform data and Technical data) which will be collected and generated during the course of the project.

It described the methodology and standards required for two data sets, Administrative data and Research data, as these data sets were identified to have substantial information development at the stage of the project. Given the early stage of the project, consulting on the remaining data sets would have been speculative and therefore not an effective information collection, resulting in a non-desirable outcome for this DMP. It also provided a detailed analysis on how these data sets will be collected, shared, exploited, re-used, or made accessible for verification, and how they will be curated and preserved. The degree of privacy and confidentiality for each data set is also detailed categorically under the data security. Herein not only the project general data security approach but the individual partners security approach is detailed for each data set.

A separate analysis of the processing of personal data was performed, where the general approach of the project with regard to processing of personal data was described and the technical and organizational measures for securing privacy, confidentiality and the protection of personal data are detailed, and in particular pseudonymisation and anonymisation techniques were identified.

The data management policy is constituted by the data cycle presented in the FAIR handling of each data set, including security and allocation of resources to ensure a positive FAIR handling of data sets.

Additionally, ethical risks and ways to prevent and minimize identified risks with respect of data handling were identified. This mapping served to outline the guidelines to be followed by the PERIVALLON consortium with respect to Open Data initiatives.

Finally yet importantly, the DMP is a living document that will be updated throughout the lifetime of the project, the first planned updates are foreseen in M15 where the remaining data sets will be categorically analyzed.

6.2 Future work and recommendations

The following steps are foreseen to guarantee a continuous and effective data management of PERIVALLON data sets:

1. perform the Update#1, which will mainly focus the FAIR analysis of the PUC, Technical and Platform data sets. Planned for M15.
2. perform an update of the declared and analyzed data sets. Planned for M15.
3. further corroborate the processing of personal data and take the necessary steps to secure a legally and ethical compliant personal data processing. Planned for M12.

This DMP reflects the initial stage of the project and recognizes the efforts of all partners to maintain a proper data management handling of PERIVALLON data sets. Nevertheless, the consortium is encouraged and will provide updates on the handling of such data sets.

6.3 Timetable for DMP Updates
























































































Updates to the DMP are planned as reflected in the timetable below. Further updates, as necessary, will be reported and decided during the lifespan of the project.

Release #	Type of Report	Month	Name	Lead Beneficiary
First release	D1.2.	M6	Initial data management plan	UNIVIE
Update#1	Internal	M15	Update data management plan	UNIVIE
Update#2	Internal	M27	Update data management plan	UNIVIE
Final release	D1.3	M36	Security, Ethics, Legal and Privacy monitoring, Final Data Management Plan & IPR management	UNIVIE

Annexes

Annex 1 - DMP Questionnaire

The Data Management Plan Questionnaire Part 1 and Part 2 are available at: <https://univie-idlaw.limesurvey.net/questionAdministration/listQuestions?surveyid=313575>.

<input type="checkbox"/>	Action	Question ID	Group / Question order	Code	Question	Question type	Group
<input type="checkbox"/>	  	481	1 / 1	G01Q73	Could you please provide the contact details of the responsible person(s) within your organization in charge of the research data?	Long free text	WELCOME
<input type="checkbox"/>	  	392	2 / 1	Q00	What type of research data will you collect and process during the Project? (Deliverables are not included in this question)	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	393	2 / 2	G01Q02	Will you re-use any pre-existing data?	List (Radio)	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	482	2 / 3	G02Q74	What is the origin of the data?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	394	2 / 4	G01Q03	What is the purpose of the data collection/processing?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	395	2 / 5	G01Q04	How will you collect and process the data? In what formats and size?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	396	2 / 6	G01Q05	How will you trace the collected data?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	473	2 / 7	G01Q69	Are you using any other national/funder/sectorial/departmental procedures for data management?	List (Radio)	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	486	2 / 8	G02Q76	Are there any existing procedures on which you will base your approach?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	474	2 / 9	G01Q70	Are there tools or software needed to create/process/visualize the data?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	397	2 / 10	G01Q06	Can you provide any examples (e.g images) of the dataset you are describing herein?	File upload	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	475	2 / 11	G01Q71	Does your institution have data management guidelines, a data protection or security policy that you will follow?	List (Radio)	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	485	2 / 12	G02Q75	What are they?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	398	2 / 13	G01Q07	Do you have any storage and backup strategy for the collected data?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	477	2 / 14	G01Q72	What task/s may tentatively benefit from this dataset?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	399	2 / 15	G01Q08	What PERIVALLON Key Result may tentatively benefit from this dataset?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	472	2 / 16	G01Q68	Is there any additional information you would like to add?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	498	2 / 17	G02Q87	Is the specific research dataset IPR protected (in particular, licensed)?	List (Radio)	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	499	2 / 18	G01Q88	Could you please specify the license / link to the license?	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	543	2 / 19	G01Q90	Is the research dataset annotated?	List (Radio)	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	544	2 / 20	G01Q91	Please detail the main labels included within the dataset.	Long free text	GENERAL INFORMATION ON RESEARCH DATA
<input type="checkbox"/>	  	400	3 / 1	G02Q09	What project and data identifiers will be assigned? What is the dataset denomination?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>	  	401	3 / 2	G01Q10	What standards will be used for documentation and metadata?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>	  	402	3 / 3	G01Q11	Where can you find the dataset? Is there any link/uri available?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>	  	403	3 / 4	G01Q12	What directory and file naming convention will be used?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>	  	404	3 / 5	G01Q13	Will you provide clear version numbers?	List (Radio)	MAKING DATA FINDABLE
<input type="checkbox"/>	  	405	3 / 6	G01Q14	Will search keywords be provided that optimize possibilities for re-use?	List (Radio)	MAKING DATA FINDABLE
<input type="checkbox"/>	  	406	3 / 7	G01Q15	How will potential users find out about your data?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>	  	407	3 / 8	G01Q16	What metadata will be created? How will you capture/create metadata?	Long free text	MAKING DATA FINDABLE

<input type="checkbox"/>				408	3 / 9	G01Q17	Can any metadata be created automatically?	List (Radio)	MAKING DATA FINDABLE
<input type="checkbox"/>				409	4 / 1	G03Q18	Which data produced and/or used in the project will be made openly available?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				410	4 / 2	G01Q19	How will the data be made accessible?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				411	4 / 3	G01Q20	What methods or software tools are needed to access the data?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				412	4 / 4	G01Q21	Where will the data and associated metadata, documentation and code be deposited?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				413	4 / 5	G03Q22	If there are restrictions on use, how will access be provided?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				414	4 / 6	G01Q23	What information needs to be retained in order for the data to be read and interpreted in the project and in the future?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				415	4 / 7	G01Q24	Who checks the correct execution of the access process?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				416	4 / 8	G01Q25	How will the identity of the person accessing the data be ascertained?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				417	5 / 1	G04Q26	Will you follow any data and metadata vocabularies, standards or methodologies to facilitate interoperability?	List (Radio)	MAKING THE DATA INTEROPERABLE
<input type="checkbox"/>				467	5 / 2	G05Q77	Please describe each of them.	Long free text	MAKING THE DATA INTEROPERABLE
<input type="checkbox"/>				418	5 / 3	G01Q27	Will you be using standard vocabularies for all data types?	List (Radio)	MAKING THE DATA INTEROPERABLE
<input type="checkbox"/>				419	5 / 4	G01Q28	Will you provide mappings to more commonly used ontologies?	List (Radio)	MAKING THE DATA INTEROPERABLE
<input type="checkbox"/>				420	6 / 1	G01Q29	Will you permit any data to be re-used?	List (Radio)	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				468	6 / 2	G06Q78	List the categories of data that will be made re-usable or openly accessible.	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				421	6 / 3	G01Q30	Who will be your audience for reuse? Who will use it now? Who will use it later?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				422	6 / 4	G01Q32	When will the data be made available for re-use? Any embargo periods to uphold?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				423	6 / 5	G01Q33	For what period of time will data remain re-usable?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				424	6 / 6	G01Q34	Any restrictions on who can re-use the data, for what purpose and how long?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				425	6 / 7	G04Q34	Will you require the data to be licensed for re-use?	List (Radio)	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				489	6 / 8	G06Q79	How will the data be licensed for re-use?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				426	7 / 1	G06Q35	What are the major risks regarding data security?	Long free text	DATA SECURITY
<input type="checkbox"/>				427	7 / 2	G01Q36	Have you prepared a formal risk assessment addressing each of the major risks to data security?	List (Radio)	DATA SECURITY
<input type="checkbox"/>				490	7 / 3	G07Q80	How you will minimize each risk?	Long free text	DATA SECURITY
<input type="checkbox"/>				428	7 / 4	G01Q37	What steps will be taken to protect privacy, security, confidentiality, intellectual property or other rights?	Long free text	DATA SECURITY
<input type="checkbox"/>				429	7 / 5	G01Q38	How will each category of data be encrypted?	Long free text	DATA SECURITY
<input type="checkbox"/>				430	7 / 6	G01Q39	What other security measures do you anticipate being required for safe data storage and management?	Long free text	DATA SECURITY
<input type="checkbox"/>				431	7 / 7	G01Q40	What, if any, data recovery protocols have you developed?	List (Radio)	DATA SECURITY
<input type="checkbox"/>				432	7 / 8	G01Q41	What, if any, protocols have you developed or outlined for the safe transfer of personal data?	List (Radio)	DATA SECURITY
<input type="checkbox"/>				433	7 / 9	G01Q42	Have you implemented or outlined any procedures to follow in the case of a data breach?	List (Radio)	DATA SECURITY

<input type="checkbox"/>				492	7 / 10	G07Q81	What are they?	Long free text	DATA SECURITY
<input type="checkbox"/>				434	7 / 11	G01Q43	Do you intend any of the data or categories of data to be preserved long-term?	List (Radio)	DATA SECURITY
<input type="checkbox"/>				493	7 / 12	G07Q82	How long should it be retained?	Long free text	DATA SECURITY
<input type="checkbox"/>				435	7 / 13	G01Q44	Where will such data be maintained? Are there data archives that are appropriate for your data (subject-based? or institutional)?	Long free text	DATA SECURITY
<input type="checkbox"/>				436	7 / 14	G01Q45	Who decides what data or what categories of data will be kept and for how long and on what basis?	Long free text	DATA SECURITY
<input type="checkbox"/>				437	7 / 15	G01Q46	Who will maintain the data for the long-term?	Long free text	DATA SECURITY
<input type="checkbox"/>				438	7 / 16	G01Q47	The GDPR requires personal data not be kept longer than necessary for the purpose for which it was stored. What protocol(s) will you put in place to ensure you delete personal data that is no longer required to be stored?	Long free text	DATA SECURITY
<input type="checkbox"/>				439	7 / 17	G01Q48	Will any third parties be involved in or responsible for data management or data storage activities?	List (Radio)	DATA SECURITY
<input type="checkbox"/>				494	7 / 18	G07Q83	Who will it be and what data category or data type each third party will store?	Long free text	DATA SECURITY
<input type="checkbox"/>				455	8 / 1	G08Q64	What are the costs you might encountered for making the Project data FAIR (Findable, Accessible, Interoperable and Reusable)?	Long free text	ALLOCATION OF RESOURCES
<input type="checkbox"/>				456	8 / 2	G01Q65	Will there be any fee(s) associated with storing data in repositories?	List (Radio)	ALLOCATION OF RESOURCES
<input type="checkbox"/>				457	8 / 3	G01Q66	Will you need to pay service fees for any specialist to assist with or maintain repositories or data access?	List (Radio)	ALLOCATION OF RESOURCES
<input type="checkbox"/>				458	8 / 4	G01Q67	What costs do you anticipate that will have to be incurred in the form of time and effort to prepare the data for sharing/preservation?	Long free text	ALLOCATION OF RESOURCES
<input type="checkbox"/>				440	9 / 1	G07Q49	What types of personal data do you intend to collect, generate or process?	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				441	9 / 2	G01Q50	What types of special categories of personal data (sensitive personal data) do you intend to collect, generate or process?	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				442	9 / 3	G01Q51	What, if any, types of personal data about criminal convictions and offences you intend to collect, generate or process?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				443	9 / 4	G01Q52	What kind, if any, of the data subjects will be vulnerable categories of individuals?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				444	9 / 5	G01Q53	Will you be collecting personal or sensitive personal data from people who have not given their explicit consent to participate in the Project?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				495	9 / 6	G08Q84	How will you acquire their consent?	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				445	9 / 7	G01Q54	If the data is personal data, as defined by the GDPR, which of the six Art. 6 GDPR bases will you rely on for the processing of each category of personal data? http://www.privacy-regulation.eu/en/article-6-lawfulness-of-processing-GDPR.htm	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				446	9 / 8	G01Q55	If the data is sensitive personal data, as defined by the GDPR, which of the ten Art. 9 GDPR bases will you rely on for the processing of each category of sensitive data? http://www.privacy-regulation.eu/en/article-9-processing-of-special-categories-of-personal-data-GDPR.htm	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				447	9 / 9	G01Q56	If the data is about criminal convictions and offences, as defined by the GDPR, how do you plan to comply with Art. 10 GDPR? https://www.privacy-regulation.eu/en/article-10-processing-of-personal-data-relating-to-criminal-convictions-and-offences-GDPR.htm	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				448	9 / 10	G01Q57	Have you already gained consent for data preservation and sharing from any data subject(s)?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				449	9 / 11	G01Q58	With which partner are you planning to share personal data, and/or which partner needs your data to perform his/her task?	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				450	9 / 12	G01Q59	How will you protect the identity of Project participants?	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				451	9 / 13	G01Q60	Will you engage in large scale or big data processing?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				452	9 / 14	G01Q61	Will you engage in web-scraping?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				453	9 / 15	G01Q62	Will any entity (including any service provider) outside of the E.U. have access to personal or sensitive data?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>				496	9 / 16	G08Q85	Who, for what purpose and where is each of these entities located?	Long free text	ETHICAL AND DATA PROTECTION ASPECTS










<input type="checkbox"/>	  	454	9 / 17	G01Q63	Do you plan to use pseudonymisation and/or anonymization for any data category?	List (Radio)	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>	  	497	9 / 18	G08Q86	How do you plan to implement this?	Long free text	ETHICAL AND DATA PROTECTION ASPECTS
<input type="checkbox"/>	  	537	10 / 1	G10Q89	Is there any further information that you would like to add on this specific dataset, e.g. Confidentiality, Guidelines for managing all generated data, e.g. Geographic coverage, quotation rules, examples of datasets (e.g. images)?	File upload	BONUS

Figure 4 - Data Management Plan Questionnaire Part 1

<input type="checkbox"/>	Action	Question ID	Group / Question order	Code	Question	Question type	Group
<input type="checkbox"/>		622	1 / 1	G01Q73	Please provide the contact details of the responsible person(s) within your organization in charge of the Administrative dataset?	Long free text	WELCOME
<input type="checkbox"/>		550	2 / 1	Q00	What type of administrative data will you collect, generate and process during the Project? (Deliverables are not included in this question)	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		551	2 / 2	G01Q02	Will you re-use any pre-existing data?	List (Radio)	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		623	2 / 3	G02Q74	What is the origin of the data?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		552	2 / 4	G01Q03	What is the purpose of the data collection/generation/processing?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		553	2 / 5	G01Q04	How will you collect and process the data? In what formats and sizes?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		554	2 / 6	G01Q05	How will you trace the collected data?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		619	2 / 7	G01Q70	Are there tools or software needed to create/process/visualize the data?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		556	2 / 8	G01Q07	Do you have any storage and backup strategy for the collected data?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		621	2 / 9	G01Q72	What task/s may tentatively benefit from this dataset?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		617	2 / 10	G01Q68	Is there any additional information you would like to add?	List (Radio)	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		688	2 / 11	G01Q53	Please describe it below.	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		636	2 / 12	G02Q87	Is this dataset IPR protected (in particular, licensed)?	List (Radio)	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		637	2 / 13	G01Q88	Please specify the license / link to the license?	Long free text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		656	2 / 14	G03Q52	Will you process personal data within administrative dataset?	List (Radio)	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		720	2 / 15	G02Q54	What type of personal data?	Multiple short text	GENERAL INFORMATION ON TECHNICAL, PLATFORM AND ADMINISTRATIVE DATA
<input type="checkbox"/>		559	3 / 1	G01Q10	What standards will be used for documentation and metadata?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>		560	3 / 2	G01Q11	Where can you find the dataset? Is there any link/url available?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>		561	3 / 3	G01Q12	What directory and file naming convention will be used?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>		562	3 / 4	G01Q13	Will you provide clear version numbers?	List (Radio)	MAKING DATA FINDABLE
<input type="checkbox"/>		563	3 / 5	G01Q14	Will search keywords be provided that optimize possibilities for re-use?	List (Radio)	MAKING DATA FINDABLE

<input type="checkbox"/>				564	3 / 6	G01Q15	How will potential users find out about your data?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>				565	3 / 7	G01Q16	What metadata will be created? How will you capture/create metadata?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>				558	3 / 8	G02Q09	What project and data identifiers will be assigned? What is the dataset denomination?	Long free text	MAKING DATA FINDABLE
<input type="checkbox"/>				566	3 / 9	G01Q17	Can any metadata be created automatically?	List (Radio)	MAKING DATA FINDABLE
<input type="checkbox"/>				567	4 / 1	G03Q18	Which data produced and/or used in the project will be made openly available?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				568	4 / 2	G01Q19	How will the data be made accessible?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				569	4 / 3	G01Q20	What methods or software tools are needed to access the data?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				570	4 / 4	G01Q21	Where will the data and associated metadata, documentation and code be deposited?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				571	4 / 5	G03Q22	If there are restrictions on use, how will access be provided?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				572	4 / 6	G01Q23	What information needs to be retained in order for the data to be read and interpreted in the project and in the future?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				573	4 / 7	G01Q24	Who checks the correct execution of the access process?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				574	4 / 8	G01Q25	How will the identity of the person accessing the data be ascertained?	Long free text	MAKING DATA ACCESSIBLE
<input type="checkbox"/>				575	5 / 1	G04Q26	Will you follow any data and metadata vocabularies, standards or methodologies to facilitate interoperability?	List (Radio)	MAKING THE DATA INTEROPERABLE
<input type="checkbox"/>				626	5 / 2	G05Q77	Please describe each of them.	Long free text	MAKING THE DATA INTEROPERABLE
<input type="checkbox"/>				576	5 / 3	G01Q27	Will you be using standard vocabularies for all data types?	List (Radio)	MAKING THE DATA INTEROPERABLE
<input type="checkbox"/>				578	6 / 1	G01Q29	Will you permit any data to be re-used?	List (Radio)	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				627	6 / 2	G06Q78	List the categories of data that will be made re-usable or openly accessible.	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				579	6 / 3	G01Q30	Who will be your audience for reuse? Who will use it now? Who will use it later?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				580	6 / 4	G01Q32	When will the data be made available for re-use? Any embargo periods to uphold?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				581	6 / 5	G01Q33	For what period of time will data remain re-usable?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				583	6 / 6	G04Q34	Will you require the data to be licensed for re-use?	List (Radio)	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				628	6 / 7	G06Q79	How will the data be licensed for re-use?	Long free text	MAKING HE DATA RE-USABLE
<input type="checkbox"/>				638	7 / 1	G10Q89	Is there any further information that you would like to add on this specific dataset, e.g. Confidentiality, Guidelines for managing all generated data, e.g. Geographic coverage, quotation rules, examples of datasets (e.g. images)?	File upload	BONUS

Figure 5 - Data Management Plan Questionnaire Part 2

Annex 2 – Data flows

This overview below should assist partners, in particular task leaders, to quickly identify the data flow between the different Research data types within the tasks.

Research data types contributing to Task					
T/DT	DT 1	DT 2	DT 3	DT 4	DT 5
T1.4	X				
T2.1	X			X	
T2.2	X	X		X	
T2.3		X		X	
T2.4				X	
T3.1	X	X			
T3.2		X		X	
T3.3		X		X	
T3.4		X		X	
T4.1	X				
T4.3			X		
T4.4		X	X		
T4.5					X
T5.1	X	X	X		
T5.2	X			X	
T5.3	X	X	X	X	
T5.4	X				
T5.5	X	X	X		
T5.6		X	X		

T6.1	X			X	
T6.2	X			X	
T6.3	X	X (at different capacity)	a	X	
T6.4				X	
T7.1				X	
T7.3	X			X	
T7.4				X	

Figure 6 - Overview of tasks and their respective contributing Research data types

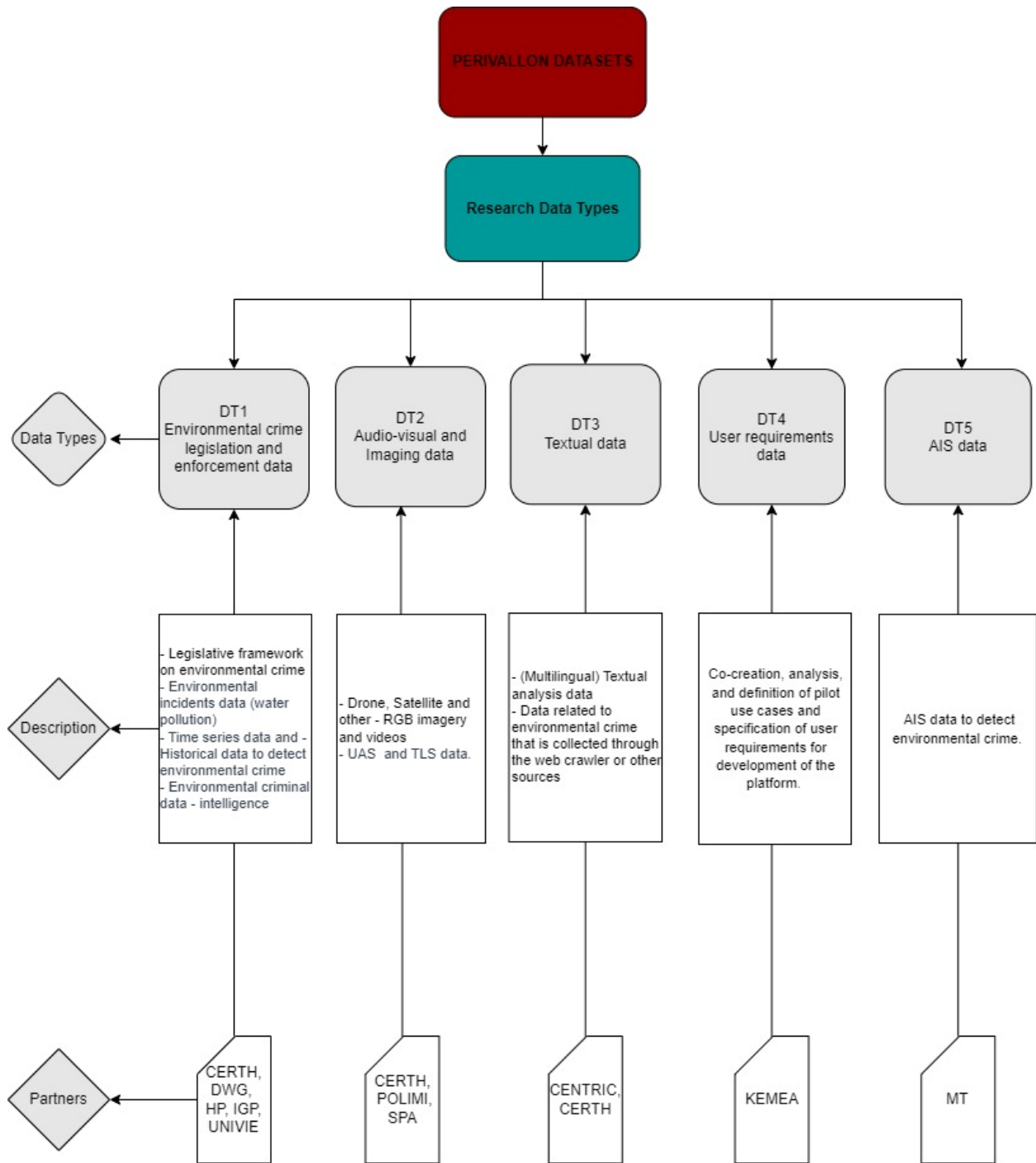


Figure 7 - PERIVALLON Research data types

Annex 3 – Informed consent form template

PERIVALLON

Informed Consent Form for processing your personal data in [...]

(Please read the PERIVALLON Information Sheet before completing this form)

Controller: [-----]

Identification Number: [-----]

Please mark with X
your selected option

YES NO

- | | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|--------------------------|
| 1. I have read the PERIVALLON Information Sheet for this research activity and have had details of it explained to me. | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. My questions about the research activity have been answered to my satisfaction and I understand that I may ask further questions at any point. | <input type="checkbox"/> | <input type="checkbox"/> |
| 3. I understand that I am free to withdraw from [the study] at any time without giving a reason for my withdrawal or to decline to answer any particular questions in the study without any consequences to my future treatment by the researcher. | <input type="checkbox"/> | <input type="checkbox"/> |
| 4. I agree to provide information to the researchers under the conditions of confidentiality set out in the PERIVALLON Information Sheet. | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. I wish to participate in [...] under the conditions set out in the PERIVALLON Information Sheet. | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. I consent to the information collected for the purposes of this research study, once anonymised (so that I cannot be identified), to be used for any other research purposes related to the PERIVALLON project. | <input type="checkbox"/> | <input type="checkbox"/> |

Name of Participant

Date

Signature

Name of person taking consent

Date

Signature

To be signed and dated in presence of the participant

Copies: The signing party should receive a copy of the signed and dated participant consent form, the information sheet and any other written information provided to the party. A copy of the signed and dated consent form should be placed in the project's main record (e.g. a site file), which must be kept in a secure location.

Annex 4 – Information sheet template

PERIVALLON

Information Sheet for engaging in a PERIVALLON related activity and
Data Processing Information Sheet

PART A - INFORMATION SHEET FOR ENGAGING IN A PERIVALLON RELATED ACTIVITY []

1. Information on the PERIVALLON activity

- Responsible contact person at PERIVALLON Consortium: [---]
- Scheduled date: [---]
- Date of Information Sheet: [---]

2. Project Information

- **Project Title:** Protecting the EuROpean terrItory from organised enVironmentAl crime through inteLLigent threat detectiON tools
- **Grant Agreement Number:** 101073952
- **Funding Programme:** HORIZON-CL3-2021-FCT-01-09

3. Project Description

PERIVALLON aims to provide an improved and comprehensive intelligence picture of organised environmental crime and develop effective and efficient tools and solutions for detecting and preventing such types of criminal activities and for assessing their environmental impact based on geospatial intelligence, remote sensing, scanning, online monitoring, analysis, correlation, risk assessment, and predictive analytics technologies, by leveraging the latest advancements in Artificial Intelligence (AI) in the fields of computer vision and multimodal analytics. As a result, enhanced investigation processes and methodologies will be derived through the capabilities provided by the developed tools and solutions, and the insights obtained through the proposed Environmental Crime Observatory.

The capacity of end users (including Police Authorities and Border Guards) will also be improved and will enable them to tackle such criminal activities in an effective manner based on advanced tools and solutions and also on the innovative training curricula developed using physical and/or digital twins of relevant environmental crime scenarios. Moreover, improved international cooperation will be facilitated through improved data sharing enabled by blockchain technologies, while improved regulation shaping and tuning will be supported through relevant policy recommendations.

PERIVALLON will be validated in field tests and demonstrations in four operational use cases. Extensive training, hands-on experience, joint exercises, and training material will boost the uptake of PERIVALLON tools and technologies. With a Consortium 5 Police and Border Guard Authorities, 3 authorities related to environmental protection, 6 Research/Academic institutions, 8 industry partners (including seven SMEs), one EU Agency, and one Foundation, PERIVALLON delivers a strong representation of the challenges, requirements and tools to meet its objectives.

4. Purpose of this Information Sheet

The purpose of this Information Sheet is to [---] and as a research participant you will take part in a research activity fulfilling this purpose.

5. Why have you as a research participant been asked to take part?

[---]

6. What will you be required to do? [e.g. talk about experiences, requirements, validate the project results]

[---]

7. Where will the respective activity take place?

[---]

8. How often will you have to take part, and for how long? [E.g. initial interview]

[---]

9. How long is the whole activity likely to last?

[---]

10. When will you have the opportunity to discuss your participation? [Debriefing]

After receiving this information sheet, you will have reasonable time to go through it, ask any related questions and raise any potential concerns. On the other hand, the responsible contact person at PERIVALLON is responsible for making this information clear, as well as for providing all the required verbal explanations and clarifications and eliminating any concern and misunderstanding.

11. Who will be responsible for all of the information when this activity is over?

[---]

12. Who will have access to it?

[The PERIVALLON project will have access to the data you provided unless the project is required to share your information with the European Commission. You will be notified in case there is a need for sharing to external partners. Personal data will be by default anonymized and a research participation number will be attributed; any deviation on the processing of personal data is otherwise stated below in the section Data Processing Information Sheet]

13. What will happen to the information when this activity is over? [How long will raw data be kept for? Will it be passed onto other people or used in other research activities?]

[---]

14. How will the project Consortium use what they find out? [Reports, publications, presentations]

[---]

15. Will anyone be able to connect you with what is recorded and reported? [Statement of confidentiality, details of coding system to protect identity]

[---]

16. How can the participant find out about the results of the activity?

[---]

17. What if you do not wish to take part?

[—]

18. What if the participant changes his/her mind during the study?

[—]

19. Details of contact persons

Role in the project	Name	Organization	Contact details
Project Manager	Mr. Eduardo Villamor Medina	ETRA	evillamor.etraid@grupoetra.com
Responsible Contact person	[---]	[---]	[---]
Ethics manager	Ms. Mariana Risetto	UNIVIE	mariana.risetto@univie.ac.at

20. Details on any foreseeable risks, discomfort or disadvantages

[describe any risk that you may encounter during the performance of your activity]

21. Details on the procedure of handling incidental findings

- *[describe the anticipated incidental findings that may arise],*
- *[inform the process by which incidental findings will be evaluated] and*
- *[inform the circumstances under which participants will be communicated]*

Should you have any further question, please address them to your responsible contact person at PERIVALLON

PART B - DATA PROCESSING INFORMATION SHEET

Following the GDPR, the following information is provided for the purposes of this research activity:

1. What type of personal data will be collected?

[---]

2. Why personal data will be processed?

[---]

3. Who is the data controller?

[---]

4. Information that the collected personal data is processed on the basis of consent.

[---]

5. Is the data going to be shared and with whom?

[---]

6. Is the data going to be transferred outside the EU, to whom and upon which legal basis?

[---]

7. How long the personal data will be kept?

[---]

8. How to get in contact with the data controller and the data protection officer?

[---]

9. Exhaustive listing of the data subject rights.

[---]

10. Is automated decision-making, including profiling, part of the data processing, and if so, what consequences are to be expected?

[---]

References

- ¹ European Commission (2016). H2020 Programme Guidelines on FAIR Data Management in Horizon 2020. [PDF Version 3.0.]. EC, page 6 – 10. Available at: https://ec.europa.eu/research/participants/data/ref/h2020/grantsmanual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf. [Accessed 16 Feb. 2023].
- ² Ec.europa.eu. Data management. [online] Available at: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.html. [Accessed 16 Feb. 2023].
- ³ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance).
- ⁴ PERIVALLON - DATA MANAGEMENT PLAN QUESTIONNAIRE PART 1: <https://univie-idlaw.limesurvey.net/313575?lang=en>, and PERIVALLON - DATA MANAGEMENT PLAN QUESTIONNAIRE PART 2: <https://univie-idlaw.limesurvey.net/712552?lang=en>.
- ⁵ In particular, PERIVALLON GA Part B p.24
- ⁶ In particular, PERIVALLON CA Art.11(5)
- ⁷ Sources e.g. <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004525>, <https://datamanagement.univie.ac.at/en/research-data-management/data-management-plans/>.
- ⁸ Data Governance Regulation, Open Data Directive, Regulation on a framework for the free flow of non-personal data in the European Union, General Data Protection Regulation.
- ⁹ Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance).
- ¹⁰ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on harmonised rules on fair access to and use of data.
- ¹¹ Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union (Text with EEA relevance).
- ¹² Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union (Text with EEA relevance).
- ¹³ Crossborderdataforum, (2023). What is 'data'? Definitions in international legal instruments on data protection, cross-border access to data & electronic evidence. [online] Available at: <https://www.crossborderdataforum.org/what-is-data-definitions-in-international-legal-instruments-on-data-protection-cross-border-access-to-data-electronic-evidence/> [Accessed 17 February 2023].
- ¹⁴ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information
- ¹⁵ Law insider. (2023). Publicly available data definition. [online] Available at: <https://www.lawinsider.com/dictionary/publicly-available-data#:~:text=publicly%20available%20data%20means%20any,under%20standard%20terms%20and%20conditions> [Accessed 17 February 2023].

¹⁶ IAPP. (2019). Publicly available data under the GDPR: Main considerations. [online] Available at: <https://iapp.org/news/a/publicly-available-data-under-gdpr-main-considerations/> [Accessed 17 February 2023].

¹⁷ European Commission, 'Open Access & Data Management - H2020 Online Manual' (ec.europa.eu) https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm.

¹⁸ PERIVALLON GA, Part B, Subsection 1.2.7, p.24

¹⁹ Open Data Foundation, 'The Open Data Foundation' (www.opendatafoundation.org) <http://www.opendatafoundation.org/>.

²⁰ PERIVALLON GA, Part B, p.6.

²¹ See Question Nr. 36 in PERIVALLON - DATA MANAGEMENT PLAN QUESTIONNAIRE PART 1.

²² Data security measures of each partner are assessed based on partners' answers provided in the DMP questionnaire.

²³ Article 32(1) GDPR.

²⁴ European Union, *Charter of Fundamental Rights of the European Union*, 26 October 2012, 2012/C 326/02, available at: <https://www.refworld.org/docid/3ae6b3b70.html>

²⁵ European Union, Consolidated version of the Treaty on the Functioning of the European Union, 26 October 2012, OJ L 326/47-326/390; 26.10.2012, available at: <https://www.refworld.org/docid/52303e8d4.html> [accessed 4 April 2023]

²⁶ Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)

²⁷ Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA

²⁸ PERIVALLON GA Part B p. 45.

²⁹ Article 4(1) GDPR.

³⁰ Art. 29 Data Protection Working Party (2014), Opinion 05/2014 on Anonymisation Techniques. Available at: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf [Accessed 27 March 2023].

³¹ Article 5(1)b GDPR.

³² PERIVALLON GA Part B Section 4.2 p. 42.

³³ Partners outside the territorial scope of GDPR: CENTRIC (UK), TAMAR (Israel) and IGP (Moldova).

³⁴ This information is explicitly required by p. 45 PART B PERIVALLON GA.

³⁵ Replies to Section – LEGAL AND ETHICAL ASPECTS- PERIVALLON - DATA MANAGEMENT PLAN QUESTIONNAIRE PART 1.

³⁶ PERIVALLON GA PART B p. 45.

³⁷ PERIVALLON GA PART A p. 8.

³⁸ Article 4(7) GDPR: ‘controller’ means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law.

³⁹ Article 4(8) GDPR: ‘processor’ means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.

⁴⁰ PERIVALLON CA Article 11(5).

⁴¹ PERIVALLON CA Article 11(5).

⁴² Article 45 GDPR.

⁴³ Commission Decision of 31 January 2011 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequate protection of personal data by the State of Israel with regard to automated processing of personal data (notified under document C(2011) 332) Text with EEA relevance.

⁴⁴ COMMISSION IMPLEMENTING DECISION of 28.6.2021 pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate protection of personal data by the United Kingdom.

⁴⁵ Detailed information on the assessing the term “likely to result in a high risk” can be found in the ARTICLE 29 DATA PROTECTION WORKING PARTY (2017) Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is “likely to result in a high risk” for the purposes of Regulation 2016/679.

⁴⁶ European Commission (2021) Ethics and data Protection. Available at: [ethics-and-data-protection_he_en.pdf \(europa.eu\)](#). [Accessed 16 February 2023].

⁴⁷ PERIVALLON GA PART A p. 32.

⁴⁸ Idem 47.

⁴⁹ Idem 47.

⁵⁰ Idem 47, p. 7 .

⁵¹ Idem 47, p. 17.

⁵² PERIVALLON GA PART B p. 30.

⁵³ ALLEA. (2017). The European Code of Conduct for Research Integrity. Available at: <https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf>; also: European Commission. (2015). *The European Charter for Researchers*. Retrieved on 3.4.2023 from https://euraxess.ec.europa.eu/sites/default/files/am509774cee_en_e4.pdf. [Accessed 06 April 2023]; Code of ethics for Researchers (WEF). Available at: <https://widgets.weforum.org/coe/index.html#code> [Accessed 07 April 2023].; good practices: <https://www.eui.eu/documents/servicesadmin/deanofstudies/researchethics/guide-data-protection-research.pdf> [Accessed 07 April 2023].

⁵⁴ Idem, 1. Principles.

⁵⁵ [Documentation of data.europa.eu \(DEU\) \(dataeuropa.gitlab.io\)](https://data.europa.eu/DEU) [Accessed 06 April 2023].

⁵⁶ European Commission (2017) Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Available at: [h2020-hi-erc-oa-guide_en.pdf \(europa.eu\)](https://ec.europa.eu/research/data/h2020-hi-erc-oa-guide_en.pdf). [Accessed 06 April 2023].

⁵⁷ European Commission (2017) Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020. Available at: [h2020-hi-erc-oa-guide_en.pdf \(europa.eu\)](https://ec.europa.eu/research/data/h2020-hi-erc-oa-guide_en.pdf). [Accessed 06 April 2023].

⁵⁸ PERIVALLON GA PART B p. 23.

⁵⁹ Zenodo. Available at <https://zenodo.org/>. [Accessed 06 April 2023].

⁶⁰ PERIVALLON GA PART B p. 23.