ENVRI-Hub
NEXT

D1.2 Data Management Plan

Status: Under EC Review
Dissemination Level: Public

Funded by
the European Union

Abstract	
Key Words	Data Management, FAIR, GDPR, metadata, re-usable data
This first Data Management Plan (DMP) introduces a report that specifies how research data will be collected, processed, monitored, and catalogued, following the FAIR principles. This deliverable is viewed as a living report that will advance throughout the life of the project.	

Revision History			
Version	**Date**	**Description**	**Author/Reviewer**
V 0.1	11/06/2024	ToC	Sjomara Specht (EGI.eu)
V 0.2	13/06/2024	Initial first input	Sjomara Specht (EGI.eu)
V0.3	04/07/2024	Ethical Aspects	Giuseppe La Rocca (EGI)
V0.3	12/07/2024	Finalised and ready for review	Sjomara Specht (EGI.eu)
V0.4	15/07/2024	External review	Ulrich Bundke (FZJ) Zhiming Zhao (UvA)
V0.5	17/07/2024	Incorporated feedback	Sjomara Specht (EGI.eu)
V0.6	19/07/2024	Approval of Board	Development Steering Board (DSB)
V 1.0	**19/07/2024**	**Final**	

Document Description			
D1.2 Data Management Plan			
Work Package Number 1			
Document Type	Deliverable		
Document Status	Under EC Review	**Version**	1
Dissemination Level	Public		
Copyright Status	 This material by Parties of the ENVRI-Hub NEXT Consortium is licensed under a **Creative Commons Attribution 4.0 International License.**		
Lead partner	EGI Foundation		
Document Link	**https://documents.egi.eu/document/4025**		
DOI	**https://doi.org/10.5281/zenodo.12784093**		
Author(s)	● Sjomara Specht (EGI)		
Reviewers	● Ulrich Bundke (FZJ) ● Zhiming Zhao (UvA)		
Moderated by:	● Sjomara Specht (EGI)		
Approved by:	Development Steering Board (DSB)		

Terminology / Acronyms	
Term/Acronym	**Definition**
API	Application Programming Interface
DMP	Data Management Plan
DocDB	Document Database
DOI	Document Object Identifier
DPO	Data Protection Officer
DSB	Development Steering Board
EB	Ethical Board
ECV	Essential Climate Variables
EOSC	European Open Science Cloud
ENVRI	Environmental Research Infrastructures
FAIR	Findable, Accessible, Interoperable, and Reusable
FDP	FAIR Data Point
GDPR	EU General Data Protection Regulation
GUI	Graphical User Interface
ORE	Open Research Europe
PMO	Project Management Office
RI	Research Infrastructure
SSO	Single Sign On
URL	Uniform resource locator
WPs	Work Packages

Table of Contents

Table of Tables

Executive summary

The ENVRI-Hub NEXT D.1.2 Data Management Plan (DMP) presents a comprehensive strategy for managing research data throughout the project's lifecycle. The purpose of this deliverable is to ensure that all data generated within the project is managed by the FAIR (Findable, Accessible, Interoperable, and Reusable) principles, promoting transparency, accessibility, and reusability.

Key points of the deliverable include:

- **Purpose and Scope**: The DMP aims to standardise data management practices across the project, ensuring that all data is collected, processed, and stored in a manner that adheres to the highest standards of scientific integrity and openness.
- **Data Management Strategies**: Detailed plans for making data findable, accessible, interoperable, and reusable are presented. This includes the use of established metadata standards, open-access repositories, and interoperable data formats.
- **Resource Allocation**: Clear delineation of roles and responsibilities related to data management, ensuring that all project partners understand their obligations and contributions.
- **Data Security and Ethical Considerations**: Robust measures are in place to protect data integrity and security, along with comprehensive ethical guidelines to address any potential issues related to data use and sharing.
- **Sustainability and Long-term Preservation**: Strategies for the long-term preservation of data, including the use of trusted repositories and adherence to open-access policies, ensuring that data remains accessible and usable beyond the project's duration.

The deliverable concludes with recommendations for ongoing data management practices and future updates to the DMP, emphasising the importance of continuous improvement and adaptation to new challenges. By following this plan, the ENVRI-Hub NEXT project aims to set a benchmark for data management in environmental research, fostering a culture of openness and collaboration.

1. Introduction

The current initial version -D1.2 Data Management Plan - is composed of preliminary information and frameworks that will be followed. Hence, it is subject to updates in the future upon developments and changes during the project. In principle, the DMP describes the standards that will be used, and how the project's research data will be stored and published for verification and reuse. ENVRI-Hub NEXT aims for full open access to data, following the FAIR principles.

1.1. Purpose and scope of the document

Open Science is at the core of RIs to accelerate the impact of scientific research. ENVRI-Hub NEXT will provide a platform for environmental services, data, knowledge, and training to support and increase collaborations and sharing of information beyond the traditional scientific community. The project will ensure that Open Science practices align with the EU's open science policy and that an Open Science strategy is implemented, followed, and reported.

Data, as described in the DMP, will be FAIR by design and made available according to the principle of "as open as possible" and as early as possible.

1.2. Structure of the document

This document is comprised of the following chapters:

Section 1: presents an introduction to the project and the document.
Section 2: presents the purpose of data collection, type and format, and origin of the data.
Section 3: outlines ENVRI-Hub NEXT FAIR data strategies
Section 4: briefly describes the allocation of resources.
Sections 5, **6** and **7**: outline data security, ethical and other issues.
Section 8: concludes this deliverable.

2. Data Summary

This section provides an overview of the potential data assets generated by the ENVRI-Hub NEXT project, including the types and origins of data involved. The data generated will adhere to the FAIR principles, ensuring it is findable, accessible, interoperable, and reusable.

2.1. Potential Data Assets Generated by the Project

The ENVRI-Hub NEXT project will generate a wide range of data assets, including:

- **Research Data**: Raw and processed data from environmental research activities.
- **Software**: Software of ENVRI-Hub, Tools, software services, scripts, workflows and computational notebooks developed for data analysis, processing, and visualization.
- **Documentation**: Manuals, guidelines, and training materials to support data use and software applications.
- **Metadata**: Descriptive information about the data assets to ensure discoverability and proper usage.
- **Publications**: Research papers and reports detailing project findings and methodologies.

2.2. Types of Data

The project will manage various types of data, including but not limited to:

- **Textual Data**: Documents, reports, and publications in formats such as .docx, .pdf, .txt, and .rtf.
- **Numerical Data**: Datasets in formats like .csv, .xls/.xlsx, and .netcdf.
- **Code and Software**: Programming scripts, software packages and APIs in formats such as .py, .R, .sh, .c/c++, .ipynb, .java and .js.
- **Multimedia Data**: Images, videos, and audio files in formats like .jpg/.jpeg, .png, .gif, .mp4, and .wav.
- **Metadata Records**: Information about the data in formats like .xml, .json, .rdf, and .ttl.
- **Presentations:** Slides and visual aids in formats such as .ppt/.pptx and .pdf.

These formats ensure that the project outputs are accessible, interoperable, and reusable, adhering to the FAIR principles. The selection of specific formats is guided by community standards and the nature of the data or document.

2.3. Existing data/software and newly generated project outputs

The project builds on both existing and newly generated data and software outputs:

Existing Data/Software

ENVRI-HUB-NEXT is not responsible for external data/software. Thus, this DMP does not cover external datasets/software listed in the following. Here individual DMPs issued by RIs and/or Projects like ENVRI FAIR are in place.

- **Content in the current ENVRI-Hub**: The pre-existing catalogue and knowledgebase include datasets, software, and services that have been curated according to FAIR principles.
- **Research Infrastructure Databases**: Databases from RIs such as **ICOS ERIC**, **LifeWatch ERIC**, and **EPOS ERIC**, among others, provide a rich source of reusable data.
- **Software**: Existing ENVRI-Hub prototype and software tools for data processing, analysis, and visualization, such as those hosted on GitLab, will be utilized and extended.
- **Metadata Standards**: Established metadata standards and schemas from prior projects and RIs.

Newly Generated Project Outputs

- **New Datasets**: Generated from ongoing research activities within the project, these datasets will be documented and published following the FAIR principles.
- **Enhanced/New Software**: Development of new features of ENVRI-Hub, software tools and significant enhancements to existing tools will support advanced data processing, interoperability, and visualisation needs.
- **Metadata Records**: Creation of comprehensive metadata records for all new datasets and software outputs willensure discoverability and reusability.
- **Documentation and Training Materials**: New documentation and training materials willsupport the use and understanding of the data and tools generated by the project.
- **Publications:** Research papers and Project Deliverables will detail the methodologies, findings, and advancements achieved through the project activities.

3. FAIR Data

The data generated will be managed according to the FAIRness in line with Open Science principles. To ensure that data and project outputs are easily discoverable, they will be made accessible through the ENVRI-Hub NEXT Community in Zenodo and other reliable open repositories. Access will be provided according to open-access policies. Interoperability will be promoted by including all essential information, data, and metadata along with the reports. Lastly, to ensure reusability, data and project outputs will be shared under permissive licenses, such as CC BY, whenever feasible.

3.1. Making data findable, including provisions of metadata

Data will be designed to be FAIR and will be made available as openly and as early as possible. Data will only be restricted when required by regulatory or legal constraints or due to the owner's legitimate interests. Raw data from consultations (like user surveys and questionnaires) will be aggregated and anonymised to ensure it can be reused for future purposes.

3.1.1. Metadata

The metadata required by data repositories will be used as outlined in **table 1**.
For documents, ENVRI-Hub NEXT has defined a standard set of metadata to ensure consistent and comprehensive documentation of data assets as shown in **table 2**.

Table 1 - Repository metadata

Element	Definition
Title	A name given to the source.
Upload type	e.g., dataset, workflow, project deliverables
Abstract	Describing the document contents and main conclusions.
Submitter	The person submitting the document to the repository
Authors	The people involved in contributing significant portions of the data.
DOI	Provided by the resource
Publication date	The date of first publication.
Version	The version number generated by the document repository for the repository identifier. Versioning rule: · +0.1 – a new version of the draft · +1.0 – a new version of the approved document
Language	A language of the intellectual content of the resource.
Keywords	A list of words that will support the search within the repository service
Communities	A specific community in which the upload will appear.
License	Specifies the copyright status under which the upload will be licensed.
Modify	The groups can modify the document. The 'EGI office' SSO group must be always marked.

Table 2 - Document metadata

Element	Definition
Title	A name given to the source. For milestones and deliverables as described in the Description of Work.
Lead Partner	The recognised short name of the lead partner within the ENVRI-Hub NEXT project
Authors	The people involved in writing significant portions of the document.
Reviewers	The people involved in reviewing the document.
Copyright status	Public material is licensed under a **Creative Commons Attribution 4.0 International License.** Confidential material may not be reproduced, distributed, or disclosed in whole or in part without the prior written consent of the ENVRI-Hub NEXT consortium
Document Type	e.g., deliverable, report, white paper
Status	· **Draft** - the document is being prepared · **Under EC review** - the document is submitted to the EC portal and has not yet been approved by the European Commission · **Approved by EC** - the document is approved by the European Commission · **Final** - status of the document
Dissemination Level	· **Public** - can be shared without restrictions · **Confidential** - can be shared only with European Commission and project partners
Document link	The URL in the document repository that provides access to the document on **DocDB**.
Digital Object Identifier	An identification number is assigned through a repository service.
Keywords	A list of words that will support the search within the document repositories.
Abstract	Describing the document contents and main conclusions.

3.2. Making data openly accessible

Textual outputs such as reports and presentations, except for administrative documents containing sensitive legal, personal, or financial information, will be made available to the ENVRI-Hub NEXT Community on Zenodo upon release. These outputs will be shared under a permissive licence, such as CC BY, and accompanied by comprehensive metadata under CC05. Tools and code will be extensively documented and preserved through open repositories like GitLab and provided using licences like GPL-3[1] or Apache-2[2].

The metadata will include details like authorship, ownership, provenance, access rights, and persistent identifiers (PIDs), all in alignment with the FAIR principles .

---

[1] https://www.gnu.org/licenses/gpl-3.0.en.html
[2] https://www.apache.org/licenses/LICENSE-2.0

3.2.1. Repositories

All documents, presentations and other materials that form an official output of the project (not just milestones and deliverables) are placed in the EGI own trusted document repository[3] to provide a managed central location for all materials.

In addition, public deliverables and publications will be shared publicly via the **Zenodo platform** to increase the discoverability of the project outputs.

All profiles, specifications, configuration files, software, workflows, and code will be deposited in Zenodo, **Code Ocean**[4] and **Papers with code**[5]. Therefore, ENVRI-Hub NEXT will use DocDB, Zenodo, and GitLab as their standard and main repositories.

3.2.2. Standardised access protocol

All data will be accessible via a uniform resource locator (URL) or Document Object Identifier (DOI). There will be no restrictions on the use of the research outputs, both during and after the end of this project.

The project will frequently interact with the Steering Committee and a board of RI representatives to ensure that outputs are all-inclusive to the ENVRI Community. It will also take into consideration public engagement from end users and citizens through interactive workshops at targeted events, consultations and social media.

3.2.3. Metadata availability

Metadata containing information to enable users to access the data will be openly available and published together with the data, in the same repositories as listed under **Repositories**. There is no time limit on metadata and data availability.

The ENVRI-Hub NEXT project acknowledges the value of documentation for interoperability purposes, increases uptake by different communities, and encourages data owners to document their research data assets. ENVRI-Hub NEXT does not enforce specific provisions on documentation as long as the data asset is hosted on one of the mentioned repositories and properly curated according to the repository's best practices.

Research data itself should not be considered self-documenting and each published asset must be associated with sufficient documentation resources accessible through a public URL. Documentation must be browsable and include hypertext references to facilitate its fruition. Recommended documentation formats include markdown, HTML, and other markup languages. The inclusion of machine-readable documentation such as OpenAPI where applicable is thoroughly encouraged. If a scientific publication is tied to the research data asset, the publication itself should be referenced and/or made available as part of the documentation.

---

[3] https://documents.egi.eu/

[4] https://codeocean.com/

[5] https://paperswithcode.com/

3.3. Making data interoperable

The ENVRI-Hub NEXT consortium acknowledges the importance of data interoperability: if a data asset cannot be compared, merged, or otherwise integrated with other assets, then its publication would bring little or no value to the community. The consortium is therefore committed to supporting and enforcing data interoperability.

3.4. Increase data re-use

The project will implement measures to increase data reuse by providing clear documentation, adhering to FAIR principles, and ensuring data is accessible and interoperable. Regular training and documentation of services, metadata or analytical workflows will be provided and made available to the training platform assisted by training events to maximize the reusability of the data.

Peer-reviewed research papers produced during the project will adhere to the European Commission's Open Access policy for publications. Preprints will be deposited in reputable open repositories, such as Arxiv or institutional archives, whenever feasible. Upon publication, these papers will be self-archived in Zenodo. Ideally, they will also be published in open-access journals or platforms, with Open Research Europe (ORE) as the preferred choice, or in selected Diamond Open Access journals. All articles will be issued with open-access licences that impose minimal restrictions. Both the papers and related materials will be made available in the ENVRI-Hub NEXT Community on Zenodo as needed.

4. Allocation of Resources

Any expenses associated with the collection/production of FAIR data during the ENVRI-Hub NEXT activities are included in the project budget. These expenditures will be required to cover a variety of particular data processing and data management operations, ranging from data collection and documentation to storage and preservation to distribution and re-utilization.

These operations are a component of the Work Packages (WPs) that processes the relevant data, hence the needed effort will be part of the relevant WP.

The expenses of long-term data preservation are minimal, by using the EGI Online Storage and Google Drive platforms. Using Zenodo and GitLab (both free of charge) ensures that costs for long-term preservation of the data are manageable. When applicable, a more accurate cost estimate will be provided at a later stage of the project.

4.1. Data Management responsibilities

Within the ENVRI-Hub NEXT project, the following roles and responsibilities are associated with Data Management, which are defined as follow:

WP leaders are in charge of organizing the data processing and quality assurance that take place inside the Work Package they are leading.

Task Leaders are in charge of the data compiled/produced throughout the operation of the task that they are responsible for. In addition to that, they also make sure that the data are properly prepared to be shared among the partners, and made publicly available, when applicable.

Data users are consortium partners who use, e.g., processing operations on the compiled/produced data.

Quality and Risk Manager monitors and supports the WP leaders, and Task Leaders/Use Case leaders in keeping the DMP confluence pages up to date. In addition, reports the changes and processes via milestones and deliverables as specified in the Grand Agreement.

5. Data Security

Any gathered data will be securely handled throughout the entire duration of the ENVRI-Hub NEXT project, to protect it from loss and unauthorized access. Personal data is only accessible to those who are authorised[6] to access it.

All partners/beneficiaries responsible for processing personal data[7] have the responsibility to ensure that the data remains protected under all necessary security controls (including backup policies and integrity checks[8]) and access controls (identification, authentication, authorization) within their infrastructure. In the unfortunate event of a personal data breach, the project partners will notify without delay their competent national supervisory authorities as well as the data subject(s) that may be affected by the breach. At the same time, they will document any personal data breaches and all related information.

Regarding open data, for security and long-term preservation, ENVRI-Hub NEXT relies on the EGI Document Repository, Zenodo, Google Drive and GitLab platforms.

---

[6] Personal contact data are collected during the EGI SSO (Single Sign On) account creation by the users. After the access is granted, the users can manage their data in autonomy.

[7] Processing, according to Regulation (EU) 2017/679 of the European Parliament (GDPR), means any operation or set of operations which is performed on personal data or on sets of personal data, whether by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure, or destruction

[8] The integrity check is the process of comparing the current state of stored data and/or programs to a previously recorded state to determine any alteration or change.

ENVRI-Hub NEXT - 101131141   www.envri-hub-next.eu

6. Ethical Aspects

From the initial assessment, the project has identified three areas where the management of ethics issues needs to be implemented: AI, Data and Humans. To perform the ethical assessment, the project is committed to using the guidelines and recommendations provided by EC in the **Ethics By Design and Ethics of Use Approaches for Artificial Intelligence** document. The document offers guidance for addressing ethics-related concerns and for demonstrating ethics compliance. Broadly speaking, this guidance covers all research activities involving the development or/and use of artificial intelligence (AI)-based systems or techniques, including robotics.

During the two periodic reports at M18 and M36, ethical assessment of the software and the datasets used by the project will be recorded in the templates provided in **Appendix II** and **III**. From a technical point of view, the ethical assessment of the software used to develop the core services of the ENVRI-Hub will be conducted under the Work Packages Leaders responsibilities (WP7/8, WP9/10, WP11/WP12, and WP13/WP14), while the ethical assessment of the environmental datasets is conducted under the ENVRIs responsibilities (**ACTRIS**, **AnaEE ERIC**, **EPOS ERIC**, **Euro-Argo ERIC**, **IAGOS AISBL**, **ICOS ERIC**, **LifeWatch ERIC**, **eLTER**, and **SeaDataNet**). The implementation of ethics requirements will be monitored by the Ethics Board and the EB.

Requirements emerging from the new horizontal data legislation, such as the European Data Governance Act (DGA) and DA will be discussed with the RI's in the context of the development of the ENVRI-Hub.

The project set up an Ethics Board defined the ToR of this Board and started to appoint members to this Board. The EAs will ensure the project's adherence to overarching EU ethical requirements and conduct a comprehensive analysis of any potential research/project ethical issues, including monitoring the potential misuse of the new AI tools being developed or the data being collected/used.

7. Other

Further details of the procedures on Data Management used in ENVRI-Hub NEXT are described in ENVRI-HUb NEXT D15.1 OEI - Requirements No.1[9].

---

[9] https://documents.egi.eu/document/4058 - access restricted to the ENVRI-Hub NEXT consortium members.

8. Conclusion and next steps

ENVRI-Hub NEXT DMP is a complete data management approach that complies with Horizon Europe recommendations and that aims to make data as findable, accessible, interoperable, and reusable (FAIR) as feasible.

Relying on robust technological solutions and standards, such as the EGI Document Repository, Zenodo, GitLab and Google Drive[10] for the execution of these processes. Additionally, this will ensure that the data created or compiled throughout the ENVRI-Hub NEXT project, including open data and public publications, will be kept and continue to be usable once the project is completed.

The plan is intended to safeguard the analysis of compiled/created data based on the privacy level and to use an alternate sharing methodology relying upon this level. Confidential information or information that raises ethical problems will not be released.

Task 1.3 will implement the DMP with the help of data managers of the RIs under the supervision of the coordinator. All RIs have ample experience in developing DMPs and FAIR management of research data in research projects and their own RI, as well as with curation and preservation issues as the reason for the existence of the RI data centres involved in the project.

Finally, the DMP is built on guaranteeing appropriately informed consent, and protecting each participant's zone of privacy, while adhering to GDPR guidelines

8.1. DMP Change Management

This DMP is considered a living document, and it will be updated during the project to reflect the most recent developments and conclusions. In actuality, the early version of the DMP will be refined at least twice on M20 and M30 of the project.

Ad hoc improvements may also be deployed if deemed necessary. In general, changes need to be fully compliant with EU laws and best practices in research data management.

Updates will be entered in the changelog table that is shown on the confluence page of the concerned DMP. Analogously, the distribution of notifications on updates will be realised via regular meetings (WPs, PMO, DSB).

---

[10] Google is listed in the DPF and thus, covered by EU-U.S. Data Privacy Framework https://www.dataprivacyframework.gov/list. All instances of Jira, GITLAB, Confluence are operated on European Servers and compliant with GDPR

Appendix I: Privacy Policy

With this privacy policy we, The EGI Foundation, inform the project members about which personal data is collected and processed when using EGI Collaboration Tools. Table 3 outlines the complete policy.

Table 3 - EGI Privacy Policy for the ENVRI-Hub NEXT project.

Name of the Service	Collaboration tools for the **ENVRI-Hub NEXT project**
Description of the Service	The EGI Collaboration Tools services for the **ENVRI-Hub NEXT** (hereinafter referred to as: "the service" or "Collaboration Tools") support the **ENVRI-Hub NEXT** project's activities. Personal data is used to provide access to the service with the proper access levels. Personal data is collected as part of the project's activities. This privacy notice describes how we, the EGI Foundation (hereinafter referred to as "we" or "the Data Controller"), collect and process data by which project members can be personally identified ("Personal Data") when the service is used.
Data controller	The EGI Foundation Science Park 140 1098 XG Amsterdam The Netherlands.
Data protection officer	The EGI Foundation Data Protection Officer Science Park 140 1098 XG Amsterdam The Netherlands E-mail: **dpo@egi.eu**
Jurisdiction and supervisory authority	Jurisdiction: NL, The Netherlands EGI Foundation's lead supervisory authority is the Dutch Data Protection Authority. They can be contacted at **https://autoriteitpersoonsgegevens.nl/en/contact-dutch-dpa/contact-us**

Personal data processed	The service may process the following personal data: **Identification data:** - Name - Identification number - E-mail address - Phone number - Address - Bank details - Other: affiliation, IP address **Behavioural data:** - Usage Data - Data on purchase or payment transactions - Working time data - Other: technical logs with timestamps, attendance at meetings **Data allowing conclusions on the personality:** - Other: membership information on groups, roles, and communities **Biographical data:** - CV data **Sociodemographic data:** - Gender
Purpose of the processing of personal data	The purpose of the collection, processing, and use of the personal data mentioned above is: - To provide the service functions, to coordinate and manage the project according to applicable requirements including projects' contracts, guidelines, funding policies, and legal requirements. - To keep evidence for audit needs. - To monitor and maintain service stability, performance, and security.
Legal basis	The legal basis for processing personal data is compliance with a legal obligation or legitimate interests pursued by the controller or by a third party according to Art. 6 (1) (f) General Data Protection Regulation (GDPR).

| **Third parties to whom personal data is disclosed** | Personal data will not be used beyond the original purpose of their acquisition. If forwarding to third parties should be necessary to answer an inquiry or to carry out a service, the consent of the data subject is considered to have been given by entering a contract when using the respective function or service. In particular, the data provided will not be used for advertising purposes.

For the purpose given in this privacy policy, personal data may be passed to the following third parties:

Within the European Union (EU) / European Economic Area (EEA):

● CESNET: resource provider, the sub-contracted data processor of EGI Foundation
● Google Ireland Limited: resource provider, the sub-contracted data processor of EGI Foundation
● EUfin: financial reporting
● Zenodo: public deliverables and publications
● European Commission
● Project's partners
● Individuals responsible for managing projects, Work Packages, and tasks.
● The records of your use and technical log files produced by the service components may be shared for security incident response purposes with other authorised participants in the academic and research-distributed digital infrastructures via secured mechanisms, only for the same purposes and only as far as necessary to provide the incident response capability were doing so is likely to assist in the investigation of suspected misuse of Infrastructure resources.

Outside the EU / EEA:

● Project partners
● Individuals responsible for managing projects, Work Packages, and tasks.

Any data transfer to a third country outside the EU or the EEA only takes place under the conditions contained in Chapter V of the GDPR and in compliance with the provisions of this privacy policy and any related policies adopted by the EGI Federation. |

Your rights	You can exercise the following rights at any time by contacting our Data Protection Officer using the contact details provided in the Data Protection Officer section:
	• Information about the data stored with us and their processing • Correction of incorrect personal data • Deletion of the data stored by us • Restriction of data processing, if we are not yet allowed to delete the data due to legal obligations • Objection to the processing of the data by us • Data portability
	Project members can complain at any time to the supervisory data protection authority (DPA). The responsible DPA depends on the country and state of residence, the project member's workplace, or the presumed violation. A list of the supervisory authorities with addresses can be found at **https://edpb.europa.eu/about-edpb/board/members_en**.
	You can contact the EGI Foundation's lead supervising authority using the contact details provided in the Jurisdiction and Supervisory Authority section.
Data retention and deletion	As per EC Grant agreements' requirements, data should be kept for at least 5 years after the end of the project.
	The data are deleted or anonymised as soon as retention periods have passed, and the data are not required anymore for any of the purposes listed above.
	The records of the project members' use and technical log files produced by the service components will be deleted or anonymised after, at most, 18 months as documented in **EGI-doc 2732: Policy on the Processing of Personal Data**.
Security	We take appropriate technical and organisational measures to ensure data security and protection against accidental or unlawful destruction, accidental loss, alteration, unauthorised disclosure, or access.
	A comprehensive overview of the technical and organisational measures taken by EGI Foundation can be found in **EGI Document 3737: EGI Foundation Technical and Organisational Measures (TOM)**
Data Protection Code of Conduct	EGI Foundation is conforming to the GEANT Code of Conduct and project members' data will be processed by the **Code of Conduct for Service Providers** and the **EGI-doc-2732-v3: Policy on the Processing of Personal Data**.
	This policy is based on **AARC Policy development kit** (licenced under **CC BY-NC-SA 4.0**)

Appendix II: Specification of objectives against Ethical Requirements

Table 4 - Respect for Human Agency

Specification of Objectives against Ethical Requirements	Project relevance	Justification
Respect for Human Agency		
End-users and others affected by the AI system are not deprived of abilities to make all decisions about their own lives, have basic freedoms taken away from them.	Applicable/ Not applicable	
End-users and others affected by the AI system are not subordinated, coerced, deceived, manipulated, objectified, or dehumanised, nor is attachment or addiction to the system and its operations being stimulated.	Applicable/ Not applicable	
The system does not autonomously make decisions about vital issues that are normally decided by humans by means of free personal choices or collective deliberations or similarly significantly affect individuals.	Applicable/ Not applicable	
The system is designed in a way that gives system operators and, as much as possible, end-users the ability to control, direct and intervene in basic operations of the system (when relevant)	Applicable/ Not applicable	

Table 5 - Privacy & Data Governance

Specification of Objectives against Ethical Requirements	Project relevance	Justification
Privacy & Data Governance		
The system processes data in line with the requirements for lawfulness, fairness and transparency set in the national and EU data protection legal framework and the reasonable expectations of the data subjects.	Applicable/ Not applicable	

Technical and organisational measures are in place to safeguard the rights of data subjects (through measures such as anonymisation, pseudonymisation, encryption, and aggregation).	Applicable/ Not applicable	
There are security measures in place to prevent data breaches and leakages (such as mechanisms for logging data access and data modification).	Applicable/ Not applicable	

Table 6 – Fairness

Specification of Objectives against Ethical Requirements	Project relevance	Justification
Fairness		
The system is designed to avoid algorithmic bias, in input data, modelling and algorithm design. The system is designed to avoid historical and selection bias in data collection, representation and measurement bias in algorithmic training, aggregation and evaluation bias in modelling and automation bias in deployment	Applicable/ Not applicable	
The system is designed so that it can be used by different groups of end-users with different abilities (whenever possible/relevant)	Applicable/ Not applicable	
The system does not have negative social impacts on relevant groups, including impacts other than those resulting from algorithmic bias or lack of Universal accessibility	Applicable/ Not applicable	

Table 7 – Individual, and Social and Environmental Well–being

Specification of Objectives against Ethical Requirements	Project relevance	Justification
Individual, and Social and Environmental Well-being		
The AI system takes the welfare of all stakeholders into account and does not unduly or unfairly reduce/undermine their well-being	Applicable/ Not applicable	
The AI system is mindful of the principles of environmental sustainability, both regarding the system itself and the supply chain to which it connects (when relevant)	Applicable/ Not applicable	
The AI system does not have the potential to negatively impact the quality of communication, social interaction, information, democratic processes, and social relations (when relevant)	Applicable/ Not applicable	
The system does not reduce safety and integrity in the workplace and complies with the relevant health and safety and employment regulations	Applicable/ Not applicable	

Table 8 - Transparency

Specification of Objectives against Ethical Requirements	Project relevance	Justification
Transparency		
The end-users are aware that they are interacting with an AI system	Applicable/ Not applicable	
The purpose, capabilities, limitations, benefits, and risks of the AI system and of the decisions conveyed are openly communicated to and understood by end-users and other stakeholders along with its possible consequences	Applicable/ Not applicable	
People can audit, query, dispute, seek to change or object to AI or robotics activities (when applicable)	Applicable/ Not applicable	

The AI system enables traceability during its entire lifecycle, from initial design to post-deployment evaluation and audit	Applicable/ Not applicable	
The system offers details about how decisions are taken and on which reasons these were based (when relevant and possible)	Applicable/ Not applicable	
The system keeps records of the decisions made (when relevant)	Applicable/ Not applicable	

Table 9 - Accountability & Oversight

Specification of Objectives against Ethical Requirements	Project relevance	Justification
Accountability & Oversight		
The system provides details of how potential ethically and socially undesirable effects will be detected, stopped, and prevented from reoccurring	Applicable/ Not applicable	
The AI system allows for human oversight during the entire life cycle of the project /regarding their decision cycles and operation (when relevant)	Applicable/ Not applicable	

Appendix III: Ethical assessment of software in ENVRI-Hub NEXT

In Appendix III, we present the template that will be used to record the results of the Ethics assessment performed by the project WPs contributing to the ENVRI-Hub platform.

This assessment will only include the Ethics requirements that the project found relevant for the ENVRI-Hub according to the analysis done and reported in **Appendix II - Specification of objectives against Ethical Requirements**.

The ethical assessment of software will be performed by the following WPs:

- WP7/8 are responsible for the Data and Service Catalogues. They provide a centralised entryway for accessing data and services from the ENVRI Science Cluster.
- WP9/10 are responsible for the ENVRI Knowledge Base and Search Engine and provide core features to enable end users to effectively discover research assets from multi-RIs providing user-centred support for semantic search using AI-based dialogue techniques.
- WP11/12 are responsible for implementing all fundamental enabling services necessary for a seamless user experience.
- WP13/14 are responsible for setting up the Analytic Framework to deliver Essential Climate Variables (ECVs) from all environmental domains as a service.

Table 10 - Respect for Human Agency

Specification of Objectives against Ethical Requirements	List of software relevant to the requirements	Requirement is implemented (Explain how implemented)	Requirement is relevant but NOT implemented (Risk identification required)	Requirement NOT relevant AT ALL (Justify the assessment)	Assessment from the Ethics Board and EB
Respect for Human Agency					
End-users and others affected by the AI system are not deprived of abilities to make all decisions about their own lives, have basic freedoms taken away from them					

End-users and others affected by the AI system are not subordinated, coerced, deceived, manipulated, objectified, or dehumanised, nor is attachment or addiction to the system and its operations being stimulated		29			
The system does not autonomously make decisions about vital issues that are normally decided by humans by means of free personal choices or collective deliberations or similarly significantly affect individuals					
The system is designed in a way that gives system operators and, as much as possible, end-users the ability to control, direct and intervene in basic					

operations of the system (when relevant)					

Table 11 - Privacy & Data Governance

Specification of Objectives against Ethical Requirements	List of software relevant to the requirements	Requirement is implemented (Explain how implemented)	Requirement is relevant but NOT implemented (Risk identification required)	Requirement NOT relevant AT ALL (Justify the assessment)	Assessment from the Ethics Board and Executive Board
Privacy & Data Governance					
The system processes data in line with the requirements for lawfulness, fairness and transparency set in the national and EU data protection legal framework and the reasonable expectations of the data subjects.					
Technical and organisational measures are in place to safeguard the rights of data subjects (through measures such as anonymisation, pseudonymisation, encryption, and aggregation).					

There are security measures in place to prevent data breaches and leakages (such as mechanisms for logging data access and data modification).		31			

Table 12 - Fairness

Specification of Objectives against Ethical Requirements	List of software relevant to the requirements	Requirement is implemented (Explain how implemented)	Requirement is relevant but NOT implemented (Risk identification required)	Requirement NOT relevant AT ALL (Justify the assessment)	Assessment from the Ethics Board and Executive Board
Fairness					
The system is designed to avoid algorithmic bias, in input data, modelling and algorithm design. The system is designed to avoid historical and selection bias in data collection, representation and measurement bias in algorithmic training, aggregation and evaluation bias in modelling and automation bias in deployment					

The system is designed so that it can be used by different groups of end-users with different abilities (whenever possible/relevant)					
The system does not have negative social impacts on relevant groups, including impacts other than those resulting from algorithmic bias or lack of universal accessibility.					

Table 13 - Individual, and Social and Environmental Well-being

Specification of Objectives against Ethical Requirements	List of software relevant to the requirements	Requirement is implemented (Explain how implemented)	Requirement is relevant but NOT implemented (Risk identification required)	Requirement NOT relevant AT ALL (Justify the assessment)	Assessment from the Ethics Board and Executive Board
Individual, and Social and Environmental Well-being					
The AI system takes the welfare of all stakeholders into account and do not unduly or unfairly reduce/undermine their well-being					

The AI system is mindful of the principles of environmental sustainability, both regarding the system itself and the supply chain to which it connects (when relevant)		33			
The AI system does not have the potential to negatively impact the quality of communication, social interaction, information, democratic processes, and social relations (when relevant)					
The system does not reduce safety and integrity in the workplace and complies with the relevant health and safety and employment regulations					

Table 14 - Transparency

Specification of Objectives against Ethical Requirements	List of software relevant to the requirements	Requirement is implemented (Explain how implemented)	Requirement is relevant but NOT implemented (Risk identification required)	Requirement NOT relevant AT ALL (Justify the assessment)	Assessment from the Ethics Board and Executive Board
Transparency					
The end-users are aware that they are interacting with an AI system					
The purpose, capabilities, limitations, benefits, and risks of the AI system and of the decisions conveyed are openly communicated to and understood by end-users and other stakeholders along with its possible consequences					
People can audit, query, dispute, seek to change or object to AI or robotics activities (when applicable)					

The AI system enables traceability during its entire lifecycle, from initial design to post-deploymen t evaluation and audit		35			
The system offers details about how decisions are taken and on which reasons these were based (when relevant and possible)					
The system keeps records of the decisions made (when relevant)					

Table 15 - Accountability & Oversight

Specification of Objectives against Ethical Requirements	List of software relevant to the requirements	Requirement is implemented (Explain how implemented)	Requirement is relevant but NOT implemented (Risk identification required)	Requirement NOT relevant AT ALL (Justify the assessment)	Assessment from the Ethics Board and Executive Board
Accountability & Oversight					
The purpose, capabilities, limitations, benefits, and risks of the AI system and of the decisions conveyed are openly communicated to and understood by end-users and other stakeholders along with its possible consequences					
People can audit, query, dispute, seek to change or object to AI or robotics activities (when applicable)					
The AI system enables traceability during its entire lifecycle, from initial design to post-deploymen					

t evaluation and audit					
The system offers details about how decisions are taken and on which reasons these were based (when relevant and possible)					
The system keeps records of the decisions made (when relevant)					