



AnaEE-ERIC

Data Management

Policy

2026

CONTACT@ANAEE.EU

WWW.ANAEE.EU

DOCUMENT INFO

Name of document	AnaEE-ERIC Data Management Policy
Status	Draft
Approved by	AoM
Written by	Lead Author: Dario De Nart Co-author: Lavanya Premvardhan, Adriano Palma
Feedback and review	Michel Boër, Ivan Nijs, Mili Barcelo, Jean-François Le Galliard
Submission to AoM	May 2026
Version	2.2
Dissemination level	Public

REVISION HISTORY

DATE	VERSION	COMMENT
May 2025	2.2	Authors: Dario De Nart, Lavanya Premvardhan, Adriano Palma, Michel Boër, Ivan Nijs, Mili Barcelo, Jean-François Le Galliard
November 2025	2.1	Authors: Dario De Nart, Lavanya Premvardhan, Adriano Palma, Michel Boër
November 2024	2.0	Authors: Dario De Nart, Lavanya Premvardhan
October 2020	1.0	Authors: Dario De Nart, Florent Massol, Flora De Natale, Chloe Martin, Giovanni L'Abate, Barbara Parrisè.

TABLE OF CONTENTS

DEFINITIONS	2
1. PURPOSE	3
2. GUIDING PRINCIPLES	3
3. DATA LIFECYCLE AND MANAGEMENT	4
4. INTELLECTUAL PROPERTY RIGHTS AND OWNERSHIP	4
5. ACCESS AND USE OF ANAEE-ERIC DATA AND SERVICES	5
6. RISK MANAGEMENT	6
7. DATA POLICY MANAGEMENT AND REVIEW	6
I. ACRONYMS	7
	8

Definitions

AnaEE data

Data generated within AnaEE-ERIC include datasets produced through the operation of AnaEE facilities and through research projects conducted at those facilities, within the experiment timeframe and in accordance with the limits, conditions, and responsibilities defined in the Service Level Agreement signed between the service provider and AnaEE. These comprise:

Core data generated systematically by AnaEE facilities through routine operations and standard measurement activities (e.g. environmental measurements, recurring monitoring variables, and standard laboratory analyses);

Project Data, generated by users conducting experiments or research projects hosted at AnaEE facilities;

Associated Digital Resources, including software, protocols, workflows, and documentation supporting the generation, processing, and interpretation of data.

Data owner

Legal entity that generates data within the framework of AnaEE-ERIC activities and, by virtue of such generation, holds the rights and assumes responsibility for the acquisition, use, and distribution of the resulting dataset. Entity who legally owns the data and decides to use AnaEE-ERIC's services to manage said data as a whole or in part. The legal entity can be an individual, a platform, an institution within AnaEE RI, or a third party involved in AnaEE-ERIC scientific initiatives. The interaction between AnaEE-ERIC and AnaEE RI data owners are typically regulated via an SLA.

Data steward

A person who *de facto* has the means to access and manage the data produced at a platform. They can be either internal or external to the platform, and a single

platform may have multiple data stewards, as well as the same data steward may be collaborating with multiple data owners.

Service Provider	A host institution of the technological platform. The legal entity that signs the Service Level Agreement
User	A person, a team, or an institution from any sector, including public and private sector, making use of AnaEE data or other AnaEE services, including access to AnaEE facilities.

1. Purpose

The role of this Data Management Policy (DMPol), part of the AnaEE-ERIC Rules of Operation, is to implement the access policy as foreseen in the statutes, art. 6.5 and 6.6. The DMPol define the organisational intent covering each aspect of AnaEE-ERIC's data-management related activities, including the ethical and legal implications of said data. The technological platforms (facilities) that make up AnaEE RI are the primary source of data. AnaEE-ERIC relies on a distributed network of technological platforms operated by Service Providers across member countries. These platforms generate and host a significant portion of the data produced through AnaEE-ERIC services, including both routine platform data and data generated by user research projects. Datasets may be hosted within AnaEE repositories, in other repositories operated by Service Providers, or in recognised national or international repositories. In order to ensure consistent curation, interoperability, and access to data generated across the distributed AnaEE infrastructure, a common AnaEE Data Policy is required.

The DMPol applies to data generated by all AnaEE-ERIC facilities and central services operating under Service Level Agreements.

2. Guiding principles

AnaEE-ERIC statutes (art. 6.5 and 6.6) states:

- Art. 6.5: "AnaEE-ERIC shall make available discovery level metadata of each project at the start, both for experimental settings in National Platforms and for data produced by service centres."
- Art 6.6: "Data shall be made available according to the open access policy of AnaEE-ERIC. Access will be regulated according to the Rules of Operation of the Statutes, in general having a common embargo period, after which the data shall become available to the public. The Access Policy shall take into account the European Data Protection legal framework (1) related to the sharing of personal data of users among the Members."

In this policy AnaEE-ERIC follows the European Commission's recommendation of being "as open as possible, as closed as necessary". Reasonable restrictions for data and metadata access consistent with open access principles shall be applied to certain datasets, especially when access could compromise potential industrial or commercial uses, breach personal data protection rules, or violate confidentiality or security requirements. Beside the standard embargo period, any departure from the open-access rule shall be granted by AnaEE-ERIC on a case-by-case basis.

The AnaEE-ERIC data policy aims to ensure that all activities related to data comply with the GDPR¹, and relevant European directives on environmental and research data. It also aim to implement FAIR principles²

¹ [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 \(General Data Protection Regulation\)](#)

² <https://www.go-fair.org/fair-principles/>

(Findable, Accessible, Interoperable, Reusable) and to promote the use of widely accepted community licensing schemes, such as Creative Common licenses.

Users of AnaEE-ERIC data and digital tools are required to make resulting publications available through open access repositories, and their results available through AnaEE-ERIC (Statutes art. 8). Open source for software is recommended, when possible. Users shall cite AnaEE-ERIC when using AnaEE data in publications.

3. Data lifecycle and management

Research data within AnaEE are generated at the platform level by: (i) service providers and installation managers through routine operations and standard measurement activities carried out by AnaEE facilities; and (ii) users of the services through experiments or research projects conducted at AnaEE facilities. For the purposes of AnaEE core data, this primarily refers to data generated during experiments timeframe or service provision carried out at AnaEE facilities, within the limits and conditions defined in the Service Level Agreement signed between the service provider and AnaEE. From a technical perspective, these data are mainly produced in three main ways: (i) automated permanent sensors (machine-based measurements), (ii) deployable non-permanent sensors requiring human intervention for data collection, and (iii) outsourced data production, such as sampling and analysis performed in external laboratories.

AnaEE-ERIC ensures a coherent approach to the entire data lifecycle. Data generated at AnaEE-ERIC platforms must be properly curated, including the creation of appropriate metadata which shall be submitted to the AnaEE-ERIC Data and Modelling Centre (DMC) through the dedicated portal within the agreed timeframe. These metadata should include a standardized description of the measurement site, facility, and experimental setting, including geographic and temporal coverage, in order to support interpretation, comparability, and reuse across platforms. Datasets may be deposited in the DMC repository or in external trusted repositories. When deposited in the DMC repository, the DMC is responsible for quality control and for ensuring the long-term preservation and availability of the datasets. When datasets are hosted in external repositories, the respective platforms shall maintain uninterrupted access to their data and services. In all cases, datasets must be exposed through the DMC catalogue to ensure discoverability and access.

Details of the data lifecycle and data management are provided in the AnaEE Data Management Plan (DMP)³.

4. Intellectual property rights and ownership

Ownership of data generated within AnaEE institutional activities resides with AnaEE-ERIC and the entities or individuals who have generated the data, unless a different arrangement is explicitly set out in a binding agreement and in compliance with the statutes and rules of operations of AnaEE-ERIC.

Data ownership depends on the type of data generated within AnaEE facilities. Core data produced through routine operations and standard measurements remain the property of the Service Provider. Where datasets are produced jointly by two or more platforms, or by platforms and users, including within hosted projects, ownership and related rights are shared and governed by specific arrangements agreed among the parties.

Where data are linked to a scientific publication or other agreed research output, they may be subject to an agreed embargo period. During the embargo period, the relevant rights holders, as applicable, shall have the right to analyse and use the data for the preparation of the related publication or research output. During this period, the Service Provider shall remain responsible for the secure storage, preservation, curation, and controlled access to the data, in accordance with the applicable agreement.

³ AnaEE-ERIC Data Management Plan, *STD Appendixes*, Annex 6. Available at: https://www.anaee.eu/wp-content/uploads/2025/09/STD_Appendixes_AnaEE_ERIC.pdf

After the expiration of the embargo period, the data shall be made openly accessible, unless an exclusion or extension applies under specific binding agreements. From that point, the Service Provider shall be responsible for:

- ensure the long-term preservation of the data;
- integrate the data into the relevant repositories;
- enable the reuse of the data;
- require appropriate acknowledgement and citation of the facility.

The DMC operates and maintains the repositories and modelling services that AnaEE-ERIC makes freely available to the platforms. Data assets are, in the first instance, managed by their owners in accordance with their internal or national policies, the AnaEE-ERIC data policy, and any specific arrangements governing access to the facility or hosted user projects. In any case, the owners shall publish research datasets, including combined datasets resulting from user-generated data and standard measurements collected by the Service Provider, in a trusted repository in accordance with the repository requirements set out in Section 5 of this policy and, where applicable, after the agreed embargo period. Owners that participate in AnaEE-ERIC activities are entitled to use DMC services to implement such management. If a dataset is found to be non-compliant with applicable policies or third party right, AnaEE-ERIC will request corrective action from the owner to ensure that third party interests are properly respected and acknowledged.

Attribution is an integral part of intellectual property management within AnaEE-ERIC. In accordance with EU best practices⁴, all AnaEE datasets must include appropriate, machine-actionable data citations so that creators and curators receive credit and users can reliably reference the assets.

5. Access and use of AnaEE-ERIC data and services

AnaEE-ERIC is committed to supporting the long-term accessibility of all published data assets. To this end, the DMC operates the institutional repository, the public catalogue and a set of data-centric services that together foster preservation, citability and reliable access. The AnaEE-ERIC repository is the preferred venue for storing Research Data produced within the infrastructure; exclusive use, however, is not required. Data owners may also deposit in other acceptable repositories or domain-specific services listed on FAIR sharing⁵, and are encouraged, where feasible, to maintain mirrored copies across multiple repositories to improve availability and resilience, provided this does not conflict with other binding data-management provisions.

An embargo period is granted to the data owner, defined in the access guidelines⁶. Any extension of this embargo period, or departure from the open-access rule shall be granted only by AnaEE-ERIC on the request of the owner, duly justified, and in case of hosted project before any access to the infrastructure is granted. AnaEE-ERIC DMC shall implement procedures to guarantee the unique access by the data owner to the data during the embargo period, or other specific rules. This does not prohibit the data owner to provide access to its data from the start, while metadata describing the dataset shall be in any case public. These metadata refer to discovery-level information about the dataset (e.g. title, description, keywords, responsible organisation, geographic and temporal coverage, and contact information), while the data themselves and detailed structural metadata remain restricted during the embargo period.

All AnaEE-ERIC research data shall be exposed in the public catalogue maintained by the DMC. Regardless of the repository used, all datasets shall be registered in the AnaEE catalogue, which provides the core metadata describing the dataset and the reference or access point (e.g. DOI, URL, or API endpoint) to the location where the dataset is hosted.

⁴ Publications Office of the European Union, Jessop, P., Data citation – A guide to best practice, Publications Office of the European Union, 2022 <https://data.europa.eu/doi/10.2830/59387>

⁵ <https://fairsharing.org/search?fairsharingRegistry=Database>

⁶ User's Guidelines for Access to the AnaEE Distributed Research Infra-structure, 2025. Available at https://www.anaee.eu/wp-content/uploads/2025/09/anaee-users-guidelines-and-access-policy-2025_compressed.pdf

Datasets shall be uniquely and persistently identifiable, structured and documented with clear, machine-actionable metadata describing their purpose, scope, provenance, quality, and applicable community standards. Each dataset shall also include a suitable license enabling reuse, with Creative Commons Attribution 4.0 (CC BY 4.0) recommended as the default.

Users of AnaEE data are required to acknowledge the infrastructure, the platform of origin and significant individual contributions in any research outputs, and to include appropriate data citations. Data owners are responsible for clarifying and identifying which parties should receive attribution.

6. Risk management

AnaEE-ERIC manages continuity risks on two fronts: keeping its own digital platforms available and helping data owners keep their datasets accessible. The DMC monitors the portal, catalogue and developer services, coordinates incident response, and, when needed, activates disaster-recovery procedures to minimise downtime. Dataset availability remains with the respective owners, who can reduce risk by depositing in the AnaEE-ERIC repository and keeping metadata and links current. If an owner plans to discontinue maintenance of any AnaEE data set, they shall give one year advance notice so the DMC can arrange, under the Service Provider SLA, the transfer or copying needed to preserve access for users.

7. Data policy management and review

AnaEE-ERIC recognises that the data and IT landscape evolves quickly and that stakeholder requirements change accordingly. The Data Management Policy is therefore reviewed as needed. As part of this review, the DMC gathers input from National Nodes (NNs), Service Providers (SPs), the Ethics/IEAB bodies, and relevant user and community forums, and analyses which policy areas are affected.

Once approved, the AnaEE-ERIC publishes the updated DMPol and communicates the changes to all stakeholders, including the date when the revision takes effect.

I. Acronyms

AnaEE-ERIC	Analysis and Experimentation on Ecosystems - European Research Infrastructure Consortium. The legal entity comprising the Central Hub, Data and Modelling Centre, Interface and Synthesis Centre, and Technology Centre.
AnaEE - RI	The AnaEE Research Infrastructure. The RI comprises the technological research platforms / facilities that provide the services defined in the AnaEE-ERIC statutes
DOI	Digital Object Identifier A proprietary Persistent Identifier (PID) managed by the International DOI Foundation (IDF), a not-for-profit membership organisation that is the governance and management body for the federation of Registration Agencies providing DOI services and registration, and is the registration authority for the ISO standard (ISO 26324) for the DOI system.
FAIR	Findable, Accessible, Interoperable, Reusable. This acronym identifies a set of properties scientific data assets should possess to enable reproducible research and thus to be relevant to the research community.
GDPR	General Data Protection Regulation The EU law that establishes the legal framework for personal data management
SLA	Service Legal Agreement The legally binding agreement between AnaEE-ERIC and, a) its NNs / members and b) Service Providers, the owners of the facilities providing AnaEE services



ANALYSIS AND EXPERIMENTATION ON ECOSYSTEMS



1, Avenue de la Terrasse
91190 Gif-sur-Yvette, France
contact@anaee.eu

WWW.ANAEE.EU