EPOS ERIC

EPOS DATA POLICY

Document information	
Status	Public
Coordinator	Carmela Freda
Authors	Daniela Mercurio, Rossana Paciello, Federica Tanlongo, Carine Bruyninx, Otto Lange, Mario Locati, Alessandro Turco, Fulvio Galeazzi, Lucio Badiali, Ornela De Giacomo, Helen Glaves
Contributors	Carmela Freda, Daniele Bailo, Jan Michalek
Reviewed and Acknowledged	EPOS ERIC Service Coordination Committee
Graphics and editorial design	Barbara Angioni
Version	1.0.0
Date	18 September 2025
How to cite	Mercurio D., Paciello R., Tanlongo F., Bruyninx C., Lange O., Locati M., Turco A., Galeazzi F., Badiali L., De Giacomo O., Glaves H., (2025). EPOS Data Policy. https://10.5281/zenodo.17287075

EPOS DATA POLICY

Table of contents

1. Introduction

The EPOS mission is to establish sustainable and long-term access to solid Earth science data and services integrating diverse European Research Infrastructures under a common federated framework.
The EPOS Data Policy is addressed to stakeholders aware that open data is mandatory within the EPOS distributed context and that data will be accessible and reused through the EPOS Platform.

EPOS relies extensively on organizations that act as Data Providers, granting EPOS the right to distribute their data. To facilitate access, the data is integrated into the EPOS Platform[1] via web services provided by Service Providers designated by the scientific communities. These Service Providers ensure seamless sharing and interoperability across the EPOS infrastructure by implementing standardized protocols.

Throughout the document, the term "data", if not differently specified, refers to a broader concept that includes not only the scientific data itself, whatever the level of generation and processing, but also all related digital assets and any components involved in their lifecycle management.

The EPOS Data Policy defines the roles and responsibilities of all key stakeholders within the EPOS Research Infrastructure who are involved in the data lifecycle. It establishes core principles governing the provision, integration, sharing, access, use and exploitation of data, ensuring its effective management and utilization. The document also includes a set of overarching, high-level principles that form the foundational framework for handling data. These principles specifically address critical aspects such as data management, interoperability and standards, data quality, data protection, and information security. The document outlines the quality framework upon which EPOS as a research infrastructure is built.

The EPOS Data Policy is intended for all individuals involved in the data lifecycle, from data collection to dissemination and reuse. This includes data managers who process data according to the FAIR principles under the guidance of a Data Provider, as well as technical personnel at the Service Provider level engaged in facilitating data distribution through the EPOS interoperability layer.

The EPOS Data Policy is also relevant to anyone interested in the reuse of EPOS solid Earth science data, particularly in relation to the EPOS perspective on Open Science and the opportunities available through the associated policies.

---

[1] https://www.epos-eu.org/dataportal

2. Values and Guiding Principles

The EPOS Data Policy is designed to establish a comprehensive framework for managing, protecting, and utilizing data within the EPOS Research Infrastructure. It aims to ensure that data is collected, stored, processed, and shared securely, in compliance with relevant regulations, while fostering openness and collaboration across disciplines.
The EPOS Data Policy is designed to be consistent with, and responds to, the EPOS ERIC Code of Ethics. In practice, all values and principles guiding how EPOS manages, protects, and makes data available must align with the ethical framework established by this Code of Ethics.

The EPOS Data Policy is guided by the following values:
1. **Openness**. The policy promotes the unrestricted sharing and accessibility of data, allowing for transparent research practices and enabling others to build upon existing knowledge. This principle ensures that data is made available to a broad range of stakeholders, encouraging reproducibility and accelerating scientific discovery. Openness is key to fostering trust and facilitating global collaboration, ensuring that knowledge is shared freely within the scientific community and beyond.
2. **Innovation**. The policy fosters technical and scientific innovation by promoting cross-fertilisation between diverse perspectives, encouraging the evaluation of alternative hypotheses and solutions, and the integration of various insights to generate creative solutions to complex challenges.
3. **Collaboration**. Recognizing the value of interdisciplinary collaboration, the policy promotes cooperation among diverse scientific communities. This collaborative approach enhances the quality and impact of research.
4. **Efficiency**. The policy emphasizes efficiency in data management to maximize the impact of investments. By preventing duplication of effort and enabling reuse and enhancement of existing data, resources can be channeled towards the most promising initiatives.
5. **Accountability**. EPOS is committed to maintaining high standards of accountability. The policy encourages independent verification of data and its provenance, ensuring that research findings are credible, reliable and reproducible.
6. **Capacity Strengthening**. The policy supports capacity building by facilitating the education of young researchers and expanding access to data for secondary analysis. This is particularly important for researchers in developing countries, enabling them to benefit from global scientific advancements.

EPOS acknowledges that scientific communities are at different stages of implementing data sharing practices and may use various methods for data distribution. To accommodate these diverse models and needs, EPOS adopts a flexible approach, recognizing that a one-size-fits-all solution is not feasible. This flexibility reinforces open scientific inquiry, encourages diverse perspectives, and promotes new research.

The EPOS Data Policy is based on the following guiding principles:
1. **Open Access**. EPOS is committed to open access to data and knowledge, ensuring that research data is widely available to the public[2].
2. **Timely Availability**. The policy mandates that data is made available promptly, without undue delay, while considering the different needs for virtual, remote, and physical access[3].
3. **Organisation for Economic Co-operation and Development Principles and Guidelines**[4] for research data from public funding, to ensure that data is managed in a way that supports transparency and accountability.
4. **Alignment with the FAIR** (Findable, Accessible, Interoperable, Reusable) Principles[5] for data access further enhancing the policy's alignment with global best practices.
5. **GEO Data Sharing and Data Management Principles**[6], which are also widely adopted throughout the geoscience community.
6. **TRUST** (Transparency, Responsibility, User focus, Sustainability and Technology) Principles[7] as a means to

---

[2] EPOS adheres to the general principle "as open as possible, as closed as necessary", so reasonable restrictions to the accessibility and reuse of data is also possible for legitimate reasons. Further details on this see Section 4.3.

[3] Exceptions to this general provision are data subject to embargo, as discussed in Section 4.3.

[4] OECD Principles and Guidelines for access to research data from public funding.

[5] https://data.europa.eu/doi/10.2777/8299402

[6] GEO Data Working Group. (2024). GEO Data Sharing and Data Management Principles. GEO Knowledge Hub. https://doi.org/10.60566/3p7vv-ykj55

[7] https://doi.org/10.1038/s41597-020-0486-7

demonstrate the digital repositories' trustworthiness and acknowledges the importance of the data sources included in the EPOS Research Infrastructure.

7. **CARE** (Collective Benefit, Authority to Control, Responsibility, Ethics) Principles for Indigenous Data Governance[8].

8. **Community Licensing**. The policy promotes the use of widely accepted community licensing schemes, such as Creative Commons, to facilitate the sharing and reuse of data.

9. **Acknowledgement and Credit.** EPOS recognizes the fundamental role of all contributors in the data lifecycle to ensure transparency, incentivize data sharing, and uphold research integrity. EPOS embraces the use of citation and acknowledgements in publications and derived research outputs, ensuring that contributors receive appropriate credit for their work.

---

[8] https://datascience.codata.org/articles/1158/files/submission/proof/1158-1-8528-2-10-20201104.pdf

3. European Legal Framework

EPOS Data Policy considers the relevant European legal framework related to environmental data, information, and databases:

- **INSPIRE Directive (2007/2/EC)**. Establishes an Infrastructure for Spatial Information in the European Community, aimed at making environmental spatial data more accessible and interoperable across Europe, including the INSPIRE implementing rules[9] and technical implementations[10].
- **Access to Environmental Information Directive (2003/4/EC)**. Implements the Aarhus Convention in the EU, ensuring public access to environmental information held by public authorities.
- **Data Protection Regulation (GDPR - 2016/679)**. While not specifically environmental, this regulation governs the processing of personal data and is relevant for environmental data involving personal information[11].
- **Open Data Directive (2019/1024/EU)**. Promotes the availability of public sector information (PSI) for reuse, which includes environmental data, and encourages making this data available in open formats.
- **Regulation (EU) 2022/868 (Data Governance Act)**. Establishes a framework for data sharing and governance across the EU, including environmental data, to facilitate reuse, promote research and innovation, and ensure responsible data handling.
- **Regulation (EU) 2023/138 (High-Value Datasets)**. Defines a set of high-value datasets, ensuring their free publication and reusability to boost transparency, innovation, and economic growth across the EU.
- **Regulation (EU) 2023/2854 (Data Act)**. Entered into force on January 11, 2024, and has applied since September 12, 2025, fostering digital innovation in the EU and facilitating data sharing and access, including for environmental data.
- **Regulation (EU) 2024/903 (Interoperable Europe Act)**. Establishes measures to enhance cross-border interoperability of digital public services across the EU, promoting seamless data exchange and cooperation among public administrations.
- **Regulation (EU) 2024/1689** of the European Parliament and of the Council, entered into force on August 1, 2024, which lays down harmonized rules on artificial intelligence (also known as the Artificial Intelligence Act or AI Act).
- **Database Directive (96/9/EC)**. Provides legal protection for databases through copyright and sui generis rights, safeguarding against unauthorized extraction and reuse of substantial parts of databases.
- **Software Directive (2009/24/EC)**. Provides legal protection for computer programs, ensuring that software, including those used for environmental data management, is protected under EU copyright law.

EPOS Data Policy also recognises relevant international observation system initiatives and national policies and legislation with the aim of full and open exchange of data, metadata and derived data products being made available with minimum delay and at no cost, except in exceptional cases where a minimal cost recovery is necessary.

---

[9] https://knowledge-base.inspire.ec.europa.eu/legislation/implementing-rules_en

[10] See in particular for geological data: https://github.com/INSPIRE-MIF/technical-guidelines/blob/main/data/ge/dataspecification_ge.adoc

[11] EPOS ERIC Privacy Policy is available at: https://www.epos-eu.org/epos-eric-privacy-policy

4. EPOS Data Lifecycle

The EPOS Data Lifecycle encompasses several stages, ensuring that data is efficiently generated, collected, processed, documented, harmonized, integrated, and made accessible to all potential stakeholders.
To support consistency and clarity across the lifecycle, EPOS employs a data taxonomy that categorizes data into different levels based on the level of processing as detailed in Table 1.

Table 1. Data taxonomy and processing levels

Level Code	Data Processing Level	Explanation and example
Level 0	Raw data generated by a data source.	Instrument data as collected from a data source, without any processing. May include noise and is not yet curated/organized. Format can still be proprietary and not interoperable. *E.g., seismograms, accelerograms, time series, full resolution satellite payload data etc.*
Level 1	Data products coming from automated or semi-automated procedures.	Level 0 data but processed, organised, and presented in a usable form for use by humans and machines. Formats are the ones adopted in the specific domain (standards or de facto standards). *E.g., earthquake locations, magnitudes, focal mechanisms, shakemaps, etc.*
Level 2	Data products resulting from scientists' investigations.	Data and data products that result from the analysis and combination of Level 1 data. Typically, at this level, data is monodisciplinary, and standards still reflect the ones adopted within a single domain/community. *E.g., crustal models, strain maps, earthquake source models, variables mapped on uniform space-time grid scales etc.*
Level 3	Integrated data products resulting from complex analyses or community-shared products.	Complex scientific outputs that are the result of combining and integrating data and methodologies from the lower processing levels, and originating from different subdomains. This level most typically features multi-, inter- or even cross-disciplinary data and products that require higher levels of interdisciplinarity. *E.g., (multi)hazards maps, catalogue of active faults, model outputs etc.*
Level 4	Software and IT tools developed to support data processing and analysis.	Reusable pieces of code and applications that can be used to automate complex scientific workflows and can be applied to different use cases. *E.g., Jupyter notebooks.*

The EPOS Data Lifecycle (Figure 1) is described in three main stages:
- **Data Provision**: data is generated, collected, and managed in accordance with established protocols, standards, and policies.
- **Data Integration**: data enters a community-driven process of standardization, harmonization, and enrichment within the EPOS Thematic Core Services[12]; at this stage, data is aligned with the standards defined by the relevant scientific communities and is made accessible via web services together with its metadata. Importantly, the metadata must remain available even in cases where the original data becomes inaccessible, thereby preserving transparency, provenance, and discoverability.
- **Data Access & Reuse**: web services are integrated into the Integrated Core Services which, through the EPOS Platform, act as the central hub for data access; the Integrated Core Services ensure that data from diverse scientific domains can be accessed, combined, analyzed and reused in a coherent and interoperable way; this integration facilitates cross-disciplinary research and analysis.

---

[12] Anthropogenic Hazards; Geological Information and Modeling; Geomagnetic Observations; GNSS Data and Products; Multi-Scale Laboratories; Near-Fault Observatories; Satellite Data; Seismology; Tsunami; Volcano Observations.
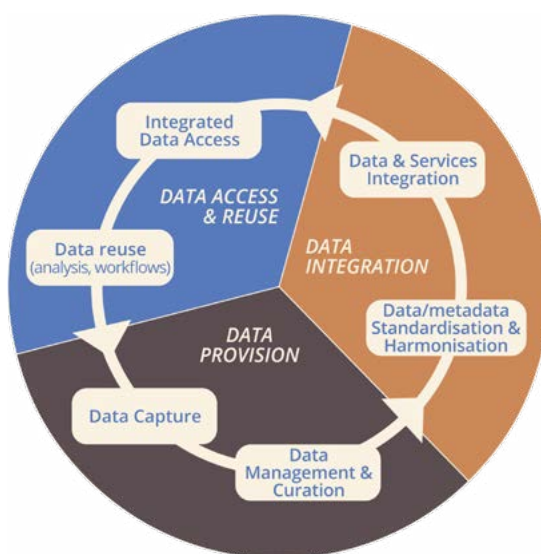
Figure 1. EPOS Data Lifecycle: key stages through which data progresses, from capture to final reuse.

Within the EPOS Data Lifecycle, two key actors are pivotal in the creation, publication and curation of the data throughout its lifecycle, and are crucial for ensuring high data quality as well as adherence to the FAIR Principles:

- **Data Provider**: a legal entity, typically a research organization, university, or private company, responsible for supplying data across various levels of the EPOS data taxonomy. As a Data Provider, the entity is committed to ensuring that data is accurate, relevant, and collected in compliance with established protocols and standards. The quality and reliability of the data supplied are fundamental to the effectiveness of subsequent processing and integration throughout the data lifecycle. Data providers allow the distribution of the data delivered through the EPOS Platform via a formal agreement (e.g., a letter of intent or supplier's letter, or in the framework of an existing agreement) aligned with the practices of the relevant Thematic Core Service (TCS).

- **Service Provider**: a legal entity operating within a TCS is responsible for aggregating, managing, and ensuring access to data from one or more Data Providers through the EPOS Platform. The Service Provider ensures that data is available through web services and that the underlying systems are available, reliable, and scalable.

Effective data management within the EPOS infrastructure relies on a set of functional roles and responsibilities, covering all stages of the data lifecycle, some of which fall outside the EPOS Delivery Framework.
As such, the definitions presented below are intended as guidelines aimed at clarifying functional responsibilities and fostering alignment across diverse contexts rather than as specific requirements, with the objective of fostering community-driven adaptation and convergence toward shared practices.

Data Producer: creates data and collaborates with the Data Manager to ensure accuracy, completeness, quality, and compliance with scientific standards and best practices for further processing and reuse.

Data Manager: oversees the operational management of data, covering collection, storage, security, and maintenance, to ensure data is accessible and readily available.

Data Curator: organizes and documents datasets, ensuring they are well-structured, high-quality, properly maintained, and augmented with metadata to guarantee usability for the broader scientific community.

Data Steward: ensures that data handling complies with relevant policies, ethical guidelines, and legal standards, and is responsible for defining data governance practices, overseeing data access protocols, and ensuring adherence to privacy and data protection laws.

IT Specialist: manages the technical infrastructure supporting data collection, processing, and accessibility, providing secure and scalable systems. This role involves maintaining network and server reliability, and implementing data security measures.

Depending on the specific needs and resources of the organization or the TCS, these roles may be assigned, either implicitly or explicitly, to an individual or a team. In some cases, a person or group may take on multiple responsibilities (e.g., acting as both Data Curator and Steward, managing tasks related to metadata quality, governance, and compliance). Their distribution can vary across organizations and Thematic Core Services, reflecting different institutional structures and community practices.

The following sections present an idealized framework of functional roles, designed to clarify responsibilities, promote consistency, and support alignment across EPOS. This framework serves as a reference for adaptation and convergence, while encouraging the adoption of shared best practices.

4.1 Data Provision

Within EPOS, data provision is carried out entirely by the Data Providers according to their mission and internal procedures. Data Providers ensure access to quality-controlled data that is generated, archived, managed and curated within the perimeter of the organization, with appropriate measures in place to guarantee its integrity and preservation, including controlled physical access to data infrastructure.

Typically, data provision also includes providing the metadata that describe the data and are an essential element for its management and curation. Both data and metadata are made available through web services (e.g., APIs), in line with the EU data regulatory framework[13] and the reuse of public sector information, which emphasizes the importance of providing data and metadata in machine-readable formats and through APIs to ensure easy access, interoperability, and facilitate automated data exchange and reuse across different sectors. Data Providers retain ownership and remain responsible for the management and curation of their data and metadata, including its quality and accuracy.

Within EPOS, Data Providers are organizations of proven quality and authority, whose participation is validated by the scientific community itself, to ensure the relevance and accuracy of the data that are then channeled into the infrastructure. The EPOS architecture ensures that new Data Providers can be integrated as they become operational and fulfill the requirements of the relevant Thematic Core Service.

Data and metadata are standardized and made interoperable through the Thematic Core Services, which represent the transnational governance frameworks where each community discusses data and services specific implementation, best practices and sustainability strategies for their provision, as well as legal and ethical issues.

The key functional roles and responsibilities at Data Provider and Thematic Core Services level are as follows:
Data Manager
> *Data Provider level*: the Data Manager, possibly in collaboration with the Data Producer, guarantee the operational management of data collection, storage, and security; this role ensures that data is properly gathered and maintained in line with national or international standards and protocols. The Data Manager focuses on maintaining the technical infrastructures, such as databases and data access systems, to guarantee that the data is ready to be shared.
> *Thematic Core Services level:* the Data Manager oversees the aggregation and integration of data from multiple Data Providers; it manages the infrastructure that enables data interoperability between various sources, ensuring that datasets are available and accessible through centralized web services. Responsibilities include ensuring that the systems for data access are reliable, scalable, and capable of meeting researchers' needs.

Data Curator
> *Data Provider level*: the Data Curator ensures that collected datasets are well-structured, accurately documented, and properly maintained. The Data Curator focuses on organizing the data, adding metadata, and ensuring that the datasets are of high quality and suitable for sharing. Primary responsibility is to guarantee that data is preserved and easily understood by any user.
> *Thematic Core Services level:* the Data Curator is responsible for harmonizing data and metadata from different Data Providers and ensures that the aggregated datasets are consistent, properly formatted, and standardized to meet the needs of the broader scientific community. The Data Curator also ensures that data is documented and ready for integration into the Integrated Core Services, prioritizing quality control and usability.

---

[13] See Section 3. European Legal Framework.

Data Steward

Data Provider level: the Data Steward focuses on local data governance and ensures that data collection complies with national policies and legal requirements, such as data privacy regulations (e.g., GDPR). Responsibilities include defining data management policies and ensuring that data is handled ethically and according to national and international standards.

Thematic Core Services level: the Data Steward assumes a broader role, ensuring that the governance of data from multiple Data Providers is aligned with European and international standards, and oversees policies for data sharing and access, guaranteeing ethical use and compliance with legal and regulatory requirements. Responsibilities also include managing data access and usage rights.

IT Specialist

Data Provider level: the IT Specialist ensures the functionality and security of local hardware, software, and networks used for data collection and storage and also provides technical support to ensure the availability of data management systems.

Thematic Core Services level: the IT Specialist focuses on the infrastructure required to aggregate and standardize data from diverse Data Providers. Responsibilities also include managing system performance, ensuring secure data access, and maintaining the central infrastructure that underpins thematic services.

EPOS supports a continuous process of review and assessment of its data systems to verify that Data Provision operates as envisioned, while actively seeking improvements and preventing or resolving problems. EPOS disseminates and promotes best practices related to data quality control and provides a mechanism to facilitate feedback from the community on all aspects of its data management systems. Feedback will be collected through the EPOS Platform and forwarded to the Thematic Core Services, who will address it internally. By fostering and disseminating good data management practices, EPOS aims to strengthen a culture of responsible data stewardship across the geoscience community.

In addition, EPOS places strong emphasis on monitoring the quality of its web services. Key indicators will include response times, the number of successful requests, and the quantity of peer-reviewed publications making use of EPOS services.

4.2 Data Integration

Within EPOS, data integration is achieved through a service-oriented approach, prioritizing flexibility and scalability. Instead of relying on the traditional method of centralizing data into a single repository, the EPOS approach focuses on integrating web services that provide direct access to the (meta) data, wherever it is located. This paradigm shift moves the emphasis from managing physical datasets to the associated mechanisms for accessing and interacting with data, enabling a more dynamic and adaptable architecture. By decoupling data storage from access mechanisms, the system ensures a scalable solution that accommodates a variety of data sources and formats.

At the core of this approach is the use of detailed metadata descriptions, which define the methods and protocols necessary to access and interact with the data through web services. These descriptions go beyond basic technical details and include communication protocols (e.g., REST), query parameters (e.g., filters, sorting), endpoints (URLs pointing to services), and the types of data available (e.g., time series, geospatial data, images). The metadata standard adopted is EPOS-DCAT-AP[14], an extension of the DCAT-AP standard tailored to the specific requirements of the EPOS Research Infrastructure, ensuring semantic interoperability between the Integrated Core Services and Thematic Core Services. To ensure effectiveness, metadata descriptions are stored in a central catalog, managed by the Integrated Core Services, with appropriate physical security measures in place to ensure its protection and long-term availability. The metadata descriptions are designed to make the web services FAIR compliant. This enhances the discoverability and usability of data, enabling users to easily identify relevant datasets and understand how to interact with them. Furthermore, the metadata includes semantic information, such as ontologies and standardized vocabularies, ensuring harmonized access to data services. These semantic elements enable service interoperability, allowing users to query and retrieve data from diverse sources in a consistent and meaningful way, regardless of the underlying format or structure.

By leveraging this service-oriented and metadata-driven approach, EPOS offers users a seamless experience when accessing, querying, and interacting with complex datasets.

---

[14] https://epos-eu.github.io/EPOS-DCAT-AP/

The key functional roles and responsibilities involved in Data Integration at the Integrated Core Services and Thematic Core Services level are:

(meta)Data Manager
- *Integrated Core Services level:* the Data Manager oversees the central metadata catalog, ensuring its robustness, security, and reliability.
- *Thematic Core Services level:* the Data Manager ensures efficient provision of the web services metadata and provides necessary technical support for seamless interaction with Thematic Core Services data sources.

(meta)Data Curator
- *Integrated Core Services level:* the Data Curator ensures that web services metadata from Thematic Core Services is well-structured, thoroughly described, and ready for integration into the central catalog.
- *Thematic Core Services level:* the Data Curator ensures the accuracy, quality, completeness, and usability of the web services metadata; they also guarantee metadata accurately describes the web services and provides clear information on data access, use, and interpretation.

(meta)Data Steward
- *Integrated Core Services level:* the Data Steward ensures web services metadata is managed according to ethical standards, including proper data sharing agreements and adherence to EPOS data governance frameworks.
- *Thematic Core Services level:* the Data Steward ensures web services metadata adheres to EPOS standards, using standardized vocabularies and ontologies for interoperability, and works closely with the Data Curator to ensure metadata is compliant and harmonized across different Thematic Core Services.

IT Specialist
Is responsible for the core infrastructure (Integrated Core Services) that integrates services and enables global access. They ensure that systems are scalable, secure, and capable of managing heterogeneous data.

4.3 Data Access and Reuse

EPOS adheres to the European Commission's data policies (see Section 3), which promote openness while balancing the need for reasonable restrictions. The guiding principle is *"as open as possible, as closed as necessary."* EPOS supports open access to data but recognizes that some datasets may require restrictions for legitimate reasons. These may include concerns related to industrial or commercial use, personal data protection, security, or confidentiality. In such cases, EPOS will apply appropriate restrictions that remain consistent with open access principles and always respect the conditions set by Data Providers for the use of their data.

When restrictions or embargoes are necessary, the responsible Thematic Core Services will evaluate and approve them. All such restrictions will be communicated clearly and transparently to users. Furthermore, information about the availability of restricted or embargoed data will also be made accessible to ensure that users remain fully aware of any limitations (see Section 6.3 for licensing conditions).

4.3.1 Access to Data

EPOS provides different types of data access, depending on the conditions stipulated by the relevant Data and Service Providers. These access types are defined as follows:
- **Open Access:** Data is freely available and accessible to users, either for download or for direct use within an EPOS service.
- **Restricted Access**: Data is available under specific conditions set out by the Data Providers. Restrictions may apply to particular categories of users and should be limited to specific datasets. In some cases, fees may be charged; however, metadata must always be available free of charge. Any applied fees must not exceed the actual cost of making the data available.
- **Embargoed Access**: Data become available only after a predefined embargo period has elapsed. As a general rule, the embargo period should be six months, but it may be extended up to three years from the date when metadata are first published. After the embargo period, the data will become either open or restricted, depending on the conditions set by the Data Providers.

Metadata (and data descriptions) are always free and available at any time, even for restricted and embargoed data. Regardless of access level, all data resources must be accompanied by a license describing the conditions for their reuse, as discussed in Section 6.3.

4.3.2 Distribution of Data

The strategy for data distribution through the EPOS Platform is defined by each Thematic Core Service and follows the access types outlined below:
- **Anonymous**. Data is accessible to users without authentication or authorization; Thematic Core Services should provide mechanisms to monitor the usage of their data even for unauthenticated users.
- **Registered/Authenticated**. Data is accessible to all registered users regardless of their authorization level; registration mechanisms may vary across specific services.
- **Authorised**. Data is accessible only to authenticated users who have been granted a specific authorization level; permissions for particular data or services are assigned to defined user group(s) in accordance with the relevant Thematic Core Services access policy.

In line with the EPOS mission to provide universal access to solid Earth science data, a certain degree of anonymous access to data will always be available to all users. However, EPOS recommends a Registered Access approach for: i. tailored services based on scientific interests and previous usage patterns, and ii. added-value services that require additional resources (e.g., AI and computing services).
Any collection of personal data in connection with Authorized Access must comply with applicable laws and regulations, with particular attention to the principle of data minimization.

4.3.3 Distribution of Software

Software disseminated via EPOS may fall into one of three categories:
- **Acquired Software**. Software obtained for use by EPOS or its users and distributed in accordance with the terms of the original license.
- **Contributed Software**. Software developed for other purposes, but relevant for EPOS, which may be subject to specific usage restrictions; redistribution is allowed insofar as permitted by the original license.
- **Generated Software**. Software developed within the EPOS infrastructure will be distributed under open source principles and released with an appropriate license.

4.3.4 Responsibilities and Liabilities

Several key actors play essential roles in ensuring proper data access within EPOS, including Data Providers, Service Providers, Thematic Core Services, Integrated Core Services, and users. Data Providers and Service Providers are responsible for setting data access policies and ensuring compliance with relevant legal frameworks. The Thematic Core Services and Integrated Core Services are responsible for evaluating and implementing data access protocols, including the approval of any restrictions or embargoes. Users, in turn, must accept the terms and conditions of use, adhere to licensing requirements, and ensure that data is used ethically and appropriately.
Before using the EPOS Platform, users are required to accept the Terms and Conditions and acknowledge the associated privacy policy. The usage of the platform is conditional on this acceptance, which includes agreement to the disclaimer that relieves EPOS, including the Data and Service Providers, of any liability related to the use of the data, as well as from any misuse of the data or associated metadata.

5. Data Security and Protection

Data security and protection must be carefully and continuously addressed during the entire data lifecycle, to ensure both the integrity and the availability of data.

Both personal and scientific data require appropriate security and protection measures. While personal information must be safeguarded to ensure privacy and compliance with regulations, scientific data must be protected against accidental loss, unauthorized handling, and corruption.

EPOS strives to ensure that all measures connected to data security and protection are implemented to the highest standards on all systems under its direct control. Data and Service Providers are also encouraged to protect the general security, robustness and resilience of the EPOS Platform.

Privacy and GDPR Compliance. EPOS and the other actors in the EPOS Data Lifecycle are committed to adhering to GDPR principles, with a particular emphasis on data minimization, which requires that the collection of personal data be limited to what is strictly necessary. Any personal data processed within the EPOS Infrastructure must be handled transparently, securely, and solely for clearly defined and legitimate purposes. Appropriate measures are applied to minimize the retention periods, and the rights of data subjects, such as access, rectification, and deletion, are upheld at every stage.

Data Integrity. Ensuring the integrity of scientific data is essential for maintaining the reliability and credibility of research findings. All actors involved in the EPOS Data Lifecycle must implement mechanisms such as checksums, cryptographic hashing, and version control to guarantee that the data remains accurate and unaltered throughout its lifecycle. Good practices include maintaining comprehensive logging and audit trails to record any modifications, access, or deletions of data, thereby ensuring full transparency and enabling review of all actions. EPOS is committed to taking all possible measures to preserve data integrity during the data provision process and will not alter data in any way throughout its lifecycle.

Protection Against Data Loss. Periodic automated backups are essential to safeguard against accidental data loss or corruption. These backups should be securely stored, ideally in multiple locations (on-site and off-site/cloud), and regularly tested for restoration. Data and Service Providers, along with Integrated Core Services managers, should maintain a disaster recovery plan to guarantee that data can be recovered quickly and efficiently in the event of hardware failure, cyberattacks, or natural disasters.

High Availability. Production data archiving and storage systems, as well as the servers hosting the Thematic Core Services and Integrated Core Services, should be designed, where possible, for high availability, with redundancy built in at multiple layers (e.g., data storage, network infrastructure). Such a design minimizes downtime and ensures continuous access to critical scientific data, even in the event of unforeseen incidents or failures affecting a specific machine or data center. For systems managing large volumes of data, load balancing and failover mechanisms should also be implemented to maintain optimal performance and mitigate potential failures without disrupting access.

Accidental Loss Prevention. Data Providers should implement safeguards to prevent accidental deletion or overwriting of data (e.g., multi-step confirmation prompts or "soft delete" features). Ideally, Data Providers should integrate validation checks at both input and processing stages to ensure data accuracy and prevent corruption caused by improper handling or entry errors.

Data Security and Access Control. While open access is promoted whenever possible, robust mechanisms for user authentication and role- or privilege-based access control are implemented where required by the nature of the data, services, or provider policies. These measures help prevent unauthorized access to, or manipulation of, sensitive or critical data.

Authentication. EPOS applies appropriate authentication measures wherever required, from both the Data and Service Provider and user perspectives. Any information obtained through authentication mechanisms is handled in accordance with the EPOS Privacy Policy.

Authorization. EPOS enforces appropriate authorization procedures to access data where necessary. Permissions are granted in line with provider policies, based on authenticated identity attributes (e.g., role, affiliation) or on a case-by-case basis according to specific requests.

6. Intellectual Property Rights

6.1 Ownership

EPOS is a complex infrastructure comprising various elements (e.g., data, services, software) provided by different organizations, each playing distinct roles in the EPOS Data Lifecycle. Unless otherwise specified for individual datasets, projects or services, no transfer of ownership, expressed or implied, occurs in this process. The same principle applies to all Intellectual Property Rights (IPR) associated with the ownership of services, datasets or other resources. Accordingly, all relevant IPR attached to a specific service, software, or dataset made available through the EPOS Platform remains with the owning organization, and in compliance with the applicable legislation. However, by making data available through the EPOS Platform, Data Providers also accept the EPOS policies related to openness, IPR, and licensing of the data, as well as requirements regarding the acknowledgement and citation of data sources (see next paragraph).

6.1.1 Jointly generated scientific outputs and access to background and side-ground

In the case of outputs that are jointly generated (for example, within EC-funded projects), the contributing parties shall have joint ownership according to the rules established in existing agreements (e.g., the Consortium Agreement of the project). If no such specific agreement exists, the parties shall agree separately in writing upon the conditions of the joint ownership. If an output is generated by drawing upon pre-existing knowledge (background) or knowledge generated in contexts other than the project or collaboration that gave rise to the output (side-ground), the party that generated the relevant background or side-ground shall retain full ownership of all related rights.

6.1.2 Third-party rights

Within EPOS, third-party rights refer to intellectual property rights (IPR) that are neither generated nor owned by any component of the EPOS Infrastructure. If EPOS, or any other component of the EPOS Infrastructure uses such third-party rights as part of its own intellectual property, it is the responsibility of that entity to obtain the authorization from the right holders, to grant access rights, and ensure that third-party IPR are not infringed.

6.2 Recognition of Data Sources

EPOS is a collective endeavor and all contributions to this effort must be properly recognised. It is recommended that researchers or organizations that publish or otherwise make available research outputs that have benefited from the resources available via the EPOS Platform cite both the EPOS Platform and the specific dataset, data product, software or service used, including any relevant Persistent Identifiers (PIDs).

EPOS recommends using PIDs for datasets, services, and research outputs to ensure proper attribution, traceability, and long-term accessibility. EPOS is agnostic with respect to the kind of PID that is used to identify datasets and other scientific products as it may depend on the type of resources to be referenced or the choices made by Data Providers and their scientific domain. However, EPOS provides a facility for minting unique persistent identifiers (DOIs) for those digital assets that are discoverable through the EPOS Platform but have no existing PID.

Each dataset available via the EPOS Platform should have a specific PID (e.g., a DOI[15]), which can be retrieved directly from the Platform. Details about the usage license and any other information relevant to the recognition of the data sources is also available, in line with the FAIR Principles (Wilkinson et al., 2016[16]).

EPOS also provides guidance on the handling of publications and their underlying data, promoting best practices for

---

[15] https://www.doi.org/
[16] https://doi.org/10.1038/sdata.2016.18

citation (see the EPOS Data, Data products, Software and Services Citation Guide[17]). This includes the provision of standardized citation strings through the EPOS Platform, as well as recommendations for data sharing and preservation to ensure the integrity, discoverability, and sustainability of research.

6.3 Licensing

Data, metadata and other scientific products should always be provided with a specific license that clearly informs users about what they are permitted to do with them and under which conditions. In the absence of a specific license, the default legal position is that no use is allowed without first obtaining permission from the individual owner. For this reason, EPOS will only distribute data where an appropriate license is affixed.

Within EPOS, Data Providers are responsible for assigning and enforcing licenses for their data, software or other scientific products. Service Providers are responsible for verifying the presence of a license in line with this Data Policy, and for ensuring the correct technical propagation of these licenses through EPOS services.

EPOS upholds the values of Open Science and the FAIR Principles and Data Providers are, therefore, required to make their scientific outputs available under valid license schemes that ensure optimal compliance with open data practices and the FAIR Principles, except where valid exceptions apply.

A wide range of open licensing schemes are available for different types of scientific outputs. When selecting a license, Data Providers should consider both international compatibility and the appropriateness of the scheme for the object being licensed. The licensing schemes most widely adopted for open science include: i) Creative Commons (CC)[18] for content, datasets and other products; ii) Open Source licenses, such as GNU/GPL, Apache or MIT, for software. Data (not curated as datasets) and metadata should ideally have an open domain license.

In the following sections, we provide an overview of recommended licensing schemes for different categories of scientific products, with examples of available options. Compliance with these licensing requirements is mandatory for data made available through the EPOS Platform.

6.3.1 Metadata

Metadata should be made as open as possible and must be released under an open license, such as the Creative Commons Public Domain Dedication (CC0) or equivalent scheme, to ensure maximum reusability[19]. Attribution-based licensing schemes (e.g., CC-BY) are also possible; however, they may impose restrictions in certain situations (e.g., Linked Data). A recommended compromise is to adopt CC0 licence, which recommends providing attribution but is non-binding.

6.3.2 Data, Databases, Datasets

An important distinction should be made between data, (curated) datasets and databases, as these are treated differently by current legislation.

Data (e.g., single observations) that are not organized into a database or published as curated datasets should, under the current legal framework, enter the public domain if produced by public organizations (including most universities and research organisations) and/or when funded with public resources. In such cases, data is not protected by IPR and should be released under a public domain license (e.g., CC0). Although not legally required, it is recommended to include information on data provenance or give appropriate credit (CC0 with attribution).

Structured data (e.g., curated databases and datasets) are protected by IPR and may be licensed under different schemes. Data Providers must carefully consider licensing options in terms of security, control, and compliance. While public domain requirements apply in many cases, restrictive conditions may also exist (e.g., where third-

---

[17] https://zenodo.org/records/14780630

[18] https://creativecommons.org/

[19] https://rea.ec.europa.eu/open-science_en

party rights are involved). In particular, some datasets may be available for reuse but not for modification for reasons such as ensuring data integrity. However, a *No Derivatives* license (e.g., CC BY ND) should only be applied in cases of a strong and fully justified need to preserve specific research outputs and documentation integrity. A generic reference to "data integrity" is not, as a rule, a sufficient justification for such a restriction.

A specific example of a clear exemption with regards to data availability and licensing conditions are the High Value datasets (EU Regulation 2023/138), which, unlike other datasets, are defined as being (cit.) *"associated with important socioeconomic benefits having a particular high value for the economy and society".* Due to their recognized socioeconomic impact, these datasets are subject to stricter openness requirements and *"shall be made available for re-use under the conditions of the Creative Commons Public Domain Dedication (CC0) or, alternatively, the Creative Commons BY 4.0 licence (CC BY), or any equivalent or less restrictive open licence [...] allowing for unrestricted reuse"*[20].

6.3.3 Content

In this context, "content" refers to a wide variety of digital objects and scientific products beyond datasets, for example, risk maps, publications, presentations, services, open educational resources (OER) and other curated objects. As a rule, all contents should include licensing and attribution information. The recommended licensing scheme for content is the Creative Commons Attribution (CC BY 4.0 International) licence, as it provides optimal compliance with the FAIR Principles. This license enables all forms of reuse and ensures licensing interoperability when combining heterogeneous material, while requiring only that appropriate attribution is given.

6.3.4 Software

Generally, all software distributed through the EPOS Platform should be free to use, modify, and share, and its license should comply with the Open Source Definition[21]. This includes free redistribution of the source code, the ability to create derivative works and the requirements to distribute the source code together with an open license. Different licensing schemes may be applied depending on specific needs[22].

Commonly used software licensing schemes include MIT License, Apache License 2.0, Mozilla Public License 2.0, GNU General Public License (GPL), Common Development and Distribution License (CDDL) and the European Free/Open Source Software Licence (F/OSS).

Organizations providing software who are concerned that companies may adopt, modify, and redistribute their software under more restrictive terms, particularly for commercial gain, are encouraged to adopt copyleft licenses (e.g., GPL). These licenses require that any modifications or derivative works remain under the same copyleft terms, thereby ensuring continued openness, reuse, and redistribution, including in for-profit contexts. Copyleft is generally considered best practice when the goal is to prevent software from being closed off or relicensed under more restrictive conditions (Free Software Foundation, 2023[23]). It supports broad reuse while safeguarding the collaborative and open nature of the software ecosystem. However, if the provider wishes to impose additional conditions on commercial use, it may consider adopting a dual licensing scheme. Using this approach, the software can be distributed under an open-source license for non-commercial purposes (e.g., academic or educational contexts), while a separate, more restrictive license, can be offered to commercial users. It is important to note that Open Source Initiative (OSI) approved licenses do not permit restrictions based on fields of use, such as "non-commercial use only". Therefore, any limitations on commercial usage must be introduced outside the scope of the OSI-approved license through a separate commercial license with a dual licensing framework.

6.3.5 Third-party content and software, cascading licensing and license compatibility

Data providers are responsible for verifying that the outputs they provide do not, to the best of their knowledge and belief, infringe any third-party intellectual property rights, and ensure that, where identified, third party interests are fully accounted for and acknowledged.

---

[20] See at http://data.europa.eu/eli/reg_impl/2023/138/oj - Article 4 "Arrangements for re-use applicable to all categories of high-value datasets.

[21] https://opensource.org/osd

[22] https://opensource.org/licenses

[23] Free Software Foundation (2023). What is Copyleft? GNU Project. Available at: https://www.gnu.org/copyleft

This principle implies that requirements discussed in this section may be difficult to fulfill when distributing an output by a third party or a derivative output, as the original license could impose limitations on sharing. In this case, more restrictive licenses are acceptable or even mandatory as applying a more open license could lead to infringements of the IPR related to the original work. Providers are responsible for making sure that the selected license is appropriate and if applicable and should check the compatibility of the derivative license with the original license[24].

Scientific output	Licensing scheme	Examples	Notes
Metadata	Public Domain Public Domain plus credits	• Creative Commons 0 (CC0) • Public Domain Mark • Public Domain Dedication • Open Data Commons Public Domain Dedication and license (PDDL)	
	Attribution	• Creative Commons Attribution (CC BY) • ODC Attribution • ISA Open Metadata licence	
Data (not organised)	Public Domain Public Domain plus credits	• Creative Commons 0 (CC0) • Public Domain Mark • Public Domain Dedication • Open Data Commons Public Domain Dedication and license (PDDL)	
Datasets and databases	Attribution	• Creative Commons Attribution (CC BY) • Open Data Commons Attribution licence	
Content	Attribution	• Creative Commons Attribution (CC BY) • Digital Peer Publishing Licences (DPPL) • GNU Free Documentation Licence (GFDL)	DPPL and GFDL only apply to textual works.
Software	Open Source	• MIT license • Apache License 2.0 • Mozilla Public License 2.0 • GNU General Public License (GPL) • Common Development and Distribution License • EUPL - F/OSS	The full list of available OS licenses can be found on the Open Source Initiative website https://opensource.org/osd
Third-party and derivative works	Depending on the original license	Same license or more restrictive than the original	Data providers are responsible for checking license compatibility before sharing third-party outputs or derivatives through EPOS.

---

[24] See for example https://wiki.creativecommons.org/wiki/Wiki/cc_license_compatibility for CC licenses.

Warranty

EPOS data is provided 'as is' to users without a warranty of any kind, whether expressed or implied, including warranties on quality or fitness for a particular purpose.

Review

The EPOS Data Policy is subject to revision in accordance with changes in legislation, as well as the evolving needs and strategic priorities of the EPOS Research Infrastructure.

EPOS ERIC Data Policy Documents

The table below provides the pointers to the documents associated with the EPOS Data Policy through their Digital Object Identifier (DOI)/Link.

DOCUMENT	Link/DOI
EPOS ERIC Platform Terms and conditions	https://www.epos-eu.org/sites/default/files/Terms_and_Conditions.pdf
EPOS Data, Data products, Software and Services Citation Guide	https://zenodo.org/records/14780630

Acronyms and Definitions

Data	Means any digital assets that are created, stored, and used for research, analysis, and interpretation (e.g. unstructured data, datasets, products, models, software)
Data Products	A processed, structured, or curated dataset derived from raw data, designed for specific analysis or visualization
Dataset	A dataset is a structured collection of data, typically organized in a standardized format, which can be accessed, processed, and interpreted as a coherent unit
Data source	Means any origin from which data is collected or retrieved (e.g., databases, sensors, computing systems)
EPOS	Means the European Plate Observing System Research Infrastructure as defined by Statutes[25]
EPOS Delivery Framework	Means the EPOS framework where the relationships among the key actors are regulated by specific rules and procedures. It includes the EPOS ERIC legal seat (represented by the ECO), the Integrated Core Services (ICS) and Integrated Core Services Central Hub (ICS-C) and the Thematic Core Services (TCS)
EPOS Platform	Means the multidisciplinary user-oriented platform built by EPOS to ensure the integration, access, use, and re-use of solid Earth science data and research outputs
FAIR	Findable, Accessible, Interoperable, and Reusable. FAIR refers to a set of guiding principles stating that data and metadata must be Findable, Accessible, Interoperable, and Reusable, thereby ensuring their effective management, long-term value, and broad usability across communities
GDPR	Means the General Data Protection Regulation (EU) 2016/679[26]
General Assembly	Means EPOS ERIC General Assembly. The General Assembly is the governing body of EPOS ERIC
ICS	Means Integrated Core Service. It represents the novel, single e-infrastructure that allows discovery, access, download, processing, and analysis of multidisciplinary data, products, and services to different stakeholders, including but not limited to the EPOS scientific community
Metadata	Means information describing data in digital form. They are structured information enabling data identification, access, interpretation, and reuse
Services	A system that provides access to data, software, or computational resources typically through an API (Application Programming Interface), enabling seamless data retrieval, integration, and interoperability across platforms
Software	A set of programs, scripts, or applications used to process, analyze, manage, or visualize data, often enhancing automation and reproducibility in research
TCS	Means EPOS Thematic Core Services. They are intended as a governance framework for each specific scientific community, linked technically to the ICS and strategically to the EPOS ERIC. They represent transnational governance frameworks where data, products, and services are provided to answer scientific questions

---

[25] https://www.epos-eu.org/sites/default/files/2024-06/EPOS ERIC Statutes Amended version 30May2024.pdf

[26] https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng