

# UNLOCKING THE FULL POTENTIAL OF THE GREEN DEAL DATA SPACE

## Making environmental data work for Europe's green transition

Environmental data are essential enablers of the European Green Deal. They support informed decision-making, underpin effective regulation, drive innovation, and empower society to act. Yet today, much of these data remains fragmented, inaccessible, or underutilised, creating significant barriers that limit Europe's progress towards its climate goals.

The Green Deal Data Space (GDSS) is envisioned as a solution to overcome these challenges and foster trusted exchange of public and private data across all areas of the green transition. To be effective, the GDSS must uphold the principles of Findability, Accessibility, Interoperability and Reusability (FAIR). Four projects—AD4GD, B-Cubed, FAIRiCUBE and USAGE—funded under the HORIZON-CL6-2021-GOVERNANCE-01-17 call have explored and documented the requirements for successfully achieving this. The GDSS will be implemented through the SAGE project (The Data Space for a Sustainable Green Europe) and will build on the results of the GREAT initiative (Governance of REsponsible innovATIOn).

After three years of intensive work, the four projects here join forces to present, in a single comprehensive document, a set of recommendations to guide the successful implementation of the Green Deal Data Space in the years ahead. These are as follows:

- 
**Strengthen data harmonisation, but stop reinventing the wheel.** Adopt and expand cross-domain standards, based on INSPIRE and other established frameworks, avoiding proprietary formats. Prioritise investment in tools that support data transformation and alignment across existing data structures.
- 
**Ensure resilient semantic interoperability on top of strong technical foundations.** Promote and enrich comprehensive controlled vocabularies with stable, well-defined concepts under established ontology frameworks, supported by technical infrastructure and governance for adaptability and long-term sustainability.
- 
**Recognise and resource the effort behind metadata.** Foster interoperability among varied metadata formats to ease the work of producers and users, and expand existing metadata standards to cover critical overlooked elements. Assure resources for metadata creation.
- 
**Enable data exchange between diverse stakeholders of the GDSS.** Develop effective strategies that motivate open data providers to participate, ensuring a balanced representation of public and commercial data within the data space, and promoting the use of standardised, GDPR-compliant, and federated technologies for data provision.
- 
**Establish inclusive, participatory and dynamic GDSS governance aligned with the European Green Deal.** Prioritise public interest, long-term sustainability, and adaptability. Leverage existing European data initiatives and state of the art solutions to preserve data sovereignty and security. Provide proper training, tools, and guidance—including real-world examples—to support effective adoption by participants of the data space.

# Data harmonisation

The European Green Deal encompasses a wide array of communities, from geospatial researchers to biodiversity conservation specialists, climate modeling experts, and circular economy businesses, across equally diverse nations, cultures and regions. Different communities have developed different standards for the representation of their data, often hindering data sharing and reuse. We suggest that the provision of all relevant datasets in the GDDS must consider:

- Investing in generic cross-domain standard data structures, vocabularies and representations that operate across a broad spectrum of complexities; avoiding too simplistic formats that ignore semantic links and metadata.

- Promoting open and well-documented formats, and ensuring easy transformation towards them, as the use of closed proprietary formats risk vendor lock-in.

- Building on the standards and guidelines for existing data formats, for instance, associating data to the vocabularies already established by INSPIRE across 34 thematic domains—many of which align with the European Green Deal—while continuously identifying gaps and ensuring their sustainable adaptation.

- Investing in tools that facilitate data transformation, including libraries and data structure mappings, enabling each community to shift between simpler, purpose-built or domain specific formats and richer, cross-domain generic formats as needed.

# Semantic Interoperability

Sharing data among the diverse community advancing the Green Deal objectives across Europe can sometimes feel like facing a digital Tower of Babel. Ensuring semantic interoperability of the data exchanged within the GDDS is critical for enabling its impactful use and reuse across all contexts and disciplines.

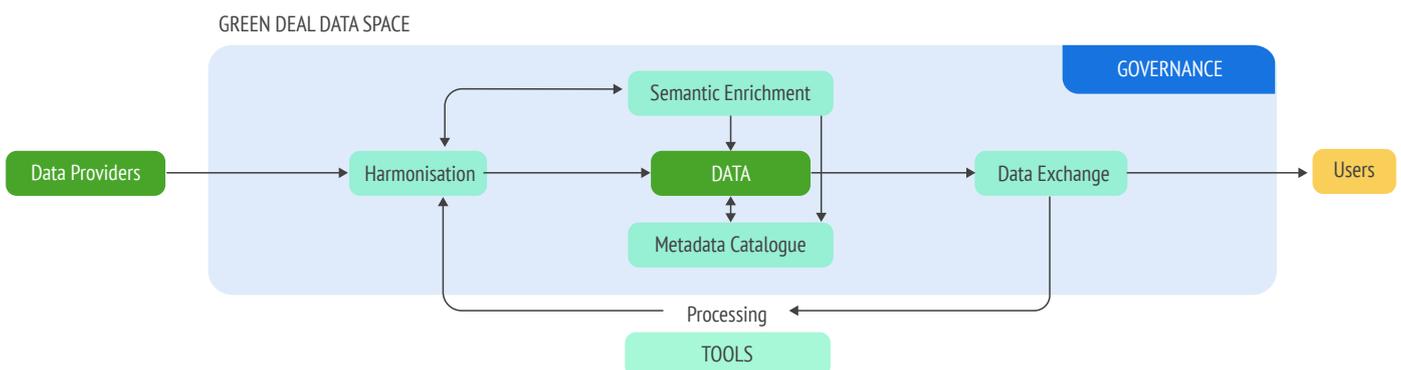
Regulation (EU) 2023/2854 (Data Act) requires participants in data spaces to provide machine-readable descriptions of datasets—covering content, usage restrictions, licenses, collection methods, quality, and uncertainty—to enable findability, accessibility, and reuse. It also mandates publicly available, consistent documentation of vocabularies, taxonomies, and code lists, among others, used in the data or services offered. We propose these requirements be taken a step further by:

- Promoting means for linking data and metadata, including variables, observed properties and units

of measure, to stable, persistent and well-defined concepts governed by shared, linkable, controlled and sustainable vocabularies and ontology frameworks that ensure robust preservation of meaning, such as the OGC/ISO Observations, Measurements and Samples standard.

- Where applicable, adopt existing standards for data structuring and annotation; when strictly necessary, extend them to cover gaps.

- Enrich internationally agreed and widely adopted vocabularies—such as the Essential Variables framework—with robust meanings, leveraging their established use within their communities of practice in each domain, e.g. utilising the I-ADOPT Ontology for the semantic representation of variables.



Simplified diagram of the Green Deal Data Space that represents the relationships between the thematic blocks of this policy brief.

# Metadata

To ensure that data shared through the GDDS can be easily found, understood and reused, it is essential that each dataset includes clear, structured metadata describing its content. Despite the critical importance of metadata, the Data Act acknowledges that poor metadata management is still one of the current barriers preventing data from delivering its full societal value. In many cases, no metadata is provided at all; in others, its content is incomplete or poorly structured, with essential concepts hidden in prose. Sometimes metadata is lost in translation between different non-interoperable structures and standards—from ISO 19115, STAC, GeoDCAT, EBV Cube Format, or the Eclipse Federated Catalogue. To ensure the findability and reusability of environmental data within the GDDS, we recommend:

- Expanding existing metadata standards to include critical elements that are currently poorly covered, such as observations, measurements and samples, I-ADOPT-compliant concept definitions, Data Product Specifications, and data provenance.

Identifying and promoting approaches—such as mappings and transformations tools—to enhance interoperability among in-use metadata standards, addressing the needs of commercial and open data alike.

- Ensuring that European research funding explicitly acknowledges the time and resources necessary for metadata creation and maintenance, and making it mandatory for successful project implementation.
- For European research projects, incorporate requirements for the creation and maintenance of a standardised metadata catalogue—documenting datasets produced from the project’s outset—into the data management plan.

# Data Exchange

The core focus of the GDDS is on the exchange of data from a diverse range of sources, both public and private, all geared toward advancing the goals of the European Green Deal. Sharing is based on data producers offering their valuable datasets under defined conditions in the data space, and automatically negotiating a usage contract or certificate with authenticated users in a secure and trusted environment. A demonstrator illustrating a potential GDDS across European projects and partners has been created, showcasing the potential of data space connectors. However, this implementation requires a specific local setup and can impose a barrier for some potential participants of the data space. Hence, we suggest:

- Adopting and promoting open, easy to use technologies that support seamless integration of both open and protected data.
- Ensuring the integration of diverse in-use APIs from data providers and user communities into the data space environment, accompanied by demonstrators and clear guidelines.
- Promoting the adoption of standardised APIs to facilitate smoother uptake, for instance the OGC APIs for geospatial assets.

- Offering conversion tools, like facades and brokers, and clear specifications to enable interoperability between APIs and data formats.

In addition, despite the success of Eclipse Dataspace Connectors, we emphasise that this technology still faces several challenges, including:

- Challenges to GDPR compliance, especially concerning the exchange of personal data in the context of contract negotiation and formation.
- Centralised architectural elements that create single points of failure, thereby compromising the overall resilience of the system.

We therefore advocate exploring alternative solutions based on blockchain catalogues, certificates, and personal wallets, such as the one implemented by the Federated Agile Collaborative Trusted System (FACTS) system.

# Governance

Establishing an inclusive, participatory and publicly-led governance framework for the GDDS that prioritises the common good is an essential step in better addressing upcoming challenges. This governance should ensure that the GDDS remains aligned with its strategic objectives and enables continuous, dynamic updates—particularly to keep pace with rapid scientific developments that may outstrip the evolution of vocabularies and standards.

## To this end, we recommend:

- Considering the global challenges related to international data governance and geopolitical tensions; it is essential to promote the use of exclusively European cloud solutions for storing and processing data within the GDDS. This approach ensures data sovereignty, enhances security, and guarantees compliance with EU regulations.
- Integrating existing open-data platforms such as the Copernicus Open Access Hub, the European Data Portal, and the INSPIRE Geoportal within the GDDS.
- Developing effective strategies to attract open data providers to join the GDDS, ensuring that the dataspace does not become exclusively commercial or paywalled.
- Mandating the citation of data in publicly funded academic publications to increase the visibility and recognition of open data, encouraging data providers to share their datasets more openly.
- Providing support to standardisation initiatives for conceptual models, underpinning vocabularies and exchange APIs, also ensuring long-term maintenance as new requirements evolve.
- Providing comprehensive training for both data providers and users on all aspects of the data provision lifecycle, including data harmonisation, semantic interoperability, metadata and data exchange mechanisms.
- Fostering cross-domain use of data harmonisation and provision tools, to prevent the emergence of new domain specific data silos, both within the GDDS as well as towards other related data spaces.

Unlocking the full potential of the Green Deal Data Space  
© 2025 by AD4GD, B-Cubed, FAIRiCUBE and USAGE is licensed under CC BY 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

**AD4GD** is conducting research and implementing FAIR building blocks for the European Green Data Space (GDDS). The project ensures interoperability of Earth observation data through the OGC SensorThings API and semantic enrichment pipelines. AD4GD also enables processing of machine learning data in the cloud and guarantees that the results are findable and securely accessible using the data space protocol implemented in the Eclipse Dataspace Connector. The building blocks are validated in three pilot cases: monitoring water quality and quantity in Berlin lakes, assessing habitat connectivity changes in Catalonia and improving the resolution of CAMS air quality maps with urban citizen science. The results will be transferred to the SAGE project to contribute to the development of the GDDS. Grant agreement ID: 101061001

**B-Cubed** aims to make data more accessible, interoperable, and usable, particularly to support evidence-based policy and deliver public value. It builds biodiversity data cubes that can be used to integrate in-situ biodiversity observations with other Earth observation data. These cubes are the basis for harmonised, policy-ready indicators to track past, current and future biodiversity scenarios. To ensure its tools meet users' needs and effectively capture relevant policy aspects of biodiversity change, B-Cubed is conducting four case studies, which launched in January 2025. These cover different locations varying in geographic extent, biodiversity richness and data availability. Grant agreement ID: 101059592

Core objective of **FAIRiCUBE** is to enable players from beyond classic Earth Observation (EO) domains to provide, access, process, and share gridded data and algorithms in a FAIR and TRUSTable manner. Within FAIRiCUBE, we demonstrate via five use cases a harmonized data space, the FAIRiCUBE Hub, where we connect all the pieces required for a data science pipeline into a user-friendly framework, where everything is FAIR and TRUSTable. In this manner, we illustrate how the Green Deal Data Space (GDDS) could be formed pertaining to gridded data and the analysis thereof. Grant agreement ID: 101059238

**USAGE** aims at providing solutions and mechanisms for making city-level environmental and climate data available to everyone based on FAIR principles, like: innovative governance mechanisms, consolidated arrangements, AI-based tools, and data analytics to share, access, and use city-level data from Earth Observation (EO), Internet of Things (IoT), authoritative and crowdsources, leveraging on standards for data and service interoperability. With 4 pilot Cities and 11 use cases USAGE highlight the importance of shared urban data spaces to enable visualization, sharing, fusion and processing of environmental and Earth Observation data from heterogeneous sources, ultimately facilitating more effective decision-making processes, besides advances in scientific research. The USAGE project has received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Grant Agreement no 101059950

10.5281/zenodo.15585325