

Data spaces: experiences from the European Green Deal

Unin. europo academy 15 March 2024 10.00 — 11.30 СЕТ

Rules of the game



The webinar will be recorded



For questions, please use the ClickMeeting chat.



Please reserve 3 min after the webinar to help us improve by filling in our feedback form



Our speakers



Carlos Martins Ferreira

Knowledge Management,

Data.europa.eu





Nevena Raczko Senior Consultant at IDC's European Government Consulting unit



Albana Kona Scientific Officer at the European Commission, Joint Research Centre



Margherita Di Leo Senior Consultant at European Commission, Joint Research Centre





Green Deal Data Space – Call for proposals -Usecases

2024 March 15 – DG DG ENV

<u>sotirios.kanellopoulos@ec.europa.eu</u> # Policy Officer



The scope of the European Green Deal Data Space





Call #DIGITAL2024CLOUDAI06GREENDEAL

Timetable and deadlines (indicative)						
Call opening:	29 February 2024					
Deadline for submission:	<u>29 May 2024 – 17:00:00 CEST</u> (Brussels)					
Evaluation:	June-July 2024					
Information on evaluation results:	July-August 2024					
GA signature:	February 2025					





Deploy an operational Green Deal Data Space (GDDS)

- The action is expected to deploy a technical infrastructure and governance mechanism for the GDDS with related use cases
- The action should enable reusing and sharing data from existing relevant private and public data ecosystems, which will feed new services and applications that contribute to reaching the objectives of the Green Deal
- The action should **take into account** the latest developments in the data strategy and green deal landscape, in particular:
 - Results of the DIGITAL <u>GREAT</u> project regarding preparatory actions for the Green Deal data space
 - Results of relevant HORIZON EUROPE projects to provide more accessible and exploitable environmental observation data in support of the European Green Deal priority actions (AD4GD, B-Cubed, FAIRiCUBE, USAGE)

Use-cases for the EGGDS

- This Call invites you to structure your proposals around use-cases
- The call has identified 3 use-cases of Interest linked to some of recent European Commission policy initiatives
- These 3 use-cases of interest are not restrictive.
- "Any Data Space use-case whose principal value lies in at least one of the four focus areas is in scope"





The "combined" DPP/ PRO-IS forms a Data Space: examples of value creation

- Life duration of textiles
- Waste flows at brand level
- Which data to feed-back to the DPP of newly manufactured products
- High granularity data in the sorting reports attached to shipments for re-use
- Ideas about how to leverage the data/ capacities of DPP in the context of a Data Space are welcome. Be it textiles or other value chains



Use Case 2: "due diligence" systems for companies supply chains

- private operators care more and more about their own environmental and sustainability performance.
- This is either due to the **imperative for compliance** to legislations like the Corporate Sustainability Reporting Directive, EU Deforestation Regulation or the Ecodesign for Sustainable Products Regulation
- Or due to the **pursuit of competitive advantage** in the sustainability race.
- This CFP aims to fund infrastructures/ applications that help private operators in their reporting/ sustainability journey and create value



Use-Case 2 : Due Diligence for deforestation free products

"Deforestation Risk" at the plot areas "Contamination Risk" with commodities of unknown origin upstream

Total Deforestation Risk



Use Case 2: "due diligence" systems for companies supply chains

- Regardless if you have already invested in a due diligence system and have already a beta version and look to invest further or you are starting to invest now, we think that the Green Deal Data Space is an interesting program for you.
- Where the imperative to disclose upstream in the value chain with other private operators collide with the need to protect commercially confidential information, applicants should propose digital solutions and protocols resolving such tension
- Applicants are encouraged to explore ways that data collected and shared in the context of due diligence can be used as a product for creating further value and propose potential business models



Use-case 3: National Forest Inventories and confidentiality preserving technologies

- The National Forest Inventories
- In contrast to the previous use-cases, in this one data (NFIs) are used as "ground-truth data" for training and validating ML models for forest monitoring
- Our proposal for a Forest Monitoring law has a provision for Confidentiality Preserving Safeguards to protect confidential attributes of NFIs and ensure data re-use
- While the decision making process is ongoing, this use-case is about the technological elements of those safeguards
- This CFP is asking for A Confidentiality Preserving Infrastructure that allows to access NFI data for training ML without compromising the confidentiality constraints



Use-case 3: National Forest Inventories and confidentiality preserving technologies





Use-cases demonstrating synergies are encouraged









Thematic Data Spaces

Destination Earth

Digital Product Lo Passport Tw

Local Digital Twins



Acknowledgments

- Directorate General Environment
 - D1 Land Use & Management
 - B3 From Waste to Resources
 - B4 Sustainable Products
- Directorate General Communications Networks, Content and Technology
 - E4 Internet of Things
- Joint Research Centre
 - D1 Forests & Bioeconomy
 - T1 Digital Economy



Outcomes and deliverables

- A governance structure for the GDDS
- A GDDS data-governance algorithmic mechanism
- A GDDS technical infrastructure for operating the selected use cases
- High-value, reusable datasets relevant for the selected use cases and more broadly the focus areas •
- A stakeholder mapping analysis of those who could join the data space after the end of the project.
- A set of concrete policy recommendations that will complement digital solutions in the implementation of the selected use cases.
- Operational / Financial Sustainability of the GDDS





• Type of action: Simple grant

Simple Grants are a flexible type of action used by a large variety of topics and can cover most activities. The consortium will mostly use personnel costs to implement action tasks, activities with third parties (subcontracting, financial support, purchase) are possible but should be limited

- Maximum grant amount: 8 mln EUR per project
- Funding rate: **50%**
- Reimbursement certain types of eligible **costs** that were **actually incurred** for the project (not the budgeted costs).

=> see budget categories and cost eligibility rules in call text

- Consortium composition: minimum 3 independent beneficiaries from 3 different eligible countries
- Project duration: **24 to 36 months**
- Ideally one big project to be funded



Other Specifications and Requirements

- Work in close partnership with the **Data Spaces Support Centre** (<u>DSSC</u>):
 - <u>ensure alignment</u> with the rest of the ecosystem of data spaces implemented with the support of Digital Europe Programme (data spaces reference architecture; common building blocks, toolboxes and standards; and data governance models)
- Use, in so far as possible and when available, the smart cloud-to-edge middleware platform <u>Simpl</u>:
 - o proof-of-concepts released by summer 2024
 - o Minimum Viable Platform released end 2024
- active integration and participation of data holders and users
 - o ensure that the data space is designed and structured to meet participant's needs
 - make the data space <u>relevant</u> and aim for high adoption rates and a strong sense of ownership of data holders and users by the end of the project's runtime
- Work from the outset towards achieving financial sustainability by the
 conclusion of the project action



Eligibility Applicants (

- Public or private legal entities (natural persons not eligible)
- Established in eligible countries (EU27, EEA, countries associated to Digital Europe => <u>list participating countries</u>)
- <u>Restricted call</u>, application art.12(6) Digital Europe Regulation:
 - Entities must **not** be directly or indirectly **controlled from a country that is not an eligible country, unless** they comply with the requirements to **guarantee** the protection of the essential security interests of the Union and the Member States, and to ensure the protection of classified documents information as set out in the relevant work programme (conditions for foreign controlled entities – guarantees)
 - => Participants* to submit a declaration to **determine the ownership and control** status
 - * Beneficiaries and affiliated entities, associated partners and subcontractors except for entities that are validated as public bodies by the Central Validation Service
- ²⁰ Project activities can only take place in the eligible countries





A 2-step process:

- 1) Register your organisation in the Participant Register
 - => you will receive a 9-digit participant identification code (PIC).
- 2) Submit the proposal in 3 parts in the Funding & Tender Portal (topic page):
 - ✓ Part A: administrative information (applicant organisations and budget for the proposal).
 - Part B: description of the action with the technical content of the proposal (<u>maximum 70</u> pages)
 - ✓ Annexes: ownership control declarations (beneficiaries and affiliated entities, associated partners and subcontractors)
 - => Check call text and related documentation for exhaustive information





All in the topic page of the Funding & Tender Portal:

- Call text DIGITAL-2024-CLOUD-AI-06-GREENDEAL
- Digital Europe Work Programme 2023-2024
- EU Funding & Tenders Online Manual
- AGA Annotated Grant Agreement
- Guidance on participation in DEP restricted calls
- Topic related FAQ
- Funding & Tenders Portal FAQ (general)
- IT Help Desk
- Partner search announcements –





Funding & tender opportunities

Single Electronic Data Interchange Area (SEDIA)

The search funding & tenders ▼ How to participate ▼ PROJECTS & RESULTS WORK AS AN EXPERT SUPPORT ▼

European Green Deal Data Space

DIGITAL-2024-CLOUD-AI-06-GREENDEAL

Topic Call for proposal								
Internal navigation	<	General information						
General information		Programme Digital Europe Programme (DIGITAL)						
Topic description		Call						
Conditions and documents		Cloud, data and artificial intelligence (DIGITAL-2024-CLOUD-AI-06)						
Partner search announcements		Type of action DIGITAL-SIMPLE DIGITAL Simple Grants	Type of MGA DIGITAL Action Grant Budget-Based [DIGITAL-AG]					
Start submission		Deadline model	Opening date					
Topic related FAQ		single-stage	29 February 2024					
Get support								
Call updates		Topic description						

Thank you for your attention







GREEN DEAL DATA SPACE FOUNDATION & ITS COMMUNITY OF PRACTICE



GREAT | The Green Deal Data Space







- Technical tools for data pooling and sharing
- Standards & interoperability (technical, semantic)
- Sectoral Data Governance (contracts, licenses, access rights, usage rights)
- IT capacity, including cloud storage, processing and services



Green Deal Data Space

A federation of data ecosystems enabling policy makers, businesses, researchers and citizens, from Europe and around the world, to jointly tackle climate change.

- Duration: 20 Months
- Running: September 2022 April 2024
- **Consortium**: 11 Partners 3 Associated Partners
- Funding: Digital Europe Programme (CSA)









Community of Practice

Technical Blueprint

Governance & Business Models

High Priority Data Sets

Roadmap

Strategic EGD Actions



Community of Practice GDDS

16 Reference Use Cases & Initiatives







500+

Stakeholders







Technical Blueprint

GDDS Digital Ecosystem

The GDDS is characterized by a high level of heterogeneity, with many already existing data sharing initiatives that offer their resources to diverse consumers, which mirrors the current state of (geospatial) data sharing globally.

Establishing a single "common format" is not possible in a multidisciplinary context like GDDS.

The challenge is how to transform a collection of disparate systems that use different technical standards into a digital ecosystem. This requires a minimal set of logical components that enable the ecosystem's digital environment.







Governance

GOVERNANCE is the process for making decisions about an entity Choosing the questions that must be decided – "Requirements"



Business models

What is the value of a data space for ?

Data Providers

Data Consumer Data

Data Intermediaries

- A lawyer: regulatory enforcement, case laws
- A policy maker: Do I have evidence/ indicators to support targets?
- A scientist: data for research, publications
- Citizens: where shall I put my data?
- **SME/Public administrations**: easy data access for value added services, public services. An opportunity to share data in a cost-effective way
- **DS Community:** Consensus that value is good use cases

- Industry: increase competitiveness in data economy. I want my data everywhere – but still under my control! Standardised data license agreements.
- **Circular economy**: Share and exchange required information with partners and regulators.
- **Existing well established data initiatives** (e.g. EMODNet, GEO): Access to cross-domain data. Promotion of their solutions in the wider community. Tech support
- **Big techs**: Access to open data (freemium models)





Priority Datasets



•





Knowledge collection from:

- **Reference Use Cases and Initiatives**
 - Interviews with specific data users and providers
 - Collection of data requirements, products and gaps
 - Data Sets Inventory: list of specific data sets required by the RUCIs
- GREAT consortium and engaged stakeholders
 - Engagement with stakeholders from different domains related to the EGD
 - Data Services inventory: list of data service/portals/catalogues



Data Sets, Services Inventory and Gaps

Inventory

Services

Data



- RUCI's data requirements and products
- Close to 100 specific data sets
- Described by mandatory or optional tags
- Categorisation in High Value Data sets and Essential Variable categories
- Data types: satellite imagery, gridded data, topographic maps, model outputs, in situ data sets, cadastral data sets and other
- FAIRness assessment using <u>F-</u> <u>UJI</u>*



- Constantly expanding inventory of data services related to the Green Deal more than 400 data services
- Data service and access information
- Prioritisation based on:
- Relevance to RUCIs
- Relevence to strategic actions
- Relevance to EGD
 programmes
- Sustainability
- Coverage

ata Gaps

 \frown

- Data sets that do not exist!
- Data sets that cannot be accessed or can be accessed or only under specific contracts
- Data sets that cannot or are hard to use in combination with other data types due to standardisation and harmonisation issues
- Data sets whose quality cannot be guaranteed, or have evidently low quality
- Data sets that are hard to use openly because of ethical implications
- Data sets that exist in insufficient spatial and temporal resolution

Data Set Inventory



Principles

Support the creation and the successful implementation of the Green Deal Data Space.

From design to proof of concept, implementation and scaling-up of the Minimum Viable GDDS.

Expand on trials and pilots involving local, regional, national, European, and global initiatives.

Validate benefits of the GDDS to vertical and horizontal domains, public sectors, businesses and citizens. Copernicus Services ecosystem (including other EU long term data services e.g., EMODnet)

> **Destination Earth** ecosystem (including e.g., EU Digital Twin of the Ocean)

Five Clusters

To streamline these ambitions and provide structure for developing concrete plans, five clusters have been identified. They provide insights into potential future pilot projects:

Biodiversity

Zero Pollution

Climate Change

VERSION 1

GDDS STRATEGIC ROADMAP





Public Deliverables

https://www.greatproject.eu/



Initial Blueprint of the GDDS Reference Architecture



Governance Requirements

Prioritized Data Sets & Gaps



Roadmap









@GreenDealDS



@Green Deal Data Space

000



00

@greendealdataspace



Public Deliverables



Funded by the European





DATA SPACES

A use case in **FORESTRY DATA**

WEBINAR

Data spaces: experiences from the European Green Deal 15 March 2024

Abana Kona, DG JRC Unit T1 - Digital Economy Margherita Di Leo, Arcadia SIT, under contract with EC JRC Manlio Bacco, DG JRC Unit T1 - Digital Economy

JRC - Centre of reference for Data Spaces

Inclusive Data Governance

- Consent management tools
- Data intermediaries
- Data altruism

Common European Data Spaces

- Support on sectoral data spaces (e.g. DGs GROW, MOVE, AGRI, ENV, CNECT)
- Collaboration with EU Data Spaces Support Centre (DSSC)
- Privacy-preserving data analysis and sharing

Standardization

- A bridge towards Interoperability
- Collaboration & Agreements













Implementation tasks:

- INSPIRE MIG group and technical subgroup.
- Yearly Monitoring & Reporting.
- Operational support: Geoportal, Validator, Registry, Helpdesk.
- Modernisation & Evolution:
 - INSPIRE good practices.
 - Open Data alignment: High-Value Datasets.
 - Policy making support: GreenData4All / Green Deal Data Space



The Directive entered into force in 2007 / Roadmap finished by December 2021.

INSPIRE and new policy context Science for policy report



INSPIRE - A Public Sector Contribution to the European Green Deal Data Space

https://publications.jrc.ec.europa.eu/repository/handle/JRC126319

- Evolution to a data ecosystem (Green Deal Data Space).
- Broadening the scope:
 - New sectors: public, private/businesses, academia.
 - New communities: developers, users.
- Widening the range of applications and use cases.
- Making INSPIRE framework more **simple, flexible and agile**.
- Lowering the **knowledge entry-level** for implementing and/or using data.
- Reusing well-adopted standards and technologies.



Privacy-preserving data analysis and sharing

Report evaluating privacy preserving techniques

- Fit for data spaces and digital twins
- Handling **sensitive** data (personal and business)
- Data intermediaries as potential trusted parties

Experimental work on Federated Learning with Differential Privacy

- Data never leave the source (privacy)
- Privacy vs utility (accuracy) tradeoff
- Interplay of the Data Act, Common European Data Spaces





A use case on forestry data (prototype)

Forest data collected by National Forest Inventories (NFIs) are used for a wide range of purposes:

- Assess the health and condition of forests, monitor changes in forest cover, and plan sustainable forest management practices.
- Information on the diversity of tree species, ecological processes, and biodiversity conservation.
- Estimating the **amount of carbon stored in forests**, for understanding the role of forests in mitigating climate change.
- Volume, quality, and distribution of timber and non-timber forest products, for forest industries.
- Governments and policy-makers use NFI data to develop policies related to forest conservation and natural resource management.
- Scientific research to understand the dynamics of forests and their ecosystems.
- **Reporting on national and international commitments** related to forest conservation, such as the UN Framework Convention on Climate Change (UNFCCC) and the Convention on Biological Diversity (CBD)







- Plot data is **collected in the field**, data collection is costly and not feasible for vast areas.
- NFIs collect forest data but it is subject to privacy constraints and not disclosed.
- Terms and conditions for its use vary according to country.
- At the regional level, it is necessary to integrate these measurements with **remote sensing data**.



Privacy vs. Utility trade-off



Pictorial representation of the privacy versus utility tradeoff. Different solid lines represent different data perturbation approaches for releasing private information © Fang Liu, 2007

- Data can be degraded before being disclosed, coordinates can be shifted and noise can be added, but makes it less useful for using in modelling, especially in remote sensing applications.
- The more privacy is preserved, the less accuracy is achieved in modelling.



Deriving biomass (AGB) maps from remote sensing Vegetation Indices (VIs)













		geometry	х	У	AGB	SR	NDVI	NDWI	GNDI	NDI45	SAVI
	0	POLYGON ((585140.000 4405760.000, 585139.904 4405758.040, 585139.616 4405756	585120.0	4405760.0	228.0	0.331338	0.502250	0.194038	0.488249	0.148449	0.251101
	1	POLYGON ((577800.000 4426580.000, 577799.904 4426578.040, 577799.616 4426576	577780.0	4426580.0	167.0	0.646074	0.215024	-0.088917	0.277205	0.071472	0.107500
	2	POLYGON ((548820.000 4453100.000, 548819.904 4453098.040, 548819.616 4453096	548800.0	4453100.0	101.0	0.651240	0.211212	-0.104627	0.253877	0.065685	0.105594
	3	POLYGON ((533860.000 4433020.000, 533859.904 4433018.040, 533859.616 4433016	533840.0	4433020.0	6.0	0.425859	0.403023	0.140343	0.384979	0.138316	0.201490
	4	POLYGON ((584720.000 4470060.000, 584719.904 4470058.040, 584719.616 4470056	584700.0	4470060.0	0.0	0.495626	0.338183	0.063042	0.349023	0.136099	0.169077
	9995	POLYGON ((529280.000 4474180.000, 529279.904 4474178.040, 529279.616 4474176	529260.0	4474180.0	79.0	0.524072	0.312648	-0.028250	0.340110	0.093602	0.156308
	9996	POLYGON ((536680.000 4480340.000, 536679.904 4480338.040, 536679.616 4480336	536660.0	4480340.0	139.0	0.407100	0.421907	0.093134	0.425397	0.173195	0.210935
	9997	POLYGON ((533420.000 4483860.000, 533419.904 4483858.040, 533419.616 4483856	533400.0	4483860.0	158.0	0.594102	0.254627	-0.011723	0.280035	0.074662	0.127300
	9998	POLYGON ((554720.000 4443360.000, 554719.904 4443358.040, 554719.616 4443356	554700.0	4443360.0	177.0	0.603680	0.247254	-0.050848	0.295968	0.093422	0.123613

Sample dataset

Machine learning models

- Multiple Linear Regression (MLR)
- Random Forest Regression (RFR)

. . .

• Gradient Boosting Regression (GBR)



Prediction (prototype)



Biomass prediction



AGB - reference GEDI

1000

0

2000

3000

4000

5000

AGB - GBR model





Federated Learning (FL)

- FL enables training models across decentralized devices or servers holding local data samples, without the need to transfer the data to a central location.
- The model is trained locally on each device using its own data, and only the model updates are shared with a central server or aggregator.
- The server then combines the updates from multiple devices to improve the global model, which is then sent back to the devices for further training.
- This approach allows for privacy-preserving training, as the raw data remains on the local devices and only model updates are exchanged.



European Commission



- Adds noise to the data to ensure that no specific record's information can be discerned from the output, while still providing accurate and useful aggregate information for the analysis.
- Can be combined with **federated learning to maximise privacy**.



Thank you



© European Union 2024

Unless otherwise noted the reuse of this presentation is authorised under the <u>CC BY 4.0</u> license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.



EU Science Hub Joint-research-centre.ec.europa.eu



Q&A



Carlos Martins Ferreira Knowledge Management, Data.europa.eu



Sotirios Kanellopoulos emerging ICT & data for environmental sustainability, statistical learning specialist Nevena Raczko Senior Consultant at IDC's European Government Consulting unit



Albana Kano Scientific Officer at the European Commission, Joint Research Centre



Margherita Di Leo Senior Consultant at European Commission, Joint Research Centre



Stay up-to-date on our upcoming activities

Unin. CUropo academy

Join our next webinar!

WEBINAR

Data sovereignty & data governance: how can data holders stay in control?

doto. europo academy

22 March 2024

 $10.00 - 11.30 \, \text{Cet}$



doto. europo academy

Please provide your feedback!







Thank you



