



D6.2

First Policy Brief

Status: Under EC Review

Dissemination Level: Public



Funded by
the European Union

This project has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement No 101188168.

Abstract

Keywords

Research Infrastructures, Data Exploitation Platform (DEP), Scalable AI, FAIR Data, e-infrastructures, EOSC, EuroHPC, Data Spaces, Virtual Access, Research Infrastructure Policy

Europe's Research Infrastructures (RIs) have become "big data factories", yet the scale and complexity of their datasets mean most of this publicly funded data cannot be reused at scale. Researchers, SMEs and even RI operators lack the ready-to-use computing environments, AI tools and skills needed to exploit it. The RI-SCALE project addresses this gap by co-designing an open, modular Data Exploitation Platform (DEP), a middleware that connects RI data holdings to large computing centres and pre-configured AI. The first DEP release delivers baseline data-lifecycle management, AI-based data processing and access control, and is being validated across four RIs (EISCAT_3D, BBMRI, Euro-BioImaging ERIC and ENES) and two national compute centres. Drawing on this first-period experience, this brief sets out recommendations to European and national decision-makers. Its central message: public funding should support not only the building of such services but their transition to sustained, shared operation.

Revision History			
Version	Date	Description	Author/Reviewer
0.1	06/07/2026	First Version	Smitesh Jain (EGI Foundation)
0.2	09/07/2026	Version shared with AMB	Smitesh Jain (EGI Foundation)
0.3	20/07/2026	Second Version	Smitesh Jain (EGI Foundation)
0.4	18/08/2026	External Review	Gergely Sipos (EGI Foundation) Ville Tenhunen (EGI Foundation)
1.0	27/08/2026	Final Version	Smitesh Jain (EGI Foundation)

Document Description			
D6.2 – First Policy Brief			
Work Package Number 6			
Document Type	Report		
Document Status	Under EC Review	Version	1.0
Dissemination Level	Public		
Copyright Status	 This material by Parties of the RI-SCALE Consortium is licensed under a Creative Commons Attribution 4.0 International License.		
AI Declaration	The authors acknowledge the use of GenAI chatbots for (non-substantial use), e.g. improving the readability and language of the text, as a basic support tool. The authors carefully reviewed and edited the proposed changes as needed and take full responsibility for the content of the publication, including the critical analysis and conclusions they developed.		
Lead partner	EGI Foundation		
Document Link	https://documents.egi.eu/document/4288		
DOI	https://zenodo.org/records/22208758		
Author(s)	Smitesh Jain (EGI Foundation)		
Reviewers	Gergely Sipos (EGI Foundation) Ville Tenhunen (EGI Foundation)		
Moderated by:	Matteo Agati (EGI Foundation)		
Approved by:	Activity Management Board		

Terminology / Acronyms	
Term/Acronym	Definition
AI	Artificial Intelligence
BBMRI-ERIC	European Research Infrastructure for Biobanking and Biomolecular Resources
DEP	Data Exploitation Platform
EHDS	European Health Data Space
EISCAT	European Incoherent Scatter Radar System
ENES	European Network for Earth System Modelling
EOSC	European Open Science Cloud
Euro-Biolmaging ERIC	European Research Infrastructure for Imaging Technologies in Biological and Biomedical Sciences
EuroHPC	European High-Performance Computing
GPU	Graphics Processing Unit
HPC	High-Performance Computing
RI	Research Infrastructure
TRL	Technology Readiness Level
TUBITAK	Scientific and Technological Research Council of Türkiye
TUW	Vienna University of Technology

Table Of Contents

1. Project Overview and Objectives..... 7

 1.1. Highlights..... 7

2. Introduction and Context..... 9

3. Policy Developments and Recommendations..... 10

List of Figures

- [Figure 1: DEP Technology connecting a Data Holding and a Compute Centre](#)

1. Project Overview and Objectives

CALL: HORIZON-INFRA-2024-TECH-01

TOPIC: HORIZON-INFRA-2024-TECH-01-01

PROJECT: Unlocking RI potential with Scalable AI and Data (RI-SCALE). www.riscale.eu

PROJECT OBJECTIVES: Empower big data-producing Research Infrastructures (RIs) with the technical means to extend their services with scalable computational platforms suited for the development and delivery of AI applications to improve, analyse and exploit their data at large. Co-design, prototype and validate this Data Exploitation Platform technology through diverse use cases from environmental and health sciences in collaboration with relevant Data Spaces. Specific objectives include:

- Co-design and develop an RI ‘Data Exploitation Platform’ (DEP) technology prototype.
- Increase the AI-based data exploitation and data mining capacity of RIs with AI frameworks and AI models integrated into the DEP framework.
- Demonstrate and validate the practical application of the AI-powered DEPs through RIs, Data Spaces and use cases that contribute to the EU policy objectives in environmental and health sciences.
- Increase the industry and university collaborations of data-intensive RIs in the area of scalable AI applications.
- Facilitate the uptake and use of AI technologies for scalable RI data analysis through competence centres, consultancy events and training.

Name and email address of the operational project coordinator: Gergely Sipos
<gergely.sipos@egi.eu>

1.1. Highlights

- **The Problem.** Research Infrastructures (RIs) have become “*big data factories of the scientific world*”, yet the size and complexity of the datasets mean most data cannot be reused at scale: researchers, SMEs and even RI operators lack the ready-to-use compute environments, AI expertise and data wrangling skills needed to exploit them.
- **Results.** RI-SCALE has released the first version of an open, modular Data Exploitation Platform (DEP) that pairs RI data with scalable computing and pre-configured AI services.
- **Lead Recommendation.** Support shared, reusable, open-standard data-exploitation middleware as a standard RI capability and back the move from prototype to sustained operational service. A common DEP approach will turn underused FAIR data holdings into

AI-ready resources, widen access for SMEs and smaller RIs through credit-based virtual access.

2. Introduction and Context

The overall objective of the RI-SCALE project is to co-design, prototype and validate this Data Exploitation Platform (DEP) technology. The DEP will enable Research Infrastructure (RI) data holdings to expand their services with an AI-ready compute arm by partnering with external compute facilities where the RI data is replicated and served for user analysis. A DEP installation fundamentally acts as an extension of an RI, connecting its data holdings to external compute centres and offering online analytical services for data exploitation. In the RI-SCALE project, the DEP will be specifically designed for the support of AI-based analysis.

Within the RI-SCALE project, the DEP technology will be validated across four RIs (EISCAT_3D, BBMRI, Euro-Biolmaging ERIC, and ENES) and with two national compute centres (TUW from Austria and TUBITAK from Turkey).

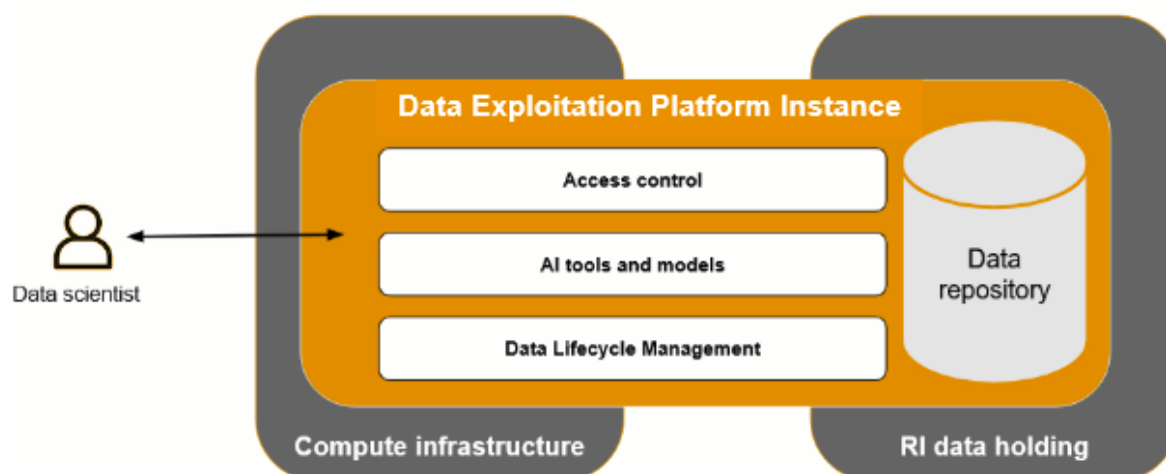


Figure 1: DEP Technology connecting a Data Holding and a Compute Centre

3. Policy Developments and Recommendations

Implementation of Research Infrastructures: The INFRA-TECH calls fund new research-infrastructure services only to demonstration maturity (around TRL 5-6); there is then no obvious follow-up instrument to operationalise those services and embed them in the RIs' own service portfolios. RI-SCALE illustrates the gap: The first DEP release¹ delivered a baseline functionality for data lifecycle management, AI-based data processing and access control, built entirely on open standards, APIs and protocols. Its modular design lets each RI adopt only the components it needs.

Recommendation: Create a follow-up funding route that carries proven research-infrastructure services from mid-TRL demonstration to operational deployment and into the RIs' permanent portfolios, complementing INFRA-TECH development calls. Without such support, each RI may end up rebuilding its own stack, leading to duplicated public spending, incompatible silos, and lock-in to single vendors or single RIs.

Access to Research Infrastructures: RI-SCALE's credit-based access mechanism will track usage across compute, storage, data transfer and AI resources, laying the basis for transnational and virtual access to DEPs; a minimum viable version is scheduled for demonstration later in the project.

Recommendation: Recognise and fund credit-based virtual access to digital services and computing as a complement to physical access. The European Charter for Access to Research Infrastructures does not yet address compute-credit-based access and could be extended to cover it. Without such mechanisms, Europe's data holdings may stay accessible mainly to those who already have computing.

Employment and skills in research infrastructures: A distributed Competence Centre delivers workshops, webinars and hands-on training to upskill RI staff and users in scalable AI and trusted data handling, using a train-the-trainer model aimed at RI national nodes so that expertise spreads beyond the consortium.

Recommendation: Fund a permanent cross-RI competence-centre network for AI and data stewardship as recurrent operational infrastructure, mirroring the EuroCC National Competence Centres model used for HPC, so the skills gap is closed durably rather than in project-length bursts.

Greening of research infrastructures (Level: operation and access): The access credits carry an energy and CO₂ green index, so that the greener computing option is the rewarded one; in parallel, energy-efficient AI accelerators are being benchmarked against GPUs for model inference.

Recommendation: Require standardised energy and CO₂ reporting for every publicly funded computing allocation, adopt a common green-index methodology, and let green credits count towards access entitlements.

¹ D5.3 – First DEP Release: <https://doi.org/10.5281/zenodo.18839895>

Technology development, data and digital services, digitalisation: The DEP is assembled from open components already used by RIs and e-infrastructures for many years rather than newly invented technology.

Recommendation: Prioritise reuse and hardening of established open components over new builds. Require provenance records for AI models served through RIs. Align with the common software layer planned on the EuroHPC, AI Factories and Data Labs side. Prioritise open, modular, interoperable building blocks over vertically integrated silos in RI technology funding.

Level of connection of the RIs to EOSC: The DEP serves FAIR data holdings for large-scale computation and is being validated for interoperability with EOSC and with three Data Spaces, each in a defined vertical - EUCAIM (cancer imaging), Destination Earth (climate) and the Copernicus Data Space Ecosystem (Earth observation). RI-SCALE integrates these initiatives using federated and decentralised identity for secure cross-domain access. This directly tests how RIs, EOSC and the Data Spaces can work as one ecosystem.

Recommendation: Align Data Space architectures, protocols and middleware so that platforms such as the DEP interoperate by design rather than through bespoke integrations. Give one body a clear mandate for a common authentication and authorisation profile across RIs, e-infrastructures, EOSC and EuroHPC.

Contribution to other research areas and to broader EU priorities: RI-SCALE contributes to the European Health Data Space (EHDS), the European Green Deal and the EU's trustworthy-AI agenda. A concrete bottleneck has emerged in the health domain: the DEP is designed to handle sensitive as well as open data, but the federated data-management layer cannot read directly from the Trusted Research Environments (TREs) used by hospitals and biobanks, so patient data must first be copied into less-protected storage before it can be processed.

Recommendation: Rather than bridging individual TREs one by one, treat Europe's TREs as a federation and connect the DEP's federated computing to it as a 'federation of federations,' so that computation moves to the data rather than the data leaving the safe environment; Co-design the governance models for cross-federation access in parallel with the technical infrastructure, not as an afterthought; and, because the EHDS Regulation mandates Secure Processing Environments (SPEs) for the secondary use of health data, reflect the same interoperability requirement in the SPE specifications being developed under EHDS implementation and fund it through the EU's health-data-infrastructure and RI programmes.