EUDAT CDI

The European e-infrastructure
of interoperable data services
supporting research

EUDAT Collaborative Data Infrastructure

*With a network of more than 20 European research organisations, data and computing centres in 14 countries, the EUDAT **Collaborative Data Infrastructure** (CDI) is the largest sustainable **infrastructure of integrated data services and resources supporting research in Europe**.*

—— It is designed to address the **full lifecycle of research data**, representing a strategic solution to the challenge of data proliferation in Europe's scientific and research communities.

—— It is realised through an ongoing **collaboration between Service Providers and Research Communities** working as part of a common framework for developing and operating an interoperable layer of common data services.

—— It is a **growing organisation based on a contractual agreement** between its members, signed for an initial period of 10 years.

Nationally and globally, the CDI:

—— **Supports research** providing a trustworthy, collaborative and interoperable data infrastructure.

—— **Assists researchers and research organisations** with essential training and consultancy on Research Data Management.

—— **Enhances** the integration between HPC and the storage infrastructure.

—— **Offers** additional business opportunities to the service providers part of the CDI.

—— **Builds the backbone** of the European Open Science Cloud (EOSC).

EUDAT Collaborative Data Infrastructure

Data shared and preserved across borders and disciplines

The **EUDAT CDI**
allows **two types of involvement**:

*research communities, data and other service providers can **use** the CDI services while trusted data repositories, computing and data centres can **join** the CDI **as generic or thematic service providers**.*

Users

Users are any researcher, research group or collaborative research organisation making use of the CDI services. Users may act on their own behalf, or on behalf of a defined User Community.

> ❝❝ Any researcher or research group (organisations or individuals) may use the integrated CDI data services

Why choose the EUDAT CDI?

- ✔ To benefit from a simple, efficient, trustworthy, affordable collaborative and interoperable data infrastructure
- ✔ To connect with Europe's most powerful supercomputers (Users interact with the EUDAT CDI through one or more of its nodes)
- ✔ To leverage on the expertise of renowned experts in data management
- ✔ To shape the data management services of the future (The CDI services are developed involving users in all phases of the service definition, implementation, configuration and operation lifecycle)

Service Providers

The CDI service providers, also commonly referred to as "CDI nodes", are generic and thematic trusted data repositories, computing and data centre providing services and resources in accordance with the CDI Service Management Framework (SMF).

> ❝❝ Joining the EUDAT CDI is your opportunity to offer the CDI Services

Why join the EUDAT CDI?

- ✔ To offer your local customers off-the-shelf solutions by deploying and running EUDAT services
- ✔ To be interoperable with other service providers and meet requirements of users & funders for increased cross-national collaborations
- ✔ To participate in new H2020 projects as EUDAT node and leverage new collaborations & funding opportunities
- ✔ To be recognized as a European leader in data management and data infrastructure services
- ✔ To be part of the CDI governance and influence the evolution of the European infrastructure landscape

Data shared and preserved across borders and disciplines

3

Data Access

Data storage

Data discovery and metadata

Data management

Authentication & Authorisation

Service management infrastructure

RDM Training & Consultancy

CDI Services

Become a user

The CDI users are research communities, data and other service providers willing to benefit from **integrated data services and resources.**

Who can access the CDI services

All individual researchers or research organisations making use of the free CDI services are referred as **Users**. **Customers**, instead, are those who request specific services from a Service Provider under certain conditions. A customer could also be a user but, if they are not, it means that the customer will purchase the services for the final user. Fees may apply for any of the CDI Services provided by a Service Provider - the amount of the applicable service fee is upon the Service Provider.

What's the added value of the CDI services?

— The EUDAT CDI services are developed according to a **service-centric strategy based on service co-design, involving users in all phases of the service definition, implementation, configuration and operation lifecycle**. Current services have been developed in close collaboration with **over 50 research communities** spanning across many different scientific disciplines and involved at all stage of the design process.

— The EUDAT CDI follows a **technology - agnostic strategy**, developing services and operational processes that can be realised across a range of complementary technologies. EUDAT promotes **harmonisation of service interfaces and underlying processes** as a way to maximise technical interoperability.

— The EUDAT CDI suite of data management services supports researchers and research communities to ensure their data is **FAIR compliant** and **promote open access to data.**

— EUDAT does not assert any rights over any of the data it holds, and **all data stored within the CDI retain their original rights**.

— EUDAT **promotes the release of data under standard** licences, with the long-term aspiration that all data in the CDI should become fully open access. **Access to data in the CDI is free at the point of use**.

The CDI service suite

The CDI services cover data access, data storage, data discovery and metadata, persistent identification, data management, authentication & authorisation, service management infrastructure and Research Data Management (RDM) training & consultancy.

The CDI service catalogue

The EUDAT CDI Services are split into seven categories: data access, data storage, data discovery and metadata, persistent identification, data management, authentication & authorisation, service management infrastructure and Research Data Mangement training & consultancy.

SOLUTIONS FOR	SERVICE	MAIN FEATURES
Data Access	B2STAGE	Data staging service to access both large-scale data storage and HPC systems
	CDI Gateway API	-
Data Storage	B2SHARE	Repository for shareable digital objects
	B2DROP	Secure and trusted cloud storage
Data Discovery and Metadata	B2FIND	Multidisciplinary metadata catalogue
	Metadata Store	-
	Data Type Registry	-
Persistent Identification	B2HANDLE	Tools for managing persistent identifiers
Data Management	B2SAFE	Data replication and long-term preservation service
Authentication & Authorisation (AAI)	B2ACCESS	Federated and multi-protocol identity provider
Service management infrastructure	Service Hosting	-
Research Data Management training & consultancy	Consultancy service & Training services	-

The EUDAT CDI will expand the current service catalogue over time by developing new technologies and services.
For more information about the CDI services visit eudat.eu/services-support

Data shared and preserved across borders and disciplines

The Service Management Framework (SMF)

The SMF is the document defining the scope, the service management related to the deployment and operation of CDI Services, the qualifications required by any Service Provider to be considered an integrated or interoperable CDI Service Provider, as well as roles and responsibilities in operating and/or delivering CDI Services.

The CDI Agreement:

This EUDAT CDI Agreement is a mutual collaboration framework between the CDI Service Providers. The Agreement is initiated within the framework of EUDAT, a research data infrastructure established by the EUDAT2020 Horizon 2020 Grant Agreement and EUDAT2020 Consortium Agreement.

How to join

Send a request to **eudat-cdi-secretariat@postit.csc.fi** specifying if you want to apply as an interoperable or as an integrated service provider

—— Become a service provider

*The CDI service providers (or "CDI nodes") are **generic and thematic trusted data repositories, computing and data centres** that have committed to provide the CDI services in accordance with the CDI Service Management Framework (SMF).*

- —— Any organisation can join as a service provider, as long as it **satisfies a set of technical requirements**. A service provider can join as an **interoperable or an integrated node** depending on the kind of integration the single organisation wants to pursue.

- —— **Membership needs to be approved by the CDI Council** and organisations can become members once they have signed the EUDAT CDI Collaboration Agreement.

- —— **Each service provider is responsible of covering the costs of the CDI services it delivers and has to pay an annual fee**. The amount of the fee depends on the type of integration and range from €5K to €20K.

- —— By joining the EUDAT CDI, **the organisation agrees to comply with the EUDAT CDI model and service management framework**.

—— Thematic vs Generic service provider

Organisations can join the CDI either as generic Service Providers or thematic Service Providers:

- —— Generic Service Providers have regional, organizational or national mandates to support scientific research, usually covering different disciplines;

- —— Thematic Service Providers are (discipline-specific) organisations mandated to support a well-defined scientific community or group of customers and users.

—— Interoperable vs integrated

Joining the EUDAT CDI can also be done in two ways that involve different levels of integration between the centre and the rest of the EUDAT CDI, as well as and different levels of responsibility:

- —— Integrated Service Providers
- —— Interoperable Service Providers

Interoperable or Integrated?

A matter of pre-requisites & benefits

Service Providers are connected to the CDI using a **CDI compliant approach** for data access and discoverability. On the basis of the pre-requisites that they can meet, service providers can be classified as Interoperable or Integrated. Different levels of integration between the service provider and the rest of the EUDAT CDI can bring different benefits.

PRE-REQUISITES

INTEROPERABLE NODES

— **must have a data repository** in which they preserve or curate data from a single research community or host data from several research communities or experiments (storage and access services).

— must identify the data hosted in the repository via some form of **persistent identifier** (persistent identification).

— it **must be possible to harvest and discover** the associated **metadata** through EUDAT's B2FIND service.

INTEGRATED NODES

— **must be interoperable nodes**.

— must **integrate** their local data infrastructure with the CDI's data management services.

— must provide **a common data access layer**.

— must integrate with a common **authentication & authorization** infrastructure (AAI).

— must **connect** their services to the common CDI service management infrastructure and operate it according to the service management framework (SMF).

BENEFITS

— **Better discoverability** of data hosted at the site.

— **Interoperability** with other service providers through CDI compliant repository.

— Possibility to **deploy and run** the **EUDAT services** and tools for local customers with support from central staff.

— **Participation to the CDI governance** and in new H2020 projects as EUDAT CDI node and leveraging new funding opportunities.

— In addition to the benefits listed for the interoperable nodes, integrated nodes benefit from **a tighter integration** with other service providers, offering seamless integration of resources to pan-European communities.

Data shared and preserved across borders and disciplines

The EUDAT CDI:

A key player for the implementation of the European Open Science Cloud

The establishment of the EUDAT CDI is timely with the imminent realization of the **European Open Science Cloud** which aims to offer open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines.

> *The European Open Science Cloud (EOSC) for Research will offer the 1.7 million European researchers and 70 million professionals in science and technology a virtual environment with free at the point of use, open and seamless services for storage, management, analysis and re-use of the data that are linked to their research activities, across borders and scientific disciplines.*

In February 2016, EUDAT signed a joint statement with LIBER (Association of European Research Libraries www.libereurope.eu), OpenAIRE (Open Access Infrastructure for Research in Europe www.openaire.eu), EGI (International Grid & Cloud Infrastructure for Research www.egi.eu) and GÉANT (European research and education networking collaboration www.geant.org) illustrating the shared vision of these organisations for the European Open Science Cloud (EOSC) for Research.

The statement details the eight key elements to the successful establishment of the Open Science Cloud, namely that it be open, publicly funded & governed, research-centric, comprehensive, diverse & distributed, interoperable, service-oriented and social.

The eight key points identified in 2016 are all reflected in the recent EOSC declaration signed by Director-General Robert-Jan Smits (Directorate-General for Research & Innovation) during the European Open Science Cloud Summit in Brussels on 12 June 2017 .

> *The EUDAT CDI services and structure are compliant with all these elements and are further evolving in that direction.*

In January 2018, the EUDAT CDI starts working with a consortium of over 74 beneficiaries including Research Infrastructures, national e-Infrastructure providers, SMEs and academic institutions within the EOSC-Hub project, a Research and Innovation Action funded by the European Commission (call H2020-EINFRA-12-2017(a)).

EOSC-hub creates the integration and management system (the Hub) of the future European Open Science Cloud and the EUDAT CDI is a key part of it.

---

[1] http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

[2] https://zenodo.org/record/32915#.WX9BgelLflU

[3] https://webcast.ec.europa.eu/european-open-science-cloud-summit

Data shared and preserved across borders and disciplines

__The CDI members (March 2020)

Generic, Integrated Service Providers

1. CSC - IT Center for Science (CSC) - FI
2. CINECA, consortium of universities (CINECA) – IT
3. Barcelona Supercomputing Center (BSC) - ES
4. Science and Technology Facilities Council (STFC) – UK
5. SURFsara Bv (SURFsara) – NL
6. Karlsruhe Institute of Technology (KIT) - DE
7. Max Planck Computing and Data Facility (MPCDF) – DE
8. Cines - National Computing Center for Higher Education (CINES) – FR
9. Greek Research and Technology Network S.A. (GRNET) – GR
10. Jülich Research Centre (JÜLICH) - DE
11. Institute of Bioorganic Chemistry Polish Academy of Sciences (IBCH PAS) - PL

Generic, Interoperable Service Providers

1. Uninett Sigma 2 As (SIGMA) - NO
2. Data Archiving and Networked Services (DANS) - NL
3. EUDAT Ltd. - FI
4. JISC LBG (JISC) – UK
5. Swedish National Infrastructure for Computing (SNIC) - SE
6. University of Edinburgh – UK
7. Danish eInfrastructure Cooperation DeIC –DK
8. The National Scientific Computing Unit of the Foundation for Science and Technology (FCTIFCCN) - PT
9. Gesellschaft fur wissenschaftliche Datenverarbeitung mbH Gottingen (GWDG) - DE
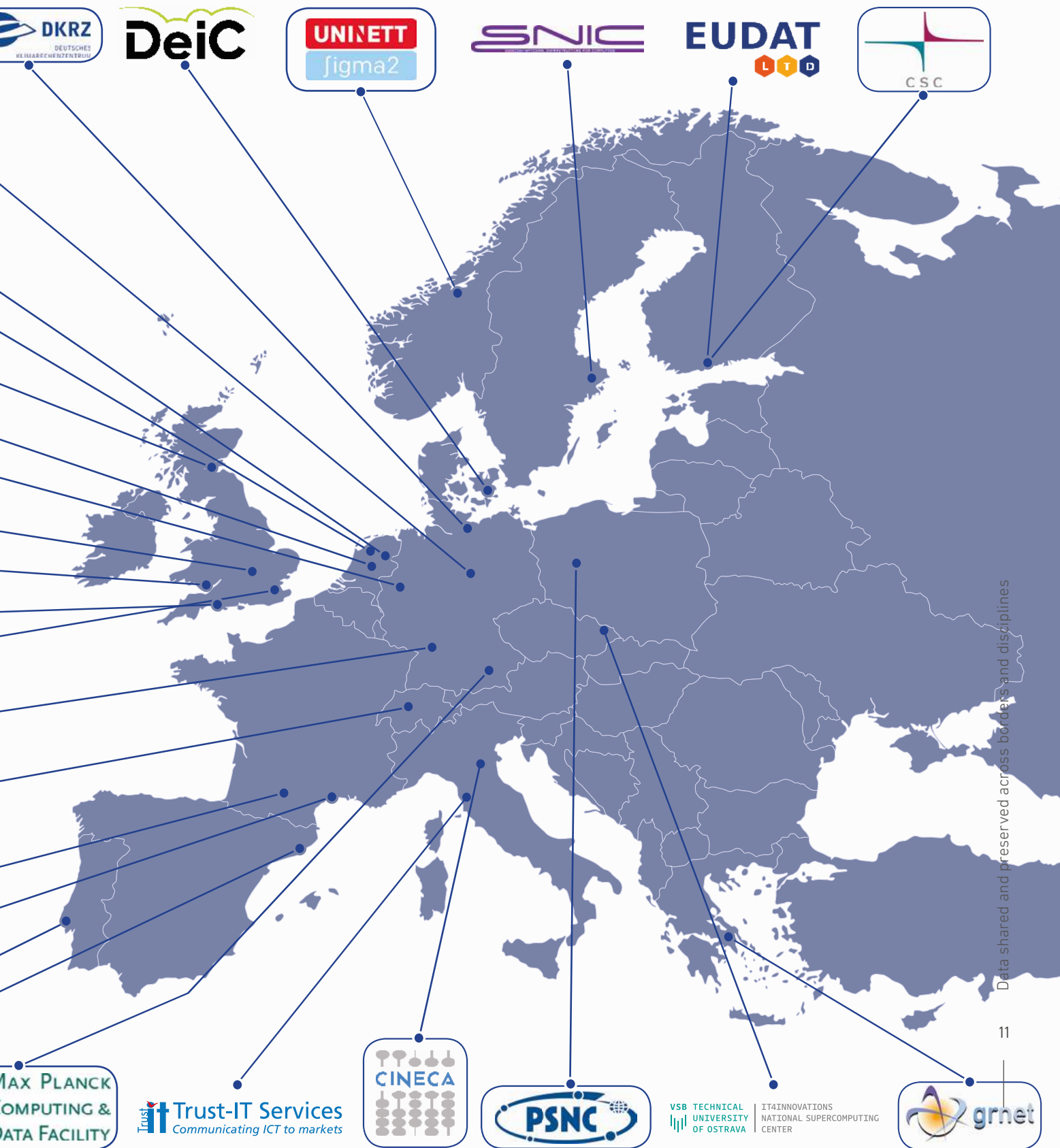10. IT4Innovations - CZ
11. Trust-IT Services - IT

Thematic, Integrated Service Providers

1. German Climate Computing Centre (DKRZ) - DEz
2. University College of London (UCL) - UK

Thematic, Interoperable Service Providers

1. European Organization for Nuclear Research (CERN) - INT
2. Culham Centre for Fusion Energy (CCFE) – UK
3. CLARIN-ERIC - NL
4. European Center in Research and Advanced Training on Scientific Computing (CERFACS) - FR

DKRZ
DEUTSCHES
KLIMARECHENZENTRUM

DeiC

UNINETT
ſigma2

SNIC

EUDAT
L T D

C S C

Data shared and preserved across borders and disciplines

MAX PLANCK
COMPUTING &
DATA FACILITY

Trust-IT Services
Communicating ICT to markets

CINECA

PSNC

VSB TECHNICAL
UNIVERSITY
OF OSTRAVA

IT4INNOVATIONS
NATIONAL SUPERCOMPUTING
CENTER

grnet

Founding signatories | Sept 2016

_ Getting involved

- Use the EUDAT CDI Research Data Management services www.eudat.eu/services-support

- Integrate & Deploy the EUDAT CDI services for your community
  www.eudat.eu/integrate-deploy-rdm-services-for-my-community

- Become a CDI service provider www.eudat.eu/become-an-eudat-service-provider

EUDAT Collaborative Data Infrastructure

www.eudat.eu

🐦 Eudat_eu