EUDAT & SeaDataCloud

SeaDataCloud Kick-off meeting

Damien Lecarpentier
CSC-IT Center for Science

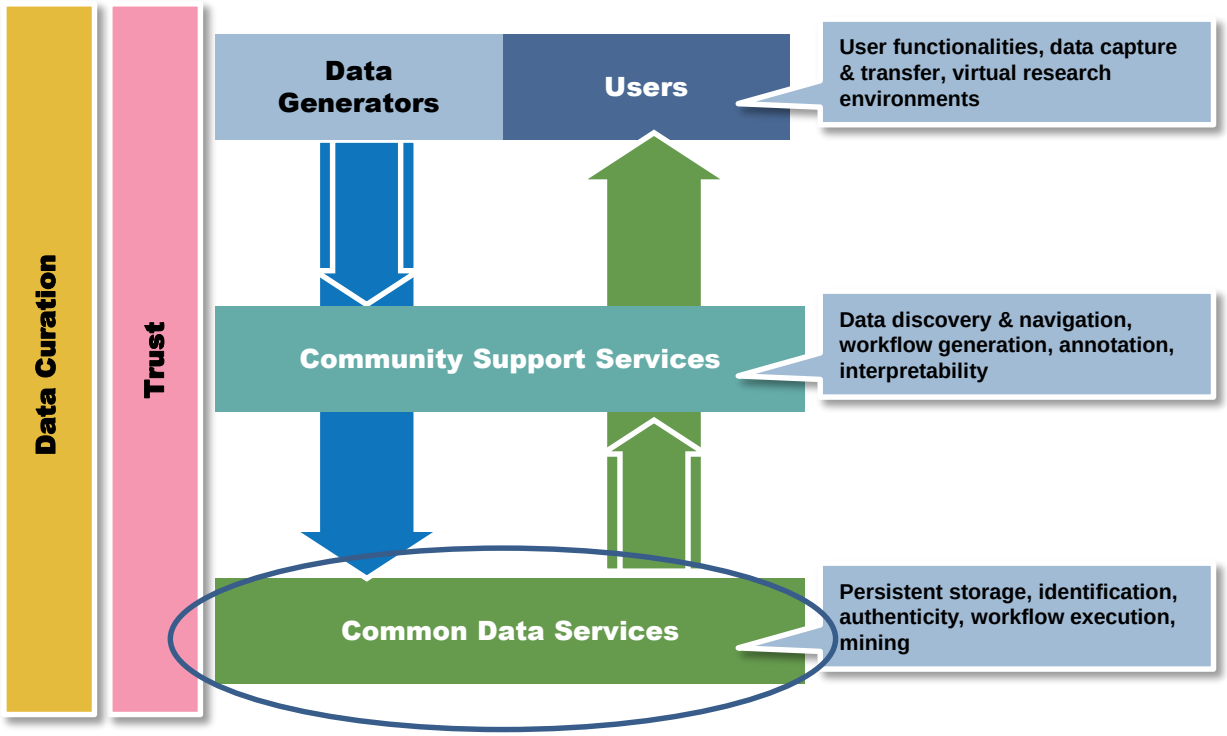
www.eudat.eu

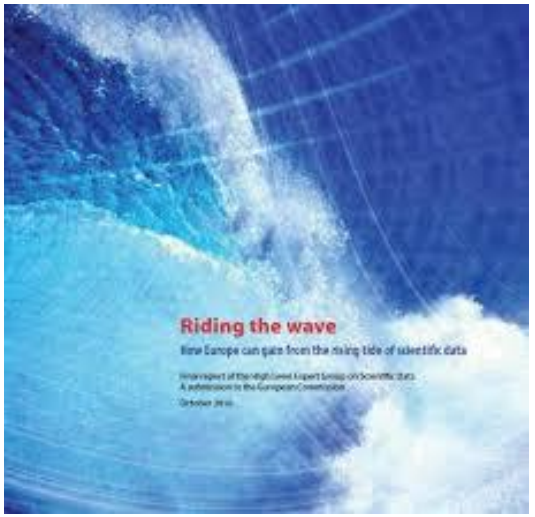
EUDAT receives funding from the European Union's Horizon 2020 programme - DG CONNECT e-Infrastructures. Contract No. 654065

EUDAT – European Data Infrastructure

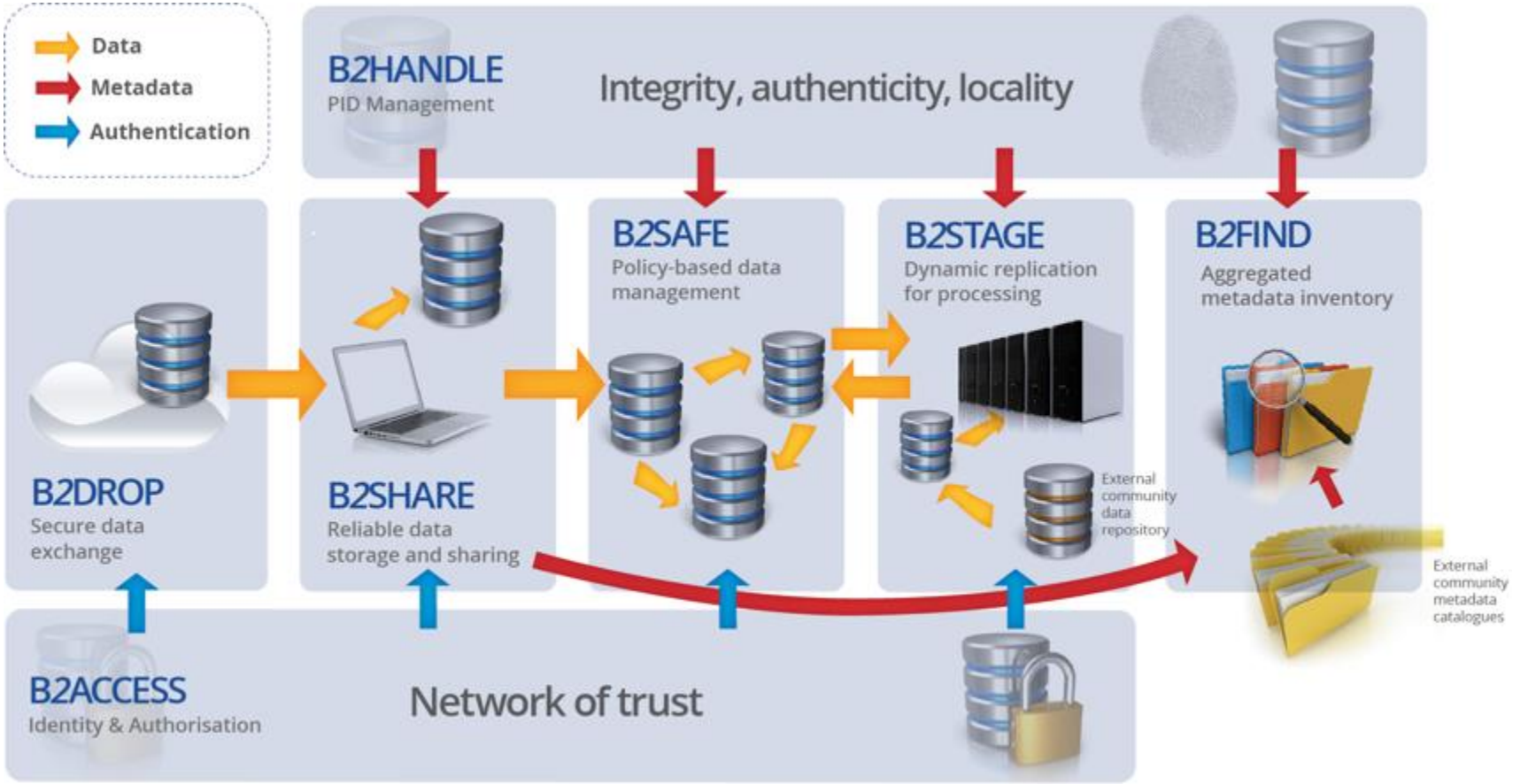
A consortium of high performance computing (HPC) / data centres, libraries, scientific communities, data scientists

Collaborative Data Infrastructure





What is the EUDAT Service offer?



B2FIND Find Research Data

B2STAGE Get Data to Computation

B2HANDLE Register your Research Data

B2SHARE Store and Share Research Data

B2SAFE Replicate Research Data Safely

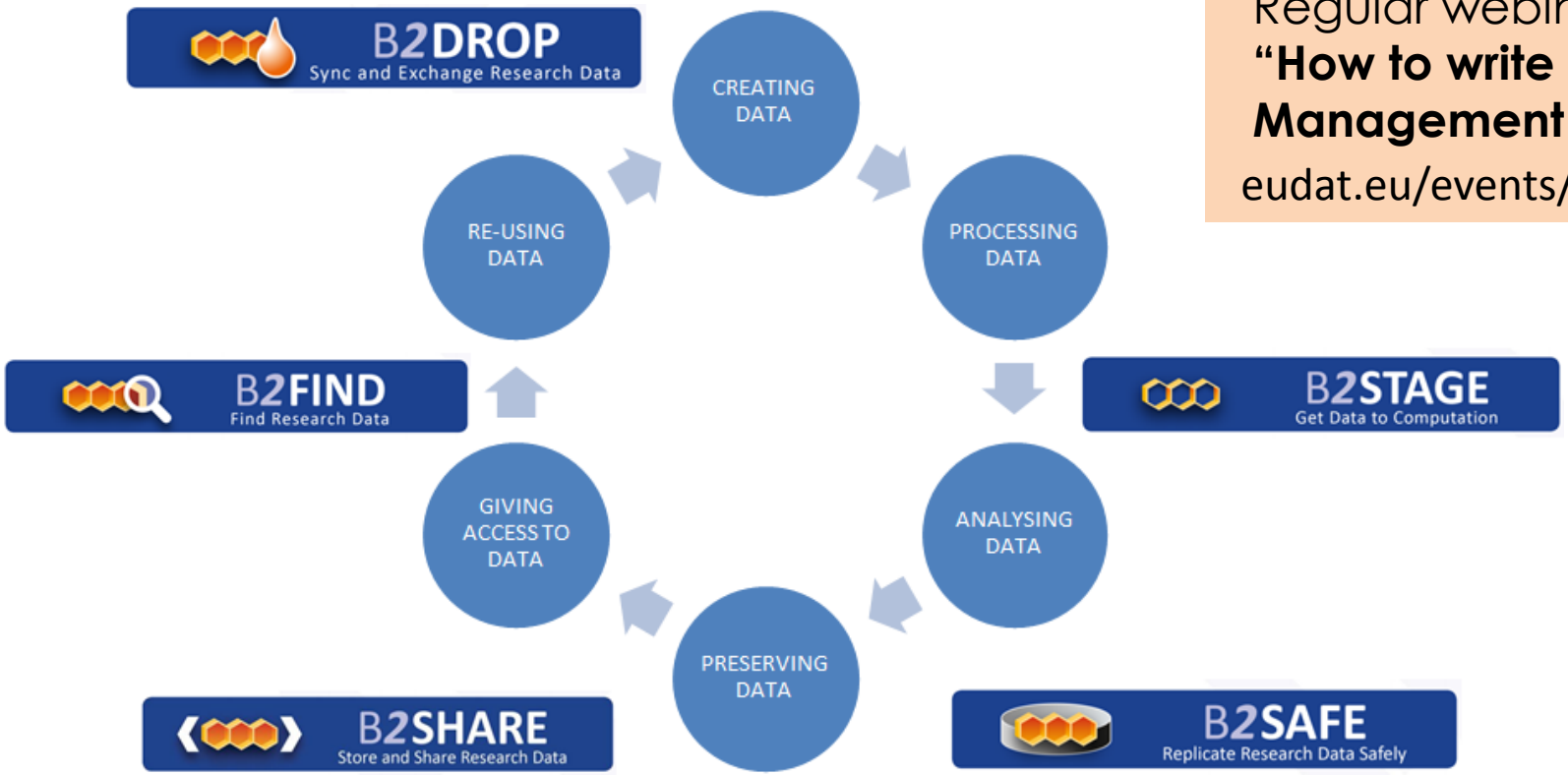
B2DROP Sync and Exchange Research Data

B2ACCESS Identity & Authorisation

B2NOTE

DATA TYPE REGISTRY

Addressing the full lifecycle of research data

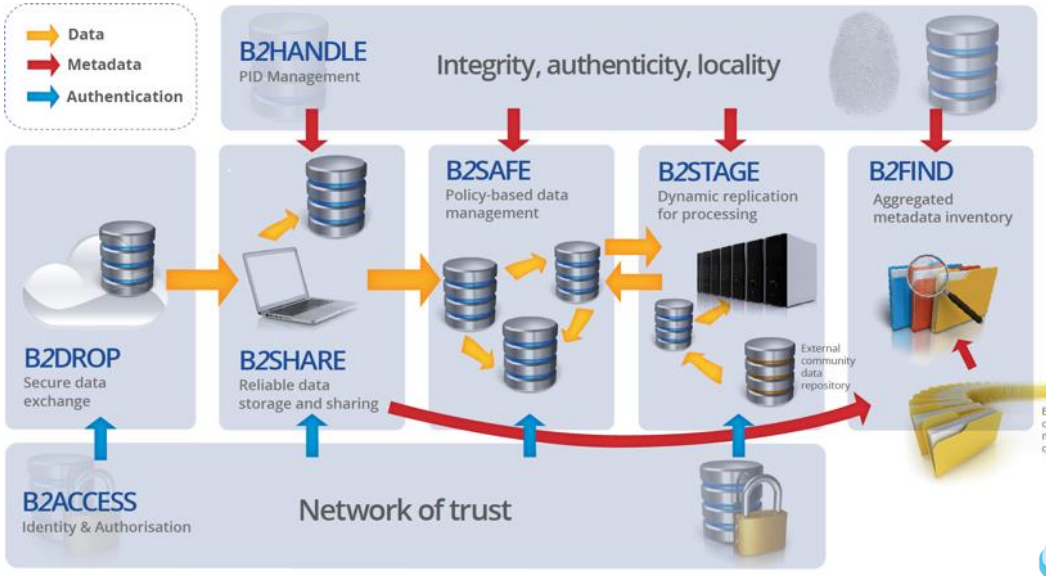


OpenAIRE & EUDAT

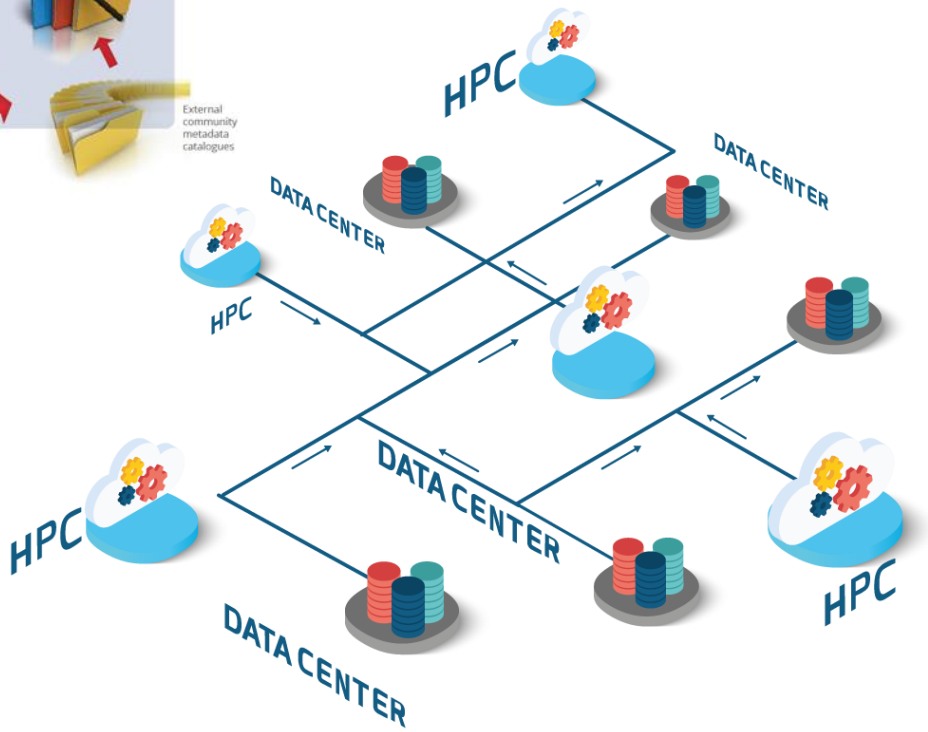
Regular webinar:
"How to write a Data Management Plan"
eudat.eu/events/webinar/

eudat.eu/training/research-data-management

The CDI – A Service Infrastructure



EUDAT Collaborative Data Infrastructure



Who are EUDAT Communities?

EUDAT ENV Core Communities

enes EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

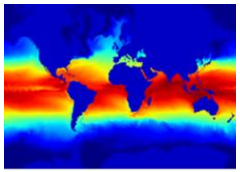
ICOS INTEGRATED CARBON OBSERVATION SYSTEM

EPOS EUROPEANPLATEOBSERVINGSYSTEM

LTER europe

SeaDataNet

EUDAT ENV Data Pilots



Support to scientific research on seasonal-to-decadal climate and air quality modelling

LIFE+RESPIRA

JADDS

LifeWatch

DATA SPHINX

Euro Argo

DRIHM DISTRIBUTED RESEARCH INFRASTRUCTURE FOR HYDRO-METEOROLOGY

Institutions and partners

INGV

GFZ Helmholtz Centre POTSDAM

LUND UNIVERSITY

ALTERNet

ENV europe

LTER europe

Gestión Ambiental de Navarra, S.A.

DRIHM DISTRIBUTED RESEARCH INFRASTRUCTURE FOR HYDRO-METEOROLOGY

Koninklijk Nederlands Meteorologisch Instituut
Ministerie van Infrastructuur en Milieu

Ifremer

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

ENVIRONMENT AGENCY AUSTRIA umweltbundesamt

Universidad de Navarra

IC3 INSTITUT CATALÀ DE CIÈNCIES DEL CLIMA

DKRZ

JADDS

Life

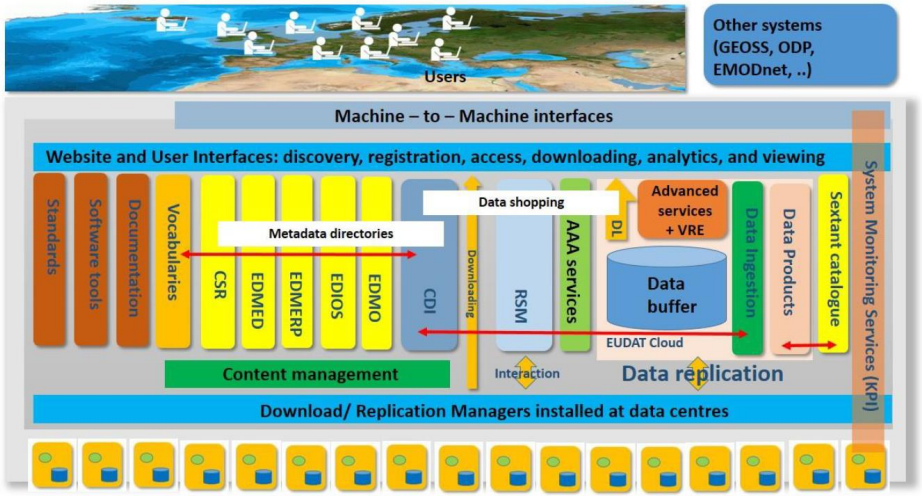
BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

DATA SPHINX

For more information visit - https://eudat.eu/use-cases

EUDAT & SeaDataCloud

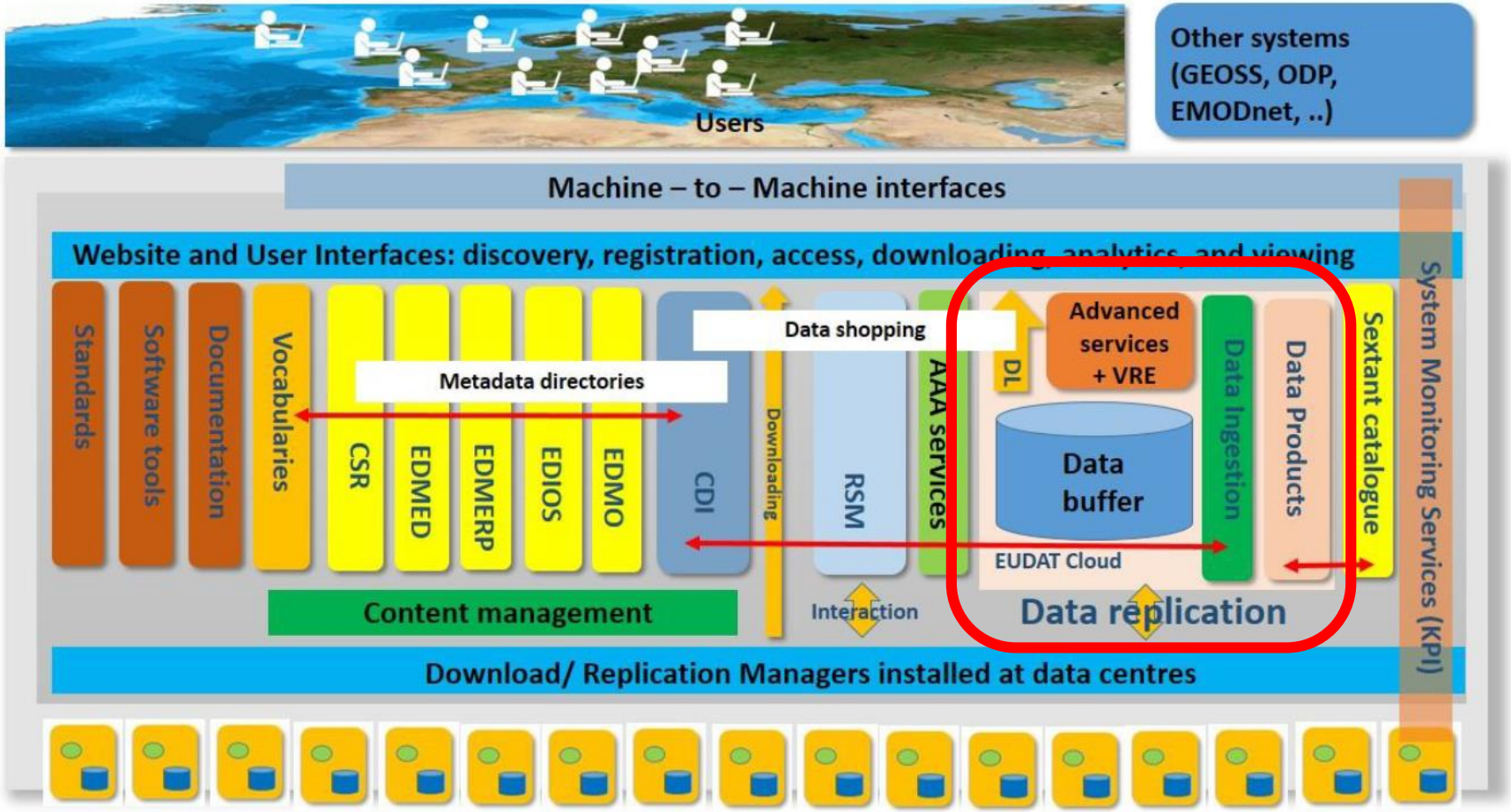
- Strategic cooperation
- Main goal:
  - Strengthening SeaDataNet underlying IT infrastructure
  - Improving access to and discovery of SeaDataNet data & products
- Output:
  - A cloud environment for storing datasets and providing cloud computing capabilities
  - A Virtual Research Environment (VRE) with various analysis and visualisation tools



Implementation phases

1) Replicate data from a subset of SeaDataNet centers to the EUDAT centers (technical cache)
2) Define relevant services for SeaDataNet (e.g. data subscription, etc.)
3) Develop a common cloud and computing environment integrated with the SeaDataNet portal and other supported services

EUDAT & SeaDataCloud



WP	Title	EUDAT Partners	PMs
WP2	**Project Network Coordination**	CSC, CINECA, DKRZ, STFC, GRNET	6
WP3	**Training of Data Providers**	CSC, CINECA	3
WP5	**Expansion and governance of metadata and data content** • WP5.4: Installing, configuring and taking into operation local components for the upgraded CDI data discovery and access service	CSC, CINECA, DKRZ, STFC, GRNET	15
WP6	**Core and Advanced Services**	CSC, CINECA, DKRZ, STFC, GRNET	35
WP7	**Tuning of requirements and overall integration**	CSC	2
WP8	**Governance of standards and development of common services** • --- • WP8.5: Upgrading the operational Monitoring system	CSC, CINECA, DKRZ, STFC, GRNET	45
WP9	**Developments of upstream services** • WP9.1: Upgrading the CDI service making use of the cloud • WP9.2 Developing and deploying additional services in the cloud for ensuring integrity and conformity of the CDI related cloud data resources • WP9.4: Developing integrated online services for ingesting autonomous observatory data • WP9.8: Development of a solution for coordinated distributed DataCite DOI minting service	CSC, CINECA, DKRZ, STFC, GRNET	76
WP10	**Developments of downstream services** • WP10.1: Development of the Virtual Research Environment (VRE) • WP10.2: Developing advanced services for search, processing, analytics, quality, visualisation • WP10.3: Developing MySeaDataCloud	CSC, CINECA, DKRZ, STFC, GRNET	75

Deliverables

Deliverable number	Title	Time	Lead
D9.1	Specification of the SeaDataNet European cloud environment	6	CSC
D9.2	Enhanced EUDAT B2SAFE replication software	12	CINECA
D9.3	Adapted EUDAT B2HOST service operational	12	DKRZ
D10.1	Specification of the Virtual Research Environment (VRE) in the cloud and development plan	12	DKRZ
D10.2	Virtual Research Environment (VRE) in the cloud operational	30	ALL
D.10.19	Specification of MySeaDataCloud and development plan	18	CSC/CINECA
D10.20	MySeaDataCloud operational	30	STFC (ALL)

EUDAT Partners in SDC

- **CSC – IT Center for Science** Ltd, established in 1971, is the Finnish national HPC centre providing high-performance computing, networking and data services to Finnish academia, research institutes, public sector and industry. CSC is a non- profit limited company whose shares are fully owned by Finnish state, and governed by the Finnish Ministry of Education and Culture.

- **CINECA,** established in 1969, is the Italian national supercomputing facility and one of the largest in Europe, providing high performance computing resources, data management and storage systems, tools, middleware services and expertise, aiming to develop and promote technical and scientific discoveries into the HPC field for the Italian and European research communities. CINECA is a non-profit consortium of 67 Italian Universities, the National Institute of Oceanography and Experimental Geophysics (OGS), the National Research Council (CNR), and the Ministry of Education, University and Research (MIUR).

- **DKRZ** is a national German facility, providing state-of-the-art super-computing data service and other associated services to the German and international scientific community to conduct top of the line Earth System and Climate Modeling. DKRZ is partner in ENES (European Network for Earth System Modelling) and is one the representatives of the Earth system research community in the EUDAT project. DKRZ is operating with its ICSU World Data Center Climate (WDCC21) a community specific long-term data archive. Connected with its WDCC

EUDAT Partners in SDC

- **STFC** is the UK public sector research organization providing access to large scale scientific facilities. The Scientific Computing Department (SCD) –- provides the STFC facilities with advanced IT infrastructure including massive data storage, high-end supercomputing, vast network bandwidth, and interoperability with other infrastructure in the UK and internationally. STFC is a major provider of resources (both compute and storage) for many large projects including WLCG, ENES (through CEDA where it also provides high performance computing and data archival services) and PanDATA providing multi-petabyte archival services for Diamond Light Source and ISIS.

- **GRNET** - the Greek Research and Technology Network - is a state–owned company, operating under the auspices of the Greek Ministry of Education. Its mission is to provide high–quality electronic Infrastructure services to the academic, research and educational community of Greece; and link these with global e-Infrastructures; GRNET is the main IT enabler for big science as well as long-tail science in Greece, and core player in international activities.

10 year initial commitment

A landmark agreement sustaining the pan European Collaborative Data Infrastructure for the next 10 years

3rd October 2016

16 major European research organisations, data and computing centres *signed an agreement to sustain the EUDAT – pan European collaborative data infrastructure for the next 10 years. The organisations stand together behind a long term sustainability plan and commit to develop, maintain and deploy pan-European research data services and to promote harmonization of research data management practices across centres.*

What does this mean?

Service providers, both generic and thematic, and research communities have joined forces as part of a common framework for developing an interoperable layer of common data services. Known as the EUDAT Collaborative Data Infrastructure (CDI), this is essentially **a European e-infrastructure of integrated data services and resources** to support research. This infrastructure and its 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 establishment of the EUDAT CDI is timely with the imminent realization of the European Open Science Cloud[1] which aims to offer open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines.

What benefits will this bring to researchers?

Researchers can rely on **innovative data services** to support their research collaboration and data management. Additionally they benefit from a common service management framework delivered by CDI service providers and the connection between sites. These 16 organizations[2] from 11 countries with a further 4[3] in the process of joining, are European leaders in the deployment of data management and data infrastructure services. By joining the EUDAT CDI they are able to better **serve their local customers across borders and support cross-national collaborations**.

CDI First Signatories

Thanks for your attention.

www.eudat.eu