



EUDAT

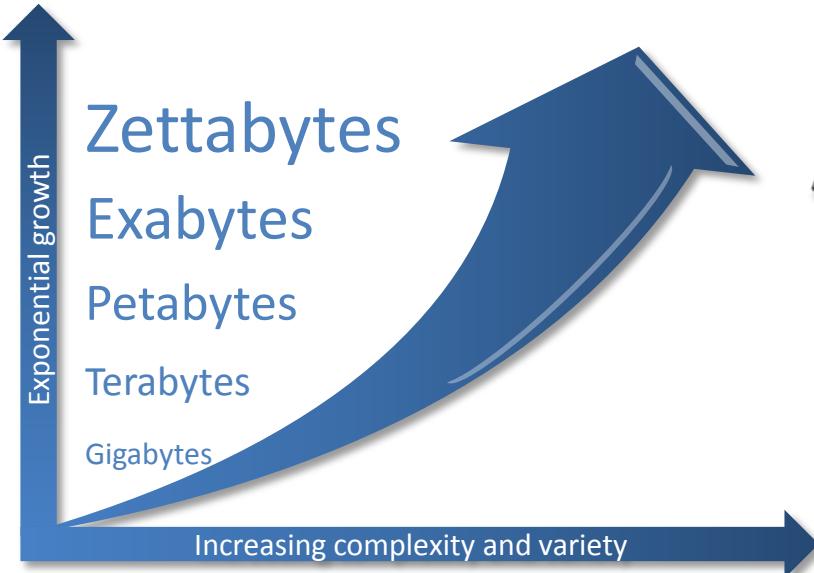
A cross-disciplinary data infrastructure in Horizon

2020

Damien Lecarpentier
EUDAT Project Manager
CSC – IT Center for Science Ltd



Data "Deluge"



- Where to store it?
- How to find it?
- How to make the most of it?



Synergies

If there are hundreds of Research Infrastructures, how many different data management systems can we sustain?

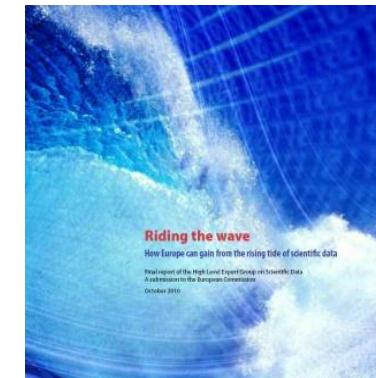
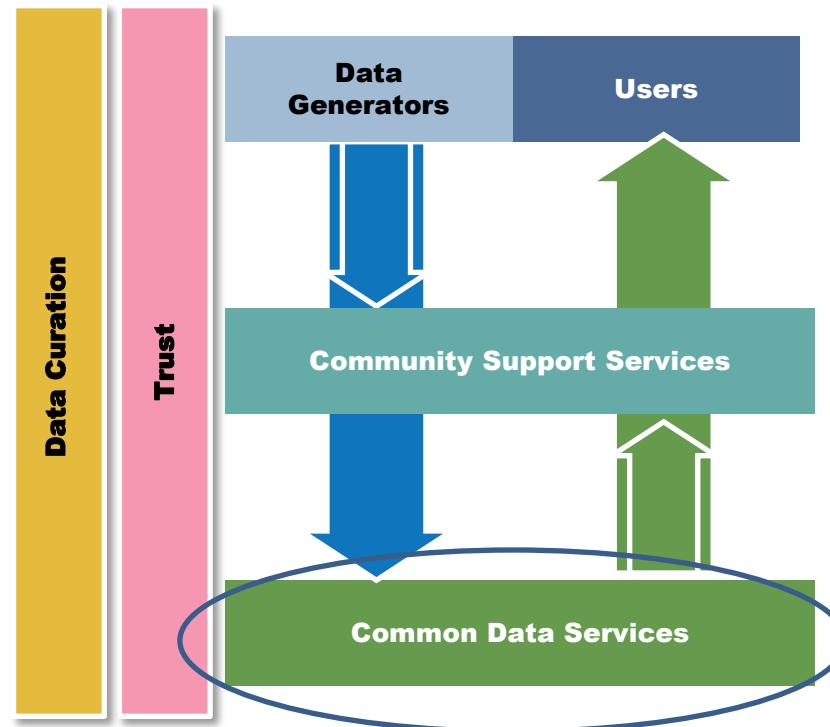




Riding the Wave

Collaborative Data Infrastructure

-A framework for the future? -



European Data



- Start date: 1st October 2011
- Duration: 36 Months
- Budget: 16.3 M€ (9.3M€ EC)
- EC Call: INFRA-2011-1.2.2
- Consortium: 25 partners from 13 countries
 - National data centers, technology providers, research
- Objectives:
 - Cost-efficient and high-quality CDI
 - Meets users' needs in flexible and sustainable way
 - Across geographical and disciplinary boundaries



<http://www.eudat.eu>

Consortium

25 European partners



EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



KIT
Karlsruhe Institute of Technology

INGV



BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

epcc

CINECA

INES
Centre Informatique National de l'Enseignement Supérieur

DKRZ
DEUTSCHES KLIMARECHENZENTRUM



CSC



RedIRIS

rzg
RECHEN-ZENTRUM GARCHING

CERFACS

Max-Planck-Institut für Meteorologie

JÜLICH
FORSCHUNGSZENTRUM

Science & Technology Facilities Council

maatG

SUPERCOMPUTING
POZNAN

SNIC

umweltbundesamt[®]
ENVIRONMENT AGENCY AUSTRIA

UNINETT sigma

SURF SARA

UCL

Trust-IT Services Ltd
Communicating ICT to markets

EUDAT

Seven Research Communities on Board

- **EPOS:** European Plate Observatory System
- **CLARIN:** Common Language Resources and Technology Infrastructure
- **ENES:** Service for Climate Modelling in Europe
- **LifeWatch:** Biodiversity Data and Observatories
- **VPH:** The Virtual Physiological Human
- **INCF:** International Neuroinformatics Coordinating Facility
- **DRIHM:** Distributed Research Infrastructure for Hydrometeorology



User Forums + 25 communities



1st User Forum
7-8 March 2012,
Barcelona



BBMRI
Biobanking and
Biomolecular
Resources Research
Infrastructure



Council of European Social Science Data Archives



EUROPEAN NETWORK
FOR EARTH SYSTEM MODELLING



DISTRIBUTED RESEARCH INFRASTRUCTURE
FOR HYDRO-METEOROLOGY



EUROPEAN PLATE OBSERVING SYSTEM



Virtual Physiological Human
network of excellence



International Neuroinformatics
Coordinating Facility



Worldwide LHC Computing Grid



Digital Research Infrastructure
for the Arts and Humanities



european
multidisciplinary
seafloor
observatory



integrated
carbon
observation
system



agInfra
A data infrastructure for agriculture



data infrastructure
for climate research
a project of the european union



EUROPEAN CLINICAL RESEARCH INFRASTRUCTURES NETWORK



BioVeL



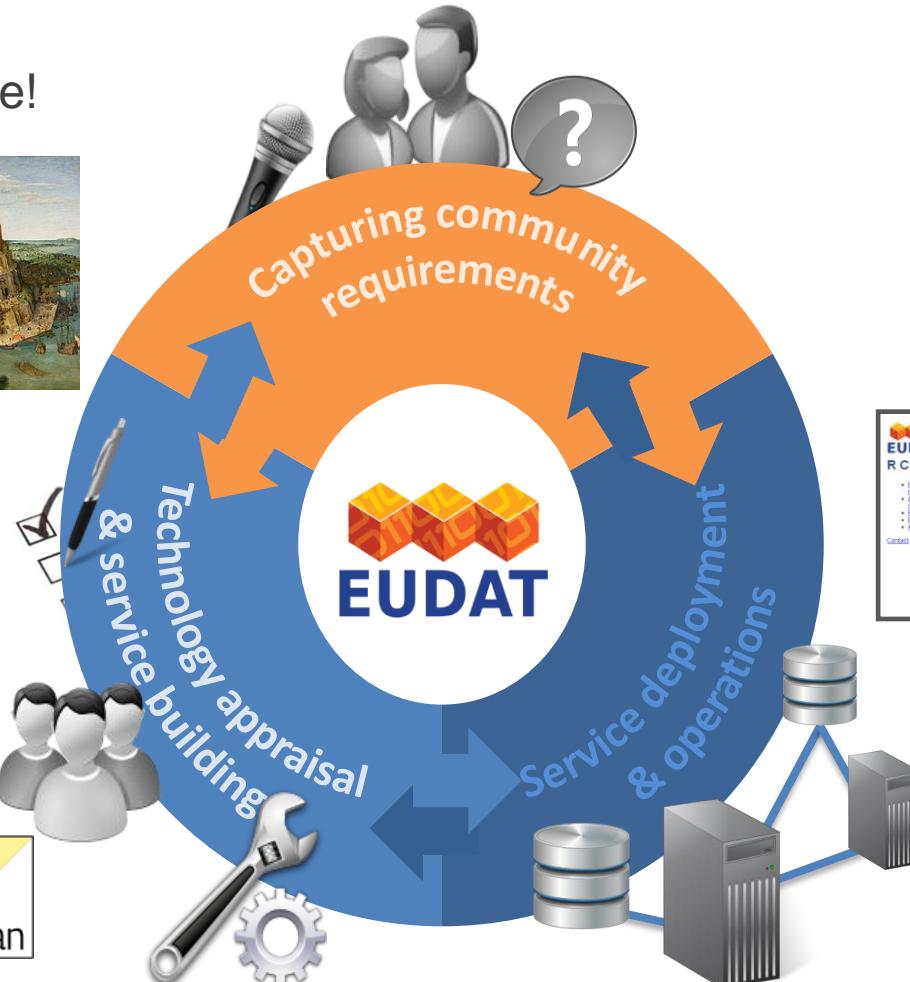
EMBRC
EUROPEAN
MARINE
BIOLOGICAL
RESOURCE
CENTRE

Service Building Process

Takes time!



Reusing existing technologies and expertise rather than reinventing everything!



Infrastructure coordination (resources, security, etc.)

A screenshot of the 'Welcome to the EUDAT Resource Coordination Tool' page. It shows a sidebar with 'EUDAT R C T' and a 'Sites' section. The main area displays a table with columns: Name, Last update date, and Certification status. The table lists numerous site entries, all marked as 'Production' and 'Certified'.

Name	Last update date	Certification status
PROD1	2018-01-01	Production
PROD2	2018-01-01	Production
PROD3	2018-01-01	Production
PROD4	2018-01-01	Production
PROD5	2018-01-01	Production
PROD6	2018-01-01	Production
PROD7	2018-01-01	Production
PROD8	2018-01-01	Production
PROD9	2018-01-01	Production
PROD10	2018-01-01	Production
PROD11	2018-01-01	Production
PROD12	2018-01-01	Production
PROD13	2018-01-01	Production
PROD14	2018-01-01	Production
PROD15	2018-01-01	Production
PROD16	2018-01-01	Production
PROD17	2018-01-01	Production
PROD18	2018-01-01	Production
PROD19	2018-01-01	Production
PROD20	2018-01-01	Production
PROD21	2018-01-01	Production
PROD22	2018-01-01	Production
PROD23	2018-01-01	Production
PROD24	2018-01-01	Production
PROD25	2018-01-01	Production
PROD26	2018-01-01	Production
PROD27	2018-01-01	Production
PROD28	2018-01-01	Production
PROD29	2018-01-01	Production
PROD30	2018-01-01	Production
PROD31	2018-01-01	Production
PROD32	2018-01-01	Production
PROD33	2018-01-01	Production
PROD34	2018-01-01	Production
PROD35	2018-01-01	Production
PROD36	2018-01-01	Production
PROD37	2018-01-01	Production
PROD38	2018-01-01	Production
PROD39	2018-01-01	Production
PROD40	2018-01-01	Production
PROD41	2018-01-01	Production
PROD42	2018-01-01	Production
PROD43	2018-01-01	Production
PROD44	2018-01-01	Production
PROD45	2018-01-01	Production
PROD46	2018-01-01	Production
PROD47	2018-01-01	Production
PROD48	2018-01-01	Production
PROD49	2018-01-01	Production
PROD50	2018-01-01	Production
PROD51	2018-01-01	Production
PROD52	2018-01-01	Production
PROD53	2018-01-01	Production
PROD54	2018-01-01	Production
PROD55	2018-01-01	Production
PROD56	2018-01-01	Production
PROD57	2018-01-01	Production
PROD58	2018-01-01	Production
PROD59	2018-01-01	Production
PROD60	2018-01-01	Production
PROD61	2018-01-01	Production
PROD62	2018-01-01	Production
PROD63	2018-01-01	Production
PROD64	2018-01-01	Production
PROD65	2018-01-01	Production
PROD66	2018-01-01	Production
PROD67	2018-01-01	Production
PROD68	2018-01-01	Production
PROD69	2018-01-01	Production
PROD70	2018-01-01	Production
PROD71	2018-01-01	Production
PROD72	2018-01-01	Production
PROD73	2018-01-01	Production
PROD74	2018-01-01	Production
PROD75	2018-01-01	Production
PROD76	2018-01-01	Production
PROD77	2018-01-01	Production
PROD78	2018-01-01	Production
PROD79	2018-01-01	Production
PROD80	2018-01-01	Production
PROD81	2018-01-01	Production
PROD82	2018-01-01	Production
PROD83	2018-01-01	Production
PROD84	2018-01-01	Production
PROD85	2018-01-01	Production
PROD86	2018-01-01	Production
PROD87	2018-01-01	Production
PROD88	2018-01-01	Production
PROD89	2018-01-01	Production
PROD90	2018-01-01	Production
PROD91	2018-01-01	Production
PROD92	2018-01-01	Production
PROD93	2018-01-01	Production
PROD94	2018-01-01	Production
PROD95	2018-01-01	Production
PROD96	2018-01-01	Production
PROD97	2018-01-01	Production
PROD98	2018-01-01	Production
PROD99	2018-01-01	Production
PROD100	2018-01-01	Production
PROD101	2018-01-01	Production
PROD102	2018-01-01	Production
PROD103	2018-01-01	Production
PROD104	2018-01-01	Production
PROD105	2018-01-01	Production
PROD106	2018-01-01	Production
PROD107	2018-01-01	Production
PROD108	2018-01-01	Production
PROD109	2018-01-01	Production
PROD110	2018-01-01	Production
PROD111	2018-01-01	Production
PROD112	2018-01-01	Production
PROD113	2018-01-01	Production
PROD114	2018-01-01	Production
PROD115	2018-01-01	Production
PROD116	2018-01-01	Production
PROD117	2018-01-01	Production
PROD118	2018-01-01	Production
PROD119	2018-01-01	Production
PROD120	2018-01-01	Production
PROD121	2018-01-01	Production
PROD122	2018-01-01	Production
PROD123	2018-01-01	Production
PROD124	2018-01-01	Production
PROD125	2018-01-01	Production
PROD126	2018-01-01	Production
PROD127	2018-01-01	Production
PROD128	2018-01-01	Production
PROD129	2018-01-01	Production
PROD130	2018-01-01	Production
PROD131	2018-01-01	Production
PROD132	2018-01-01	Production
PROD133	2018-01-01	Production
PROD134	2018-01-01	Production
PROD135	2018-01-01	Production
PROD136	2018-01-01	Production
PROD137	2018-01-01	Production
PROD138	2018-01-01	Production
PROD139	2018-01-01	Production
PROD140	2018-01-01	Production
PROD141	2018-01-01	Production
PROD142	2018-01-01	Production
PROD143	2018-01-01	Production
PROD144	2018-01-01	Production
PROD145	2018-01-01	Production
PROD146	2018-01-01	Production
PROD147	2018-01-01	Production
PROD148	2018-01-01	Production
PROD149	2018-01-01	Production
PROD150	2018-01-01	Production
PROD151	2018-01-01	Production
PROD152	2018-01-01	Production
PROD153	2018-01-01	Production
PROD154	2018-01-01	Production
PROD155	2018-01-01	Production
PROD156	2018-01-01	Production
PROD157	2018-01-01	Production
PROD158	2018-01-01	Production
PROD159	2018-01-01	Production
PROD160	2018-01-01	Production
PROD161	2018-01-01	Production
PROD162	2018-01-01	Production
PROD163	2018-01-01	Production
PROD164	2018-01-01	Production
PROD165	2018-01-01	Production
PROD166	2018-01-01	Production
PROD167	2018-01-01	Production
PROD168	2018-01-01	Production
PROD169	2018-01-01	Production
PROD170	2018-01-01	Production
PROD171	2018-01-01	Production
PROD172	2018-01-01	Production
PROD173	2018-01-01	Production
PROD174	2018-01-01	Production
PROD175	2018-01-01	Production
PROD176	2018-01-01	Production
PROD177	2018-01-01	Production
PROD178	2018-01-01	Production
PROD179	2018-01-01	Production
PROD180	2018-01-01	Production
PROD181	2018-01-01	Production
PROD182	2018-01-01	Production
PROD183	2018-01-01	Production
PROD184	2018-01-01	Production
PROD185	2018-01-01	Production
PROD186	2018-01-01	Production
PROD187	2018-01-01	Production
PROD188	2018-01-01	Production
PROD189	2018-01-01	Production
PROD190	2018-01-01	Production
PROD191	2018-01-01	Production
PROD192	2018-01-01	Production
PROD193	2018-01-01	Production
PROD194	2018-01-01	Production
PROD195	2018-01-01	Production
PROD196	2018-01-01	Production
PROD197	2018-01-01	Production
PROD198	2018-01-01	Production
PROD199	2018-01-01	Production
PROD200	2018-01-01	Production
PROD201	2018-01-01	Production
PROD202	2018-01-01	Production
PROD203	2018-01-01	Production
PROD204	2018-01-01	Production
PROD205	2018-01-01	Production
PROD206	2018-01-01	Production
PROD207	2018-01-01	Production
PROD208	2018-01-01	Production
PROD209	2018-01-01	Production
PROD210	2018-01-01	Production
PROD211	2018-01-01	Production
PROD212	2018-01-01	Production
PROD213	2018-01-01	Production
PROD214	2018-01-01	Production
PROD215	2018-01-01	Production
PROD216	2018-01-01	Production
PROD217	2018-01-01	Production
PROD218	2018-01-01	Production
PROD219	2018-01-01	Production
PROD220	2018-01-01	Production
PROD221	2018-01-01	Production
PROD222	2018-01-01	Production
PROD223	2018-01-01	Production
PROD224	2018-01-01	Production
PROD225	2018-01-01	Production
PROD226	2018-01-01	Production
PROD227	2018-01-01	Production
PROD228	2018-01-01	Production
PROD229	2018-01-01	Production
PROD230	2018-01-01	Production
PROD231	2018-01-01	Production
PROD232	2018-01-01	Production
PROD233	2018-01-01	Production
PROD234	2018-01-01	Production
PROD235	2018-01-01	Production
PROD236	2018-01-01	Production
PROD237	2018-01-01	Production
PROD238	2018-01-01	Production
PROD239	2018-01-01	Production
PROD240	2018-01-01	Production
PROD241	2018-01-01	Production
PROD242	2018-01-01	Production
PROD243	2018-01-01	Production
PROD244	2018-01-01	Production
PROD245	2018-01-01	Production
PROD246	2018-01-01	Production
PROD247	2018-01-01	Production
PROD248	2018-01-01	Production
PROD249	2018-01-01	Production
PROD250	2018-01-01	Production
PROD251	2018-01-01	Production
PROD252	2018-01-01	Production
PROD253	2018-01-01	Production
PROD254	2018-01-01	Production
PROD255	2018-01-01	Production
PROD256	2018-01-01	Production
PROD257	2018-01-01	Production
PROD258	2018-01-01	Production
PROD259	2018-01-01	Production
PROD260	2018-01-01	Production
PROD261	2018-01-01	Production
PROD262	2018-01-01	Production
PROD263	2018-01-01	Production
PROD264	2018-01-01	Production
PROD265	2018-01-01	Production
PROD266	2018-01-01	Production
PROD267	2018-01-01	Production
PROD268	2018-01-01	Production
PROD269	2018-01-01	Production
PROD270	2018-01-01	Production
PROD271	2018-01-01	Production
PROD272	2018-01-01	Production
PROD273	2018-01-01	Production
PROD274	2018-01-01	Production
PROD275	2018-01-01	Production
PROD276	2018-01-01	Production
PROD277	2018-01-01	Production
PROD278	2018-01-01	Production
PROD279	2018-01-01	Production
PROD280	2018-01-01	Production
PROD281	2018-01-01	Production
PROD282	2018-01-01	Production
PROD283	2018-01-01	Production
PROD284	2018-01-01	Production
PROD285	2018-01-01	Production
PROD286	2018-01-01	Production
PROD287	2018-01-01	Production
PROD288	2018-01-01	Production
PROD289	2018-01-01	Production
PROD290	2018-01-01	Production
PROD291	2018-01-01	Production
PROD292	2018-01-01	Production
PROD293	2018-01-01	Production
PROD294	2018-01-01	Production
PROD295	2018-01-01	Production
PROD296	2018-01-01	Production
PROD297	2018-01-01	Production
PROD298	2018-01-01	Production
PROD299	2018-01-01	Production
PROD300	2018-01-01	Production

Selected Services

Metadata Catalogue

Aggregated EUDAT metadata domain.
Data inventory



PID

Identity
Integrity
Authenticity
Locations

AAI

Network of trust
among
authentication
and
authorization
actors

Data Staging

Dynamic replication
to HPC workspace
for processing



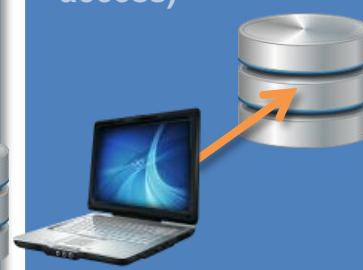
Safe Replication

Data curation and
access optimization

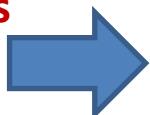


Simple Store

Researcher data
store (simple
upload, share and
access)



New services
to come



EUDAT Box
dropbox-like service
easy sharing
local synching



Semantic Anno
checking & referencing

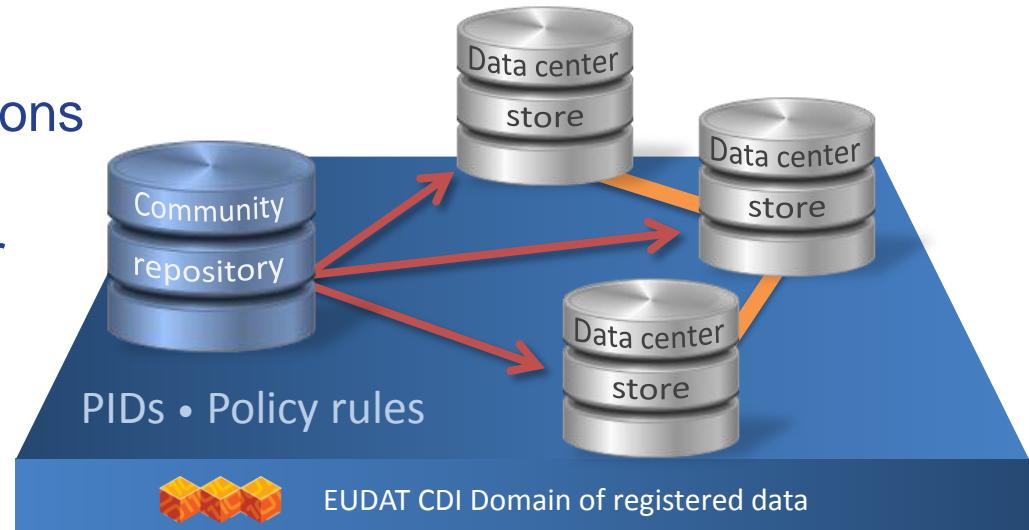


Dynamic Data
immediate handling



Safe Replication Service

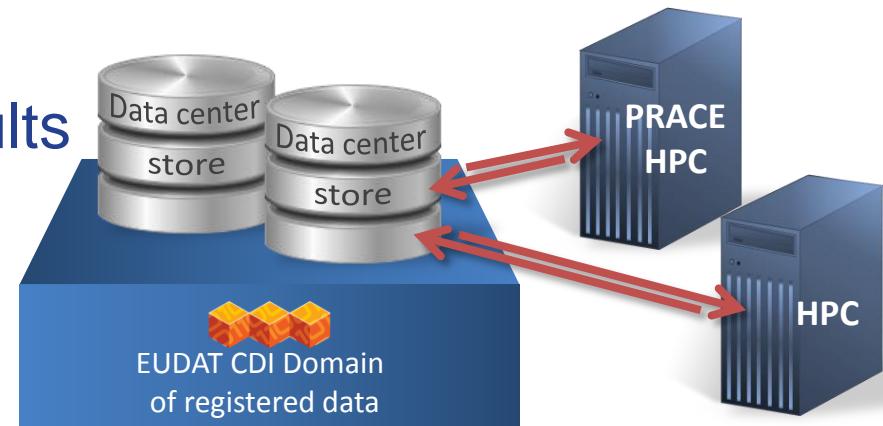
- Robust, safe and highly available data replication service for small- and medium- sized repositories
 - To guard against data loss in long-term archiving and preservation
 - To optimize access for user from different regions
 - To bring data closer to powerful computers for compute-intensive analysis



<http://eudat.eu/safe-replication> | eudat-safereplication@postit.csc.fi

Data Staging Service

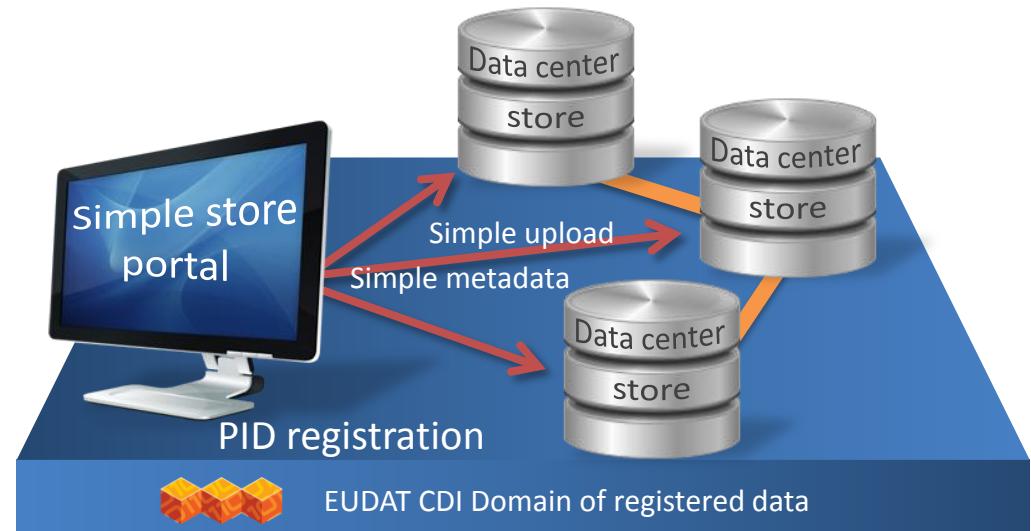
- Support researchers in transferring large data collections from EUDAT storage to HPC facilities
- Reliable, efficient, and easy-to-use tools to manage data transfers
- Provide the means to re-ingest computational results back into the EUDAT infrastructure



<http://eudat.eu/datastaging> | eudat-datastaging@postit.csc.fi

Simple Store Service

- Allow registered users to upload "long tail" data into the EUDAT store
- Enable sharing objects and collections with other researchers
- Utilise other EUDAT services to provide reliability and data retention



<http://eudat.eu/simplestore> | eudat-simplestore@postit.csc.fi

1. Upload files (drag and drop or select files)

Deposits are done in 3 steps:

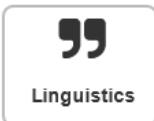
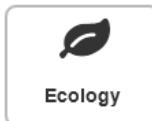
1. you select and upload one or several data resources
2. you then select the metadata set you want to use and describe the resource by metadata
3. you deposit the selected resource including its metadata elements. The deposit takes a moment to be processed. You will get a reference URL at once.

[Select files](#) [Start upload](#)

Filename	Size	Status	
EUDAT-poster03.jpg	324 KB	<div style="width: 100%; background-color: #0072BD; height: 10px;"></div>	

2. Select a domain or project

Please select a metadata description set which best fits to your data. The EUDAT Generic set includes a common set of elements which are inspired by Dublin Core, MARC and DataCite and this set is included in all sets as basis. So if you just want the common set select Generic, if you want in addition a specific community set select that community. Soon there will be more communities. You can always revise your decisions and select another set, however you may lose what you have already entered.

**Generic****Linguistics****Ecology****DRIHM**

3. Add some basic details

Generic

Add more details?

drihm

Title	<input type="text" value="DRIHM Images"/>
Description	<input type="text" value="Images uploaded as part of test work on the DRIHM project."/> ...
Creator	<input type="text" value="Quillon Harpham"/>
Open Access	<input checked="" type="checkbox"/> ON Open Access items may be downloaded by anyone
Licence	<input type="text" value="Open"/>
Publisher	<input type="text" value="HR Wallingford"/>
Publication Date	<input type="text" value="30th August 2013"/>
Tags	<input type="text" value="Hydrology, Meteorology, Hydraulics, Impact"/> Comma separated list of keywords associated with item

Reference date	<input type="text"/>
Reference System	<input type="text"/>
Topic Category	<input type="text"/>
Responsible Party	<input type="text"/>
Geographic Location	<input type="text"/>
Spatial Resolution	<input type="text"/>
Vertical Extent	<input type="text"/>
Lineage	<input type="text"/>

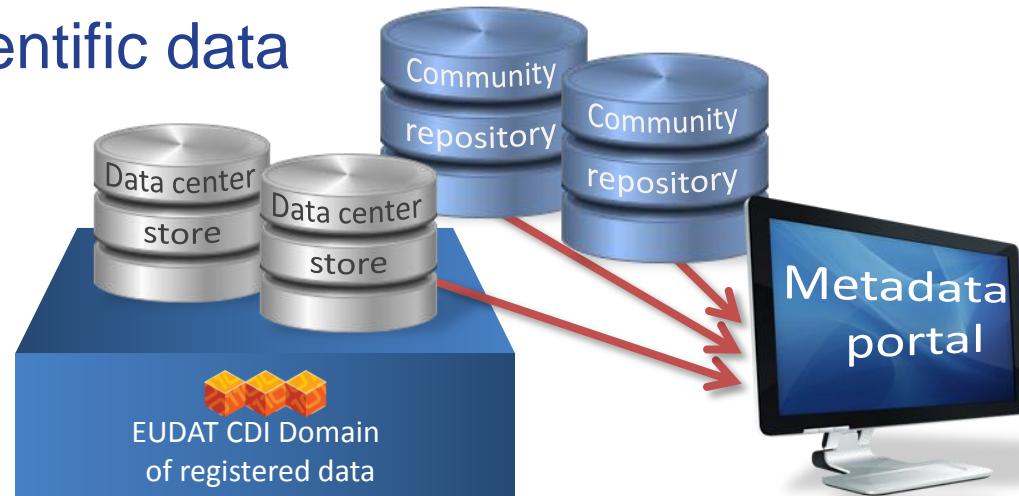
4. Deposit

Simple Store Basic/Premium

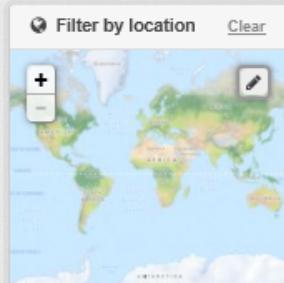
Properties/functions	Basic	Premium
Upload Capacity	< 2GB per file/deposit	On-demand
Storage Capacity	Faire share	Unlimited
Center Selection	No	Yes
Replication	No	Yes
Customized interfaces (MD fields, logo, etc.)	Yes	Yes
Access management	Standard (open/not open)	Extended (restricted access to groups, etc.)
Duration	TBC	Based on SLAs

Metadata Service

- Easily find collections of scientific data – generated either by various communities or via EUDAT services
- Access those data collections through the given references in the metadata to the relevant data stores
- Europeana of scientific data



<http://eudat.eu/metadata> | eudat-metadata@postit.csc.fi

[Home](#) / Datasets

Map data © OpenStreetMap contributors
Tiles by MapQuest

Communities

[Clear All](#)[DataCite \(45789\)](#)[CLARIN \(41759\)](#)[GBIF \(11090\)](#)[ENES \(265\)](#)[Show More Communities](#)

Tags

[Clear All](#)[Meertens collection... \(29257\)](#)[Dataset \(22392\)](#)[Text \(21247\)](#)**98,903 datasets found**

Order by: Relevance

**B2FIND**[Datasets](#)[Communities](#)[About](#) 

Hunting

This dataset has no description

Toy Game 1

This dataset has no description

radio: (national) radio news

Responsibility.Type = SAMPLING

TA_10F_NR_CTMD

This dataset has no description

FasciasStoryLA67

This dataset has no description

Communities

CLARIN



CLARIN

The Common Language Resources and Technology Infrastructure (CLARIN) project...

41759 Datasets



DataCite

DataCite is a not-for-profit organisation formed in London on 1 December...

45789 Datasets



ENES

The European Network for Earth System modelling (ENES) provides information...

265 Datasets



GBIF

The Global Biodiversity Information Facility (GBIF) was established by...

11090 Datasets



EUDAT Simple Store

The European Data Infrastructure (EUDAT) will offer a Simple Store service...

0 Datasets



Towards Horizon 2020

User driven services

Sustainability

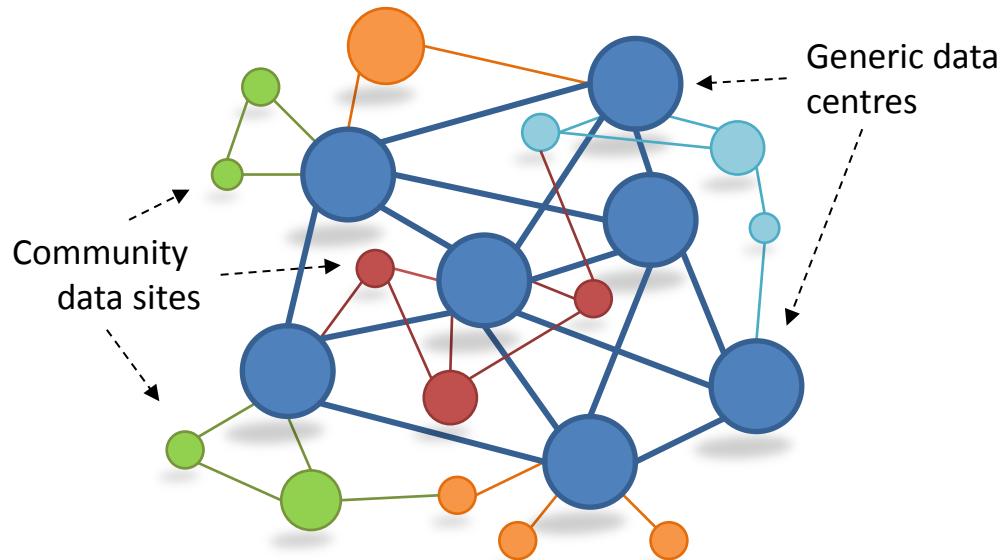
Trust

Synergy

Joint e-infrastructure roadmaps

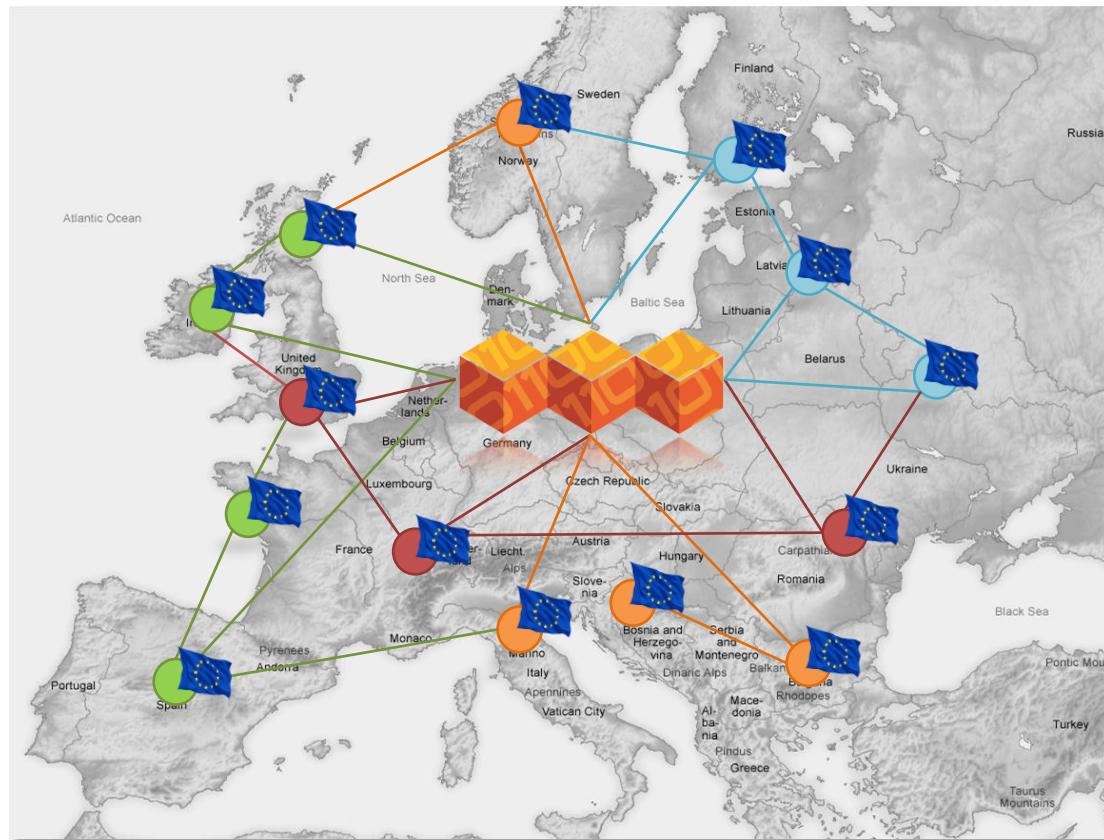
Global collaboration

A Network of Trusted Centers



- Strong and sustainable generic data centers with existing trusted relationships
- Each having specific relationship with research communities
- EUDAT is about providing solutions in a federated environment

Bridging National and European solutions



- Strong requirement from researchers and funders
- Path to Sustainability



HORIZON 2020



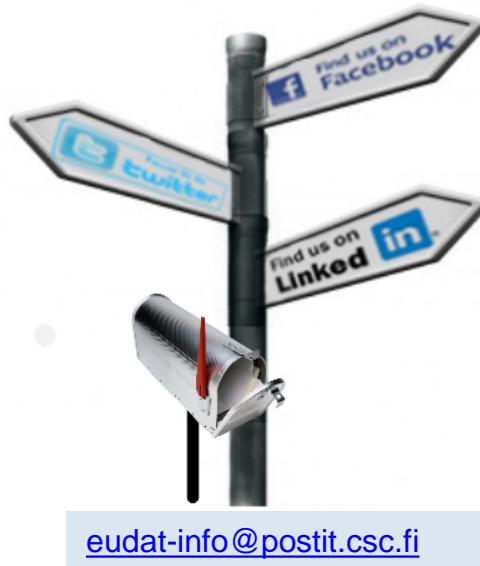
EUDAT Priorities in H2020

- Consolidation of Core Services
 - Increased performance, new functionalities, AAI, etc.
 - Develop tools and policies to facilitate usage: data management plans, licensing, training, etc.
 - Development of new services
- Financial Sustainability
 - Cost and funding models
 - Framework and mechanisms for sharing resources across sites and across communities (juste retour, etc.)
- Interoperability
 - E-Infrastructures → a joint roadmap?
 - National initiatives → service portfolios
 - RDA → EUDAT as a driver and implementer

Contact Us



<http://www.eudat.eu>



Project Coordinator: Kimmo Koski
kimmo.koski@csc.fi

Scientific Coordinator: Peter Wittenburg
peter.wittenburg@mpi.nl

Project Manager: Damien Lecarpentier
damien.lecarpentier@csc.fi

Dissemination Manager: Nagham Salman
naghamsalman@bsc.es

Industry Task Force: David Manset
dmanset@maatg.fr



EUDAT