



OUTCOME OF THE PRACE-EUDAT SEMINAR

*Giuseppe Fiameni | CINECA@EUDAT.CDI
PRACE Days 2018
Ljubljana, May 31st*



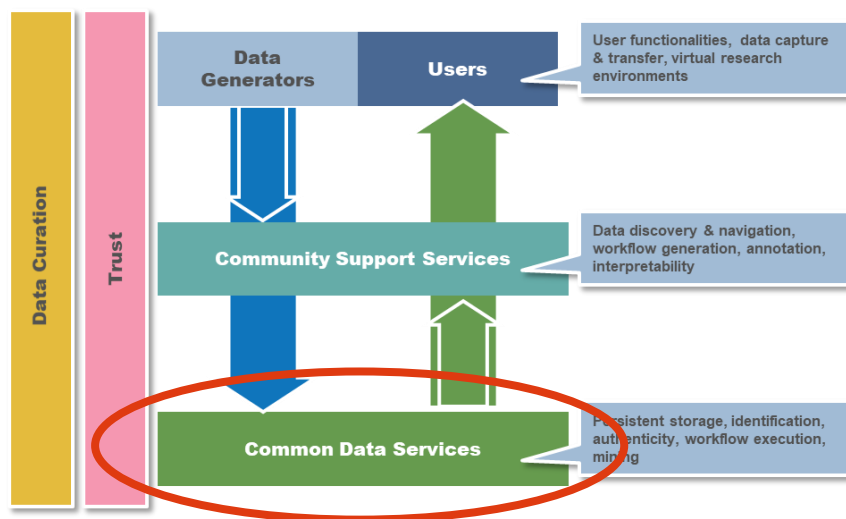
eudat.eu



@eudat_eu

The Collaborative Data Infrastructure

EUDAT is the European e-infrastructure of providing data services and resources to support scientific research.



The CDI members (January 2018)

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 (JULICH) - DE
- 11 Institute of Bioorganic Chemistry Polish Academy of Sciences (IBCH PAS) - PL

Generic, Interoperable Service Providers

- 1 Uninet Sigma 2 As (SIGMA) - NO
- 2 Data Archiving and Networked Services (DANS) - NL
- 3 JISC LBG (JISC) - UK
- 4 Swedish National Infrastructure for Computing (SNIC) - SE
- 5 University of Edinburgh - UK
- 6 Danish eInfrastructure Cooperation (DeiC) - DK

Thematic, Integrated Service Providers

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

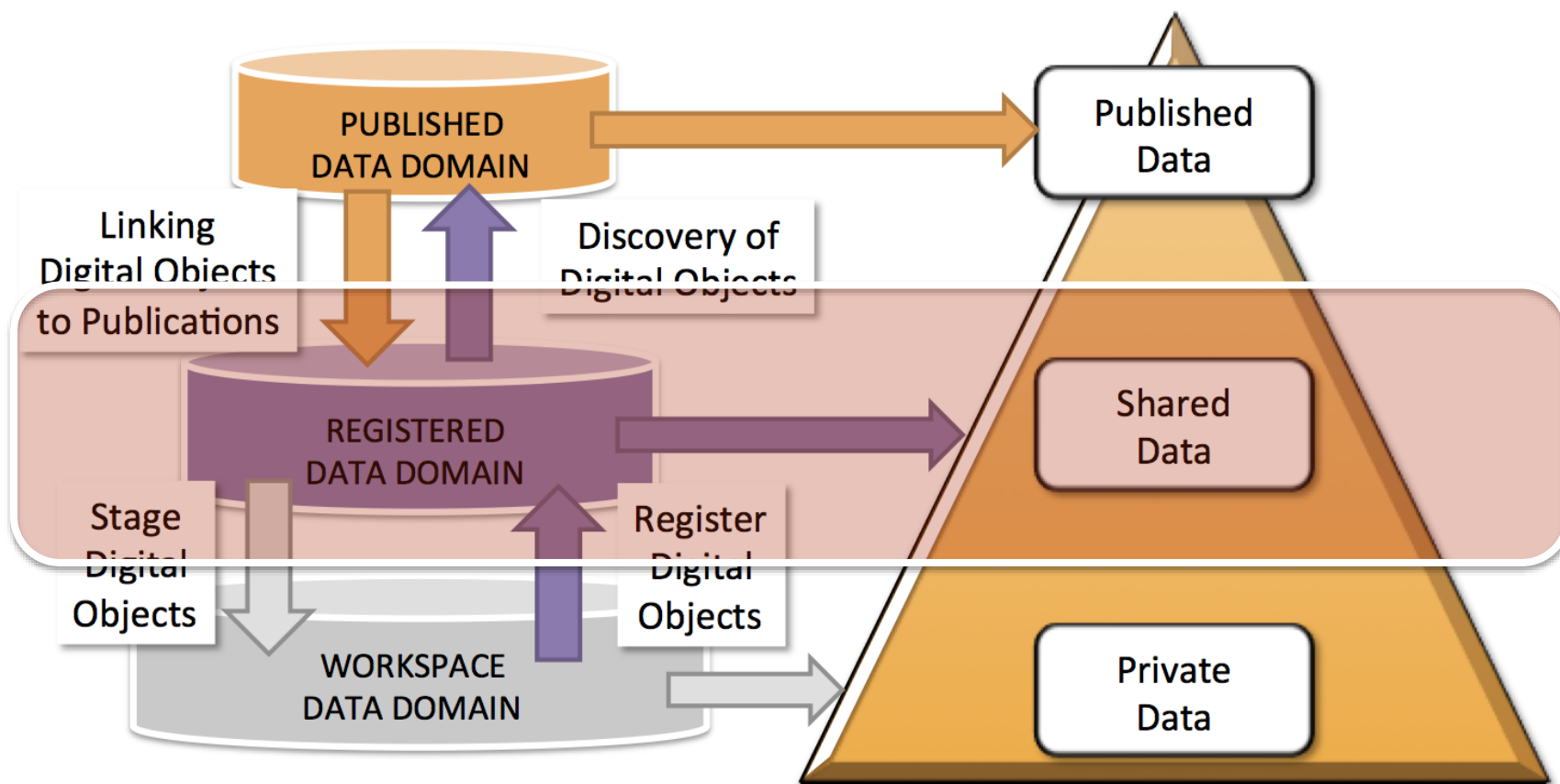
Thematic, Interoperable Service Providers

- 1 European Organization for Nuclear Research (CERN) - INT
- 2 Culham Centre for Fusion Energy (CCFE) - UK
- 3 Meertens Institute (MEERTENS) - NL
- 4 CLARIN-ERIC - NL



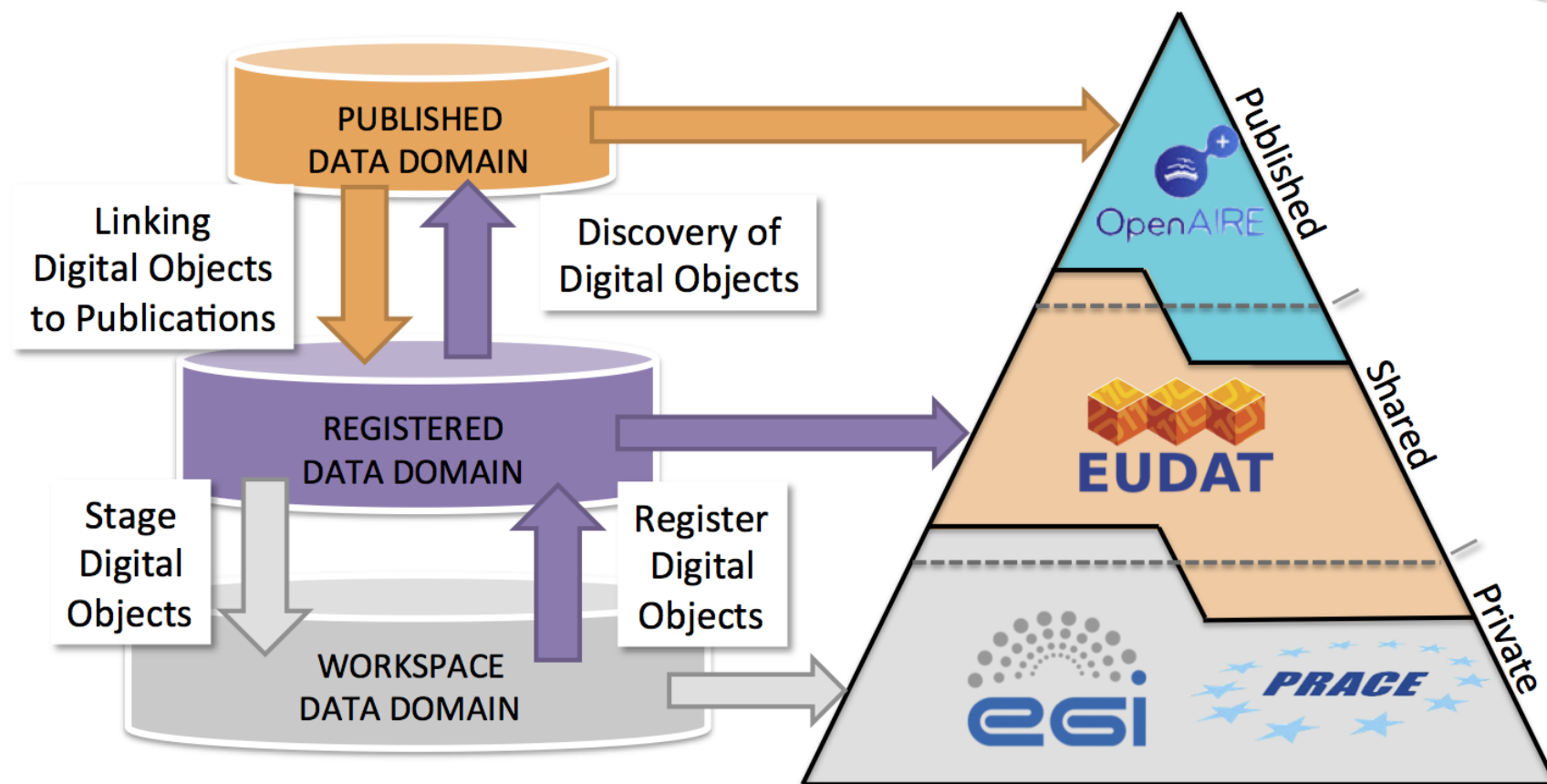
*more than 20 European
research organisations,
data and computing
centres in 14 countries*

CDI Data Domain







































EUDAT Data Domain modeled on the ANDS¹ Data Curation Continuum

CDI Data Domain

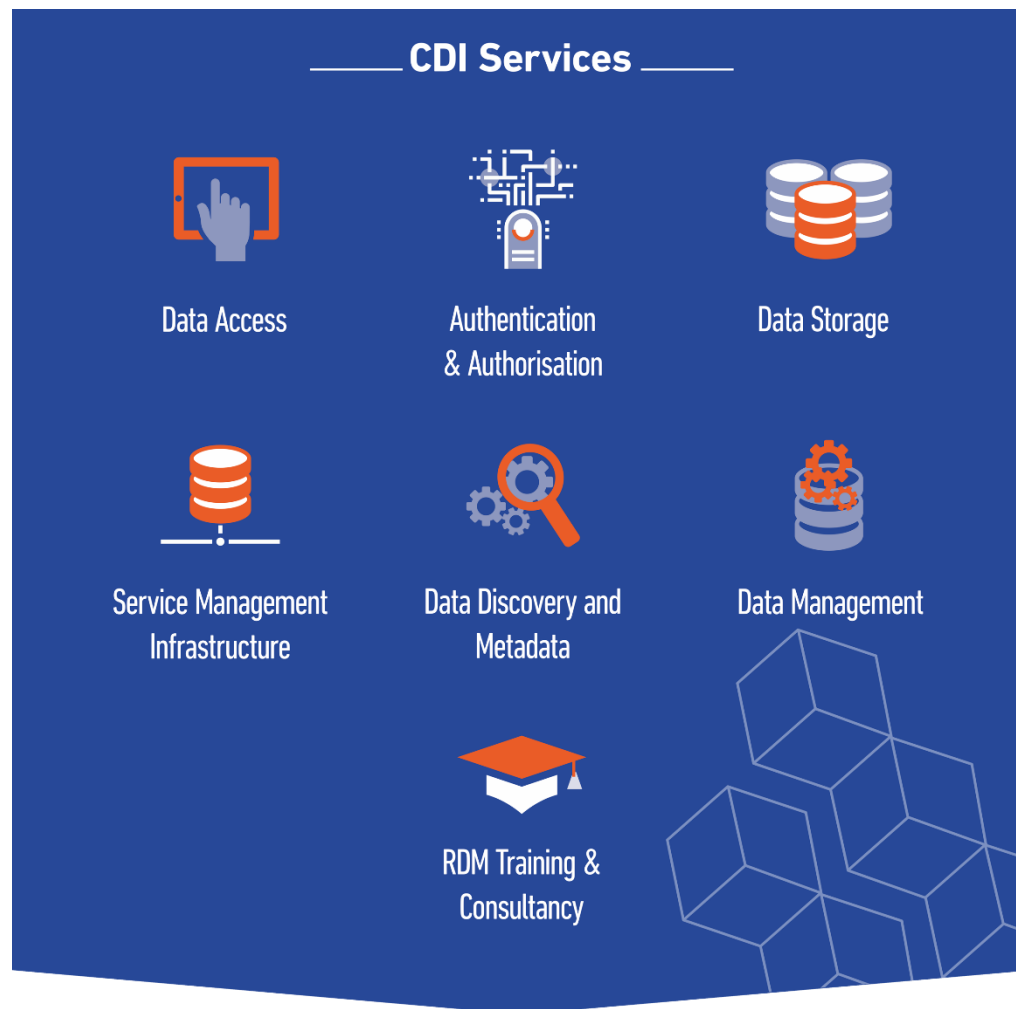


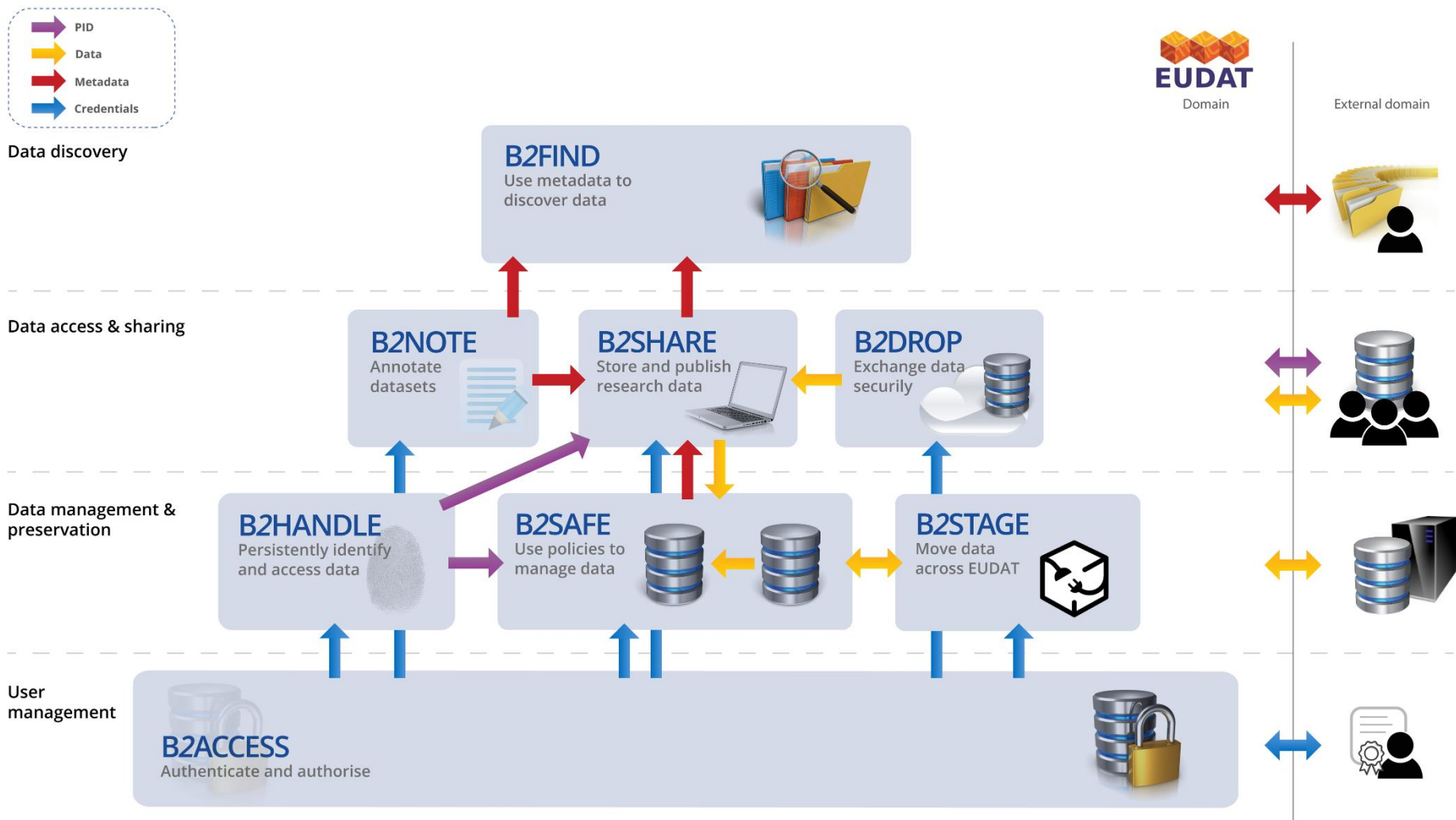
EUDAT Data Domain modeled on the ANDS¹ Data Curation Continuum

 Direct simulation data of turbulent flows	 High resolution simulations for Precision Cosmology	 Tokamak data mirror for JET and MAST data	 Unified Access to EISCAT radar data	 Research data repository for students' own results	 Data preservation and standardization in computational fluid dynamics
 European Network for Earth System Modelling	 European Long Term Ecological Research Network	 Data storage and preservation of high resolution climate experiments	 European Plate Observing System	 Distributed Research Infrastructure for Hydro-Meteorology	 Support to scientific research on seasonal-to-decadal climate and air quality modelling
 Jülich Atmospheric Data Distribution Service	 Integrated Carbon Observation System	 Public access to fine-grained city air quality data from roving sensors	 Infrastructure for accessing marine data	 Virtual Humans	 Data Infrastructure for Chemical Safety
 EUDAT as a long-term repository for the University of Porto	 Central online location for data sharing for all Aalto University researchers	 Enriching European Newspapers	 Cloud-like services to improve the preservation of digital cultural heritage	 Storing, Cataloguing, Relating, and Exposing OCR Objects from the Open Philology Project	 Common Language Resources and Technology Infrastructure
 Data infrastructure for biodiversity and ecosystem community	 Supporting data management for structural biologists	 An EUDAT-based FAIR Data Approach for Data Interoperability	 European Long Term Ecological Research Network	 Linked Data service pilot	 Long-term preservation of herbarium specimen images
 A distributed infrastructure for life-science information	 EUDAT services to guarantee long time archiving and visibility to the repository of IST Austria	 The use of the EUDAT repository to store clinical trials in a secure and compliant way	 International Neuroinformatics Coordinating Facility	 Information and Data Management Repository Platform for nanoscience in Europe	

	Physical Sciences
	Earth and related environmental sciences
	Medical biotechnology
	Multi-disciplinary
	Media and communications
	Languages & Literature
	Biological Sciences
	Mathematics and computer sciences
	Clinical Medicine
	Basic Medicine
	Nano-technology

EUDAT services are designed, built and implemented based on user community requirements.





The collaboration with PRACE

- ✦ Join forces to ensure the long-term interoperability between the two infrastructures to enable cross-infrastructure projects
- ✦ Enable EUDAT users fast access to computing resources for data intensive processing
- ✦ Enable PRACE users to manage results of their computations on the medium long period





- **Joint Open Calls for proposals**


- EUDAT offering data services and resources through regular PRACE calls
- Review process is transparent to users

- **Joint training activities**

- PRACE project investigators involved into EUDAT Data Management webinars and courses

- **Continuous technical discussion and developments of new components**

- Definition of the EUDAT Workspace area
- Synchronization of authentication credentials for single sign-on
- EUDAT clients as part of the PRACE Common Production Environment



Acronym	Title	Field	Country	Data requirements in TB during the PRACE project	Data requirements in TB after the PRACE project	Duration of the access to this data service after the PRACE allocation	PRACE Site	
HybTurb3D	Hybrid 3D simulations of turbulence and kinetic instabilities at ion scales in the expanding solar wind	Astro Sciences	IT	140 TB	140 TB	24 months	SurfSARA	CINECA
MULTINANO	Multiscale simulations of nanoparticle suspensions	Engineering	IT	30 TB	30 TB	24 months	MPCD	CINECA
HiResClimate	High Resolution EC-Earth Simulations	Earth Sciences	IE	150TB	150TB	12 months	KTH	EPCC
AFID	Effect of rotation and surface roughness on heat transport in turbulent flow	Engineering	NL	11TB	10TB for 24 months 1TB for long-term storage and publication	10TB for 24 months 1TB for long-term storage and publication	EPCC	SurfSARA
CHARTERED	Charge transfer dynamics by time dependent density functional theory	Materials Science	SE	30TB	30TB	24 months	IT4I	KTH/PDC
				Total	448TB			

Acronym	Title	Field	Country	Data requirements in TB during the PRACE project	Data requirements in TB after the PRACE project	Duration of the access to this data service after the PRACE allocation	PRACE	EUDAT Site
CHARTERED2	Charge transfer dynamics by time dependent density functional theory	Materials Science	SE	30TB	30TB	24 months	IT4I	KTH/PDC
subgridEoR	The effects of small-scale structure and halo stochasticity on Cosmic Reionization	Astrophysics		100TB	100TB	36 months	TBD	BSC or JSC
				Total	130TB			

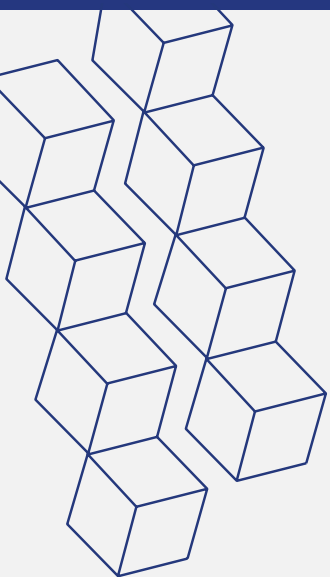
The goal of the workshop

- ✦ Present the EUDAT services and the interoperability possibilities between the two infrastructures
- ✦ Pilots status and achievement
- ✦ Hands-on on how to write a Data Management Plan (DMP)

Looking forward

- Collaboration formalised with EUDAT Ltd partner in PRACE-6IP
- **Data lifecycle coverage**
 - Support PRACE users in developing plans to cover the entire data lifecycle from creating, processing, analysing, to preserving data (data management plans)
 - Leverage new types of services which can be handled at Tier-1 level: data analytics, machine/deep learning, etc.
- **Resources access model**
 - Investigate suitability of peer-review access for computing & data & alternative models for well-organised communities
 - Investigate on-demand HPC / HPC as a service model (e.g. for AI or Deep Learning)





www.eudat.eu

 [@eudat_eu](https://twitter.com/eudat_eu)