Open Data, Open Infrastructures and Open Tools

The road ahead for the EOSC community
20 June 2019, Tallinn

eosc-hub.eu

@EOSC_eu

Tiziana Ferrari
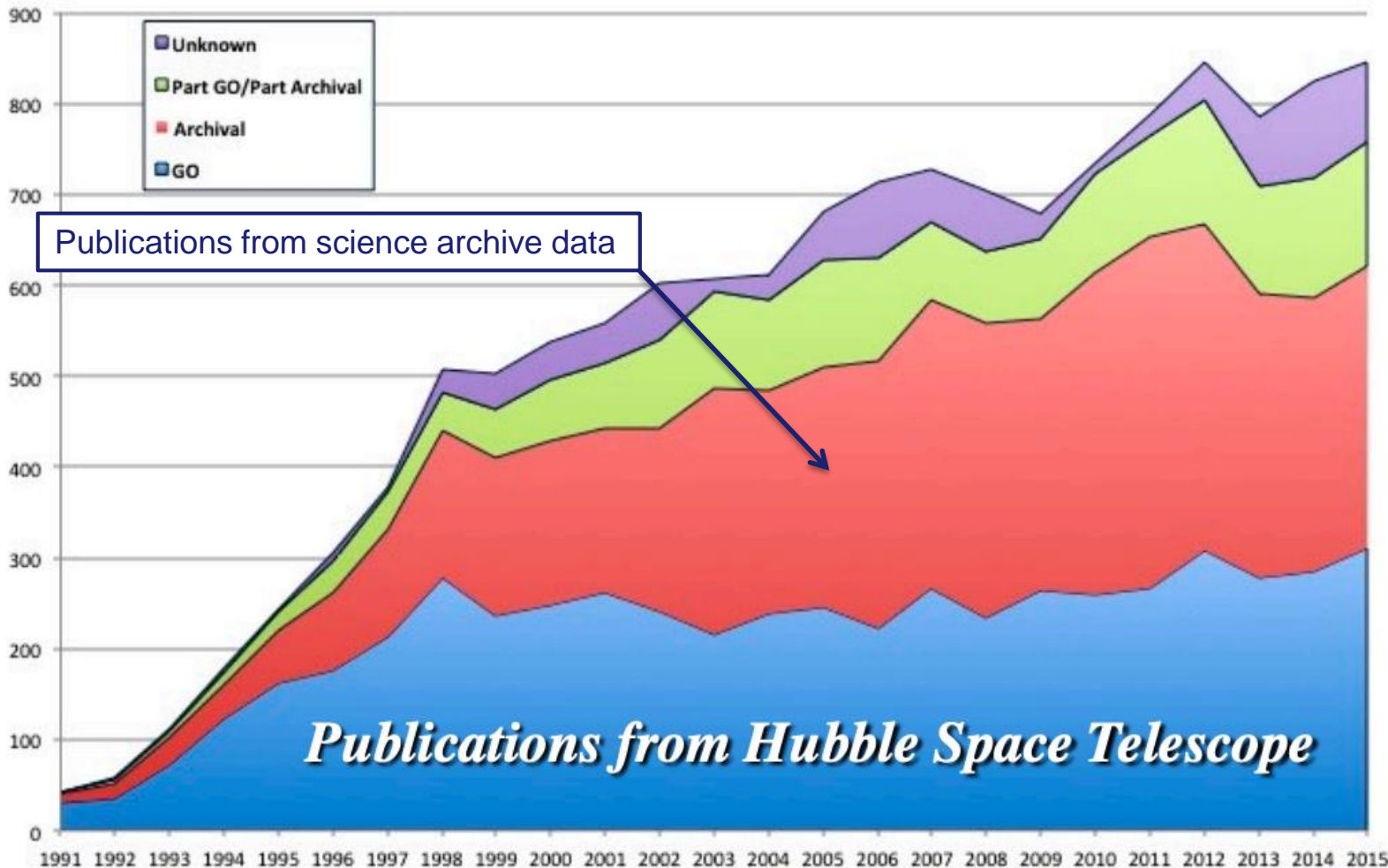*Tiziana.Ferrari@egi.eu*

EOSC-hub

- The movement to make scientific research and its dissemination **accessible to all levels** of an inquiring society, amateur or professional.

- Scientific products (publications, data, physical samples, and software) and capacities for their access and exploitation

All linked together

- A continuation of, rather than a revolution in research practices

EOSC-hub

Science archives are a multiplier for total science output

Publications from science archive data

Legend: Unknown, Part GO/Part Archival, Archival, GO

Publications from Hubble Space Telescope

- Assumes the archives are persistent and maintained
- Assumes archival data is open and accessible to users
- Assumes data products stored are appropriate for general use
- Assumes users retrieving data have resources to process to a science result

3

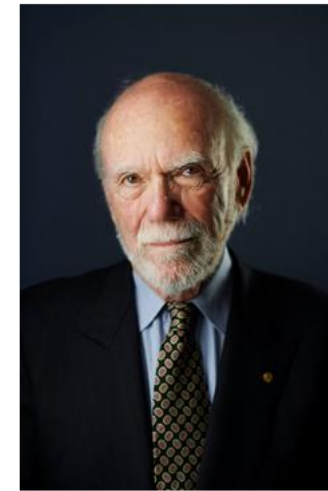# Open Data for Basic Science and Scientific Excellence

The LIGO-VIRGO Collaboraiton:
The discovery of gravitational waves

Expanding and exploring or explore the knowledge in physics with the LIGO-VIRGO collaboration

The Nobel Prize in Physics 2017

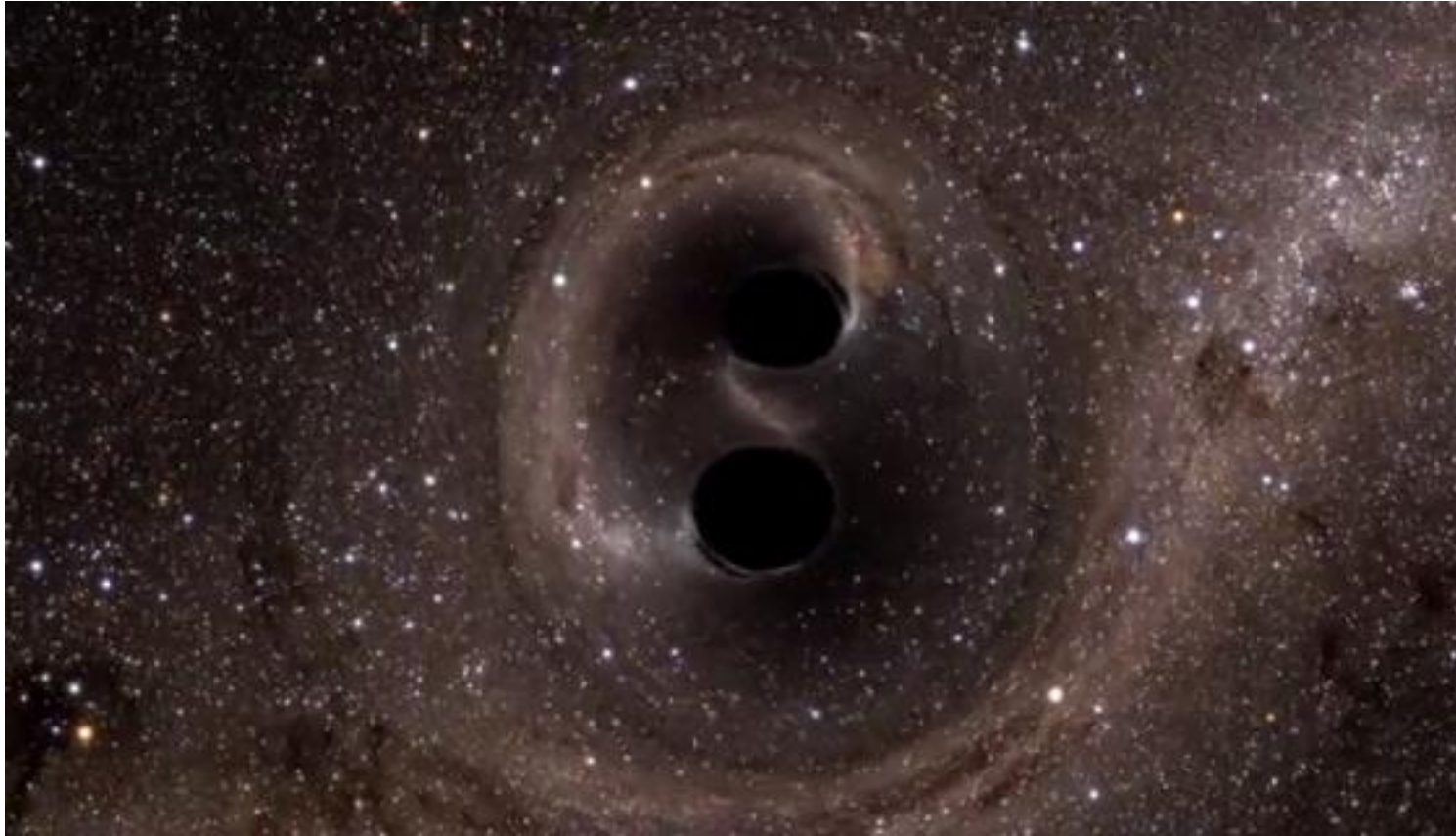© Nobel Media AB. Photo: A. Mahmoud
Rainer Weiss
Prize share: 1/2

© Nobel Media AB. Photo: A.Mahmoud
Barry C. Barish
Prize share: 1/4

© Nobel Media AB. Photo: A.Mahmoud
Kip S. Thorne
Prize share: 1/4

# Evidence 100 years after Einstein's theory of relativity

The masses of the two merging black holes — 29 and 36 solar masses, respectively. The single black hole that resulted from this smashup contains just 62 times the mass of the sun.

Sharing data across international user groups

The Virgo detector is located in Italy, within the site of the **European Gravitational Observatory.**


Elapsed time * Number of Processors (hours) by Year

2018
2019
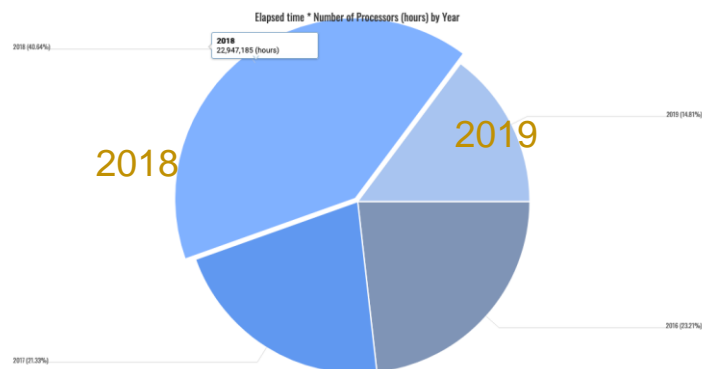
- The LIGO Scientific Collaboration
  - \> 1000 scientists from the USA and 14 other countries. The two LIGO detectors are located in Hanford and Livingston in the US.
- The Virgo Collaboration
  - \> 250 physicists and engineers affiliated with European
- Data from the LIGO and VIRGO detectors is shared and used for cross-correlation and validation of observation results
- 56 Million CPU hours in EGI (2016-2019)

Open Infrastructures
to Understand the Pathways to Disease

The WeNMR international community of practice

- *Proteins and biomolecules interact in a complex network*
- *Glitches in this network can cause diseases like cancer*
- *Researchers use 3D models to study how protein and biomolecules interact*
- *This structural information is key to understand the origins of disease and to develop new drugs*

Technical and Policy Challenges
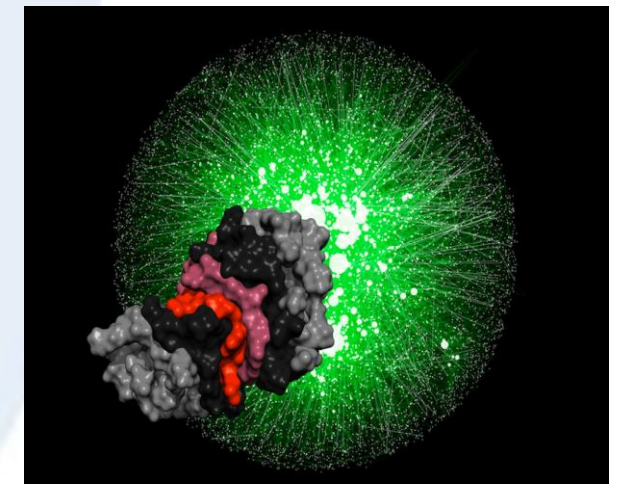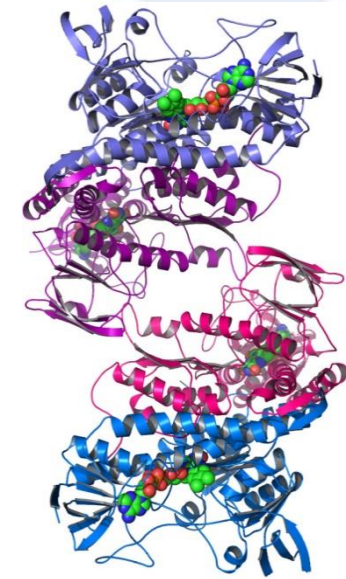
- The analysis and interpretation of the results relies on specialised software

- Standalone software applications are difficult to install and maintain

- The complex workflows involved in 3D structural modelling are difficult to manage

- National digital infrastructures are not open by default to international research groups ➜ Need for a long-term resource, data and platform infrastructure

- *Access to compute resources is difficult to obtain, negotiate and/or guarantee for extended periods of time ➜ EGI Federation (EU) and Open Science Grid (USA)*

Sharing and federation of compute and storage capacity worldwide

The worldwide resource platform
hosting WeNMR applications, computing and data

Shared and federated resources for the Structural Biology WeNMR community via

USA

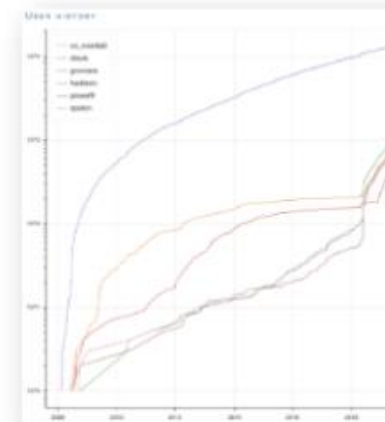Europe, Africa and Asia Pacific Region

| 60 Million CPU Hours | 160 cloud CPU cores | 250 TB storage | > 13500 registered users |

WeNMR: An international User Community of > 13,000 researchers

Sustained growth of our user base

https://wenmr.science.uu.nl/user_map

	Country	All_Users ▽	HADDOCK	DISVIS	POWERFIT	SPOTON	CS_ROSETTA3	GROMACS
1	Total Users	13,647	13,108	861	611	688	643	421
2	EU Users	3,085	2,873	244	141	154	158	84
3	India	2,947	2,897	119	98	133	97	107
4	United States	2,166	2,081	147	86	111	93	45

Open Tools for Evidence-based Policy Making
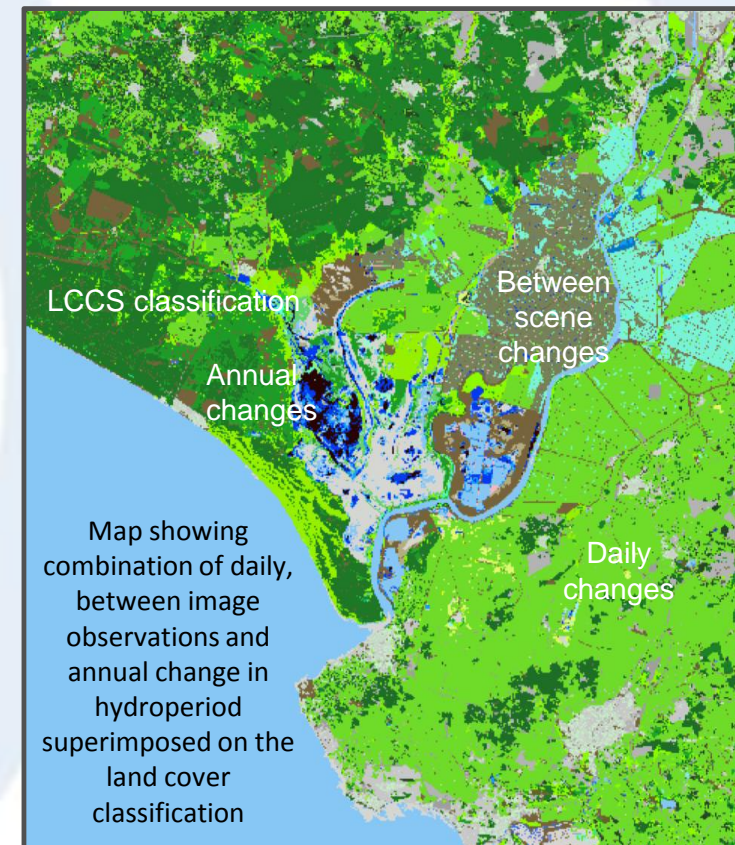
Protecting European Landscapes

The Earth Observation Data for Ecosystem Mapping System

Europe's protected areas and surrounds are vulnerable to human-induced events and processes including those associated with climate change.
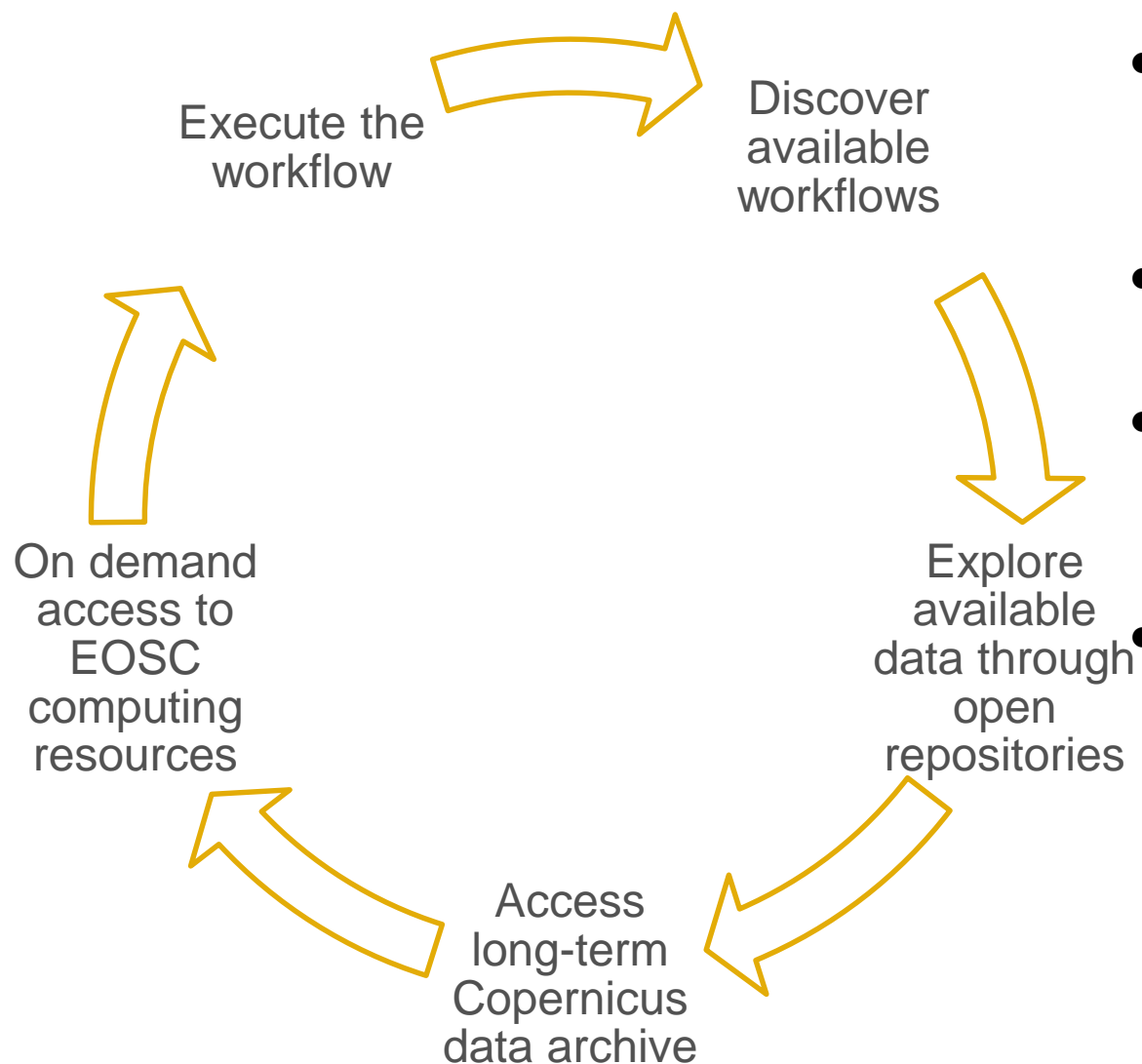
As an example, the biodiversity of Doñana National Park in Spain is affected by agricultural expansion, and urban development and climatic variability with all impacting on water supply.

Whilst large amounts of Earth Observation data are available, these are vastly underexploited for monitoring changes and implementing management actions aimed at nature conservation and sustainable and wise use.


LCCS classification

Between scene changes

Annual changes

Daily changes

Map showing combination of daily, between image observations and annual change in hydroperiod superimposed on the land cover classification

Barriers and the EODESM approach

EOSC-hub

Execute the workflow → Discover available workflows → Explore available data through open repositories → Access long-term Copernicus data archive → On demand access to EOSC computing resources → (cycle)
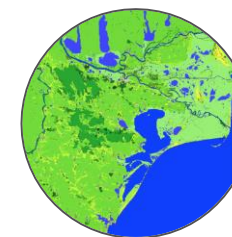
- Ready and easy access to the large volumes of an increasingly diverse Earth Observation resource and processing systems.
- Lack of capacity for discovery and integrated analyses of EO and other spatial data.
- Difficulty in grasping the concepts behind EO data and techniques for classifying land covers and monitoring change.
- Sharing idea and concepts across domains and between teams.
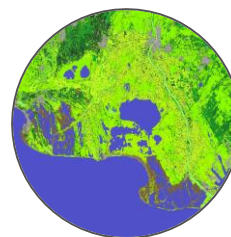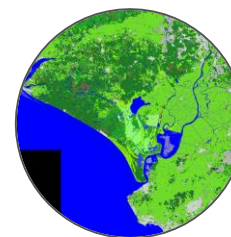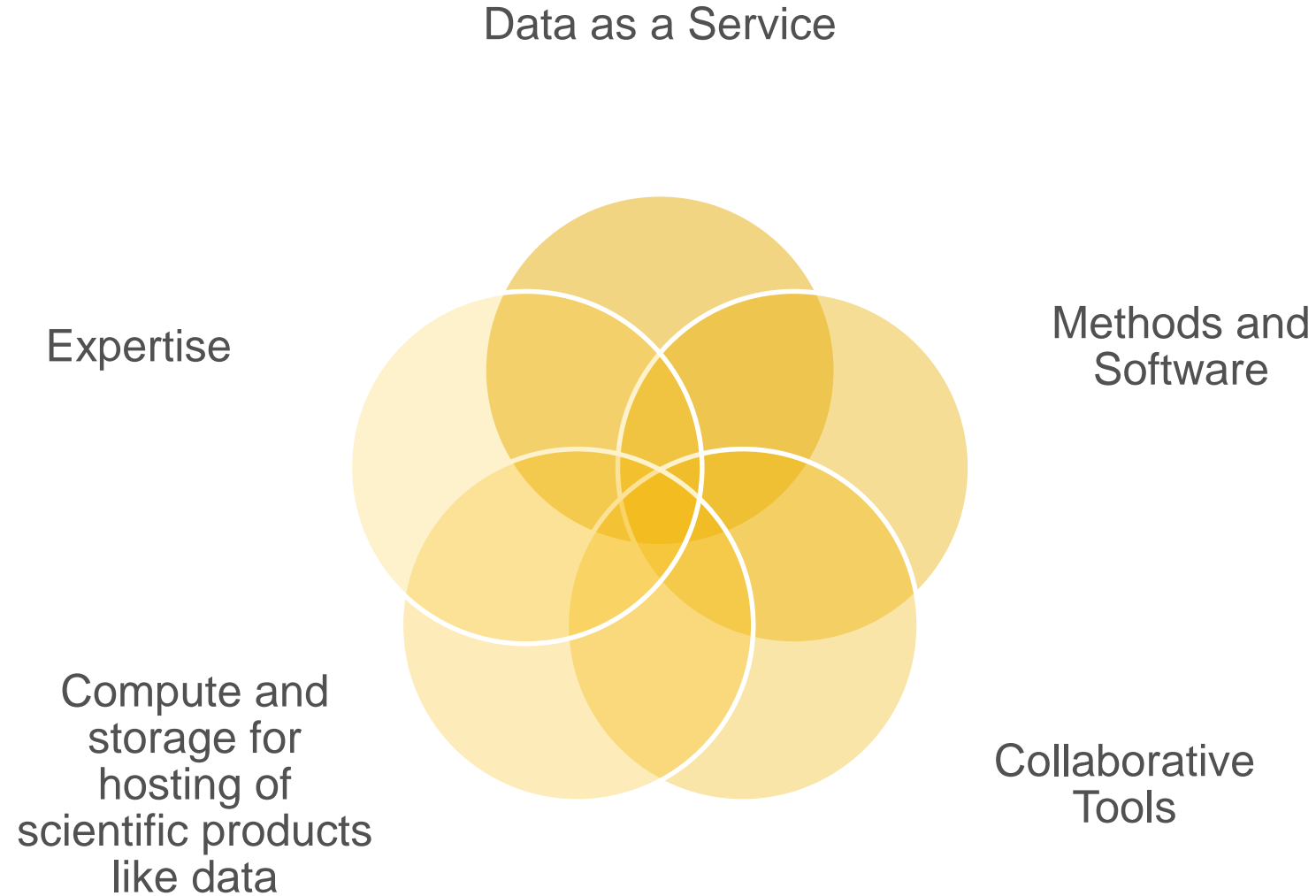
Camargue, France   Donana Spain   Danube Romania/Ukraine
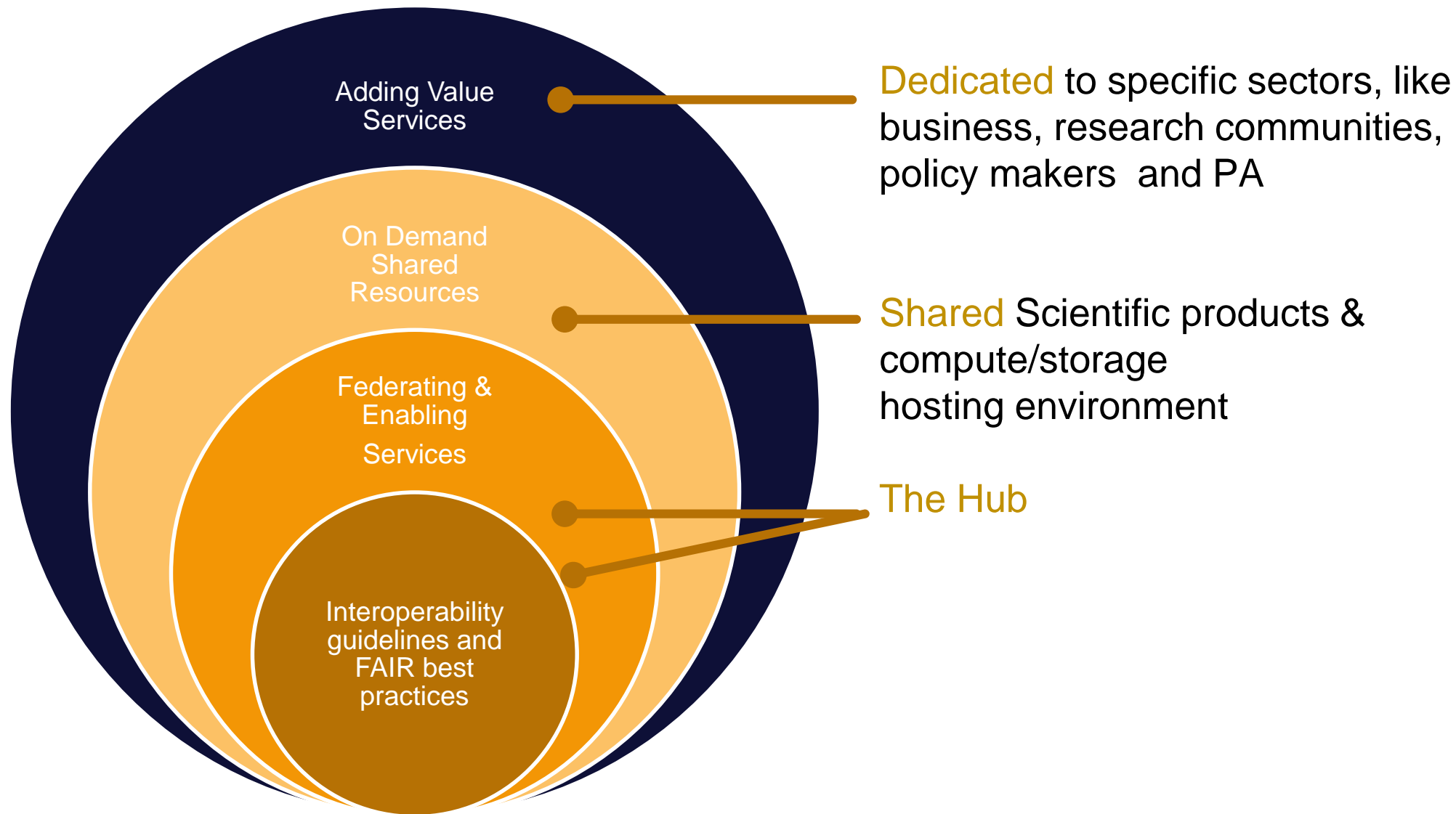
EOSC from design to implementation

About the European Open Science Cloud
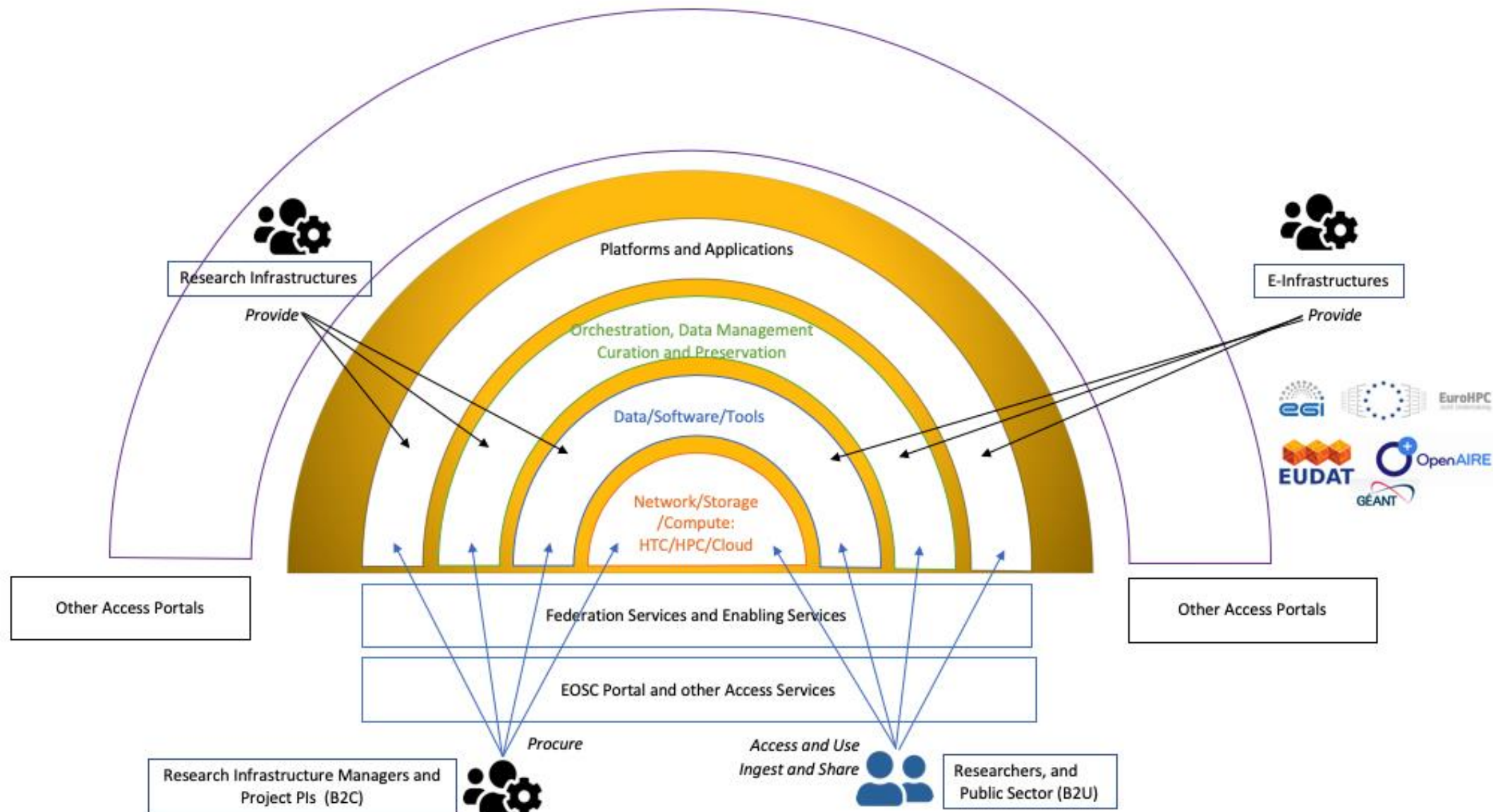
The federated infrastructure and supporting initiative providing all researchers, innovators, companies and citizens with seamless access to an open-by-default, efficient and cross-disciplinary environment

for storing, accessing, reusing data, tools, publications and other scientific outputs for research, innovation and educational purposes

Credits: EOSCpilot (https://eoscpilot.eu/)

EOSC-hub contributions for Open Science

Data as a Service

Expertise

Methods and Software

Collaborative Tools

Compute and storage for hosting of scientific products like data

EOSC From Design to Implementation

EOSC-hub

Dedicated to specific sectors, like business, research communities, policy makers and PA

Shared Scientific products & compute/storage hosting environment

The Hub

- Adding Value Services
- On Demand Shared Resources
- Federating & Enabling Services
- Interoperability guidelines and FAIR best practices

EOSC-hub contributions to EOSC

- EOSC Portal Website, Marketplace and service provider onboarding
- Portfolios
  - Hub Portfolio
  - EOSC Service Portfolio
  - Enabling and integration activities
- Business plans and roadmaps on service sustainability
- The Hub Service Management System

- Rules of Participation
  - For Users and Providers
- Technical Interoperability Guidelines
- Digital Innovation Hub
- Competence Centres
- Knowledge Management & Training
- Procurement (EOSC-hub – OCRE collaboration)

Thank you
for your attention!

_____

Questions?

EOSC-hub

🔗 eosc-hub.eu 🐦 @EOSC_eu