



Overview of the European Open Science Cloud

Jan Wiebelitz
e-IRG support programme

Historical view

- Oct 2010 **Riding the wave report**, HLEG on Scientific Data
- Sept 2012 COM(2012) 529 final **Unleashing the Potential of Cloud Computing in Europe**
- Dec 2012 **e-IRG Roadmap 2012** proposed an e-Infrastructure Commons
- Sep 2015 EGI published its **Open Science Commons**
- Sep 2015 **HLEG report on the European Open Science Cloud**
- Apr 2016 COM(2016) 178 final **European Cloud Initiative – Building a competitive data and knowledge economy in Europe**
- Work Programme 2016-2017 Call: **INFRADEV-04-2016: European Open Science Cloud for Research**
- May 2017 COM(2017) 228 final on the **Mid-Term Review on the implementation of the Digital Single Market**
- Oct 2017 **EOSC Declaration**



Realising the European Open Science Cloud

- Report and recommendations of the HLEG on the European Open Science Cloud
 - Build on existing infrastructure and expertise
 - Revise Rules of Engagement
 - Federate the gems and amplify good practise
 - EU contribution to an Internet of FAIR data and services
 - Develop core data expertise
 - Make data stewardship mandatory



EOSC Political drivers



"We must create infrastructure. Europe's final transition must be one from fragmented datasets to an integrated European Open Science Cloud. By 2020, we want all European researchers to be able to deposit, access and analyse European scientific data through a European Open Science Cloud"

Commissioner Carlos Moedas, Amsterdam, 4 April 2016

EC Communication "European Cloud Initiative" of 19/4/16

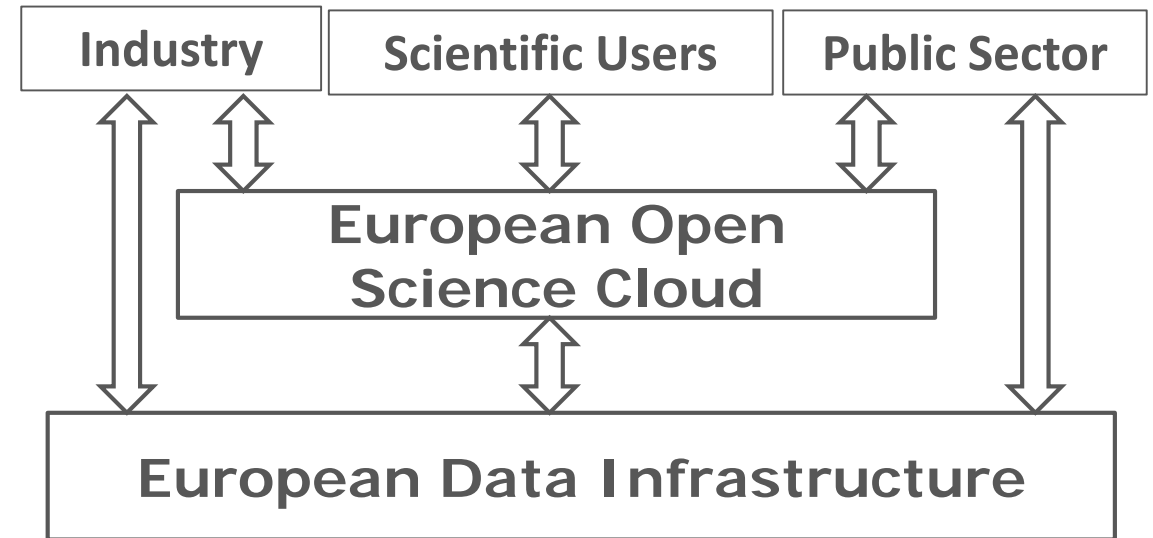
Three pillars:

- EOSC: the European Open Science Cloud
- EDI: the European Data Infrastructure (Development and deployment of large-scale EU HPC, data and network infrastructure)
- Widening access & building trust (SMEs, Industry, Government)

Slide adapted from Lorenza Saracco, European Commission, presented at the EOSCpilot kickoff meeting

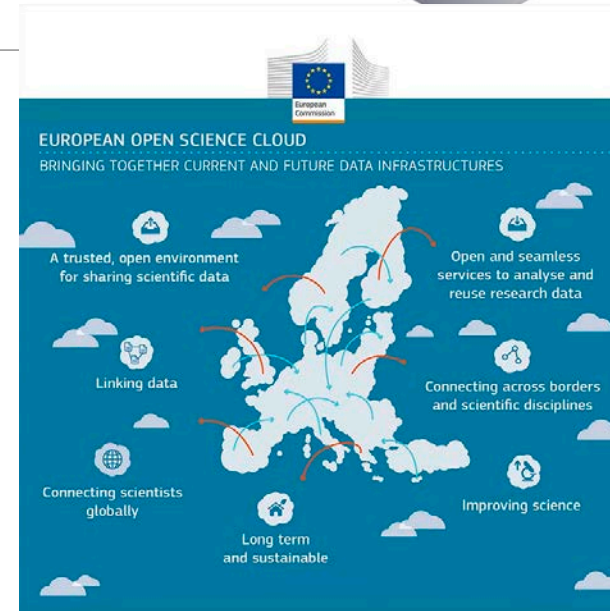
European Cloud Initiative

- develop a trusted, open environment for the scientific community for storing, sharing and re-using scientific data and results, the **European Open Science Cloud**
- deploy the underpinning super-computing capacity, the fast connectivity and the high-capacity cloud solutions they need via a **European Data Infrastructure**



European Cloud Science Cloud

- Make all scientific data Produced by H2020 Programme **open by default**
- Raise awareness and change **incentive structure**
- Develop specifications for **interoperability and data sharing** across disciplines and infrastructures
- Create a fit-for-purpose pan-European **governance structure**
- Develop **cloud-based services** for Open Science
- **Enlarge the scientific user base ...** to researchers and innovators from all disciplines



COM (2016) 178 final European Cloud Initiative

INFRADEV-04-2016: European Open Science Cloud for Research



The aim of this topic is the **launch of a pilot action** that should demonstrate how wide availability of scientific data and data-analysis services for European researchers can be ensured through a cloud infrastructure.

Expected Impact:

- Facilitate **access of researchers** across all scientific disciplines **to** the broadest possible **set of data and to other resources needed for data driven science** to flourish.
- Provide a **governance and business model** that **sets the rules for the use of data**, deals with issues related to privacy, sensitivity, copyright and security and **oversees the provision of services** (business and governance layer).
- **Create** a cross-border and multi-disciplinary **open innovation environment for research data, knowledge and services** with engaged stakeholders and organisations.
- Foster the establishment of global standards, ontologies and interoperability for scientific data.

EOSCpilot project

The EOSCpilot project will support the first phase in the development of the EOSC.

- Propose and trial a **governance framework** for the EOSC
- Contribute to **policy and best practice in open science**
- Develop a number of **demonstrators** functioning as high-profile pilots that integrate services and infrastructures **to show interoperability and its benefits** in a number of scientific domains
- Engage with a broad range of stakeholders, crossing borders and communities, **to build the trust and skills** required for adoption of an open approach to scientific research.



<https://eoscpiot.eu/>

Science Demonstrators in EOSCpilot

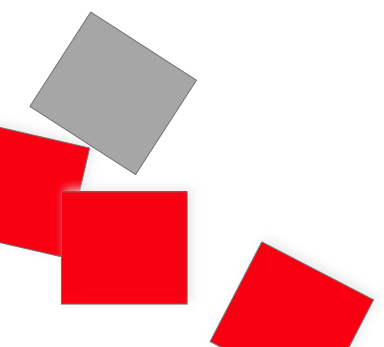


- **Environmental & Earth Sciences** - ENVRI Radiative Forcing Integration to enable comparable data access across multiple research communities by working on data integration and harmonised access
- **High Energy Physics** – DPHEP/WLCG: large-scale, long-term data preservation and re-use of physics data through the deployment of HEP data in the EOSC open to other research communities
- **Social Sciences** – TEXTCROWD: Collaborative semantic enrichment of text-based datasets by developing new software to enable a semantic enrichment of text sources and make it available on the EOSC.
- **Life Sciences** - Pan-Cancer Analyses & Cloud Computing within the EOSC to accelerate genomic analysis on the EOSC and reuse solutions in other areas (e.g. for cardiovascular & neuro-degenerative diseases)
- **Physics** - The photon-neutron community to improve the community's computing facilities by creating a virtual platform for all users (e.g., for users with no storage facilities at their home institutes)

EINFRA-12-2017: Data and Distributed Computing e-infrastructures for Open Science



- ***integrate at European level** the geographically and disciplinary dispersed **resources** to achieve economies of scale and efficiency gains in providing the **best data and computing capacity and services** to the research and education communities*
- **EOSC-hub project**
 - Start in early 2018
 - builds on **existing technology** already at TRL 8
 - address the need for **interoperability** by promoting the adoption of open standards and protocols.
 - integrated **catalogue of services, software and data**



OpenAIRE

- Series of projects (Dec 2009 – Dec 2020)
- Support and services for Open Access policies, for both publications and data,
- Implementation of an Open Access Pilot
- Implementation of National Open Access Desks
- Open Science as a Service (OSaaS)
 - publish research artefacts
 - monitor their research impact



<http://openaire.eu>

eInfraCentral - European e-Infrastructure Services Gateway



Mission: to ensure that, by 2020 a broader and more varied set of users (including industry) **discovers and accesses the existing and developing e-infrastructure capacity**.



A common approach to defining and monitoring e-infrastructures services will increase the uptake of and enhance understanding of where improvements can be made in delivering e-infrastructure services.

- **eInfraCentral portal**
 - single entry point (one-stop shop) for end users
 - main entry point to European e-infrastructures services
 - monitoring of key performance indicators (KPIs)
- **Service catalogue alignment**
- **KPIs for e-infrastructures services**

<http://einfracentral.eu/>

EDISON - Education for Data Intensive Science to Open New science frontiers



EDISON was a 2-year project (started September 2015) with the purpose of accelerating the creation of the Data Science profession.

- create a foundation for establishing a new profession of Data Scientist
- vision was to build a curriculum for Data Scientist
- EDISON Data Science framework (EDSF)
 - targeted education and training,
 - professional certification,
 - organizational and individual skills management and
 - career transferability.



EDISON

building the data
science profession

<http://edison-project.eu/>

Go-FAIR Initiative

Go-FAIR is a proposal for the practical implementation of the EOSC

Go-FAIR consists of three interconnected pillars:

- **GO CHANGE** aims to instigate cultural change to make the FAIR principles a working standard in science and to reform reward systems to incorporate open science activities.
- **GO TRAIN** is about locating, creating, maintaining, and sustaining the required data expertise in Europe through training and education. The aim is to have core certified data experts and to have at least one certified institute in each Member State and for each discipline to support implementation of data stewardship.
- **GO BUILD** deals with the need for interoperable and federated data infrastructures. In addition, it is about the harmonisation of standards, protocols, and services, which enable all researchers to deposit, access, and analyse scientific data across disciplines.



Support on national Secretary of State level by the Netherlands and Germany (joint position paper)

<https://www.dtls.nl/fair-data/go-fair/>

H2020 Mid-Term Review

Conclusion:

[...]

This Mid-term Review makes clear that there is **no more time to lose to turn political commitments into reality**. It calls for the focus to stay on the big things that require a common response, and substantial investments in infrastructures and skills

[...]



COM (2017) 228 final Mid-Term Review on ...

EOSC Declaration



EOSC Summit of 12 June 2017

- more than 80 key scientific stakeholders
- strong sense of commitment, dedication and intellectual rigour with respect to the implementation of the EOSC



RECOGNISING the challenges of data driven research in pursuing excellent science;

GRANTING that the vision of European Open Science is that of a research data commons, widely inclusive of all disciplines and Member States, sustainable in the long-term,

CONFIRMING that the implementation of the EOSC is a process, not a project, by its nature iterative and based on constant learning and mutual alignment;

UPHOLDING that the EOSC Summit marked the beginning and not the end of this process, one based on continuous engagement with scientific stakeholders, the European Commission,

PROPOSES that all EOSC stakeholders consider sharing the following intents and will actively support their implementation in the respective capacities:

EOSC Declaration - Data Culture and FAIR data



- **Common culture of data stewardship** for long term re-use of research data
- **Open-by-default** access to an research data environment
- **Necessary skills and education** in research data management, data stewardship and data science should be provided throughout Europe
- **Rewarding Open Data and FAIR Data** in career assessment and project evaluation
- **FAIR principles** should be applied to research data but also to data related algorithm, tools, workflows, protocols, services, etc.
- The use of **Data Management Plans** should become obligatory

EOSC Declaration - Research data services and architecture



- The EOSC will be developed as a **data infrastructure commons** serving the needs of scientists.
- The EOSC will **federate existing resources** across national data centres, European e-infrastructures and research infrastructures
- **Users should contribute** to define the main common functionalities needed by their own community.
- The EC plans to propose, by end of 2017, a legal instrument that provides a **procurement framework** for the exascale supercomputing and data infrastructure.

EOSC Declaration - Governance and funding

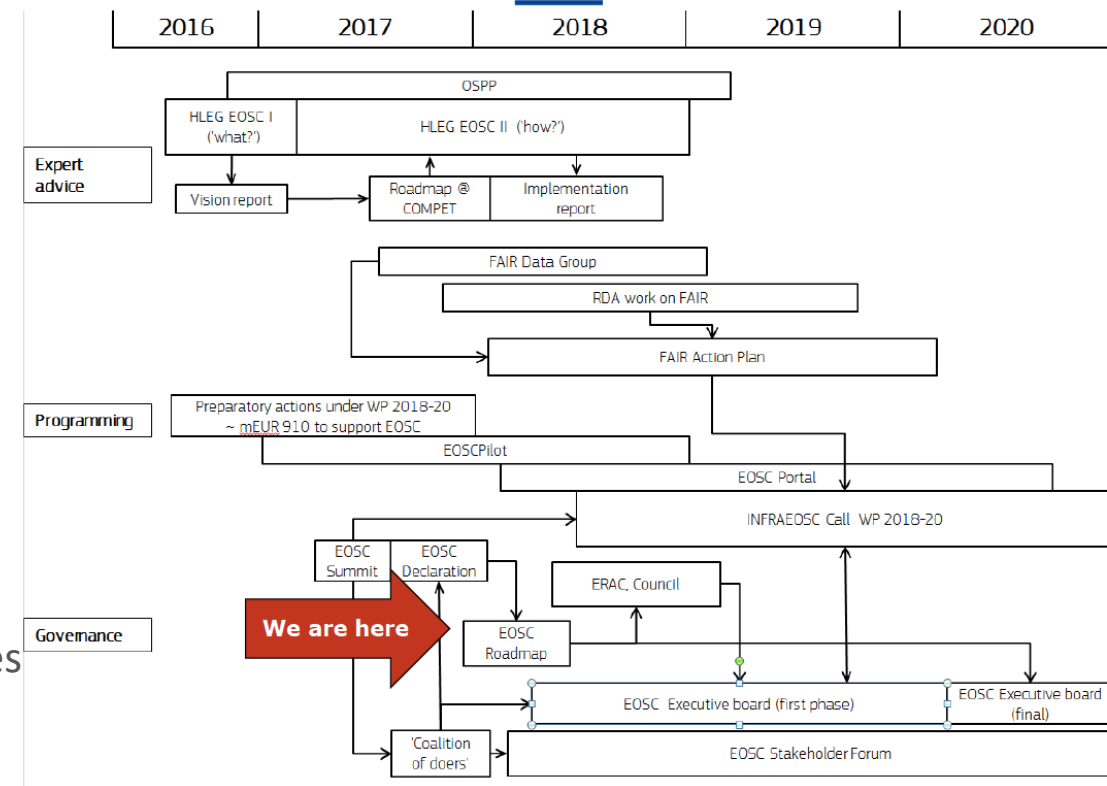


- The EOSC governance model should be based on **representativity, proportionality, accountability, inclusiveness and transparency**
- EOSC governance framework will be **co-designed, stakeholder driven** and composed of three main layers:
 - institutional, including EU Member States and European Commission
 - operational, including a governance board and relevant working committees (e.g. thematic and functional)
 - advisory, including a stakeholder forum
- The European Commission, Member States and Research Funders will use existing and future resources strategically, to ensure **long-term sustainability** of open research data and research infrastructures, facilitating inter-disciplinarity

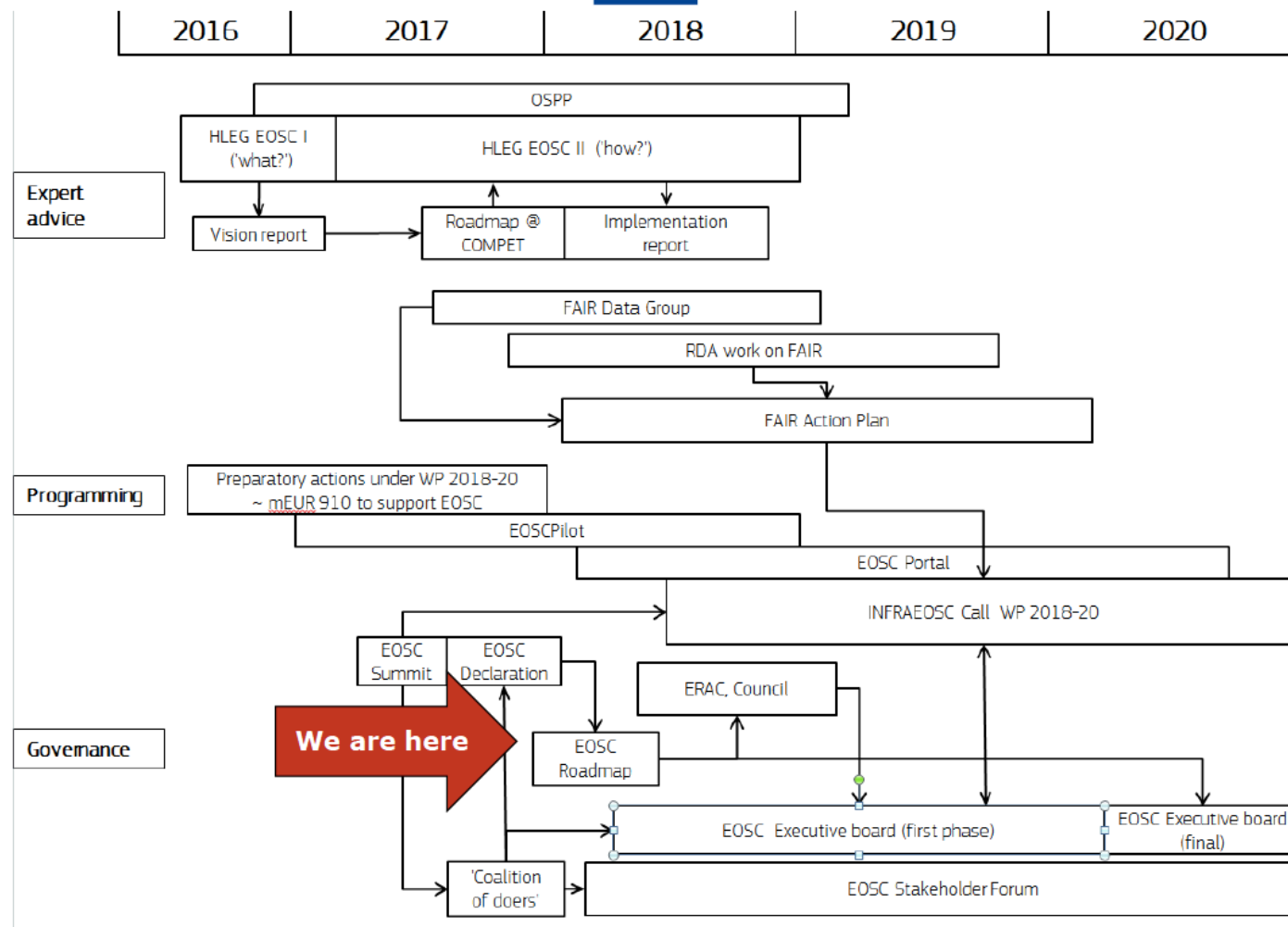
EOSC Timeline

- EOSC Declaration (Sept 2017)
 - key input for the Roadmap
 - for agreement and poss. specific commitments by stakeholders
 - basis for follow-up discussion with MS (ERAC SWG 5+)
- EOSC Stakeholders Forum (Nov 2017)
 - open to ALL categories represented at the Summit & endorsing EOSC Declaration
 - by end August, EC to define application procedure & working modalities
 - initially based on EOSCpilot project (Stakeholder Engagement Event 28-29 Nov)
- EOSC Roadmap (Dec 2017)
 - Governance structure, incl. mandate & selection procedure for Executive Board
 - Broad (federated) architecture, incl. EOSC core services
 - Financing

EOSC Timeline



EOSC Timeline



EOSC Roadmap

		2017	2018	2019	2020	2021	2022
EOSCpilot							
EOSC-Hub							
Openaire							
RDA/FREYA/etc							
INFRAEOSC-01							
INFRAEOSC-04							
INFRAEOSC-05 (Gov and Fair)							
INFRAEOSC-02							
INFRAEOSC-05 (RIA)							
INFRAEOSC-06							

Events: EOSC stakeholder meeting

- Communicate the early results of the EOSCpilot project,
- Identify possible EOSC shapes and features
- Discuss the practical possibilities of the EOSC and identify minimal viable products to deliver European added value
- Prioritize the implementation roadmap
- Present and discuss possible governance, funding and business models
- Share Good Practices and local initiatives fostering Open Science, which could be generalized to the EOSC



<https://eosc-pilot.eu/events/eosc-stakeholder-forum-shaping-future-eosc>

Events: DI4R 2017 conference

“Connecting the building blocks for Open Science”

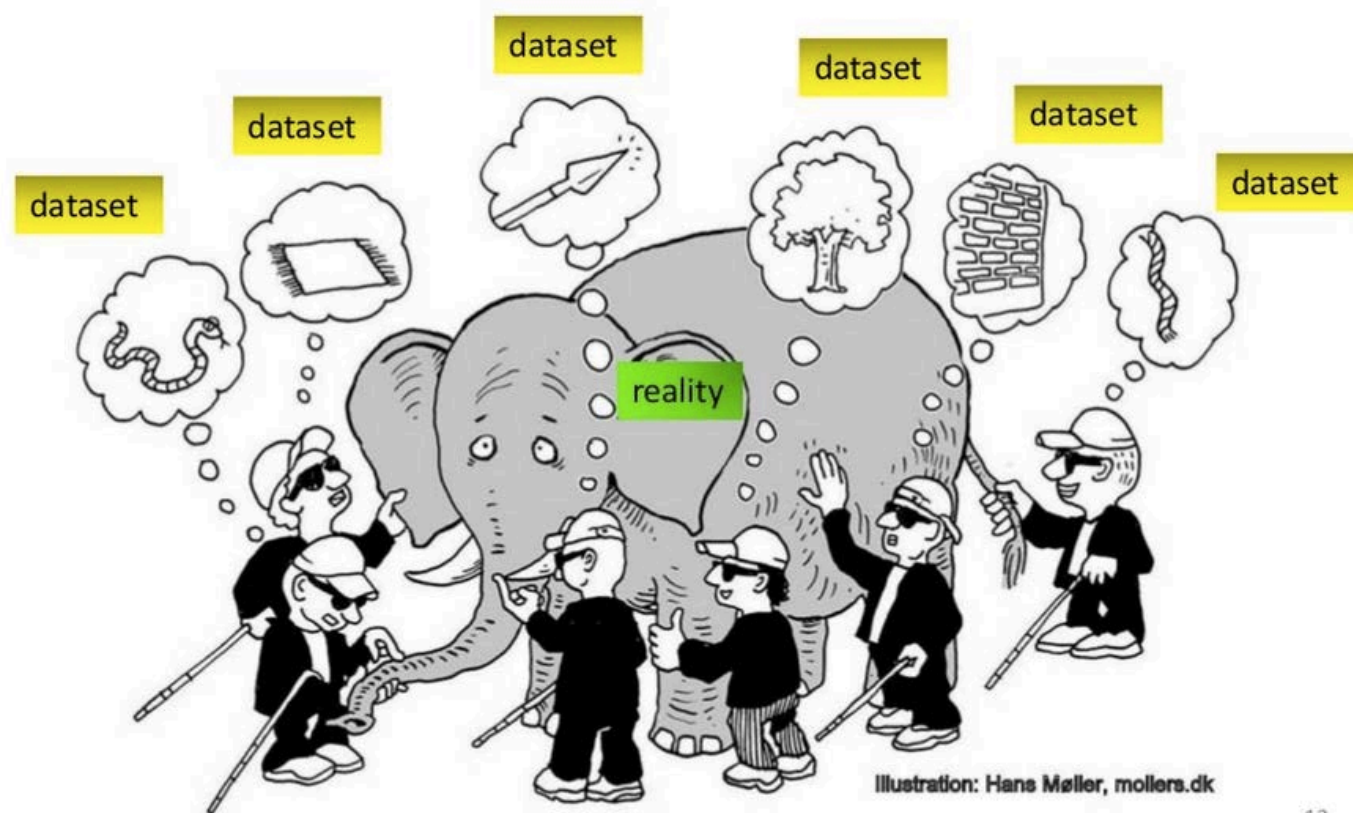
- jointly organised by EGI, EUDAT, GÉANT, OpenAIRE, PRACE and RDA Europe
- showcase the policies, processes, best practices, data and services
- The overarching goal is to demonstrate how open science, higher education and innovators can benefit from these building blocks, and ultimately to advance integration and cooperation between initiatives.



Why integration is difficult



Why integration is difficult?
.... complementary views



12

Yannis Tzitzikas, EUDAT Summer School, 3-7 July 2017, FORTH

References

- Riding the wave - How Europe can gain from the rising tide of scientific data
http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=707
- The Data Harvest Report – sharing data for knowledge, jobs and growth
<https://www.rd-alliance.org/data-harvest-report-sharing-data-knowledge-jobs-and-growth.html>
- e-IRG Roadmap 2012, White Paper 2013, Roadmap 2016
<http://e-irg.eu/publications>
- COM(2016) 178 final European Cloud Initiative – Building a competitive data and knowledge economy in Europe
http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15266
- Cloudy, increasingly FAIR; revising the FAIR Data guiding principles for the European Science Cloud
<https://espace.curtin.edu.au/bitstream/handle/20.500.11937/53669/253006.pdf?sequence=2&isAllowed=y>
- Go-FAIR joint position paper of the Netherlands and Germany
<https://www.government.nl/documents/reports/2017/05/30/joint-position-paper-on-the-european-open-science-cloud>
- Open Science Commons
<http://go.egi.eu/oscwp>
- Realising the European Open Science Cloud (HLEG report)
https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf
- EOSC Declaration
https://ec.europa.eu/research/openscience/pdf/eosc_declaration.pdf#view=fit&pagemode=none