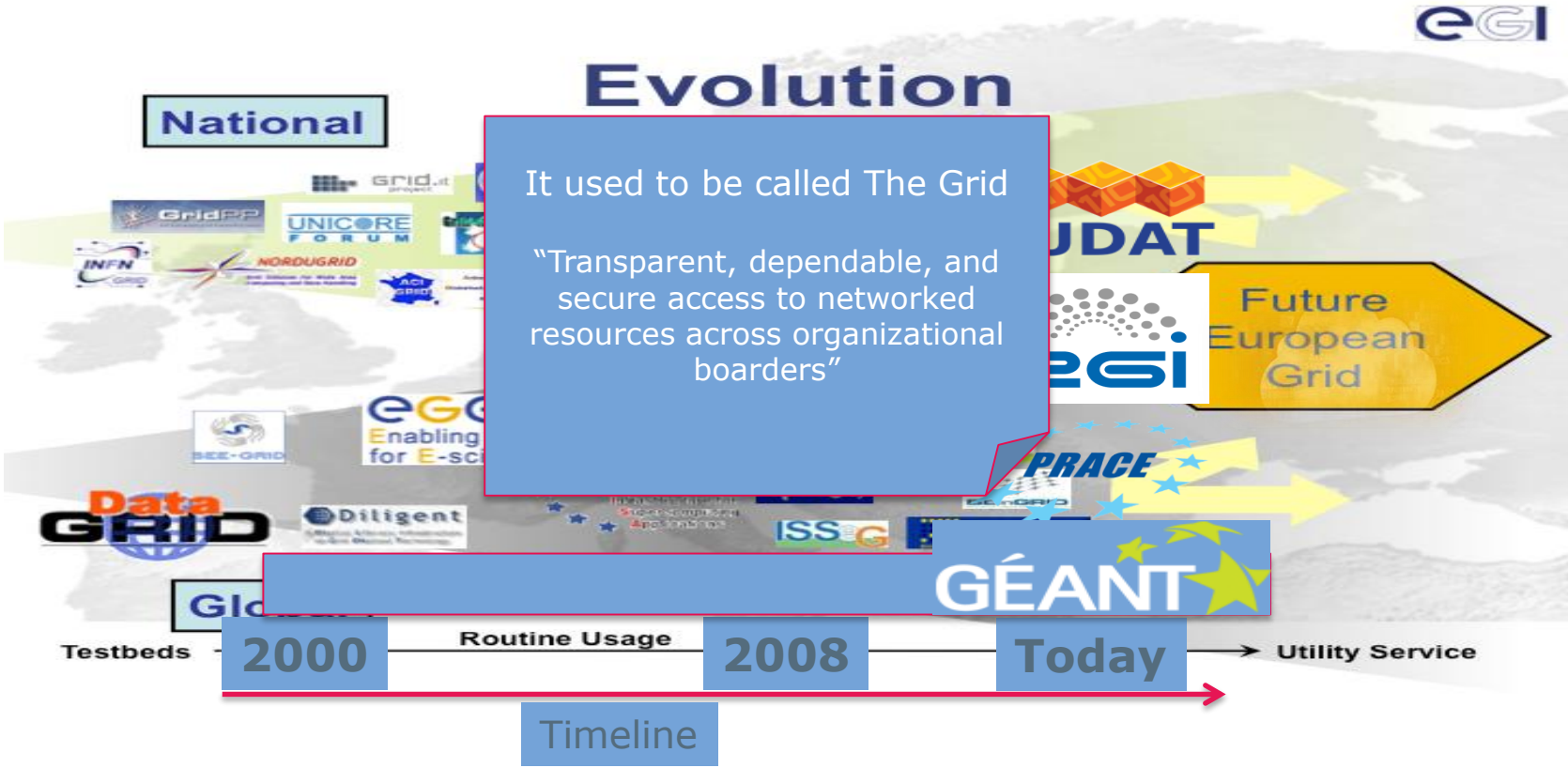
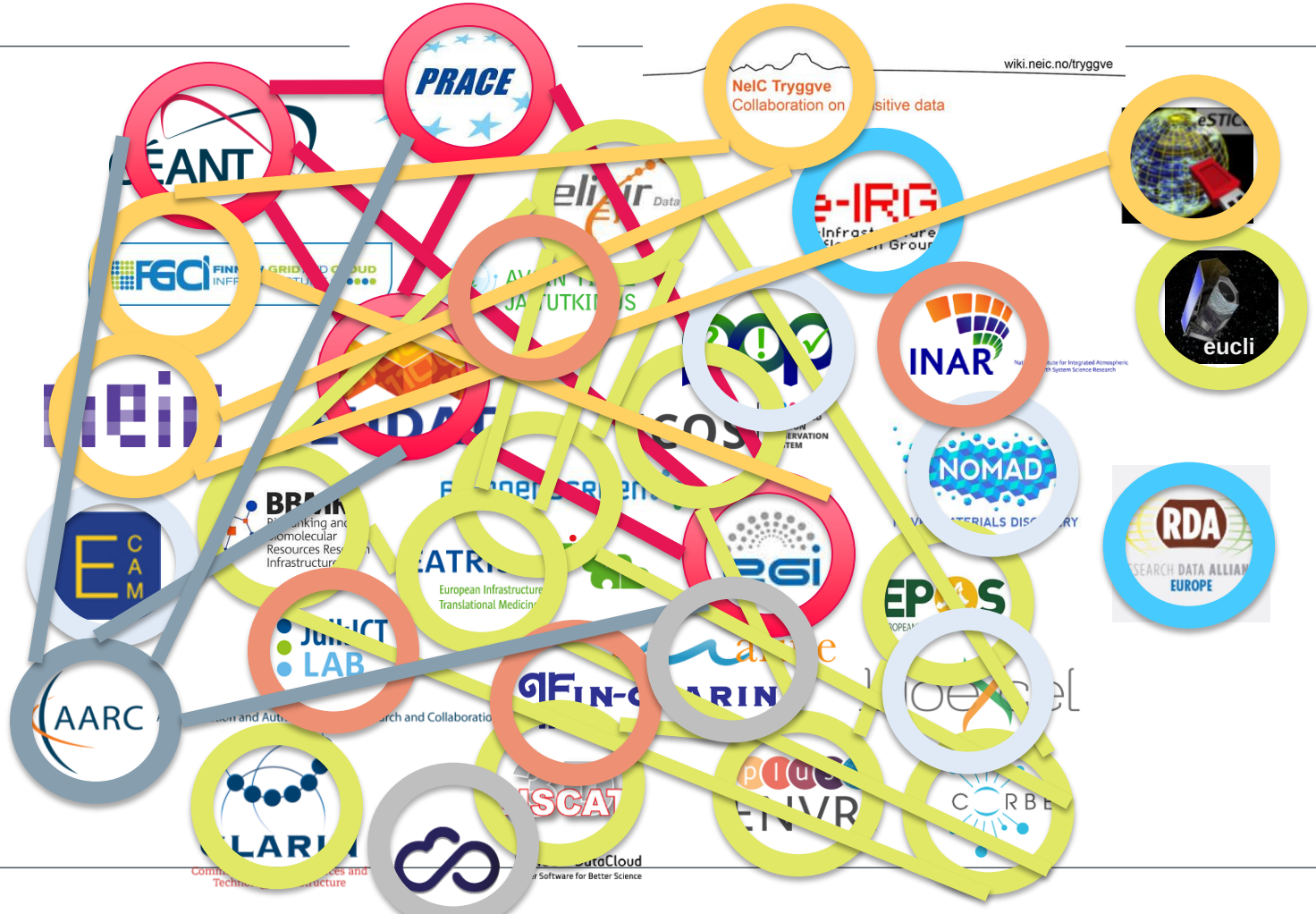
Jisc

i

25 April 2016

European Open Science Cloud


Evolution
National
It used to be called The Grid
"Transparent, dependable, and secure access to networked resources across organizational boarders"
Future European Grid
PRACE
GÉANT
Testbeds
2000
Routine Usage
2008
Today
Utility Service
Timeline


wiki.neic.no/tryggve
PRACE
NeIC Tryggve
Collaboration on
sitive data
eSTIC
eucli
INAR
NOMAD
RDA
EUROPE
FGCI
AARC
LAB
European Infrastructure
Translational Medicin
Biomolecular
Resources Res
Infrastructure
CORBE
Software for Better Science

A Common Endeavour (EU Perspective)

Context: European Research Area

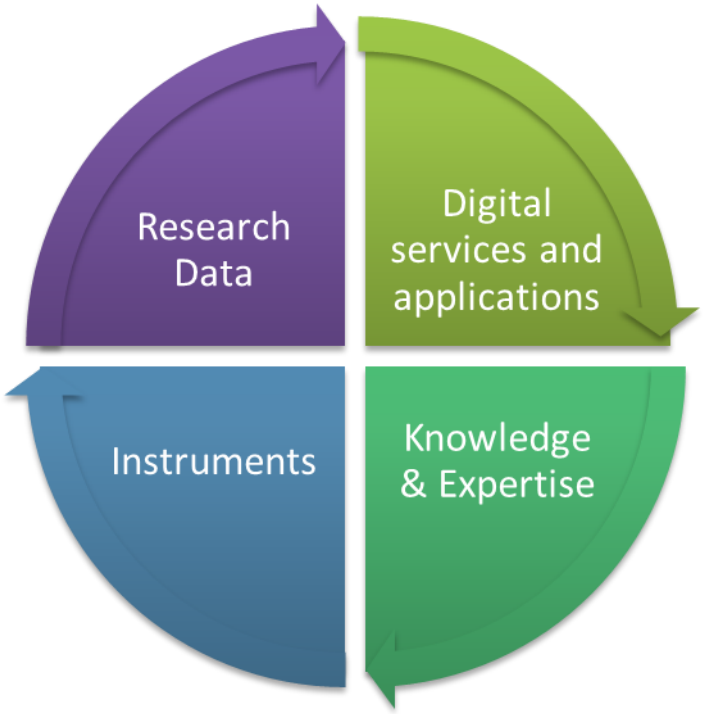
http://ec.europa.eu/research/era/

Context: Open Science

***Opening** of the **creation** and **dissemination** of **scholarly knowledge** towards a multitude of stakeholders, from professional researchers to citizens*

It needs:

- Shared resources
  - Integrated, easy and fair access
- Engaged communities
  - Participating in the process
  - Collaborating in the management and stewardship
- Governance
  - Rules to access/exclude
  - Rules to resolve conflicts
- Financial support
  - For long-term availability



Context: Philosophy of Open Science

School of thought	Involved groups	Central assumption	Central Aim	Tools & Methods
Democratic	Scientists, politicians, citizens	The access to knowledge is unequally distributed	**Making knowledge freely available for everyone**	Open access, intellectual property rights, Open data, Open code
Public	Scientists & citizens	Science needs to be made accessible to the public	**Making science accessible for citizens**	Citizen Science, Science PR, Science Blogging
Infrastructure	Scientists & platform providers	Efficient research depends on the available tools, applications and shared infrastructures	**Creating openly available platforms, tools and services for scientists**	Collaboration platforms, tools, computing platforms
Pragmatic	Scientists	Knowledge creation could be more efficient if scientists collaborated	**Opening up the process of knowledge creation**	Wisdom of the crowds, network effects, Open Data, Open Code
Measurement	Scientists & politicians	Scientific contributions today need alternative impact measurements	**Developing an alternative metric system for scientific impact**	Altmetrics, peer review, citation, impact factor

Source: Opening Science book, 2013

Enter the "Commons"

Institutionalised community governance of the production and/or sharing of a particular type of resource (from natural to intellectual)

GÉANT: European Communications Commons

Open Science Commons

Constructing Genome Commons

Internet

...

e-Infrastructure Commons

Jisc: National e-Infrastruture Commons

Wikipedia

Linux

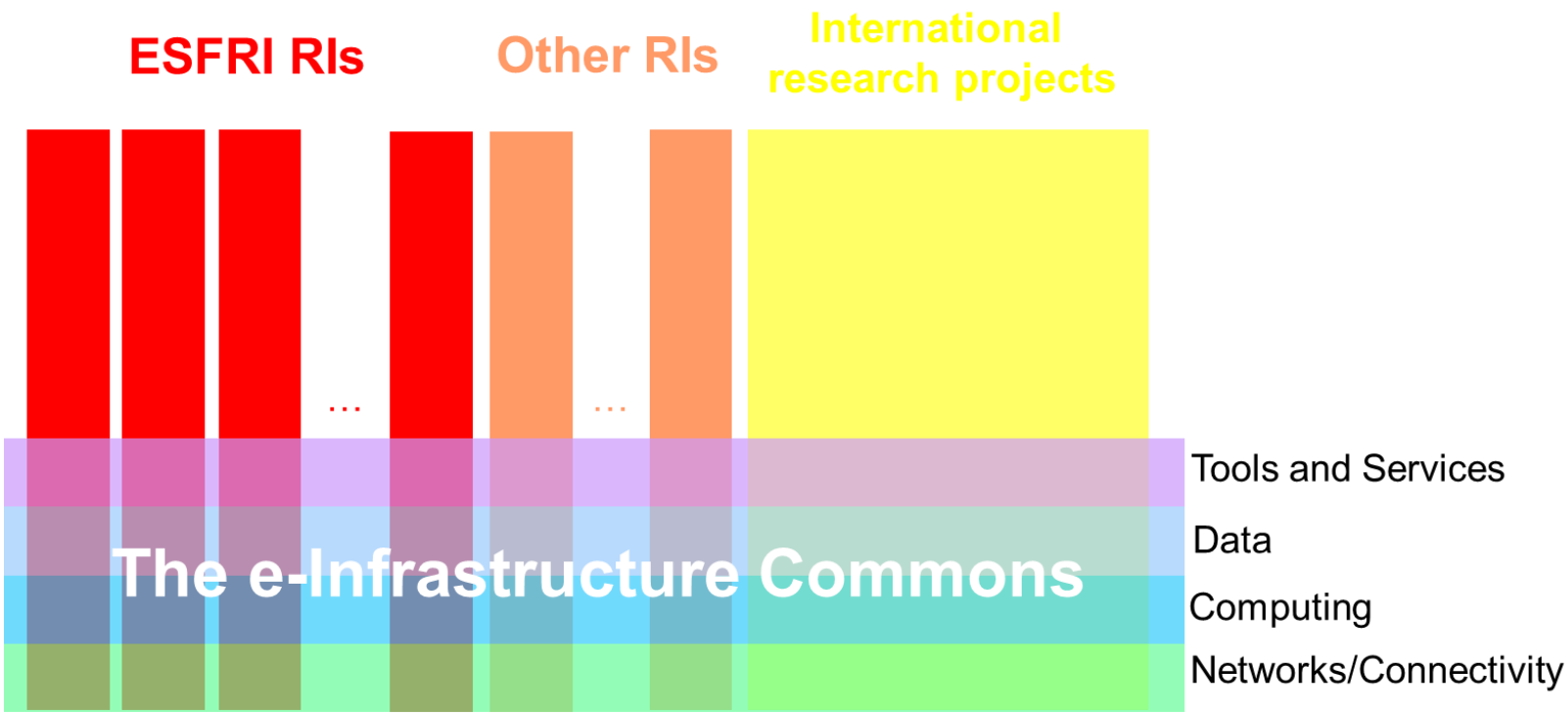
e-IRG: Open Science and e-Infrastructure

Enabling components for Open Science

- Definitions, policies, rules and standards
- **e-Infrastructure** and e-Science tools for enabling discovery, easy access, and use of the results
- Funding schemes for the costs for providing access to and storing/maintaining the results
- A refined system for giving credit to researches that provide access to their results to others

Sometimes full open access to research results can not be implemented. Note that the components above still are essential for efficient progress of science!

e-IRG: e-Infrastructure Commons

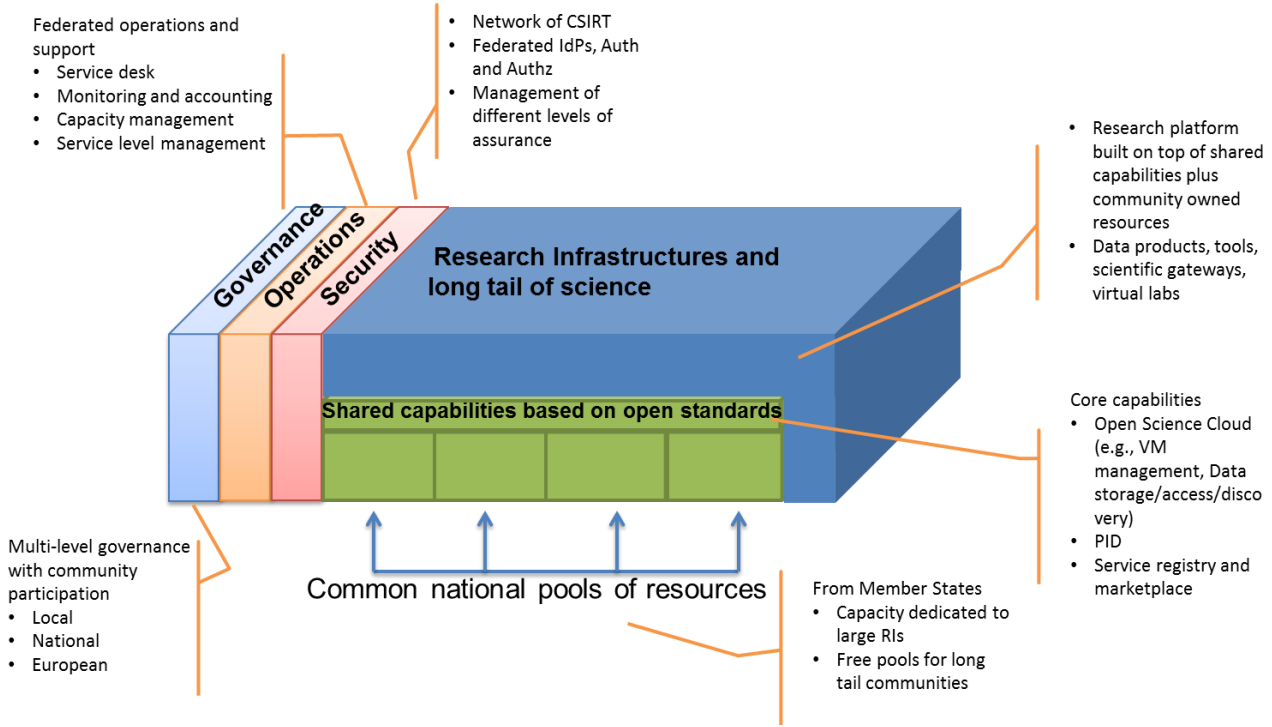


EGI.eu: Open Science and Commons

Principles of the Commons	What it means to the Open Science Commons
Shared community resources	Research data, scientific instruments, digital services, software, scientific publications, educational and training, expertise
Community-based rules and procedures in place with built-in incentives for responsible use	Access modes are well defined and non-discriminatory for all members of the ERA (e.g. see charter for open access to RIs); clear points of access and support
Governance: the community is part	Governance model with multiple stakeholders, including research communities, scientific infrastructures, resource providers, national and European infrastructures, etc.
Long-term, persistent care for a given resource for the benefit of oneself and others	Long-term support of funding agencies to allow for infrastructures to take a long-term view and build for a common European future

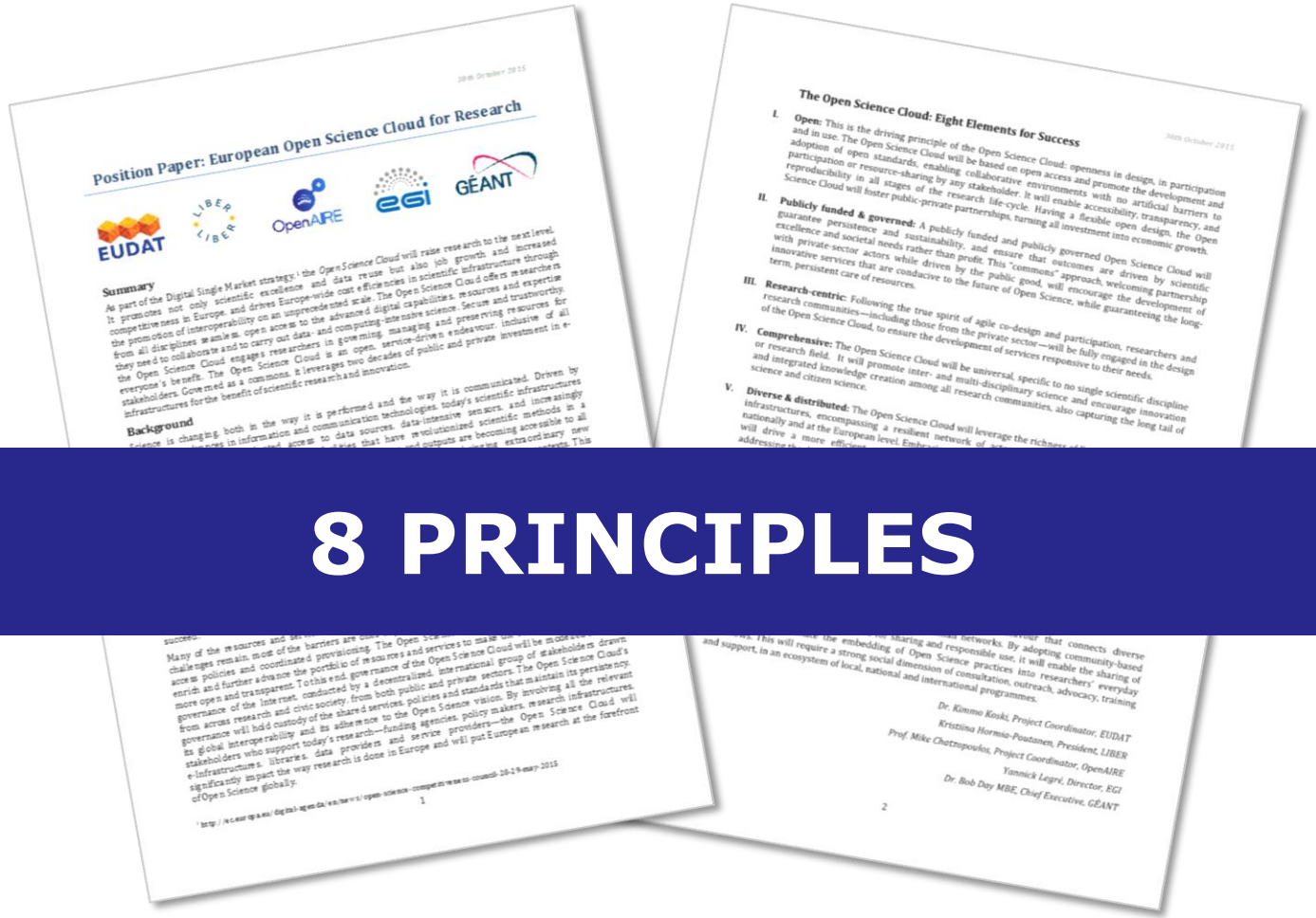
EGI.eu: Developing an Open Science Commons

Shared Digital ERA Backbone



Joint e-Infrastructure Vision

The Open Science Cloud offers researchers from all disciplines seamless, open access to the advanced digital capabilities, resources and expertise they need to collaborate and to carry out data- and computing-intensive science. Secure and trustworthy, the Open Science Cloud engages researchers in governing, managing and preserving resources for everyone's benefit.



8 PRINCIPLES

EOSC is not only e-infra

Sustainability need also to be ensured for Open Access policies and infrastructure

1. **Adopt sound governance structures** with greater representation from funders and policy makers, promoting the wider use of crucial identifiers and standards
2. **Ensure the financial sustainability of critical services**, particularly the DOAJ and SHERPA services
3. **Create an integrated infrastructure for OA repositories** based on central 'nodes', interoperability across the broader landscape, and increased engagement with the European Commission's OpenAIRE project and the work of the Confederation of Open Access Repositories (COAR)
4. **Invest strategically in OA services in order to create a coherent OA infrastructure** that is efficient, integrated and representative of all stakeholders

From Science 2.0 to Open Science Cloud

- 2014 – EU Consultation on "Science 2.0" – Science in Transition
- Q2 2015 – Launch of Digital Single Market Strategy
  - European Open Science Cloud as part of the European Cloud Initiative
- Q4 2015 High Level Expert Group formed
- 19 April 2016 – EC Communication -"Path to digitise European industry" including "European Cloud Initiative - Building a competitive data and knowledge economy in Europe"
- 22 June – INFRADEV-04 Pilot Proposal Deadline

- European
- Open
- Science
- Cloud

EOSC: **Framing**

- **Trusted access** to **services** & **systems**
- **Re-use** of shared **data**
- **Across** disciplinary, social and geographical **borders**
- **Federated** environment, across Member States

EOSC: ‘Internet approach’

- **Minimal** international guidance and governance
- **Maximum** freedom to implement.
- **Globally** interoperable and accessible
- **Globally** embedded in a ‘**Commons**’

EOSC: **Scope**

- **Human expertise**
- **Core resources**
- **Standards, Best Practices**
- **Underpinning technical infrastructures**
- **A web of Data and Services**

EOSC: **Supports**

- **Open Science**
- **Open Innovation**
- **Systematic and professional data management**
- **Long term data stewardship**

EOSC: **Challenges and Observations**

- The majority of the challenges are **social** rather than **technical**
- Not just t**he size of data,** but in particular **complex data** and **analytics across domains**.
- Shortage of **data experts** globally and in the European Union
- **Archaic system of rewards** and **funding** of science and innovation
- '**Valley of death**' between **(e-)infrastructure providers** and **domain specialists.**
- **Short funding cycles** of **core research infrastructures** are **not fit for purpose**
- **Fragmentation** between domains causes **repetitive** and **isolated** solutions
- Distributed data sets increasingly **do not move** (**size & privacy** reasons)
- Centralised HPC is **insufficient** to support **distributed meta-analysis and learning**.
- However, the **major components** for a **first generation EOSC** are largely 'there'
- But '**lost in fragmentation**' and spread over 28 Member States.

EOSC: **Key requirements**

- **New modes of scholarly communication**
- **Modern reward and recognition practices need to support data sharing and re-use**
- **Innovative, fit for purpose funding schemes for sustainable underpinning infrastructures**
- Core **data experts** need to be trained and their career perspective significantly improved
- Cross-disciplinary **collaboration-specific measures** for review, funding and infrastructure
- Support for the transition from **scientific insights** towards **societal innovation**
- The EOSC needs to be developed as an **eco-system of infrastructures**
- Key Performance Indicators should be developed for the **EOSC**
- The EOSC should **enable automation of data processing: machine actionability** is key.
- **FAIR principles** [1, 6] (http://www.nature.com/articles/sdata201618)

EOSC: **Policy Recommendations**

- P1: Take immediate, affirmative action in close concert with Member States
- P2: Close discussions about the ‘perceived need’
- P3: Build on existing capacity and expertise where possible
- P4: Frame the EOSC as supporting Internet based protocols & applications

EOSC: **Governance Recommendations**

- G1: Aim at the lightest possible, internationally effective governance
- G2: Guidance only where guidance is due
- G3: Define Rules of Engagement for formal participation in the EOSC
- G4: Federate the Gems across Member States

EOSC: **Implementation Recommendations**

- I1: Turn this report into an EC approved document to guide EOSC initiative
- I2: Develop, Endorse and implement a Rules of Engagement scheme
- I3: Fund a concentrated effort to locate and develop Data Expertise in Europe
- I4: Install a highly innovative guided funding scheme for the preparatory phase
- I5: Make adequate data stewardship mandatory for all research proposals
- I6: Install an executive team to deal with international coherence of the EOSC
- I7: Install an executive team to deal with the preparatory phase of the EOSC

EOSC: 7 Immediate actions based on feed back

- II1: **Publish** the report (final draft available)
- II2: Develop, Pilot and implement a **Rules of Engagement** scheme
- II3: Detail the **transition and sustainability model** (and pilot it)
- II4: Train the **data experts** to bridge between ‘e-INFRA’ and ‘ESFRI’
- II5: **Assist data stewardship planning** and exec. tools for all researchers
- II6 Develop the plan for what ‘**minimal essential governance**’ means in practice
- II7 **Federate** Interoperability standards and best practices (key RDA-role)

European Cloud Initiative – Communication 19th April 2016

Building a competitive data and knowledge economy in Europe

- to develop a trusted, open environment for the scientific community for storing, sharing and re-using scientific data and results- the **European Open Science Cloud**
- to deploy the underpinning super-computing capacity, the fast connectivity and the high-capacity cloud solutions they need via a **European Data Infrastructure**
- Focussing initially on the scientific community, the user base will be expanded to the public sector and to industry, creating solutions and technologies that will benefit all areas of the economy and society

Key Activities

- **Make all scientific data produced by the Horizon 2020 Programme open by default**.
- **Raise awareness and change incentive structures**
- Develop specifications for **interoperability and data sharing** across disciplines and infrastructures
- Create **a fit-for-purpose pan-European governance structure** to federate scientific data infrastructures and overcome fragmentation
- Develop **cloud-based services for Open Science.** Supported by the European Data Infrastructure, they will allow researchers to find and access shared research data, to employ advanced analytical software, to use high-performance computing resources and to learn about best data-driven science practices from leading disciplines
- **Enlarge the scientific user base** of the European Open Science Cloud to researchers and innovators from all disciplines and Member States, as well as from partner countries and global initiatives

European Data Infrastructure

Exascale supercomputers infrastructure by 2022, based on EU technology, which would rank in the first 3 places of the world

- foster an HPC ecosystem capable of developing new European technology such as **low power HPC chips**
- integrate technologies into system prototypes, co-designing50 solutions and procuring HPC systems; the resulting HPC infrastructure will focus on **supercomputers of top-range capabilities** connected to mid-range EU national computing centres and to pan-European data and software infrastructure to offer supercomputing as a service
- provide **seamless, high-speed, reliable and secure connectivity** to make HPC accessible across the EU; the trans-European high speed network (GÉANT) and National Research and Education Networks (NREN) already connect 50 million researchers and students; these infrastructures will be upgraded to match the increase of data volumes to be transferred and the extension of the user base.
- complemented by an ambitious, long-term and large-scale flagship initiative to unlock the full potential of **quantum technologies**, accelerate their development and bring commercial products to public and private users

Widening Access and Trust

- **European Data Infrastructure will be widened to the public sector**, for example through large-scale pilots involving eGovernment62 and public sector stakeholders and by progressively opening the European Data Infrastructure to **users from industry** and the public sector to achieve a European dimension
  - create a platform for public authorities to open their data and services, creating a "Government as a Service" (GaaS) base for the EU.
- meet **high standards of quality, reliability and confidentiality**, to ensure protection of personal data and intellectual property, **and security**
  - The Commission working with industry and Member States will promote the use of existing relevant certifications and standards, and – where appropriate – the creation of European-level certification and labelling, in particular to support public procurement of cloud services.
- adapt HPC and Big Data solutions to a cloud environment in order to enable broad access, notably for SMEs;
  - develop an ecosystem to strengthen the cloud industry in Europe, using the European Open Science Cloud as a testbed for innovative cloud technology solutions;

Sustainability and funding

- In cooperation with Member States and stakeholders, the Commission will explore appropriate governance and financing mechanisms for the Open Science Cloud and the European Data Infrastructure and define an implementation Roadmap. As of 2016
- The Commission will put forward approaches for blending different funding streams, for discussion with Member States and stakeholders, in order to realise the objectives of this Communication.

Acknowledgements

- Includes content from:
  - Rachel Bruce – Jisc
  - Per Oster – CSC
  - Tiziana Ferrari, Sergei Andreozzi – EGI.eu
  - Barend Mons – EC EOSC High Level Expert Group

Find out more…

Except where otherwise noted, this work is licensed under CC-BY-NC-ND