Making Open Science the new normal

15th International Digital Curation Conference

19 February 2020

Kostas Glinos
Open Science
European Commission

Why do we need Open Science?

- Open Science means sharing
  - knowledge and tools
  - as early as possible
  - between researchers and between disciplines, and also with society at large
- Open Science has the potential to increase:
  - Quality and efficiency of R&I, if all the produced results are shared, made reusable, and if their reproducibility is improved;
  - Creativity, through collective intelligence and cross-disciplinary research that does not require laborious data wrangling;
  - Trust in the science system, engaging both researchers and citizens.

Health R&I: a reproducibility crisis?

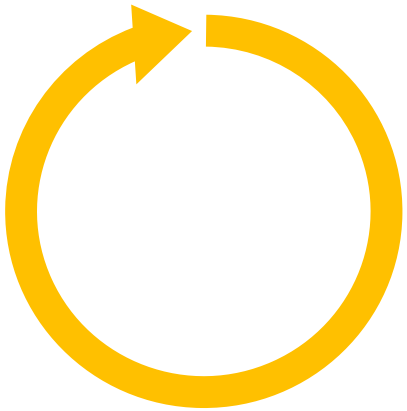
- Close to €300bn/year for Health R&I (worldwide)
- As much as 85% of the research investment may be wasted
  (Chalmers & Glasziou, 2009; Macleod, 2014)

Unusable research reports

- Methods and codes unavailable
- Inadequate info on medical interventions in trials, etc..

Scientific question not pertinent

- Lack of awareness of existing evidence
- Not relevant to clinicians, carers and patients, etc.



Biased reporting of results

- Selective reporting
- Reported data not made comparable with other studies
- Conflicts of interest
- Fraud, etc.

Poor study design, conduct & analysis

- Not replicated enough
- Not enough collaborative efforts
- Poor training/mentoring of researchers, etc.

Results not fully accessible

- 'Disappointing' or negative results less likely to be shared
- Trials not registered, etc..

Commissioner Gabriel

Hearing at the European Parliament, Sept. 30th, 2019

- "The **Open Science** issue is […] an issue that is dear to my heart"
- "Today, more than ever, we need researchers to **share the results** of their projects […] and to **capitalize on the research of others**"
- "I will insist on having **data** that are […] **reusable, accessible, of quality**"

European Commission

Main challenges and priorities

Improve the practice of research and innovation

- Openly accessible scholarly publications
- Early sharing of all research outputs
- All data FAIR, RDM
- Reproducible results
- Societal engagement and responsibility
- Towards a new ethical conduct for researchers?

Develop proper enablers

- *Rewards and incentives to adopt Open Science practices, with appropriate metrics*
- *Infrastructures (including EOSC)*
- *Appropriate skills and education, including for research integrity*

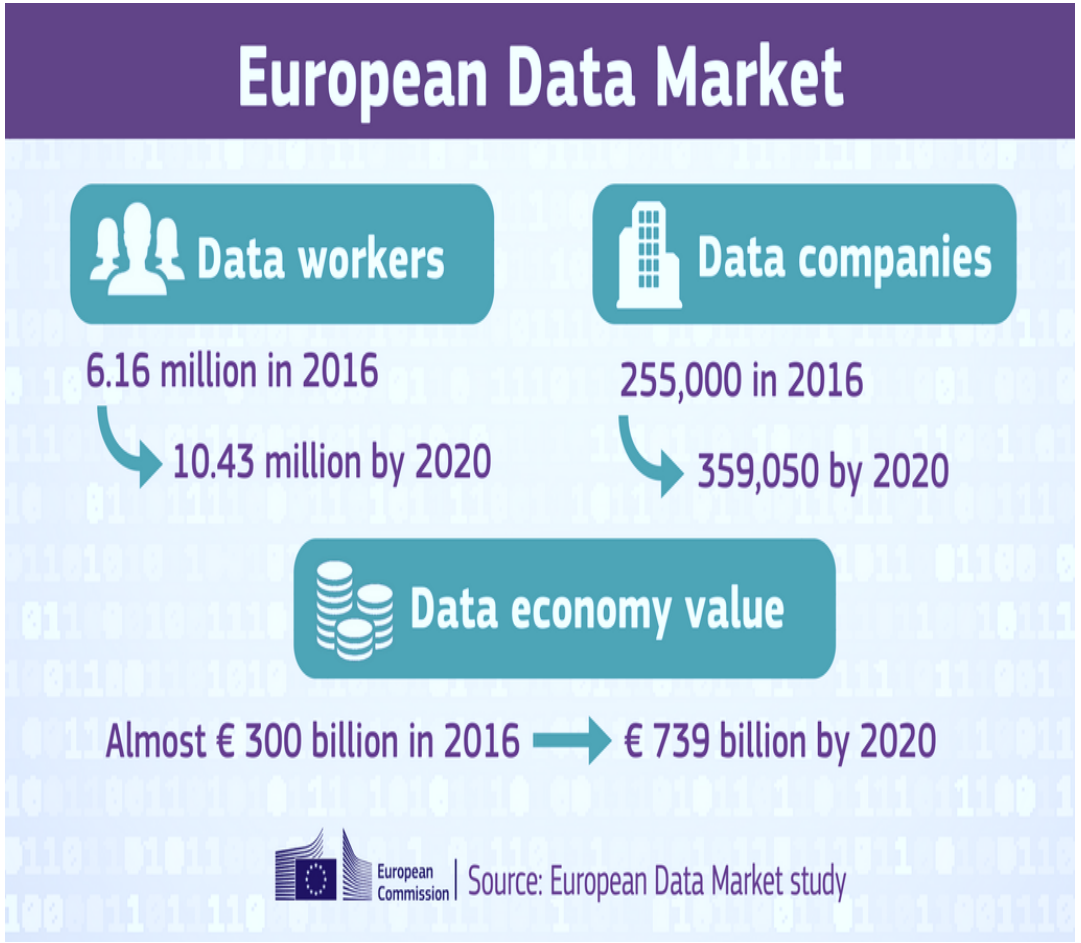
Practices

Open Access

- *OA to publications has been the cornerstone of EU Open Science policy since 2007*
- *Encouraged open access policies in EU member states through the Recommendation on access to and preservation of scientific information (2012, revised in 2018)*
  - *Implementation monitored through the national points of reference. New report to be released March 2020*
- *Implementation in EU research programmes*
  - *From pilot on OA to publications in FP7 to a pilot on open data in H2020 to mandatory OA to publications and “as open as possible as closed as necessary” access to data today*
  - *Above 90% OA to publications uptake in Horizon 2020*
- *The European Commission supports Plan S*

Data

- Data is the **main asset** of the digital economy
- Data production continues to **grow exponentially**→ from 33 zettabytes in 2018 to 175 zettabytes in 2025



- Poor data management incurs an **opportunity cost** in the billions of euros
- **Research/Science** is one of the main data producing/consuming domains

European Commission

Responsible Research Data Management

- The Commission embraces the **FAIR principles** (Findability, Accessibility, Interoperability, and reusability)
  - **FAIR research data** to be complemented by additional requirements on **identifiers**, **trusted repositories**, **machine-readable licences**, etc.
- The **opportunity cost** of having research data that is **not FAIR** has been estimated at **€10.2bn** (Source: Cost-benefit analysis of FAIR research data, 2017)
- The Commission, as a funder, plans on introducing **RDM requirements** under **Horizon Europe**
- A usual **misconception** is that **FAIR = open**

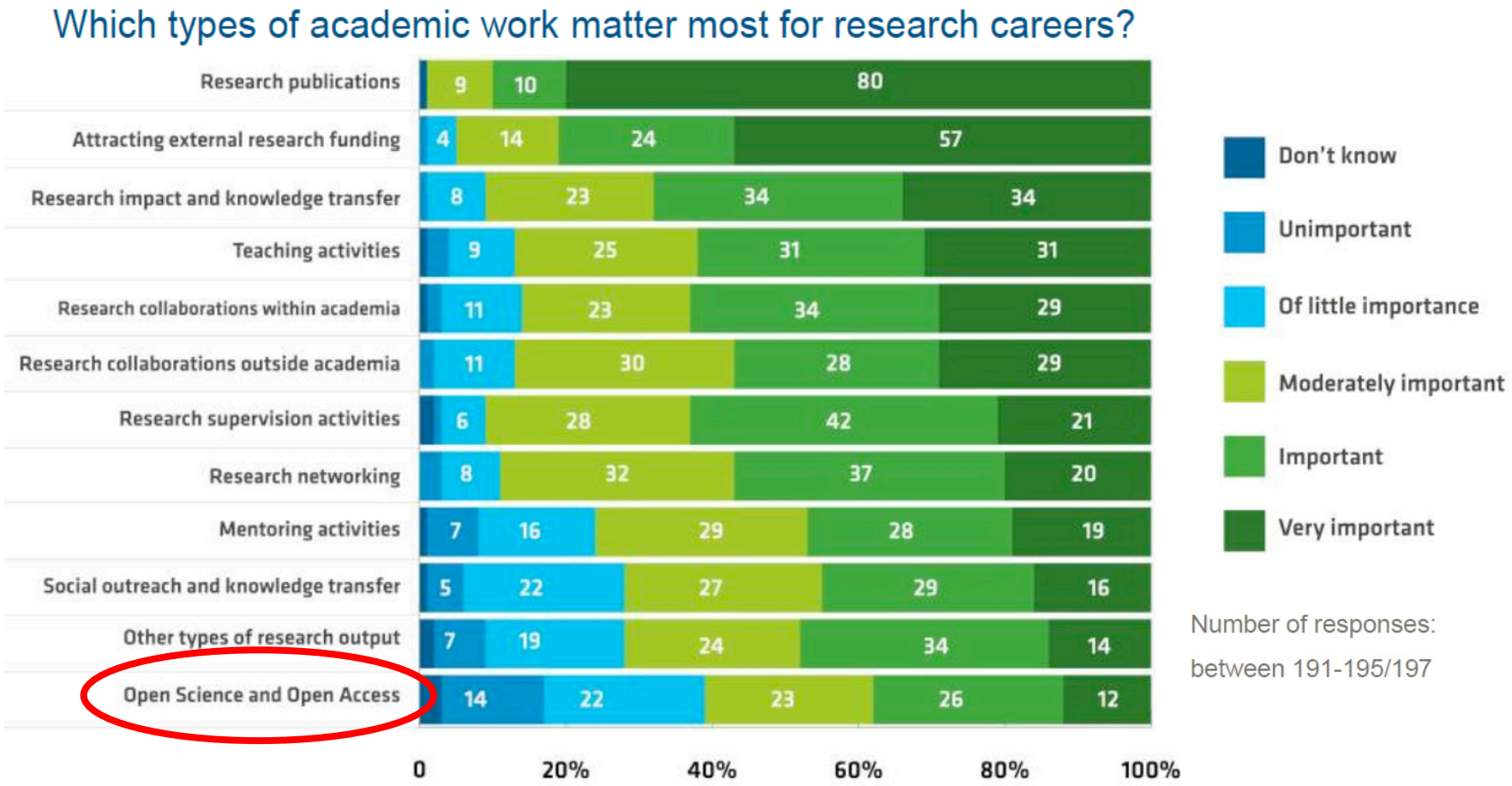
Science & society

- Improving **trust** between science and society
  - Engaging citizens and civil society in co-designing & co-creating responsible R&I agendas and contents
  - Facilitating participation by citizens and civil society organisations in research activities
  - Improving science communication, also in local settings
  - Ensuring the integrity of research and cultivating a new ethos in science

- *Horizon 2020 work programme 2018-2020 included:*
  - “Exploring and supporting citizen science”
  - “Accelerating and catalysing processes of institutional change”, including research integrity
  - “Building the knowledge base for SwafS”, including science communication

Enablers

OS practice and research careers



Source: EUA, 2019 Open Science survey

European Commission

Rewards and incentives

[From various expert group reports and position papers:]

- Extra credit to individuals, groups and projects, that practice Open Science, based on appropriate indicators
  - All evaluation steps at all stages of careers are concerned: from PhD thesis examination, recruitment, promotion, project proposal assessment, funding allocation systems, prizes and recognitions, etc.
  - In the evaluation processes, more emphasis needs to be put on the quality of the outputs, including on reproducibility of results.
- Journal impact factors should not be used as a proxy for quality. Article-level and societal impact metrics should be developed and used.
- There are already good practices that need to be collected and further exploited.
- Incentives for life-long training on Open Science practices and research integrity are required.

Skills and education

[From various expert group reports and position papers:]

- Open Science principles and practice, together with data literacy, ethics and research integrity, should be an integral part of researcher's education and career development
- Training on Open Science needs to be tailored for different disciplines and levels of researchers, but also for students and research managers and administrators
- Life-long training needs to be incentivised, sustainably resourced and its impact be monitored
- There are already good practices that need to be collected and further exploited.

- Integrate Open Science principles and baseline skills in
  - the Innovative Doctoral Training Principles?
  - the European Charter for researchers and Code of conduct for recruitment?
  - the Human Resources Strategy for Researchers?

European Open Science Cloud

EUROPEAN OPEN SCIENCE CLOUD

- The European Open Science Cloud (EOSC)—European Cloud Initiative Communication, 2016—is an **enabler of Open Science**
- ESOC will be a **trusted** and **open distributed** system for the scientific community, providing **seamless access to data** and **interoperable services** addressing the **whole research data lifecycle**
- EOSC design principles: **co-creation**, **research-led**, **community-driven**, **scalable**, **incremental and iterative**, **continuous engagement**, **consultation** & **user testing**.

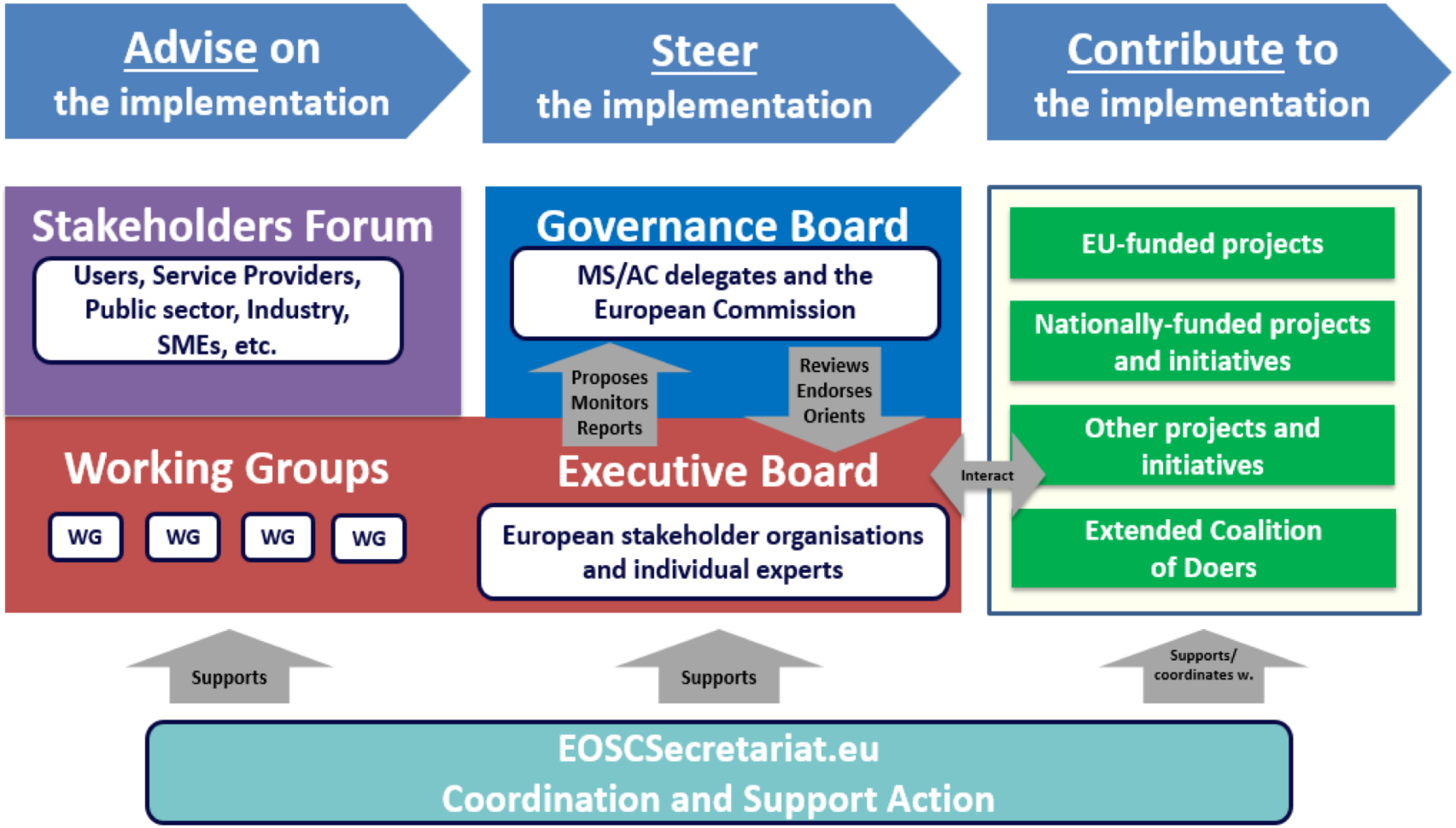
European Commission

President U. von der Leyen

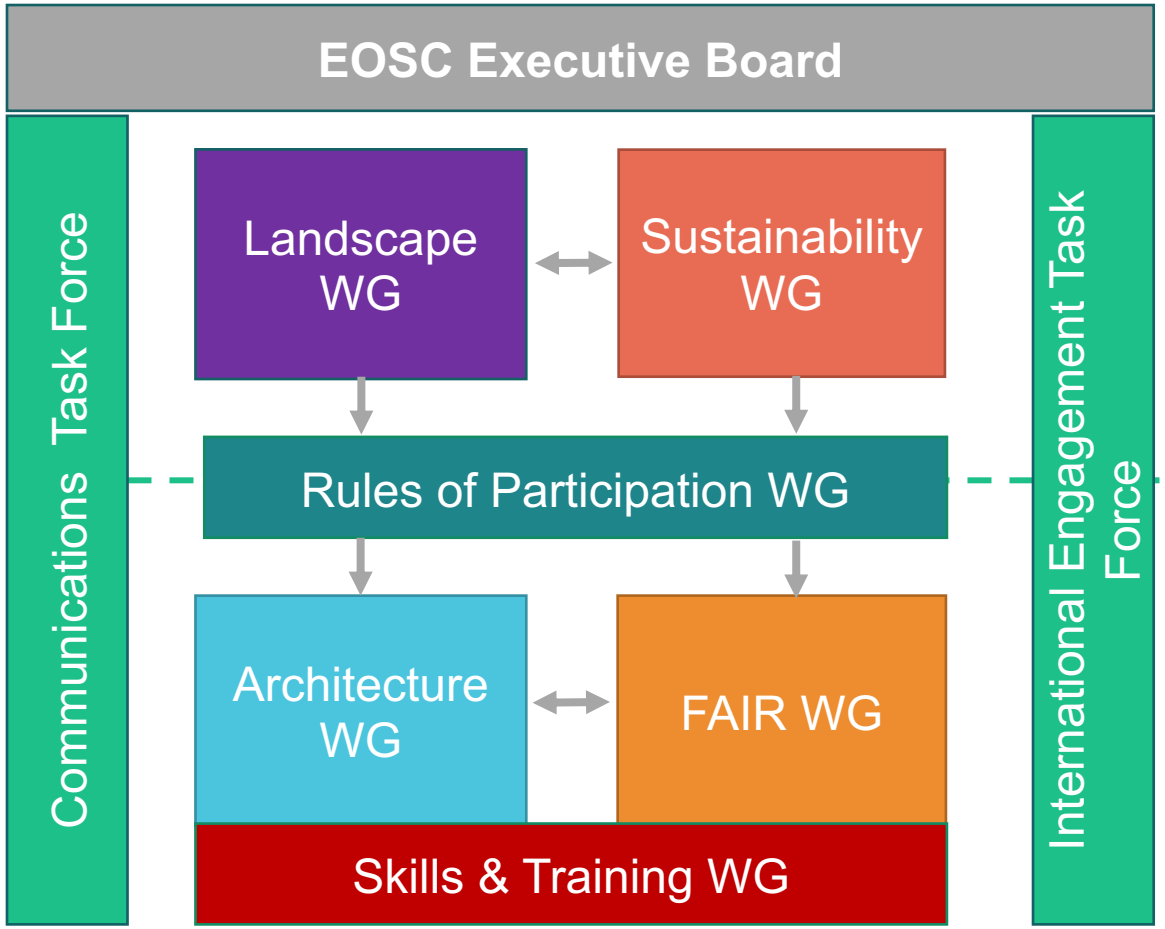
World Economic Forum, Davos, Jan. 22nd, 2020

- "We are creating a **European Open Science Cloud** now"
- "It is a **trusted space** for researchers to **store** their data and to **access** data from researchers from **all other disciplines**"
- "We will create a **pool of interlinked information**, **a web of research data**"

EOSC governance structure



EOSC EB Working Groups and Task Forces



Strategic focus

Close liaison with Governance Board

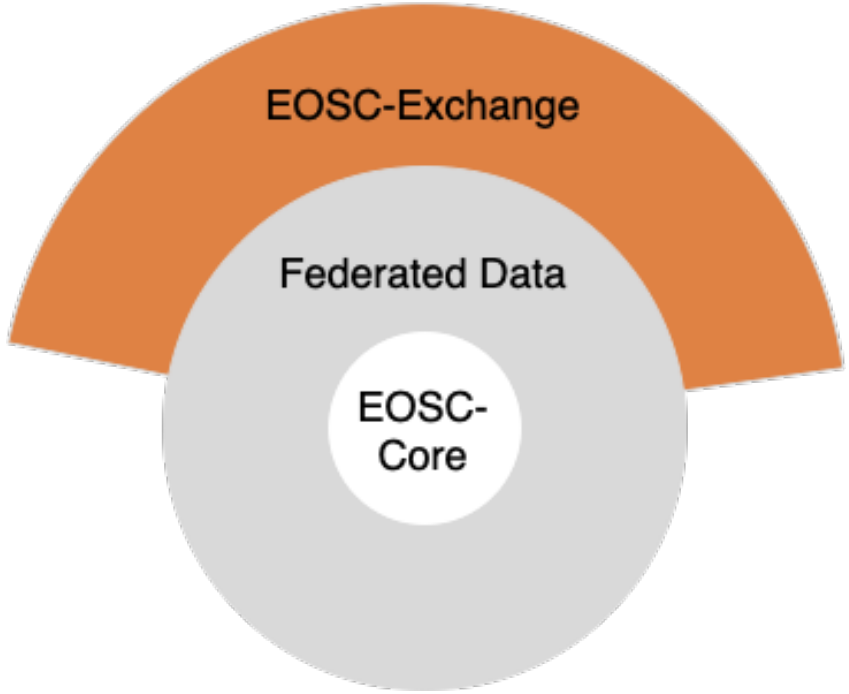
Practical implementation focused WGs

A first edition of EOSC by end 2020

- **Rules of Participation WG**:
  - Agreed and tested Rules of Participation
- **Landscape WG**:
  - Analysis of existing national infrastructures and policies
- **Sustainability WG**:
  - Financing model, legal entity, and post 2020 governance structure
- **Architecture WG**:
  - Functioning federated core
  - Initial/minimum set of EOSC data and services
  - EOSC Interoperability Framework (with FAIR WG)
- **FAIR WG**:
  - Persistent Identifier policy (with Architecture WG)
  - Metrics for FAIR data and certified services
- **Training & Skills WG**:
  - EOSC MVP for training & sustainability model (with Architecture WG and Architecture WG)
  - Recommendation for EOSC skills/training in national digital skills policies

First edition: minimum viable EOSC (MVE)

- The MVE includes **EOSC-Core** (layer of discovery and interoperability) and **EOSC-Exchange** (layer of enriched services), both of which work with **federated FAIR datasets**
- MVE must enable the **federation** of existing and planned **research data infrastructures**
- Step 1: Federate the **thematic clusters** and **regional projects**
- Advance incrementally: Begin with simple use cases (open data, rather than sensitive or closed).



European Commission

EOSC-Core: functions and proposed coverage

- Provides the means to **discover**, **share**, **access** and **re-use** data and services (layer of discovery and interoperability)
- In its first iteration, **not planned to store, transport or process data**
- To be **as widely used as possible**: accessible to any authenticated user to promote uptake of EOSC

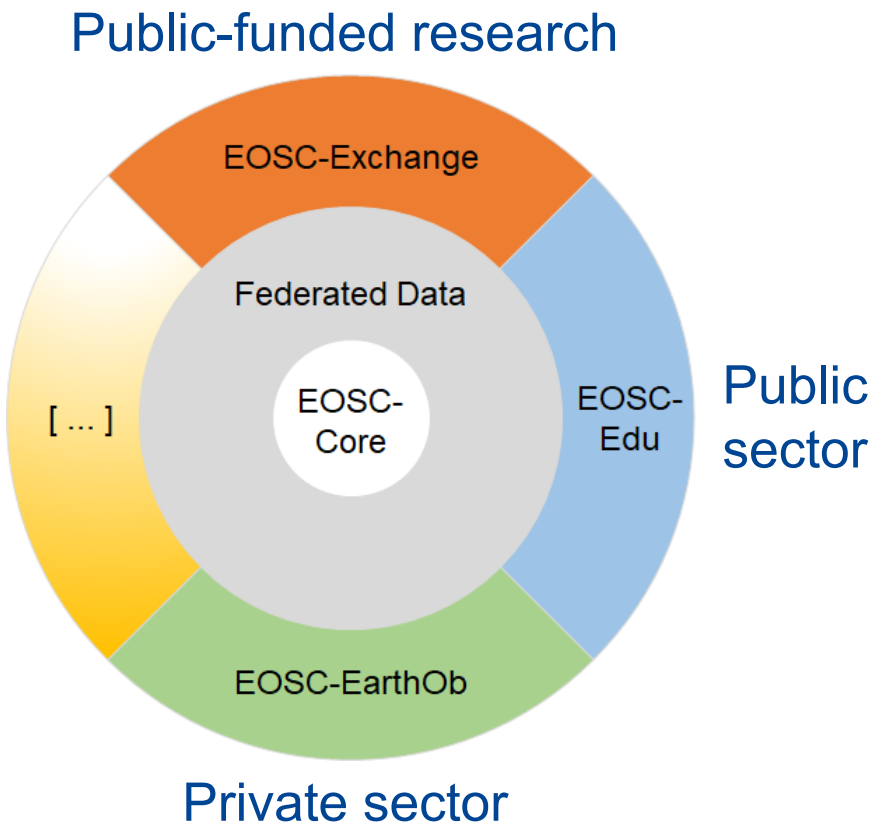
Current discussion on **minimum requirements** vs. **nice to have** (Minimum Viable EOSC vs. Maximum Valuable EOSC)

Some services under consideration:

- AAI framework
- PID services
- Shared open science policy framework
- Data access framework
- Service management & access framework
- Minimum legal metadata framework
- Help-desk

Proposed second and third iterations

- **All research communities** to be tackled by the first iteration under EOSC-Exchange layer
- **Incremental extensions** to target **public administration** and **industry**
- Vision: **One 'marketplace'**.
- Differing requirements/legislation may require linked but alternately governed spaces



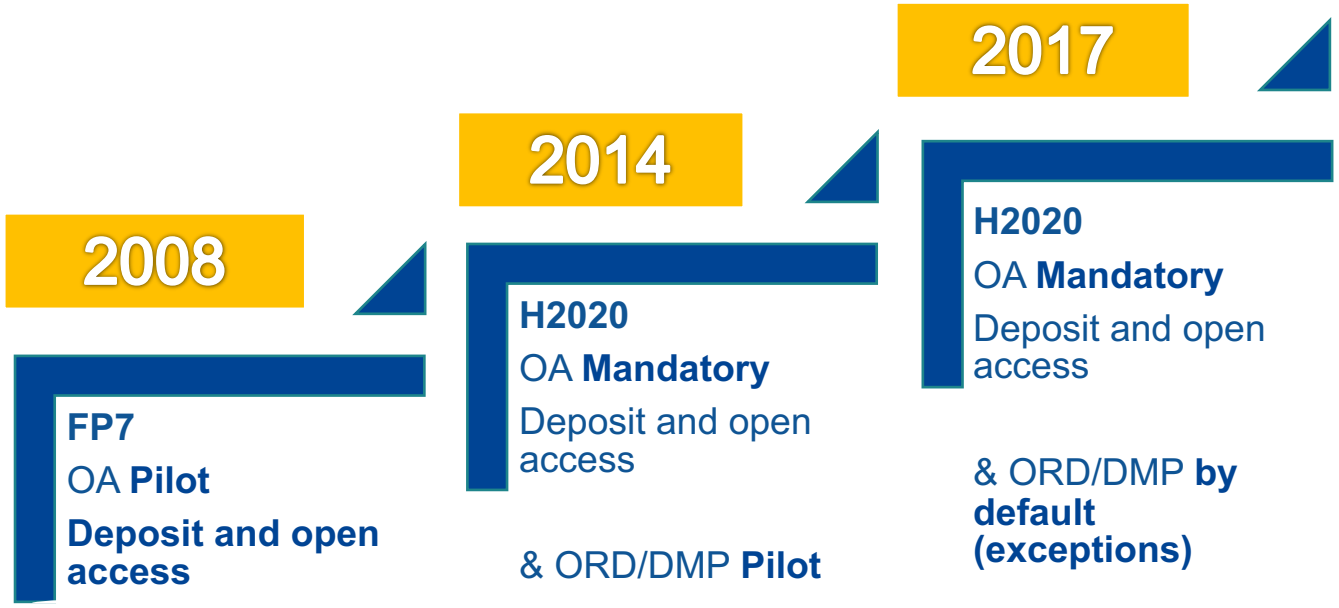
Ongoing partnership proposal and SRIA

- Current governance of EOSC will present a proposal to the European Commission for a **co-programmed partnership** under Horizon Europe
- The proposed partnership will make the initiative evolve from a project-based approach to an **all-encompassing ecosystem** where the different stakeholders take the necessary commitments to accomplish it.
- Counter-signatory of the Commission will be a **legal entity** to be set up before the end of the year **representing key stakeholders in the EOSC process**
- Discussions are taking place on how the European Commission, Member States and Associated Countries, and this legal entity of organisations **will steer EOSC post-2020**
- A **Strategic Research and Innovation Agenda** will set the **long-term strategic priorities** and will serve as **a handbook for implementing the EOSC vision**

How do we move forward together?

- Fully implementing Open Science practice requires **engagement of all stakeholders**
  - Researchers, research performing organisations, funders, policy makers, citizens, etc.
  - At all levels (regional, national, European and international).
- **One** EU policy, **three** sets of complementary actions:
  - Develop the necessary **infrastructures and services**
  - Induce and support **coordinated paradigm shifts**
  - **Walk the talk** in European R&I programmes and **lead by example**

Evolution of our policies across the FPs



Likely under **Horizon Europe (2021)**:

- Open Science embedded throughout the FP
- **OA mandatory**: deposit and open access. Probably more stringent conditions
- **Open data** by default
- **RDM** considerations to be **evaluated** at proposal stage
- **Mandatory DMP**, in line with FAIR, for **all** projects that generate/collect/reuse research data

We are pioneers with respect to other funders!

RDM and DMPs in Horizon Europe

- **DMP for all projects** that generate/collect research data:
  - No need to reinvent the wheel—we will seek alignment and build on already existing initatives (e.g. Science Europe core requirements, DMPOnline/ARGOS DMP tools, etc.)
  - Data is **not the only digital output** to be managed (software, code, workflows, protocols, models, etc.). The Wellcome Trust devotes a section of their DMP to 'Other research output'
  - **Machine-actionability** is key to allow computers to integrate external services to interlink basic building-blocks: author (ORCID), data (DataCite), organisations (RoR, FundRef), etc.

- **Initial RDM considerations** to be provided at **proposal level**

Open Science in EU legislation

- Revision of the **Recommendation on access to and preservation of scientific information** (2018)
  - Also covers infrastructure, metrics, rewards, skills, etc.
- Revision of the **Open data and the re-use of public sector information directive** (2019)
  - Applies to publicly funded research data that are publicly available through repositories
- Revision of the **EU Copyright Directive** (2019)
  - Provides for an exception for research organisations to carry out Text and Data Mining (data analytics)

- Is there something more needed?

Potential actions for system changes (1)

- "**Open Science coordinators**" at national level, and "**Open Science officers**" at institutional level, would facilitate Open Science policy coherence and implementation, and learning from good practice
- Monitor the implementation of the Commission **recommendations on access to and preservation of scientific information**, with the National Points of Reference, and incentivize good practice
- Support the further development of **infrastructures** for Open Science, including EOSC
- Open Science in a **revitalised European Research Area** (ERA), in synergy with the European Higher Education Area (EHEA)
- Recognition of Open Science in **regional R&I Smart Specialisation Strategies** (RIS3s)

Potential actions for system changes (2)

- *Support institutions including Universities and their alliances on the modernisation of research and researchers' assessment, including by identifying, assessing and piloting the validity of metrics and quantitative/qualitative indicators*
- *Support policies and conditions (e.g. infrastructures) that enable open access as the default way of disseminating research (publications, data and other contributions)*
- *Develop a strong narrative for the role of Science in Society and promote specific actions, e.g. in science communication, science skills, rewarding engagement with citizens, research assessment pilots, promoting reproducibility, etc.*
- *Mainstream OS practices in thematic areas, e.g. FAIR data projects in specific domains*
- *…*

Keep in touch

ec.europa.eu/

europa.eu/

@EU_Commission

@EuropeanCommission

European Commission

europeancommission

@EuropeanCommission

EUTube

EU Spotify

European Commission

Thank you

European Commission

Relevant updates from the WGs

- **Persistent Identifier Policy** (FAIR WG and Architecture WG):
  - Out for **public consultation** since **mid December 2019**
  - **Popularly received document**: Over 900 downloads and comments from Europe, Australia, USA and China
  - **New version due in March** and **implementation guidelines in Summer 2020**
  - **PID policy** openly available:

    https://doi.org/10.5281/zenodo.3574203
  - Join the international discussion and provide your **feedback**:

    https://www.pidforum.org/t/please-give-us-your-feedback/775

Relevant updates from the WGs

- **Draft Rules of Participation** out for consultation **this week** for a period of 4 to 6 weeks
- The **primary targets** of this consultation are:
  - INFRAEOSC projects
  - Executive Board Working Groups
  - Member States and Associated Countries
- National input to be compiled by WG members nominated by the governance board
- A **wider consultation** will be launched **later this year** with the objective of having a completed version before the EOSC symposium (end of October)
- A **final version** consolidating any feedback from the symposium will be presented in **December**

Relevant updates from the WGs

- Interim **metrics and certification** recommendations:
  - Metrics based on output from RDA FAIR Maturity Working Group & deliverables from the Horizon2020 project FAIRsFAIR
  - Based on **community feedback**, tbd what can be required in EOSC
  - **Repository certification** based on CoreTrustSeal, with phased entry
  - Workshop planned end of April to define other services/elements to be certified
  - Reports: FAIR metrics for EOSC and EOSC service certification
  - We welcome your feedback at the EOSC Liaison Platform or by email at inform-FAIR-wg@eoscsecretariat.eu
  - **Full recommendations** will be out for consultation early Summer 2020 & finalised community-approved version out before the EOSC Symposium in Autumn

RDA FAIR Data Maturity Model WG

https://www.rd-alliance.org/groups/fair-data-maturity-model-wg

- Context: **Ambiguity/wide range of interpretations** surrounding FAIR (for data)
- **18-month** Working Group launched in January 2019, co-chaired by Edit Herczog, Keith Russell and Shelley Stall
- **Supported** by the European Commission (**DG R&I**) through a tender with PricewaterhouseCoopers (**PwC**). PwC acts as a secretariat for the WG and performs editorial work
- Output:
  - **RDA recommendation**
  - A common set of **core assessment criteria** for FAIRness (**indicators**) and **guidelines/checklists** to help create assessment methodologies. Currently, indicators and guidelines **under testing** by **user communities**, whose feedback will be integrated into the WG's output.

Indicators: https://docs.google.com/spreadsheets/d/1mkjElFrTBPBH0QViODexNur0xNGhJqau0zkL4w8RRAw/edit#gid=1210743571

Guidelines: https://docs.google.com/document/d/1pDGGL3-BbBJu18KlfZUI3AizKLHXGXdIi_mPtpEWmeg/edit