EOSC (Enabler of Opening Science Commons) Expectations Regarding Data from Science

Karel Luyben

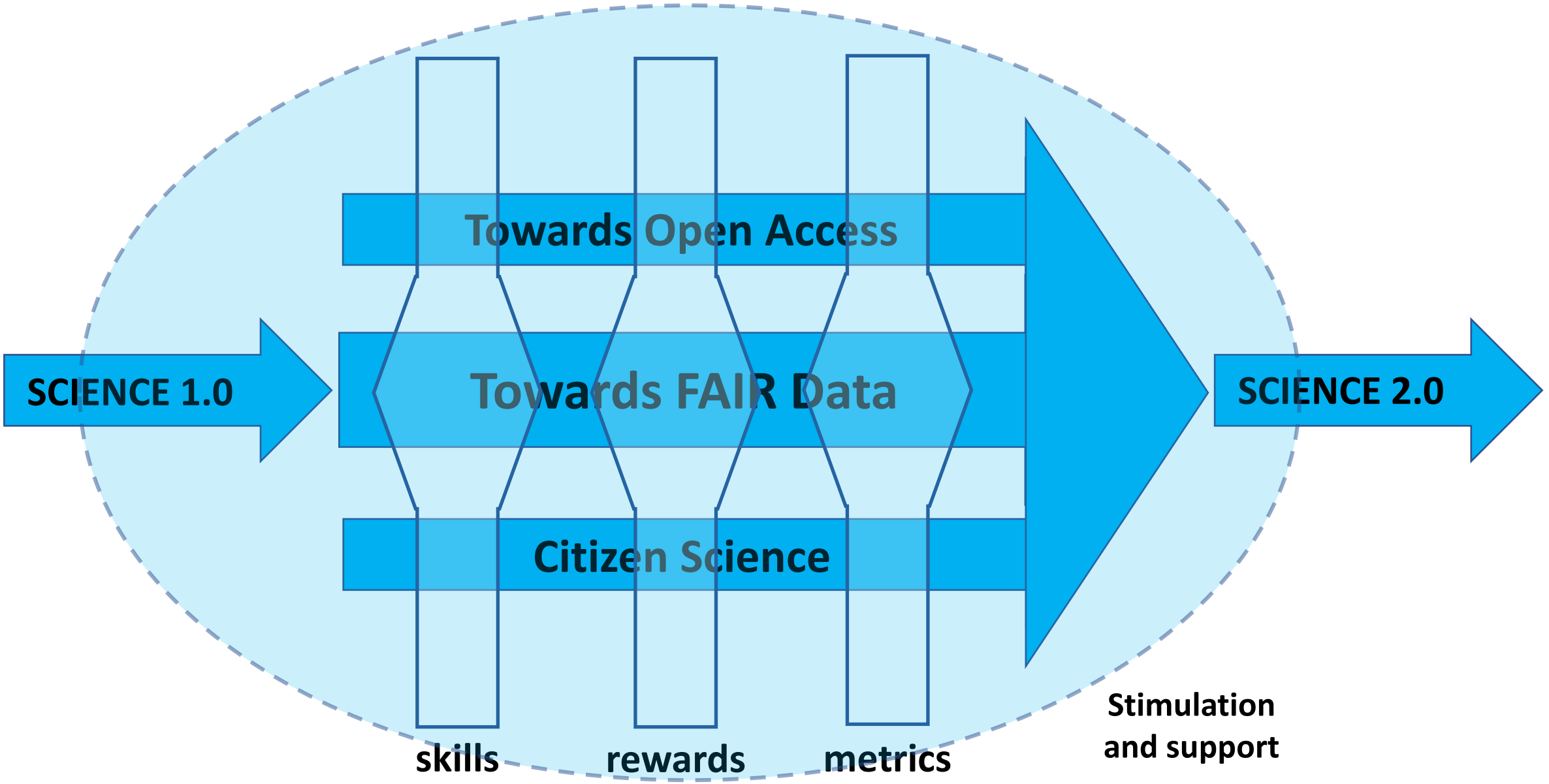
Chair of Executive Board of EOSC

April 14, 2020

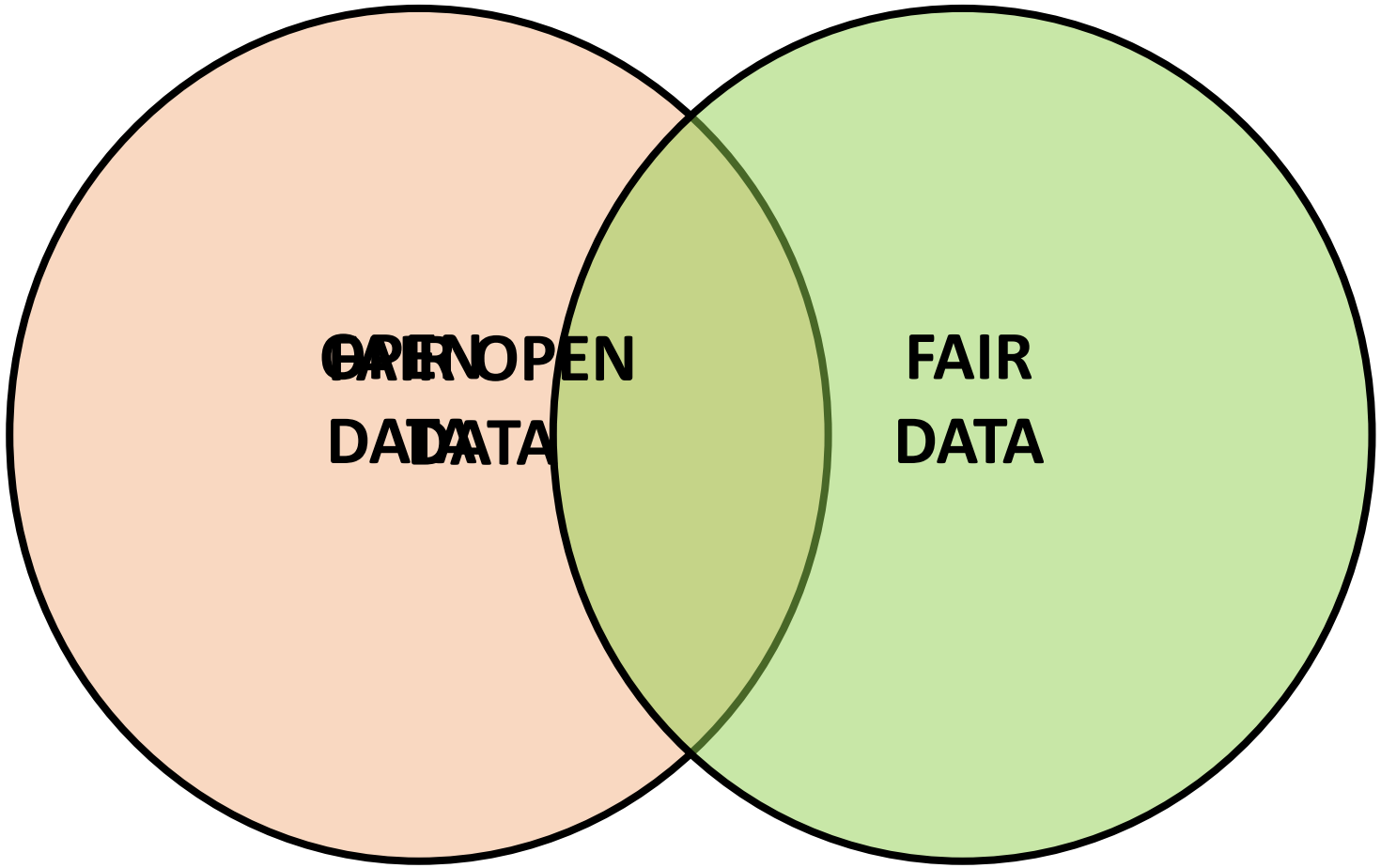
GEDE workshop, online

EUROPEAN OPEN SCIENCE CLOUD

OPEN SCIENCE

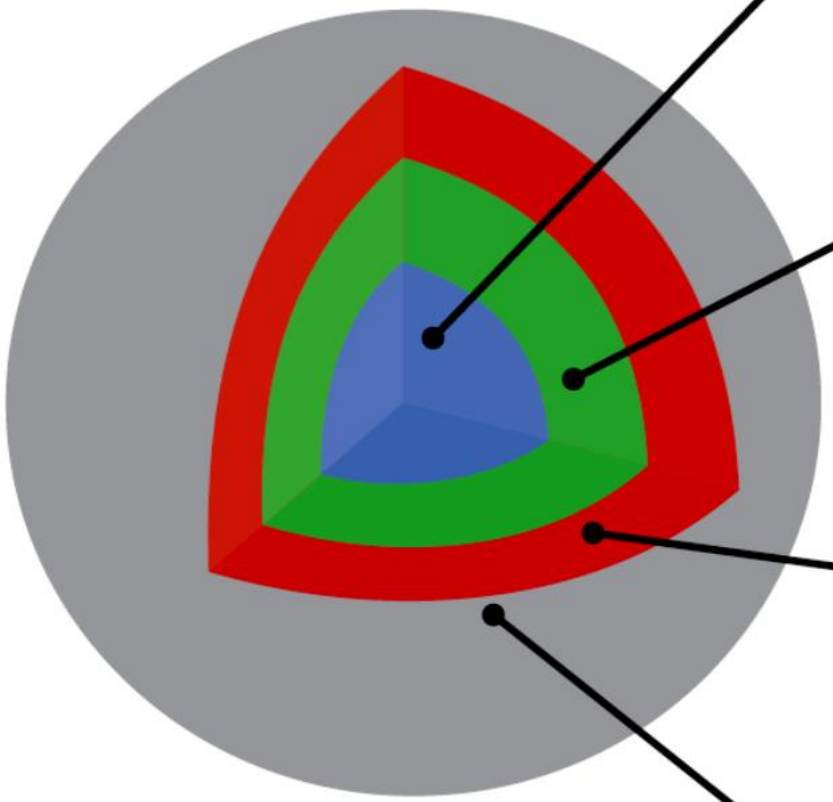


OPEN DATA and/or FAIR DATA



FAIR ≡
Findable
Accessible
Interoperable
Reusable

Towards “as FAIR as possible” and “as open as possible”



DIGITAL OBJECT

Data, code and other research outputs

At its most basic level, data or code is a bitstream or binary sequence. For this to have meaning and to be FAIR, it needs to be represented in standard formats and be accompanied by Persistent Identifiers (PIDs), metadata and documentation. These layers of meaning enrich the object and enable reuse.

IDENTIFIERS

Persistent and unique (PIDs)

Digital Objects should be assigned a unique and persistent identifier such as a DOI or URN. This enables stable links to the object and supports citation and reuse to be tracked. Identifiers should also be applied to other related concepts such as the data authors (ORCIDs), projects (RAIDs), funders and associated research resources (RRIDs).

STANDARDS & CODE

Open, documented formats

Digital Objects should be represented in common and ideally open file formats. This enables others to reuse them as the format is in widespread use and software is available to read the files. Open and well-documented formats are easier to preserve. Data also need to be accompanied by the code use to process and analyse the data.

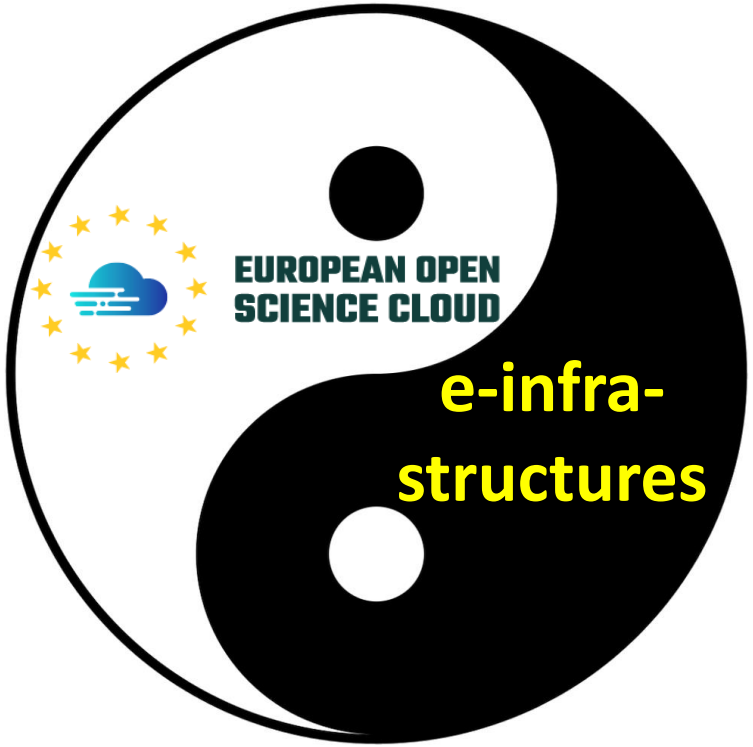
METADATA

Contextual documentation

In order for Digital Objects to be assessable and reusable, they should be accompanied by sufficient metadata and documentation. Basic metadata will enable data discovery, but much richer information and provenance is required to understand how, why, when and by whom the objects were created. To enable the broadest reuse, they should be accompanied by a plurality of relevant attributes and a clear and accessible usage license.

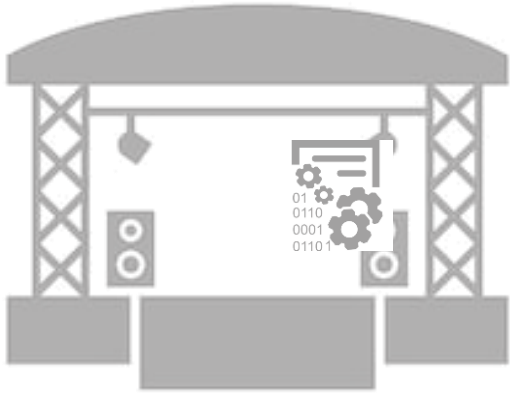
Twinning the data- to the e-infrastructure

EOSC could be seen as a twin sister (or brother) of the e-infra-structure organisations. One offering the store, compute and connect services and the other servicing the data and the interoperability.



“A web of scientific insight”

- Web of FAIR research data (and services)
- Federation of existing services
- Virtual space where science producers and consumers come together
- An open-ended range of content and services
- Quality mark « Data made in Europe »



Issues in the transition to 2nd stage

- Boundary conditions
- Core activities 2020+
- Data Services infrastructure in countries
- Legal entity
- Business model
- Implementation timeline / workplan

Possible core functions for EOSC 2020+

- Develop and govern federating core
- Manage compliance framework
- Manage trusted certification
- Manage the AAI
- Manage PID policies
- Develop outreach to stakeholders
- Monitor services and transactions
- Manage 'EOSC' trademark(s)
- Contribute to Horizon EU policy

Boundary Conditions for EOSC

- Core funding for EOSC from EU
- Inclusiveness of all stakeholders
- Core follows subsidiarity principle
- Providers with a shared purpose
- Countries have different structures
- Self-inclusivity as much as possible
- Hardware agnostic infrastructure
- Focus on FAIR data and services

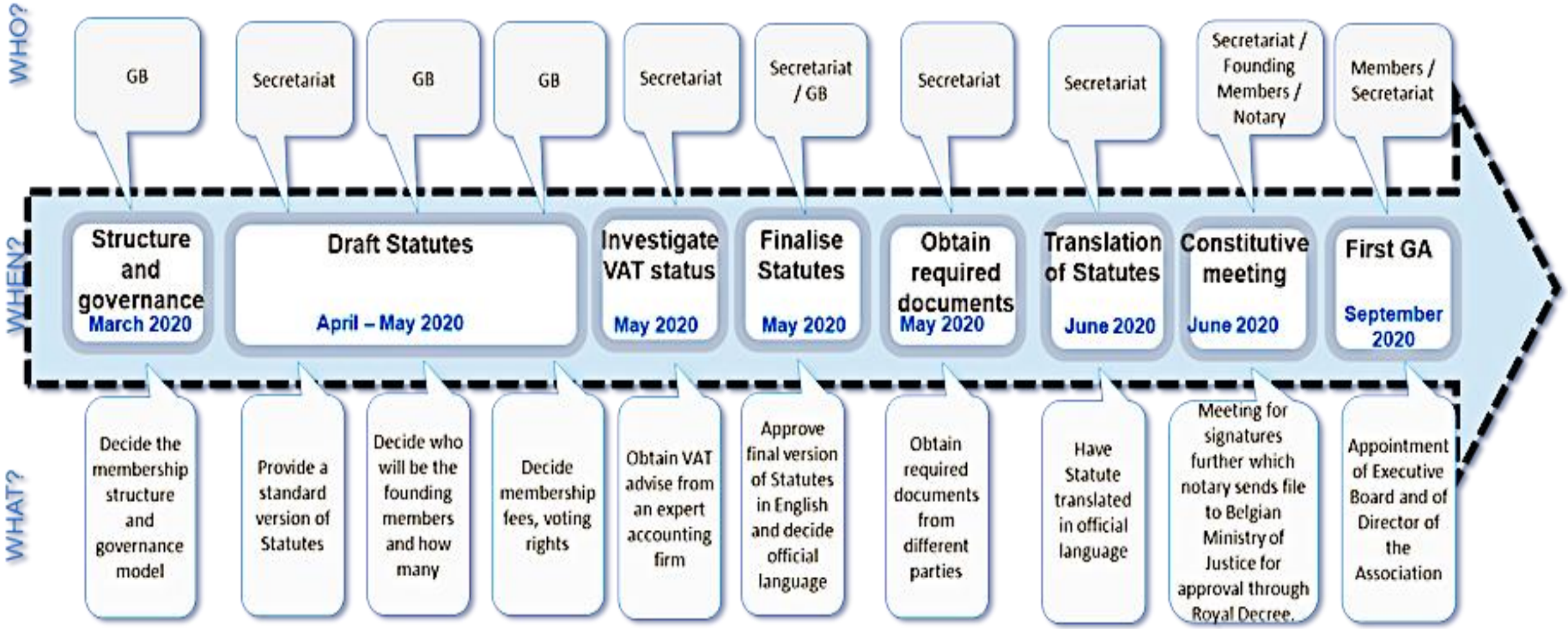
Towards a Legal Entity for EOSC

- EOSC Governance works towards an EOSC Legal Entity
  - a Belgian A(I)SBL (association international sans but lucratif)
- Criteria for Membership are currently being discussed at in the EOSC Governance – potentially to be agreed by May 2020
- A European Partnership Agreement, between the EC and said Association, is set to begin from the beginning of 2021
  - After December 2020, the Governance and Executive Boards, along with the EOSC Working Groups, will cease to exist
- Joining the EOSC Legal Entity = Joining the EOSC Partnership!

EOSC Legal Entity at inception(mission and role)

- At the inception, the EOSC legal entity will be a relatively small-sized association that establishes, in an open and transparent manner, a vision and roadmap for the implementation of the European Open Science Cloud through the development of a Strategic Research and Innovation Agenda (SRIA). The EOSC Legal Entity will then monitor and coordinate the implementation of the SRIA following a two-tier approach:
  - The tier implemented by the Commission: based mainly on grants following calls for proposals set out in Commission Horizon Europe Work Programmes;
  - The tier implemented by the members of the EOSC Association: under their own responsibility and rules.
- At a later stage, when a sustainable business model is determined, the EOSC Legal Entity may also be able to take on, or outsource some EOSC operational core functions directly itself (assuming it is provided the resources to do that). The ambition of the EOSC legal entity is to expand its membership in such a way that it will progressively become a wide, inclusive and representative partnership of EOSC entities committed to implement the EOSC vision.

The what, when and who
A proposed timeline



Items worked on at the moment

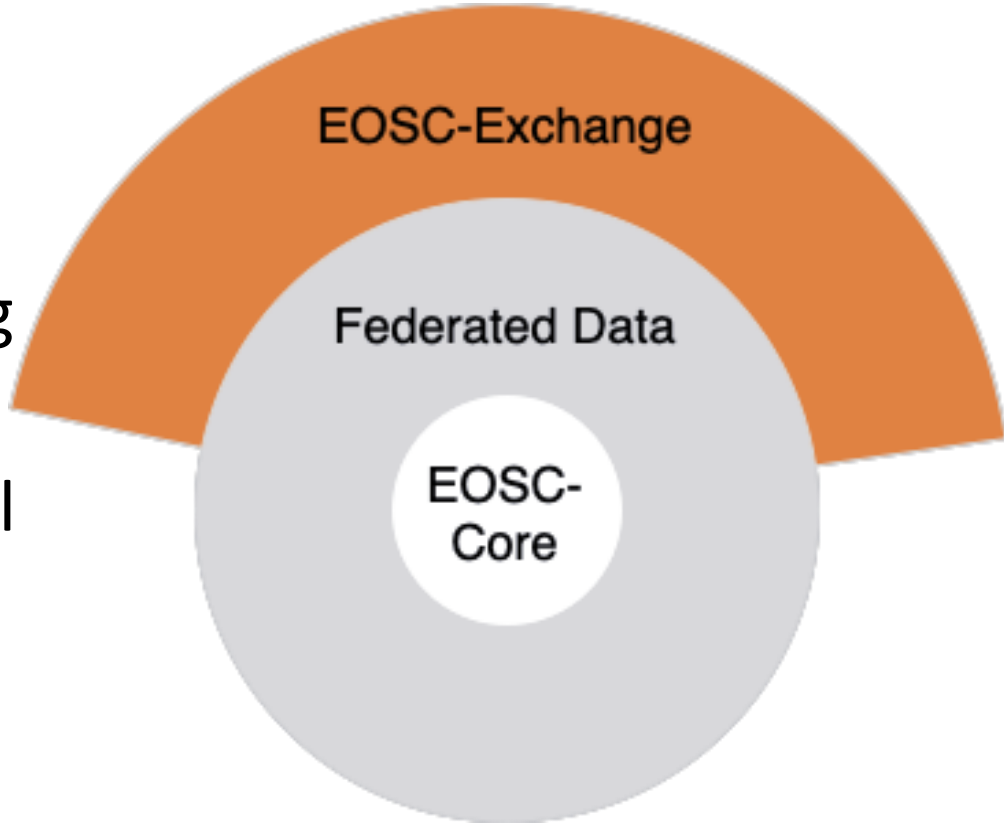
- EOSC Partnership Proposal
- EOSC Strategic Research and Innovation Agenda
  - Update in May during EOSC-hub week
  - Consultation on the draft SRIA by June 2020
- EOSC 2020 Workplan
- RoP Consultation
- Landscape analysis report

A first iteration of EOSC by 2020

- Agreed and tested Rules of Participation *(RoP WG)*
- Analysis of the existing national infrastructures and policies. *(Landscape WG)*
- Financing model, legal entity & post 2020 governance structure *(Sustainability WG)*
- Operational federated core *(Architecture WG)*
- Initial set of EOSC data and services *(Architecture WG)*
- EOSC Interoperability Framework *(Architecture and FAIR WGs)*
- Persistent Identifier policy *(FAIR and Architecture WGs)*
- Metrics for FAIR data and certified services *(FAIR WG)*

First iteration-minimum valuable EOSC (MValE)

- The MValE includes EOSC-Core and EOSC-Exchange which work with federated FAIR datasets
- MValE must enable the federation of existing and planned research data infrastructures
- Federate the disciplinary cluster and regional projects as a critical first step
- Begin with simple use cases – open data not sensitive or closed



EOSC-Core: functions and proposed coverage

Functions

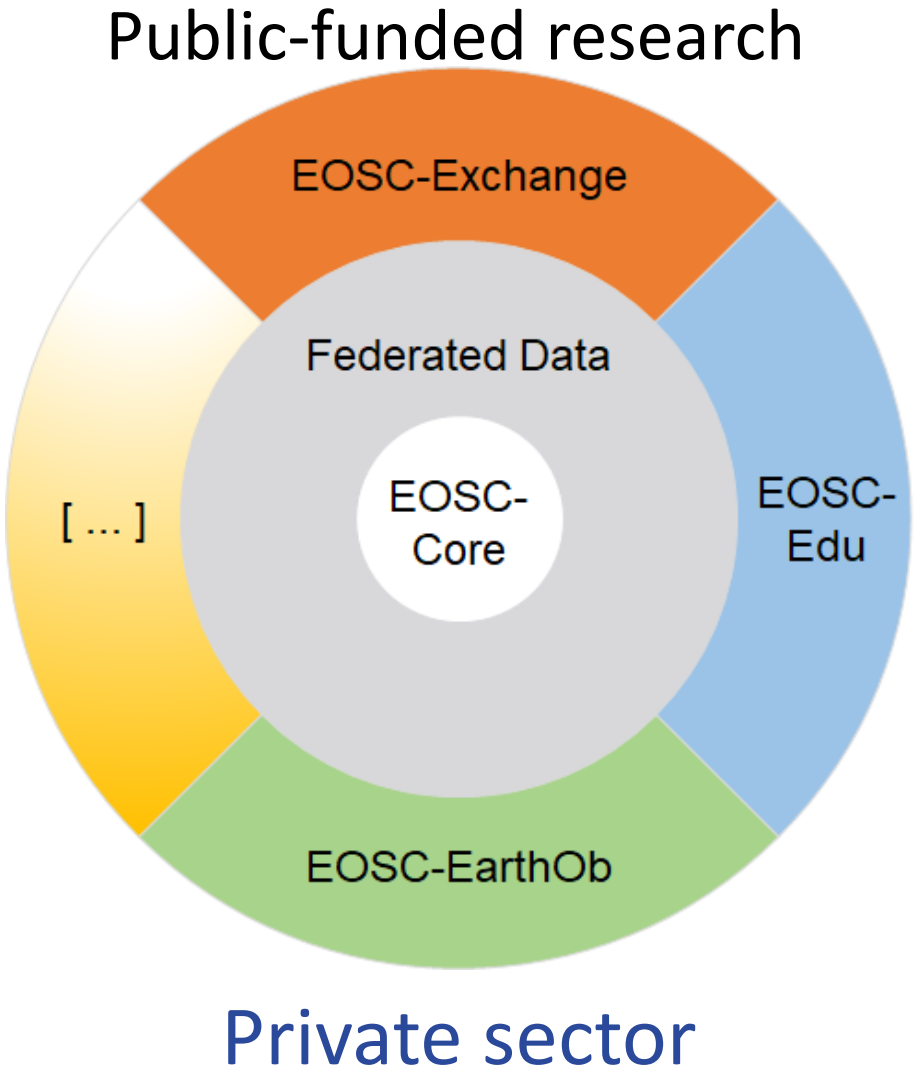
- Provides the means to discover, share, access and re-use data and initial services
- Will not store, transport or process data, at least initially
- Should be used as widely as possible → will be accessible to any authenticated user to promote open research across Europe

Proposed coverage

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

Proposed second and third iterations

- Extensions to serve public sector and industry
- These are not completely new users as some public sector and industrial partners will already use MValE
- Would ideally be one ‘marketplace’ but differing requirements and legislation may require linked but alternately governed spaces

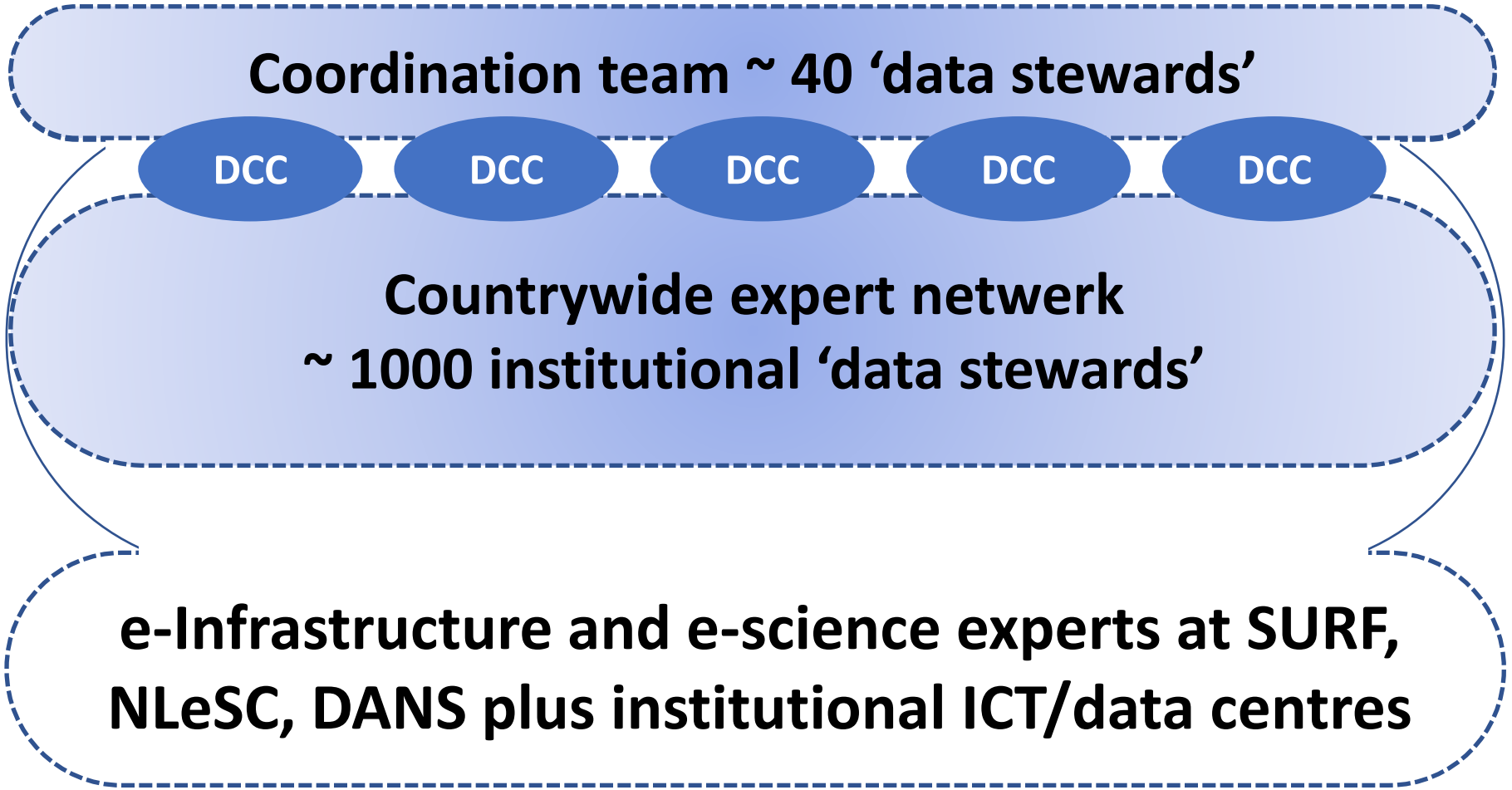


THANK YOU

EUROPEAN OPEN SCIENCE CLOUD

BACKUP

NL-National Data Services Infrastructure



EOSC EB WG Architecture

WHAT? Propose technical framework to enable and sustain an evolving EOSC including interoperability to build federation of systems

WHY? Define interoperability layer necessary for seamless operation of EOSC federation core > **Identify the federated core of EOSC!**

HOW? Review current offerings and define:
(1) EOSC core services and their interfaces
(2) EOSC OS APIs for reuse by thematic services
(3) EOSC portal components and catalogues
(4) EOSC data description standards
(5) Other standards and best practices
(6) Converge towards globally-accepted frameworks

https://www.eoscsecretariat.eu/working-groups/architecture-working-group

Chair: Jean-Francois Abramatic
Independent Expert

WG ARCHITECRURE

WG members have identified 5 key areas of work

- **Authentication & Authorization Infrastructure (AAI)**
- **Persistent Identifier policy (with FAIR WG)**
- **Glossary (with Landscape WG)**
- **Metadata**
- **Data Access**

Chair: Jean-Francois Abramatic
Independent Expert

Two initial task forces launched – AAI & PID

EOSC EB WG Landscape

WHAT? Provide options for progressive convergence and alignment of structures and initiatives in Europe to facilitate the EOSC

WHY? **Consolidate the existing landscape!** > National RIs, eIs, ESFRIs, open science policies, cluster projects, and thematic initiatives

HOW?

(1) Build on outcomes of H2020 projects and input from MS
(2) Map EOSC-relevant national infrastructures and spending
(3) Take stock of federation constraints and opportunities
(4) Propose mechanisms and best practices for alignment
(5) Analyse MS preparedness for financial resources, political stability and infrastructural planning

Chairs: Jan Hrušák & John Womersley

https://www.eoscsecretariat.eu/working-groups/landscape-working-group

23

WG LANDSCAPE

Introduction Background

Inventory - stocktaking and synthesis from existing documents

A) EOSC related infrastructures with reference framework
B) policies + legislation

Mechanisms, modus operandi, trends and developments

A) operations of infrastructures
B) policies evolvement

Analysis and generalization, good practice examples, preparedness of MS

in collaboration with the other WGs and the projects (external subcontractor)

- **Initial document drafted. Presented to EB on 25 October & GB thereafter**
- **Second phase inputs to come in Q4 2019 from the INFRAEOSC-5B projects**
- **Plan to subcontract analytical work on MS preparedness**
- **Validation workshop foreseen for January 2020**

Chair: Jan Hrusak
Independent Expert

EOSC EB WG Sustainability

WHAT? Provide strategic, legal and financing recommendations for an operational, scalable and sustainable EOSC federation after 2020

WHY? A clear understanding of key strategic, financial and legal aspects for 2nd stage of EOSC > **Towards a sustainable EOSC!**

HOW?

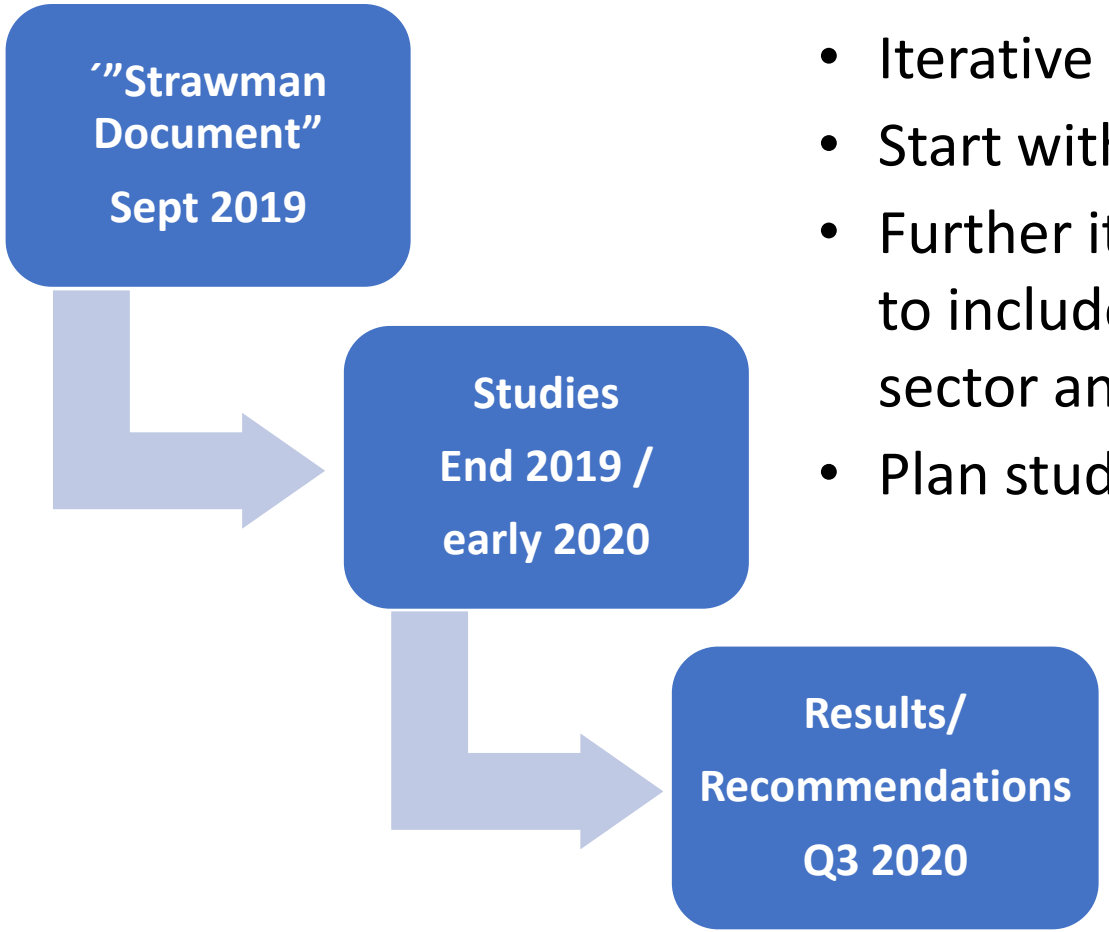
(1) Analyse business models and implications on legal entity, costs, regulations, financial strategies, and NDI alignment
(2) Map potential legal entities, taking into account legislation
(3) Examine options for a governance framework for EOSC
(4) Analyse regulatory/policy environments and assess impact of proposed structures and funding streams

Chair: Rupert Lueck

https://www.eoscsecretariat.eu/working-groups/sustainability-working-group

25

WG SUSTAINEBLITY



Strawman document

- Iterative approach – gradual growth of EOSC
- Start with MVP for publicly-funded researchers
- Further iterations, establishing Digital Marketplaces to include additional services for researchers, public sector and industry
- Plan studies on legal entity and business models

Co-Chair: Rupert Lueck
EMBL representative

Co-Chair: Lidia Borrell-Damian
Science Europe representative

EOSC EB WG Rules of Participation

WHAT? Recommend rules of participation defining rights, obligations, and accountability to govern EOSC users, providers, and operators and also ientify common requirements

WHY? Guarantee open, secure and cost-effective federated EOSC with services of documented quality > **Assessable Quality!**

HOW? Identification of different rules taking into account:
(1) Differing roles and services
(2) Scientific disciplines
(3) Preparedness of infrastructures, services and MS
(4) Growing range of EOSC users and providers
(5) Changing needs and practices including regulations

https://www.eoscsecretariat.eu/working-groups/rules-participation-working-group

Chairs: Juan Bicarregui & Wolfram Horstmann

27

WG RoP

Working on scoping and partitioning the scope

Different RoP will be required for different roles (users, providers)

Questions for discussion:

1. How should the RoP constrain what service providers can require of users?
2. How should users discover and access EOSC services?
3. How should EOSC minimise the barriers to service providers taking part?
4. To what extent should all of the above apply to Non-European users and service providers?

Chair: Juan Bicarregui
RDA representative

EOSC EB WG FAIR

WHAT? Provide recommendations on implementation of FAIR including service requirements for cross-disciplinary interoperability

WHY? Connect people, data, services via standards > **Be the glue!**

HOW?

(1) Data standards & sharing agreements
(2) Upscale best-practice solutions
(3) EOSC interoperability framework
(4) Identify service requirements for FAIR
(5) Persistent identifier policy for EOSC
(6) Frameworks for FAIR data and services
(7) Converge towards globally-accepted frameworks

https://www.eoscsecretariat.eu/working-groups/fair-working-group

Chair: Sarah Jones
Independent Expert

WG FAIR

Key outputs / progress to date

Defining matrix map FAIR practice / maturity

Drafting initial PID policy to consult at events

Liaising with FAIRsFAIR and RDA on metrics & certification

Considering what the EOSC Interoperability Framework should cover and how it should be formed

Chair: Sarah Jones
Independent Expert

WG FAIR

Dynamic duo
- Rob Hooft
- Marta Teperek

Semantic specialists
- Oscar Corcho
- Krzys Kurowski

FAIR practice → Interoperability

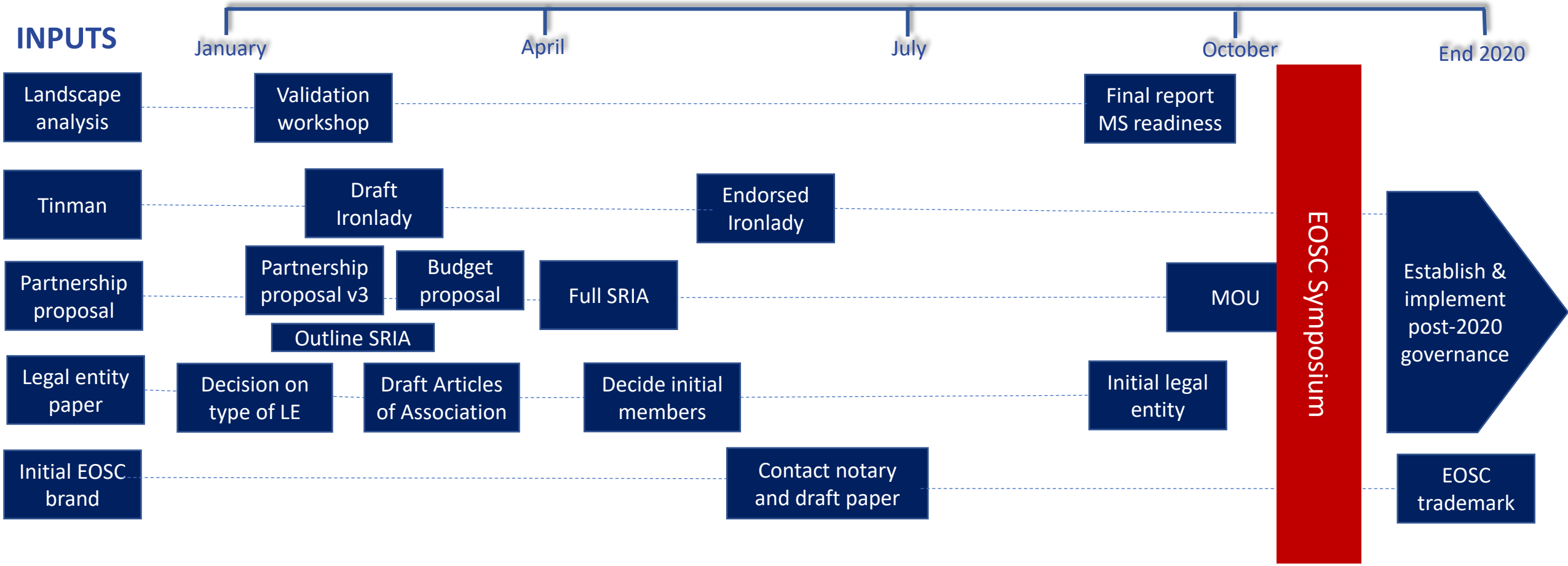
PID policy

Metrics & certification

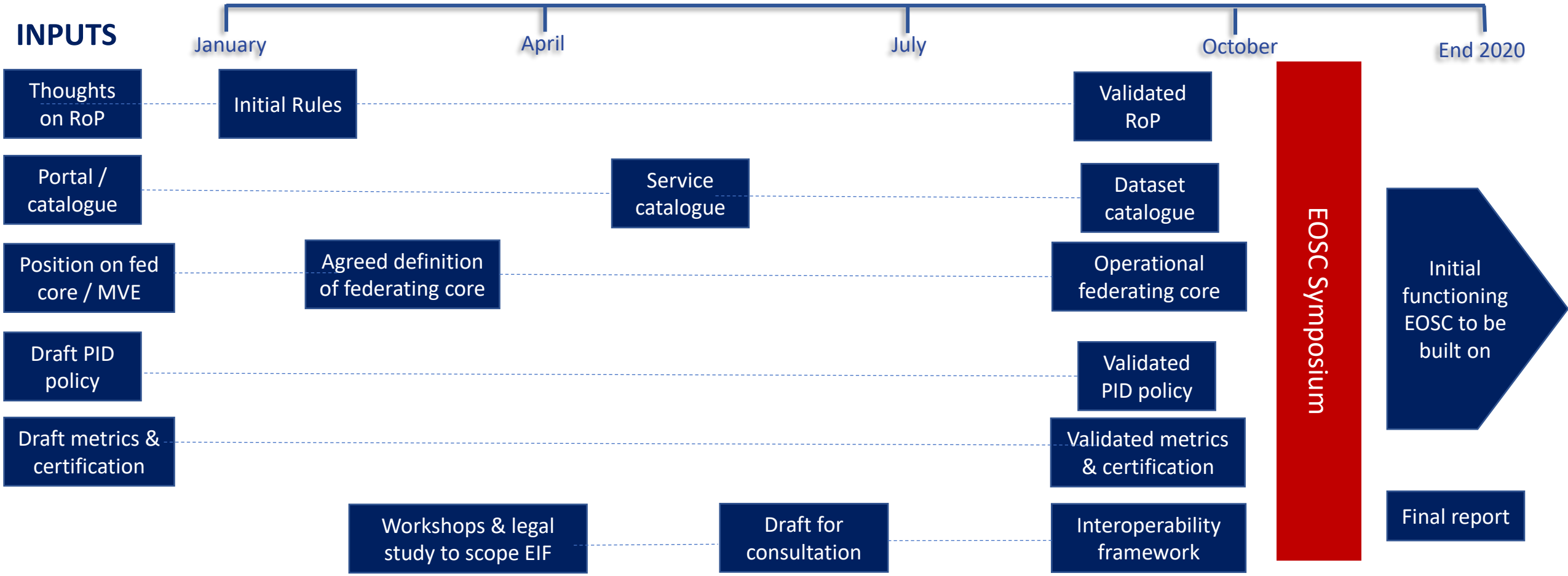
Identifier nerds
- Peter Wittenburg
- Rachael Kotarski

Research leads
- Francoise Genova
- Magnus Aronsen

Timeline mapping out strategic aspects for 2020



Timeline mapping out implementation aspects for 2020



First phase of EOSC

EOSC

- present stage of EOSC: 2018-2020
- H2020 projects: M€250 - M€600
- Address six roadmap action lines:
  - ◆ data ◆ services ◆ architecture
  - ◆ access ◆ rules ◆ governance

"start of the second phase of the implementation is dependent on an evaluation by the Commission and the Member States of the first phase and without prejudice to the Multiannual Financial Framework after 2020"

Current EOSC governance tasks

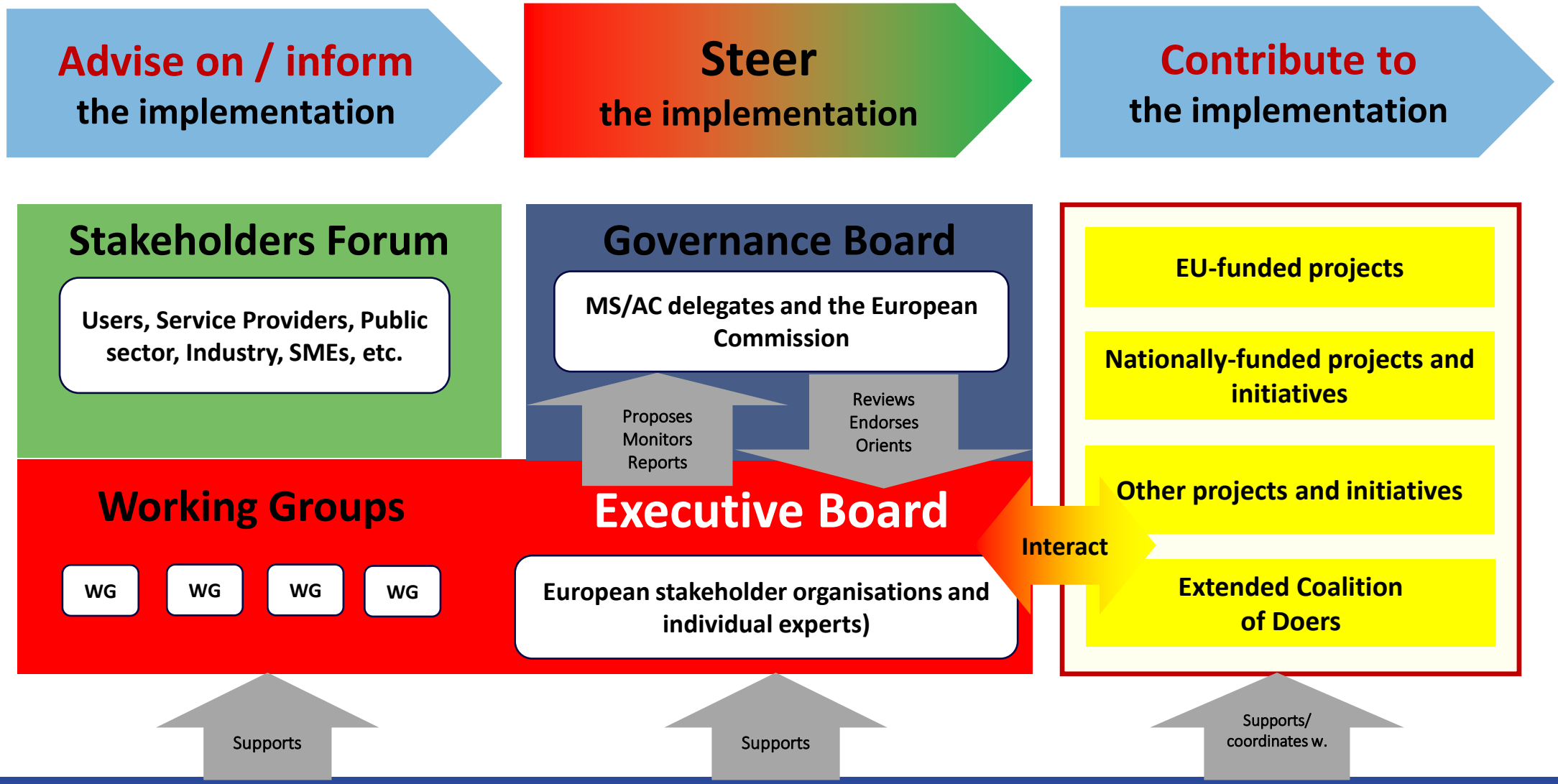
- steer initial implementation
- involve all R&I stakeholders
- transition to 2nd stage 2020+

Some activities

- 35+ active EOSC projects
- EOSC SIP published
- EOSC Workplan
- Six active Working Groups
- About 10 interim reports
- EOSC Concertation Days
- EOSC @ RDA Event – Oct.
- EOSC Symposium – Nov.
- Partnership proposal draft4

EOSCsecretariat

Governance of EOSC 2019-2020



EOSCSecretariat.eu
Coordination and Support Action

Issuess relevant for forming the association

- Strategic Research and Innovation Agenda
- Alignment between national and EU level policies
- Decision-making (depends on final governance structure of AISBL)

- Loss of Focus
- Potential fragmentation
- Legal, tax & VAT compliance considerations for members