



# EOSCPilot Status Update

Matthew Dovey, [matthew.dovey@jisc.ac.uk](mailto:matthew.dovey@jisc.ac.uk)

**EOSC**<sub>pilot</sub>  
The European Open Science  
Cloud for Research Pilot Project  
[www.eoscpilot.eu](http://www.eoscpilot.eu)



# *EOSCpilot*: High Level Aims

The *EOSCpilot* project will support the first phase of development of EOSC:

- 🔗 **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
- 🔗 **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
- 🔗 **Establish the governance framework** for the EOSC and contribute to the development of European open science policy and best practice



# EOSCpilot Challenges

Three types of challenges:

**Scientific Challenges are really *Opportunities***

🔗 **Scientific Challenges:** deploying the EOSC to deliver Open Science

**Technical Challenges are *Barriers to overcome***

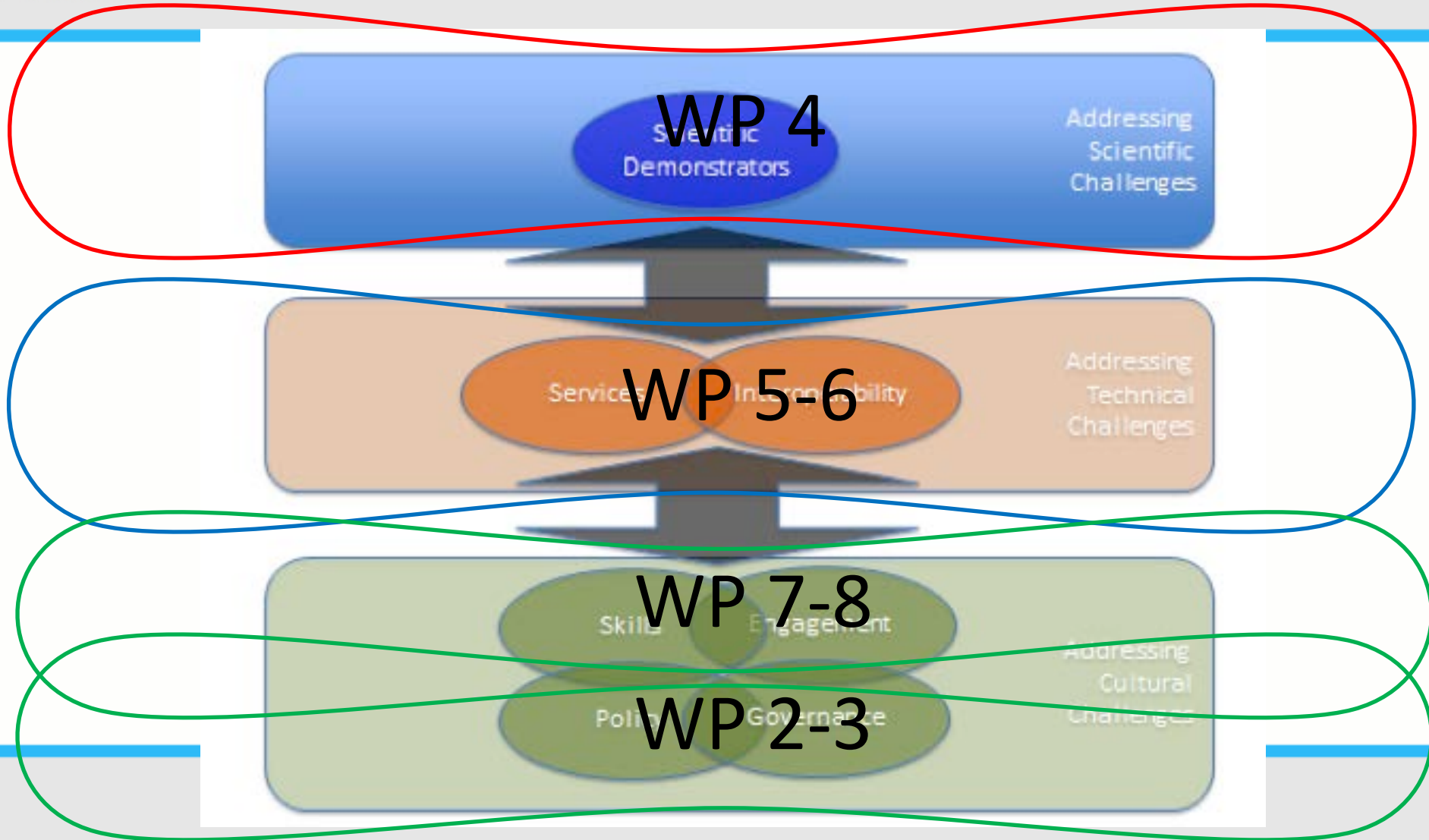
🔗 **Technical Challenges:** developing technical solutions that meet the scientific needs

**Cultural Challenges are also *Barriers***

🔗 **Cultural Challenges:** adopting new, more open ways of working



# Workpackage Challenges







# Workpackage Level Objectives

## Science Demonstrators Objective

- To develop a number of Science Demonstrators ... to drive the development of the EOSC.

## Services Objective

- To create a number of EOSC pilot services that federate data, infrastructure and services ...

## Interoperability Objective

- To define and implement specifications, interfaces, standards and processes that ...underpin interoperability and sharing ...

## Governance Objective

- To design and trial a stakeholder-driven governance framework ...

## Policy Objective.

- To establish the policy environment required for the effective operation...

## Skills Objective.

- To develop common standards and assessment frameworks to ensure ...

## Community Engagement Objective.

- To identify and bring together ... the major groups of stakeholders ...



# Scientific Demonstrators

**The science areas targeted in EOSCpilot have been covered as follows:**

First five Science Demonstrators (pre-selected)

**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 (including materials science):** 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)



# Scientific Demonstrators

## Second five Science Demonstrators:

**Energy Research – PROMINENCE:** HPCaaS for Fusion - Access to HPC class nodes for the Fusion Research community through a cloud interface

**Earth Sciences – EPOS/VERCE:** Virtual Earthquake and Computational Earth Science e-science environment in Europe

**Life Sciences / Genome Research:** Life Sciences Datasets: Leveraging EOSC to offload updating and standardizing life sciences datasets and to improve studies reproducibility, reusability and interoperability

**Life Sciences / Structural Biology:** CryoEM Workflows: Linking distributed data and data analysis resources as workflows in Structural Biology with cryo Electron Microscopy: Interoperability and reuse

**Physical Sciences / Astronomy:** LOFAR Data: Easy access to LOFAR data and knowledge extraction through Open Science Cloud



## Third five Science Demonstrators:

**Generic Technology: Frictionless Data Exchange** Across Research Data, Software and Scientific Paper Repositories

**Life Sciences – Genome Research - Bioimaging:** Mining a large image repository to extract new biological knowledge about human gene function.

**Astro Sciences: VisIVO:** Data Knowledge Visual Analytics Framework for Astrophysics

**Earth Sciences – Hydrology:** Switching on the EOSC for Reproducible Computational Hydrology by FAIR-ifying eWaterCycle and SWITCH-ON.

**Social Sciences and Humanities: VisualMedia:** a service for sharing and visualizing visual media files on the web





## Objectives

### Infrastructure interoperability:

facilitate the most adequate infrastructures for the treatment of extensive amounts of data. Demonstrate with multi-infrastructure, multi-community pilots (science demonstrators from WP3 & WP4)

### Research and Data Interoperability:

Data & services to be findable, accessible, interoperable and reusable (FAIR)

### Testbeds for interoperability:

Put to work Science Demonstrator and learn about interoperability issues and solutions

### Key Output:

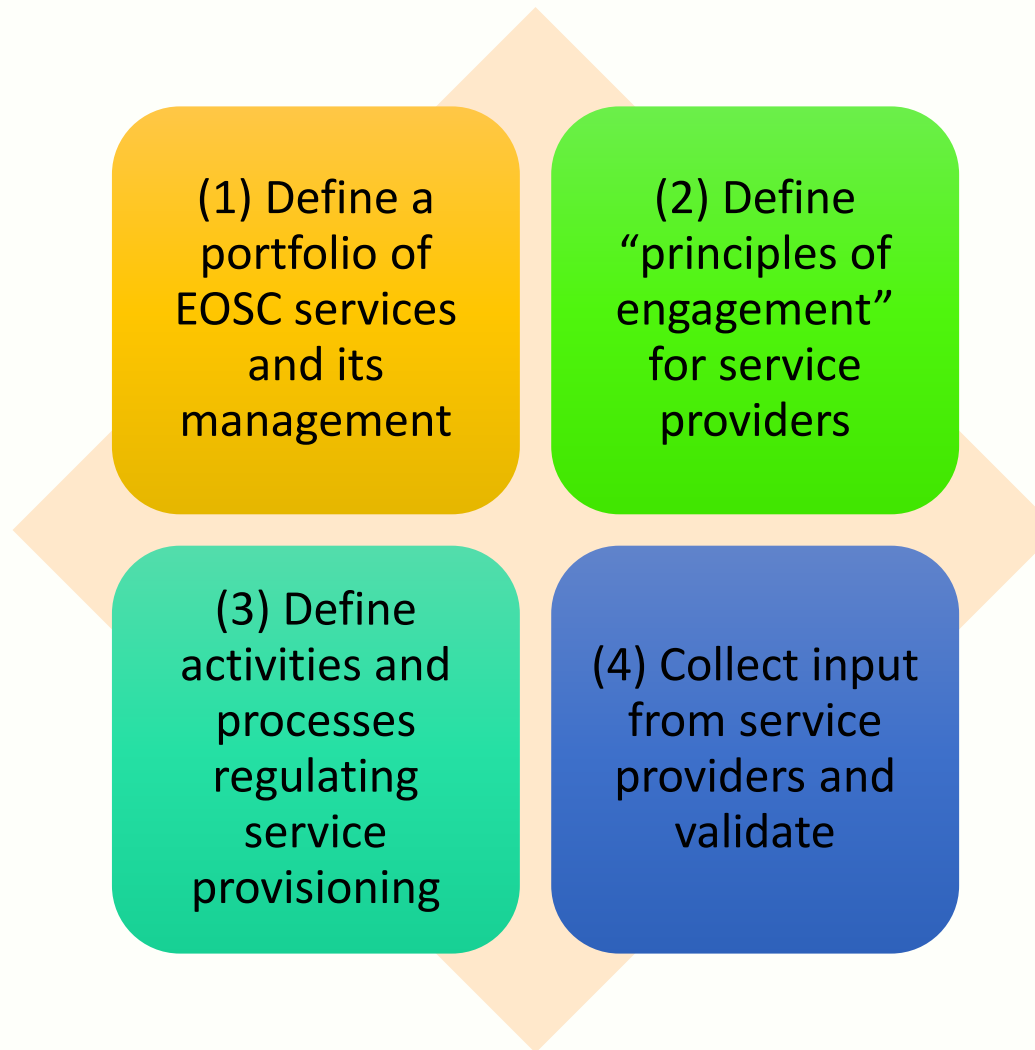
*The design of a future EOSC based on **federated interoperable services** meeting the **needs** of the thematic research domains and wider user base.*

#### WP Outputs so far:

- Gap analysis
- EOSC architecture
- Report on data interoperability
- Requirements for testbeds

#### More to come:

- Testbed reports
- Final report on data interoperability
- Final EOSC architecture





## Revise & enhance EOSC architecture

- Actor roles and activities
- Services
- Assumptions

## Apply architecture to specific services

- Validate their deployment model
- e.g. service catalogue: centralised, distributed, replicated

## Establish links

- between roles, activities, EOSC system of system components, policies
- Compare to current service provisioning models

## Revise glossary and apply across all WPs



# Service Portfolio Work plan

Determine the 'core services'

- and their functionality

Define for each service category

- Principles of Engagement
- Service components

Define service portfolio management for each service category

- Work through examples



## M1-6: Review of existing service management frameworks

- Initial survey focusing on existing e-Infrastructures & Research Infrastructures
  - EGI, EUDAT, PRACE, OpenAIRE, GEANT, ESFRIs, etc.

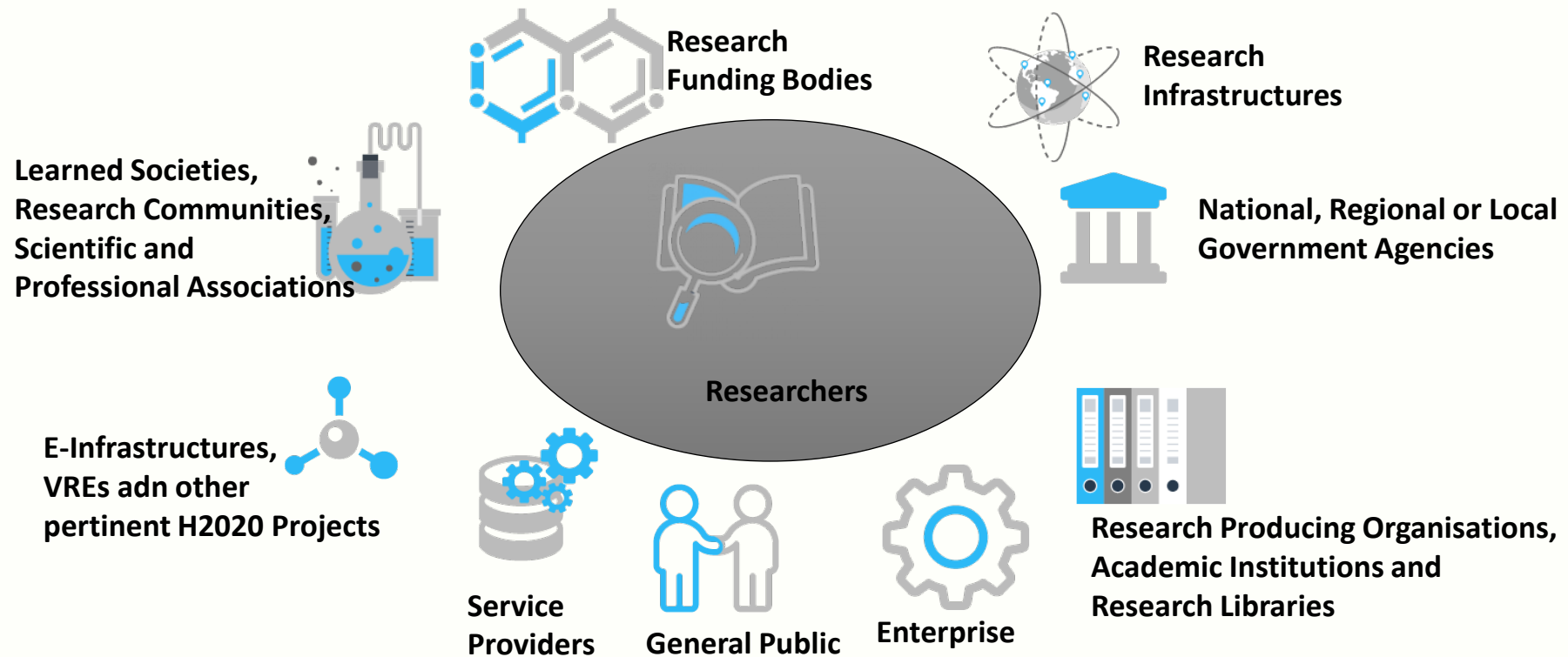
## M6-12: Investigation on Rules of Engagement (with WP2)

- Initial framing of the issue, discussion with stakeholders

## M12-M18: Design of EOSC Service Management Framework

- Define framework to organize and manage services within a future EOSC







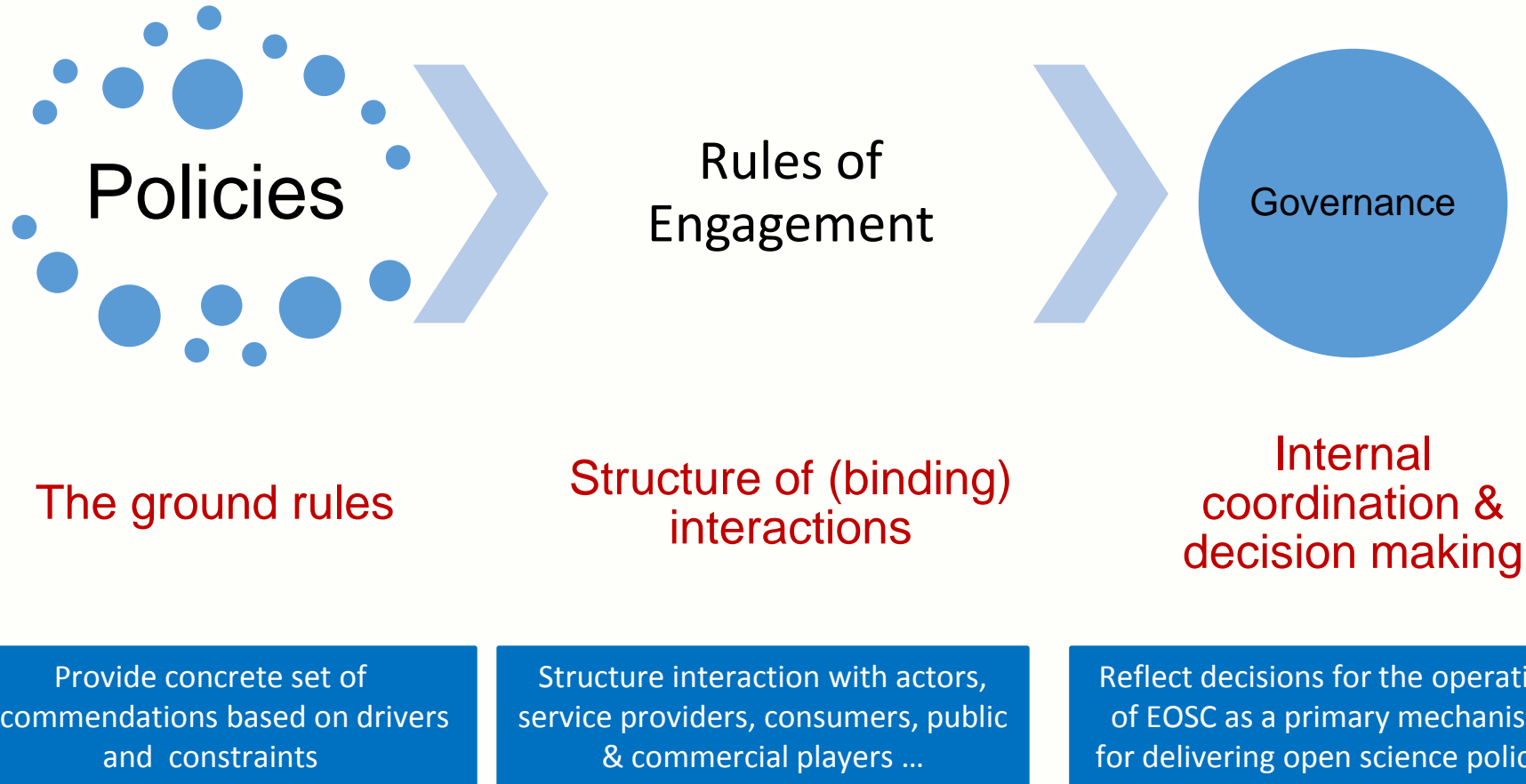
- 👤 Mapping skills landscape – focus on data stewardship
- 👤 Definition of stewardship; extension of EDISON competence model to this area
- 👤 Used the competence model to highlight gaps in skills provision (informed by science demonstrators)
- 👤 Extensive analysis of alternative models for cross-domain training provision



- 70+ competencies needed by researchers and research groups for data stewardship
- Organised around research data lifecycle
- Helps answer:
  - What skills are needed to build, operate, support, use a particular EOSC service?
  - What capabilities should my organisation build, through recruitment, training, staff development?
  - What should our job descriptions contain relating to data stewardship?
- Utilises EDISON, FitSM, CoreTrustSeal, FOSTER+ and other work



# Policies in EOSC: the overall picture





# Priorities for EOSC : a matrix of policies

*Focus on specific policy areas around*

Infrastructure  
and services

Free flow of  
data

Data & open  
science skills

Public  
procurement

*In the context of*

Data  
Economy

Digitising European  
industry

*Paying attention to*

IPR

Personal  
Data

PSI

*...as they are key to implementation for data  
and information/ content flow*

Open Science, Open Innovation, Open to the world

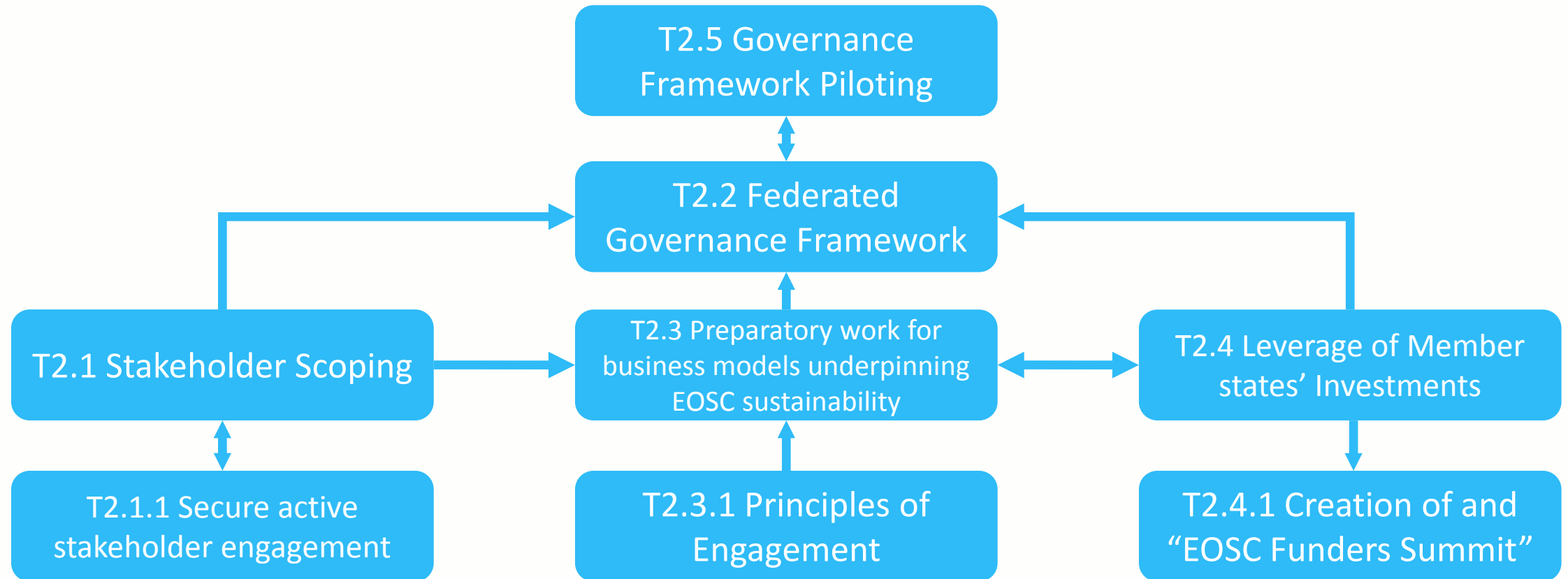




## First result: high level recommendations

1. Produce **consistent policies** at the EU, the Member State and the institutional level
2. **Standardise interactions** at the organisational and institutional (micro) level
3. Focus on the **interactions with the industry**, where the greater inefficiencies currently exist
4. Focus on **interactions with platforms** in order to maximize value, protect data ownership and portability and avoid vendor lock-in
5. **Automate the application of policies supporting OS** by design and default, as well as data sovereignty for the user
6. Support the development of the e-infrastructures services that could **use the EU GDPR as a competitive advantage**

# Governance



# Principles of Engagement

- Investigate organisational Rules of Engagement for scientific users and service providers in the EOSC
- The aim is to recommend a minimum set of compatible organisational rules and practices, necessary for EOSC participation and function



# EOSC Resource Model: Principles of Engagement

*EOSC resources = technical, middleware, knowledge, access and facilitation services*

***EOSC Resource = Services + Data + People***



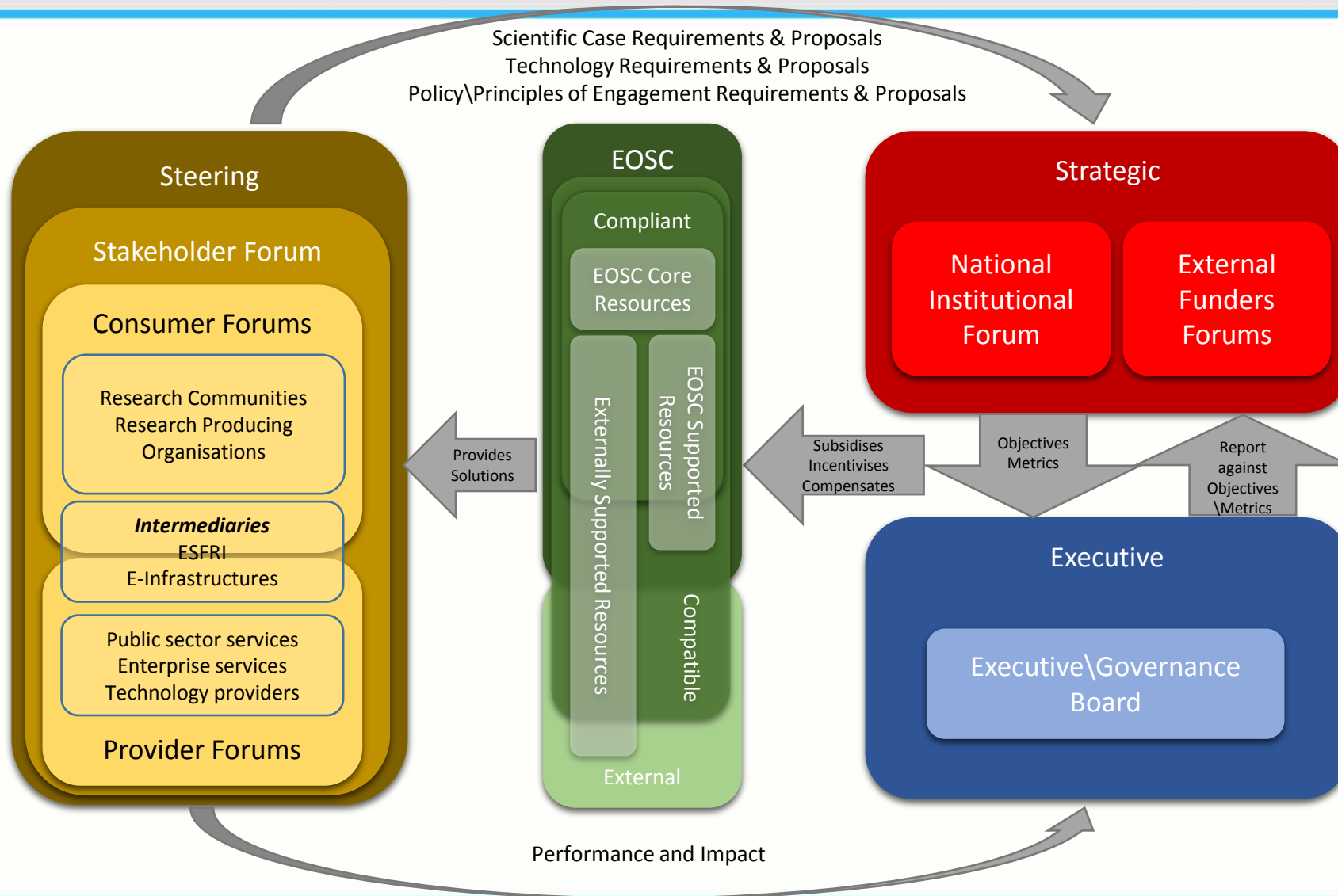
**Compliant** Most of the resources within EOSC fully compliant with the Principles of Engagement

**Compatible** Not fully compliant (yet), technically compatible with the EOSC, of value to EOSC Consumers

**External** Resources outside of EOSC, of value to EOSC consumers, may or not be technically compatible with EOSC resources, “non-EOSC approved players are free to explore any role in the Open Science ecosystem they wish, even if they do not adhere to the RoE”



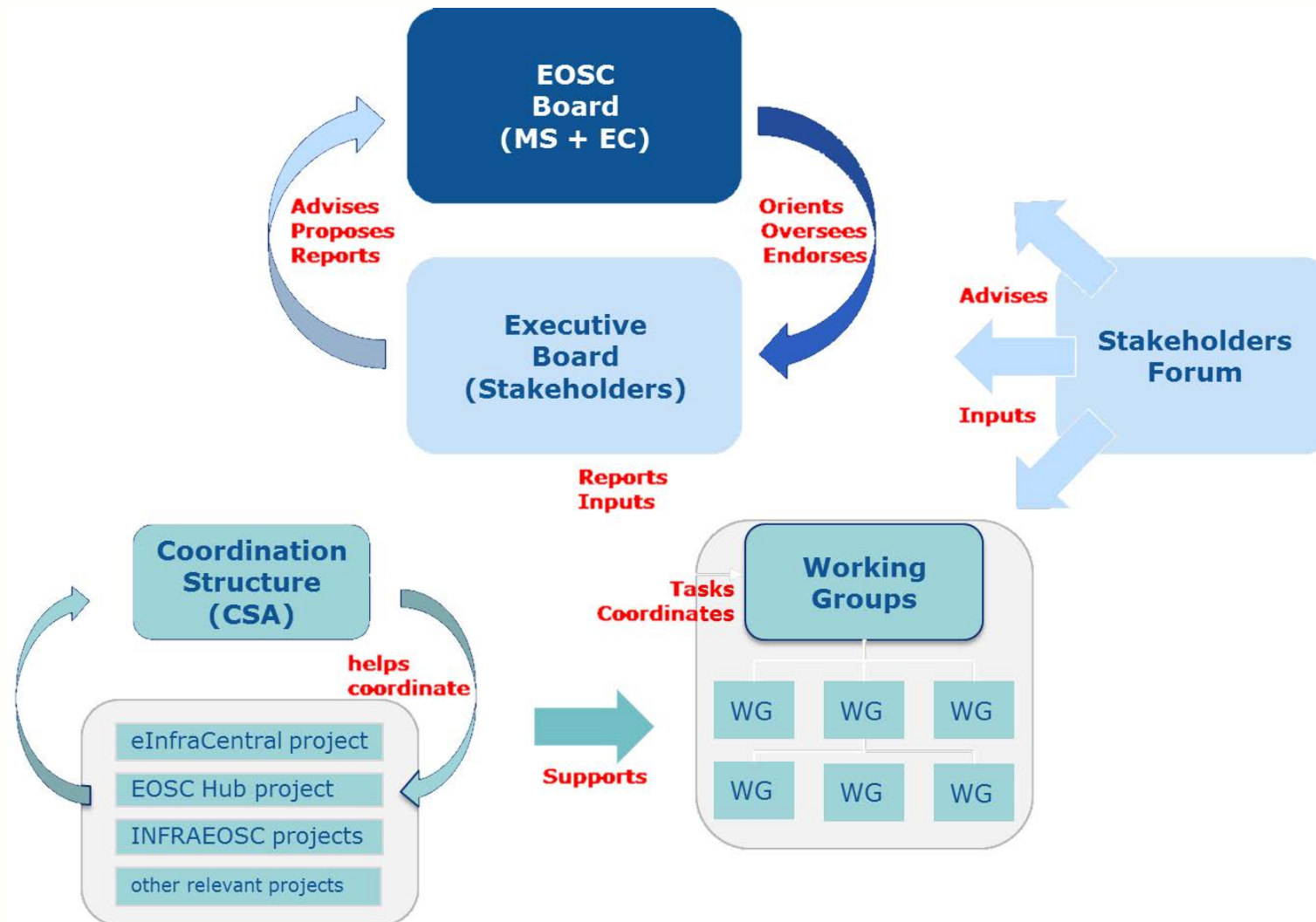
# EOSC Governance Decision Flow Model



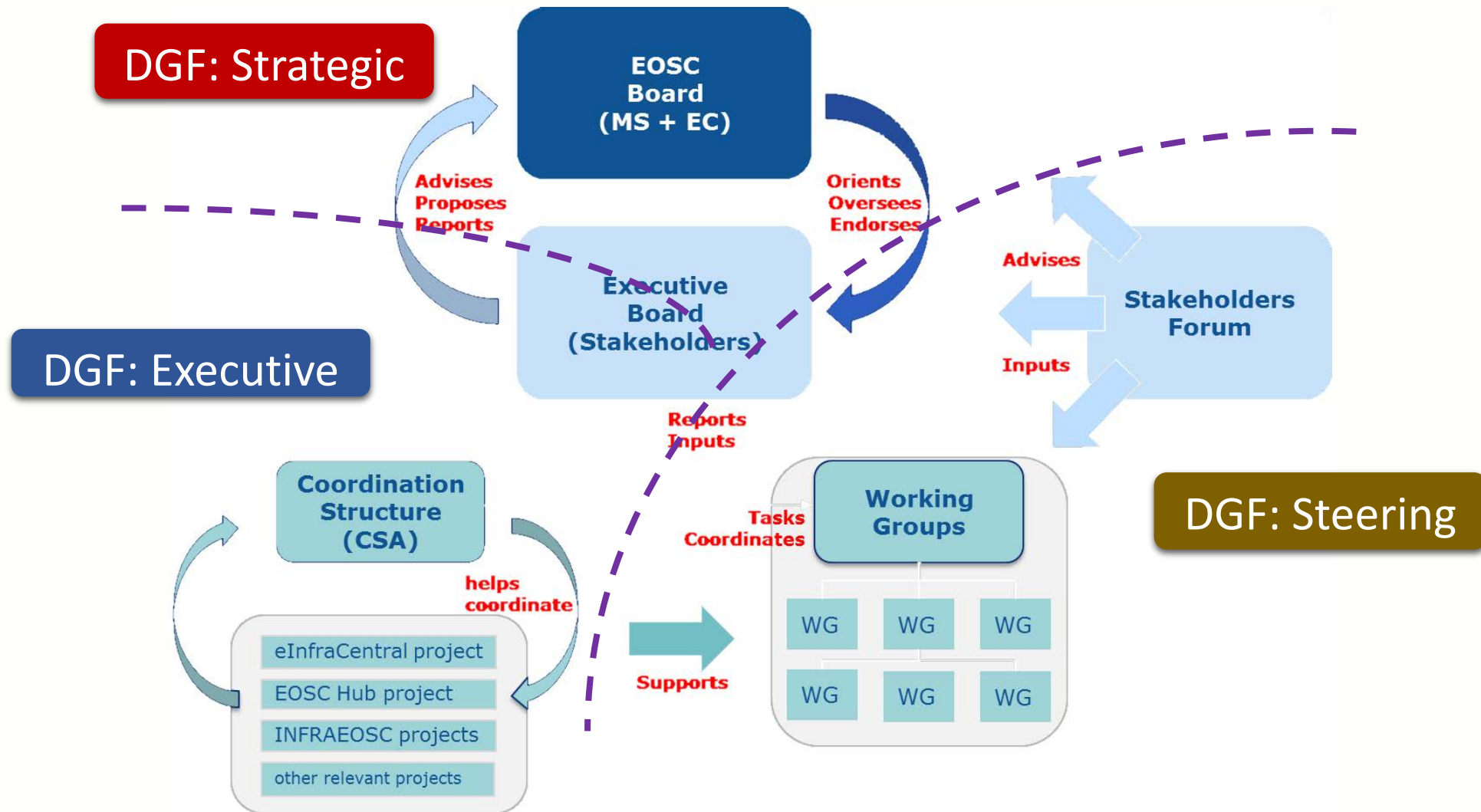




# SWD - Governance



# Governance Crosswalk



<https://eoscpilot.eu>

