



European
Commission

Final report and
recommendations of
the Commission 2nd High
Level Expert Group on the
European Open Science
Cloud (EOSC)



PROMPTING AN EOSC IN PRACTICE

*Research and
Innovation*

2018

Prompting an EOSC in practice

European Commission
Directorate-General for Research and Innovation
Directorate B – Open Innovation and Open Science
Unit B2 – Open Science

Contact Maarja Adojaan
E-mail maarja.adojaan@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu

European Commission
B-1049 Brussels

Manuscript completed in November 2018.

This document has been prepared for the European Commission however it reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

More information on the European Union is available on the internet (<http://europa.eu>).

Luxembourg: Publications Office of the European Union, 2018

Print	ISBN 978-92-79-94835-0	doi:10.2777/620195	KI-03-18-339-EN-C
PDF	ISBN 978-92-79-94836-7	doi:10.2777/112658	KI-03-18-339-EN-N

© European Union, 2018.

Reuse is authorised provided the source is acknowledged. The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39).

For any use or reproduction of photos or other material that is not under the EU copyright, permission must be sought directly from the copyright holders.

Prompting an EOSC in practice

Final report and recommendations
of the Commission 2nd High Level Expert Group [2017-2018]
on the European Open Science Cloud (EOSC)

2018

Members: Silvana Muscella (Chair), Isabel Campos Plasencia, George A. Komatsoulis, Andreas Mortensen, Toivo Räm, François Robida, Linda Strick, Klaus Tochtermann, Žiga Turk, Ross Wilkinson

The Expert Group operates in full autonomy and transparency. The views and recommendations in this report are those of the Expert Group members acting in their personal capacities and do not necessarily represent the opinions of the European Commission or any other body; nor do they commit the Commission to implement them.

Acknowledgements

The authors are grateful to many experts for their input and private conversations. In particular, they would like to thank:

Carmela Asero	Former Policy analyst at the European Commission - DG Research and Innovation (DG RTD)
Alberto Azevedo	Oceanographer at National Laboratory for Civil Engineering (LNEC) for OPENCoastS
Paolo Balboni	Founding Partner of ICT Legal Consulting and Professor of Privacy, Cybersecurity, and IT Contract Law at the European Centre on Privacy and Cybersecurity within the Maastricht University Faculty of Law
Juan Bicarregui	Head of the Data Services division in the Scientific Computing Department at Science & Technology Facilities Council (STFC)
Paolo Budroni	Head of Department of e-Infrastructure at Vienna University Library and National Delegate at e-IRG
Donatella Castelli	Senior Researcher Consiglio Nazionale delle Ricerche (CNR-ISTI)
Ron Dekker	Chief Executive Officer and legal representative of Consortium of European Social Science Data Archives (CESSDA)
Matthew Dovey	Head of e-Infrastructure Strategy at JISC
Willem Elbers	Software Engineer / Infrastructure Specialist for CLARIN
Tiziana Ferrari	Technical Director at the EGI Foundation and Coordinator of the EOSC-hub project
Jorge Gomes	Jorge Gomes researcher at Laboratory of Instrumentation and Particle Physics (LIP) for OPENCoastS
Silvia Gomez Recio	Secretary General at Young European Research Universities Network (YERUN)
Annabel Grant	Senior Business Development Officer at GÉANT
Stefan Hanslik	Head of Unit Technical Science BMWFW
Simon Hodson	Executive Director of CODATA
Robert Jones	Responsible for Externally Funded Projects at CERN IT Department and Coordinator of Helix Nebula initiative
Sarah Jones	Associate Director of Digital Curation Centre (DCC)
Paolo Lombardi	Director, Trust-IT Services
Hermann Lederer	Deputy director, Max Planck Computing & Data Facility
Damien Lecarpentier	Project Director at CSC– IT Centre for Science
Thomas Lichtenwöhner	Federal Ministry of Education, Science & Research, Austria
Jesús Marco	Research Professor at Instituto de Física de Cantabria (IFCA)
Eva Mendez	Representative of YERUN, Deputy Vice-President for Strategy and Digital Education of Universidad Carlos III de Madrid
Natalia Manola	Research Associate University of Athens
Barend Mons	Professor in Biosemantics at the Leiden University Medical Center (LUMC) and also a board member of the Leiden Centre of Data Science, Director of the International Support and Coordination office of GO FAIR
Steven Newhouse	Head of Technical Services at European Bioinformatics Institute for ELIXIR
Gareth O'Neill	President of Eurodoc
Anabela Oliveira	Head of Division at National Laboratory for Civil Engineering (LNEC) for OPENCoastS
Per Oster	Director of the Research Infrastructures unit at CSC and Project Director of the EOSC-Hub project
Federico Ruggeri	Director Consortium GARR
Simone Sacchi	E-Infrastructure Senior Strategist and Project Manager, Trust-IT Services
Gergely Sipos	Customer and Technical Outreach Manager at Stichting EGI

Dieter Van Uytvanck Technical Director of CLARIN ERIC
Susheel Varma Technical Coordinator at European Bioinformatics Institute for ELIXIR

We would like to thank the following individuals, groups, universities, organisations, alliances and EU projects who contributed to the Open Consultation Platform June 2018:

AIB, CNRS, CESSDA, CERN, DI4R, DANS, DCC, Digital Science, EGI.eu, EMBL- European Molecular Biology Laboratory, EOSC-Hub, EOSCPilot, EIROForum, EURODOC, EUA — European University Association FREYA, GÉANT, GGP, ISAC-CNR, JISC, LIBER, Maastricht university, Max Planck Institute for Software Systems, Monmouth University, NMIS- Networked Multimedia Information, PLAN-E, RCN, Research Council of Norway, Research Data Alliance (RDA), Systems Software Sustainability Institute, Trust-IT services, TU Delft, UGOE, UK Atomic Energy Authority, Utrecht University, YERUN.

And, at the European Commission, for their continued guidance and support:

Maarja Adojaan, Emanuele Barbarossa, Patrick Brenier, Jean-Claude Burgelman, Athanasios Karalopoulos, Wainer Lusoli, Corina Pascu, Michel Schouppe, Katarzyna Szkuta, Victoria Tsoukala.

Legal Notice

This document has been prepared for the European Commission however it reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

More information on the European Union is available on the internet (<http://europa.eu>).

Luxembourg: Publications Office of the European Union, 2018

TABLE OF CONTENTS

GLOSSARY	7
FOREWORD BY COMMISSIONER CARLOS MOEDAS	8
PREFACE	9
EXECUTIVE SUMMARY	10
1. INTRODUCTION	14
1.1 Objectives of the 2nd High Level Expert Group	14
1.2 Scope and structure of the document	14
1.3 Building upon the 1st EOSC HLEG	15
2. THE POLICY LANDSCAPE: FROM COMMUNICATION (COM/20160178final) TO THE STAFF WORKING DOCUMENT (SWD 2018 83final) Council Conclusions TO LAUNCHING OF THE EOSC	16
2.1 Impact of current regulation	17
2.2 The European Open Science Cloud, FAIR & OSPP Synergies	18
3. MAKING EOSC A VIABLE ECOSYSTEM	23
3.1 The EOSC Minimum Viable Ecosystem	23
3.2 Identification of those involved, roles, and results of their work	23
3.3 Features of the Minimum Viable Ecosystem (MVE)	25
3.4 Governance	26
4. EOSC BUSINESS MODEL: FINANCING THE EOSC	28
4.1 Business model	28
4.2 Governance, transparency and accountability	29
4.3 Funding model and payment mechanisms	29
5. RULES OF PARTICIPATION	33
5.1 Federating the existing infrastructures	33
5.2 Eligibility criteria for those involved	34
5.3 Participation according to the business model	35
5.4 Liability related to service provision	35
5.5 Data quality	36
5.6 Data security	36
6. THE ROAD AHEAD TO COMPLETION OF THE 2nd HLEG	37
6.1 2 nd EOSC summit 2018	37
6.2 Support to the Implementation Plan and roadmap for the EOSC	37
6.2 Endorsements & commitments from the coalition of doers	39
6.4 Long term challenges	40

LIST OF FIGURES

Figure 1 — Selection matrix for the MVE.....	25
Figure 2 — Three-layer governance model.....	26
Figure 3 — The 70+ signatories of the EOSC declaration.....	39

LIST OF TABLES

Table 1 — Analysis of those involved.....	25
Table 2 — Comparison among possible funding/revenue models.....	31

GLOSSARY

Item	Description
CoC	Code of Conduct
CSPs	Cloud Service Providers
EDI	European Data Infrastructure
eIDAS	electronic Identification, Authentication and trust Services
EIROforum	Combines the resources, facilities and expertise of its member organisations to support European science in reaching its full potential
EOSC	European Open Science Cloud
EOSC Portal	The term applied for the benefit of the current report to imply the EOSC is a common portal giving consolidated access to existing e-infrastructures
ERAC	European Research Area Committee
ERIC	European Research Infrastructure Consortium
ESFRI	European Strategy Forum on Research Infrastructures
EURODOC	The European council of Doctoral Candidates and Junior Researchers
FAIR	Findable, Accessible, Interoperable & Reusable
GO FAIR	A bottom-up international approach. for the practical implementation of the European Open Science Cloud (EOSC) as part of a global Internet of FAIR Data & Services
HLEG	High-Level Expert Group
IaaS	Infrastructure as a Service
ICANN	Internet Corporation for Assigned Names and Numbers
IETF	Internet Engineering Task Force
IG	Interest Group
IPR	Intellectual Property Rights
LERU	League of European Research Universities
MVE	Minimum Viable Ecosystem
OSPP	Open science policy platform
PaaS	Platform as a Service
PPP	Public Private Partnership
RDA	Research Data Alliance
SaaS	Software as a Service
SWD	Staff Working Document
WG	Working Group
YERUN	Young European Research Universities

FOREWORD BY COMMISSIONER CARLOS MOEDAS

One year ago, the European Commission published a declaration, inviting national governments, industry and the scientific community to participate in establishing the European Open Science Cloud – a trusted environment for sharing and analysing data from all publicly funded research.

The response to the declaration has been strong and positive, enabling good progress on the complex tasks facing us. We have just launched the first version of the Cloud's portal, the governance structure is in place and we are well on track to having the Cloud operational by 2020.

In all this work, we have benefitted extensively from the advice of high-level experts groups. I am therefore pleased to receive the recommendations laid out in this report and in the report 'Turning FAIR into reality'. They will help guide us when developing a Cloud that is open to all researchers, and which will function as a user-friendly, collaborative tool for data sharing and re-use.

The authors of the two reports touch upon a number of key issues for the Cloud. They discuss the definition of what constitutes a minimum viable research data ecosystem in Europe, its main rules of participation, governance framework, and possible financing models. They also look at how the Cloud can effectively interlink people, data, services and trainings, publications, projects and organisations. In addition, they present an action plan to make research data findable, accessible, interoperable and reusable (FAIR): attributes which are essential to extract the full scientific value from data resources and to unleash the potential for large-scale, machine-driven analysis.

Europe's decision to develop the European Open Science Cloud reflects the willingness to embrace change, but also to empower 1.7 million European researchers and 70 million professionals in science and technology. The ultimate goal is to achieve a fundamental transformation of the whole research lifecycle and to make it more credible with increased integrity, more efficient, collaborative and more responsive to societal challenges.

I am convinced that the Cloud will allow a new generation of scholars to find, combine and analyse data and discoveries in a way that supersedes anything we have ever seen before. It will accelerate the transition to Open Science and Open Innovation and bring science and research closer to societal needs.



A handwritten signature in black ink, appearing to read 'Carlos Moedas'.

Carlos Moedas,

Commissioner for Research, Science and Innovation.

PREFACE

‘Musical harmony is based on physical principles, while in cooking, ingredients must be weighed out with precision. At the same time, you have to be able to invent because if one follows the same recipe all the time, you never create anything new’.

Fabiola Gianotti

The work of any high-level expert group is always complex, as they are requested to harmonise, at times, very different points of view and especially considering the complexity and ambition of the EOSC vision, we found ourselves, at times, making choices on delivering a fair and objective analysis from multiple points of view and we have enjoyed finding a common vision together.

With this report we aim to mark a transition towards the practical implementation of the EOSC, based on the European Commission’s implementation roadmap, and to set the scene to the practical launch of the EOSC, by placing focus on the governance structure, rules of participation and business model options. The ideas presented in this report bring together, deliberate, and further expand various policy papers and recommendations contributing to the establishment of the EOSC that have been published by ongoing Horizon 2020 projects and national initiatives, European Commission reports, as well by the Commission FAIR data expert group and by the Open science policy platform.

We, (very probably) just like you, want all European researchers in science and technology, to reap the full benefits of data-driven science, for the benefit of society and the public and with due respect for their privacy. We want the unprecedented production of research data, in terms of quality, quantity and variety, to be accessible and usable in productive, ethical and user-friendly ways. We have learned over one and a half years that this ambition will only succeed if it is shared and inclusive, and if it is based on the accumulated knowledge and practices from all of our stakeholders.

A sentence, which we see also highlighted in this report and which we think truly does sum up the EOSC of tomorrow is that the EOSC intends to **‘interlink people, data, services and trainings, publications, projects and organisations across borders and scientific disciplines’**.

Our recommendations invite our stakeholders to go that step further and, through engagement to EOSC, get the research infrastructures, ESFRIs, ERICs, etc. to do what they could not do before and spell it out as part of their key performance indicators (KPIs) in their future projects. We have also done our best to listen to our stakeholders and weaved considerations made by them throughout this final report and into the recommendations it spells out.

We have found ourselves at the very heart of significant change with regards to European legislations including the EU general data protection regulation (GDPR), the EU copyright directive, as well as other initiatives such as the Coalition-s for the acceleration to full and immediate open access to scientific publications; these will all help the open science cloud fulfil its potential. By pushing the boat out further with our recommendations for how the EOSC portal should evolve based on the rules of participation that have been submitted to us through the open consultation platform over the summer of 2018, we have tried to capture the pain points, the challenges and the must-haves throughout. We do not want to reinvent the wheel, but we wish to capitalise on existing vehicles and tools that make up our strong scientific base and investments made in infrastructures.

We hope that the report will make an impact as part of the EOSC launch for the end of the year for all EU Member States and we are extremely grateful to have been given the opportunity to work on such a challenging yet stimulating process where we have been engaged, especially over the past 10 months, alongside relevant, and timely EU milestones to support the Digital Single Market which, going forward, will have direct implications on the EOSC. We know Europe has the skills, knowledge and capacity to drive its new, open science cloud, pragmatically forward and hope that with this report we can contribute to reaching this ambitious goal.

The group would like to extend its gratitude to everyone who actively contributed to it; there were many who did, and we hope that we have been true to their input – while assuming responsibility for choices made in composing this report.

The 2nd EOSC High-Level Expert Group (HLEG) [2017-2018]

EXECUTIVE SUMMARY

'Science is the captain and practice the soldiers'

Leonardo Da Vinci

With the staff working document (SWD) on the EOSC at its very heart, this final report picks up from the recommendations of the first EOSC HLEG report, the findings from the 2nd EOSC HLEG interim report with the aim of providing truly practical considerations and pointers for the timely implementation of the EOSC, based around the concept of a 'Minimal Viable Product' i.e. a product with just enough features to satisfy early customers, and to provide feedback for future product development. It also highlights the ground-breaking work taking place in Europe, guided by the increasingly prevalent principles of research data sharing.

The Commission presented its vision for the European Open Science Cloud (EOSC) in its April 2016 Communication on the 'European Cloud Initiative',¹ as a part of the digital single market strategy. This will be attained through policy action and financial support to integrate and consolidate e-infrastructure platforms, to federate existing research infrastructures and scientific clouds and to support the development of cloud-based services for open science. In close collaboration with Member States to connect the priority European research infrastructures to the EOSC, the Commission will also work towards an action plan for scientific data interoperability, including 'metadata', specifications and certification.

To help drive forward and implement the EOSC, the main thread of the report is to understand how the EOSC, as highlighted also in the foreword, can effectively interlink **people, data, services and training, publications, projects and organisations**. Formation of these bonds would not only produce valuable data, but also provide visibility and networking space, with obvious incentive mechanisms for the recognition of the work of scientists and the value of supporting infrastructures. The EOSC should be a user-friendly, collaborative tool for data sharing and reuse.

The group took stock of work that is underway in Europe. Existing use cases could work as **common service working models** illustrating what the European scientific community is already achieving. The latter, indeed, formed a sample of 12 **EOSC in practice stories** that accompanied the interim report.

As Vincent Cerf summed up recently: 'I hope our computer science community will find or invent ways to engage, using powerful computing, Artificial Intelligence (AI), machine learning, and other tools to enable better quality assessment of the **ocean of content contained in our growing online universe**'².

This, this group believes, is what the EOSC is trying to do and reflects the change in the way scientific research is carried out. With a twist: scientists are not alone to navigate in this ocean of data, but rely on an increasingly sophisticated set of tools, practices and networks that enable data sharing and reuse. As a central element of this report the 2nd EOSC HLEG has given a set of practical considerations and recommendations for implementation of the EOSC, with an eye to the mechanisms behind the EOSC portal, and considerations around skills, monitoring and policy in the areas of implementation, with subsequent revised engagement and steering recommendations.

The following recommendations are the result of a combination of extensive discussion exchanges amongst expert findings and the stakeholders cited at the beginning of the report, first-hand experience from some of the Science Demonstrators (SDs) involved in EOSC Pilot project³, results from the open consultation launched at the EOSC summit in June 2018, and from evidence gathered from the EOSC in practice stories.

Please notice that the recommendations are not ranked by importance, rather they are thematically clustered, and, within each recommendation group, they are listed in the order they have been discussed among the HLEG.

Implementation recommendations, the EOSC should:

1. **Serve all researchers and all research support units from all research domains** to do their research competitively.

1 COM (2016) 178 final

2 'Unintended Consequences' By Vinton G. Cerf Communications of the ACM, Vol. 61 No. 3, Page 7 doi: 10.1145/3184402

3 <http://eoscipilot.eu/science-demonstrators>

2. **Have KPIs and/or other metrics in the new INFRA-EOSC⁴ projects workplan** that respond to how precisely EOSC will benefit them and what they can do in EOSC that could not have been carried out previously.
3. Implement ‘whatever works’ to **increase the availability and volume of quality & user-friendly scientific information on-line** in ways that (i) maximise the range of communities served and (ii) serve the three fundamental values of open science, namely rights of use, right of modification and right of redistribution.
4. **Select standards and community-endorsed best practices** to ensure interoperability and composability of EOSC services and resources (separate for all elements of the ecosystem – data, data access services, software, etc. – but with a systemic perspective to develop a trustable ecosystem), and promote and enforce their adoption across research communities, leveraging existing international initiatives and collaborations.
5. Use **international fora**, such as the RDA, ICSU-WDS, GO-FAIR, Research Infrastructures, etc., as vehicles to support the implementation and adoption of standards/best practices.
6. Submit, as part of EOSC’s mission, a **landscape analysis carried out on a national level within the Member States** for assessing statistics and key assets around the composition and relevant clustering of the community of users.

For the EOSC Portal

7. Define **two sets of rules of participation relevant for providing a minimum EOSC Minimal Viable Ecosystem (MVE) life ecosystem**:
 - a. the first for data, service and infrastructure providers;
 - b. the second for users and produce a short list of licences, covering both, data and applications that would meet the needs of all;
8. The universal entry point to the EOSC should provide access to a **marketplace of efficient and effective services**, with lightweight integration (authentication and authorisation infrastructure, order management, etc.) where service providers can find service users and viceversa. Other entry points, e.g. through already well-established and well-functioning platforms, are also desirable and should be seen as a plus and should be integrated, avoiding fragmentation.
9. Involve **industry** in EOSC, utilising the **data and services market place** where research and industry can interact together. Industry could see their involvement by adding value services on research data of commercial relevance but may also consume scientific outputs (viz., software and data).
10. Promote the development of services as **independent, interoperable and exchangeable building blocks** to foster the future accreditation of innovative and/or efficient alternatives, connected and built upon a base of well-defined and stable core services, notably in areas that are science-agnostic and provide an opportunity to use common solutions to access, manage and process big data.
11. Promote the development of **open, sustainable, versioned, documented and energy consumption aware software** for all elements of the EOSC, with a goal to render maintenance possible at minimum cost and technical effort.
12. Ensure **EOSC services meet the user needs** and are effectively and efficiently delivered (e.g. carry out regular assessments and comparing EOSC to other similar publicly and commercially funded initiatives).
13. **Simplify early (beta) participation** in the EOSC by potential data providers, service providers, and infrastructure providers, by relaxing initial constraints without relaxing (i) quality standards for data and services, (ii) respect for EU & national regulations and laws, or (iii) the need for consistent identification and referencing of research entities and resources.
14. Define a process to **Access management to Services/Resources** during their lifecycle through a service portfolio management to be improved and fully adopted at the EOSC governance phase 1..
15. Ensure EOSC provides an **environment for co-development, testing and innovation**.

For Skills, Monitoring, Business Models & Policy

16. Build a workforce able to execute the vision of the EOSC by **ensuring data stewards, data and infrastructure technologists and scientific data experts who are trained and supported adequately**.

⁴ Include Research Infrastructures (RIs) including eInfrastructures, ESFRIs, ERICs

17. **Monitor data access and re-use** in the EOSC by potential data providers, service providers, and infrastructure providers, by relaxing initial constraints without relaxing (i) quality standards for data and services, (ii) respect for EU & national regulations and laws, or (iii) the need for consistent identification and referencing of research entities and resources.
18. Introduce an effective combination of different types of business models as well as direct financing and EOSC vouchers, with **all service providers developing a clear business model** and, in the EOSC rules of participation, a requirement for participation demonstrating sustainability.
19. Consider introducing **funding instruments in Horizon Europe** to minimise operational risks in completion of the EOSC implementation vision, such as an EOSC-PPP (Public Private Partnership), incentives for pan-European infrastructure development and **capacity building** programmes.
20. Define and enforce **data management policies** aligned with the requirements of EOSC.
21. Define an **EOSC Helpdesk** with online facilities to lower barriers to entry and ensure transparency and support to user engagement. Additionally, capacity building measures & services could be envisaged after 2020.

Engagement recommendations:

1. Create **career-enhancing incentives** for researchers who open the science that they produce (e.g. for researchers who lodge high quality, curated data in trusted repositories, share data services to their peers, or develop open software and services, and make the EOSC an attractive, efficient and useful – yet not imposed – portal to those incentives) and ground these incentives by transparent open science metrics and indicators.
2. Develop, both at Member State and also EU level, appropriate **engagement schemes** whereby publicly-funded research infrastructure providers, EU public sector repositories and research communities take part in the EOSC.
3. EOSC should take **national and international developments into account** in an inclusive federated approach and should be connectable to national and international frameworks by establishing mechanisms to ensure collaboration between national and European services and service providers as well as interoperability between existing infrastructures, avoiding overlap and duplication of services..
4. **Stimulate the ‘supply side’** of the EOSC, by ensuring creation of economic incentives for demonstrably viable research infrastructure providers to use and co-develop shared facilities and data repositories through the EOSC, in ways that make data and services accessible beyond their intended constituency (scientific community or geographical scope). This would be supported by the Commission through the European Data Infrastructure (EDI).
5. **Stimulate the ‘demand side’** of the EOSC, by ensuring establishment of dedicated funding for demonstrations of the EOSC at EU level that maximise return on investment in terms of generating use of EOSC by researchers and their infrastructure providers (e.g. EOSC in practice stories, cross-disciplinary success of EOSC).

Steering recommendations:

1. Base research support along all lifecycles of data and services around the **concept of trust**, by adopting different forms of technology (i.e.: block chain etc.).
2. Ensure that the working groups and the other advisory structures cover well for the executive board **the latest scientific and organisational trends and novel ideas** for the necessary decisions in those areas.
3. Harness **inputs** and support their activities in the context of EOSC implementation.
4. Ensure, both at pan-European and international level, that research communities pursue **advanced partnerships**, also supporting those grants and other incentives, so that EOSC’s progress is not dependent on the slowest but on the fastest movers. Lagging countries may gain access to research infrastructures and data which they did not have access before.
5. Guidelines and rules should be clearly separated into (i) domains for which **stability and trust** are important and (ii) domains for which progress must take place rapidly. The former should have rules and **instructions** that remain stable in time, whereas the latter should be run by **living documents**, facilitating innovation and

change while minimising negative effects of phase-outs. All guidelines and rules should be accounted for in the **establishment and progress of the working groups**.

6. Develop a strong synergy with the cybersecurity competence centres in Europe in 2019 and the Wise Information Security for e-Infrastructures (WISE) trust community to support the framework of a **shared security model** in EOSC.

1. INTRODUCTION

1.1 Objectives of the 2nd High Level Expert Group

The 2nd HLEG EOSC was set up to add value and complement the mission of existing Commission expert groups (e.g. the Horizon 2020 Commission expert group on Turning FAIR data into reality). It has been charged with overseeing the practical, technical implementation of the strategic policy recommendations of 1st HLEG on EOSC, which concluded its activity in February 2018, in the context of Horizon 2020 programming. Furthermore, the group has taken over the work of the cloud subgroup of the Open science policy platform (OSPP) to ensure continuity on the EOSC topic and to guarantee an impartial, level playing field for all stakeholders. Specifically, the HLEG EOSC will take the recommendations of the OSPP on cloud governance into account and help articulate the Commission's position at the EOSC summits in June 2017 and 2018.

The 2nd HLEG EOSC therefore zones in on two crucial aspects of the EOSC:

- » governance of the EOSC (including rules of participation and management as such), as well as
- » financing of the EOSC (including the development of novel ideas like the 'cloud coins and credits').

The group's goal is to consider these aspects and take a fresh look into the subject; it is composed of experts of varied backgrounds who act in their personal capacity and do not represent vested interests in research infrastructures, have a thorough understanding of the specificities of scientific research, fully understand the value of sharing and re-use of research data and/or have relevant knowledge on the value and functioning of scientific data/cloud infrastructures.

Additionally, the 2nd HLEG EOSC was asked to do the following:

- » Provide support with the Rules of Participation
- » Advise the Commission on actions from the implementation roadmap of the EOSC initiative (by end of 2017), notably its governance and financing aspects

Draft and take forward the stakeholders' agreed conclusions resulting from the EOSC summit held in June 2017. Collaborate with the project EOSCPilot, by providing views on the project's deliverables and by providing input.

1.2 Scope and structure of the document

This report aims to dig deeper into the EOSC roadmap and deliver insights to help build the EOSC beyond the SWD in light of the imminent MS plus EC mandated board due to be launched in November 2018.

The EOSC is a key enabler of continued European excellence in science⁵. The present document builds upon the work carried out by the First EOSC HLEG⁶. It states the objectives of this 2nd HLEG, and recommends a way forward to implement the EOSC, supplementing the SWD⁷, which provided the bedrock for this HLEG to build on. The 2nd EOSC HLEG has produced practical considerations and recommendations for implementation of the EOSC. Finally, the document also includes several practical examples selected from the EOSC 'coalition of doers'.

The approach taken in drafting this report has been that of a 'minimum viable ecosystem', suggesting the steps to practically implement the EOSC as the effective product of such an ecosystem. The report identifies those involved and roles in this MVE, based on the idea of a minimum viable product; it helps define the EOSC features and governance structures and review the legal context.

The document is structured as follows:

- » Section 1 gives an overall introduction on the EOSC HLEG;

⁵ The EOSC: <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

⁶ The 1st EOSC HLEG (High-Level Expert Group) – 'Realising the European Open Science Cloud': https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none

⁷ European Commission, 'Commission Staff Working Document' – March 14th, 2018, http://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf

- » The policy landscape for EOSC is briefly summarised in Section 2. As part of the work conducted by the 2nd HLEG EOSC, synergies among EOSC and the OSPP⁸ and the FAIR Data Expert group⁹ have been identified.
- » Section 3 tackles issues related to viability of the EOSC ecosystem.
- » Some options for the EOSC business models are provided in Section 4, with due consideration of their impact on the governance structure.
- » Rules of participation to the EOSC are dealt with in Section 5.
- » Section 6 looks at the road ahead from 2018, bridging the way into the next phase of EOSC development, also linking with stakeholders and defining the remaining challenges.

1.3 Building upon the 1st EOSC HLEG

The first report and recommendations of the Commission High Level Expert Group on the European Open Science Cloud 'Realising the European Open Science Cloud'¹⁰ were published on 11 October 2016. They emphasised that immediate action had to be taken to implement a federated, globally accessible environment – where researchers could publish, find and re-use data and tools for research. The 1st EOSC HLEG supported the Commission in outlining a general vision for the EOSC during the timeframe of the preparation and initial follow-up of the Cloud Initiative Communication.

In recognition of the significance of the EOSC declaration issued in 2017, which delivers key statements on data culture and FAIR data, research data services and architecture, together with governance and funding, the HLEG wish to add some recommendations to strengthen the EOSC in the areas of Policy, Governance, and Implementation.

In particular, from the **policy** point of view, some Member States have taken immediate affirmative action on the EOSC. Those qualifying as early movers developed an ERAC SWG OSI (standing working group on open science & innovation) providing their 'Draft Opinion on EOSC Governance models and Strategic Implementation Plan'¹¹.

A number of proposals, which closed in April 2018 and June 2018 under H2020, to define an EU framework for FAIR research data (the INFRA-EOSC-2018 calls)¹², to develop initial catalogue of datasets accessible via the EOSC and develop rules of participation in consultation with stakeholders have been issued. Preparation of a FAIR data action plan, currently under the remit of the FAIR Data Expert Group is expected by the end of 2018¹³.

As regards **governance**, EOSCPilot and the 2nd EOSC HLEG are helping provide guidance on the creation of an internationally effective governance, as suggested also by the SWD. Moreover, on the 'Amplify good practice' in the EOSC recommendation, the 2nd EOSC HLEG managed to collect and describe a number of 'EOSC in practice' stories, which are reported in found in the interim report produced by the 2nd EOSC HLEG: 'Prompting an EOSC in practice'¹⁴.

As to **implementation**: Rules of Participation are indeed being developed, for subsequent endorsement and implementation. Moreover, regarding the 'funding scheme' recommendation of the 1st HLEG, this final report describes several possible options for practical business models (see Sec. 4). Finally, to 'develop a concrete plan for the architecture of data interoperability in the EOSC', the EOSCPilot initiative is delivering a detailed architecture plan¹⁵.

8 OSPP: <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform>

9 The FAIR Data Group: <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetail&groupID=3464>

10 The 1st EOSC HLEG (High-Level Expert Group) – 'Realising the European Open Science Cloud': https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none

11 ERAC Secretariat, 'Assessment of the Amsterdam Call for Action', February 2018, <http://data.consilium.europa.eu/doc/document/ST-1202-2018-INIT/en/pdf>

12 <https://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/calls/h2020-infraeosc-2018-2020.html#c,topics=callIdentifier/t/H2020-INFRAEOSC-2018-2020/1/1/1/default-group&callStatus/t/Forthcoming/1/1/0/default-group&callStatus/t/Open/1/1/0/default-group&callStatus/t/Closed/1/1/0/default-group&+identifier/desc>

13 The FAIR Data Group: <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetail&groupID=3464>

14 'Prompting an EOSC In Practice Interim report https://ec.europa.eu/info/sites/info/files/conferences/eosc_summit_2018/prompting_an_eosc_in_practice_eosc_hleg_interim_report.pdf

15 EOSC Pilot, <http://eoscpilot.eu/>.

2. THE POLICY LANDSCAPE: FROM COMMUNICATION (COM/20160178final) TO THE STAFF WORKING DOCUMENT (SWD 2018 83final) Council Conclusions TO LAUNCHING OF THE EOSC

There is increasing agreement among Member States and major research infrastructures and e-infrastructure stakeholders that an integrated approach to research data infrastructures is needed, going beyond layers (computing, data storage, use) and national and disciplinary silos.

In recent years significant development in open science, including open access to research publications and open data, has taken place in Europe. E-infrastructure commons were proposed. The creation of an EOSC as a federated, globally accessible environment, where researchers, innovators, companies and the general public can publish, find and reuse each other's data and tools for research, innovation and educational purposes under well-defined and trusted conditions has received enthusiastic support from the scientific community. The set-up of the EOSC is one of the ambitions announced by the European Commission in the Communication on the European cloud initiative launched in April 2016. The EOSC vision is: 'to give Europe a global lead in scientific data infrastructures and to ensure that European scientists reap the full benefits of data-driven science'¹⁶. EOSC will operate as a trusted, open environment for the scientific community for storing, sharing and re-using scientific data and results, supported by high-capacity cloud solutions with super-computing capacity via an EDI. The EOSC summit and EOSC declaration serve as a bottom-up component.

On 14 March 2018, European Commission produced the 'Commission staff working document' (SWD)¹⁷, an agreed-upon, preliminary concept of the EOSC and forwarded it to the European Parliament and to the EU Council.

The main characteristics of the EOSC and its mission, as planned in the SWD, can be summarised as follows.

- » EOSC aims to provide members of Europe's research community with 'a virtual environment with free at the point of use, open and seamless services for storage, management, analysis and reuse of research data, across borders and scientific disciplines' European Cloud Initiative COM (2016) 178 final.
- » EOSC as a concept builds on the observation that the current rate of progress, in both practices and infrastructure build-up, for the generation, storage, and sharing of scientific data, is very high; yet progress and initiatives are highly fragmented, across several dimensions (practices, disciplines, countries). EOSC therefore has a main goal to federate existing initiatives and structures within a single, consolidated and seamless platform.
- » EOSC aims to achieve this through six action lines, comprising:
 - (i) provision of an architecture for federated infrastructures;
 - (ii) FAIR data management and tools to ensure data stewardship across borders and disciplines;
 - (iii) services designed from the user's perspective;
 - (iv) access mechanisms and interfaces;
 - (v) rules of participation for those involved, and;
 - (vi) a governance structure that can pilot EU leadership in data-driven science.

EOSC is to respond to the needs of the community; its construction therefore incorporates the participation of a 'coalition of doers' defined and assembled at meetings in June 2017 and subsequently in June 2018. EOSC is to federate existing and planned research data infrastructures, connecting them with a soft overlay and build upon existing large-scale EU scientific networks including ICANN, IETF, AIOTI, GÉANT and ELIXIR.

¹⁶ The EOSC: <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

¹⁷ European Commission, Commission staff working document, 14.3.2018, http://ec.europa.eu/research/openscience/pdf/swd_2018_83_f1_staff_working_paper_en.pdf

2.1 Impact of current regulation

EOSC activities will need to comply with the current EU legal framework and Member State laws and take any planned changes into account. The EC is aware of the importance of the free flow of non-personal data for the emergence of the data economy. It claims that 'Free flow of non-personal data are a pre-requisite for a competitive data economy within the Digital Single Market. To fully unleash the data economy benefits we need to ensure a free flow of data, allowing companies and public administrations to store and process non-personal data wherever they choose in the EU¹⁸.'

A proposal for regulation was defined¹⁹ in September 2017 'to ensure the free movement of data other than personal data within the Union' and 'laying down rules relating to data localisation requirements, the availability of data'. It states specifically that 'Location of data for storage or other processing within the Union shall not be restricted to the territory of a specific Member State, and storage or other processing in any other Member State shall not be prohibited or restricted, unless it is justified on grounds of public security.'

The regulation is limited to data 'other than personal data'. data in EOSC will conform with article 4²⁰ and article 157 of the GDPR²¹ (in force from May 2018 the Regulation (EU) 2016/679 of the European Parliament and of Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data) and guarantee ethical and responsible handling of personal data. In autumn 2017, the Commission conducted a public consultation reviewing the directive on the re-use of public sector information (PSI Directive). This Directive requires that public sector bodies make data re-usable by third parties but does not currently include research data in its scope. It is still not clear whether bringing research data into the scope of the PSI Directive would lead to more benefits, as opposed to continuing with other 'soft law' measures that are showing some results (e.g. the Open Data pilot under H2020).

Privacy or commercial sensitivity concerns could justify exclusion of the reuse of certain data. In the case of EOSC, the rules of participation will define the rights, obligations and accountability of those involved from EOSC including data producers, service providers, data/service users, vis-à-vis the applicable legal frameworks (e.g. GDPR, copyright rules, Data Security and Cybercrime, dispute resolution and redress mechanisms, e-commerce directive).

The above and other regulations, as well as advances made in cloud computing together with AI has also created uneasiness among some scholars who have misgivings²² that these developments challenge 'traditional legal principles and increase legal uncertainty of various rights protection in the information society'. They conclude, and this panel agrees, that legal certainty is of paramount importance if new technologies are to contribute to economic and societal growth.

18 <https://ec.europa.eu/digital-single-market/en/free-flow-non-personal-data>

19 http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=46830..

20 <http://www.privacy-regulation.eu/en/article-4-definitions-GDPR.htm> 'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'

21 <http://www.privacy-regulation.eu/en/recital-157-GDPR.htm> 'In order to facilitate scientific research, personal data can be processed for scientific research purposes, subject to appropriate conditions and safeguards set out in Union or Member State law'.

22 <http://www.kluwerlawonline.com/abstract.php?id=EULR2017029>

2.2 The European Open Science Cloud, FAIR & OSPP Synergies

The best way to identify synergies between the three advisory bodies of the European Commission OSPP, the FAIR data expert group, and the HLEG EOSC is to categorise them first. Useful aspects to take into consideration are the individual scope and time frame of each.

The OSPP paints the broader picture of **open science** as a goal and gives policy advice to the Commissioner encompassing dimensions such as reward systems, measuring quality and impact (altmetrics), changing business models for publishing, FAIR open data, EOSC, research integrity, citizen science and open education and skills.

The FAIR Data Expert Group, initiated by the European Commission, looks closely at the FAIR data principles and formulates recommendations on how to implement them. The focus here lies on cultural change, incentives and metrics, as well as the skills and capacity that need to be built to make **FAIR data**, and the services supporting them, a reality.

Both the OSPP and FAIR areas are going to facilitate development of relevant skills to support EOSC at large. More capacity building activities should be encouraged in other adjacent areas in order to build the workforce that EOSC needs to implement its vision.

Finally, the HLEG EOSC is expected to make recommendations on how to shape and implement the **EOSC** as a federated infrastructure for data-driven research, based on open standards and best practices that support the open creation, curation and dissemination of scholarly knowledge and FAIR scientific data. According to the European Commission, the EOSC aims to create a trusted environment to host and process research data in support of EU science in its global leading role²³.

While the OSPP has the mandate to support the development and implementation of open science policy in Europe, the European Commission sets the agenda with the proposal of the EOSC. The FAIR data expert group is one of the first groups to produce more detailed recommendations on how to translate the FAIR data principles into practical guidelines, based on ongoing effort and best practices in the EU Member States. The broader picture of the OSPP goes together with a long-term perspective, while the FAIR data expert group recommendations are to be seen from a mid-term perspective, as is the case with the work of the HLEG EOSC to build a federated infrastructure.

The EOSC will only be successful if the federated infrastructure is embedded in a cultural change, thus the work of the HLEG EOSC and the FAIR Data Expert Group is recommended to closely link the two together. Only when researchers value curated data as a first class research results, rather than sticking to publications (of results), will the full potential of the EOSC be unleashed. This is where the OSPP and the other two expert groups come in. The OSPP should support the European Commission and the Member States to provide guidance on open science principles. This is exactly what the OSPP did by publishing eight broad recommendations for the implementation of the EOSC that can be seen as guiding principles for the HLEG EOSC and the FAIR Data Expert Group²⁴.

Wisely, the OSPP embeds five assumptions related to the federated infrastructure in three principles addressing issues like awareness, skills development and ethics. For example, to optimally use the EOSC in the future, and to implement it accordingly, the skills set related to research data has to be improved in the scientific system to provide FAIR data services. Data administrators could provide mediation between science and the developers of infrastructure. Nevertheless, researchers will at least need some core competencies in data management and will ideally be supported by data stewards and data scientists.

Hence, the eight OSPP principles built the ideal framework for a balanced division of effort between the HLEG EOSC, which addresses the first five and the FAIR Data Expert Group, which focuses on the latter three. However, it is noteworthy that the somewhat top-down approach of these three (Commission initiated) expert groups and the Commission's work have to be supported and complemented by bottom-up processes. Some examples, (in a non-exhaustive list) are: RDA, the GO-FAIR initiative, H2020 implementation projects and coordination actions.

The Research Data Alliance (RDA) is precisely that, mainly a bottom-up initiative that focuses on research data management. It takes a long-term international approach to the topic and aims to provide guidance on research data standards that are helpful not only 'for some', but also 'in sum' for all research communities. While, the RDA helps underpin the work of the three different working groups related to research data management, effort is still required to support the infrastructural aspects and its federated approach. With the mid- to long-term scope of ESFRI and the various ERICs encompassed in RDA, the European landscape is well-prepared to

23 <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

24 https://ec.europa.eu/research/openscience/pdf/ospp_euro_open_science_cloud_report-.pdf

meet these needs. The discipline-specific ERICs could help to distinguish between the generic and research community-specific infrastructure needs while ESFRI provides the perfect forum to summarise the findings.

Beyond Europe, and pursuing the global impact envisioned for the EOSC by the Commission, infrastructure development should take into consideration how interoperability can be achieved between European service provisioning and other international initiatives operating in the same areas of development. An example in this sense is reflecting on how a global federation of federated services can be achieved. Of relevance for this discourse is work conducted at the USA National Institute of Standards and Technology (NIST²⁵) within the NIST Big Data Interoperability Framework (NBDIF) Architecture subgroup where standards for federated architectures of services have been developed.

Synchronising the European initiatives with global endeavours while maintaining momentum, requires a fast-track implementation process which harmonises both research data management and infrastructure needs. Therefore, to fully complete the picture with a relatively short-term, bottom up initiative, GO FAIR is worthy of mention. According to a joint position paper, Germany and the Netherlands, together with France, see GO FAIR as a fast-track implementation initiative to boost the further development of the EOSC²⁶. It aims to build on existing initiatives with critical mass, at the Member State level, such as the German National Research Data Infrastructure (NFDI) or the Dutch National Open Science Platform, among other science-driven initiatives, and to identify the early implementation needs of existing networks and consortia. GO FAIR adopts an implementation approach to FAIR research data and services. It builds upon the recommendations of the first HLEG on the EOSC with the objective to ‘federate the gems’ across the Member States²⁷. GO FAIR follows a bottom-up, open implementation strategy for the technical governance and funding needed to establish the first phase of the EOSC, as part of a broader global Internet of FAIR Data & Services. GO FAIR establishes Implementation Networks (IN) of research communities that will bolster and specify the principles and recommendations of the Commission’s expert groups from the bottom up. For instance, funding of data curation, preservation & exploitation, and analysis could be facilitated in order to match the exponentially growing needs of European Research Infrastructures for data driven research.

All activities mentioned above have a stronger focus on research data as opposed to services for research data management. For example, the H2020 implementation projects and coordination actions (e.g. eInfraCentral, EOSCPilot, EOSC-Hub, INFRA-EOSC-2018 calls, FREYA, OpenAire-Advance, RDA Europe 4 etc.) offer service platforms with services at different levels (generic vs community-specific) for different stakeholders involved in research data management. New calls in 2019 will cover prototyping new innovative services, support to the EOSC Governance, Enhancing the EOSC portal and connecting thematic clouds all to support the EOSC vision. Still, some challenges remain and should be addressed soon, as indicated in Sec. 1.6.4. A synthesis table comparing the relevant initiatives with differences and commonalities may be found below.

Projects	Comparable areas of development within the EOSC	Targeted stakeholder groups
AARC2 Start: May 2017 End: Apr 2019	» ICT technical specifications development » Interoperability and service architecture » Authentication and Authorisation infrastructure » Discipline agnostic approach NB:: Develop and pilot an integrated cross-discipline authentication and authorisation framework, building on existing authentication and authorisation infrastructures (AAs)	» Research performing organisations » Research Infrastructure and E-infrastructures » Service providers

25 <https://www.nist.gov/>

26 <https://www.dtls.nl/wp-content/uploads/2017/05/DE-NL-Joint-Paper-FINAL.pdf>

27 <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud-hleg>

Projects	Comparable areas of development within the EOSC	Targeted stakeholder groups
<p>eInfraCentral Start: Jan 2017 End: Jun 2019</p>	<ul style="list-style-type: none"> » EOSC service catalogue and marketplace » Policy development » Discipline agnostic approach <p>NB: Structure an open and guided discussion between e-infrastructures to consensually define a common catalogue for their services. Develop a single entry point (one-stop shop) - the eInfraCentral portal - for end users to browse the service catalogue, and enhance the monitoring of KPIs that focus on availability and QoS and user satisfaction. Draw policy and sustainability lessons for the future development of a European e-infrastructure 'market place' as an extension of the common service catalogue and portal so that it includes a much broader range of e-infrastructures and services.</p>	<ul style="list-style-type: none"> » Research performing organisations » Research Infrastructure and E-infrastructures » Service providers
<p>EOSCPilot Start: Jan 2017 End: Dec 2018</p>	<ul style="list-style-type: none"> » Interoperability and service architecture » Community building » Policy development » Skills & capabilities development » Discipline agnostic approach <p>NB: Propose and trial the governance framework for the EOSC and contribute to the development of European open science policy and best practice. 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. 15 SDs to pilot EOSC service provisioning from disciplinary and interdisciplinary perspectives</p>	<ul style="list-style-type: none"> » Research performing organisations » Scientific communities and individual researchers » Research Infrastructure and E-infrastructures » Service providers » Funding bodies » Enterprise » Policy makers
<p>EOSC Hub Start: Jan 2018 End: Dec 2020</p>	<ul style="list-style-type: none"> » EOSC service catalogue and marketplace » Unified discovery and access of services and resources » Service portfolio management » Interoperability and service architecture » Service provisioning » Data provisioning » Skills & capabilities development » Discipline agnostic approach 	<ul style="list-style-type: none"> » Research performing organisations » Scientific communities and individual researchers » Research Infrastructure and E-infrastructures » Service providers » Enterprise

Projects	Comparable areas of development within the EOSC	Targeted stakeholder groups
	<p>NB: Bring together multiple service providers to create the Hub. Provision of unified discover, access, use and reuse of resources for advanced data-driven research (services, data, training)</p>	
<p>FREYA Start: Jan 2018 End: Dec 2020</p>	<ul style="list-style-type: none"> » ICT technical specifications development » Interoperability and service architecture » Persistent identifiers » Discipline agnostic approach <p>NB: Extend the infrastructure for persistent identifiers (PIDs) as a core component of open research, in the EU and globally. The PID Graph connects and integrates PID systems, creating relationships across a network of PIDs and serving as a basis for new services. The PID forum promotes engagement with the global community via the RDA and through organising conferences, workshops and other PID-themed events. The <i>PID Commons</i> addresses the sustainability of the PID infrastructure resulting from FREYA beyond the lifetime of the project</p>	<ul style="list-style-type: none"> » Research performing organisations » Research Infrastructure and E-infrastructures » Service providers
<p>OpenAIRE-Advance</p>	<ul style="list-style-type: none"> » Unified discovery and access of services and resources » Service provisioning » Data provisioning » Community building » Policy recommendation » Skills & capabilities development » Discipline agnostic approach <p>NB: Provision of a range of service around research outputs (unified catalogue, metrics dashboard, open data). Community building and mobilisation around Scholarly Communication, Open Access, RDM. Unified discovery of scholarly outputs. Training on Open Access, RDM, policies. Leverages a network of National Nodes to connect with local and disciplinary communities.</p>	<ul style="list-style-type: none"> » Research performing organisations » Scientific communities and individual researchers » Research Infrastructure and E-infrastructures » Funding bodies » Policy makers

Projects	Comparable areas of development within the EOSC	Targeted stakeholder groups
RDA Europe 4	<ul style="list-style-type: none"> » ICT technical specifications development » Community building » Policy recommendation » Skills & capabilities development » Discipline agnostic approach (RDA WGs and IGs may focus on specific domain challenges) <p>NB:</p> <p>Mandated to become the centrepiece for an EU Open Science Strategy, providing skilled, voluntary resources from the EU investment to address Digital Single Market issues, also through a dedicated open cascading grant process.</p> <p>Support European participation in RDA activities through the cascading grant model, including the establishment of RDA National Nodes in Europe</p>	<ul style="list-style-type: none"> » Research performing organisations » Scientific communities and individual researchers » Research Infrastructure and E-infrastructures » Funding bodies » Policy makers
Domain cluster projects (H2020-INFRAEOSC-04-2018 call)	<ul style="list-style-type: none"> » Interoperability and service architecture » Service provisioning » Data provisioning » Community building » Policy development » Skills & capabilities development » Discipline specific approach <p>NB:</p> <p>Ensure the connection of the research infrastructures identified in the ESFRI Roadmap to the EOSC. Support will be provided through cluster projects gathering ESFRI projects and landmarks in each of the following large thematic domains: Biomedical Science, Environment and Earth Sciences, Physics and Analytical Facilities, Social Science and Humanities, Astronomy, Energy. Proposals will address the stewardship of data handled by the involved research infrastructures according to the FAIR[1] principles and in line with the objectives of open science.</p>	<ul style="list-style-type: none"> » Research performing organisations » Scientific communities and individual researchers » Funding bodies
Prototyping new innovative services (H2020-INFRAEOSC-04-2018 call)	<ul style="list-style-type: none"> » Service provisioning <p>Develop an agile, fit-for-purpose and sustainable service offering accessible through the EOSC-Hub that can satisfy the evolving needs of the scientific community by stimulating the design and prototyping of novel innovative digital services</p>	<ul style="list-style-type: none"> » Research performing organisations » Scientific communities and individual researchers » Research Infrastructure and E-infrastructures

3. MAKING EOSC A VIABLE ECOSYSTEM

The European Commission and the EU Member States need to put the required mechanisms in place for the EOSC coordinated effort to materialise, from a policy, technical and user perspective. At this initial stage, a significant step forward is an MVE. It would enable EOSC to emerge as a collaborative effort, in an iterative way.

This report addresses the aspects of implementation, engagement and steering, user engagement and proposed value proposition for them. It includes motivational mechanisms and levers to be used to convince users to change their practice and use the EOSC versus other potential options

3.1 The EOSC Minimum Viable Ecosystem

Providing resources to support the EOSC will take place in a very heterogeneous landscape of e-infrastructures and service providers, with dispersed users at best aggregated around disciplinary poles and national infrastructures. Addressing this challenge requires the definition of a smallest common denominator, referred to here as the **Minimum Viable Ecosystem (MVE)** of the **EOSC**.

An MVE will emerge and thrive only if some basic technical, political and human resources conditions are met. From a technical perspective, interoperable services and open data must be guaranteed. On a human resources level, a coordinated effort has to be made to put the right incentives in place for all of those involved (researchers, software developers and infrastructure managers, research managers) to design, contribute to and exploit the system. Policies are needed for the technical and human-oriented conditions to emerge. In modern science, recognition comes mostly from scientific outputs, and this is a fundamental factor to consider for the MVE to emerge. Policies have the greatest influence in setting incentive and obligations from publicly funded research. The MVE will possibly evolve according to the future expectations of EOSC, and the related success criteria.

3.2 Identification of those involved, roles, and results of their work

Those involved in the EOSC MVE, as outlined by the EOSCPilot (see for instance the EOSCPilot booklet²⁸), are largely:

- » European researchers structured in small or large teams based around institutions;
- » software developers: from individuals, to small to medium teams;
- » infrastructure managers;

The engagement of all of those involved heavily depends on the rules of participation. In addition to its explicit rules, one of its key functions is to create relationships of trust with users. We outline here the basic description from the point of view of their role in EOSC and explain how the RoP of the different actors will contribute to incentivize and generate the MVE.

EUROPEAN RESEARCHERS

The researchers' job is based on data and on computational resources. They need to produce or find data relevant to the inquiry, find an appropriate service or hosting for their own data and services, do the necessary transformations, run the analysis, publish the results and make data available to others.

Flexible ways to access and share data and direct access to fast networks to do so are at the top of the agenda for researchers. One of the main problems that the EOSC needs to solve is the fact that researchers in Europe still have insufficient access to e-infrastructures.

²⁸ EOSC Booklet <https://eoscpilot.eu/stakeholder-forum/eoscpilot-booklet>, November 2017.

Giving access to all the researchers in Europe to the Pan-European research infrastructures facilities in a straightforward way would provide sufficient incentive for scientists to adopt EOSC as a platform to access computing and storage services, or share their scientific data.

SOFTWARE DEVELOPERS/SERVICE PROVIDERS

Interoperable services and open data rely on the principles of software openness. The software used in EOSC services should guarantee interoperability and comply with standards, be they de facto or by right (de Jure). Data produced and handled with EOSC software services should respect the FAIR principles; services within EOSC should be secure and comply with the European authorisation and authentication policies; as a general policy, the software elements are provided upstream to open source projects, to guarantee the required level of sustainability; to provide persistent identifiers, identification scheme and machine-readable metadata about the resources.

For this key activity to be successful in terms of engaging human talent, breakthrough ideas leading to innovation need to be awarded with the proper recognition. Putting in place transparent mechanisms to recognise successful software development, such as creating an **'EOSC-Ready' certification for software** products, would have a positive impact on the software development ecosystem in Europe. The successful development of an *'EOSC-Ready'* branded software product, would improve the reputation of researchers and technologists and dynamically harness the potential of European developers, across academia and industry.

Software could have different levels of service management integration. Highly integrated services are operated according to the EOSC service management system. Medium integrated services run with a more mature service management framework. Low integrated run with a less mature service management framework.

Software sustainability needs to be addressed in the EOSC ecosystem as it is in any environment meant to sustain production work: proper versioning and documentation, as well as a clear sustainability path through community support, are among the desired features. Resource consumption optimization regarding storage, deployment or energy consumption are also factors worth considering when services are deployed on a large scale.

INFRASTRUCTURE MANAGERS

From the incentive point of view, infrastructure managers can perceive EOSC as a big opportunity. Firstly, integration of their resources in the EOSC is a way to achieve a higher, more efficient usage of those resources. The coordination effort should translate into economy of scale for infrastructures, but also into a potential expansion in their user base. EOSC may also create an incentive to upgrade and enlarge the capacities.

More importantly, at the level of increasing competitiveness, integrating in EOSC opens the opportunity to gather experience in running innovative services needed by cutting-edge research projects, supporting multi-disciplinary teams, etc.

Both managerially speaking and from a human point of view, the generation of expertise in deploying and running advanced services to support frontier research creates *know-how* in the resource centres. Often this type of service is only available in research infrastructures as prototype, long before it becomes commercially viable or profitable, if ever. When fed back into industry, in the form of trained people, it is this *know-how* that deliver the added value necessary for economic growth.

RESEARCH FUNDING ORGANIZATIONS

They are under pressure to deliver value for money, societal impacts and availability of publicly funded research. Their incentive is two-fold. (1) the EOSC lowers the cost of infrastructure for open science which has bearing both on political priority of openness as well as the need for efficiency. And (2) EOSC provides a downstream implementation opportunity of cloud and related technological research thus again increasing relevance of some specific topics of IT research.

OVERALL USER ANALYSIS

A preliminary schematic user analysis for EOSC is summarised in the table below

User / Provider	Actions	Story ending
End User	Register for use* Describe Data Discover service* Find data Transform data Run analysis* Store results Pay for service Sponsored to use a service	Evidence based on research accomplished, followed and cited
Service developer	Identify user needs Create access enabling services (e.g. marketplace, helpdesk, authorisation, identification, workflow, blockchain ...)* Create some research enabling services* Publish service* Provide consulting about service Charge for service	Investment into development of service returned
Research funding organisation	Identify user needs* Recognition opportunities Aggregation services	Acknowledgement of the EOSC as central reference in research funding themes
Core Infrastructure provider	Attract service hosting* Charge for hardware resource use	Well exploited, secure, interoperable and searchable infrastructure

Table 1 - Analysis of those involved

3.3 Features of the Minimum Viable Ecosystem (MVE)

The MVE consists of services listed noted with a star in the table above. In its minimal rendition they would comply with the 'low' level integration with the service management framework (as defined in Section 3.2) and satisfy the minimum compliance level. Access enabling services would satisfy the requirements about open metadata and persistent identifiers. These would be part of the 'EOSC federating core'. The selection has been made using the matrix, such as the one in the figure below. Note that the elements with a star are what is needed for bare-bones MVE and that it is quite likely that even from the onset other features may be offered as well with some services, for example helpdesk.

Further criteria – in addition to openness – to shape the MVE include sustainability of resources, documentation and energy consumption.

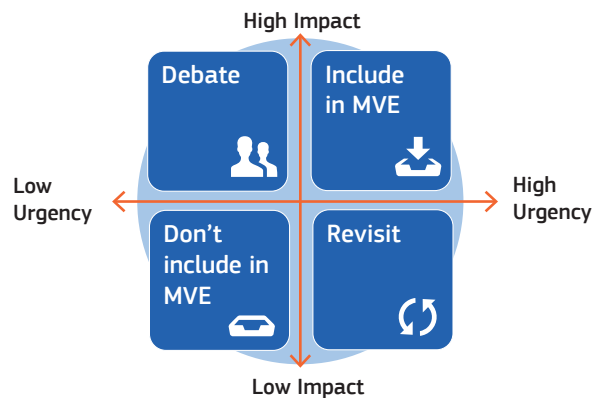


Figure 1 – Selection matrix for the MVE

3.4 Governance

EOSC is a concept of an important core scientific infrastructure for European science that will need to lead the future development of infrastructures far into the future. From a legal point of view, the EOSC vision could benefit from the involved stakeholders defining a legal entity to take forward a number of practical matters from handling finances to hiring dedicated staff etc. This could be facilitated for instance by the establishment of an EOSC PPP or a similar vehicle.

On an operational basis, the governance of EOSC should enable the MVE process. Furthermore, the proposed structure needs to be capable of running a fully-fledged EOSC. As mentioned above, the users have to play a central role in the design and implementation of the EOSC. Requirements, standards, operating procedures, etc., should be defined through close collaboration between:

- » the end user - scientific community;
- » the service providers – developers, intermediaries and operators;
- » funding agencies and scientific policy makers.

Rapid realisation of an EOSC governance is needed to move from vision to implementation. The two-stage approach for the implementation of EOSC has received general support, the first stage being the process of developing the EOSC and the second stage its management, operation and development.

In full conformance with what presides and the SWD, a three-layer governance model, based on the EOSC Declaration is proposed, as depicted in the following Figure. The three layers are:

1. Strategic Layer in the form of an EOSC board to combine state-of-the-art expertise on scientific cloud infrastructures with the Funders and Policy Makers. The Board will therefore include EU Member States and Associated Countries representatives. The EOSC board will mainly make strategic decisions on the development and evolution of the EOSC.

2. Executive Layer in the form of an executive board to manage day-to-day operation of the EOSC and procurers designing and planning work-related future developments. This, the only full-time staffed layer, will be supported by Working Groups, and will have the responsibility of ensuring that user needs are met and strategic requirements addressed.

3. Stakeholder layer organised in the form of a stakeholders forum to provide a medium for stakeholders: Users (Consumers), Providers and Intermediaries of EOSC Resources. This would have the main role to discuss, supervise and channel communication between the EOSC and the communities across all three layers. .

A **Coordination Support Action Structure** is providing additional support to the Executive layer.

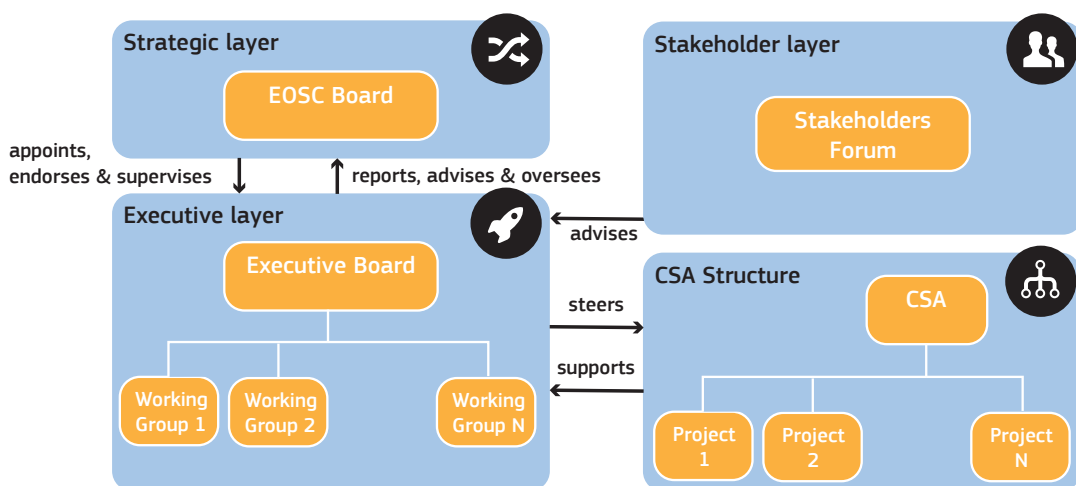


Figure 2 – Three-layer governance model

The three-layer model above provides a good basis for discussion and decision-making between Member States at Council level, enabling the EOSC to move forward. The structure must ensure an appropriate flow of information between the different governance layers. It should also accommodate the transparent sharing of responsibilities, as well as their supervision.

The **stakeholders forum** must grant all stakeholder groups the possibility to determine the requirements, policies and principles of participation. The stakeholders forum should on the one hand advise the executive board, while appraising the work of it on the other hand. The stakeholders forum must have an organisational structure that enables consumers, providers and intermediaries of all sizes to participate and it will be interacting with the boards.

A number of different **Working Groups (WGs)**, formed on the basis of interoperability contexts, the stakeholder roles, or broad scientific or infrastructure domains, could be created. The WGs should be time-based (e.g. up to 2-year duration) and should work on specific areas, whereby the priorities are set by the stakeholders forum, in conjunction with the strategic and executive layers. The forum must be open to discussing new activities, which may, in turn, lead to the creation of new groups. Some ideas already for potential Working Group identification which merit further analysis and discussion are:

- 1) The rules of participation WG
- 2) reference architecture WG
- 3) open standards in service development and seamless deployment WG
- 4) resource allocation WG
- 5) governance & legal structure WG
- 6) incentives & business models WG
- 7) FAIR principles over data & services WG
- 8) global scientific research WG
- 9) data management policies WG
- 10) quality management of data WG
- 11) data security & compliance WG
- 12) monitoring & indicators WG.

The **EOSC board** will involve Member State, Associated Countries and Commission representatives.. The EOSC board will define the strategic aim of the EOSC, review, agree and prioritise the EOSC proposals and requirements from the strategic vision point of view. The Board should monitor and assess the achievements of EOSC, approve the list of the executive board members and executive board work plan. The EOSC board has to ensure the coordination of Member States and Commission initiatives.

On 11th September 2018 the Commission launched an open call for applications for the selection of members of the Commission expert group - **executive board** of the EOSC reviewed by an identification committee. The expert group will assist the Board and Commission in the first phase of development of the EOSC until end of 2020 and will help to prepare the transition to the second phase of development of the EOSC. The group will be composed of up to 11 members and will include up to three independent experts - programme managers with international experience, good knowledge of institutional framework of research and innovation (R&I) funding and technical knowledge of data infrastructures — appointed in a personal capacity and up to eight pan-European organisations of R&I stakeholders most relevant for the EOSC implementation such as the large pan-European research infrastructures (RIs) including e-Infrastructures, public research organisations ('PROs'), universities, public research funding organisations and industry organisations. It is vital that all voices are heard and that the executive board will include not only representatives with skills of infrastructure(s) but also active and leading scientists, and a good representative pool of users to guarantee a user centred approach.

The executive board would provide the EOSC board with background and options for the way forward and propose its work plan for approval. To support the implementation of the work plan the executive board could work with stakeholders or set up ad hoc WGs. The executive board could then select WG members from the stakeholders forum or map its WGs to stakeholder groups.

During the first phase of the EOSC, the executive board will receive support from the CSA structure (effort coordinated through a specific H2020 CSA project). The CSA project(s), funded by Horizon 2020, will help the executive board coordinate all relevant European Commission-funded projects and support the implementation of the work plan while also acting as a secretariat to deliver the EOSC main functionalities. Online services such as an EOSC user helpdesk that continue post 2020, as well as e-learning facilities for capacity building, could prove beneficial to expand the user base and lower barriers to entry for transparency.

EOSC should embrace those who have produced something valid should have the opportunity to invite others to share what has been produced, to allow lower barriers to entry.

4 EOSC BUSINESS MODEL: FINANCING THE EOSC

As highlighted in the Draft Council conclusions of the EOSC May 2018²⁹ this chapter provides some considerations around developing a sustainable business model. The council invites the Commission to elaborate, in consultation with the Member States, on the future financing of the EOSC. This section sets out the business model, considers the Governance implications, gives examples of different business models; a hybrid business model that uses a mix of more conventional grant/contract acquisitions and a ‘cloud coin’-like environment.

Conventional grants and contracts could be a logical way to manage the fixed costs of an EOSC, for providing resources to academic institutions and the private sector, and to also include models such as pre-commercial procurement for the development of new services for the development of capabilities or making large public data sets available on widely accessible infrastructure. Cloud coins are most useful for managing the marginal (or operating) costs of the EOSC, providing individual investigators with the ability to effectively move among the variety of resources that constitute the EOSC. They are also convenient when it comes to monitoring.

4.1 Business model

The EOSC business model is a critical non-technical element that will determine the success of the EOSC vision. Given the dispersed nature of scientific research and the variety of tools and processes required by scientists in different fields and locations, a federated environment requires a similarly decentralised business model to support the technical environment that will be developed. Such a business model must do the following:

1. Support the rapid acquisition and delivery of a variety of cloud services and other technologies needed by investigators.
2. Allow International, National and private funding entities to maintain key electronic data and software resources for the benefit of the scientific community.
3. Enable reuse *in situ* of high value scientific digital objects, by which we mean data, software, metadata, workflows and other digital artefacts of scientific research.
4. Ensure sufficient interoperability to consent movement of digital objects between environments for reuse elsewhere, where unique capabilities exist elsewhere.
5. Maintain appropriate requirements on providers (of both digital objects and services) so that it is simpler for investigators to utilise FAIR principles for digital objects.
6. Limit the scope of the federation of scientific clouds to a reasonable number to help ensure that critical masses of digital objects exist in locations where they can be aggregated and reused.
7. Ensure that the private sector re-invests in R&D to stimulate innovation and create new markets³⁰.
8. Provide the necessary human support for the technology delivered to scientific users.

The current model for provisioning access to RIs is based on the guidelines contained in the Charter for Access, where three main models are described:

A. Excellence-Driven Access: exclusively dependent on the scientific excellence, originality, quality and technical and ethical feasibility of an application, evaluated through peer review conducted by internal or external experts. This enables Users to gain access to the best facilities, resources and services wherever located. This mode enables collaborative research and technological development efforts across geographical and disciplinary boundaries.

B. Market-Driven Access: is defined through an agreement between the User and the e-Infrastructure that will lead to a fee for the Access and that may remain confidential. This is on a grant-based option.

29 <https://ec.europa.eu/digital-single-market/en/blogposts/european-open-science-cloud-council-conclusions>

30 Source: (<https://marianamazucato.com/uncategorized/mariana-mazzucato-appointed-as-special-advisor-for-mission-driven-science-and-innovation-to-eu-commissioner-for-research-carlos-moedas>)

C. Wide Access mode: guarantees the broadest possible gateway to scientific data and digital services provided by the e-Infrastructure to Users, wherever they are based. Adopting this mode maximises availability and visibility of the data and services provided.

In practical terms, a model based on the Wide Access mode modulated by a negotiated, agreeable Access restriction, is the pragmatic way to start moving with the EOSC. Private providers willing to provide resources within the EOSC framework will envision a Market-Driven approach to support users.

Considering the approaching transition between H2020 and Horizon Europe a number of instruments should be considered, besides what has already been planned, to ensure a continuous support of resources to the implementation phase, to reduce the risk of interruptions in completing the operational vision of EOSC. For instance, an EOSC-PPP (Public Private Partnership) could be envisioned; additionally, incentives for pan-European infrastructure development programmes could be designed as well as capacity building programmes for people's competences, (à la COST³¹ Actions, Marie-Curie³² Programme, etc).

4.2 Governance, transparency and accountability

The EOSC board will have the ultimate responsibility for defining the business model and the requirements for participation, as well as oversee that EOSC service providers operate according to the rules of participation. The key is complete transparency of the marketplace for cloud services for all stakeholders: CSPs, academic providers and users of the EOSC.

Ensuring transparency and accountability are essential elements of any business model and must be enforced by the governance framework. Transparency dictates that all participants and providers in the EOSC have the same basic access rights to the marketplace and that actual costs are clear to end users, so that they can make appropriate decisions on the use of the various EOSC capabilities. Transparency also demands open interfaces³³ that are specified for all users. The EOSC executive board should provide an annual report that provides insights into the relative use of various CSPs and services.

Any processing of personal data has to comply with the EU General Data Protection Regulation (GDPR) and has to follow the principles relating to the processing of personal data. The Regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data. Compliance to GDPR also includes to follow 'Data protection by design and by default', thus meaning that whenever personal data is handled it must be design and built with consideration of the principles and provide safeguards to protect data (for example, using pseudonymization or full anonymization where appropriate). For accountability reasons, the data controller might need to demonstrate compliance with the principles laid down in the regulations.

4.3 Funding model and payment mechanisms

Regarding funding models, or how money transits from public and private coffers to ensure that data is open (opened up) and shared (and reused), a payment model is how this money actually is distributed in the ecosystem, e.g. how can it be transferred to ultimate data and service providers. Several possibilities need to be envisaged in this heterogeneous landscape.

The most obvious model involves member states and multistate entities (e.g. the EC, EMBL, etc.) directly supporting elements of the EOSC that exist locally or in areas of scientific expertise, via institutional funding, direct grants and contract funding.

In order for such support to create an EOSC MVE, rather than simply another set of digital silos, recipients of support would need to meet a set of technical and operational standards that would ensure that these resources (storage, computing, and higher order services such as software, pre-defined workflows, etc.) are accessible to scientists outside of the host institution and across Member States. Computing assets and RIs

³¹ European Cooperation in Science and Technology

³² <https://ec.europa.eu/research/mariecurieactions/>

³³ Note that open interface is distinct from an open source interface. An open interface has a public defined specification but does not necessarily have publicly available source code. An open interface can therefore be either open or closed source

such as ESRF³⁴ in Europe and at many US National Laboratories have been successfully funded using similar models. In those cases, the institute receives a grant from the funding entity (the US Department of Energy in the latter) to build/operate the resource and make it available to other grantees of the funding entity. These grantees apply for capacity at the resource, which is tasked to identify high value research projects, and indeed, the resource operator is evaluated on the quality of the science that it provides resources to. The limitations of such a payment model is that the amount of resources available is by definition pre-determined; that is, there is often little ability to increase or decrease capacity in the event that the resource is incorrectly sized. In addition, restrictions placed by law or policy at funding agencies can restrict the ability of certain researchers to access these resources – in the example cited above, non-Department of Energy grantees cannot get access to the resource.

Another option is a funding and payment model that is based on a certification programme for commercial and non-commercial providers of computing services that can provide scientifically useful services, that will meet EOSC-defined standards to ensure minimum levels of access and interoperability and that will accept specific, EOSC-defined financial transactions in payment for these services ('cloud coins', see below). Ideally, this will create a competitive marketplace for these services that could, with appropriate governance, become an implemented EOSC. The basic outline of such a model for the EOSC is relatively straightforward.

The EOSC board defines a series of **minimum requirements for commercial or non-commercial entities** that wish to participate in the EOSC as service providers. These requirements have been preliminarily discussed in the EOSCPilot 'Principles of Engagement' draft document³⁵ and shall be further specified in the upcoming months by the various EOSC initiatives. The EOSC should be designed to maximise the ability of data and other digital objects to meet FAIR criteria. At a minimum, these requirements would likely include compute/storage/network capacity, accessibility, interfaces (these could be as simple as basic upload/download/execute commands or value-added Application Programming Interfaces), Identifiers and Metadata, Authentication/Authorisation, Information Assurance, compatibility with various regulatory requirements where necessary³⁶ and a willingness to accept euro-denominated EOSC vouchers, distributed to investigators through grants, or supplementary awards, that can be used to purchase services from these compliant providers. This model will dictate that a series of providers of services will have to appeal to thousands of individual scientists and research groups, incentivising competition among them, which should in turn yield better service levels at lower prices. To coordinate acquirement, the EOSC and member states would also certify one or more brokers to manage the acquisition, distribution and payment for EOSC vouchers. These brokers could be government agencies in member states, entities within member states, transnational governments or private firms.

Such a model would have several advantages. First, it incentivises both providers and investigators to converge to highest value services at the lowest possible price, and for providers to compete to develop new services of interest to the scientific community. Secondly, by requiring that providers meet certain requirements for broad access to the resources being provided and using relatively standard transaction types, this model promotes easy reuse of data and other digital objects by other investigators. To the extent that is appropriate, cloud technology is utilised such that individual resources can be scaled up or down according to usage. If the transaction method is created appropriately, it provides funding agencies with a unique level of insight into the utilisation of various types of scientific computing and particular data and software assets via the pre-existing reporting capabilities of the global financial network. Finally, such a model simplifies acquisition by a variety of categories of entity. As long as the brokers have appropriate agreements in force with the full provider network (which should be mandatory to become a broker), the various Member States can acquire, distribute and pay for cloud coins under their own acquisition regulations, with no additional complexity burdening the individual investigator.

The National Institutes of Health in the United States has conducted a cloud credit experiment. Eight service providers (e.g. Infrastructure, Platform, Software), became conformant providers. Approximately \$3m were distributed as cloud credits. This project will complete at the end of the 2018 Fiscal Year.

How the transaction is carried out, e.g. the payment model proper, is largely up to the entity funding the EOSC vouchers. Options could include pre-purchase of services, escrow of funds with investigators given withdrawal rights up to certain amounts, or distribution of credits via a pre-paid debit system. In

34 <http://www.esrf.eu/>

35 'Principles of Engagement' EOSCPilot, 2018 - https://docs.google.com/document/d/1JBd30DMG_KhG8vIh1rOIM4VpYFjOnk4sS59whjalEI/edit

36 Potentially optional in certain contexts.

addition, in this model, funding agencies or participating states could choose to directly fund the costs of maintaining/archiving key scientific data sets or other electronic resources in environments that scientists regularly select for research purposes. Moreover, this model provides for a dispute resolution scheme. For example, if a user pays with EOSC vouchers but does not obtain the promise QoS (i.e. the service is 'down' for a period of time in breach of the Service Level Agreement) a mitigation is submitted.

The question remains as to the management of a coordination centre/gateway/marketplace. Such a site(s) would be helpful to assist in the integration of the disparate pieces of the EOSC federated infrastructure. Several sources of revenue could be envisaged to support the management of these resources. One possibility could be to require conformant commercial service providers to pay an annual subscription fee to offer their services in the marketplace. Another option could be a transaction tax on vouchers (i.e. a service provider is taxed a small percentage on each voucher that is redeemed against their services), perhaps combined with a transaction tax on private sector users. The investment recovered could then **contribute to funding the operational costs of the EOSC Gateway/ marketplace** and providing key open data sets. Ideally, the marketplace would keep track of how frequently a dataset is used and the provider of that data set compensated accordingly, similarly to how YouTube pays people who upload videos based on how many times they are viewed.

The following table provides a pro vs con - based comparison of the three major options: the conventional 'direct support' funding and payment model, a pure EOSC vouchers model and a hybrid model where some resources are supported via direct funding and other resources are accessed with cloud coins.

Model	PROS	CONS
Direct Support: Elements of EOSC receive direct payments from funding agencies	<ol style="list-style-type: none"> 1. Simple to operate and implement, based on current funding mechanisms 2. Cost effective and tested mechanism to maintain core data assets for broad use 3. Provides EOSC nodes with core revenue to develop services that might not be economical otherwise 	<ol style="list-style-type: none"> 1. Frequent capacity mismatch (either too large or too small) leading to cost inefficiencies or access issues 2. Resources can have internal focuses, reducing access from outside stakeholders 3. Constant funding stream reduces the incentive to innovate to attract users and ensure adequate support 4. Burdensome for commercial entities, even where they could provide significant cost savings and be incentivised to innovate.
Pure EOSC vouchers: Researchers use 'cloud coins' to support EOSC services/data	<ol style="list-style-type: none"> 1. Enables maximum choice for researchers by operating in a competitive marketplace for services 2. Enforces innovation by requiring services to support themselves via 'cloud coin'-based cost recovery 3. Provides simplified access for commercial providers to the marketplace 4. Expected to simplify acquisition of 'just enough' services and place significant pressure on costs leading to a migration of data management costs from infrastructure to ensuring FAIR principles 	<ol style="list-style-type: none"> 1. Much more complicated to implement, requiring the creation of significant trust and payment management networks. 2. More difficult for non-commercial organisations to directly interface with the EOSC, due to a lack of venture capital in such environments. 3. Poor model for maintenance of crucial public data sets – would require that end users of data to 'vote' with credits to maintain data. 4. External focus (to attract investment via 'cloud coin' users could reduce the ability of these resources to support internal stakeholders)

Model	PROS	CONS
Hybrid: EOSC vouchers and direct support	<ol style="list-style-type: none"> 1. Provides necessary resources for the management of well defined, but extremely high value data sets. 2. Supports access by both commercial and non-commercial organisations into market, allowing both groups to innovate in areas where they have specialised abilities 3 Provides clear opportunities for innovation and rapid development via the creation of a competitive market in scientific information technology services. 4. Provides maximum choice to researchers, who can utilise multiple categories of resources 5. Expected to simplify acquisition of 'just enough' services, thereby ensuring needed capacity and reducing expenditure on unnecessary services. 	<ol style="list-style-type: none"> 1. Increased complexity due to the nature of a mixed market. 2. Implementing the 'cloud coin' market will remain complex but can be built organically even as other resources utilise more conventional funding.

Table 2 – Comparison among possible funding/revenue models

5. RULES OF PARTICIPATION

As its bottom-line, the EOSC aims to support three objectives: (1) to increase value of scientific data assets by making them easily available to a greater number of researchers, across disciplines (interdisciplinarity) and borders (EU added value) and (2) to reduce the costs of scientific data management, while (3) ensuring adequate protection of information/personal data according to applicable EU rules (e.g. REGULATION (EU) 2016/679). Indeed, cost reduction is a key element to ensure that resources remain available to carry out the first objective without requiring reductions in the resources necessary to carry out cutting-edge research that will generate the next generation of scientific data. Moreover, an EOSC needs to be developed with the fundamental principles of data protection by design and information security in mind. Further ‘ancillary’ objectives of the EOSC are inclusiveness, bearing a clear value proposition that is simple to adopt, among other attributes

The idea of creating an EOSC PPP for the future could consider tackling the issue of inclusiveness and assessing the role of commercial entities offering paid services which may run the risk of excluding researchers/universities/institutes that may face economic/funding issues.

A key part of the governance framework is a set of rules of participation that support these objectives by setting the rights and responsibilities of participants to the EOSC. Collectively, they will help ensure that the objectives of the EOSC, described above, are met.

In relation to the rules of participation, the FAIR principles can be mapped into different parts of the EOSC MVE. **Findability** may be a function that is best dealt with centrally, by the ‘core’ of the EOSC. Data producers are responsible for the adoption of a globally unique and persistent identifier assigned to their (meta)data. For harvesting machine-readable metadata and for building up a central index in the MVE, data providers can facilitate findability via the adoption of standardised interfaces to their data repositories.

If (meta)data are not open and free, **accessibility** may be best embedded in the EOSC MVE via the integration of an existing, generic (i.e. independent from any discipline) authentication and authorisation infrastructure. If for data protection issues, e.g. for highly sensitive data, access to data cannot be fully automated, this infrastructure might offer services to connect the data scientists with the data repository to negotiate the access to the data.

The EOSC MVE may best support **interoperability** via services which operate with standardised vocabularies (e.g. thesauri) or standardised knowledge representation models (e.g. for an ontology) used to formally describe the (meta)data in a machine-understandable way.

Re-Usability of (meta)data, is a function of the EOSC MVE which is of utmost importance to unfold the full potential of EOSC, particularly on reproducibility. Services of the EOSC MVE may support data producers to provide best descriptions of their (meta)data so that they can be replicated and/or combined in different settings.

5.1 Federating the existing infrastructures

A key element of an EOSC MVE is the incorporation of resources that have been widely used, that have demonstrated value for the scientific community and that provide suitable guarantees in terms of data protection and information security. It is impossible to imagine a reduction of costs if existing capabilities are not included or are re-developed as part of an EOSC. That is not to say that the federation must be limited to those already involved. Robust ecosystems require a constant influx of new resources and, likewise, the departure of those that are no longer widely useful. Existing providers who wish that some of their offering become part of the EOSC will need to follow requirements as per the eligibility criteria for those involved; however, their presence will be essential to creating the critical mass of data and services needed to drive users into the EOSC.

How simple or laborious it will be for existing entities to join the EOSC will depend on a series of choices made during its implementation. For example, if it is determined that a service provider has to provide computing infrastructure in addition to data, sites that are primarily data repositories would most probably have to invest in additional infrastructure or migrate to a commercial cloud infrastructure in order to participate. Similarly, rules and requirements around Application Programming Interface (API) access and information assurance, particularly liability for data breach, might restrict the participation of a certain type of entity in the EOSC due to the costs associated with API development, identity and access management or security audit. One option is to phase-in

certain parts of the eligibility criteria, to enable changes to take place as part of a normal system development life cycle, or to provide funding for conversion of certain high value assets to more scalable infrastructures.

Additionally, the EOSC should serve all possible users and not just academic users. Private sector users should be considered stakeholders in the EOSC as well as participants from the start, not added after the fact. By participating, private sector entities may want to invest in the long-term development and sustainability of the EOSC, along with the public sector and not just serve to exploit public data for free. Considerations such as access for those involved in a non-scientific capacity as well as non-EU/associated countries researchers are not addressed here. The other category of participants are the funding entities and the brokers utilised by the funding entity to distribute cloud coins. For both of these groups, it is essential to ensure that they follow practices that enforce fair market conditions. For funders, this means treating all providers as equal partners and allowing the users of the EOSC maximum discretion to select resources. Brokers would be obliged to behave in a disinterested fashion with all providers. Entities that establish brokers must require that the broker does not establish a monopoly or fall under the control of a service provider that then uses its influence to exclude other service providers from the marketplace.

5.2 Eligibility criteria for those involved

There are two sets of eligibility criteria relevant to delivering an EOSC MVE: the first for data and service providers, the second for users. This report focuses primarily on users. These criteria need to meet two objectives: ensure that the data respect the policy goals, specifically FAIR guidelines and information assurance/personal data protection guidelines/regulations, and further, create an EOSC in practice that is appropriately sized and defined to increase data sharing and reuse.

Key rules for participants therefore will include:

- 1. Capacity:** The ability to reuse data *in situ* (as opposed to making additional copies of data), and the need to create an appropriately sized EOSC suggests that participating services must possess an appropriate amount of computational, storage and network capacity available to external users. The exact minima may depend on the nature of the node, the types of analysis likely to be required and the size(s) of data sets likely to be deposited. Such capacity needs to be available to external users, capacity that is only available to local users (for example, members of a specific university) should not be counted. Such capacity could be made available by implementing nodes in commercial or non-commercial clouds or by providing access to dedicated computing resources associated with a particular node in the environment.
- 2. Accessibility:** In line with work underway by the FAIR Data experts Group, a key element of FAIR principles is that data must be accessible; hence participating entities will need to have appropriate mechanisms to ensure that this is so. This will likely include a minimum set of interfaces for data deposit and download, as well as capabilities to launch analytic tools against data deposited at the site. These requirements should not require complex APIs or simple Graphical User Interfaces (GUIs), although providers could certainly choose to provide them. Over time, an EOSC MVE may deliver such capabilities, but the initial requirements should be more basic; premised on ensuring that there are understood technical means of interaction with the resource. Further, these requirements are not an assertion that all data and services must be directly available for all at all times: several categories of resource will be subject to data, and the resource must provide means to implement relevant restrictions on access and reuse.
- 3. Identifiers and Metadata:** The ability to find, interoperate and reuse scientific content is dependent on the ability to understand the data, software or workflow that is being evaluated for reuse. Maintenance of this metadata is fundamentally the responsibility of the submitter of data or other digital objects. Resources must cooperate with such indexing capabilities (which themselves are likely to be part of the EOSC) which are ultimately developed under the auspices of the programme.
- 4. Information assurance and data protection by design:** Resources will need to support appropriate information assurance activities in cooperation with providers of content. Given that liability for improper access will probably reside primarily with the user/organisation that submits the data or that maintains access controls, a shared security model will be essential. Moreover, compliance with the fundamental principle of data protection by design will be needed. In this respect, appropriate technical and organisational measures, designed to implement data-protection principles (such as data minimisation and protection of data subjects' rights) should be applied in an effective manner to integrate the necessary safeguards in the processing activities.

Participants that access the EOSC to deposit or use data will mainly need to agree to the RoP of the use of the data or other digital objects, and to FAIR principles. In practice this likely means that data that are deposited in the EOSC must meet minimum metadata including appropriate identifiers and links to information assurance descriptions. These requirements will vary depending on the scientific discipline and will need to be defined by experts in that field. The role of nodes in the EOSC will also need to be clarified– should the node that receives the data pass metadata to indices or should it be delivered directly by the depositor? Efficiency would seem to demand for the former, but this remains an open question.

5.3 Participation according to the business model

The development of novel capabilities, long-term storage/maintenance of data resources and fixed cost capabilities are likely to be provided using direct payments to organisations setting up nodes in the EOSC. By contrast, numerous research activities by individual investigators may be supported via EOSC vouchers. Nodes in the EOSC will have to be able to engage with the business model. This will probably imply a business arrangement with the brokers set up by funding agencies in order to accept these vouchers as payment.

One important question about participation by providers of services is whether certification should be based on a self-certification model or certification by a third party. Self-certification has obvious advantages in that it generally reduces costs and hence barriers to entry as compared to third-party certification. During the early phases of an EOSC, when it will be difficult for providers to assess the size of the market, a third-party system could prevent many smaller or newer (but potentially technically innovative) organizations from choosing to participate. However, third party certification provides greater assurance that providers meet the standards required for effective participation in a broad EOSC that properly implements FAIR principles, as compared to self-certification where ‘word of mouth’ is the only effective deterrent against unscrupulous actors. Given the above, the HLEG recommends that the EOSC begin with a reviewed self-certification; that is, that organizations perform a self-certification and submit the results to a body (perhaps under contract to the Executive Committee) that reviews the assertions made by the provider. This, combined with a forum for users of services to share their experiences, should provide sufficient certainty for the early stages of the EOSC. Once the EOSC has achieved a critical mass and it is possible to perform market research for EOSC services, certification can move to a more formal third-party process.

All of the above should always keep in mind the importance of retaining transparency and lowering barriers of entry to ensure inclusive engagement by all.

5.4 Liability related to service provision

The rules of participation will need to set general principles and define minimum standard procedures with respect to possible liabilities resulting from the provision and the use of EOSC services.

To minimise such liabilities, the EOSC MVE should support identity management and access controls on digital objects that enable appropriate access to data.

Providers should be liable for failures in areas under their responsibility, such as compliance with data-protection rules, data security, security overall and employees with disproportionate access rights. As the submitters control access, they retain liability for data leakage and to ensure that relevant individuals accessing information meet the necessary requirements.

As regards to data quality and warranties as to fitness for purpose, the EOSC MVE would need to operate under the principle of let the buyer beware (*caveat emptor*). That is, while submitters may be liable for outright fraudulent data, the nature of scientific research data determines that EOSC data should probably be provided with no warranties for any particular purpose, (although Section 5.5, on assessing data quality, should be also taken into consideration). However, submitters should provide statements of quality – not all data can or should be of the highest quality, but it should be clear to users, what the quality of the data are determined to be by the provider. This may be an element related to the FAIRdata metrics efforts as they are being developed.

SWD

Data should be:

- » processed lawfully, fairly and in a transparent manner in relation to the data subject (principle of ‘lawfulness, fairness and transparency’);
- » collected for specified, explicit and legitimate purposes;
- » adequate, relevant and limited to what is strictly necessary in relation to the purposes for which it is processed;
- » accurate and, where necessary, kept up to date.

5.5 Data quality

Data quality is likely to be the most difficult element to standardize in any given set of rules of participation, considering that the usual standard of ‘fit for purpose’ varies so much from use case to use case. There are two mechanisms to ensure appropriate data quality. The first derives from FAIR principles, as interoperable and reusable data implies that the data set has a given minimum amount of metadata. Defining an appropriate standard for metadata that can be efficiently defined by data depositors and implemented by repositories and index/search services, will be key to implementation of the EOSC. It would be helpful if all data deposit is accompanied by a data and metadata quality statement, describing the processes used leading to the data deposit.

The second mechanism is that of peer-review and collective filtering, e.g. individually based reviews provided by users. As data become more accessible, it may be useful to provide mechanisms in search systems/indices for users to provide reviews that could be used to supplement citation counts. These reviews could make it easier for data sets that were found to be insufficient for the needs of one project or publication to be reused elsewhere. This would also reduce the time taken by users looking for data to examine a data set before deciding to take a deeper look or look elsewhere. Such reviews would also provide an opportunity for groups to indicate issues with data analysis performed in the initial publication(s) and provide an important part of the replication infrastructure currently needed in science today.

5.6 Data security

Data that will be distributed via the EOSC will have different levels of access control depending on intellectual property (IP) issues, embargoes prior to publication and personal data-protection considerations. In addition, certain types of research may have national security implications that require additional levels of access control. The only model viable in such an environment is one whereby data security, or more accurately, access control remains with the entity that is ultimately legally responsible for ensuring that the data are properly restricted. This implies a very flexible access control regime, as some data (such as, for example the information underpinning a conventional research publication that does not involve human subjects or touch on national-security issues) should be made open after publication, while information such as human-subject research data may need to be explicitly controlled by a data access committee at the organisation that carried out the research. In other cases, a holding entity (for example a data repository) could assume the legal burden for ensuring appropriate access control.

From a more conventional information security standpoint, the EOSC, like most clouds, will need to operate under a shared security model. That is, the provider of IaaS, PaaS and SaaS, (See Glossary for acronyms) will share the information security duties with the entity that is utilising the service(s). Again, this could be an individual or a research institution using a generic storage service, or a much more sophisticated environment, where a repository provides value-added services on top of basic IaaS provided by a third party. In these cases, the EOSC entity should follow appropriate best practices for cloud security, including a risk assessment, a data-protection impact assessment (as the case may be), implementation of appropriate controls and auditing commensurate with the risk assessment³⁷.

³⁷ View ENISA frameworks at: <https://www.enisa.europa.eu/topics/threat-risk-management/risk-management>

6. THE ROAD AHEAD TO COMPLETION OF THE 2nd HLEG

Several activities around the RoP and its implications for the operation of the EOSC under the EOSC ecosystem have been discussed. Moreover, the RoP define the rights, obligations and accountability of those involved from the EOSC (such as: data producers, service providers, data/service users) against:

- » Agreed tools, specifications, catalogues and standards (EOSC shared resources) and applicable methodologies (framework for FAIR research data);
- » Adopted principles for regulating transactions in the EOSC (e.g. financial mechanisms and procedures, agreements/bylaws established by the EOSC governance framework);
- » Applicable legal frameworks (e.g. GDPR, copyright rules, Data Security and Cybercrime, dispute resolution and redress mechanisms, e-commerce directive);

However, there is room and need for differentiating the rules applicable to those involved from the EOSC depending on their maturity and role and taking into consideration:

- » Specificities of different scientific disciplines;
- » Diversity and level of readiness of infrastructures and services at discipline, Member States and EU level (RIs, e-infrastructures) and the differences in their established rules and processes;
- » Variety of service providers and users that will be involved in the EOSC (e.g. public vs private; horizontal vs specialised);
- » Changing needs and practices on the implementation of the rules, in particular on compliance with existing legal frameworks (e.g. GDPR) and emerging ones (e.g. free flow of data)..

The 2nd EOSC HLEG have provided a more detailed description of the RoP, and supported the preparation of the June 2018 EOSC summit. Moreover, the implementation roadmap is to be incorporated, and the coalition of doers will be stimulated to serve a role in the governance structure.

6.1 2nd EOSC summit 2018

One year on from the EOSC Summit 2017, this 1-day event offered as an opportunity to build upon the achievements and progress being made towards an EOSC, and to continue to build momentum.

The EOSC summit 2018 served as an opportunity to **launch an open stakeholder consultation** on the draft '*Rules of Participation of EOSC*', as a key input for the future EOSC governance.

The aim of the Consultation was twofold.

- » To allow stakeholders as any other person involved in the EOSC to rank, vote and discuss the three sets of Recommendations (implementation, engagement & steering recommendations) that the 2nd EOSC HLEG presented during the EOSC summit and which will guide the practical development of the EOSC.
- » Contribute to discussions on relevant topics for European researchers and research e- infrastructures and actively contribute to the shaping of the EOSC RoP.

6.2 Support to the Implementation Plan and roadmap for the EOSC

The EOSC is expected to grow into an eco-system providing an open environment for the scientific community to use, provide and share scientific data and results. Basic components of the open cloud environment are being addressed in the various H2020 projects currently funded, and in: eInfraCentral, EOSCPilot, EOSC Hub, successful INFRA-EOSC-2018 proposals, FREYA, OPENAire-Advance, RDA Europe 4, EOSC-Hub³⁸, which integrates and consolidates services, software and data from key e-infrastructures with common access mechanisms for the scientific community. The EOSC portal will provide a single entry point enabling consolidated access to

38 Horizon 2020 Grant no 777536

federated e-infrastructures for data-driven science. In addition, the pilots³⁹ will demonstrate trials crossing borders and scientific domains. All aspects of the implementation phase should be carefully monitored and measures also with the support of agreed upon KPIs and or other metrics. This will positively support execution on the various projects' workplan.

The pilots provide a first step towards growth and additional integration of services, resources and data. A current requisite is a set of rules for participation in the EOSC eco-system through establishment of a compliance framework. This needs to take national science infrastructure roadmaps into account to reach out into a collaborative eco-system instead of fragmented, single, national solutions.

The compliance framework is to outline the rules for participation, how to meet FAIR data principles, how services can be integrated, address security and privacy and include governance practises and risk controls. Its approach and commitments given by strategy and objectives will be overseen by the governance board of the EOSC (as part of the governance model). The policy and process set the rules for tools and processes which can be embedded into the EOSC eco-system. The organisational structure, human resource mechanisms, change management, performance measurement, training, communication and education cover the usage for the resources and access for people.

There are several challenges which need to be addressed before a compliance framework can be defined:

Challenge-1 enable integration of and access to resources that will be federated in the EOSC

As opposed to developing yet another interoperability framework, it is important to follow an open interface approach, which permits a seamless integration/adaptation to existing Hubs. For resources and services, it is fundamental to build services based on open standards APIs and protocols.

Rules are be set to build services based on open standard APIs and protocols for resource and service portability reasons.

Access to services is to follow legal security and privacy requirements. This implies that the EOSC provider needs to implement the risk controls for GDPR and eIDAS compliance.

Rules are to define the compliance statement⁴⁰.

Challenge-2 Cross-border and cross-discipline research collaboration.

It is not only necessary to access and find services or data sets for collaborative work, but also to enable collaboration between ad-hoc communities in the future. These communities might work in a protected space and therefore necessitate individual settings. It is important that the EOSC eco-system provides a collaboration mechanism for ad-hoc research groups with the respective settings for security and privacy. Depending on the needed resources for computing, the system also needs to provide ad-hoc resources for intensive computations.

Rules for collaboration are to include the potential of building dynamic ad-hoc groups with respective security settings, if necessary. The rules are to also include publication of the results as generated in the collaboration project.

The portal is to include social media to allow the researcher to build cross-country and cross-border collaboration.

Scientific data repositories have to follow FAIR data principles. This also includes a FAIR data certification. Here it is important to note that the European Commission is working on a European Cloud Certification⁴¹ scheme as part of the Free Flow of Data regulation draft. As several certification schemes already exist, which can be found on the European Union Agency for Network and Information Security (ENISA) webpage⁴², it is important to work together and not define yet another scheme. Furthermore, certification for services, infrastructure and resource providers generally involves costs. Therefore, it has to be determined who is pays for the certification, or whether self-certification is sufficient. Resources offered by smaller research entities also have to be taken into consideration, for example, does this have to be the same certification process required by the EOSC marketplace (via the EOSC Hub)?

39 <https://eoscpilot.eu/>

40 For GDPR: either CoC, art. 40, or certification art. 42

41 Digital Single Market Stakeholder Meeting December 2017

42 <https://resilience.enisa.europa.eu/cloud-computing-certification>

Rules need to be set for FAIR data compliance but must also take into consideration risk controls already set by resource or service providers.

A new role may have come into play for the federated EOSC eco-system, or rather a kind of national or scientific communities Hub, which will be responsible for compliance and clearing of new services and data sets.

6.2 Endorsements & commitments from the coalition of doers

The first EOSC summit was held on 12th June 2017. It brought together 110 players, including scientists from a wide spectrum of scientific fields, representatives from scientific infrastructures, research funders and officials of Member States and Associate Countries acknowledging strong support for the implementation of the EOSC.

The second EOSC summit organised on the 11th June 2018 brought together 180 key players for the implementation of the EOSC, representing all categories and scientific fields. Fifteen research funders and about 30 officials attended from ministries of Member States and Associate Countries. The summit aimed at taking stock on progress and achievements towards the EOSC, a year on since the 2017 Summit, while allowing participants to share information on relevant activities and commitments and to reflect on the next steps of the implementation. With a view to providing input for the future EOSC governance and set-up, a stakeholder consultation on the draft *Rules of participation of EOSC* and on the draft *FAIR data action plan* was launched.

Those involved who are key to the implementation of the EOSC, confirmed the relevant role played by the ‘coalition of doers’ and encouraged momentum to be maintained; this led to the submission of over 70 endorsements/commitment letters as a result of the EOSC declaration, see the figure below.



Figure 3 – The 70+ signatories of the EOSC declaration

Participation was highly representative, including scientific fields, national scientific infrastructures, research funders and ministries of Member States and Associated Countries. The 2nd EOSC HLEG stimulated the ‘coalition of doers’ to move to the next level and describe their vision with practical examples of an EOSC situation, practice or implementation. This could be an existing initiative, or a vision for future implementation within a timeframe of three to five years.

The current coalition of doers should maintain their momentum. Their individual incentives and suggestions for cooperation with EOSC, they have made to date, should be showcased at the yearly stakeholder's forum event.

6.4 Long term challenges

A number of long-term challenges have been identified by the HLEG, among which are the following.

- » Private sector involvement in the EOSC.
- » Human capacity development for open science.
- » Further development of procurement models for EOSC delivery.
- » Addressing Green IT requirements.
- » Emerging technological areas and priority needs (e.g. Blockchain, ethical AI, cybersecurity).
- » FAIR-data-related challenges:
 - » What are the criteria for 'EOSC-compliant' services, particularly if they are offered from third parties?
 - » Are FAIR services needed to deal with FAIR data and if so, what are the criteria for FAIR services?
 - » Could criteria for FAIR services define 'EOSC-compliance' of services?

A timely aspect to be explored in EOSC is to look at the blockchain research. It allows scientists to share digital information, but not to copy it. With that the IPRs can be kept with the source of origin. Information held on a blockchain exists as a shared - and continually reconciled - database. It helps to guarantee the validity of a transaction by recording it not only on a main register, but connected distributes registers, which are all connected via a secure validation mechanism. This kind of distributed ledger can be applied to the data registries'. This can have relevant impact on the way researcher build their reputation and get recognised. The concept of the blockchain can introduce a new concept of trust for communities to work together in an open or closed/restricted way.

Getting in touch with the EU

IN PERSON

All over the European Union there are hundreds of Europe Direct Information Centres.

You can find the address of the centre nearest you at:

https://europa.eu/european-union/contact_en

ON THE PHONE OR BY EMAIL

Europe Direct is a service that answers your questions about the European Union.

You can contact this service:

- by freephone: **00 800 6 7 8 9 10 11** (certain operators may charge for these calls),
- at the following standard number: **+32 22999696**, or
- by email via: **https://europa.eu/european-union/contact_en**

Finding information about the EU

ONLINE

Information about the European Union in all the official languages of the EU is available on the Europa website at:

https://europa.eu/european-union/index_en

EU PUBLICATIONS

You can download or order free and priced EU publications from

<https://publications.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see **https://europa.eu/european-union/contact_en**)

EU LAW AND RELATED DOCUMENTS

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: **<http://eur-lex.europa.eu>**

OPEN DATA FROM THE EU

The EU Open Data Portal (**<http://data.europa.eu/euodp/en/data>**) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.



To help drive forward and implement the EOSC, the main thread of the report is to understand how the EOSC can effectively interlink **People, Data, Services and Training, Publications, Projects and Organisations**. The ideas presented here bring together, reflect on and further expand on various policy papers and recommendations contributing to the establishment of the EOSC that have been published by ongoing Horizon 2020 projects and national initiatives, as well as by the Commission FAIR Data expert group and by the Open Science Policy Platform, with whom the group have collaborated actively. The report shows how Europe, with its strong scientific base and investments made in infrastructures has the skills, knowledge and capacity to turn EOSC into a reality in less than a year from now.

Studies and reports