**Open Consultation**

**for the**

**Strategic Research and Innovation Agenda (SRIA)**

**of the**

**European Open Science Cloud (EOSC)**

Version 25 June 2020



|  |
| --- |
| * This document sets the draft general framework for future strategic RDI activities to be further defined in the context of the candidate [EOSC European Partnership proposed under the Horizon Europe Programme.](https://ec.europa.eu/info/files/european-open-science-cloud-eosc_en) * It reflects elements of the candidate Partnership proposal document proposed by the current EOSC Governance as well as further work by the Executive Board, in order to develop by October 2020 a first full version of the Strategic Research and Innovation Agenda (SRIA) for the EOSC. * With the consultation launched on 20 July, we are seeking the views and contributions of different stakeholders on the content of this document through the accompanying questionnaire. The consultation will remain open until September 15. |

Table of Contents

[1 EOSC Partnership Vision and Approach 3](#_Toc43662331)

[*1.1 The vision 3*](#_Toc43662332)

[*1.2 Main drivers 4*](#_Toc43662333)

[*1.3 Value of a European Partnership approach 5*](#_Toc43662334)

[*1.4 Guiding principles to implement the EOSC vision 5*](#_Toc43662335)

[*1.5 Overall Timeframe 6*](#_Toc43662336)

[*1.7 Synergies with other programmes 8*](#_Toc43662337)

[2 Strategic Objectives and Areas of Action 8](#_Toc43662338)

[*2.1. Strategic Objectives 8*](#_Toc43662339)

[*2.2. Areas of Action 11*](#_Toc43662340)

[3 Priorities per Strategic Objective 16](#_Toc43662341)

[*3.1. Strategic Objective 1: 18*](#_Toc43662342)

[*3.2. Strategic Objective 2: 19*](#_Toc43662343)

[*3.3. Strategic Objective 3: 22*](#_Toc43662344)

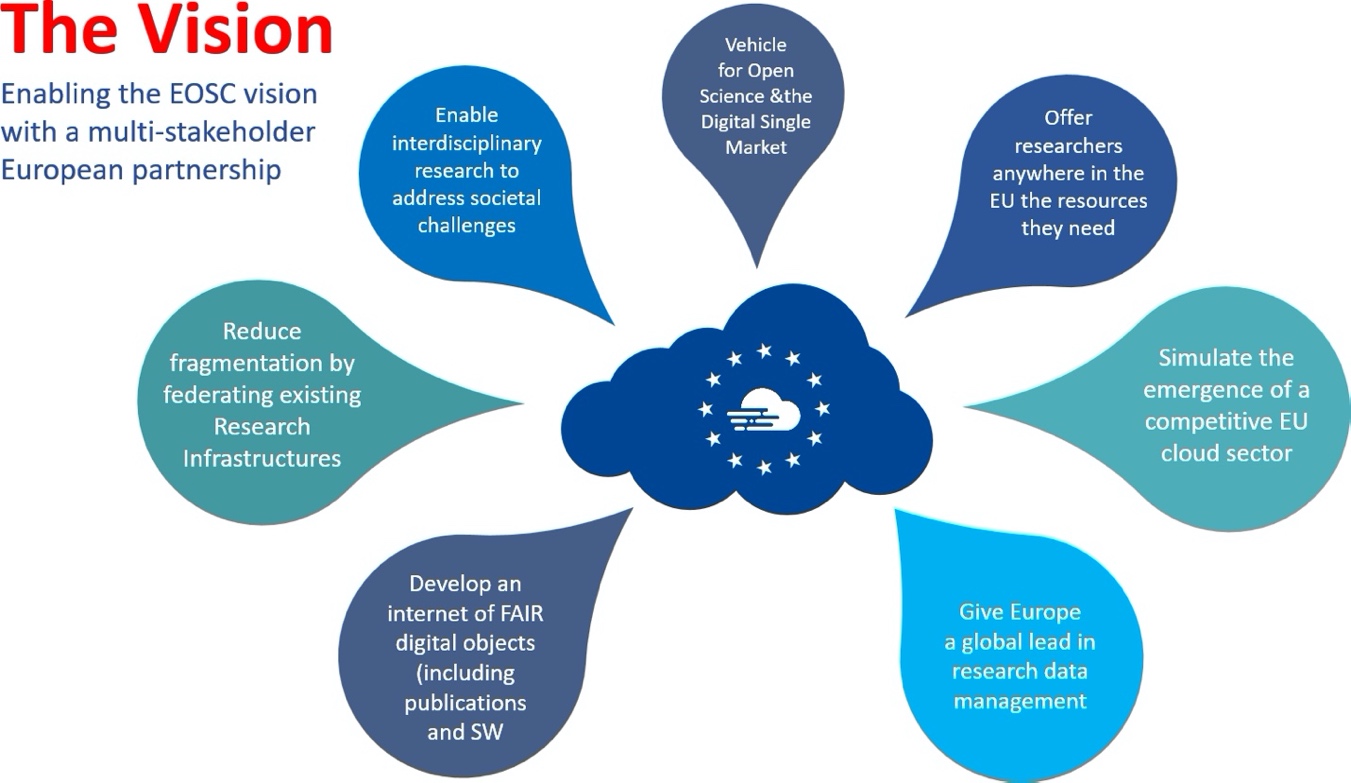
[4 Expected impacts 24](#_Toc43662345)

# EOSC Partnership Vision and Approach

The science system in Europe is in transition. It faces increasing challenges related to trust in science and reproducibility, rewards and incentives to become more open, collaborative and effective. It needs to allow data sharing and management to trigger reuse and enable new scales of multi-disciplinary research and frontier science. It is expected to open up to major industrial and social innovations while adapting to cutting-edge digital technologies and interconnecting the European landscape of Research Infrastructures and e-Infrastructures. Accelerating this transition of the European science system requires a multi-stakeholder European Partnership to enhance the circulation of research data and knowledge in digital form across borders and disciplines and to allow scientists and machines to collaborate in storing, processing, finding, accessing and reusing scientific data.

## 1.1 The vision

The [European Open Science Cloud (EOSC) Partnership](https://ec.europa.eu/info/files/european-open-science-cloud-eosc_en) will enable a trusted, virtual, federated environment in Europe to store, share and re-use digital output from research (publications, data and software) across borders and scientific disciplines. The Partnership will bring together institutional, national and European initiatives and engage all relevant stakeholders to co-design and deploy a European Research Data Commons where data are Findable, Accessible, Interoperable, Reusable (FAIR). This European contribution to a “Web of FAIR Data and Related Services for Science” will enhance the possibilities for researchers to find, share and reuse publications, data, and software leading to new insights and innovations, higher research productivity and improved reproducibility in science. The challenge is not limited to linking datasets, federating infrastructures or aligning policies; it starts by linking people and organisations across the EOSC ecosystem.

****

## 1.2 Main drivers

* ***The exponential growth in the quantity of research data****.* Today’s data generation is growing exponentially and is measured in Zettabytes (1021 bytes). The vast increase in data production equally applies to the domain of research, whereby researchers are already unable to read (in terms of amount) or access (in case of data) all relevant digital knowledge in their field. In particular, the underlying research data remains predominantly unpublished and are therefore unfindable and inaccessible.
* ***Creating interoperable data commons for data-intensive science and innovation*.** Scientific data is in dire need of openness, better handling, careful management, machine actionability and sheer re-use[[1]](#footnote-1). This cannot be realised without specifications and standards for common components to enable interoperability across the research data ecosystem. Moving to ‘FAIR-by-design’ digital research outputs requires further efforts to develop, refine and adopt shared vocabularies, ontologies, metadata specifications and standards, as well as increasing the supply and professionalisation of data stewardship, data repositories and data services in Europe. The science, research and innovation data space ambitioned by the EOSC shall be properly articulated within the new sectoral dataspaces foreseen in the European Data Strategy to enable cross-border access and use of data and data enabled services by the researchers, public authorities and the commercial sector.
* ***Bridging existing European research infrastructures and creating interoperable data commons*.** Over the last decade or so, there have been significant investments across Europe in data-oriented research infrastructures and e-Infrastructures. The outcome is a vast quantity of infrastructure components of various scales and scopes, centralised or distributed, being generic, domain-specific or cross-disciplinary. Many of these components have not been designed initially to work together.
* ***Connecting research data with robust computing technologies and fast connectivity networks.***At the same time, Europe needs a robust European-scale environment for storage and data analysis and computing. Access policies for networking, data storage, computing and processing differ across Europe. This makes scientific cooperation in the EU more difficult, especially if we want to leverage on existing e-infrastructures across countries and disciplines. Shareable research data, open data analysis tools and connected computing facilities need to become available for all researchers.
* ***The policy drive for Open Science and Open Innovation in Europe.*** Open Science and Open innovation are widely recognised as key transformative, enabling elements of the European R&I policy driving a renewed European Research Area (ERA). While the majority of the Member States and Associated Countries have policies in place regarding open access to scholarly publications, less countries have yet defined policies regarding FAIR data and only few mention EOSC in their policies (although some plan to do so in the future). There is a need to progressively operationalise and open up the EOSC ecosystem to allow not only European academics and researchers but also public authorities, innovators and companies and society at large to contribute and exploit open data, publications, software and associated services in Europe.

## 1.3 Value of a European Partnership approach

A European Partnership is identified by the current EOSC Governance as the best instrument to overcome fragmentation and to provide a framework for collaboration and pooling of resources at European, national, regional and institutional levels. An open and inclusive European Partnership will help ensuring directionality (common vision and objectives) and additionality (complementary commitments and contributions at all levels). A Partnership helps providing a focal point and a framework to reach consensus amongst committed ‘Doers’. By agreeing on a common Strategic Research and Innovation Agenda (SRIA) reflecting the views of the EOSC communities, the Partnership will demonstrate a shared vision and operational objectives for the next seven years and beyond.

## 1.4 Guiding principles to implement the EOSC vision

* ***Multi-stakeholder approach.***The EOSC vision is not limited to linking datasets, federating infrastructures or aligning policies; it starts by linking multiple stakeholders along the data lifecycle and across the European research ecosystem. This includes typically the Member States (for instance through national organisations having received a mandate from their ministry to represent their national research landscape), research performing organisations (including universities and research centres), research infrastructures and e-Infrastructures (e.g. related to storage, computing and communications), research libraries, research associations, international research centres, as well as (at a later stage) other entities from the public and private sectors.
* ***“As open as possible as closed as necessary”.***The primary focus of the implementation will be on Open Data that are as FAIR as possible. This will allow the deployment of services and applications that will become examples to demonstrate the benefits of Open Science. In parallel, efforts will be put on enabling new types of incentives and skills to share data openly, improve data FAIR management, develop literacy and data stewardship skills. As EOSC matures, it will gradually expand towards more complex and specific challenges in areas where FAIR data cannot be kept open by default. Data protection, for example GDPR compliance, will have to be ensured in order to extend the reach of the EOSC to disciplines in which personal data are processed.
* ***Towards a ‘Web of FAIR data and related services for Science*’.** The Partnership will engage all relevant stakeholders to co-design and deploy a European Research Data Commons where data are Findable, Accessible, Interoperable, Reusable (FAIR). This European contribution to a ‘Web of FAIR Data and Related Services for Science’ will enhance the possibilities for researchers to find, share and reuse all kinds of research digital artefacts such as publications, data, and software. The FAIR guiding principles and related metadata standards will bring the minimum soft overlay required to enable common implementation and form a Web of FAIR Data and services for science. At the same time, it will offer maximum freedom of implementation and will allow participation from all stakeholders, including research infrastructure providers, research performing organisations and businesses.
* ***Federating existing research infrastructures*.** The European landscape of Research Infrastructures is largely influenced by national roadmaps. Preliminary findings of the ongoing mapping by the EOSC Governance has confirmed that a majority of Member States and Associated Countries are making significant investments in national infrastructures that could in principle be federated as part of the EOSC ecosystem. The same applies to many of the more thematic data intensive ESFRI landmark infrastructures, also supported by several Member States. The scale of such investments over the past decade is certainly in the order of billions of euros. The Partnership envisions enabling a federating core (or EOSC-Core) assembling all the basic elements to operate and provide the means to discover, share, access and re-use data and services across the EOSC ecosystem as well as access to a service layer comprising Common Services and Thematic services (EOSC-Exchange).
* ***Machines & People*.** Computers have long surpassed individuals in their ability to perform pattern recognition over large data sets. Ultimately, actionable knowledge and translation of its benefits to society will be handled by humans in the 'machine era' for decades to come. As computers have become indispensable research assistants, we need to make what we publish understandable to them. Scientists need both in concert to form social machines in order to be able to visualize, analyse and predict in facing complex problems. The availability of data that are ‘FAIR by design’ and shared Application Programming Interfaces (API) will allow new ways of collaboration between scientists and machines to make the best use of research digital objects of any kind.

## 1.5 Overall Timeframe

Three stages of implementation are foreseen serving the following two major milestones[[2]](#footnote-2):

* Deploy European Open Science Cloud operations to serve EU researchers by 2025;
* Open up, connect and articulate EOSC beyond the research communities, with the wider public sector and the private sector from 2024 onwards.

|  |
| --- |
| **Stage 1: period 2021-2023**  Creating the **European Open Science Cloud operations (EOSC-Core)** to provide authentication & authorization infrastructure (AAI) and other necessary core functions of the Minimum Viable EOSC.  **Stage 2: period 2024-2025**  Expanding the Minimal Viable EOSC with **access to** **added value services, applications and tools** **(EOSC-Exchange)** supporting the full cycle of scientific workflows.  First pilots / demonstrators on linking EOSC beyond the research communities with the wider public sector and the private sector from 2024 onwards, for addressing societal challenges  **Stage 3: period 2026-2027 and beyond**  Deployment of the **Web of FAIR Data and related Services,** including framework conditions for interoperability and machine-actionability of data. Connection of the European contribution to a Web of FAIR Data and Services to other Open Science Commons across the world.  Continuous support to enhancing the ecosystem of the Web of FAIR data and related services for the research community. |

*1.6 Synergies with other initiatives: boosting the impact of EOSC through collaboration and alliances*

The EOSC partnership is a key partnership in a wide network of other partnerships, initiatives and infrastructures that conform the digital and the research ecosystems in Europe and globally. The necessary shared activities and alliances should be built in order for Europe to optimize the outcomes of its efforts for the benefit of the research community and society at large.

***European Partnerships and Missions***

The EOSC Partnership aims to accelerate the transition to more effective Open Science and Open Innovation with new opportunities for multi-disciplinary research, frontier science and data-intensive science. Therefore, the Partnership has a unique transversal role (cross-Pillar, cross-Cluster) in Horizon Europe, bringing prospects of collaboration with most of the proposed European partnerships.

* All vertical partnerships can take benefit from a successful development of the EOSC as this will equip them with minimal, rigorous standards and protocols and maximum freedom of implementation to share and reuse data and other digital objects across relevant domains of research.
* The EOSC partnership will join forces with partnerships working on specific technologies: the European Partnership for High Performance Computing, European Partnership on Artificial Intelligence, Data & Robotics.

***European Data Spaces[[3]](#footnote-3)***

The EOSC Partnership will also seek synergies with the European Data Spaces. While EOSC focuses on federating research data infrastructures, the research data made accessible through EOSC will be relevant for several dataspaces (e.g. health, Destination Earth). EOSC will build primarily on FAIR data. Where appropriate, access to such data will be provided.

*Open Science organisations*

Open Science has become a shared goal all over the world. While Europe is in a leadership and pioneering position, there are many other initiatives all around the world.

* Existing Open Science Commons initiatives include the Australian Research Data Commons (ARDC), the African Open Science Platform (AOSP), NIH Data Commons, Canada's National Data Services Framework (NDSF), the Chinese CSTCloud and the Association of Southeast Asian Nations (ASEAN) initiative.
* In the thematic realm, examples are the Australian Biocommons, the Data Commons for Food Security and CSIROs Managed Data Ecosystem.
* The Research Data Alliance (RDA) has become a recognised actor in the development of standards and good practices for managing research data. CODATA, as the standing committee of the International Science Council works closely with the RDA and GO FAIR working groups and implementation networks, while the World Data System focuses on high quality data services with an emphasis on repositories and their certification.

## 1.7 Synergies with other programmes

The EOSC partnership can play a strategic role in supporting the EU policy priority of having a Europe fit for the digital age. Therefore, at European level, synergies with other programmes, in particular the Digital Europe Programme and the Connecting Europe Facility, in areas such as the development of digital skills, standards or tools should be exploited in order to pool the necessary resources and avoid overlaps.

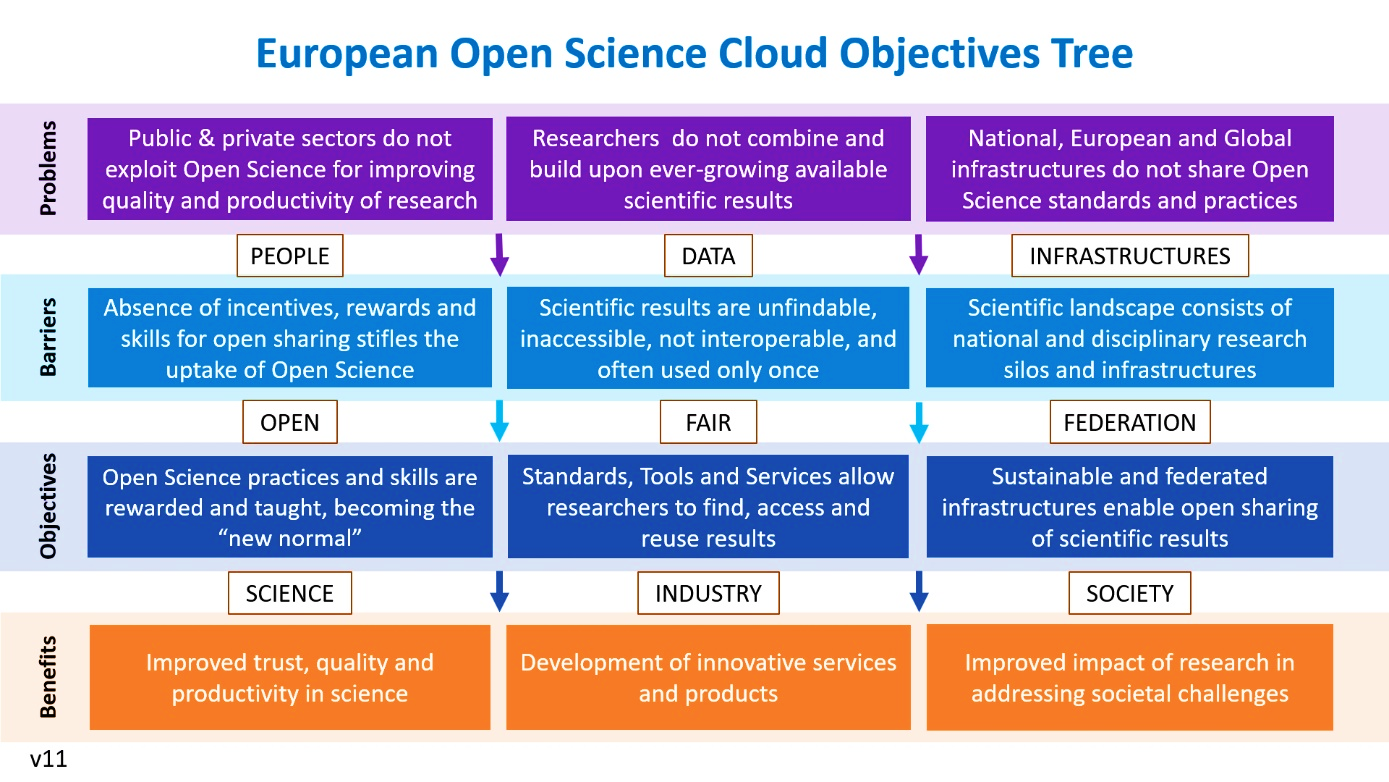
Synergies with the [European Structural and Investment Funds](https://ec.europa.eu/regional_policy/en/policy/what/glossary/e/esif) should also be exploited in order to leverage the investments through structural funds (Smart Specialisation Strategy) in research data infrastructures that can be federated as part of the EOSC ecosystem. Moreover, a majority of Member States are making significant investments in national infrastructures and Open Science programmes that could in principle be federated as part of the EOSC environment. There is, therefore, a strategic European value of improving alignment and compatibility of national Open Science policies and national plans for data infrastructures in the EOSC context.

# Strategic Objectives and Areas of Action

We have identified three main *Strategic Objectives* and fourteen *Areas of Action* that will accelerate the development and deployment of the EOSC ecosystem along the vision, principles and timeline described in Chapter 1. These Strategic Objectives and Areas of Action are set out in more detail below.

**NOTE:** *Please note that the numbering of the Strategic Objectives (SO1, SO2 and SO3) and the Areas of Action (AA1, AA2, etc.) is not related in any way to their level of importance. They are numbered so that they can be then recognized in the accompanying questionnaire.*

## Strategic Objectives

The Objectives tree below sets the foundations for the EOSC partnership intervention logic along three main Strategic Objectives.

**Strategic Objective 1:**

***Open science practices and skills are rewarded and taught, becoming the “new normal”***

**Main milestone:** The EOSC ecosystem underpins the reward of open science practices and data stewardship thus improving trust, quality and productivity in science.

**Scope:** A key goal of the EOSC Partnership is to help in shifting the research enterprise in Europe towards an Open Science model as there is a political will towards the notion of open research and many European countries are implementing national programmes that are aligned with the European Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information[[4]](#footnote-4).

EOSC will be established as the EU wide infrastructure for open research. Assuming the European Commission and national funders will require from researchers to use ‘EOSC-approved’ infrastructures, this will lead to output in line the “new normal”.

**Strategic Objective 2:**

***Standards, tools and services allow researchers to find, access and reuse results***

**Main milestone:** Making data and any other digital research object (such as algorithms, tools, and workflows) as FAIR as possible across all European research infrastructures will be key for sharing/reusing research results.

**Scope:** The FAIR principles provide high-level answers to the challenges in managing research data:

* ***Heterogeneity:*** Data are defined by the scientific discipline that produces or studies them. They are coming from equipment, surveys or any other data production /collection process. Data are also the results of processing steps developed after the raw data collection. As a consequence of this multifaceted landscape, a very wide variety of data frameworks have been designed by research communities.
* ***Granularity:***While raw data may be produced in large volumes, users may not need access to raw data but only to the results of post-processing steps. Deciding at what stage data should be preserved and made available to users is also specific to the scientific domain.
* ***Versioning:***Data are collected as time goes by. The sampling rate, the importance of preserving knowledge on the sequence of events also depends on the domain.
* ***Disambiguation:*** When it comes to representing data in digital form, the choice of identifiers is a problem in itself. It may occur that one identifier points to multiple objects introducing one more level of complexity for the user.
* ***Diffusion******prevention:*** On the other hand, it may also occur that multiple identifiers point to the same object introducing a different level of complexity.

The availability of data that are FAIR by design will allow scientists to make the best use of new data by leveraging the power of machines. FAIR data being machine-actionable allow to develop software services, applications and tools that deliver the proper information for scientists to nurture their research.

Researchers are increasingly relying on computational and machine-assisted support to deal with research data as a result of the increase in volume, complexity, and creation speed of data. There is thus currently a scientific and policy consensus that research data must be made machine-actionable, when applicable, to allow computational systems to find, access, interoperate, and reuse research data.

Digital objects must be described with rich metadata, assigned a globally unique persistent identifier, and be released with a clear and accessible usage license. There is an onus on researchers to adopt relevant community standards and select appropriate data services which enable digital objects to be discovered and retrieved using standard protocols, applicable for both humans and machines. Research communities need to define data standards, sharing agreements and services to enable FAIR digital objects.

**A screenshot of a cell phone

Description automatically generated**

**DATA**

**The core bits**

**IDENTIFIERS**

**Persistent and unique (PIDs)**

**STANDARDS & CODE**

**Open, documented formats**

**METADATA**

**Contextual documentation**

*Figure 3: FAIR ecosystem*

The strength of the FAIR principles is in defining a set of common characteristics required in all digital objects, irrespective of type, discipline and content. This enables machines to act across a broad set of content, enabling interdisciplinary research. Many aspects of the FAIR principles however speak to community standards and practices. The principles will be applied differently according to the needs and requirements in the different fields of knowledge. Crosswalks and brokering are needed to support interoperability across disciplinary standards. The interoperability frameworks should be articulated in common ways and adopt global standards where relevant. Intelligent crosswalks, brokering mechanisms and semantic and other cutting-edge technologies like artificial intelligence, should all be explored to break down silos and allow for cross-disciplinary data exploration, analysis and visualisation.

**Strategic Objective 3:**

***Sustainable and federated infrastructures enable open sharing of scientific results***

**Main Milestone:** The EOSC offers a federation of infrastructures, forming a **Web of FAIR Data** **and related Services for Science**, supporting research by the academic, private and public sector.

**Scope:** The system will be based in three layers: (1) The federating core (or **EOSC-Core**), (2) the **federation** of existing and planned research data infrastructures; and (3) a service layer comprising Common Services and Thematic services (**EOSC-Exchange**). Building on existing research data infrastructures, EOSC will grow through a series of iterations. Each iteration will add more functionalities and services for a wider user base and satisfying a broader range of use-cases.

(1) The **EOSC-Core** assembles all the basic elements to operate and provide the means to discover, share, access and re-use data and services. These elements address key technical, cultural and policy decisions of EOSC and they must be maintained over the long term.

(2) **Federation**: the FAIR principles and metadata standards enable the federation of existing and planned research data infrastructures, adding a soft overlay to connect them and forming a Web of FAIR Data and services. This is a distributed system that needs an underlying frame based on commonly agreed, minimal standards and maximal freedom to operate with agility, whilst still ensuring global and interdisciplinary interoperability.

(3) The **EOSC-Exchange** builds on the EOSC-Core to ensure a rich set of services (common and thematic) exploiting FAIR data and encouraging its reuse are available to publicly funded researchers. It is expected that rivalrous services, such as those that store, preserve or transport research data as well as those than compute against it, will be made available via EOSC-Exchange. Service providers that participate in EOSC-Exchange will be required to conform to predefined Rules of Participation.

## Areas of Action

We have identified fourteen Areas of Action (AA) to help deploying the EOSC ecosystem. Some of them are more technical in nature, others relate more social dimensions (financial, legal, educational, cultural, policy). In the table below, these are listed in two columns:

|  |  |
| --- | --- |
| ***‘Implementation challenges’*** | ***‘Boundary conditions’*** |
| AA1: Identifiers  AA2: Metadata & Ontologies  AA3: FAIR metrics and Certification  AA4: Authentication & Authorisation Infrastructure  AA5: Users Environments  AA6: Resource Providers Environments  AA7: EOSC Interoperability Framework | AA8: Rules of Participation  AA9: Landscape Monitoring  AA10: Business Models  AA11: Skills & Training  AA12: Rewards & Recognition  AA13: Communication  AA14: Widening to the public and private sectors |

***AA1 - Identifiers***

Persistence of the identity of entities and stability of references to those entities are essential to the European Open Science Cloud.  Only if researchers can be assured that entities (including publications, data and software resources) do not alter over time and are continuously accessible via linking mechanisms, can a trusted distributed research ecosystem be sustained which supports verifiable and reusable research. The use of Persistent IDentifiers (PIDs) has been specifically recognised within the FAIR data principles as a key feature supporting findability and accessibility of research entities. Therefore, PIDs form a stable, trusted structure which can be used to make the research infrastructure a reliable source of verifiable and reproducible research. In the EOSC, we should seek to support a shared policy for the use of PIDs both for the management and analysis of data, and also the publication, curation and tracking of research outputs.

***AA2 - Metadata and Ontologies***

Metadata and ontologies are essential to realise Open Science, and thus need to be addressed by EOSC. Metadata and ontologies have evolved organically over time, addressing the needs of specific communities. Because of these main drivers, an overarching, coordinated approach to metadata and ontologies for scholarly resources has for the most part been missing so far. This both explains the most important gaps we are seeing today, and provides a unique opportunity for EOSC.

The ultimate goal of this coordinated approach going forward should be to improve the adoption and interoperability of existing metadata schemata through registries that describe these metadata schemata in a standardized and machine-accessible way, better researcher-focussed tools and services working with these metadata, crosswalks between existing metadata schemata, and training and documentation. The drivers for all work around metadata and ontologies should be use cases from and adoption by the researcher community, and the work be based on existing infrastructure and communities.

***AA3 - FAIR Metrics & Certification***

[To be completed]

***AA4 - Authentication & Authorisation Infrastructure***

[To be completed]

***AA5 - Users environments***

User environments are the digital platforms which users go to in order to interact with EOSC and EOSC resources. These include Portals, through which the EOSC infrastructure is accessed. It may also include other environments yet to be created, as part of the central part of EOSC or created by thematic or regional communities.

***AA6 - Resource providers environments***

These comprise the interface between the resource provider, a community and EOSC. This interface will be a framework of processes, tools, approved standards, APIs and other elements which enable resource providers to bring their resources into EOSC. Hence, Resource Provider environments are key to EOSC in that they are the way that EOSC brings in the ‘supply-side’ from the broader community, from generic e-infrastructure services through to thematic services coming from the ESFRI clusters and RIs. By interfacing with EOSC, their resources are then available within EOSC, and are able to add value to European research.

***AA7 - EOSC Interoperability Framework***

Achieving interoperability within EOSC is essential to federate services and provide added value for users. The [draft EOSC Interoperability Framework](https://www.eoscsecretariat.eu/sites/default/files/eosc-interoperability-framework-v1.0.pdf) (which is currently open for comment) identifies general principles and organises them into the four layers: technical, semantic, organisational and legal. The framework also contains a proposal for how the management of FAIR Digital Objects should be done in the context of EOSC.

The initial draft has been developed by members of the [FAIR](https://www.eoscsecretariat.eu/working-groups/fair-working-group) and [Architecture](https://www.eoscsecretariat.eu/working-groups/architecture-working-group) Working Groups. The authors conducted an extensive review of related literature and interviewed key stakeholders from ERICs, ESFRI projects, service providers and research communities. This helped to identify problems and requirements in each aspect of interoperability to provide recommendations for EOSC. Legal issues will be included in the next version, based on recommendations from a [commissioned study](https://www.eoscsecretariat.eu/funding-opportunities/open-calls/experts-study-legal-regulatory-issues-fair-principles).

***AA8 - Rules of Participation***

The Rules of Participation (RoP) for the European Open Science Cloud aim to provide standards for policy, processes and procedures that provide assurance of quality and trust in the services offered through EOSC. The RoP apply to all digital resources made accessible via EOSC, including data and services. They define a minimal set of rights, obligations and accountability governing the activities of all those participating in EOSC, such as data and service users, data and service providers, and the operators of EOSC itself.

***AA9 - Landscape monitoring***

A first Landscape report is being developed by the Landscape Working Group of the EOSC Executive Board and summarising and describing activities relevant to EOSC in the European Member States (MS) and Associated Countries (AC) as well as some border countries. It compiles existing policies and investments based on inputs from the MS and AC, on the expert knowledge of the WG members and delegates to the EOSC Governance Board. It is further complemented by information from open sources (validated or extended by countries’ authorities when possible) as well as information gathered through Horizon 2020 research projects. Furthermore, the Landscape report also provides an overall description of the landscape of the European infrastructures because EOSC in its core is foreseen as a federated structure of data services linked to data and services provided by research infrastructures at the national and institutional level.

A clear outcome of the Validation workshop of the Landscape report was the finding that the information gathered in the analysis must be kept up-to-date and regularly updated. The monitoring of evolution of national infrastructures and initiatives and the development of respective national policies, supported by a set of relevant Key performance indicators (KPIs), is required in order to allow founded decisions on EOSC. These KPIs must be designed, selected and approved with all the major players as they have also a formative effect and influence the development of national environments. KPIs cannot replace the expertise and knowledge of the Evaluation/Monitoring panel, and the monitoring cannot be reduced to only administrative procedures.

***AA10 - Business models***

[To be completed]

***AA11 - Skills and training***

In order to leverage the potential of EOSC for open and data-intensive research, a key challenge for Europe is to ensure the availability of highly and appropriately skilled people with an excellent knowledge of standards and best practices for delivering, using, sharing, analysing open and FAIR data as well as applications and tools.

Aligned with the New Digital Skills Europe objective (REF), EOSC sets out to develop a large digital talent pool equipped with adequate digital skills embracing a wide range of data related profiles, but also anticipating skills needed to manage software, code and workflows related to AI/HPC. EOSC is expected to facilitate and constitute an open research labour force of data scientists who have expertise in analytics, statistics, machine learning, data mining and data management; data stewards who have strong domain knowledge and the ability to apply this know-how within organisations to create value; research engineers who are able to develop and embed analytical tools in data intensive workflows; open science specialists and  data  stewards who provide their expertise for publishing research in an open and FAIR way, while taking care of legal/IPR issues and ethical issues.

***AA12 - Rewards and recognition***

Presently the Rewards and Recognition (R&R) systems, although not uniform, are too much geared towards number of publications and number of citations (often in relative terms) for as well (parts of) institutions as for individual scientists. Often these systems are working country wide stimulated by ministries, but also the competion between scientists stimulate these systems. Generalising: we would need to move from a R&R system based on: “me”; “culture of shame”; “technical issue” to a R&R system based on: “we”; “culture of care”; “social issue”. The future of evaluating scientists should be on a broader basis, with items like: **Education** (if appropiate); **Research**; **Influence** (on science and/or society and/or economy and/or teaching); **Organisation** and **Leadership**. Different levels of evaluation (Institution, Department, Individual) sound be based on different and appropiate criteria.

***AA13 - Communication***

The EOSC will serve many stakeholders. The Landscape Working Group distinguishes nine different groups. For communication purposes, these groups can be aggregated into three main categories:

1. **Research Service Providers**

* e-Infrastructures, like PRACE, GEANT, OpenAIRE, EUDAT, EGI, also referred to as delivering horizontal services
* Research Infrastructures, like ESFRIs, also referred to as delivering vertical or thematic services.
* Data & Research initiatives, like RDA, offering global platforms for sharing expertise.
* Cloud Providers, incl. commercial parties, offering services to research.
* Cloud Community

1. **Research Performers**

* Research Communities
* Research Performing Organisations

1. **Research Funders**

* Research Funders
* Policy Makers

Each stakeholder group may have different expectations and perceptions on EOSC. Even if these interpretations differ, they can still be consistent. For example, funders may focus on governance issues, whereas researchers and providers focus on functionalities. Despite these differences, we want to provide clarity on the *why*, *how* and *what* of EOSC in a consistent way, and at the same time address these different stakeholders.

*AA14 - Widening to the General Public and the Private Sector*

[To be completed]

# Priorities per Strategic Objective

In chapter 2, the fourteen Areas of Action (AA) have been presented. For each of these areas **an analysis is currently being carried out by the EOSC Executive Board and its Working Groups** of what is the *Status* in the area, what are the *Gaps* that will need to be addressed and what are the *Priorities* in which action will need to concentrate at different levels (EU, national and regional) in order to overcome the gaps and bring the area at the status needed to support the three main EOSC Strategic Objectives (SO).

|  |
| --- |
| **Example of the analysis currently being developed**  **Which objective are we targeting? SO1:** Open science practices and skills are rewarded and taught, becoming the “new normal”  **In which area of intervention?**  *AA5 - Users environments*  **What is the gap(s)?** **Discovery of EOSC:** In the next phase of building EOSC there must be ways to expose the wider community to EOSC.**Discovery of resources:** EOSC portal to offer an integrated meta catalogue by pulling resources from other catalogues, but also for other catalogues to pull resource listings from the central meta catalogue. This interaction must be based on common agreement to use shared formats for resource description, and on APIs.**Ordering, access and use:** Ordering systems must be strengthened and the opportunities for providers must be clarified, as well as the work implied for providers, such that they take up these opportunities and users can benefit from integrated ordering**Composing resources in a user space:** To compose resources in a user space from autonomous and distributed service provider federations requires a legal and organisational framework. This is needed for ensuring the position of the users and their work**Technical support:** Training and support of the open science principles, methods and technologies are essential part of the success of the EOSC. Technical support delivered by service providers of the EOSC is a part of this**Community of practice of EOSC researchers:** More serious attempts must be made to support the creation of communities of practice, as they offer some of the clearest added value of EOSC much as European funding drives the creation of communities of practice in research across the European Research Area. **What is the priority?**  ***First set of priorities:***   * Advanced discoverability of portals * Meta catalogues to aggregate information from the resource catalogues of the service providers * The licenses, usage terms and conditions and user authentication and authorisation methods set by service providers * Legal and organisational framework and its implementation in the distributed architecture * Portals and other EOSC services support this kind of functionality   ***Second set of priorities :***   * Open interfaces of catalogues * The payment principles are transparent and as easy as possible to use along the lifecycle of the research project * Collaboration with service provider and local level support functionalities and resources. * EOSC helpdesk functionalities   **Additional details:**  Users, user groups and service providers have various expectations and requirements like:   * Seamless and easy access to resources: scientific services, research infrastructures, data and others * A space to share and re-use scientific data * A networking and innovation space to open new opportunities of collaboration * A high bandwidth service for transferring data between distributed facilities * A solution for the long-term preservation of large quantities of open data * High-performance storage and compute resources for the data analysis, accessible through cloud technologies like containers or FaaS technologies * A federated search capability for searching and finding scientific data * A set of services for data simulation and analysis ranging from generic services like Jupyter notebooks to domain specific applications per scientific application, including cloud and HPC resources * Service registration, helpdesk, monitoring and accounting * Clear exposure of conditions for accessing the resources   These services will be provided from the distributed service architecture based on offerings of the EOSC service providers. These current expectations mix the realistic and achievable in a short time, to the more aspirational and long term. The current status is outlined in the following sections  **Expected output in 2021-2023:**  The expected output for this first stage (2021-2023) has to come almost exclusively from the INFRAEOSC-03-2020 and INFRAEOSC-07-2020 projects coming out of the present evaluation process. |

The preliminary results of the analysis allow us to showcase an initial set of priorities in this section. Therefore, the following tables show which of the Areas of Action are considered relevant to achieve each of the EOSC objectives and include examples of some of the already identified priorities/activities. Additional information is provided, when available, to allow the reader to understand better how the activity should be developed: tentative year/s, level of intervention (EU, national, regional), by who, under which programme (if relevant), type of activity (coordination, RDI, etc.). A non-exhaustive list of outputs / deliverables that are envisioned along the different stages of development is then displayed together with a set of Key Performance Indicators that will allow the future EOSC association to monitor the implementation of the strategy.

The future final set of priorities will lead to activities that will be implemented by the different stakeholders at different levels. While some of the priorities will need to be translated into actions funded at EU level under Horizon Europe or possible other EU funding programmes, other activities will be designed to foster the commitments at national and regional level as well as at the institutional one. Regarding the actions at EU level, it is important to highlight that currently the evaluation of proposals on INFRAEOSC-03-2020 and INFRAEOSC-07-2020 is ongoing. Regarding the EOSC Strategy as it is expected that the projects resulting from these proposals will be carrying out some of the activities proposed in the following tables.

***NOTE:***

* *The numbering of the priorities (P1.1., P1.2., etc.) is not related to their level of importance or the timeframe for their implementation. Priorities are numbered so that they can be then recognized in the accompanying questionnaire.*
* *The 14 Areas of Action have been divided over the three Strategic Objectives. Several of them are relevant for more than one Strategic Objective. For sake of simplicity in the consultation process we have added them to only one of the Strategic Objectives.*

## Strategic Objective 1:

**Open science practices and skills are rewarded and taught, becoming the “new normal”**

|  |
| --- |
| **SO1 - Priorities** |
| ***AA5 - Users environments***  PA5.1: Advanced discoverability of portals  PA5.2: Meta catalogues to aggregate information from the resource catalogues of the service providers  PA5.3: The licenses, usage terms and conditions and user authentication and authorisation methods set by service providers  PA5.4: Legal and organisational framework and its implementation in the distributed architecture  PA5.5: Portals and other EOSC services support this kind of functionality |
| ***AA11 - Skills and training***  PA11.1: Developing the next generation of data/EOSC professionals.  PA11.2: Educating students and researchers.  PA11.3: EOSC to become a trusted and long-lasting knowledge hub of learning materials and tooling  PA11.4: Developing an EOSC leadership programme to foster the right policy environment for skills and training |
| ***AA12 - Rewards and recognition***  PA12.1: Draw up a national starting document outlining the most important items for evaluation  PA12.2: Discuss the item in (all) the institutions in the country  PA12.3: Discuss these development between the countries and learn from each other  PA12.4: Draw up as outcome of the discussions a new Rewards & Recognition system |
| ***AA13 - Communication***  [To be completed] |

***Potential outputs:***

**Stage 1 (2021-2023):**

* EOSC supports the training and deployment of professional data stewards together with European curriculum frameworks
* encourage the recognition and reward of Open Science, data skills and data stewardship among members of the EOSC

**Stage 2 (2024-2025):**

* Members of the EOSC Association recognise and reward adoption of FAIR, strongly incentivising data reuse.
* Researchers are incentivized to explore and implement FAIR-by-design practices in their laboratories / observatories / analytical resources.

**Stage 3 (2026-2027):**

* Provide alternative or new infrastructure that enables rewards and incentives for Open Science
* Support research institutions in capacity development for FAIR sharing within disciplines and for interdisciplinary research through interoperability frameworks.
* EOSC ecosystem facilitates and provides the capacity for Open Science for the majority of researchers of the EOSC Association members

***Potential KPIs:***

**Researchers doing publicly funded research make relevant results available, as openly as possible**

* KPI (2023) 90 % of publications from EOSC Association members are open access
* KPI (2025) KPI for software to be determined

**Professional data stewards available in research performing organisations in Europe to support Open Science:**

* KPI (2025) European curricula for data stewards are defined
* KPI (2027) All research done by EOSC Association members is supported by professional data stewards

**Researchers are incentivized to perform Open Science**

* KPI (2025) All association members endorse a commitment to recognize open science activities in research career assessments

**The scope of EOSC is widened to serve the public and private sectors**

* KPI (2025) a large majority of research funding members of the EOSC Association require data sharing and incentivise reuse.

## Strategic Objective 2:

**Standards, tools and services allow researchers to find, access and reuse results**

|  |
| --- |
| **SO2 - Priorities** |
| ***AA1 - Identifiers***  PA1.1: Prioritise Identifiers for: Instruments, Services, Organisations and Software, although there is a need for particular domains to provide their own community standards.  PA1.2: A “meta resolver” that can deal with any identifier type of interest.  PA1.3: Specifications (schemata) for PID records/kernel information to support machine actionable PIDs.  PA1.4: Type definitions for the most common data formats or building blocks.  PA1.5: Standardised Interfaces and protocols for exchanging information on PIDs to support the creation and use of a PID Graph.  PA1.6: Tools to support the certification of PID infrastructures against the EOSC PID Policy. |
| ***AA2 - Metadata & Ontologies***  PA2.1: Provide or embrace/stimulate existing registries with information about existing metadata schemas, ontologies, crosswalks, and tools for metadata management.  PA2.2: Develop guidelines for a minimum metadata description for data discovery and metadata exchange, based on existing metadata schemas and tools.  PA2.3: Develop services that build on metadata registries and can facilitate the diffusion of metadata schemas across communities, sharing and community maintenance of crosswalks, measure the uptake of metadata resources across communities, validating data sources against metadata schemes, etc. |
| ***AA3 - FAIR metrics & certification***  PA3.1: Support communities to clarify their requirements with respect to FAIR.  PA3.2: Test of the FAIR data maturity model in a wide range of communities, in a neutral forum and seeking for international agreement, to fine-tune and customize the recommendations, identify adverse consequences and apply corrections.  PA3.3: Thorough evaluation and comparison of assessment tools (manual, automated) and identification of their biases and applicability in different contexts.  PA3.4: Support communities to align certification schemas with FAIR.  PA3.5: Support data repositories and services to progress towards certification.  PA3.6: Support to data and service providers to progress in the FAIRness of their holdings.  PA3.7: Support to the definition of FAIR for software and of the assessment framework for key elements of the FAIR ecosystem, in particular PID services.  PA3.8: Definition and implementation of governance of the principles, assessment frameworks and metrics, adapted to each specific case |
| ***AA4 – Authentication & Authorisation Infrastructure:***  PA4.1: Conduct a gap analysis for the draft EOSC AAI Architecture  PA4.2: Evaluate use cases  PA4.3: Establish a first set of Rules of Participation  PA4.4: Produce final EOSC AAI Architecture as input for ongoing EOSC work on AAI |
| ***AA7 - EOSC Interoperability framework***  **Technical Interoperability**  PA7.1: Promote the use open specifications, where available, to ensure technical interoperability when establishing EOSC services.  PA7.2: Define a common security and privacy framework and establish processes for EOSC services to ensure secure and trustworthy data exchange between all involved parties.  PA7.3: Define an AAI process for EOSC that is common across communities, easy to implement by resource providers and easy to understand by users.  PA7.4: Service-Level Agreements for all EOSC resource providers are easy to understand by users from different communities.  PA7.5: Enable discovery of data sources available in different formats, either generic or community-based, to facilitate overcoming their heterogeneity and provide easy access and tools to integrate data across communities, enabling the usage of these data.  PA7.6: Tools for quality validation of metadata records and content of digital objects.  PA7.7: Search tools for coarse-grained and fine-grained datasets (and other research objects). There will be a range of general-purpose and domain-specific/specialised search tools, exploiting general-purpose and domain-specific metadata.  PA7.8: Need for a PID policy, accommodating any appropriate PID usage, recognising that established practices are at different levels of maturity for different resources and new PID types may emerge.  **Semantic Interoperability**  PA7.9: EOSC should provide support for the **maintenance of repositories of semantic artefacts, and governance frameworks** for such repositories, depending on common practices and stages of semantic resource development and usage of different communities.  PA7.10: EOSC should define clear protocols and building blocks for the **federation/harvesting of these repositories of semantic artefacts**.  PA7.11: Research communities  should be well supported (independently of their current state of semantic artefact adoption) so as to generate **clear and precise definitions for the terms** that they use, as well as **for their metadata and data schemas** (and incorporate those that they are already using) and **their documentation**. EOSC should provide support to make these definitions publicly available and referenceable by persistent identifiers for machine readability.  PA7.12: Urgent, additional resources (financial, but also skills and training) should be dedicated specifically to communities with less developed or with no community standards to mitigate the risk of the EOSC becoming inaccessible to the majority of researchers within academic institutions.  PA7.13: EOSC should propose a minimum vocabulary to allow **discovery over federated research artefacts (data, software, publications)** across scientific communities, and based onexisting metadata models (e.g., DCAT, DDI 4 Core, DataCite core schema). There should be some alignment among them, and this vocabulary should be extensible, so as to allow for disciplinary metadata that is typical for some research communities.  **Organisational Interoperability**  PA7.14: The **current set of rules of participation recommendations should be completed with aspects related to interoperability**. For instance, for data providers this may include asking explicitly that data is published according to specific data formats and/or vocabularies for a specific community.  PA7.15: The same is applicable to **services**, which may be recommended to ingest or output data according to such standardised data formats and/or vocabularies, and with their corresponding metadata, with some level of quality.  **Legal Interoperability**  PA7.16: Need for a common set of machine-readable licenses for all types of research outputs.  PA7.17: Need for metadata schemas for Service-Level Agreements. |

***Potential outputs:***

**Stage 1 (2021-2023):**

* Establish a common framework for managing user identity and access.
* Establish a mechanism for discovery of and access to data and services across the federated EOSC ecosystem.
* Further define FAIR for implementation within the different disciplines.
* Implement Interoperability Framework for FAIR sharing within disciplines and for interdisciplinary research to support FAIR Digital Objects.
* Interoperability frameworks are available to facilitate research collaboration and networking.
* Develop metrics for FAIR Digital Objects to measure and increase FAIRness.
* Provide assessment frameworks to certify that repository services enable FAIR.

**Stage 2 (2024-2025):**

* Adopt a standard mechanism for naming and locating data and services.
* Assemble the components of the EOSC architecture (AAI, APIs for access by machines service management, minimum metadata framework, open metrics, PID services, helpdesk, portal, etc.).
* Ensure a feedback mechanism to engage with users and further develop the EOSC Core and EOSC Exchange tools and services.
* Semantic and other technologies have been implemented to facilitate automated processing.

**Stage 3 (2026-2027):**

* The EOSC ecosystem integrates all types of digital objects, including publications, data and software.

***Potential KPIs:***

**Research data produced by publicly funded research in Europe is FAIR by design**

* KPI: The number of research communities with documented standards and protocols for data sharing and reuse increases by 20% during each stage of Horizon Europe
* KPI (2023) 40% of the repositories in the EOSC will have a certification (e.g. CoreTrustSeal); by 2027 60% will be certified.
* KPI (2025) significant amount of data are FAIR-by-design.
* KPI (2025) 60% of the members of the EOSC Association have policies which require FAIR to be implemented in project design via Data Management Plans
* KPI (2027) 70 % of research data from EOSC Association members that researchers deem to be of relevance are made as FAIR as possible, ideally open

**The EOSC interoperability framework supports a wide range of FAIR digital objects including data, software and other research artefacts.**

* KPI (2023) the EOSC Interoperability Framework is adopted by at least 5 major research infrastructures in Europe, enabling their data to be federated into EOSC; by 10 in 2025 and by 15 in 2027
* KPI (2023) data from EOSC cluster projects is made FAIR and can be accessed through EOSC.
* KPI (2025) data and other digital objects from a wide spectrum of disciplines are FAIR and available in EOSC.

**European research is increasingly discovered and reused across disciplines as a result of EOSC**

* KPI (2023) data and other digital objects from domains with interoperable data strategies are available through EOSC-2
* KPI (2023) a significant proportion of European research data is discoverable through EOSC.

## Strategic Objective 3:

**Sustainable and federated infrastructures enable open sharing of scientific results**

|  |
| --- |
| **SO3 - Priorities** |
| ***AA6 - Resource providers environments***  PA6.1: Onboarding of resources: Thematic classification from interoperability framework for metadata  PA6.2: Onboarding of resources: Definition of standards for EOSC compliance  PA6.3: Onboarding of resources: Automation, via APIs  PA6.4: Access to resources: Implement AAI recommendations  PA6.5: Access to resources: Portal service catalogue assessment  PA6.6: Composability of resources: Using semantic interoperability framework |
| ***AA8 - Rules of Participation***  PA8.1: Encouraging open science skills, recognising and encouraging practitioners, and rewarding open science   * + Encouraging the growth of new professions of research supporters and EOSC enablers by recognising certificates   + Requiring professional certificates as part of services’ EOSC-approved accreditation   + Encourage provision of professional data stewards to support research activities   + Encourage recognition of open science activities in research career assessments.   PA8.2: Consistent and wide (or increasing) application of the FAIR principless   * + Requiring data management planning and documented standards and protocols for data sharing and reuse including accessibility through EOSC   + Encouraging adherence to an open charter for data sets incorporating rich, community defined and FAIR metadata including granularity levels, versioning, provenance, sustainability commitments, quality goals and licencing terms   + Encouraging research publications produced to be open access   + Encouraging research data outputs to be as FAIR as possible and ideally open   + Requiring use of EOSC-recognised PID services.   PA8.3: Harnessing the potential of RDIs more fully and effectively   * + Requiring compliance with the EOSC service description template and onboarding processes   + Encouraging RDIs to work towards certification, with community standards, helping to produce FAIR outputs   + Encouraging compliance with the EOSC Interoperability Framework to ensure services achieve a minimum level of interoperability (semantic, legal, technical and organisational)   + Encouraging FAIR certification of repositories (e.g. with CoreTrustSeal)   + Standardisation of access policies   + Encouraging adoption of standards   + Requiring compliance with the EOSC AAI framework for managing user identity and access. |
| ***AA9 - Landscape monitoring***  *{to be processed further}*   * Existence of standardised national Open Science and FAIR data strategies, including the description of these policies (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Existence of central/national contact point for Open Science (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Existence of national policy on Open Access publishing and Open Access to publications, and if YES, does it include financial incentives and support schemes? (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Existence of national policy on Data and Services, and if YES, does it include Open Access to data including financial incentives and support schemes? (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Existence of national policy on Open learning including financial incentives and support schemes (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Existence of national, regional, or sectorial research evaluation schemes of universities and RPOs, which account for existing institutional implementation of Open Science principles and Open Access schemes (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Existence of reference to EOSC in current national, regional, or sectorial policies (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Progression of the of the institutional structure(s) at national level accountable for defining and implementing EOSC related policies and strategies including their hierarchical structure * Existence of EOSC dedicated funding streams or criteria in national funding mechanisms or programs (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Existence of dedicated funding streams or other meassures (Programs, Grant schemes, project support, financial and other incentives), which target the promotion and/or implementation of Open Science principles at institutional level (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Evolution (investment) of the backbone infrastructure(s) at national level, which is contributing to EOSC services (**KPI** - number of relevant infrastructures in the individual categories i.e. data infrastructures, e- infrastructures, HPC Infrastructure, data repositories, and other services) * Sustainable funding of operation cost for the backbone infrastructure(s) at national level, which is contributing to EOSC services (**Binary/ descriptive KPI:** YES/NO/In planning + text) * Evolution of the Open science mindset at the national or regional level (**KPI** - number of universities, public and private RPOs, thematic infrastructures, which have adopted OA policies, Number of OA repositories, …) * Impact indicators (**KPI** – e.g. share of OA users and publications per country, including users associated with industry, number of publicly available datasets, usage of data provided by global) |
| ***AA10 - Business models***  [To be completed] |
| ***AA14 - Widening to the General Public and Private Sector***  [To be completed] |

***Potential outputs:***

**Stage 1 (2021-2023):**

* An online dashboard will be operational to visualize the evolving landscape of policies, infrastructures, and initiatives relevant to the EOSC.
* Facilitate stepwise European alignment of key elements (such as national policies and standards) towards open and “FAIR by design” for all digital research objects.
* Ensure an efficient and smooth onboarding process for service providers.
* Monitoring system that will gather usage data, evidence and best practices valuable to academics, industry, the public sector and the policy-makers.

**Stage 2 (2024-2025):**

* Ensure that service providers that participate in ​EOSC-Exchange conform to predefined Rules of Participation.
* Establish a monitoring system for the uptake of the EOSC core services, and the access to EOSC exchange services.
* The EOSC ecosystem (e.g. data, software, services, tools) is equipped to serve entities from the private sector and public authorities so that they can access and re-use results of publicly funded research, innovating new data services.
* Develop models for the long-term sustainability of the EOSC-Core and EOSC-Exchange.

**Stage 3 (2026-2027):**

* Implement the essential additional functionalities dedicated to the requirements of end-users from the public and private sector.

***Potential KPIs:***

**EOSC is operational and provides a stable infrastructure supporting research addressing societal challenges**

* KPI (2023) core functions of Minimum Viable EOSC are developed and the EOSC ecosystem is accessible to researchers across disciplines and countries.
* KPI (2025) incentives and models for commercial providers to join EOSC Exchange
* KPI (2027) several commercial service providers provide research related services through EOSC
* KPI (2027) EOSC Core and EOSC Exchange become accessible to users from the private sector, the public sector and society at large.
* KPI (2027) significant number of tools and services from several national research infrastructures also become accessible through EOSC

**EOSC is populated with a valuable corpus of interoperable data**

* KPI (2025) significant amount of data from the major national research infrastructures is available through EOSC.
* KPI (2027) significant amount of data from publicly funded research is available through EOSC

**EOSC is a valuable resource to a wide range of users from the public and private sectors**

* KPI (2025) several use cases of researchers engaging with stakeholders from the public and private sector to address societal challenges.
* KPI (2027) EOSC in use by public sector outside of academia.

# Expected impacts

Realising the ambitions of the European Open Science Cloud will transform the broader research and innovation ecosystem along multiple dimensions (cultural, technical, organisational, educational, policy) and multiple levels (international, European, national, institutional). Ultimately, when the EOSC ecosystem will be fully deployed and when new research datasets, tools and services will become systematically FAIR-by-design, then researchers will be able to deliver research output such as data and software as they are delivering publications today. Publications, data and software will be shared as open as possible and accessed seamlessly so as to optimise reuse and research efficiency.

As presented in the Objectives Tree, in the beginning of this document the impact that these developments will have are major and will lead to the following benefits:

***Improved trust in science through increased openness and quality of scientific research in Europe***

By realising the objectives of the European Open Science Cloud, it is expected that a transformation in the research culture will take place in which open research will become “the new normal”. This will lead to widely recognized incentives for academics, industry and public services to share their data and other research digital objects, and improve data management training, literacy and data stewardship skills. This will ultimately help researchers to exploit high quality data in new and efficient ways, whether that be in the form of managing, sharing, computing, analysing or storing their data. This increased support at all stages of the research data lifecycle will help researchers make maximum use of both their data and time so that they achieve the best possible results.

The opening of the research data lifecycle will also positively lead to a change in the research incentives and rewards systems and structures. As more FAIR data are made available, the possibilities of rewarding researchers, not only for publishing their articles but also for publishing their data sets and analytical tools, increases. This opening, in turn, leads to a more meaningful monitoring and better reproducibility and validation of research results.

This will also allow researchers to produce more reliable science at the service of the society at large, and minimise the impacts of false claims or unsupported statements. Trust in scientific insights will thus increase.

***Stimulate the development of innovative services and products in Europe***

The implementation of the EOSC ecosystem data will enable European research to make its digital transition while ensuring transparency, reproducibility and societal impact. By providing seamless access to increasing volumes of research data, EOSC will stimulate the uptake of different services, both from public and commercial providers, that align with the principles of EOSC. Therefore, the Web of FAIR Data will provide the ideal ground for building a wide range of new innovative and value-added services (from visualisation and analytics to long-term preservation). It will be as transformative as the World-Wide Web has been to business and our everyday life.

EOSC will be instrumental in stimulating the European private sector, as for example, the cloud industry, that is willing to align to these principles while, at the same time, it ensures that European researchers remain in control of their data, stored in trusted and FAIR certified European repositories, and that scientific knowledge will stay “as open as possible as close as necessary”.

***Better quality and more productive research and innovation, addressing societal challenges***

“*The European Open Science Cloud will make science respond better and faster to what society needs*”

***Ursula von der Leyen***, president of the European Commission[[5]](#footnote-5)

These words of the European Commission president in the context of the COVID-19 crisis and the development of a COVID-19 data sharing platform[[6]](#footnote-6) by bringing together infrastructure, resources and data from different sources, are a good example of the potential impact that EOSC will have in addressing the societal challenges.

Scientists are increasingly being asked to help in developing solutions to the global societal challenges of the 21st century and must employ new and innovative technologies on research data to be able to deal with these complex and interconnected issues. EOSC will lead to a fundamental revolution in the way researchers, companies and administrations share and exploit research data, somewhat similar to how the internet revolutionised our sharing and exploitation of information**.** Ultimately, each and every scientist will do research differently than the way that it used to be performed. When the ecosystem of new tools and services will be available, and as many as possible FAIR-by-design new datasets will be generated, researchers will be able to deliver much more rapidly the outputs of each part of the research life cycle, including data and software, with the same level of precision as they are delivering publications today. For research teams and laboratories, publications, data and software will not be managed in a holistic way, as interrelated digital objects, in order to optimise the use of research results.

The impact of EOSC in the capacity of research to address the current and future challenges of the world, will be based on the following elements:

* **More multidisciplinary research through data sharing and cross usage of services**: within each discipline, the communication and collaboration between scientists will benefit from faster and seamless sharing of publications, data, software, services, tools and other digital research outputs.
* Better and faster data and results sharing will **strengthen collaboration among researchers and disciplines** and open opportunities for new levels of integration. For the cooperation between teams to address multidisciplinary challenges, the use of data and software across research silos will allow the exploration of new avenues as has never been possible before.
* **Increased added value of the services** in the EOSC ecosystem: A European-scale environment for computational, storage, data analysis and other data-related services and tools will facilitate scientific multi-disciplinary cooperation, leading to discoveries and solutions in key areas such as environment and health.

1. “Realising the European Open Science Cloud”, [Report](https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf) of the Commission High Level Expert Group on the European Open Science Cloud, 2016. [↑](#footnote-ref-1)
2. COM(2020) 66 final, European data strategy [↑](#footnote-ref-2)
3. https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020\_en.pdf [↑](#footnote-ref-3)
4. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018H0790> [↑](#footnote-ref-4)
5. <https://ec.europa.eu/commission/presscorner/detail/en/ip_20_680> [↑](#footnote-ref-5)
6. <https://www.covid19dataportal.org/> [↑](#footnote-ref-6)