

EOSC Multi-Annual Roadmap (MAR) 2025 and 2026–2027

Advisory notes

The Multi-Annual Roadmap defines a set of priorities for future investment in EOSC. Within this MAR we have highlighted activities for 2025 as it is expected that two Work Programmes will be developed - one for 2025 and one for the 2026-2027 period.

The priorities under each objective are not placed in order of value. They are thematically clustered on top. They have been labelled purely for ease of referencing in the consultation.

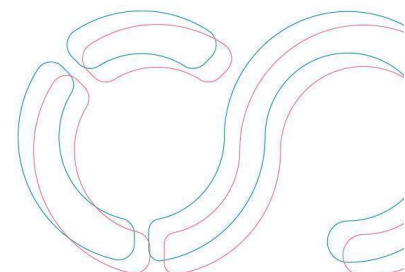
1. Objective 1 – Ensure that Open Science practices and skills are rewarded and taught, becoming the ‘new normal’

1.1 European level priorities

- A. Align on core curricula for data stewardship, Open Science and FAIR for both support staff and researchers. (2025)
- B. Support the development of networks for data stewards, research software engineers, semantic artefact curators and Open Science communities, leveraging on existing national networks and coordinating these at European level.(2025)
- C. Collect ongoing national best practices and case studies for the research and career assessment of open science practices in collaboration with the Coalition for Advancing Research Assessment (CoARA). (2025)
- D. Support and complement the CoARA movement with developing tools and processes for the research and career assessment of open science practices.
- E. Pilot indicators/metrics at the European level for the research and career assessment of open science practices and disseminate them Europe-wide.
- F. Support and embed appraisal and preservation, and monitor their value and costs. Practices should be transparent, aligned (i.e. differences for national or disciplinary reasons are justified) and interoperable (i.e. through machine actionable rights management).
- G. A dedicated provision for supporting the EOSC Partnership in its concertation actions. This covers stakeholder engagement, coordination, governance and monitoring. A cascading grant mechanism should support widening activities with under-represented countries and national competence centres, amongst other things.

1.2 National level priorities

- H. Support the implementation of aligned European curricula for data stewardship and encourage the inclusion of these as a core element in research programmes. The establishment of proper career paths for research support professionals is also needed.



- I. Establish the cost of data management, data stewardship, maintenance and preservation of research outputs (including software and semantic artefacts) as eligible within national funding schemes.
- J. Leverage existing national Competence Centres, such as those run by Skills4EOSC and the Dutch Digital Competence Centres, and strengthen their participation in coordination networks.
- K. Use cases from the public sector, citizen science and industry should be encouraged in the next period in parallel to the research sector. Also use cases running at national or regional level are interesting. Emphasis should be to support uptake (2025).

1.3 Institutional level priorities

- L. Offer core Open Science and FAIR training to researchers at all levels and recognise these skills as important for professional development.
- M. Support professional development programmes to ensure research support staff have the required data stewardship skills.
- N. Engage in competence centres and networks for data stewards, research software engineers, semantic artefact curators and Open Science communities to share models and harmonise best practises.
- O. Adjust, follow up, and evaluate research review mechanisms to ensure FAIR research outputs and Open Science are appropriately recognised and rewarded.
- P. Define and implement training and procedures to select data, software, and other research outputs that retain the value necessary to be preserved via EOSC.

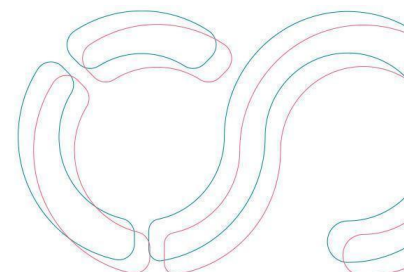
1.4 Expected outcomes

- Q. Research assessment in Europe is more harmonised, supporting Open Science as the new normal.
- R. Researchers and support staff have appropriate data stewardship skills and access to peer networks so a base-level of knowledge is reached across the European academic community, including countries where skill gaps exist.
- S. Researchers recognise the value of data stewardship and routinely include this as an element in grant proposals where possible.
- T. Researchers are supported by assessment procedures and metrics (fully recognized by their organisations) that incentivise open science practices.
- U. Research communities proactively engage in decisions about which data, software and other outputs to preserve for the long-term and guidelines are available per community.
- V. The EOSC Partnership transitions to a new governance model by 2027.
- W. Engagement in EOSC is widened resulting in integration of content from key data spaces, broader uptake from research performing organisations and new initiatives in previously under-represented countries.
- X. The value of EOSC is demonstrated via different (sectoral) use cases, increasing the range of applications supported and growing the user base.

1.5 Clarification text

EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu
Reg. number: 0755 723 931 | VAT number: BE0755 723 931



Data stewardship

Data stewardship curricula for both support staff and researchers should be aligned as much as possible at the European level to ensure a base standard of Open Science and FAIR practice is reached. The curricula should build on existing work conducted by EOSC projects, institutions and the global community, for example in the context of RDA Interest Groups. In particular we recommend an emphasis on semantic artefact curation to bridge between domain specific and domain agnostic skills. There is a lack of awareness and skills when it comes to the curation of semantic artefacts that future data stewardship curricula should address.

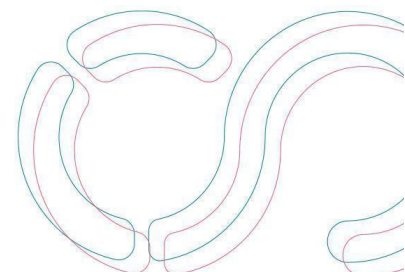
Both Member States and institutions have a responsibility to roll out the curricula and recognise the importance of these skills, particularly by establishing permanent positions and opportunities for career development for professional data stewards. Outreach programmes should target young researchers by connecting at the regional level. In addition to the curricula, professional networks for data stewards and a recognition of the associated costs as eligible in national grant schemes are also recommended to incentivise the change of culture. Building competence centres is key to knowledge transfer on Open Science and FAIR in the EOSC context. While the structure, operational mode and organisation of Competence Centres may vary very widely, they are usually associated with excellence, training, interdisciplinarity, standardisation and a collaborative approach that ensures harmonisation and alignment at European level.

Research assessment

In order to support the transition to Open Science and FAIR practice, the work done by researchers relevant for this transition need to be recognized and rewarded. New measures need to be defined and tested to assess uptake and ensure they lead to the desired culture change. A recent major development in the landscape is the coalition for advancing research assessment (CoARA). This initiative will entail a period of experimentation by coalition members for several years. This experimentation needs to be supported by further EOSC projects aiming at development, piloting, implementation, and medium-to-long-term evaluation of new methods for research assessment. Research funding and performing organisations, whether CoARA members or not, should implement new Open Science metrics in research review mechanisms.

Data preservation

Investment is needed to embed preservation systems and appraisal practices and to make them machine actionable for efficiency, interoperability and transparency. Disciplinary expectations for data and metadata preservation must be combined for cross-disciplinary needs. Transparency of appraisal ensures broad input on decisions to changing levels of preservation or transfer to alternate repositories. Retention and preservation should be transparent in digital object metadata as well as at the repository level. Semantic artefacts - as key knowledge organisation tools - also require special attention concerning long-term preservation.



Engagement and widening

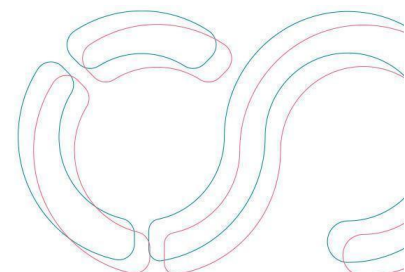
Expanding on use cases should be a key emphasis in this period to support uptake and widening of EOSC. The Minimum Viable EOSC (MVE) should be tested by different communities, including the data spaces and industry to validate applicability beyond academic researchers. Member State involvement is key to incentivise other sectors to trial EOSC.

A dedicated provision is also recommended to assist the EOSC Partnership in evolving its work with different stakeholders, specifically research performing organisations, the data spaces and countries less engaged in EOSC. The actions could be supported by cascading grants to facilitate uptake where it is lacking, or dedicated funding lines to support the countries and regions. Dedicated calls like WIDERA-CSA or other types of instrument are expected. Many countries have existing national open science strategies or initiatives and these should be cooperated with to align with EOSC. Activities in the dedicated provision should address topics such as competence centres, the transition to a new governance structure, service delivery models and monitoring of Open Science activities via the EOSC Observatory. A greater degree of funding should be directed towards the EOSC Association over time as part of sustainability plans when adopting responsibility for project outputs.

2. Objective 2 – Enable the definition of standards, and the development of tools and services, to allow researchers to find, access, reuse and combine results

2.1 European level priorities

- A. Promote and sustain the use of Persistent Identifiers (PIDs) that are already common practice. Support activities where PID usage is not yet a common practice.
- B. Advance technical specifications to support digital objects to be FAIR and evaluate viable options to adopt the concept for operational data services. The publication of FAIR digital objects should also be promoted and incentivised. (2025)
- C. Continue to invest in the creation, adoption and governance of community-based metadata and data standards to support the discovery, interoperability and reuse of research data and software.
- D. Continue to develop and maintain open interfaces, alignments, crosswalks, and APIs that enable interoperability and foster adoption of EOSC. The EOSC Interoperability Framework should facilitate the automated composition of EOSC data and services.
- E. Evaluate the viable options to provide European and international researchers with a platform for software through development, preservation and reuse, most likely as a federated network of software repositories.
- F. Focus on community-specific FAIR metrics and data quality constraints (especially on the interoperability of these) by engaging with research clusters to apply FAIR in their contexts. (2025)
- G. Support the emergence of a widely representative governance body around FAIR metrics and assessment.



2.2 National level priorities

- H. Promote and sustain the use of PIDs that are already common practice. Support activities where PID usage is not yet a common practice.
- I. Integrate widely used and adopted PIDs into national services. Incentivise usage of PID technologies being developed in EOSC (like PID Meta Resolver, Data Type Registry, PID graph, PID Policy Compliance Assessment Toolkit).
- J. Support and incentivize the development, maintenance and adoption of open standards and APIs to enable resource composability and to achieve interoperability across communities, including in the private sector.

2.3 Institutional level priorities

- K. Integrate widely used and adopted PIDs into institutional services and incentivise usage of PID technologies being developed in EOSC (like PID Meta Resolver, Data Type Registry, PID graph, PID Policy Compliance Assessment Toolkit).
- L. Support research communities to adopt both general and domain-specific standards to increase adoption of FAIR practices and reuse.
- M. Develop procedures and support for publishing semantic artefacts through institutional or vocabulary specific thematic repositories.
- N. Encourage sharing of software through institutional or thematic repositories.

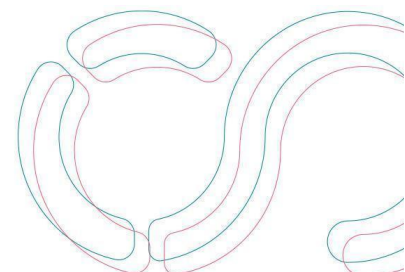
2.4 Expected outcomes

- O. A clear increase in the usage of PIDs and exploitation of links between them via typing, contextualisation and knowledge graphs to visualise and track research activity.
- P. The integration of PIDs into data services is enabled by supporting the interoperability of PID systems. This reduces the current overhead and confusion caused by re-registering the same research outcome at different PID providers.
- Q. Infrastructures for research software and source code are promoted at national and institutional level, increasing the amount of software and code accessible via EOSC.
- R. Metrics are defined to measure FAIR on a community level and evaluator tools enable reliable and trusted comparisons between these.
- S. Ensure the use of existing or community driven recommendations via assessments regarding standardisation and data quality.
- T. Cross-domain reuse of data increases due to higher levels of standards adoption and alignments/crosswalks between these.
- U. Quality control mechanisms for semantic artefacts and other emerging vocabularies (vocabulary recommenders) are more and more in place and increase quality.

2.5 Clarification text

PIDs

Persistent Identifiers (PIDs) are a critical component of the underlying infrastructure to allow resources to be accurately identified, accessed and linked to other objects. Many types of PID are already in common usage and this should be sustained via support at the European and national level, encouraging adoption by communities that are not already active. The



integration of different actors supporting PID services should also be coordinated at European level for sustainability. Support should be directed towards national and institutional services, enabling their adoption of PIDs to ensure content federated into EOSC is well contextualised. Services being supported at the European level such as PID Meta-Resolver, Data Type Registries, PID graph technologies and the PID Compliance Assessment Toolkit should also be incentivised by Member States and institutions to promote uptake and researchers should be encouraged to use PIDs for the whole research lifecycle, as well as for their research activities. The interoperability of PID systems should also be encouraged as objects often get assigned two or more different PIDs since few interactions are possible between the providers. As noted under objective 3, a long-term sustainability plan should also be developed within EOSC for PIDs to help the research community understand their long term financial commitment related with PID adoption.

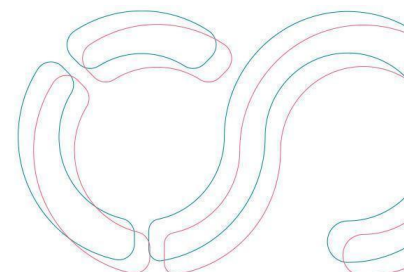
Interoperability

The absence of quality metadata continues to be a major obstacle to the discovery and reuse of research outputs. As such, continued investment in developing, evaluating and maintaining community-based metadata standards is recommended at the European level and work to encourage adoption is suggested within institutions. The adoption of standards is also critical to support interoperability. Work related to an EOSC Interoperability Framework is ongoing in EOSC Future and relevant parallel activities are planned in the SIMPL procurement for the Common European Data Spaces. The EOSC Interoperability Framework should be compatible with the SIMPL guidelines for the Data Spaces, requiring collaboration as both develop. To enable interoperability, work on the development of open interfaces, alignments, crosswalks and APIs must continue, and research communities need to be encouraged and supported to dedicate time to this important work. Interoperability should cover all the outputs of research as well as the services, enabling composability of resources so research workflows can run smoothly without manual intervention from the user. Research institutions also play a key role in encouraging researchers to adopt relevant standards and implement FAIR.

An emphasis has been placed on further support for FAIR digital objects since their implementation encourages the usage of PIDs and agreed metadata schemas, and as such facilitates interoperability. They provide a conceptual and implementation framework to develop scalable cross-disciplinary capabilities, deal with the increasing data volumes and their inherent complexity, build tools that help to increase trust in data, create mechanisms to efficiently operate in the domain of scientific assertions, and promote data interoperability. FAIR digital objects are identified via PIDs and can be regarded as contextualised sub-graphs of general scientific knowledge graphs (SKGs), which allow to account, describe and share everything about a particular research.

Source-code repositories

Further support is needed to support the development, archiving, sharing and reuse of research software. At the European level, we need to evaluate the viable options to provide European and international researchers with a platform for software development and sharing, most likely as a federated network of software repositories. A study should be



launched to build a Pan-European software catalogue. Stress should be put on reusability, and the use of standards that facilitate the deployment and uptake of software.

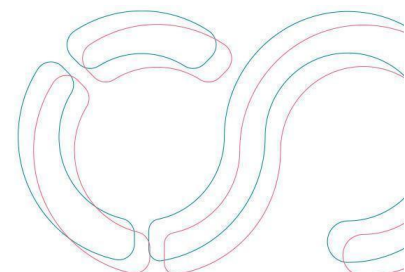
FAIR metrics

Building on the work conducted to define a canonical set of FAIR metrics based on a harmonized workflow for metadata harvesting, these should be tested in different communities to evolve community-specific metrics and evaluation procedures. The research clusters will play a pivotal role in this to help build community consensus. A sufficient degree of interoperability across community-specific metrics must also be achieved to be able to compare evaluations at an EOSC-wide level. Use of metadata schemas, standardised distribution formats, data models and controlled vocabulary enables data quality assessment. A community forum where tool providers could come together would help in this aim. To enhance community trust in FAIR Assessments, an emergent widely representative and transparent governance body around FAIR metrics should be supported and encouraged, such that the tools being applied to judge stakeholder's data holdings have undergone an open peer-review process.

3. Objective 3 – Establish a sustainable and federated infrastructure enabling open sharing of scientific results

3.1 European level priorities

- A. By the end of 2025 a sustainable initial EOSC Core should be up and running and thereafter further enhanced and extended to support emerging and challenging use cases. The new technologies implemented need to be continually tested to ensure EOSC is fit-for-purpose.
- B. Enhance the federation model of EOSC by defining a harmonised operational and legal framework to facilitate the sharing and the access of data and services across European countries.
- C. Develop a user driven mechanism to determine a selection of most needed horizontal and thematic services (2025) which can serve as a pilot for the centrally financed service portfolio for the period after Horizon Europe.
- D. EOSC should continue to utilise and build on existing AAI, as provided by national and regional AAI Federations. From 2027 onwards, EOSC will remain a major stakeholder in the pan-European AAI for research and education and will actively contribute with requirements, use cases and participation in standardisation activities.
- E. Continue to maintain and enhance the EOSC common search and access engine for FAIR research objects, including multilingual functionality and Search Engine Optimisation.
- F. Develop a maturity model for cross-domain semantic interoperability and support for semantic artefact catalogues.
- G. Dependable semantic artefact catalogues should be part of the EOSC core, akin to how vocabularies are published by the EC publications Office at EU Vocabularies.
- H. Compliance monitoring of the EOSC Rules of Participation shall be machine-actionable as much as possible to facilitate the scalability as well as reducing the cost of implementing and monitoring the rules.



- I. A transition towards machine-actionable repositories should be encouraged by adding this as a requirement to the EOSC Rules of Participation. (2025)
- J. Provide a PID sustainability plan for EOSC to help the research community understand their long term financial commitment related with PID adoption.

3.2 *National level priorities*

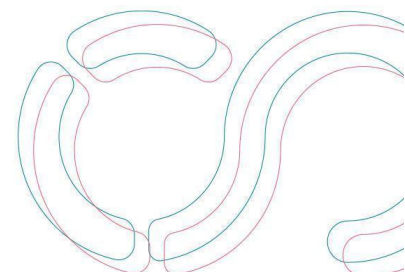
- K. Member States should work with Research Infrastructures and national service providers to establish and operate national EOSC support nodes to help integration with EOSC with regards to the provision as well as usage of data and services.
- L. Encourage Member States to review and adjust national policies and regulations to enable services to be used in a cross-border context, and engage in a series of in-depth trials to evaluate the utility of proposed resourcing models to agree approaches for EOSC.
- M. Member states should work with the National Research and Education Networks to ensure the AAI is available to national research infrastructures, resource providers and users and can be used within and across borders. (2025)
- N. Acknowledge European and domain specific semantic artefact catalogues in national infrastructures and guidelines.
- O. Support the establishment of last resort data repositories/stores to support European and national open access data policies.
- P. Support repositories to become machine-actionable for greater reuse of content via funding programmes and technical support actions. (2025)

3.3 *Institutional level priorities*

- Q. Research institutions should encourage and support researchers to adopt the data and services federated by EOSC to scale up usage.
- R. By 2027 all relevant institutions should be integrated into the National AAIs, enabling their users to collaborate seamlessly at the national and European levels. (2025)
- S. Make sure that researchers are aware of and reference existing European and domain specific semantic artefact catalogues.
- T. Encourage institutional repositories to become machine-actionable and recommend that research communities deposit data in machine-actionable repositories. (2025)

3.4 *Expected outcomes*

- U. Consolidation and further enhancement of the EOSC-Core, strengthens the federation aspects and integration leading to greater uptake by national, regional and institutional nodes, data spaces and other partnerships.
- V. Sustainability of EOSC is secured by co-investment at European and national levels with agreed cost-recovery models that incentivize participation of institutions and cross border activity, and by consolidating and harmonising EOSC-Core and the national nodes.
- W. Levels of FAIRness increase and more content is accessed by machines due to federated repositories supporting such interactions.
- X. Interoperability is enhanced due to a greater number of semantic artefact catalogues and maturity models to help the sector advance.



EOSC Federation and MVE

As noted in the *EOSC operations and evolution post-2027* reflection Paper, “EOSC should be a federation of existing and planned research data infrastructures, adding a soft overlay to connect them and making them operate as one seamless European research data infrastructure”. There are several basic elements of the Federation, including AAI, PIDs, Interoperability Frameworks, Rules of Participation and common search, all of which are addressed in this MAR. The annexes describing these are included as a reference on vision and terminology.

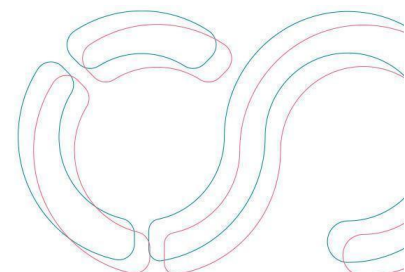
Much work has already been undertaken to develop the MVE and integrate Research Infrastructures into this. In addition, further work is planned via the forthcoming INFRAEOSC calls and EOSC procurement call, which intends to establish a European EOSC node to which national, regional and institutional nodes can connect. As such we expect to have an initial EOSC Core up and running by 2025. The emphasis in the 2025-27 work programme should therefore be to enhance and extend the other components of the MVE, namely the EOSC Exchange and the Data Federation, aiming for a completion of the MVE structure by 2026. Ongoing investment is needed to ensure continuity and sustainability. The EOSC Federation model should also be further enhanced by defining a harmonised operational and legal framework to facilitate the sharing and the access of data and services across European countries. Centres of expertise should also support stakeholders with onboarding.

Actions are highlighted at national and institutional level to encourage further integration by other nodes, research infrastructures and adoption by researchers. The sustainability question falls to the European and National level. It is foreseen that the EOSC Core needs to be centrally supported, however contributions to this could transition from European funds to national inputs. Outcomes will be dependent on the future governance structure of EOSC and nature of the legal entity.

To enable the adoption of the EOSC platform we recommend an EOSC-wide approach that provides all European-authenticated users with adequate and easy access to high demand services which can serve as a pilot for the centrally financed service portfolio for the period after Horizon Europe. Alternative means to grants could be explored and trialled to determine feasible options for the EOSC Exchange in the future. The timing needs to be aligned with the outcomes of the EOSC procurement call and forthcoming INFRAEOSC-01-04 call which will develop next generation services for the EOSC Core.

EOSC resourcing models

Enabling nationally funded facilities to be used in a cross-border context continues to be a challenge in the uptake of EOSC. While a variety of current projects and EOSC Association Task Forces are working on this issue to identify potential resourcing models to use, greater engagement is needed at the Member State level to review and where appropriate adjust



local policies and regulations. Support from Member States is needed to encourage inputs from national infrastructures and to permit facilities to be used in a cross-border context. The proposed resourcing models also need to be trialled and evaluated to determine which set of approaches to implement in the context of EOSC.

Authentication and Authorisation Infrastructure (AAI)

In the last two decades, research collaboration has been a driving force for the evolution of AAI in Europe and across the globe. With the work in the national academic federations, eduGAIN, and then AARC, the foundation has been established for building a robust EOSC AAI that can meet the requirements of scientific collaboration. With the development and implementation of AAIs across Europe and with the new building block provided by the European Commission and the Member States in eIDAS 2.0 and the European ID wallet, digital identities are becoming fundamental for the digital life of the European citizens. In the period 2025 - 2027, EOSC should expect that there is no further need to develop, operate and maintain its own AAI, but rather to be able to use pan-European AAI that will be provided by the National and Regional AAI Federations and which will meet all its requirements, enabling EOSC to focus on the seamless delivery of services. Efforts should be focused on activities at national and institutional level to ensure AAI is available to national infrastructures and resource providers and that all institutions are integrated into national AAIs.

EOSC search

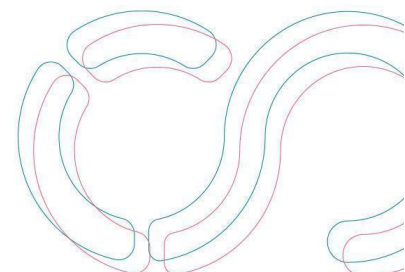
Although a cross-search has been developed for EOSC enabling data, software and other research resources from a variety of repositories to be collectively queried, more work is still needed to enhance search capabilities and expand the number of resources and types of research outputs included. In particular, Search Engine Optimisation, multilingual tools, and cross-domain mappings of search terms to facilitate uptake by researchers across Europe using their native languages are recommended.

Semantic interoperability and semantic artefacts

Projects such as WorldFAIR and FAIR Impact continue to work on domain standards and developing models for cross-domain interoperability. Within the forthcoming work programme, emphasis should shift to a maturity model for cross-domain interoperability, enabling and encouraging communities to increase the potential for reuse of data. Catalogues of semantic artefacts are also needed in addition to domain standards. Semantic artefacts are machine readable models of knowledge such as controlled vocabularies, thesauri, and ontologies which facilitate the extraction and representation of knowledge within data sets using annotations or assertions. Dependable semantic artefact catalogues should be part of the EOSC core and acknowledged in national infrastructure and guidelines to ensure researchers are aware of and using existing catalogues.

Rules of Participation

A first iteration of a compliance monitoring framework should be developed for the EOSC Rules of Participation. This should include the preparation of a certification framework, training materials and toolsets necessary for the experts to perform the machine-aided



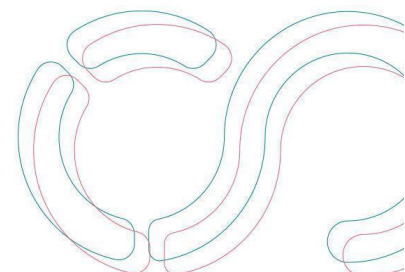
certification of EOSC resources and catalogues. Activities are also needed to increase awareness of the RoP among the community. This should include early engagement with the EOSC stakeholders to facilitate adoption of the RoP and to avoid them becoming a barrier to federation.

Machine-actionable repositories

In line with FAIR principles and data quality which are key fundamentals of EOSC, repositories should be machine-actionable so the data, software and publications they hold can be queried and reused at scale. Actions are needed at all levels to support this. At the European level, policies should encourage the transition to machine-actionable repositories by including requirements in the EOSC Rules of Participation. Support programmes are also needed at Member State level to enable national, thematic and institutional repositories to make the transition. Institutions also play a key role in encouraging research communities to choose machine-actionable repositories. It is also recommended to support the establishment of last resort data repositories/stores to ensure researchers always have a place of deposit.

Widening

In conjunction with the widening actions emphasised under objective 1, actions are needed at the European level by the operators of EOSC to target other data spaces and partnerships to be part of the federation and integrate into EOSC. The policy and governance actions led by the EOSC Partnership to enable collaboration with these stakeholders can support activity here.



Annex

This Annex comes from the EOSC operations and evolution post-2027 reflection Paper and serves as background information on the core concepts and vision for EOSC.

EOSC Federation

EOSC is expected to serve approximately 2 million researchers in Europe, many of them working at more than 800 European universities, and progressively to expand its user base to include the wider public sector and the private sector. An EOSC that offers added value to researchers was taken as a starting point with its scope as described in the Strategic Implementation Plan¹:

'EOSC should be a federation of existing and planned research data infrastructures, adding a soft overlay to connect them and making them operate as one seamless European research data infrastructure.'

The gradual expansion to the public and private sectors will create solutions and technologies that will benefit all areas of society, e.g. science, economy and education. Building on existing research data infrastructures, EOSC will grow through a series of iterations. Each iteration will add more functionality and services for a wider user base and satisfy a broader range of use cases, gradually adding extra value for the end users. These added values include primary scientific advantages, such as enhanced data and service connections, a better ability to address interdisciplinary and societal challenges, and improved e-infrastructure services and tools for RIs and their data consumers. Furthermore, persistent qualifying factors, such as transparency, high-quality data, research acknowledgement/credit and training, are important factors in each iteration.

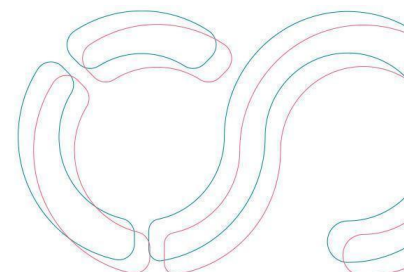
Thus, EOSC will essentially involve the federation of existing research data infrastructures and the realisation of a Web of FAIR Data and Related Services for Science. For this, the research data have to be made interoperable and machine-actionable following the FAIR guiding principles². This web of data will allow researchers to find, exploit and combine linked datasets, providing a basis for artificial intelligence (AI) tools, leading to new discoveries and research paradigms. EOSC will initially focus on traditional research data but will also include research publications and research code. EOSC will encourage FAIR datasets to be made fully open and will follow the principle of 'as open as possible, as restricted as necessary'. This is typically important for medical, military, private, and commercial datasets which may not be opened immediately or fully or indeed ever released.

Establishing such a federated infrastructure and enabling its transformation to a self-sustained, cross-border capacity to openly share scientific results, needs a collaborative concerted action of many stakeholders and a collaborative organisation with the appropriate mandate.

¹ European Open Science Cloud (EOSC) Strategic Implementation Plan', 2019, ISBN 978-92-76-09175-2, doi 10.2777/202370, KI-03-19-507-EN-N

<https://op.europa.eu/en/publication-detail/-/publication/78ae5276-ae8e-11e9-9d01-01aa75ed71a1>

² FAIR Principles: <https://www.go-fair.org/fair-principles/>



Elements of the Minimal Viable EOSC (MVE)

The Core

The Core is the prerequisite service component enabling the large-scale brokering of research data and services between the participants in the EOSC Federation³. This federating hub consists of various technology components and organisational processes which can be operationalised in various ways⁴ but should be viewed as one single element⁵ from a financing perspective, without which the EOSC Federation cannot exist as an operational reality. A large-scale federation comprising thousands of organisations requires operational stability over time of its federating core, which in turn necessitates stable financial boundary conditions. Both the Commission and Member States have a strong interest in maintaining strategic control over the EOSC Federation, ensuring inclusiveness, national relevance and synchronisation with other European and national initiatives. Thus, it befalls the Commission and Member states combined to ensure long term stable funding for the Core as a prerequisite component for the EOSC Federation. The required longer-term budget for the Core should come as financial rather than in-kind contributions.

Challenges related to the Core:

- Identify the appropriate vehicle and governance model for pooling and channelling EU, Member States and Associated countries funding to the entity(ies) providing the Core as an operational service
- Use of EOSC by everyone regardless of particular countries' financial contribution to the Core

The Data Federation

FAIR Data need to be federated in such a way that metadata on research outputs is harvested into a cross-search to enable greater discovery and reuse of data residing in multiple institutional, domain-specific and national repositories across Europe.

The Exchange (marketplace)

The Exchange is the EOSC Federation's pan-European marketplace for EOSC resources (research data and services). It enables the brokering of community services between Federation participants, gives access to procurement contracts⁶ with research-relevant commercial services and Federation-funded access to selected services⁷ of the two aforementioned categories where this is deemed in the overall interest of the EOSC

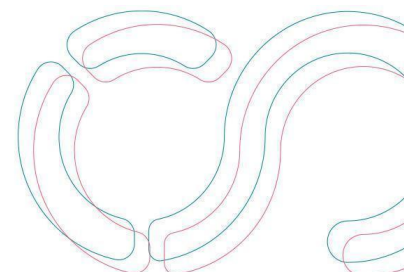
³ EOSC participation is assumed to be large-scale: consumption and production of data and services by millions of researchers through thousands of organisations in a very dynamic environment. Large-scale federations take time to build and impose longevity requirements on the federation infrastructure

⁴ See Appendix B Summary of Core components for details

⁵ That the Core as the hub of the federation should be viewed as one coherent entity from the point of view of the federation participants does not mean it is not dependent on other infrastructures, such as nationally deployed identity federations enabling seamless pan-European login

⁶ Access to procurement ready-to-use contracts for the use of commercial services

⁷ While the Exchange provides a standard brokering mechanism to make services available from a wide array of service providers to the 1000s of federation participants and millions of end users, it is not envisioned that the financing of all service consumption of all services will be channelled through the EOSC Federation



Federation⁸. It is open for discussion if the financing for the portfolio of funded services can be similar to the financing of the Core, except that for funded access to community services it could realistically also be implemented through in-kind contributions (i.e. service delivery). For the remuneration of access to brokered community services a limited number of standard remuneration mechanisms could be facilitated by the EOSC Federation, e.g. pay-for-use (financial transactions), earmarked grants, vouchers or virtual access. This will serve as a catalyst to lower the barrier of service provisioning and consumption. Current service aggregators (European, national and institutional level) are expected to play a role in the delivery of brokered services to the end user. The cost for the portfolio of funded services depends on the contents of the funded portfolio and the desired level of consumption funded directly through the EOSC-Federation (as opposed to funding consumption through project grants, RIs or national service providers).

Challenges related to the Exchange:

- National service providers and institutions need a mandate to provide services to organisations in other countries, and mandate of organisations to consume services from those in other countries
- VAT on transactions, VAT treatment of cross-border service provisioning are not the same in all relevant jurisdictions (EU, EEA, Associated countries)
- Procurement-free usage of community services offered through marketplace
- (Governance of) Portfolio management of collection of EOSC-Federation funded services

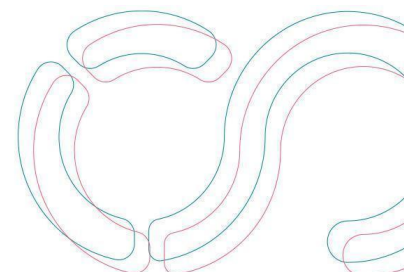
Basic elements of the EOSC Federation

The EOSC Authentication and Authorization Infrastructure

The goal of the EOSC Authentication and Authorization Infrastructure (AAI) is to manage user identity and harmonised access based on a common framework across the EOSC Federation. This is notably implemented by connecting multiple EOSC service providers to existing AAI systems. In practice, each EOSC service provider should sign an agreement, giving the power of attorney to an EOSC coordination structure responsible for the operations in the EOSC Federation. This new structure would be entitled to sign subsequent agreements with the involved Identity Federations. This construction avoids the cumbersome situation where each EOSC service provider would need to sign an agreement with each identity federation. The technical/operational responsibilities of the EOSC AAI registry service and proxy functions (protocol translations) could be delegated/distributed to suitable existing operator(s). Still, the ultimate EOSC AAI service coordination, ownership and accountability shall stay with the coordination structure in charge of the operations of the EOSC Federation.

The Framework for assessing compliance with the EOSC Policy on Persistent Identifiers

⁸ One can imagine different reasons for making a service “centrally” available through the EOSC Federation, e.g. ensuring a minimum basic service availability in the entire EOSC Federation’s geographic area regardless of decisions made on country or institution level, stimulating the uptake of new types of services, economies of scale etc.



Many types of Persistent Identifiers (PID) are already in active use in the international research domain. Numerous technologies, databases and cloud systems use sometimes local identifiers. To create a functioning PID ecosystem for the EOSC and to implement the FAIR principles, PID services should be interoperable with each other and between PID service providers, as well as across research infrastructures. The EOSC Policy for Persistent Identifiers⁹ seeks to accommodate mature and established PID practices, schemes, technologies and providers which have a global presence. PIDs in the EOSC Federation shall be used to a maximal extent to enable the management and analysis of FAIR data but also the publication, curation and tracking of digital research outputs. A flexible framework is required to assess compliance of PID service providers and PID minting organisations and services with the EOSC PID Policy.

The EOSC Interoperability Frameworks

The EOSC Interoperability Frameworks (EIF) set of standards and guidelines¹⁰ to support interoperability and composability of EOSC resources across borders and disciplines while respecting privacy and security¹¹. They act as the 'glue' to connect services and digital research artefacts from different EOSC providers across the federation. A robust governance and coordination structure is required to manage, update and promote the EOSC Interoperability Frameworks required to support the expansion of the EOSC Federation and the execution of cross-community use cases. The EOSC Interoperability Frameworks shall recognise that research infrastructures have been working on interoperability within and across their infrastructures for many years. The update of these frameworks shall happen in close interaction with the various research communities at international level and support, where necessary, the development, adoption and certification of relevant standards. A governance process is needed to define and apply the same criteria for inclusion for all interoperability guidelines that will be part of the EOSC Interoperability Frameworks.

The EOSC Rules of Participation

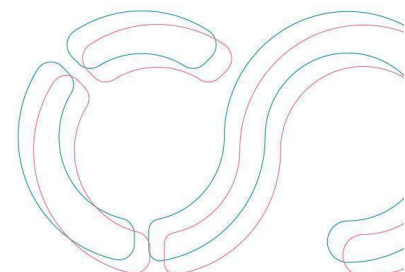
The EOSC Federation requires a minimal set of clear Rules of Participation (RoP) that shall define the rights, obligations and accountability governing all transactions by the various EOSC users, providers and operators across the federation. In this activity, current emphasis is on further operationalising the high-level general principles¹² released by the EOSC governance in 2020 and being updated in the context of the new EOSC-FOCUS project. The challenge is to set common requirements across the very heterogeneous European landscape of Research Infrastructures and service providers while preserving inclusiveness. The EOSC Rules of Participation shall guarantee an open, secure and cost-effective federated EOSC with services of documented quality. There will be room for differentiating the rules applicable to different EOSC users, service providers and contributing nodes at Community,

⁹ doi: 10.2777/926037

¹⁰ Of technical, policy, legal or administrative nature

¹¹ Such as copyright status, disclosure limitations, patents pending, other IPR on the datasets or workflows, the existence of personal data, designation of data as public sector information (PSI).

¹² doi: 10.2777/30541



European, National, Regional, or institutional levels. These rules will need to be maintained, enforced and revised according to user feedback.

The EOSC common search and access engine for FAIR research objects shared across the federation

To support cross-disciplinary research, European researchers navigating the EOSC Federation shall benefit from a mechanism for discovery of and access to FAIR research data shared by different communities across scientific disciplines. A flexible and scalable federated search and access service across the EOSC repositories shall exploit the PID meta-resolver and the EOSC Interoperability Frameworks (including EOSC guidelines for a minimum metadata description for data discovery and metadata exchange) currently developed in Horizon Europe. Such a discovery engine should ensure customisability of querying (e.g. interfaces, APIs and scripts) and, when possible, ranking of queries results according to FAIRness of data and other prioritisations responding to the users' needs. Such a data discovery and access engine shall provide open interfaces where data consumers (users and machines) are able to discover and use data. In some cases, controlled (authenticated and possibly authorised) access might be required to respect ethical, legal, social or commercial aspects.

The EOSC Security Coordination

An EOSC Security Incident Response Team (ESIRT) is required to manage the security baseline for the EOSC AAI Federation, provide security coordination across the federated services, ensure coordination and investigation of security incidents in all participating EOSC nodes/platforms¹³, coordinate security incident response activities between the EOSC Service Providers and provide guidelines for security risk assessment. The Security framework goes beyond an AAI Federation. This is an independent domain that relies on AAI for access control but also includes, encryption, repository security, intrusion detection and mitigation.

The EOSC monitoring and accounting modules

EOSC monitoring and accounting shall provide the capability to check the status, availability, and reliability of EOSC resources (research products and services) but also to track and record usage and usage satisfaction of those resources across the EOSC Federation based on a commonly agreed open metrics framework. This should also contribute to providing statistics regarding Open Science uptake (such as the production and/or reuse of open/FAIR research artefacts) which is a fundamental driver for the EOSC and the European Research Area (ERA). This shall also bring the possibility to define new indicators based on citations, usage data and links or by analysing content of the EOSC resource catalogue graph. This monitoring component shall link to the EOSC Observatory which shall continue to gather and present results of the monitoring of EOSC readiness, implementation, and uptake for Member States and Associated countries and the level of the EOSC European Partnership.

¹³ Being thematic or generic and at any level (European, national or institutional)

