

EOSC Declaration

Brussels, 26 October 2017

European Open Science Cloud
New Research & Innovation Opportunities



The EOSC Declaration and its principles, guiding the implementation of the EOSC, are the tangible result of the EOSC Summit of 12 June 2017. They have been endorsed by the undersigning stakeholders, found in the List of Signatories, who also committed to specific actions to implement it (Action List). As such, the Declaration does not commit the European Commission and Union institutions.

EOSC Declaration

RECOGNISING the challenges of data driven research in pursuing excellent science;

GRANTING that the vision of European Open Science is that of a research data commons, widely inclusive of all disciplines and Member States, sustainable in the long-term,

CONFIRMING that the implementation of the EOSC is a process, not a project, by its nature iterative and based on constant learning and mutual alignment;

UPHOLDING that the EOSC Summit marked the beginning and not the end of this process, one based on continuous engagement with scientific stakeholders, the signatories,

ENDORSE the following intents and will actively support their implementation in their respective capacities:

Data culture and FAIR data

- **[Data culture]** European science must be grounded in a common culture of data stewardship, so that research data is recognised as a significant output of research and is appropriately curated throughout and after the period conducting the research. Only a considerable cultural change will enable long-term reuse for science and for innovation of data created by research activities: no disciplines, institutions or countries must be left behind.
- **[Open access by-default]** All researchers in Europe must enjoy access to an open-by-default, efficient and cross-disciplinary research data environment supported by FAIR data principles. Open access must be the default setting for all results of publicly funded research in Europe, allowing for proportionate limitations only in duly justified cases of personal data protection, confidentiality, IPR concerns, national security or similar (e.g. 'as open as possible and as closed as necessary').
- **[Skills]** The necessary skills and education in research data management, data stewardship and data science should be provided throughout the EU as part of higher education, the training system and on-the-job best practice in the industry. University associations, research organisations, research libraries and other educational brokers play an important role but they need substantial support from the European Commission and the Member States.
- **[Data stewardship]** Researchers need the support of adequately trained data stewards. The European Commission and Member States should invest in the education of data stewards via career programmes delivered by universities, research institutions and other trans-European agents.
- **[Rewards and incentives]** Rewarding research data sharing is essential. Researchers who make research data open and FAIR for reuse and/or reuse and reproduce data should be rewarded, both in their career assessment and in the evaluation of projects (initial funding, review of performance and impact). This should go hand in hand with other career policies in universities and research institutions (appointments, promotions etc.).
- **[FAIR principles]** Implementation of the FAIR principles must be pragmatic and technology-neutral, encompassing all four dimensions: findability, accessibility, interoperability and reusability. FAIR principles are neither standards nor practices. The disciplinary sectors must develop their specific notions of FAIR data in a coordinated fashion and determine the desired level of FAIR-ness. FAIR principles should apply not only to research data but also to data-related algorithms, tools, workflows, protocols, services and other kinds of digital research objects.

- **[Standards]** The EOSC must be underpinned by minimal and rigorous global standards for open research data, as well as standards for EOSC based services for collaboration through the EOSC (e.g. to facilitate inter-disciplinarity and avoid fragmentation). These standards (technical, semantic, legal and organisational) must combine long-term sustainability with optimal freedom of local implementation. They should be jointly defined by the research communities, taking into account existing instruments (e.g. EU Rolling Plan on ICT Standardisation). Cross-disciplinary agreements/protocols will lead to specific standards, inspired directly by relevant domain specific needs. Variations across scientific disciplines and their specific efforts of making research data open and FAIR should be respected.
- **[FAIR Data governance]** The design and implementation of FAIR principles must be built upon inclusive stakeholder participation (e.g. researchers from different scientific disciplines, EU Member States and the European Commission). Policy will go hand in hand with the implementation of technical and human resources, and a social infrastructure including education and training. To make FAIR data a reality, it is imperative to engage stakeholders and relevant multipliers, based on a solid stakeholder engagement strategy, on inter-institutional arrangements, well-established frameworks and decision making flows. Data governance needs to be agreed upon and the division of responsibilities be charted, ensuring transparency, representativity and accountability. European and national scientific research organisations, publishers and other actors must align their data-related business processes, responsibilities and expectations to achieve commonly agreed goals.
- **[Implementation & transition to FAIR]** Implementation of FAIR principles requires careful prioritisation and orchestration. The FAIR Data Action Plan 2018-2020 is an important collaborative instrument for the embedding of FAIR principles in the first phase of the EOSC. The plan will not necessarily suggest any specific technology, standard or implementation solution. For an even transition of data from different levels of maturity to FAIR, existing activities to make data FAIR (e.g. GO-FAIR) must be complemented by new initiatives that embed FAIR principles in all the phases of data life cycle.
- **[Research data repositories]** Trusted research data repositories play a fundamental role in modern science. Scientist must be able to find, re-use, deposit and share data via trusted data repositories that implement FAIR data principles and that ensure long-term sustainability of research data across all disciplines. Data repositories must be easy to find and identify, and provide to users full transparency about their services.
- **[Accreditation/certification]** Scientists must be assured that the European and national scientific research infrastructures where they deposit/ access data conform to clear rules and criteria (e.g. certified) and that their data is FAIR compliant. An accreditation or certification mechanism must be set in place based on agreed processes and an accreditation or certification body must maintain an up-to-date and accessible catalogue of certified repositories. Experience from existing accreditation processes must be taken into account.
- **[Data Management Plans]** A key element of good data management is a Data Management Plan (DMP); the use of DMPs should become obligatory in all research projects generating or collecting publicly funded research data, based on online tools conforming to common methodologies. Funder and institutional requirements must be aligned and minimum conditions for DMPs must be defined. Researchers' host institutions have a responsibility to oversee and complete the DMPs and hand them over to data repositories.
- **[Technical implementation]** While FAIR data must be implemented as part of good data governance at this highest possible level (e.g. certification, institutional implementation and support, as describe above), researchers also need handy tools to make data FAIR. These include:

[Citation system] A data citation system should be put in place to reward the provision of excellent open data. This will assist both the assessment of researchers and their

projects, and help implementing the findability, accessibility, interoperability and reusability of research data.

[Common catalogues] There must be catalogues (e.g. for datasets, services, standards) based on machine readable metadata and identifiable by means of a common and persistent identification mechanism that will make research data findable via an 'EOSC Portal'.

[Semantic layer] Research data must be both syntactically and semantically understandable, allowing meaningful data exchange and reuse among scientific disciplines and countries.

[FAIR tools and services] Easy access must be available to a common set of FAIR tools and services, to guide the curation of FAIR data for re-use and to assess FAIR compliance.

- **[Data expert organisations]** The Research Data Alliance, CODATA, DDI Alliance and other organisations active in the research communities must be used as forums to reach consensus on practical implementation of FAIR data principles at European and global level.
- **[Legal aspects]** It is essential for the success of EOSC to clarify and address the legal uncertainty of Open Access to research data, as well as the correct legal implementation of the FAIR principles. Legal barriers to access and reusability of research data must be identified and overcome and the underpinning legal framework must be made simpler and more coherent. Conversely, issues of ownership must be addressed, particularly where institutions have created services and resources. All these measures should allow easier integration of research data across different legal frameworks, policy implementation plans and strategies.

Research data services and architecture

- **[EOSC architecture]** The EOSC will be developed as a data infrastructure commons serving the needs of scientists. It should provide both common functions and localised services delegated to community level. Indeed, the EOSC will federate existing resources across national data centres, European e-infrastructures and research infrastructures; service provision will be based on local-to-central subsidiarity (e.g. national and disciplinary nodes connected to nodes of pan-European level); it will top-up mature capacity through the acquisition of resources at pan-European level by EOSC operators, to serve a wider number of researchers in Europe. Users should contribute to define the main common functionalities needed by their own community. A continuous dialogue to build trust and agreements among funders, users and service providers is necessary for sustainability.
- **[Implementation]** Resources, components and initiatives of pan-European relevance will be federated on the basis of objective criteria, agreed by stakeholder-driven governance, such as organisational readiness and technical capacity to deliver EOSC main functionalities: provision of core common services, certification activities, joint-procurement initiatives, definition of minimum quality standards of service (based on clear Service Level Agreements SLAs), identity provisioning and management, common cataloguing data and computing/analytic services and tools.
- **[Legacy]** The EOSC should incentivise the re-use of existing building blocks, state-of-the-art services and solutions delivered by past and ongoing projects, local, national and European, as opposed to subsidizing actions aiming at reinventing the wheel. It should facilitate learning from the past, adopting best practices, tailoring scientific community needs through live use cases and leveraging the network effect.

- **[User needs]** Users should see the EOSC as a one-stop-shop to find, access, and use research data and services from multiple disciplines and platforms. Services and functionalities shall be user driven and determined by clear use cases. Intermediary users and other brokers of end-users' demand – IT departments, umbrella associations, community networks – should assist data scientists and ICT specialists in the identification of key requirements for EOSC services.
- **[Service provision]** Research Data Infrastructures, e-infrastructures and commercial operators will develop and provide services based on user needs, and discontinue provision when not justified by the level of adoption. Services will be offered at highest Technology Readiness Levels (TRLs) and kept future-proof based on a cutting-edge cloud based environment. In order to avoid lock-in by individual service providers, the EOSC should foster fair competition of public, PPP and private providers on clear value propositions of highly professional services.
- **[Service deployment]** The EOSC shall support different deployment models (e.g. Infrastructure as a Service, Platform as a Service, Software as a Service), to meet the needs of communities at different levels of maturity in the provision and use of research data service. The EOSC shall support the whole research lifecycle by strong development at platform level that facilitate the provision of a wide set of software, infrastructure, protocols, methods, incentives, training, services. Software sustainability should be treated on an equal footing as data stewardship.
- **[Thematic areas]** The EOSC shall promote the co-ordination and progressive federation of open data infrastructures developed in specific thematic areas (e.g. health, environment, food, marine, social sciences, transport). The EOSC will implement a common reference scheme to ensure FAIR data uptake and compliance by national and European data providers in all disciplines.
- **[Research infrastructures]** The role of ESFRI and EIROFORUM research infrastructures and organisations in the EOSC will be enhanced, Member States and the European Commission made significant investment; research infrastructures should be 'the steward of the community of standards' and provide scientists with a ramp-up for the utilisation of the EOSC.
- **[EU-added value and coordination]** The EOSC must implement policy hand in hand with technology. Condition of national and European measures is required to link the initiative to national strategies, to maximise the added value of inter-disciplinarity by making data FAIR, and to preventing duplication of efforts and investments. Over time, coordination will provide European added value by minimizing overlap and reducing fragmentation of infrastructures and services, helping long-term sustainability.
- **[High Performance Computing and the EOSC]** European commitment to HPC is clearly demonstrated by the signature of the EuroHPC Declaration by twelve Member States(*) and one Associated State(**) since March 23, 2017. The Member States agreed to work together and with the European Commission in the context of a multi-government agreement called EuroHPC for acquiring and deploying by 2022/2023 a pan-European integrated exascale supercomputing infrastructure that will support data-intensive advanced applications and services. It is a response to the surging demand from scientists, industry and the public sector for access to leading-edge computing capacity to cope with vast amounts of data produced in almost all scientific and engineering domains. This supercomputing and data infrastructure could support the European Open Science Cloud by providing data access and advanced computing and data management services. The EC plans to adopt in January 2018 a proposal for the establishment of a legal instrument (the EuroHPC Joint Undertaking) that provides a procurement framework for the exascale supercomputing and data infrastructure, and supports the development of a full European HPC ecosystem.

(*)Belgium, Bulgaria, Croatia, France, Greece, Germany, Italy, Luxembourg, Netherlands, Portugal, Slovenia, Spain

(**)Switzerland

- **[Innovation]** The EOSC should create a level playing field for businesses and innovative SMEs to develop, and co-develop with publicly funded institutions, added-value services for researchers. Funding should support the migration of cutting-edge solutions to the EOSC, increasing European added value by fostering innovation.

Governance and funding

- **[Governance model]** A long-term, sustainable research infrastructure in Europe requires a strong and flexible governance model based on trust and increasing mutuality. As interdisciplinarity is one of the main objectives of the EOSC, the governance model should be based on representativity, proportionality, accountability, inclusiveness and transparency.
- **[Governance framework]** The EOSC governance framework will be co-designed, stakeholder driven and composed of three main layers: 1) institutional, including EU Member States and European Commission 2) operational, including an executive board and relevant working committees (e.g. thematic and functional) and 3) advisory, including a stakeholder forum.
- **[Executive board]** An executive board will coordinate the efforts of stakeholders endorsing the EOSC Declaration, with the broad mandate to reach practical agreements for the implementation of an EOSC Roadmap by 2020. The board will have an advisory role and an implementing role of decisions concerning the programming, financing and towards the setting up of a long-term governance and business model for the EOSC. It will make best use of the outcomes of past and current projects (e.g. EOSCpilot, eInfraCentral and EOSChub) and independent expert advice and studies.
- **[Coordination structure]** A coordination structure, funded by Horizon 2020, will help the executive board to manage the implementation, according to agreed rules and methods of stakeholder participation. The structure and its participating entities should be accountable for the responsibilities assumed, based on an objective assessment of their level of readiness in delivering the EOSC main functionalities.
- **[Long-term sustainability]** Research Funders will use existing and future resources strategically, to ensure long-term sustainability of open research data and research infrastructures, facilitating inter-disciplinarity.
- **[Funding]** Over time, a co-funding mechanism mixing different revenue streams for the EOSC will be set up, to increase the accountability of the governance, building trust, sharing resources and building long-term capacity for European research data. Early implementation of the EOSC will pilot innovative business models and support an integrated data and service platform for European research.
- **[Global aspects]** The EOSC will be European and open to the world, reaching out over time to relevant global research partners. It will increase the global value of open research data and support stakeholder engagement, including researchers and citizens. It will gradually widen the initiative to federated network of infrastructures and nodes from global research partners. The EOSC Stakeholder Forum will have an important role in this sense.

 <http://ec.europa.eu/research/openscience/eosc>

 #EOSC