Check for updates

OPEN LETTER

REVISED The European Open Science Cloud as a common good

Potentials and limitations of this endeavour

[version 2; peer review: 3 approved]

Anna Bertelli, Melania Acciai, Giorgio Rossi

University of Milan Department of Physics Aldo Pontremoli, Milan, Lombardy, Italy

v2 **First published:** 22 Jan 2025, **5**:19
https://doi.org/10.12688/openreseurope.19170.1

Latest published: 11 Apr 2025, **5**:19
https://doi.org/10.12688/openreseurope.19170.2

Abstract

The European Open Science Cloud (EOSC) is envisioned as a transformative platform for advancing Open Science, aimed at benefiting a diverse array of stakeholders, including researchers, innovators, institutions, and the broader public. To fully harness EOSC's potential as a common good, capable of delivering services to the research community such to potentially transform the way scientific production and communication is done, we address critical barriers that may actually restrict the equitable access and the optimal use of such services. In particular, we emphasize that key resources as required to access and exploit EOSC's advanced FAIR-data services – such as data-processing algorithms – are, in fact, intrinsically limited and the access will be competitive. Governance and funding of EOSC present challenges associated with its effective openness in terms of accessibility to resources for its advanced exploitation.

Keywords

European Open Science Cloud, FAIR Data, AI, Science Policies and Funding, Common Good, Inequality.

Open Peer Review

Approval Status ✓ ✓ ✓

	1	2	3
version 2 (revision) 11 Apr 2025			✓ view
version 1 22 Jan 2025	✓ view	✓ view	✓ view

1. **Sonia Stefanizzi**, Università di Milano Bicocca, Milan, Italy
2. **Kristijan Armeni**, Johns Hopkins University, Baltimore, USA
3. **Rodrigo Vargas Honorato**, Utrecht University, Utrecht, The Netherlands

Any reports and responses or comments on the article can be found at the end of the article.

HE This article is included in the Horizon Europe gateway.

Corresponding authors: Anna Bertelli (anna.bertelli@unimi.it), Melania Acciai (melania.acciai@unimi.it)

Author roles: Bertelli A: Conceptualization, Investigation, Writing – Original Draft Preparation, Writing – Review & Editing; **Acciai M**: Conceptualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Rossi G**: Conceptualization, Investigation, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101058092 – HORIZON-INFRA-DEV-01 (Support to Reinforce the European Strategy Forum on Research Infrastructures 3 [StR-ESFRI3])
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Bertelli A *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License , which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Bertelli A, Acciai M and Rossi G. **The European Open Science Cloud as a common good Potentials and limitations of this endeavour [version 2; peer review: 3 approved]** Open Research Europe 2025, **5**:19 https://doi.org/10.12688/openreseurope.19170.2

First published: 22 Jan 2025, **5**:19 https://doi.org/10.12688/openreseurope.19170.1

REVISED **Amendments from Version 1**

In the latest version, emphasis has been placed on accentuating the distinctions between the products offered by EOSC (data and services). The definition of the Federation of Nodes has been refined, and useful references have been appended to AI-related energy costs and consequences. Additionally, the "fair share scheduler" was mentioned as a potential mechanism for the equitable distribution of infrastructure resources. Finally, we have removed an ambiguous hint to potential funding schemes that necessitated further study and can be adequately addressed elsewhere.

Any further responses from the reviewers can be found at the end of the article

Introduction

The European Open Science Cloud (EOSC) is a pioneering initiative intended to enhance European research by facilitating open access to data, analytical tools, and advanced algorithms for data processing and analysis. By embracing the FAIR (Findable, Accessible, Interoperable, Reusable) principles (Wilkinson *et al.*, 2016), EOSC aims to foster a high-quality research ecosystem that supports both Open Science practices and advanced data services, essential for driving collaborative, transparent, and reproducible research.

EOSC faces the challenge of meeting the high expectations of the research community and policymakers, who view it as a catalyst for transparent, unbiased research and a driver of innovation across diverse research domains. EOSC is expected to support both curiosity-driven and mission-oriented investigations, making high-quality data and analytical tools accessible across fields.

As an instrument to promote Open Science, a critical objective of EOSC is to prevent "any particular actor from locking research into a proprietary, consolidated environment, which would recreate, at a different level, the enclosures of knowledge that Open Science seeks to eliminate" (Mounier, 2022, p. 3).

To ensure that EOSC can fulfil its vision of inclusive and equitable access to research resources, it is essential to structure it in a way that maximises societal benefit. This will require comprehensive assessments of both scientific and economic impact, as well as a sustainable funding model that encourages broad use while minimising access barriers.

In this paper, we analyse the current efforts to establish EOSC as a tool to accelerate the implementation of Open Science. The EOSC should uphold the idea of science as a global common good, effectively contributing to the universally available, accessible, and usable character of knowledge. However, to benefit from EOSC it is necessary to access data processing services, including high power computing, that are intrinsically finite and rivalrous. These issues are explored in detail and potential solutions to address these challenges, promoting an inclusive and effective use of EOSC are proposed.

EOSC as a common good for science

The European Open Science Cloud (EOSC)[1] is an ambitious, long-term initiative launched in 2015 by the European Commission, aimed at developing a web of FAIR Data and Services to support science across Europe. This initiative is, therefore, the result of a deliberate political will, responding to the growing need within the scientific community to reflect on its practices and limitations, and to align science increasingly with the principles of Open Science. It is also meant to support social and economic innovation in accordance with the European Research Area (ERA)[2] policies and the Sustainable Development Goals (SDGs)[3].

The term Open Science (OS) has become increasingly common in conferences, universities, and policy discussions related to science. But the exact definition of Open Science remains vague and somewhat a "moving target". Following Fecher and Friesike (2013), OS is an umbrella term that captures the emerging trend and need within the scientific community, as well as in critical literature on scientific practices, to align science more closely with the idealized image we intuitively associate with it. In other words, it encapsulates all advocacies for a science that is more open – specifically, more accessible, collaborative, interdisciplinary, transparent, and inclusive.

Why Open Science is good for science

The emergence of OS can be attributed to various factors, all rooted in scientific practice. Broadly speaking, these factors can be categorized into pragmatic reasons and value-driven reasons, with significant interconnection between the two[4].

The first group of reasons emphasizes how the scientific enterprise as a whole can benefit from implementing practices of transparency, openness, and collaboration. The second group illustrates how a more open science practice aligns more closely with the ideals and expectations of science as a democratic enterprise, with results that are equally accessible and usable by all.

The EOSC initiative has the potential to advance the scientific community on all the aforementioned fronts. That it is an initiative capable of increasing productivity, efficiency, and transparency within the scientific community is already evident, even at the current embryonic stage of network of FAIR data and data services, offering a clear entry point and a searchable catalog.

[1] European Open Science Cloud (EOSC): https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science/european-open-science-cloud-eosc_en

[2] European Research Area (ERA): https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/european-research-area_en

[3] The United Nations' Sustainable Development Goals (SDGs): https://sdgs.un.org/goals

[4] For a review of the arguments in favour of open science in the literature, see Vicente-Saez & Martinez-Fuentes (2018) and Fecher & Friesike (2013).

On the other hand, its potential to enhance the real and perceived trustworthiness of science, as well as the accessibility and democratic nature of scientific practice, is both highly likely and desirable, provided that certain potential challenges in the rapidly evolving technological setting are addressed effectively. That it is desirable is clearly stated, as an example, in the EOSC Steering Board's Policy Paper on "Advanced Digitalisation of Research":

> Research reproducibility and data usability will be enhanced if the advanced digitalization of data collection, validation, analysis, and simulation will be developed and become commons of the research community enforcing the Open Science principles and policies, creating a critical mass of Quality Assessed FAIR Data (QAFAIRD) and research-objects enabling reliable and secure Artificial Intelligence, Machine Learning and Virtual Research Environments (European Commission, 2024, p. 2).

The risks associated with the implementation phase of EOSC, however, need adequate attention and analysis.

The construction of the EOSC is currently envisioned as a Federation of Nodes, initiated by the European Commission through the procurement and launch of the EU-Node[5] as the foundational cornerstone. The EOSC Federation is defined as a "system of systems", where the federated components, or nodes, can either be purpose-built, like the EU-Node, or derived from existing research data systems, such as those developed and operated by Research Infrastructures and e-Infrastructures, or by other research data service organisations of national or institutional character. The nodes will represent ensembles of FAIR-data services (archives, search and retrieval software, quality assessment, cybersecurity, statistical tools) interoperable with other federated nodes to provide a seamless system for researchers to store, share, process, analyze and reuse FAIR research data. The EOSC Federation does not yet have a formal constitution; instead, it is evolving through a "learn-by-doing" approach through the aggregation of nodes that should find a first configuration by the end of 2025. This iterative process will lead to the development of a shared set of principles and practices, as outlined in a Handbook[6], which is intended to be a dynamic reference document.

It is at this early stage of development of the EOSC Federation that that the associated risks connected with inclusiveness, sovereignty, robustness, security, and trust need to be understood and addressed.

EOSC as a common good for society at large

The EOSC is designed to serve as a powerful instrument for open collaboration, transparent data lifecycles, and support for social and economic innovation. In this sense, we could say that the EOSC can be seen as an initiative aimed at the benefit of society, thus characterising itself as a common good. The concept of the "common good" can be understood through a philosophical lens:

> […] the common good is best understood as part of an encompassing model for practical reasoning among the members of a political community. The model takes for granted that citizens stand in a "political" or "civic" relationship with one another and that this relationship requires them to create and maintain certain facilities on the grounds that these facilities serve certain common interests (Hussain & Kohn, 2024).

This definition, however, is quite generic; moreover, some other times, the EOSC is defined instead as a "public good". For instance, in the "Opinion paper on EOSC and commercial partners" by the EOSC Steering Board expert group, we read: "The focus of the EOSC is to become a public good. Key contributions towards construction and future operation of EOSC are expected from public research organizations and services also from the private sector" (European Commission, 2022, p. 3).

The terms "common good" and "public good" address different aspects. The common good encompasses resources that may not serve every individual's immediate interest but contribute to collective well-being – such as universities, arts, and cultural institutions. The resources that constitute the common good address a unique set of shared interests that are fundamental to all citizens – interests that form the core of civic engagement and mutual responsibility (Hussain & Kohn, 2024).

The term "public good", on the other hand, specifically tackles the economics perspective rather than the philosophical one. In everyday speech, public goods are understood as collective goods that are provided by the state. But in economics a public good is defined as something that is non-excludable – meaning it is difficult, if not impossible, to prevent anyone from accessing it, and non-rivalrous – one person's use of it does not diminish its availability to others (Samuelson, 1954). According to the public choice school of economics, the state should provide public goods only when the benefits are non-excludable and the enjoyment of the goods is non-rivalrous (Olson, 1965; Samuelson, 1954).

This conceptual framework is useful for highlighting one aspect: if EOSC is composed of both non-excludable and excludable parts, in order for it to be a tool for OS, it is important to adopt policies that limit potential excludability. The excludability in fact may emerge from the conditions of access to EOSC by the research community at large. We will analyse these aspects in a more detailed manner in the paragraph dedicated to "EOSC and its components".

From this characterization of EOSC as a common or public good, it is evident that issues of capacity, inclusion and equality need to be taken into account. This could be trivial, since the EOSC was explicitly designed to broaden access for

[5] EOSC Federation: https://open-science-cloud.ec.europa.eu

[6] EOSC Federation Handbook: https://eosc.eu/eosc-federation-handbook

under-resourced researchers and, in theory, to reduce disparities among those with different levels of access to resources. Yet, this assumption warrants close examination against the real-world conditions EOSC faces.

Inclusion and equality in the Open Science framework

In the scientific landscape – much like in society at large – inequality and exclusion remain pervasive. The ability to conduct research is contingent on access to material resources and enabling conditions which vary significantly across countries, regions, gender, and thematic sectors.

According to DiPrete and Eirich (2006), academia is particularly susceptible to what they term "cumulative advantage", defined as "a general mechanism for inequality across any temporal process (...) in which a favorable relative position becomes a resource that produces further relative gains" (p. 271). This concept refers to a process in which a favorable starting position leads to further advantages, perpetuating inequality. Unequal access to knowledge, infrastructure, and resources exemplifies this dynamic.

OS policies must acknowledge and address these realities. Power imbalances and historical inequalities in knowledge production will not vanish simply because OS initiatives are implemented (Albornoz *et al.*, 2020). As Ross-Hellauer and colleagues caution, "uncritical narratives of openness may fail to address structural barriers in knowledge production, thereby perpetuating the cumulative advantage of dominant groups and the knowledge they produce" (Ross-Hellauer *et al.*, 2022, p. 3).

EOSC must address issues of inclusion and equality, as these are core to the principles of the OS movement, which EOSC seeks to embody and advance. For instance, the definition given in the UNESCO Recommendations on Open Science (2021) pose inclusion in a prominent position:

> Open science is defined as an inclusive construct that combines various movements and practices aiming to make multilingual scientific knowledge openly available, accessible and reusable for everyone, to increase scientific collaborations and sharing of information for the benefits of science and society, and to open the processes of scientific knowledge creation, evaluation and communication to societal actors beyond the traditional scientific community (UNESCO, 2021, p. 7).

According to Sabina Leonelli, the most appropriate way to manage Open Science practices would be to prioritise inclusion (Leonelli, 2023). If inclusion is not considered as the starting point of this endeavour, it would then be too late to reach it. Leonelli points out that in most OS policy documents, the ultimate goal of enhancing transparency and quality (often defined as reproducibility) is to create an inclusive and equitable research process (see for instance Burgelman *et al.*, 2019; European Commission, 2016; European Commission, 2018b; National Academies, 2018; United Nations, 2019). However, this ideal is rarely implemented in practice.

If the disparities resulting from cumulative advantage are not adequately addressed before the full implementation of EOSC, the initiative itself may result in another vehicle for reinforcing cumulative advantage for those in more privileged positions. This, in turn, would widen the gap between researchers with abundant resources and those with fewer, contradicting EOSC's purpose.

Below, we examine why and how EOSC could fail reducing inequalities if the issue of inclusion is not properly addressed. One key factor is that global standards can intensify discrimination (Leonelli, 2024): researchers who lack the resources to participate in the development and governance of data infrastructure are disproportionately impacted by OS. However, rather than on researchers' participation in decision-making processes, our focus centers on issues more closely related to their material starting conditions.

We analyse the different components of the EOSC in view of their accessibility and usability for all researchers and identify potential barriers that may limit equitable access.

EOSC and its components

From a structural point of view, the EOSC presents aspects that in the economic theory would be defined as excludable and non-excludable, rivalrous and non-rivalrous. From the combination of these variables, it is possible to identify three components of the EOSC.

Non-rivalrous and excludable component: scientific knowledge. Scientific knowledge is often classified as both a common and a public good (Becerril Garcia, 2024; Boulton, 2021; International Science Council, 2024), though it fits the economic definition of a public good only in theory (Stiglitz, 1999). In practice, various obstacles – such as restricted access and paywalls – limit universal access and equitable benefits from scientific knowledge.

In the case of EOSC, we observe that access to EOSC FAIR data by one researcher may not diminish access for others, characterising it as non-rivalrous.

EOSC is designed to support an unfiltered, unrestricted number of users – initially targeting the research and educational sectors but ultimately expanding access to innovation-driven organisations and broader society. This open-access approach may necessitate scaling up infrastructure costs to accommodate growing demand, but it also promises medium- to long-term benefits by fostering unconventional discovery, innovation, and a vibrant scientific culture.

Yet, EOSC's principle of "as open as possible" – which replicates an OS principle – may be subject to excludable aspects, including barriers like access fees to some services or restricted-access platforms, based on own rules of the participating nodes, which limit accessibility. Other examples include insufficient infrastructure for data analysis owned by the users, and and a lack of expertise, which can limit the

broad impact of EOSC FAIR data on research. These barriers often stem from material conditions that the (mainly methodological) shift associated with OS marginally tackle. Moreover, the knowledge generated through scientific research is frequently privatised, further exacerbating inequalities (Florio, 2024).

Given these considerations, EOSC is better described as a "club good" (Buchanan, 1965). While EOSC data and some EOSC services may indeed be non-excludable, others services may restrict access to a subset of users. EOSC strives to embody the ideal of a common good, aiming to offer services accessible and beneficial to all. However, this remains an aspirational goal, and access to the resources and services within EOSC is not automatically open to all; various conditions and barriers can influence who is able to exploit its offerings. Therefore, policies about EOSC implementation need to effectively minimize its excludable component.

Rivalrous and excludable aspects: the infrastructural level. To understand the rivalrous side of the EOSC, it is useful to draw a parallel with the Research Infrastructure (RI) paradigm[7]. Public investments in establishing and operating RIs create a shared resource for competitive researchers, enabling them to leverage unique facilities to generate data and knowledge more efficiently, with a higher return on investment than typical research environments. However, access to RIs is inherently limited; they are rivalrous because their finite capacity requires access to be regulated, often through peer-reviewed selection processes that prioritise high-impact projects.

Prima facie, as previously noted, this is not the case for EOSC, as access by one researcher does not hinder access by others. However, at a deeper level, EOSC exhibits a rivalrous side too, precisely in what makes RIs rivalrous: the infrastructural component, the machines, and tools essential for generating and interpreting data. Despite being publicly funded and managed, these elements function as private goods. This is because their ownership and most likely their use is typically restricted to a selected group of individuals or organizations that have access to the necessary resources, whether financial, technical, or institutional. As a result, while the infrastructure itself might be publicly supported, its practical availability is often exclusive. This exclusivity extends beyond the tools themselves to include the energy required to power the entire research ecosystem, and particularly to support data generation, storage, and processing.

[7] Research Infrastructures are – according to an agreed-upon definition that has also been endorsed by ESFRI (the European Strategy Forum on Research Infrastructures) – "facilities that provide resources and services for the research communities to conduct research and foster innovation in their fields. These include: major equipment or sets of instruments; knowledge-related facilities such as collections, archives or scientific data infrastructures; computing systems; communication networks" (https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/research-infrastructures_en).

Current computing and data storage infrastructures, which are critical to scientific research, are largely dependent on national investments, and their capacity is often constrained. As demand for computational resources increases, there is a growing concern that these infrastructures may not scale quickly enough to meet the needs. The rise of technologies like Artificial Intelligence (AI), especially Generative AI tools and Large Language Models (LLMs), is exacerbating this situation. These tools require massive computational power and memory, alongside specialized ancillary software, to function. As the need for high-performance computing (HPC) escalates, the inequality in access to these resources does it too, as only some institutions with substantial funding or access to state-of-the-art own infrastructure, can deploy or maintain these technologies at scale.

Furthermore, as these technologies evolve, the focus of interoperability within the research ecosystem is shifting from ensuring compatibility between data formats to ensuring "machine actionability". This shift means that the emphasis is no longer just on whether data can be shared or used by different researchers but on whether it can be acted upon directly by machines.

In light of these considerations on the material resources required for research, it can be stated that EOSC services, which are designed as open platforms for sharing scientific data, find themselves in a paradoxical position. In fact, they exhibit characteristics of rivalrous and excludable goods – much like private goods – due to their reliance on proprietary, resource-intensive technologies. The limited availability of critical resources, such as computing power and storage capacity, means that only those with the necessary financial and technical resources will be able to fully benefit from EOSC data. This disparity risks undermining the core principles of openness, accessibility, and equality that EOSC wants to foster. To address these challenges ad-hoc models must be developed, taking into account models such as the "fair share scheduler" proposed by Kay and Lauder (1988).

A focus on Artificial Intelligence

Artificial intelligence as a tool for scientific research represents an area that demands careful and forward-thinking consideration, particularly as technological advancements continue to reshape the landscape of research infrastructure. As we move further, it is increasingly likely that machines will become the primary users of the EOSC, while their human owners or operators will assume the role of meta-users. This shift in the nature of EOSC usage introduces significant challenges, particularly in the realm of data querying, which will increasingly involve AI systems capable of making automated, complex requests, perhaps bypassing the FAIR paradigm.

This raises important questions about what constitutes a legitimate query for machines exploiting the EOSC and what does not. It will be essential to explore the criteria that define such legitimacy – especially when these queries are driven by AI algorithms rather than human intent. In particular, we

must address how to prevent that queries that do not meet certain ethical, technical or scientific standards, or, in other words, that push the boundaries of what is legally permissible or technically/scientifically feasible, trigger the EOSC. The use of deep learning algorithms within the context of EOSC, trained on data that adhere to FAIR principles, represents a significant opportunity for the scientific community and a major step towards the development of secure and reliable AI instruments. Because of this EOSC must develop robust strategies to protect its data against malicious intrusions or harmful use. Generative AI creates synthetic data that will integrate into the EOSC, with predictable effects of both positive and negative sign. The increase of information exchange of EOSC has intrinsically the same nature as any other communication technology of the past (e.g. printing books) which contributed to the spread of good knowledge in case of scientific books, but also created the background for the most negative episodes of humanity in case of fake-news and books spreading fake-stories.

One key element in addressing this issue will be the development of an expanded concept of metadata. Metadata primarily focus on the methods and attributes of data collection and are mandatory for reproducibility and reuse by the specialists. However, as AI becomes more integrated into research workflows, it will be essential for metadata to also capture the potential uses of the data, the contexts in which the data might be applied, the impacts that may result from their use, and the correlations that might emerge in different scenarios. We could refer to this expanded form of metadata as "hyper-metadata" – a set of data descriptors that complement domain specific metadata by offering a more comprehensive view of how data could be utilized and the broader consequences of its application. In well confined cases AI can run on abundant raw data and obtain reliable results, as e.g. in pulsar discovery from astronomical underexploited archives. A massive use of AI on raw data would on the other hand escape the FAIR principles and makes it necessary to be regulated by hyper-metadata filters to understand the output of the algorithms.

Funding and sustainability challenges for the EOSC

The absence or the excessive delay of the EOSC would have significant adverse effects for the European research landscape. Studies have underscored the high social and economic costs of lacking an ecosystem for open and FAIR data (European Commission, 2018a; Stott, 2014)[8]. However, there remains a concern that existing inequalities in research infrastructure and resources across thematic, institutional, regional, and national levels could widen as a result of the EOSC. Those with an initial advantage in research capacity may be better positioned in the competition for knowledge production once EOSC services are in operation.

Consider this analogy: street lighting is a public service funded through previous investments in the production and distribution of electricity, with installation, maintenance, and energy costs integrated into a general municipal or national budget. Street lighting offers widespread benefits, though some people benefit more than others; for example, those in remote areas may receive less light and therefore rely more on their own sources. In much the same way, users of the EOSC benefit unevenly depending on their material resources and infrastructure.

A related issue is the potential for a free-rider problem. Although a combination of high demand, reciprocity, and altruism might encourage stakeholders to invest in the EOSC, there remains a risk that some entities could gain considerable benefits from its data and services without contributing proportionally. This could lead to an imbalance, where those who invest – such as taxpayers through European and national funding and research institutions via in-kind contributions – shoulder the costs, while non-contributors gain disproportionate advantages.

To address these challenges, it is essential to establish a sustainable public funding model for the EOSC and the infrastructure required to harness its resources effectively. A European strategy for the EOSC should guarantee equitable access to digital and AI services for all researchers, regardless of their institution, nationality, or field of study.

The need of a sustainable public funding model

The current EOSC model assumes that pre-existing national investments will fulfill its initial needs, with national networks expected to provide adequate storage and communication capacity in its early phase. However, the EOSC Federation's reliance on self-funding nodes is unlikely to be sustainable in the medium to long term, as it risks exacerbating competition and exclusion among researchers.

Within the current model, national or EU funding is generally restricted to early-to-middle-stage innovation: this is not ideal to keep pace with the rapid changes in the landscape. According to the "Financial Sustainability Task Force Progress Report" mandated by the EOSC Association (Robertson *et al.*, 2022), for EOSC to be sustained, the European Commission, Member States, Research Funding Organizations and Research Performing Organizations need to review their existing funding mechanisms and explore and experiment with new funding mechanisms to help sustain a growing and maturing research ecosystem for science and society.

Overall, the statement contained in the report on "Solutions for a sustainable EOSC", written in 2020 – when the "Minimum Viable EOSC" (MVE), which is now under construction, was being imagined, remains true:

> […] there are many potential funding schemes and mechanisms that could fund the different components of the MVE but each comes with its own constraints and

[8] A European Commission study has shown that the cost of not having research data available to the European economy is at least €10.2 billion annually. Moreover, the researchers behind this study also identified several unquantifiable consequences of the absence of FAIR data, including potential impacts on research quality, economic turnover, and the machine readability of research data. By drawing a comparison with the European open data economy, they estimated that these unquantifiable effects could add another €16 billion to the cost, bringing the total to at least €26 billion per year (European Commission, 2018a).

> integrating them into a comprehensive funding plan will be a challenge requiring effort of an entrepreneurial nature to actively seek funding opportunities and secure them (European Commission, 2020, p. 17).

For instance, there are barriers[9] connected to in-cash contributions that can be overcome through a mixture of in-cash and in-kind, as done by Research Infrastructures. The sustainability of the full potential of an operational EOSC appears to require a hybrid economic solution with a mix of general public funding for research projects and of specific investments to cope with the energy-costing rivalrous services.

The rise in AI-related costs and related funding

As an important quota of HPC lined to AI and FAIR-data exploitation will develop, the cost of supercomputers, memories and networks and the related energy cost will become an explicit element of the cost of the EOSC[10]. This inevitably leads to rivalrous and excludable services to the research community, likewise the large Research Infrastructures.

To mitigate these challenges, the public sector must develop targeted strategies based on a comprehensive analysis of the specific needs of diverse research communities across different contexts and fields, and a continuous cost monitoring and outlook. Achieving this vision will require significant capital investment and a well-coordinated operational budget among Member States, Associated Countries, and the European Commission, especially as EOSC data will be accessible to all, irrespective of their contribution.

Let us have a look to the Research Infrastructure funding paradigm, particularly the one used by the European Strategy Forum on Research Infrastructures (ESFRI) for RIs[11]. Open, competitive access to RIs has greatly contributed to the success and accessibility of these investments; however, no standardized funding model has been universally adopted, and research inequality has not been fully addressed through open access alone. Funding models for RIs often include mechanisms such as juste retour (where benefits are proportional to each country's investment), reserved quotas for national users, and limited access quotas for users from non-contributing institutions or countries (ESFRI Long-Term Sustainability Working Group, 2017).

These models inherently give researchers from well-resourced institutions in RI-owner countries a competitive edge in securing access, particularly when it comes to leading projects as principal investigators. Researchers from lower- and middle-income countries often face additional barriers in this competitive landscape, with fewer opportunities to leverage these high-cost facilities. This disparity illustrates how open, competitive access does not overcome inequalities based on institutional and national resources, in spite of the fact that scientific merit of proposals from all provenances brings value to the RI investment.

With the increasing reliance on AI and the vast infrastructure required to support its development within EOSC, a hybrid funding approach may be necessary, balancing open access with strategic funding models to support equitable participation. Careful planning is needed to prevent the unintended consequence of reinforcing existing research inequalities and to enable EOSC to support a genuinely inclusive research environment across the EU and beyond.

Conclusions

The geography of the digital realm is rich with both possibilities and challenges. It offers the potential for unprecedented collaboration, with data that can be revisited and exploited infinitely, enabling continuous innovation. Yet precisely because this digital landscape lacks tangible boundaries, it can sometimes create hard to cross barriers. Digital tools have the power to unlock significant opportunities for growth, but they do not reduce intrinsically the disparities between people.

To ensure that EOSC achieves its intended impact and limits the risks of exclusion, a coordinated effort is essential across multiple governance levels, particularly on resources and regulatory frameworks. This implies understanding the necessary capital investments and operational budgets required from Member States, Associated Countries, and the European Commission, as EOSC data will be accessible to all, regardless of individual investment levels. On the other hand, mitigation measures must be introduced to reduce the resulting exclusiveness of costly services needed to exploit EOSC data. Establishing a sustainable public funding model and a robust governance of EOSC infrastructure is critical to meet the scope of maximising the impact on Open Science research, and to approximate the beneficial concept of EOSC as a common good.

Disclaimer

The views expressed in this article are those of the authors. Publication in Open Research Europe does not imply endorsement by the European Commission.

Ethics and consent

Ethics and consent were not required.

Data availability statement

No data are associated with this article.

Acknowledgements

We acknowledge insightful comments on the economic aspects of our work by Massimo Florio (University of Milano) and Roberto Marchionatti (University of Torino).

[9] "These barriers include access policies and required levels of assurance; variations in Value Added Tax rules; data protection and intellectual property restrictions" (European Commission, 2020, p. 17).

[10] A growing body of research explores the significant energy demands of AI and its environmental impact. For example, see Kaack *et al.* (2022), Luccioni *et al.* (2024), and Strubell *et al.* (2019).

[11] European Strategy Forum on Research Infrastructures (ESFRI): https://www.esfri.eu/

References

Albornoz D, Huang M, Martin IM, *et al.*: **Framing power: tracing key discourses in Open Science policies.** In: *The 22nd International Conference on Electronic Publishing. Connecting the Knowledge Commons: From Projects to Sustainable Infrastructure.* ELPUB, 2018, Toronto, Canada, June, 2018. 2020.
Publisher Full Text

Becerril Garcia A: **The unique opportunity to advance science as a global public good: Open Science in a world of contrasts.** 2024; Retrieved September 23, 2024.
Reference Source

Boulton GS: **Science as a global public good, international science council position paper.** 2021.
Reference Source

Buchanan JM: **An economic theory of clubs.** *Economica.* 1965; **32**(125): 1–14.
Publisher Full Text

Burgelman JC, Pascu C, Szkuta K, *et al.*: **Open Science, open data, and open scholarship: European policies to make science fit for the Twenty-First century.** *Front Big Data.* 2019; **2**: 43.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

DiPrete TA, Eirich GM: **Cumulative advantage as a mechanism for inequality: a review of theoretical and empirical developments.** *Annu Rev Sociol.* 2006; **32**: 271–297.
Publisher Full Text

ESFRI Long-Term Sustainability Working Group: **Long-term sustainability of research infrastructures.** ESFRI Scripta Series, 2017; **II**.
Reference Source

European Commission: **Open innovation, Open Science, open to the world: a vision for Europe.** Publications Office of the European Union, 2016.
Reference Source

European Commission: **Cost of not having FAIR research data.** Publications Office of the European Union, 2018a.
Publisher Full Text

European Commission: **Mutual learning exercise. Open science: altmetrics and rewards.** Publications Office of the European Union, 2018b.
Publisher Full Text

European Commission: Directorate-General for Research and Innovation: **Solutions for a sustainable EOSC: a FAIR lady (olim iron lady) report from the EOSC Sustainability Working Group.** Publications Office of the European Union, 2020.
Publisher Full Text

European Commission: Directorate-General for Research and Innovation: **Opinion paper on EOSC and commercial partners.** Publications Office of the European Union, 2022.
Publisher Full Text

European Commission: Directorate-General for Research and Innovation: **Opinion paper on advanced digitalisation of research.** Publications Office of the European Union, 2024.
Publisher Full Text

Fecher B, Friesike S: **Open Science: one term, five schools of thought.** RatSWD_WP_ 218, 2013.
Publisher Full Text

Florio M: **The privatisation of knowledge a new policy agenda for health, energy, and data governance.** Routledge Studies in Public Economics and Finance, New York, Routledge, 2024.
Reference Source

Hussain W, Kohn M: **The common good.** *The stanford encyclopedia of philosophy.* (Fall 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.). 2024.
Reference Source

International Science Council: **Science as a global public good.** International Science Council, 2024; (Accessed: 4 September 2024).
Reference Source

Kaack LH, Donti PL, Strubell E, *et al.*: **Aligning artificial intelligence with climate change mitigation.** *Nat Clim Chang.* 2022; **12**(6): 518–527.
Publisher Full Text

Kay J, Lauder P: **A fair share scheduler.** *Commun ACM.* 1988; **31**(1): 44–55.
Publisher Full Text

Leonelli S: **Philosophy of Open Science.** *Cambridge Elements in the Philosophy of Science.* Cambridge University Press, Cambridge (UK), 2023.
Publisher Full Text

Leonelli S: **Human-Centric Open Science.** speech presented at the EOSC Tripartite Event, Brussels, April 16, 2024. 2024.
Reference Source

Luccioni S, Jernite Y, Strubell E: **Power hungry processing: Watts driving the cost of AI deployment?** *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency.* Association for Computing Machinery, 2024; 85–99.
Publisher Full Text

Mounier P: **Connecting the building blocks of Open Science: an ecological approach.** *Septentrio Conference Series.* 2022; (1): Article 1.
Publisher Full Text

National Academies of Sciences, Engineering, and Medicine: **Open science by design: realizing a vision for 21st century research.** The National Academies Press, Washington, DC, 2018.
PubMed Abstract | **Publisher Full Text**

Olson M: **The logic of collective action: public goods and the theory of groups.** Cambridge, MA Harvard University Press, 1965.
Reference Source

Robertson D, Meijer J, Roi A, *et al.*: **Towards sustainable funding models for the European Open Science Cloud.** Financial Sustainability Task Force Progress report, EOSC Association, *Zenodo*. 2022.
Publisher Full Text

Ross-Hellauer T, Reichmann S, Cole NL, *et al.*: **Dynamics of cumulative advantage and threats to equity in Open Science: a scoping review.** *R Soc Open Sci.* 2022; **9**(1): 211032.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Samuelson PA: **The pure theory of public expenditure.** *Rev Econ Stat.* 1954; **36**(4): 387–389.
Publisher Full Text

Stiglitz JE: **Knowledge as a global public good.** In: I. Kaul, I. Grunberg, & M. Stern (Eds.), *Global Public Goods.* 1st ed., Oxford University Press, New York, 1999; 308–325.
Publisher Full Text

Stott A: **Open data for economic growth.** World Bank Group, Washington, D.C., 2014.
Reference Source

Strubell E, Ganesh A, McCallum A: **Energy and policy considerations for deep learning in NLP.** In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics.* ACL Anthology, 2019; 3645–3650.
Publisher Full Text

UNESCO recommendation on Open Science. UNESCO Digital Library, 2021.
Reference Source

United Nations: **Towards global Open Science: core enabler of the UN 2030 Agenda.** 2019.
Reference Source

Vicente-Saez R, Martinez-Fuentes C: **Open Science now: a systematic literature review for an integrated definition.** *J Bus Res.* 2018; **88**: 428–436.
Publisher Full Text

Wilkinson MD, Dumontier M, Aalbersberg IJJ, *et al.*: **The FAIR guiding principles for scientific data management and stewardship.** *Sci Data.* 2016; **3**(1): 160018.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔

Version 2

Reviewer Report 16 April 2025

https://doi.org/10.21956/openreseurope.21517.r53069

© **2025 Honorato R.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Rodrigo Vargas Honorato**
Utrecht University, Utrecht, The Netherlands

No further comments

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Open Science, EU Projects

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 22 February 2025

https://doi.org/10.21956/openreseurope.20746.r50341

© **2025 Honorato R.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Rodrigo Vargas Honorato**
Utrecht University, Utrecht, The Netherlands

Overview: The article evaluates the European Open Science Cloud (EOSC) as a key initiative for Open Science, emphasizing its role as a common good while identifying challenges in governance, accessibility, and funding. It brings as key points; The risk of cumulative advantage, benefiting well-funded institutions disproportionately; discusses the AI's impact on computational resources

and ethical concerns; details the financial challenges, including the feasibility of a hybrid funding model and further discusses the need for inclusive governance to ensure equitable access.
Rationale for the Open Letter: The rationale is provided in sufficient detail.
Engagement with Different Perspectives: The open letter adequately balances different views.
Support for Factual Claims: The factual statements are strongly supported.
Accessibility of Language: The language is clear and accessible.
Clarity of Recommendations: The recommendations for the next steps are clearly explained.
The open letter is a well-researched contribution to the Open Science debate with a strongly supported statements.

Is the rationale for the Open Letter provided in sufficient detail? (Please consider whether existing challenges in the field are outlined clearly and whether the purpose of the letter is explained)
Yes

Does the article adequately reference differing views and opinions?
Yes

Are all factual statements correct, and are statements and arguments made adequately supported by citations?
Yes

Is the Open Letter written in accessible language? (Please consider whether all subject-specific terms, concepts and abbreviations are explained)
Yes

Where applicable, are recommendations and next steps explained clearly for others to follow? (Please consider whether others in the research community would be able to implement guidelines or recommendations and/or constructively engage in the debate)
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Open Science, EU Projects

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 12 February 2025

https://doi.org/10.21956/openreseurope.20746.r50358

© 2025 Armeni K. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Kristijan Armeni
Johns Hopkins University, Baltimore, Maryland, USA

The authors discuss the European Open Science Cloud (EOSC) initiative as a common/public good and the implications of the framework for EOSC's research equity. They emphasize that equity is a core goal behind reproducibility requirements in open science and that initiatives like EOSC might fail to deliver equitable outcomes if the implementation and funding schemes do not follow suit.

The authors turn to the concepts of common and public good and their exclusive nature. Due to the documented phenomenon of cumulative advantage, even open research infrastructure risks aggravating digital inequity: those who have more resources, have competitive advantage and can leverage more open resources that those who have less.

The crux of the argument pertains to infrastructure and funding models. The authors argue that by their nature, infrastructure and materials goods (that would comprise the EOSC) are excludable (those who use it, take up space for others) and cannot guarantee equitable outcomes if equitable policies and funding models are not put in place.

I find this commentary to be timely, well argued for, and for the most part clearly written. I think it would be stronger if it included concrete example(s) of EOSC usage (if it exists) that could be used throughout the commentary to highlight the problems the authors are addressing and possibly policy solutions. In general, a greater focus and discussion on potential solutions, which are currently outlined towards the end, would make the commentary stronger. I indicate these and some minor points in my comments below.

Comments
1. Clarifications. "Federation of nodes" -- are these actual computational nodes or are these metaphorical components? Could Google contribute a Node? Meta its data centers? It would be helpful to illustrate perhaps with a short example. You do provide links to the EOSC documents, but a self-contained example, if available, would be valuable. Likewise for the term "Research Infrastructures" which has a specific meaning, but will remain opaque to non-policy experts.

2. Adding references to work on energy and AI. Specifically, when discussing relationship between AI and energy demands, consider adding references to some of the important work on AI computing infrastructure and energy, these references are a good starting point:

- Strubell et al, 2020 [ref 3]
- Kaack et al, 2021 [ref 1]
- Luccioni et al, 2024: [ref 4]

3. Excludable infrastructure and fair usage. You state: "*The limited availability of critical resources, such as computing power and storage capacity, means that only those with the necessary financial and technical resources will be able to fully benefit from EOSC services. This disparity risks undermining the core principles of openness, accessibility, and equality that EOSC wants to foster.*"
--> In high-performance computing (HPC), a common procedure for fair allocation of computing resources to its users already having access takes into account any user's history of usage etc. (Kay & Louder, 1988. A fair share scheduler [ref 2]). The EOSC cannot make infrastructure goods non-excludable, but to run with a practical example, would your argument imply that EOSC at the infrastructure level must implement something akin to the fair share policy? Would that even be

possible in the EOSC participation/funding schema?

4. Discussion of hybrid funding schemes. "*A framework that combines investment returns with access quotas might help ensure that researchers from diverse [...]*". Funding schemes have the highest leverage point to ensure equity and this seems a good place to provide a practical example of this solution. If authors know of and can provide a reference where this concept (access quotas etc.) and its practical feasibility is discussed, it would make this argument stronger.

5. Typo. "characterising it as non-rivalrous," --> missing either a sentence or a period.

References

1. Kaack L, Donti P, Strubell E, Kamiya G, et al.: Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*. 2022; **12** (6): 518-527 Publisher Full Text
2. Kay J, Lauder P: A fair share scheduler. *Communications of the ACM*. 1988; **31** (1): 44-55 Publisher Full Text
3. Strubell,E.,Ganesh,A.,McCallum,A: Energy and Policy Considerations for Deep Learning in NLP. *ACL Anthology*. Publisher Full Text
4. Luccioni,S.,Jernite,Y.,Strubell,E: Power Hungry Processing: Watts Driving the Cost of AI Deployment?. *Association for Computing Machinery*. Publisher Full Text

Is the rationale for the Open Letter provided in sufficient detail? (Please consider whether existing challenges in the field are outlined clearly and whether the purpose of the letter is explained)

Yes

Does the article adequately reference differing views and opinions?

Yes

Are all factual statements correct, and are statements and arguments made adequately supported by citations?

Partly

Is the Open Letter written in accessible language? (Please consider whether all subject-specific terms, concepts and abbreviations are explained)

Yes

Where applicable, are recommendations and next steps explained clearly for others to follow? (Please consider whether others in the research community would be able to implement guidelines or recommendations and/or constructively engage in the debate)

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: cognitive neuroscience, AI, open science, data sharing

I confirm that I have read this submission and believe that I have an appropriate level of

expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 07 February 2025

https://doi.org/10.21956/openreseurope.20746.r50345

© **2025 Stefanizzi S.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sonia Stefanizzi
Università di Milano Bicocca, Milan, Italy

This article addresses a crucial issue in the context of European open science, that is, the role of the European Open Science Cloud (EOSC) as a common good for research and society. The authors provide a detailed analysis of the potential and limitations of EOSC, emphasizing aspects of inclusiveness, sustainability, and accessibility. The article is an important contribution to the open science debate, offering interesting insights into the potential and limitations of EOSC. However, more attention to the practical details of the proposed solutions and a closer look at the ethical and social implications would improve the robustness of the analysis. In particular, it is suggested that the criteria of inclusiveness be developed more. Although the issue of inclusiveness is central, the article does not explore in depth how existing inequalities can be concretely mitigated in the context of EOSC. It is suggested to provide practical examples or case studies of already successfully implemented inclusiveness policies that could be adapted to the EOSC context.
The article touches on the topic of AI in a promising way, but does not fully develop the ethical and technical implications of its use within EOSC. It would be useful to elaborate on how AI may affect data governance, accessibility, and potential discrimination, and to propose strategies to manage these issues. In addition, proposals to ensure sustainable funding of EOSC should be expanded. In this regard, it would be fruitful to include a more in-depth discussion of hybrid funding models (e.g., public and private contributions), specifying potential practical applications and barriers. Stakeholder involvement (e.g., researchers, institutions) is mentioned, but the process through which these groups could contribute to EOSC governance could be better clarified. It is suggested that concrete mechanisms be described for including stakeholders in the decision-making process, such as the use of public consultations or participatory platforms.

Finally, the importance of EOSC for improving scientific communication and collaboration is mentioned in the article but not explored in depth. It would be useful to discuss how EOSC can foster the dissemination of knowledge.

Is the rationale for the Open Letter provided in sufficient detail? (Please consider whether existing challenges in the field are outlined clearly and whether the purpose of the letter is explained)
Yes

Does the article adequately reference differing views and opinions?

Yes

Are all factual statements correct, and are statements and arguments made adequately supported by citations?
Yes

Is the Open Letter written in accessible language? (Please consider whether all subject-specific terms, concepts and abbreviations are explained)
Yes

Where applicable, are recommendations and next steps explained clearly for others to follow? (Please consider whether others in the research community would be able to implement guidelines or recommendations and/or constructively engage in the debate)
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research infrastructures, open data, open government data, open science

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 11 Feb 2025

Anna Bertelli

We would like to thank Prof. Stefanizzi for her overall apprecitiation of our work and for the stimulus to further explore the potential criticalities of EOSC.

Our article highlights key aspects of EOSC that should be part of the public discourse during this crucial phase of the EOSC Federation's initial implementation, as well as decisions on funding strategy within the Framework Programme for research.

Among the suggestions put forward by Prof. Stefanizzi we definitely strive to delve deeper into the criteria of inclusivity, the overall impact of AI, including in data management, in future works.

We will not forget to include concrete examples and case studies, as suggested.

Competing Interests: No competing interests were disclosed.