Achieving Open Science in the European Open Science Cloud

Setting out OpenAIRE's vision and contribution to EOSC

Position Paper – September 2019

Authors and Contributors

Natalia Manola, OpenAIRE Managing Director (Athena Research & Innovation Centre)

Najla Rettberg, OpenAIRE Network and Outreach Director (Univ. of Göttingen)

Paolo Manghi, OpenAIRE Technical Director (CNR-ISTI)

Mike Mertens, Univ. of Göttingen

Birgit Schmidt, Head of Knowledge Commons, Göttingen State and University Library

Tobias Steiner, OpenAIRE Scientific Advisor, Göttingen State and University Library

Prodromos Tsiavos, OpenAIRE Policy and Legal Advisor (Athena Research & Innovation Centre)

Eloy Rodrigues, Director Documentation Services, University of Minho

Nina Karlstrom, Norwegian Directorate for ICT and Joint Services in Higher Education and Research (UNIT)

Mojca Kotar, University of Ljubljana

Version 1.2

26 September 2019

About OpenAIRE

OpenAIRE is a socio-technical infrastructure for scholarly communication and Open Science. For over ten years it has been supporting Open Science at national levels via its network of experts from key national organizations (National Open Access Desks – NOADs) who support policy development for Open Science within the research realm. Together they train and guide today and tomorrow's researchers in the many aspects of Open Science.

At a technical level, OpenAIRE has developed a range of innovative services, powered by its Open Science Graph - a unique and vast body of knowledge, an interoperable dataset of research results that promotes discoverability and relationships between research artefacts. OpenAIRE's B2B (business to business) and B2C (business to client) services consist of a services layer, a data interoperability layer and an access interoperability layer. By gathering knowledge from all its different data sources, OpenAIRE also plays a role of promoting interoperability - via its technical guidelines - which in turn facilitate the integration of diverse datasets for discovery in EOSC.

Given this backdrop, OpenAIRE is now in a position to bring established and future forms of scholarly communication to EOSC and to move from an abstract, technological infrastructure closer to EOSC being meaningful for every individual researcher in Europe. OpenAIRE is an infrastructure in itself and, by this nature, primarily focuses on B2B processes behind the scenes. But OpenAIRE takes the human part of infrastructure and the local roots of research very seriously and – with topics such as publishing, reputation and assessment -- addresses questions that are of utmost importance to individual researchers.

OpenAIRE AMKE, a not-profit civil partnership company founded in 2018 with currently 16 members, is the legal entity behind OpenAIRE organization overseeing the operations of the OpenAIRE infrastructure, network and services.

OpenAIRE-Advance is a project funded by the European Union (Grant Agreement No 777541).

Table of Contents

Executive Summary

The European Open Science Cloud (EOSC) is Europe's undertaking to address the rising field of data-driven research in a more efficient way, to break down walls in existing siloed approaches for access to services and all scientific content (publications, data, software, methods), and to further evolve the ways in which researchers are able to collaborate. This paper has a twofold role: (i) to strengthen and reinforce Open Science as the modus operandi in EOSC, i.e., embedding openness by default in all stages of research, and (ii) to highlight the important role of *open* scientific communication and OpenAIRE participatory infrastructure as a vehicle to deliver Open Science in a trusted and reliable way.

EOSC as a *"Research Commons"*

EOSC is best envisioned to be a **co-created space,** in which research information is produced and transformed into knowledge**.** A *"Research Commons"* that is collaboratively developed and managed by Europe's research communities. Interconnecting both digital and human resources, EOSC adds *new* value to existing infrastructures, via **frictionless data flow**, **intelligent discovery** and **retrieval of scientific output**, as well as **homogenized and secure access to services** that transfer, store and analyse data.

Its key characteristics are:

Openness by default: A trusted infrastructure with Open Science practices for transparency, reproducibility and accountability embedded at every phase of the research life cycle.

A network of digital resources: A System of Systems, focussing on *business-to-business* sharing and access. A fit-for-purpose, lightweight *regulatory* framework, to ensure that different players have the best possible conditions to develop on their own, with freedom of contract as a cornerstone.

Researcher-centric: An abstraction layer for data manipulation, including reporting, workflows, parallelism, and persistence, enabling a reproducible 'digital laboratory' as a standard way of working, allowing them to communicate with peers and non-peers alike.

Open governance: A shared arena, maintained and preserved by a range of players in collaboration, governed openly at a national level (via national nodes to strengthen EOSC structures), and guided by the EC.

Training tomorrow's data experts: A coordinated, professional and certified training infrastructure, embedded in institutions and research groups, as part of the European Digital Skills agenda.

Achieving Open Science in EOSC

For EOSC to fulfil its mission and vision, Open Science and specifically Open Access to research results should be a strategic priority. First, this would bootstrap EOSC uptake as it would achieve a critical mass of openly accessible content for researchers to experiment with, stimulating the demand. Second, it would be the key mechanism for linking research to the critically needed in the emerging AI ecosystem, by turning data into *smart* and *actionable data:* connecting, organizing, and packaging it into making insights available to all (accessible), easy to apply (understandable), and focused on the task at hand (actionable).

To achieve this, we consider three focus areas:

Holistic perspective on data flow: Equally addressing large scale science and scientific research carried out on a smaller scale, is critical for the success of open science in EOSC, both in terms of value and investments, and in terms of researcher uptake.

Researcher in the centre: From gaining traction (what we have now) to gaining widespread adoption of open science, EOSC should pay attention to:

Easiness: Offer services and tools *seamlessly embedded* in research workflows.

Reward: Help with restructuring incentives for data-driven science by keeping track of research production and use.

Skills: Coordinate targeted training and support, addressing all aspects and all levels of the research life cycle.

Commitment: Influence the efficiency of processes and increase sustainability of resources and business models for openness.

Scientific communication as a vehicle for Open Science: The research and cultural diversity of Europe will accept no "one-size-fits-all" solutions for open science. Scientific communication, a long-established research practice, is an apt choice as a common language for understanding and applying open practices across borders and disciplines, with all stakeholders in the loop.

EOSC in this regard should be the catalyst to ensure the delivery of required changes and to establish openness as the new norm in open Scientific Communication:

- Publishing all kinds of scientific products
- Publishing semantic links
- Publishing experiment products as digital packages of workflows
- Innovation in publishing and dissemination practices and methods
- Quality control for securing the quality, reproducibility, FAIRness of research results
- Assessment and rewarding by intelligently combining diverse/open/auditable metrics.
- Monitoring Open Science to ensure its implementation and adjust its aims.

EOSC 'Deconstructed'

EOSC, similar to any other digital infrastructure, will require a seamless interaction between technological and human elements to ensure its effective operation and pervasiveness within researcher communities. As a multi-tier structure EOSC interconnects three main key players: institutions which implement policies, deliver services and support researchers; national structures which shape and align policies and coordinate their connection to EOSC; the EC who guides and coordinates efforts.

Technological layer: Key to EOSC architecture is its sustainability, which will *primarily* be driven by: 1- a participatory design and governance to accommodate different needs and requirements, 2- shared investments as they are being developed by member states (MS/AC), 3- the ability to adapt to new technologies and foster innovation. It is therefore critical that EOSC architecture avoids at any cost a monolithic and centralized approach, and follow a "System of Systems" approach, where resources are brought together at different levels to deliver data and data services. Emphasis is therefore placed on a business-to-business (B2B) sharing (data, services, people) and access, with agreements on:

1- A *shared policy compliance framework* (i) dictating and applying the rules of how the data elements are published, shared and re-used,

and (ii) implementing an interlinked data space where every research result comes with its context (related entities), provenance (full data and science path) and usage.

2- A ***uniform delivery of services and data*** starting from non-regulatory measures, such as recommended *standard contract terms* for services and data access, *specifications* for quality (certification), security and privacy (including trust and identity), accompanied by functions for *categorization* (to simplify browsing), *curation* (to deliver quality), *cataloguing* (to offer intelligent discovery), *crowdsourcing* (to engage researchers for active participation).

The human infrastructure reflects a **balance of policy, training and support implementation** to ensure take-up at MS/AC level and among research communities. It is crucial to highlight the importance of this network, which helps in shaping, adopting and adapting the *shared policies* on national and institutional levels, and ensures that the next generation of researchers is trained in Open Science practices.

OpenAIRE – a pillar for Open Science

OpenAIRE has been supporting Open Science since 2009. It embodies a participatory, service-driven infrastructure anchored on the triplet of **policies – services - training**, contributing the following assets in EOSC:

Policies & Governance: A network of **34 National Open Access Desks (NOADs)**, which comprises experts working on transferring and translating EU policies to a local level. NOADs and their organizations are the *de-facto national nodes for Open Science* in the majority of their countries, already highlighting the connection to EU policies, and ready to step up and become an integral part of **national EOSC structures**.

Infrastructure & Services: A **data infrastructure** enabling open scientific communication, connecting and federating repositories and services across institutions, national settings and RIs, via two key services:

The ***OpenAIRE Guidelines for metadata exchange***, a common interface and a key access mechanism (RoP) used by content providers (literature, data, software, methods, workflows) from national infrastructures around the world.

The ***OpenAIRE Research Graph***, a decentralized, trusted, interoperable dataset containing all interlinked EOSC resources: research results, who has produced them, under which grant, using which facility and e-science service, how open and fair are they, how has it been used.

Training & Support: An Open Science Helpdesk bringing coherence to the EOSC training and support landscape, by leveraging the unique potential and placement of NOADs to train stakeholders and build local support networks for researchers and data practitioners.

OpenAIRE MAKE, founded in 2018 (counts 16 members as of Aug 2019), is a legal vehicle to engage with MS/AC on committing to setting aside resources for Open Science coordination and activities, as an integral component of their national research infrastructures.

Key Focus Areas & priorities for EOSC

We have identified a set of key recommendations for core activities and functions for an effective implementation of open science and EOSC.

1 - Policies and Governance

EOSC will converge **national support structures** bringing all players together in a collaborative arrangement. No single organization is able to fulfil the Commons approach alone and

implementation of Open Science requires specific handling, as most of the barriers are cultural and organizational.

OPEN SCIENCE IN NATIONAL EOSC STRUCTURES: Include authoritative open science experts, such as the OpenAIRE NOADs as integral components in national EOSC structures to formulate, align and implement policy: from macro to micro, from EU to local.

OPEN SCIENCE IN THE WIDER NATIONAL AGENDA. Place open science policy for EOSC as part of a collective set of national innovation strategies and overall governmental agendas to gain: (i) stronger commitment as national agendas are translated into concrete initiatives by other innovation system actors, and (ii) better coordination among different national players, as science transcends ministries' jurisdiction.

2 - Infrastructure and services

For Open Science to succeed in EOSC, we need to: i) provide services for all stakeholders involved in the research life cycle, ii) ensure data federation for both small and big data to become an integral part of EOSC, iii) embed services in institutional settings, and iv) link to international infrastructures.

SCIENTIFIC COMMUNICATION FOR OPEN SCIENCE. Identify commonalities across disciplines and develop / operate horizontal publishing and research rewarding services. Avoid duplication by finding the touching points of national vs. cross-RI data federation.

LINKED OPEN SCIENCE: Develop and maintain a curated interlinked graph of all research outcomes and make it a **core, public shared resource (asset) of EOSC**. A global and trusted research graph under which many efforts and services converge, and which can be used as an overarching EOSC catalogue used in contextual discovery and assessment of research.

OPEN ANALYTICS FRAMEWORK & INFRASTRUCTURE: Embed an infrastructure for scientific reward in EOSC from early stages by collecting all types of usage data for all types of resources (citations, usage events for data, services, software). As part of the *Research Commons* and an international undertaking, this infrastructure should place emphasis on B2B aspects (specifications, formats, APIs) rather than indicators themselves.

3 - Support and training

EOSC should develop and rollout a robust and **pervasive training and support Infrastructure**. Grounded on a B2B approach, EOSC must dedicate efforts for **coordination** and **federation,** mixing national and thematic training activities (RIs).

CORE EOSC TRAINING OFFICE FOR OPEN SCIENCE: Operate a coordination office for the production and collection of materials (courses, modules, curricula, etc.) to be distributed and adapted by members who will perpetuate/propagate the training.

FEDERATION OF LEARNING MATERIAL: Develop and operate a **Learning Material Registry** as a core supporting mechanism for both trainers and trainees in EOSC. Collecting "certified", classified and reusable (open and FAIR) training materials (OER), as these are produced across communities and countries.

TRAINING CERTIFICATION: Establish a credential mechanism that offers possibilities for official recognition (for both trainers and learners).

EOSC OPEN SCIENCE HELPDESK: Organize and operate a distributed and hierarchical structure with the country in the centre, addressing the technical-organizational and legal aspects of data and service Interoperability. Augment with RI/domain specific support.

Acronyms

AAI	Authentication and authorisation infrastructure
AI	Artificial intelligence
B2B	Business-to-business
B2C	Business-to-customer
DMP	Data management plan
e-IRG	e-Infrastructure Reflection Group
EC	European Commission
EOSC	European Open Science Cloud
FAIR	A set of principles to make data (F)indable, (A)ccessible, (I)nteroperable and (R)eusable
GDPR	The European Union's General Data Protection Regulation
IaaS	Infrastructure as a Service
ICT	Information and communications technology
INSPIRE	The INSPIRE Directive aims to create a European Union spatial data infrastructure for the purposes of EU environmental policies and policies or activities which may have an impact on the environment
IPR	Intellectual property right
KPI	Key performance indicator
LE	Legal entity
NFDI	"Nationale Forschungsdateninfrastruktur", Germany's National Research Data Infrastructure
NOAD	National Open Access Desk, the OpenAIRE network's national first points of contact
RDM	Research data management
RI	Research infrastructure
RPO	Research performing organisation
PaaS	Platform as a Service
SoS	System of Systems

1 | INTRODUCTION

This paper has a twofold role: (i) to strengthen and reinforce Open Science as the modus operandi in EOSC, i.e., embedding openness by default in all stages of research, and (ii) to highlight the important role of open scholarly communication and OpenAIRE participatory infrastructure as a vehicle to deliver Open Science in a trusted and reliable way.

1 – Defining EOSC as a **'Research Commons'**. A collaborative and community-led initiative such as EOSC builds on **existing infrastructures** to enable access to research data – going beyond discipline, geographical, political and social boundaries – and by doing so delivers great potential for Europe's researchers. This combination of socio-technical elements might be viewed as a 'Research Commons'[1] and should include the following key features:

- **Open Science by default**: embedding Open Science practices in all research activities to enable reproducibility, transparency and accountability.
- **Applying an Open Scientific Communications realm:** supporting scholarly publishing services for all types of research outputs at different research lifecycle phases, making publishing an *ex-ante* rather than an *ex-post* activity of research. As scholarly communication is currently largely 'closed' and this is a mismatch to the EOSC vision, EOSC could be the environment where open policies across disciplines and countries are developed, aligned and reinforced[2].
- **A holistic perspective on Data Flow:** devising an architecture and machinery which makes data smart and actionable, by (i) ensuring context is maintained along the data value chain, (ii) merging and combining small-scale data into bigger datasets, and (iii) accompanying data with services that provide secure and meaningful (intelligent) access.

2 – Successfully delivering the **technical and the human** aspects of EOSC. These are two important and complementary factors of delivering Open Science and this is where OpenAIRE can bring its experience to address the challenge of supporting a sustained revolution in science.

- **The technological aspect** includes not only traditional large-scale, pervasive research infrastructures, but also the various layers of machine intelligence required to ensure that there is seamless access to research outputs. All systems, whether national infrastructures, research infrastructures or research performing organisations and those supporting them such as libraries and data centres could operate seamlessly in a European research landscape devised by the Commission and, vitally, the Member States.
- **The human aspect** anchors the massive potential of the tools and services that technology provides through a combination of sound **governance, shared policies** and the **progressive and systematic training for Open Science.** This involves both researchers and data scientists, librarians and research administrators, repository managers and data centre staff - in short of all the active inhabitants of the entire ecosystem of knowledge creation in the European Research Area

[1] cf. e.g. Bosman and Kramer, 'The European Open Science Cloud as a Commons?'

[2] E.g. mandating articles are immediately available with a Creative Commons license, complementing the current activities of PlanS and OA2020[2].

2 | OUR VISION FOR EOSC

EOSC is about **sharing and re-using scientific content.** In an increasingly data-driven European research ecosystem, researchers should have unhindered and seamless access to resources to support not only access but tools for data analysis, as well as allowing other research support players (funder, institutions, libraries) mechanisms for monitoring and evaluation. These resources should be easily re-usable, facilitate effective collaborations across countries and continents, across disciplines, and not least support the opening up of science to shape and establish the connections to industry and society (for innovation and wider public involvement in decision-making processes). Europe being one of the top science production loci in the world has only to benefit from the sharing and re-use of scientific content. It needs to focus not just on providing the necessary infrastructure for enabling such sharing but also on the improvement of governance and sustainability, and on improving of the scientists' digital and science skills.

EOSC represents a **Research Commons**: a **co-created space in which new research information is produced and transformed into knowledge[3,4].** This in turn promotes new research questions, encourages re-use not only of formally published conclusions but the re-use of all underlying research outcomes (data, software, algorithms, lab books, etc.). It is a shared arena, maintained and preserved by a whole range of players in collaboration. This builds on previous concepts of the *digital commons*[5] and *infrastructure commons*[6] and thereby firmly sets the context that excellent research is the responsibility of all, and that sound research is also accountable to all those who own or have a share in this knowledge ecology, including citizens and industry.[7,8]

EOSC is about an **increase in transparency and accountability:** this allows a more efficient and effective way of monitoring research, in order to move from a top-down process of funding and fostering research and innovation to giving researchers greater power to determine accountable research. Researchers are not simply customers of various off-the-shelf service providers, whether public or commercial, but are able to shape and govern these services themselves[9]. In these ways, the scientists acquire ownership of both the data, publications and the relevant services, governance and infrastructure and are more willing to further share their research results. The creation of **trusted data spaces where such sharing may take place is expected to reconcile openness with business use of open data** and will further support partnerships between European RPOs and businesses.

---

[3] "Research Commons" is informed by fundamental discussions of the Commons in science and research as presented in e.g. Ostrom et al., *The Drama of the Commons*; or Frischmann, Madison, and Strandburg, *Governing Knowledge Commons*.

[4] Helfrich, 'The Logic of the Commons & the Market: A Shorthand Comparison of Their Core Beliefs | The Wealth of the Commons'.

[5] Kluitenberg, 'Constructing the Digital Commons: A Venture into Hybridisation'; da Rimini, 'Social Technologies and the Digital Commons'; and Ferrari, Scardaci, and Andreozzi, 'The Open Science Commons for the European Research Area'.

[6] Frischmann, Madison, and Strandburg, Governing Knowledge Commons; or Little, 'Tending the Infrastructure Commons'.

[7] cf. Fuster Morell, 'Governance of Online Creation Communities: provision of infrastructure for the building of digital commons'.

[8] Grossman, 'What Is Data Commons and How Can Your Organization Build One?'

[9] e-Infrastructure Reflection Group; Holmgren et al., 'National Nodes - Getting Organised; How Far Are We? Implementing e-Infrastructure Commons and the European Open Science Cloud'.

3 | OPEN SCIENCE IN EOSC

Openness, a success factor for EOSC. The European Open Science Cloud (EOSC) will only be successful if there is enough researcher engagement and uptake. Free (at the point of use), uncomplicated and intelligent access to data should be a core EOSC Unique Selling Point, adding value to existing EC and MS/AC investments. To reach a turning point for EOSC to fulfil its mission and vision, Open Science and specifically Open Access to research results (data, publications, software, protocols, as well as tools and services) should be a strategic priority, a key enabler into achieving a critical mass of accessible content for researchers to experiment with, stimulating the demand.

3.1 Researchers in the center

Until now we have primarily been focusing on data producers via mandates, where we need to turn it to a gain for the data consumers. EOSC should intrinsically support Open Science through a researcher-focused approach, taking into account the diversity of research cultures and maturity of infrastructures. We have identified five elements as key enablers:

1. Focus on researchers needs
2. Provide services through the eyes of a data scientist
3. Foster data-driven science, wherever it comes from
4. Reach out to researchers via a common language: scientific communication

Researcher needs

Open Science has different meanings to different people[10], but the emergent principles consistently emphasize openness and sharing of all research resources and outcomes. Whether this is about big science or science performed on a smaller scale, by individuals or small teams, the problems faced in gaining widespread support for Open Science[11] are about:

- Incentives (what is in it for the individual)
- Services and inclusiveness (availability and use of tools)
- Sustainability (where to invest resources)

Services provided through the eyes of a data scientist

The data science community has matured immensely over the last few years in automating processes and pipelines in general. The ultimate success for EOSC, apart from providing them with frictionless data flow and intelligent discovery, is to i) optimize the time a data scientist spends redoing the same workflow by re-using services and data from others, and ii) give them the facilities and capabilities to develop new data analytics services. In brief, EOSC should act as the facilitator of Open Science and

[10] Tennant et al., 'Foundations for Open Scholarship Strategy Development', section 4.2.2.
[11] National Academies of Sciences et al., Broadening Access to the Results of Scientific Research.

collaborative practices[12]. By combining the two EOSC creates a user-friendly environment for researchers with the following offerings:

- enable a reproducible 'digital laboratory' as a standard way of working
- offer collaborative tools (VREs, Scientific gateways), applications capable to offer an adequate level of transparency to EOSC users, based on their IT skills and domains.
- create abstraction of data manipulation, including reporting, pipelines, parallelism, persistence
- deploy solutions without concerns about dev-ops / cloud ops

Data-driven science: small science/data vs. big science/data

Open Science on one hand, which focuses on the communication, i.e. sharing and reproducibility of data, and computing cloud on the other, are interchangeable components of data driven research. Their complementarity builds bridges within, but also between big and long tail of science, and is essential into engaging researchers in the diverse European landscape:

1. From big data to smart data, processes and outcomes: To shape the services offered to researchers, research infrastructures have based service development on well-defined and homogeneous user communities, using large-scale facilities that generate big volumes of data, stored and processed in dedicated data centres, subsequently accessed via agreed-upon rules. But the true value of big data lies in its application to a scientific process or purpose: the way it is actioned on a research environment and subsequent context. Challenges, once limited to huge machines managed by international teams, are now beginning to crop up in small devices used by single researchers. Consequently, more researchers need to assume the role of "data scientist". A big challenge therefore is to turn big data sets into actionable and smart data, i.e., connecting, organizing, and packaging it into making insights available to all (accessible), easy to apply (understandable), and focused on the task at hand (actionable).[13]
2. From small fragmented data to integrated big data: A large portion of scientists' work is done in a small science environment, i.e., individuals and small teams collect data for specific projects, who often contribute to more radical and disruptive innovations[14]. These localized environments produce small volumes of un-homogenized, often unstructured data. This data is able to answer targeted queries whose value is becoming all the more important to artificial intelligence (AI) solutions[15]. In a wider context the development of new data infrastructures enable these small data to be merged, scaled and linked into larger datasets with the help of data analytics, ultimately enabling sharing and reuse.[16, 17] In that sense, the role of scholarly communications, particularly in terms of a mass-micro scale, becomes quintessential for the promotion of actual Open Science.

[12] Open Science relies on collaborative practices, but at the same time we may have collaborative practices that do not constitute Open Science.
[13] Fortin and Currie, 'Big Science vs. Little Science: How Scientific Impact Scales with Funding'.
[14] Howe, 'Listen: Small research teams are better than big ones at disrupting science'.
[15] https://channels.theinnovationenterprise.com/articles/the-role-of-unstructured-data-in-ai
[16] Many RIs are trying to achieve this. Examples: OpenMinted & OpenAIRE for textual scientific publications
[17] Use case on neuroscience with long-tail data compared to Human Brain Projects (US & EU): Ferguson et al., 'Big Data from Small Data'.

A common understanding, a common language for Open Science

The diverse research communities and cultural variety of Europe will accept no "one-size-fits-all" solutions for doing open science and yet there is a need to find a common basis, a **common language around scientific communication** to bridge views, perspectives and gaps. The objective is to engage researchers as the main consumers of EOSC, and to enable them to share and re-use results and resources, they should be able to rely on sustainable infrastructures that support the **publishing of, and reporting on, research results at any stage of the research life-cycle**.

Use of scientific communication practices as integral components of Open Science are necessary to bridge the different perspectives and research phases in EOSC. While scholarly communication, along with scholarly publishing, has been attributed with slow adaptation to change, many areas are now beginning to employ Open Science practices[18] to enable transparency, reproducibility and accountability, across all phases of the research life cycle, from the production of raw data, to its reuse and exploitation. This is explicitly pointed out by the consistent use of identifiers for all research related results and entities, metadata descriptions, citation practices, licenses, certification.

> Deep changes have affected scholarly publishing, but the process itself has remained remarkably stable. It includes four key functions that have accompanied scientific publishing since the 17th century: registration (attribution), Certification (Peer review), dissemination (distribution, access), preservation (scholarly memory and permanent archiving). Evaluation is another function that has been associated to scholarly publishing in the last few decades, ... Digital technologies do not disrupt the publishing functions, but they allow for their distribution among different actors...
>
> *EC High Level on the 'Future of Scholarly Publishing and Scholarly Communication' 2019*

3.2 Transforming & embedding open scientific communication in EOSC

Despite the hype, the effective implementation of Open Science is hindered by several cultural and technical barriers. Researchers embraced digital science, use "digital laboratories" (e.g., research infrastructures, thematic services) to conduct their research and publish research data, but practices and tools are still far from achieving the expectations of transparency and reproducibility of Open Science. **Publishing is still a post-experimental**, tedious, manual process, too often limited to journal articles, in some

> Today's scientific communication infrastructure practices, workflows, services, and philosophy still follows traditional paper-centric paradigms, where datasets are published in a very similar fashion.

[18] We use the term Scholarly and Scientific Communication interchangeably: "Scholarly communication [...] refers to any form of exchange used by scholars and researchers to participate in the elaboration of knowledge through critical discussions and conversations with fellow humans. This encompasses all the procedures, from the purely informal conversation to the highly formalised stage of "publishing". In fact, scholarly publishing can be defined as the formalised sub-set of scholarly communication." Directorate-General for Research and Innovation (European Commission), 'Future of Scholarly Publishing and Scholarly Communication', p.14.

contexts semantically linked to datasets, rarely to software, generally disregarding digital representations of experiments.

For open science to be in the centre of EOSC, and for researchers to be truly engaged, we need to build upon and introduce key changes in existing scientific communication practices as follows (Figure 1):

1. Publishing all kinds of scientific products, literature, data, software, other research artefacts.
2. Publishing semantic links: links between scientific products are increasingly as essential as the products themselves in navigating, understanding, evaluating, and re-using the research life cycle and its results.
3. Publishing experiment products, intended as digital representations of the steps of an experiment, including input products, RI services adopted, service configuration, etc.; experiment products need to enable the reproducibility of the experiment itself, to facilitate re-use, assessment and evaluation of scientific results.
4. New publishing/dissemination practices and methods, including automated publishing, but also looking into other innovative dissemination methods/practices[19]
5. Quality control and securing the quality, reproducibility, FAIRness, etc. of research results, with innovative methods and practices, beyond traditional peer-review such as open peer review, commenting and social networking, automated/AI tools, etc.
6. Assessment and rewarding which combines more intelligent use of more diverse / open / auditable metrics, with qualitative assessment[20]
7. Monitoring Open Science as this is key for Open Science itself, to ensure its implementation and adjust the aim to establish better policies and incentives.

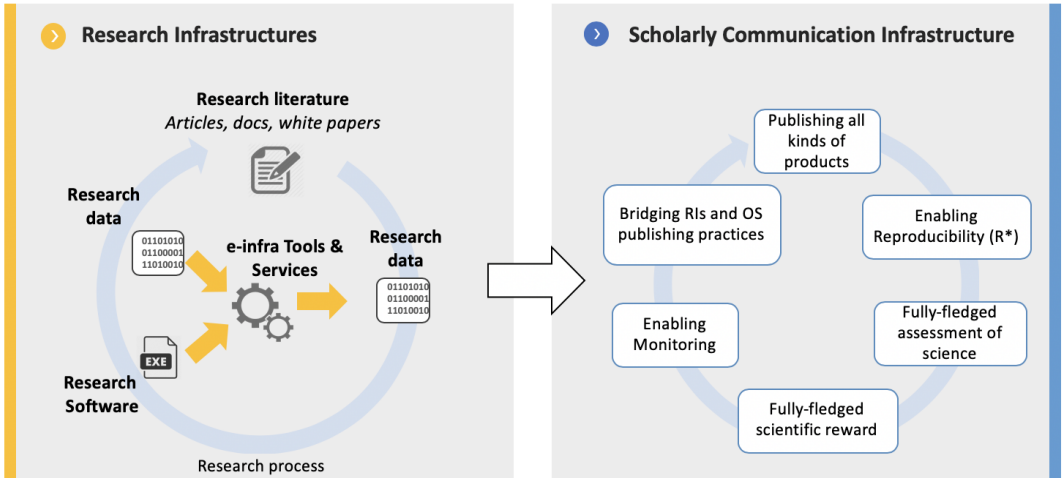


FIGURE 1. RESEARCH LIFE-CYCLE MAPPED TO SCIENTIFIC/SCHOLARLY COMMUNICATION PROCESSES

[19] https://comments.coar-repositories.org/5-pubfair-architecture/ - see paragraph 5.3
[20] OSPP recommendations for https://ec.europa.eu/research/openscience/pdf/integrated_advice_ospp_recommendations.pdf

EOSC comes in a timely manner and can become the catalyst that transforms current scientific communication practices towards openness. This transformation should also help into breaking down artificial fragmentation of three existing semi-autonomous areas of infrastructure (Figure 2), unify these and present a common interface to researchers and other actors involved in the research ecosystem.

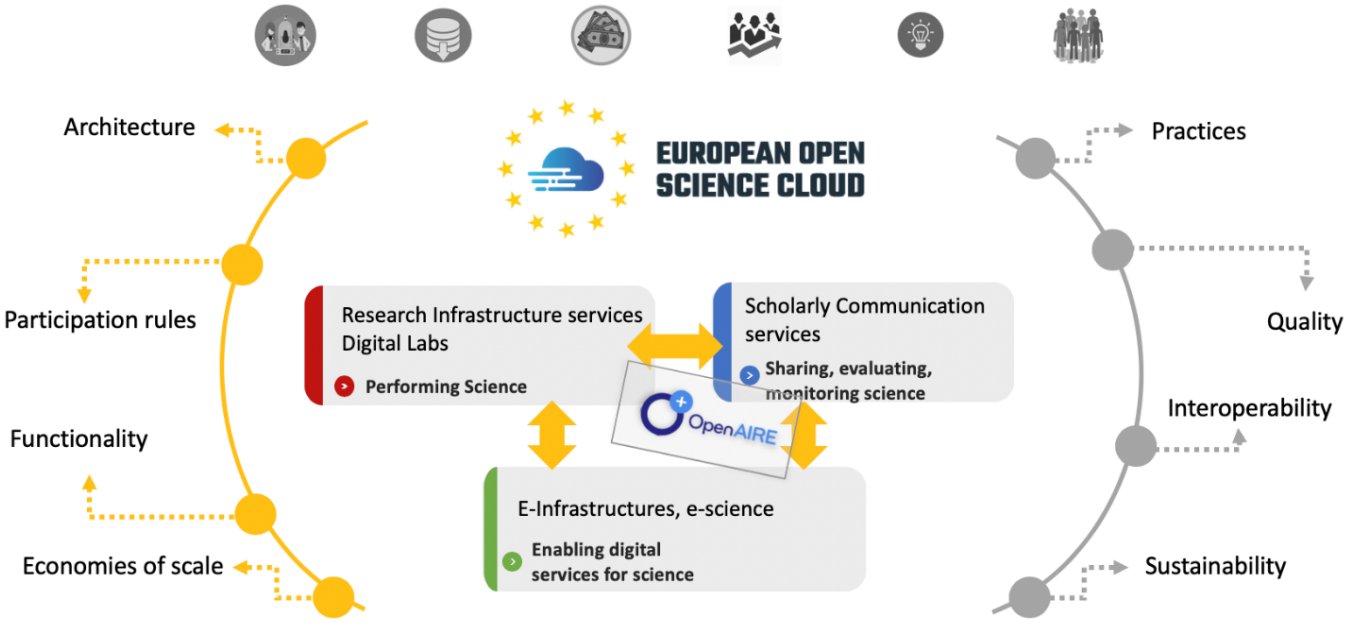


FIGURE 2. SCIENTIFIC COMMUNICATION: A COMMON INTERFACE FOR EOSC.

4 | EOSC DECONSTRUCTED

EOSC, as with any other digital infrastructure, needs to comprise a solid and dynamic interaction between **technological** and **human** elements to ensure its effective operation and pervasiveness within researcher circles. Each of these elements has internal hierarchies and sublayers addressing primarily different phases of the research life-cycle and different needs across borders and disciplines. In the end, these are bound together by a **participatory, open and transparent governance** layer that is upheld by all those in the Research Commons, which includes an agreed organizational, technical, legal and ethical framework, or else policies that are expressed in the form of Rules of Participation.

> EOSC as a Research Commons is an open resource, with standards that seem permissive in contrast to proprietary and commercial enterprises, but which still are subject to regulation. Oversight and management of this research commons are crucial elements of governance to ensure sustainability and adherence to the community's values and collective interest.

4.1 The Technological Layer

Given the diversity in requirements, the variety of business models, and the need for agility and flexibility to foster innovation, the EOSC architecture *should avoid at any cost monolithic and centralized approaches*. An EOSC architecture based on strict service participation rules cannot be regarded as a viable solution, and it should therefore be driven by **"System of Systems"** (SoS) principles[21]. SoS's conceive integration of systems by respecting community system practices, facilitating convergence of such systems towards an integrated system by the incremental introduction of recommendations, policies, and participation rules based on worthiness and convenience rather than on membership obligations.

Overall EOSC can be seen as a set of interconnected digital resources which allows frictionless data flow, (intelligent) discovery of scientific content, and production of new knowledge (insights) via homogenized and secure access to resources that transfer, store and analyse data. With the research commons approach, we envision a distributed and participatory environment with **emphasis on B2B (business to business)** data sharing and resource accessing, with **non-regulatory measures**, such as (i) fostering the use of APIs for simpler and more automated access to and use of datasets; and (ii) developing recommended standard contract terms.[22]

To better understand and communicate EOSC to researchers and policy makers, we have deconstructed it into four functional layers, taking into consideration the following preconditions:

[21] See for example a vision of how the different component parts of EOSC should fit together: Romier and Fede, 'D6.8: Final EOSC Architecture | eoscpilot.eu'.

[22] European Commission's 'Communication "Towards a Common European Data Space"', https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space

1. Making **existing, funded localized services** the cornerstones of EOSC architecture, which can efficiently mitigate risks of EOSC becoming an artificial *de novo* construction. Making efficient use of the investments that governments already provide for institutional services in Universities (HEI) and RPOs (libraries, IT Services, data centres) and solving the ubiquitous well-known communication problem of the "last mile". Fundamental IT and repository services are successfully serving science today and the technological, financial and governing interface between those and EOSC need clear specification.
2. **An EOSC "user" is contextual**: the primary end-user may be defined as the "researcher" – there are, however, local infrastructure services (e.g. libraries, IT Services, funders) who are also users of EOSC, in that they provide or support EOSC-related services locally in the institutional context of Universities and RPO's with technology adaptation, training or helpdesk activities.
3. It is **architecturally key to include the dimension of *openness*** across all functional areas of the system. Openness of science should ensure its reproducibility, transparency, and collaborative flavour, as such it cannot be considered as post-production feature, a mere access policy issue. Scientific services should be "aware" of being part of a production chain that ends with the publishing of products to also **provide assessment, discovering, monitoring, and re-use**. In this context service classification and service interoperability touch on Open Science-specific aspects, with a further dimension of cross-discipline and discipline specific perspectives.

The four layers of EOSC

The **Services Layer** includes the diversity of services developed and operated by research, national or institutional infrastructures. These four dimensions cover both sectoral and geographical/ regulation elements, i.e. the specific research disciplines or areas and the national or European governance and regulation ecosystem within which Open Science take place. The **Data Interoperability Layer**, and the **Access Interoperability Layer** are two key conjoined layers fostering *federation* and *compliance* for Data and Access. Their role is to break down siloed approaches and enable seamless access for researchers through technical, semantic, organizational (business) and legal interoperability[23]. The **Monitoring layer** is essential in

1. Services Layer. This layer consists of all services provided in various settings: institutional, national, European/RIs. It is the nucleus of the decentralized EOSC, reflecting the diversity of Europe with local practices and investments, providing the right climate of agility and flexibility needed for innovation. Services can be classified into:

- Foundation services: storage, computation, networking and data management provided by data centers. This includes the infrastructure layer that is based on virtualization technology where the service providers offer virtual machines as a service to the end-users (IaaS), as well as Platform as a Service (PaaS) offer where users are enabled to create applications for low- level data management aspects
- Research enabling services: analytics, publishing, sharing and discovery services allowing data driven science at the global and discipline level. Services at this level may be horizontal (e.g. publishing in

[23] The European Interoperability framework https://ec.europa.eu/isa2/eif_en

OS, analytics from specialized to generalized AI), but many need to take into account the idiosyncrasies of the scientific disciplines or the underlying data.

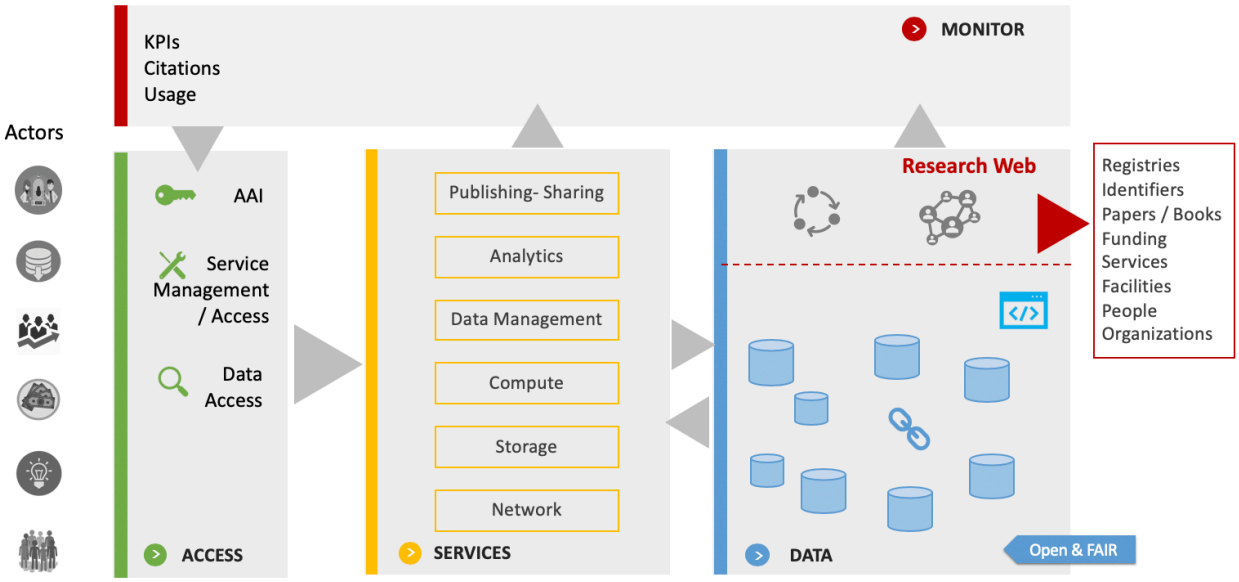


FIGURE 3. EOSC DECONSTRUCTED: SERVICES, DATA, ACCESS

2. Data Interoperability Layer. At the centre of EOSC there is an ever expanding, on-demand body of data and knowledge content providers that contains over 250 million publications, datasets, experiments, workflows and protocols[24]. The data generated, stored, transferred, analysed, consumed is an integral part of the European common research data space[25], i.e., a seamless digital area with the scale that aspires to enable the development of new products and services based on data. In order for EOSC to fulfil a frictionless data flow, it must embed a data compliance framework for open / FAIR data which dictates and applies the rules of how the data elements are published, shared and re-used both at the discipline-specific level and at cross-discipline level. This multi-tiered view of data must be taken into account while building directives for interoperability as different fields bear specific community requirements, may share some of them with other fields (e.g. INSPIRE), and may share some with all the others (e.g., citation level, FAIR Maturity RDA WG). This compliance framework, which includes standards, agreed upon protocols and APIs, vocabularies for semantics, licenses, etc., can be viewed in a similar way to other EU initiatives such as INSPIRE[26] for geodata, GDPR[27] for personal data, PSI for Public Sector Information.

Interconnecting data & science: the Research Web. This added value layer contextualizes research by recording the provenance and usage of all digital records and interlinking them to all entities involved in the research lifecycle (people, organizations, facilities, funding, equipment) and beyond (innovation,

[24] For example, scholexplorer lists all content providers continuously updated.
[25] European Commission - 'Communication "Towards a Common European Data Space"', https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space
[26] https://inspire.ec.europa.eu/
[27] https://eugdpr.org/

society). In a linked digital world this layer is paramount in EOSC as it effectively delivers an *interlinked* Research Commons enabling:

- Transparency, reproducibility and scientific reward;
- Interrelations of big science data infrastructures or smaller interlinked environments[28];
- Interdisciplinarity;
- Policy monitoring (openness, research impact. and innovation).

3. Access Interoperability Layer. To converge the diverse solutions in Europe, and to create EOSC as a system of systems, it is important to assure a *uniform delivery of services and data*, across borders, disciplines and commercial/public sector. Respecting the diversity and inclusiveness of existing investments and practices and ensuring that specialized infrastructures/services retain their unique function, this layer implements the EOSC *digital commons framework.* A B2B mechanism towards creating trusted environments for researchers and service providers alike. It brings measures from the world of the digital marketplaces, such as *standard contract terms* for services and data access, *specifications* for quality (certification), security and privacy (including trust and identity). It supports *categorization* (to simplify browsing), *curation* (to deliver quality), *cataloguing* (to offer intelligent discovery), *crowdsourcing* (to engage researchers for active participation). It may be broken down to the specific frameworks:

- Authentication and Authorization Interoperability (AAI) framework, a trust and identity service for researchers to seamlessly access all EOSC resources.
- Data access framework, whose primary role is to offer data as a service, enabling open interfaces where data consumers are able to discover and retrieve best-fit data to meet their needs. Of particular interest is the architectural model conceived in industry from the International Data Spaces Association[29] which details the different roles and addresses data sovereignty, parts of which may apply in sharing (open) research data.
- Service management and access framework, for bringing a consistent and agreed upon understanding of e-science services, enabling their intelligent discovery and access (what they offer, which science problem they address, what is their operational capacity, how they are accessed, who pays for them).
- A minimum legal metadata framework, for ensuring openness and interoperability, privacy and security. These will include among others, copyright status, disclosure limitations, patents pending, other IPR on the datasets or workflows, existence of personal data, designation of data as PSI, etc.

4. Monitoring Layer. Essential for the success and sustainability of EOSC is the assessment of its components (policies, access framework, services, data) through their usage. This overarching layer is a core function, where we collect in a normalized and agreed upon manner all types of usage and performance related data for all individual components of EOSC. This would, on one hand, reflect researcher uptake and interests (innovation), and on the other it would provide a most beneficial feedback on how to assess and shift where necessary policies, business, funding and usage models.

[28] Cf. Bateson, 'Warm Data: Contextual Research and New Forms of Information', and Demchenko et al., 'Addressing Big Data Issues in Scientific Data Infrastructure'.

[29] https://www.internationaldataspaces.org/wp-content/uploads/dlm_uploads/2018/04/InternationalDataSpacesAssociation_ReferenzArchitecture2.0.pdf

4.2 Who is the Human Layer?

Ideally the implementation of EOSC will ensure a balance of policy, training and support implementation to ensure take up at Member State level and among research communities. A network of professionals who will support the uptake of services and training is an essential infrastructure to make EOSC succeed.

Who Comprises the Human Layer?

The Actors: *The growth in data science requires relevant data professionals to help researchers choose, evaluate and navigate the plethora of tools and resources available to get the most from their data and how to publish data.*	
Experts in Open Science	National nodes who are able to develop policies, support policy implementation, and national level training around EOSC services.
Experts on legal and ethical issues	Mediating on issues such data ethics, IPR, copyright and licensing. Support for these issues actively encourage data sharing.
Domain specific experts	Expert data curators & data scientists and data stewards to support different disciplines with the institution.
ICT Operations Experts	Service providers for data management services
Data Scientists / Engineers / Computer science experts	Leading the community to the next level of data management services, whose role is to take EOSC to its full AI potential
Business experts	Defining new models for revenues and money pathways
Information/Social scientists	Assessing user behaviours and guide the design and processes for the adoption of services.
Intermediaries: *This layer of the Human infrastructure comprises the decision makers who set the framework and policies in place to implement and support EOSC*	
The research performing institution	At research-intensive institutions, academic professionals hold a balanced position in the research chain, working as mediators and stewards for EOSC services, encouraging research standards, advocating Open Access policy as well as tools for Open Science practice, promoting technical interoperability, demonstrating the research incentives and rewards in doing Open Science
Policy makers	Role to engage with setting regulations and liaising with bodies on competition and research
Research Infrastructures	A key entry point into EOSC, secure uptake of EOSC vertical services
Scholarly communication ventures	Publishers and learned societies will be consumers of a shared research data environment. Learned societies have to manage their data assets and ownership workflows and EOSC data storage and curation services will be of use.

TABLE 1. ACTORS AND SKILLS FOR OPEN SCIENCE

4.3 EOSC - A multi-level endeavour

"Too often in the past we have used technical approaches, such as federation, to combat the fear that a system can be co-opted or controlled by unaccountable parties. Instead we need to consider how the community can create accountable and trustworthy organisations. Trust is built on three pillars: good governance (and therefore good intentions), capacity and resources (sustainability), and believable insurance mechanisms for when something goes wrong. These principles are an attempt to set out how these three pillars can be consistently addressed."[30]

Open Science is a complex issue: it is perceived in different ways by different groups. With a diverse Europe as a backdrop, we envision different layers of implementation and coordination within EOSC, with accompanying responsibilities. The three basic elements for delivering Open Science, **policies – services - training**, need to be tackled at all levels, with varying attention to details, from *macro* to *micro,* from *federating* to *implementing*.

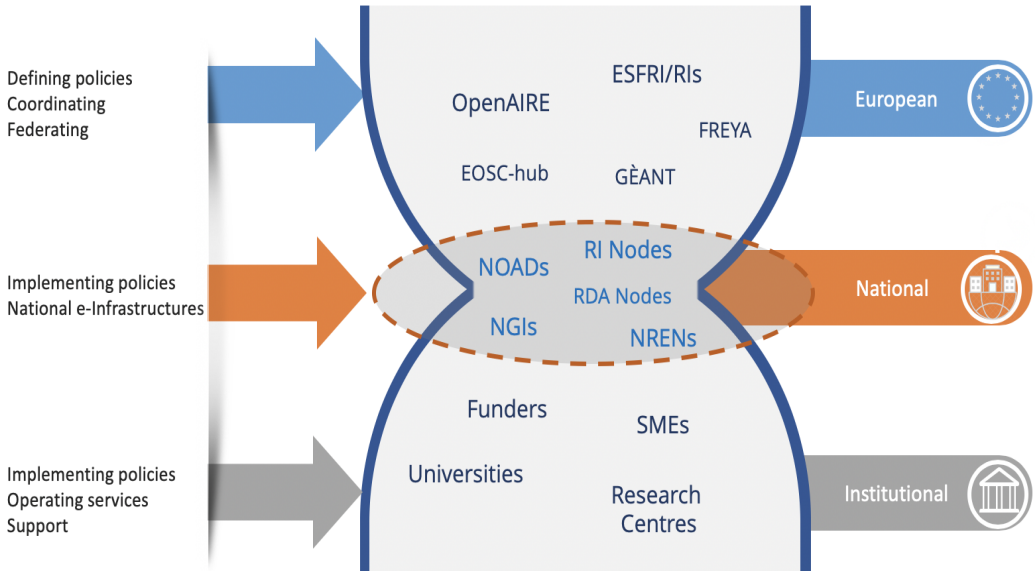


FIGURE 4. EOSC PLAYERS, ROLES AND INTERPLAY

- **Institutional** settings deliver implementation and support to researchers (where research happens) via detailed policies, services (e.g., repositories, DMP, eduGain) and support (e.g., trained personnel in library and research office).
- **National** settings set the policies and legislation (if needed) at member state level and play a coordination role, matching the needs of local interests, idiosyncrasies and investments, ensuring operations of national infrastructures and their linkage to EOSC.
- **The European Commission**, with its funded digital infrastructures, **guides** and **coordinates** the efforts on policy implementation and facilitates access to services. This also includes EU agencies that are located in one country but maintained by many (e.g., CERN).

[30] Bilder, Lin, and Neylon, 'Principles for Open Scholarly Infrastructures-V1'. For further reading, see also Neylon, 'Sustaining Scholarly Infrastructures through Collective Action'.

In this respect, EOSC is not meant to own and operate everything, but rather to actively create a space for bottom-up innovation. In particular, the national nodes who are already set up by the major e-Infrastructures and Research Infrastructures (RIs), have a central role to play in the new era brought by EOSC. Seeking synergies and convergence under a national EOSC structure model, as no one of these nodes alone is able to address all aspects of EOSC, these *National EOSC Nodes* will be able to address policies, services and training in a holistic and sustainable way, and:

- provide fit for purpose services to researchers, embedding open science practices and applying economies of scale by default;
- act as a seamless bridge to national (e.g., PSI), EU and world infrastructures;
- bridge open science, cloud and HPC;
- be the long-term, trusted organization for ministries and research performing organizations alike.

The *National EOSC Nodes* may follow different models from country to country, ranging e.g., from one organization encompassing all EOSC "departments", to partnerships (i.e., a virtual office with people from different organizations[31]), and they may cater to different priorities based on the needs of national research strategies.

FROM A 'FEDERATING CORE' TO A 'FEDERATED COMMONS'

In order for EOSC to gain traction and gain positive value in the eyes of the Member States and Research Infrastructures is to avoid fragmentation and push forward the message that EOSC provides value added services that would simply not be affordable at national level.

This is where the concept of the 'federating core'[32] should be complemented by a framework of services, policies, training and shared governance, in short, a *Federated Commons*. To reach the Open Science goals, such a 'Federated Commons' would need to minimally comprise of:

- Shared Policies for open science, including rewarding the practice of Open Science across countries and funders. These need to be consistent, standardised and machine readable.
- A robust and pervasive training programme for researchers that can be tailored to individual national needs and make use of local expertise: Train-the-trainers approach
- An increasingly common legal framework expressed in common and interoperable policies and RoP's on IPR, Personal Data, and for data re-use, and the eventual straightforward sharing of all research outputs (software, algorithms, lab books, etc.). This should also include clear and simple guidelines and platforms in relation to procurement as well as competition law/state aid rules guidance.
- Access policies, whose fundamentals are accepted and embedded in research policy and national strategy by Member States.[33]

[31] An example is the Danish e-Infrastructure cooperation https://www.deic.dk/
[32] 'D2.3: EOSC-Hub Briefing Paper - EOSC Federating Core Governance and Sustainability Public | EOSC Hub'.
[33] e-IRG Reflection Group, Holmgren et al., 'National Nodes - Getting Organised; How Far Are We? Implementing e-Infrastructure Commons and the European Open Science Cloud', p.4.

This would be more of a long-term asset to the EOSC than for example one providing a single-task service. Member States are increasing doing this, desiring best practice information to solidify their own confidence in undertaking long-term investment in research infrastructure.[34]

[34] The National Research Data Infrastructure (NFDI) in Germany is an excellent example of this, whereby the authors of the NFDI tender documents clearly specify where reference to international practices and standards will concretely underpin their own endeavours, and will be necessary to them, and make them more successful in operation, despite being largely defined as an infrastructure serving the needs of a national policy and its research base. See Deutsche Forschungsgemeinschaft, 'Guidance Notes on Funding Criteria National Research Data Infrastructure (NFDI)'.

5 | OPENAIRE AS A KEY PILLAR FOR OPEN SCIENCE IN EOSC

OpenAIRE contributes to different components of EOSC's strategic implementation plan[35], while **leveraging existing national scholarly communication infrastructures** as EOSC providers. The following table illustrates how OpenAIRE builds Open Science structures addressing all levels of EOSC, national, institutional (or thematic) and international, following a three-pillar approach of **policy** formation & alignment, **training**, and **services** (B2B, B2C).

	Policy	Training	Services
National	Support national funder alignment with pan-European policies for Open Science, open access, research data management, FAIR via NOADs and national Open Science Helpdesks	Facilitate train-the trainer approaches for all stakeholders, synthesising existing national training initiatives.	Connect and Integrate national infrastructures as key providers. Facilitate publishing of all research artefacts (data, software, citations, publications)
Institutional / Thematic	Align discipline specific policies for FAIR, RDM, Open with EU/national	Training on OS topics: FAIR, Copyright, IPR, Policy	**RIs:** Research Community Dashboard, Repositories: Metadata Guidelines and Exchange **Scholarly Communication:** publishing and linking of research artefacts. Researchers: Explore - OpenDMP
International	Facilitate alignment with international OS policies and their implementation: research data management, open access	Create and contribute to new training practices to support Open Science implementation via the Community of Practice[36]	Connect to global research systems: global metadata for research outcomes, research via transparent systems supported by the Open Science research graph.

TABLE 2. OPENAIRE MODES OF OPERATION

The following sections present what we consider to be the key priorities in EOSC, and how OpenAIRE contributes to the adoption of Open Science in this.

[35] Jones and Abramatic, 'European Open Science Cloud (EOSC) Strategic Implementation Plan'.
[36] https://www.openaire.eu/cop-training

5.1 Policies and Governance at the national level

EOSC will converge at national support structures, and this is at the heart of OpenAIRE's participatory design with **34 National Open Access Desks (NOADs)** who coordinate Open Science activities and interoperation of services already existing at institutional and national levels. Given their established role and expertise, NOADs have become an integral part of the national EOSC structure and culture for supporting **Open Science implementation** of EOSC.

KEY FOCUS AREAS AND PRIORITIES FOR EOSC

OPEN SCIENCE IN NATIONAL EOSC STRUCTURES: Building on theoretical underpinnings around management of the commons that highlight the need for polycentricity and adaptive governance[37], EOSC should embed a variety of different national infrastructure models becoming connected via interoperable standards and the fostering of interdisciplinary personal exchange. OpenAIRE's national first points of contact, the National Open Access Desks (NOADs) will play a fundamental role in the formation of this linked network of national EOSC structures. Acting as nodes of Open Science support, policy information, training, and coordination, they will ensure open collaboration with existing initiatives and organizations. Through NOADs, scholarly communication experts will help seek connections between the different fields, acting as expert translators who are able to help communicate between researchers, funders, technologists, educators, students, and the general public. OpenAIRE, as the articulating medium, can be both a horizontal conduit (EOSC Clusters, Thematic and discipline-specific RIs), as well as a two-way channel between the Member States' research goals and interests on the one hand, and EOSC's oversight role on the other.

OPEN GOVERNMENT PARTNERSHIP. With the EU's established stance towards Open Science as maybe *the* key factor to define public funding in Europe for the foreseeable future, it is clear that Open Science should now play an important role in national policies, and on the agenda of all governments involved. Correspondingly, EOSC – as a key pillar of Europe's strategy towards a common data space – requires strong commitment of all Member States. That said, many Member States have already committed to the principles of the Open Government Partnership (OGP), which since 2011 has had as its main aims, to "promote accountable, responsive and inclusive governance."[38] This commitment is not merely a signalling of good intentions; the framework has been actively used by a majority of OpenAIRE NOADs[39,40] as a way of embedding and strengthening their support for both Open Access and Open Science by ensuring that research policies also make explicit reference to the needs of the wider society. OpenAIRE is adopting and adapting the OGP paradigm as one of the foundations for expanding and increasing the social and network effect of its NOADs.

[37] Ostrom et al., 'The Future of the Commons - Beyond Market Failure and Government Regulation', p.60; Dietz, Ostrom, and Stern, 'The Struggle to Govern the Commons'; and Johnson, 'From Coalition to Commons'.
[38] https://www.opengovpartnership.org/about/. Accessed on 25.07.2019.
[39] Currently 25 of 36 NOADs are signed up to the OGP.
[40] Example: France https://www.opengovpartnership.org/countries/france. The National plan for Open Science, launched in July 2018 refers to OGP explicitly.

5.2 Infrastructures and services for federating data, for federating Open Science, for monitoring Open Science

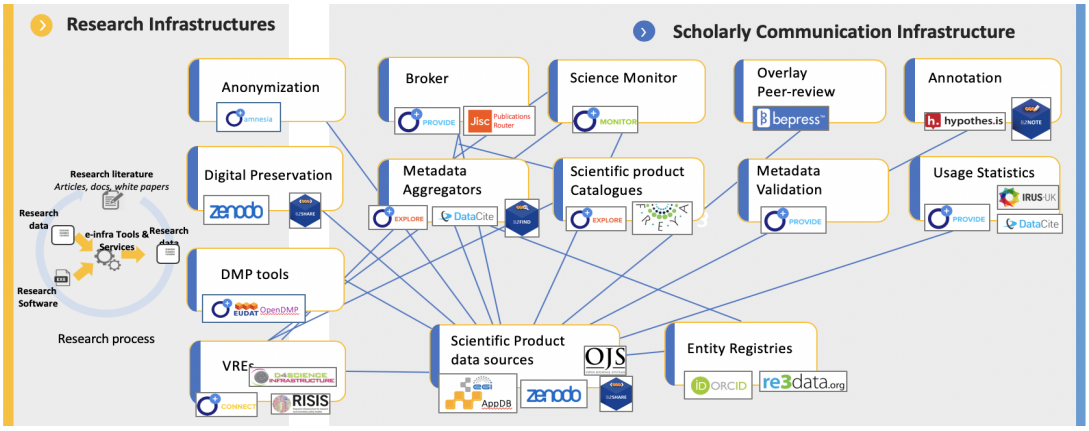
For Open Science to succeed in EOSC, we need to i) provide services for all stakeholders involved in the research life cycle: researchers, support personnel at RPOs, funders and policy makers, ii) ensure all services are interlinked via a data federation mechanism, allowing both small and big data to become an integral part of EOSC, iii) embed services in institutional settings where research happens and where support is given, and iv) take into consideration the international context and global research and scholarly communication infrastructures.

KEY FOCUS AREAS AND PRIORITIES FOR EOSC

SCHOLARLY COMMUNICATION SERVICES AS A MEANS OF OPEN SCIENCE FOR RIS AND NATIONAL E-INFRASTRUCTURES. A key added value of EOSC is to identify commonalities across disciplines for implementing Open Science and develop/operate services in a horizontal way to avoid duplication of efforts. OpenAIRE's scholarly communication approach should be the **touching points of national vs. cross-RI data federation**. Emphasis should be placed on **common OpenAIRE-RI efforts** so that RIs adopt and adapt to the emerging scholarly communication reality. The following figure illustrates existing efforts that bridge scientific/thematic RI services to scholarly communication services, to seamlessly publish research data and their semantic links at real-time in the scholarly communication infrastructure (incl. experiments, software, etc.) on behalf of the scientists. EOSC needs to further strengthen these ties, under well-defined rules, roles and responsibilities.

OPEN RESEARCH GRAPH: A crucial step in breaking down the fragmentation in EOSC is to create an interlinked graph of all its resources: research results, who has produced them, under which grant, using which facility and e-science service, how open and fair it is, how it has been used. OpenAIRE's Open Research Graph, developed as a decentralized, trusted, interoperable graph is a **core shared resource (asset) of EOSC** that converges many efforts and services. With complementary effort from RIs to include discipline domain knowledge this graph is envisioned to be the foundation for:

- An overarching EOSC scientific catalogue enabling contextual discovery of research
- A link to PSI (open government data) and other knowledge sources



OPEN ANALYTICS FRAMEWORK AND INFRASTRUCTURE: For Open Science to gain traction, EOSC must embed an infrastructure for scientific reward which gathers all types of usage data for all types of resources (citations, usage events for data, services, software). Part of the *Research Commons*, this infrastructure should traverse all EOSC layers, with emphasis given on B2B aspects (specifications, formats, APIs) rather than indicators and services/tools for self-assessment, as the latter need additional processes for bringing consensus among the different players and may transform over time. As this is an international undertaking, EOSC governance should foresee mechanisms how these are operated in EOSC and linked to international structures.

6 | OPENAIRE OFFERS A SERIES OF B2B AND B2C SERVICES THAT SUPPORT, ACCELERATE AND MONITOR OPEN SCIENCE (SEE DETAILS IN 1 |

Appendix) corresponding to the four conceived EOSC layers in Figure 3:

- **Services Layer:** Zenodo, Amnesia, OpenDMP, Research Monitor/Analytics, Research Community Dashboard
- **Data Interoperability Layer:** Guidelines for Metadata Exchange, ScholExplorer, Validator, Usage Statistics, OA Broker
- **Access Interoperability Layer:** Content Provider Dashboard (PROVIDE), Open Research Graph (EXPLORE)

6.1 Coordinated Training and Support at the EU level

Digital and Open Science skills are a cornerstone in EOSC's operations and future. One of EOSC's great opportunities is to build long lasting capacities through a pan-European coordinated training infrastructure that empowers individuals and institutions to come up to speed with Open Science developments. In accordance to the New Skills in Europe Agenda[41], EOSC should develop and rollout a **robust** and **pervasive Training and Support Infrastructure** shadowing the EC's adopted steps for establishing *human capital - employability – competitiveness*.[42] Grounded on a B2B approach, EOSC should dedicate efforts for training on **coordination** and **federation,** mixing national (e.g., NOADs, NGIs, NRENs) and thematic (RIs) training activities.

OpenAIRE brings coherence to the EOSC training and support landscape by leveraging the unique potential and placement of its NOADs, present in every EU country, to train stakeholders and build the relevant support networks for researchers and data practitioners. The training programme[43] addresses the current skills and content gaps and takes a multiplicator *Train the Trainer* approach for face-to-face training activities and support efforts.

[41] https://ec.europa.eu/digital-single-market/en/news/digital-skills-core-new-skills-agenda-europe

[42] Communication "A New Skills Agenda for Europe: Working Together to Strengthen Human Capital, Employability and Competitiveness"

[43] A Train the Trainer programme to enhance and sustain a trainers' network (building upon FOSTER Open Science trainers' network) https://www.fosteropenscience.eu/

KEY FOCUS AREAS AND PRIORITIES FOR EOSC

CORE EOSC TRAINING OFFICE FOR OPEN SCIENCE: A coordination office for the production and collection of materials (courses, modules, curricula, etc.) to be distributed and adapted by national members who will perpetuate/propagate the training. Building on the emerging European Training Community of Practice[44] this office must bring together trainers from various EU related initiatives to develop a consolidated cross-infrastructure training package for data skills, data science and data stewardship.

FEDERATION OF LEARNING MATERIAL: The diversity of communities, learning methods, data relevant and Open Science topics assumes that learning materials come in many shapes and sizes. Complementary to the EOSC portal and administered by the *EOSC Training Core Office on OS*, a Learning Material Registry will be a core supporting mechanism for trainers and trainees in EOSC alike. Collecting "certified", classified[45] and reusable (open and FAIR) training materials such as these are produced across communities and countries, it will serve as an online information network designed to organize and vet academic content for trainers.

CERTIFICATION: As every EOSC service, training must be owned and operated by individual national members making use of local expertise, investment and organizational characteristics. EOSC, as a connecting and coordination mechanism for all things related to research data and e-science services, must professionalize the underlying training infrastructure (as set out by the three steps of the New Skills in Europe Agenda[22]) and establish a credential mechanism (certification) that offers possibilities for official recognition (for both trainers and learners).

EOSC OPEN SCIENCE HELPDESK: A distributed and hierarchical structure with the country level in the centre addressing the technical-organizational and legal aspects of the Data and Access Interoperability layers, augmented with RI/domain specific helpdesk support.

[44] A grass-roots network coordinated by Open AIRE which maps out the training activities of various EOSC-related initiatives and research infrastructures https://www.openaire.eu/cop-training.

[45] Classification of training material and users based on skills: EOSCPilot,https://www.eoscpilot.eu/eoscpilot%E2%80%99s-contributions-eosc-skills; FOSTER, https://www.fosteropenscience.eu/; The Data Management Training (DMT) Clearinghouse, http://dmtclearinghouse.esipfed.org/

7 | OPENAIRE CURRENT OUTLOOK AND SUSTAINABILITY WITHIN EOSC

OpenAIRE has been supported by four subsequent lines of EC project funding since 2009, its current project phase, OpenAIRE Advance, is funded via the Research and Innovation Horizon 2020 programme. It has affiliated projects, OpenAIRE Connect, and close relationships with other EC projects, namely FOSTER and OpenMinTed. The OpenAIRE technical infrastructure (technical coordination) and the NOAD network (scientific coordination) have equal balance of funding. A significant part of the resourcing goes to fund the 'human' aspect of OpenAIRE, the NOAD network, which has been a successful instrument to develop, grow, support and align national Open Science activities with the EC's move to Open Science.

To facilitate a shift from the reliance on European funding, and in order to develop its independent business model for the growing suite of services, the OpenAIRE legal entity was founded in 2018. This shift will also tackle the issue of how to engage with Member States on committing to setting aside resources for Open Science coordination and activities that will be an integral component of their national research infrastructures. The NOAD activities are a natural part of these engagement activities in order to align with national Open Science endeavours and a committed resourcing model at national level would ensure a permanent continuation of OpenAIRE's activities.

7.1 OpenAIRE Legal Entity

The OpenAIRE Legal Entity (LE), or OpenAIRE AMKE,[46] was established in September 2018. It provides a robust, well-established legal footing for the development of OpenAIRE into the future. Its mission is to work with research and education communities, public and private organisations in order to:

- Support, implement and monitor Open Science in Europe
- Study, analyse and promote Open Science and e-science
- Connect open research with society, economy and the public sector
- Connect Europe to global open research

The Activities of OpenAIRE AMKE are realised by the following:

- It supports a *distributed infrastructure*, which is de facto federated across Europe, whether in the hands of Member States or RIs.
- It *facilitates* the network effects of existing e-science organisations and their individual researchers
- In terms of policy and governance, we focus on *collaboration* among the Member States.

OpenAIRE AMKE[47] is a membership body that is organised in a multilateral way, with 16 signatories as of August 2019 (and 5 more pending to sign). It is governed by a number of standing committees which are committed to oversee implementation of open science in their countries. Table 3 illustrates how the

[46] From Greek law, a Civil Non-Profit Company - Astiki Mi Kerdoskopiki Eteria
[47] https://www.openaire.eu/organization

internal governance of OpenAIRE AMKE complements the EOSC Working Groups[48], reflecting the fact that all players in the scholarly communication landscape are working together making the Research Commons as outlined above a reality[49].

Operations	
	Finance and Budget Standing Committee The committee provides guidance with regards to budgeting, financial planning and reporting, as well as monitoring of adherence to budgets.
	Outreach, Communications and Engagement The committee implements a communications, outreach and engagement strategy in order to address and engage relevant communities through OpenAIRE services and activities.
	Policies and Ethics Standing Committee The committee serves as a guide for ethical decision-making, and provides clear rules and regulations informing operational staff and board guidelines.
Operations reflecting EOSC Working Groups	
	Training and Support Standing Committee The committee oversees all aspects of OpenAIRE's training and support strategy, including planning, implementation, and strategic advice on collaboration.
	Services and Technology Standing Committee The committee provides the strategic framework to define, assess, expand, maintain and improve the OpenAIRE services and enhance interoperability with all relevant services.
	Open Science Foresight The committee analyses trends in emerging and existing policies, business models, services and tools, and works to establish OpenAIRE LE as a key European organization in open scholarly communication.

TABLE 3. OPENAIRE AMKE STANDING COMMITTEES

[48] https://www.eoscsecretariat.eu/eosc-working-groups

[49] The standing committees are being populated and managed by the diverse members of the LE many of whom are active at the national level on EOSC working groups

8 | CONCLUSIONS

OpenAIRE contributes to different components of EOSC's strategic implementation plan[50], while maintaining the underlying theme of **leveraging existing national scholarly communication infrastructures** to be contributing components of pan European providers.

EOSC will be the cornerstone of Europe's research landscape via the following:

- **A trusted infrastructure:** Creating an open, free, trusted and transparent infrastructure that allows data sharing across Europe. Provenance of data sources and reliability of services, and standards will be guaranteed.
- **A network of services:** EOSC is a System of Systems. Therefore, the community can rely on B2B services and available data via APIs.
- **Researcher-centric:** EOSC will reach out to researchers and use a language that resonates with their research needs. This includes for example fostering a FAIR research environment, with relevant training and seamless services for access and data sharing. For other research support individuals and funders, EOSC will provide assessment and monitoring mechanisms. For this to happen, an open, free and distributed body of research information needs to be maintained and updated, namely an interlinked research graph.
- **Democratic governance:** EOSC will be governed openly at national level. It will need experts, such as national nodes to take on this task to strengthen EOSC structures. A sub-department of EOSC (namely the OpenAIRE AMKE) can make it its task to coordinate these national entities for scholarly communication and Open Science.
- **Training tomorrow's data experts:** Training is the cornerstone of EOSC. Data stewards and experts need to be embedded in institutions and research groups to guide researchers towards Open Science across the research life-cycle. This process has to be professionalized and certified and needs to step up to tomorrow's research needs, and a body of training knowledge has to be stored and updated.

Critical elements of OpenAIRE's contribution to EOSC are the following:

- **Technical Architecture:** OpenAIRE's main business is to connect international infrastructures and networks via an interoperability layer and provision for thematic infrastructures for visibility and to enable publishing open research.
- **Open Science Helpdesk:** To support the availability of trustworthy and FAIR research data, training and Support for implementing FAIR data is a key part of OpenAIRE's activities via outreach at national level via the Helpdesk and Support arm of the service portfolio. This can be achieved by embedding data management practices across institutions and research communities.
- **Services for EOSC:** OpenAIRE provides a range of scholarly communication services which are both B2B and B2C. These include publishing services for all scholarly artefacts (Zenodo), discover services (Explore), monitoring and analysis for funders and institutions (Monitor), and support services for content providers to connect their content to EOSC (Connect).

[50] Jones and Abramatic, 'European Open Science Cloud (EOSC) Strategic Implementation Plan'.

- Governance: EOSC implementation at national level will work provided there is a trustworthy network of support at member state level. OpenAIRE NOADs have experience of aligning local policies with European level policies and can apply this experience to infrastructure and Open Science policy alignment. The application of the Open Science Policy framework such as that which is happening in OpenAIRE Advance will certainly assist in linking key national players in EOSC via action plans (26 countries in OpenAIRE are currently members of this Open Government Partnership which is a first step to creating national Open Science partnerships. [51] The critical instrument to foster Open Science policies is via the OpenAIRE AMKE and the work via the standing committees will pave the way to converge national EOSC structures.
- Communication: OpenAIRE will work concertedly to ensure that effective and clear outreach mechanisms are made to communicate its role in EOSC. Via the EOSC portal, the OpenAIRE EOSC services will be co-branded

[51] https://www.opengovpartnership.org/members/

9 | APPENDIX – OPENAIRE ACTIVITIES FOR EOSC IMPLEMENTATION

Architecture

WG ARCHITECTURE

OpenAIRE adds an **interoperability overlay in EOSC** by connecting all research outcomes (publications, data, software, methods, protocols) and linking them to people, organizations, funders, projects, research facilities - the OpenAIRE Research Graph - which provides the foundation to:

1. Connect institutional and national infrastructures in Europe and links to them to similar infrastructures around the world, via the international adoption of the OpenAIRE Guidelines for metadata exchange[52] for *Literature, Data, Software and other products repositories* and well-established APIs.
2. Provide a layer for thematic infrastructures, i.e., emerging research communities, Clusters and RIs, to align publishing of Open Science-oriented scientific products by applying best OS practices and the OpenAIRE Guidelines in off-the-shelf manner (**Research Community Dashboard**).
3. Work with repositories around the world (regional and international networks) to identify common services and functionalities to establish repositories as the basis of a global Open Data infrastructure, and provides web services where content providers can benefit from enhanced repository collections for more visibility and better institution research assessment and compliance with EOSC rules (**Content Provider Dashboard**).

KEY DELIVERABLES BY END OF 2020

- Q2/3 2019: EOSC-Hub data services to be compatible with OpenAIRE Guidelines
- Q4 2019: Enabling ResourceSync and schema.org as emerging API technologies (e.g., DuraSpace)
- Q1 2020: Roadmap for internationalization of guidelines

Data

WG FAIR

OpenAIRE fosters, promotes and embeds a FAIR and open research data management culture via **systematic training** and **services** (see next):

- Support materials for RDM (openness, FAIRness, stewardship, preservation, secure and trusted deposit)

[52] 1450 content providers are registered in OpenAIRE; Latin America and Canada have fully adopted the OpenAIRE Guidelines; Asian countries are implementing many elements, adding to support local requirements; work is in progress with African and Arab countries on the use of an aligned schema.

- Implementation within national research communities and universities as promoted by the 34 OpenAIRE network of OS experts (NOADs)
- Targeted webinars for different stakeholders and f2f national training events (162 training events, 6.6K attendants)
- Co-development of the EOSC training infrastructure: established a Community of Practice for training coordinators and managers from research and e-infrastructures

KEY DELIVERABLES BY END OF 2020

- Q1 2019 – Q1 2020: Enhancement of the Open Science collection of support materials (guides, factsheets, FAQs)
- Q3 2019 – Q4 2020: Continuous delivery of webinars and training events for RDM/OS
- Q3 2019: Training materials adapted for international consumption focused on FAIR practical implementation
- Q4 2019: Recommendations for the EOSC Portal Training Services (Community of Practice of Trainers)
- Q4 2019: Adopting good practice recommendations on metadata for training materials

Services

WG ARCHITECTURE & FAIR

OpenAIRE offers B2C and B2B services to find, access and re-use research data, as well as services to make data FAIR, to store them and ensure long-term preservation.

B2C	
Explore	A catalogue for contextualized access to interlinked scientific products (literature, research data, software, protocols, etc.) and research life-cycle assets (projects, funders, organizations, data sources).
Zenodo	A catch-all repository that sharing, preservation and showcasing of multidisciplinary research products (research data, software and publications).
Monitor	An evidence-based monitoring service. Monitors the uptake of OS and explores pervasiveness of approaches, identifies clusters, scientific trends. Targets funders, institutions, projects.
Amnesia	Anonymizing of sensitive data.
Research Community Dashboard	A dashboard for research communities to gather, link, search and deposit their research artefacts according to Open Science best practices and OpenAIRE/EOSC guidelines.
Content provider Dashboard	One-stop-shop for content provider (repositories, data archives, aggregators, CRIS systems) managers to interact with OpenAIRE/EOSC to register, validate, enrich and measure their collections and content.
OpenDMP	A Data Management Plan Service, configurable for domain, national, institutional settings and linked via APIs to EOSC and global registries and vocabularies.
OS OBSERVATORY	A core service for EOSC providing a set of indicators on Open Science for Europe/EOSC.
B2B	

APIs	Access of the OpenAIRE Open Research Graph (OAI-PMH, REST, SPARQL)
ScholExplorer	Populates and provides access to a graph of datasets-literature and datasets-datasets relationships (implementation of RDA/Scholix: 620K literature objects, 2.6 mi datasets, 18 mi bi-directional links).
Usage Stats	Collects and analyses usage data from repositories via an interoperable format to exploit usage metrics (106 providers, 45 mi usage events)
Validator	Verifies compliance of repositories to metadata schemata via OAI-PMH. OpenAIRE instance validates for OpenAIRE Guidelines (902 literature repos, 425 OA Journals, 100 data repositories validated).
Broker	Enables metadata exchange among repositories, publishers or aggregators and allows the enriching of local metadata (170 providers, 20 mi events).

KEY DELIVERABLES BY END OF 2020

- Q2 2019: Roll out Catch-all Broker service (including research data records)
- Q3 2019: Release 1.0 of Research Community Dashboard
- Q3 2019: Release 1.0 of OpenDMP
- Q4 2019: Release 2.0 of Funder Monitoring & Project Dashboards; Beta release of Institutional Dashboard
- Q4 2019: Beta release of Open Science Observatory
- 2019: Update of D4.3 Use cases of OpenAIRE OS services for EOSC stakeholders
- Q1 2020: Calculation of unit costs for all services and expose to the EOSC Portal

Access and Interface

WG ARCHITECTURE

OpenAIRE builds a global **Open Research Graph** which provides cross-over mechanisms for seamless access to data and scientific artefacts. This is envisioned to be one of the entry points for the EOSC data catalogue(s). It is currently accessible via in beta hosting a graph linking 94 mi publication records (access to 11 mi full texts), 8 mi records of data and scientific related products, from 1.5K content providers linked to 28 funders.

KEY DELIVERABLES BY END OF 2020

- Q3 2019: Roll out in production the new enriched OpenAIRE Open Research Graph and embed in the EOSC Portal.

Rules

WG ROP

OpenAIRE considers two aspects for rules for participation in EOSC: **a. technical compliance** as this is set out by the OpenAIRE Guidelines for content providers, and **b. the implementation of Open Science** practices.

Working on the policy front, OpenAIRE harmonizes Open Science policies across all Europe and support their implementation at the local level via hands on support, training and key materials.

We are there with expert advice to both researcher communities and their national policy makers on how to be OA mandate, Plan S, EOSC and FP9 ready, being the link between top-down and bottom up approaches.

KEY DELIVERABLES BY END OF 2020

- Q1 2019: Toolkit for policy makers on Open Access and Open Science which includes OS policy templates and checklists for research funding (RFOs) and research performing organizations (RPOs)
- Q1 2019: D3.2 Toolkit for researchers on legal issues, which includes Legal support materials to gain traction within research communities: GDPR, Copyright, IPR

Governance

WG SUSTAINABILITY & LANDSCAPE

Paving the way for EOSC at national level: OpenAIRE 34 National Open Access Desks contribute to and lead on local community activities, including policies and governance. For the past three years EOSC has been on their agenda, establishing dialogue about how to form national coordination points and prepare action plans for EOSC, building synergies with other national players (infrastructure and service providers, Open Science policy makers) in order to advise national partners on how best to obtain value-added*ness* for their local infrastructure within the EOSC vision and implementation.

NOADs getting 'EOSC' ready: 8 are on national OS WGs, 15 have contributed to OA/ OS policies at institutional level (e.g., Turkey, Switzerland, Romania, Serbia, Greece, Spain).

The OpenAIRE newly formed Legal Entity (16 national organizations from 14 countries and 2 international organizations) is set out to coordinate the open scholarly/scientific communication in Europe (services, policies, training).

KEY DELIVERABLES BY END OF 2020

- Q2 2019: White Paper: OpenAIRE Legal Entity operations
- Q2 2019: White Paper: OpenAIRE vision towards EOSC
- Q3 2019: Guidance for National Open Government Partnership Activities (Open Science part of the wider agenda)
- Q3 2019: OpenAIRE Sustainability Report on status and demand of infrastructure services involving national/European research and academic partners (via NOADs/National Open Science Partnerships, collaboration with Clusters, ESFRI RIs and SMEs), with the primary focus being on drivers for obtaining optimal national commitment and value for research activity and impact to and through EOSC.

10 | REFERENCES

[1] Bateson, Nora. 'Warm Data: Contextual Research and New Forms of Information'. Transform (blog), 8 January 2019. https://thetransformseries.net/2019/01/08/warm-data-contextual-research-and-new-forms-of-information/.

[2] Bilder, Geoffrey, Jennifer Lin, and Cameron Neylon. 'Principles for Open Scholarly Infrastructures-V1', 23 February 2015. https://doi.org/10.6084/m9.figshare.1314859.v1.

[3] Bohlen, Vincent, Lisa Bruns, Nadja Menz, Fabian Kirstein, and Sonja Schimmler. 'Open Data Spaces - Towards the IDS Open Data Ecosystem'. Fraunhofer Institute for Open Communication Systems, December 2018. https://www.internationaldataspaces.org/wp-content/uploads/2019/07/Open-Data-Spaces-IDSA.pdf.

[4] Bosman, Jeroen, Ian Bruno, Chris Chapman, Bastian Greshake Tzovaras, Nate Jacobs, Bianca Kramer, Maryann Martone, et al. 'The Scholarly Commons - Principles and Practices to Guide Research Communication'. Preprint. Open Science Framework, 15 September 2017. https://doi.org/10.31219/osf.io/6c2xt.

[5] Bosman, Jeroen, and Bianca Kramer. 'The European Open Science Cloud as a Commons?', 25 October 2017. https://doi.org/10.6084/m9.figshare.5537899.v1.

[6] 'Communication "A New Skills Agenda for Europe: Working Together to Strengthen Human Capital, Employability and Competitiveness"'. European Commission. Accessed 22 August 2019. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016DC0381.

[7] 'Communication "Towards a Common European Data Space"'. European Commission, 25 April 2018. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52018DC0232.

[8] 'D2.3: EOSC-Hub Briefing Paper - EOSC Federating Core Governance and Sustainability Public | EOSC Hub'. EOSC-hub Consortium. Accessed 13 August 2019. https://www.eosc-hub.eu/deliverable/d23-eosc-hub-briefing-paper-eosc-federating-core-governance-and-sustainability-public.

[9] 'Dataspaces'. Wikipedia, 18 June 2019. https://en.wikipedia.org/w/index.php?title=Dataspaces&oldid=902374549.

[10] Demchenko, Yuri, Paola Grosso, Cees de Laat, and Peter Membrey. 'Addressing Big Data Issues in Scientific Data Infrastructure'. In 2013 International Conference on Collaboration Technologies and Systems (CTS), 48–55. San Diego, CA, USA: IEEE, 2013. https://doi.org/10.1109/CTS.2013.6567203.

[11] Dietz, T., Elinor Ostrom, and Paul C. Stern. 'The Struggle to Govern the Commons'. Science 302, no. 5652 (12 December 2003): 1907–12. https://doi.org/10.1126/science.1091015.

[12] Directorate-General for Research and Innovation (European Commission). 'Future of Scholarly Publishing and Scholarly Communication: Report of the Expert Group to the European Commission.' PDF, 30 January 2019. https://doi.org/10.2777/836532.

[13] 'EUGDPR – Information Portal'. Accessed 13 August 2019. https://eugdpr.org/.

[14] Ferguson, Adam R, Jessica L Nielson, Melissa H Cragin, Anita E Bandrowski, and Maryann E Martone. 'Big Data from Small Data: Data-Sharing in the "long Tail" of Neuroscience'. Nature Neuroscience 17, no. 11 (November 2014): 1442–47. https://doi.org/10.1038/nn.3838.

[15] Ferrari, Tiziana, Diego Scardaci, and Sergio Andreozzi. 'The Open Science Commons for the European Research Area'. In Earth Observation Open Science and Innovation, edited by Pierre-Philippe Mathieu and Christoph Aubrecht, 43–67. ISSI Scientific Report Series. Cham: Springer International Publishing, 2018. https://doi.org/10.1007/978-3-319-65633-5_3.

[16] Fortin, Jean-Michel, and David J. Currie. 'Big Science vs. Little Science: How Scientific Impact Scales with Funding'. Edited by Vincent Larivière. PLoS ONE 8, no. 6 (19 June 2013): e65263. https://doi.org/10.1371/journal.pone.0065263.

[17] Frischmann, Brett M., Michael J. Madison, and Katherine J. Strandburg, eds. *Governing Knowledge Commons*. Oxford: Oxford Univ. Press, 2014.

[18] Fuster Morell, Mayo. 'Governance of Online Creation Communities: Provision of Infrastructure for the Building of Digital Commons', 2010. http://cadmus.eui.eu//handle/1814/14709.

[19] Ganck, Alain de. 'The New European Interoperability Framework'. Text. ISA² - European Commission, 16 February 2017. https://ec.europa.eu/isa2/eif_en.

[20] Grossman, Robert. 'What Is Data Commons and How Can Your Organization Build One?' Data & Analytics presented at the Molecular Medicine Tri Conference, 12 February 2018. https://www.slideshare.net/rgrossman/how-data-commons-are-changing-the-way-that-large-biomedical-datasets-are-analyzed-and-shared.

[21] 'Guidance Notes on Funding Criteria National Research Data Infrastructure (NFDI)'. Deutsche Forschungsgemeinschaft, June 2019. http://www.dfg.de/formulare/nfdi120/nfdi120_en.pdf.

[22] Helfrich, Silke. 'The Logic of the Commons & the Market: A Shorthand Comparison of Their Core Beliefs | The Wealth of the Commons', n.d. http://wealthofthecommons.org/essay/logic-commons-market-shorthand-comparsion-their-core-beliefs.

[23] Holmgren, Sverker, Arjen van Rijn, Ulrike Jaekel, Josva Kleist, Gabriele von Voigt, Fotis Karayannis, Michael Maragakis, and Jan Wiebelitz, eds. 'National Nodes - Getting Organised; How Far Are We? Implementing e-Infrastructure Commons and the European Open Science Cloud', 14 June 2019. http://e-irg.eu/catalogue/eirg-1006.

[24] Howe, Nick. 'Listen: Small Research Teams Are Better than Big Ones at Disrupting Science'. Nature, 13 February 2019. https://doi.org/10.1038/d41586-019-00581-4.

[25] Johnson, Rob. 'From Coalition to Commons: Plan S and the Future of Scholarly Communication'. Insights the UKSG Journal 32 (30 January 2019): 5. https://doi.org/10.1629/uksg.453.

[26] Jones, Sarah, and Jean-François Abramatic. 'European Open Science Cloud (EOSC) Strategic Implementation Plan'. Directorate-General for Research and Innovation (European Commission), 24 July 2019. https://publications.europa.eu/s/mGCw.

[27] Jürjens, Jan. 'The Industrial Data Space: Digital Industrial Platform Value Chains in All Sectors of the Economy'. presented at the Workshop on Digitising European Industry Working Group 2 on Digital Industrial Platform, 4 May 2017. https://ec.europa.eu/futurium/en/system/files/ged/industrial_data_space.pdf.

[28] Kluitenberg, Eric. 'Constructing the Digital Commons: A Venture into Hybridisation'. Media Art Net, 2003. http://www.mediaartnet.org/source-text/163/.

[29] 'Le Plan national pour la science ouverte : les résultats de la recherche scientifique ouverts à tous, sans entrave, sans délai, sans paiement'. Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. Accessed 13 August 2019. http://www.enseignementsup-recherche.gouv.fr/cid132529/le-plan-national-pour-la-science-ouverte-les-resultats-de-la-recherche-scientifique-ouverts-a-tous-sans-entrave-sans-delai-sans-paiement.html.

[30] Little, R. G. 'Tending the Infrastructure Commons: Ensuring the Sustainability of Our Vital Public Systems'. Structure and Infrastructure Engineering 1, no. 4 (December 2005): 263–70. https://doi.org/10.1080/15732470500103708.

[31] Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. *National Plan for Open Science / Le Plan national pour la science ouverte* (2018). http://cache.media.enseignementsup-recherche.gouv.fr/file/Recherche/50/1/SO_A4_2018_EN_01_leger_982501.pdf.

[32] National Academies of Sciences, Engineering, Policy and Global Affairs. *Toward an Open Science Enterprise. Broadening Access to the Results of Scientific Research*. National Academies Press (US), 2018. https://www.ncbi.nlm.nih.gov/books/NBK525412/.

[33] National Academies of Sciences, Engineering, and Medicine. *Open Science by Design: Realizing a Vision for 21st Century Research*. Washington, D.C.: National Academies Press, 2018. https://doi.org/10.17226/25116.

[34] 'New Skills Agenda for Europe - Employment, Social Affairs & Inclusion - European Commission'. Accessed 22 August 2019. https://ec.europa.eu/social/main.jsp?catId=1223.

[35] Neylon, Cameron. 'Sustaining Scholarly Infrastructures through Collective Action: The Lessons That Olson Can Teach Us'. KULA: Knowledge Creation, Dissemination, and Preservation Studies 1, no. 1 (27 December 2017): 3. https://doi.org/10.5334/kula.7.

[36] Ostrom, Elinor, Christina Chang, Mark Pennington, and Vlad Tarko. 'The Future of the Commons - Beyond Market Failure and Government Regulation'. SSRN Electronic Journal, 2012. https://doi.org/10.2139/ssrn.2267381.

[37] Ostrom, Elinor, Thomas Dietz, Nives Dolšak, Paul C. Stern, Susan Stonich, and Elke U. Weber, eds. *The Drama of the Commons*. Washington, DC: The National Academies Press, 2002. https://doi.org/10.17226/10287.

[38] Publications Office of the European Union. 'Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data.' Website, 26 November 2018. https://publications.europa.eu/s/mGCS.

[39] Rimini, Francesca da. 'Social Technologies and the Digital Commons'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 23 September 2006. https://doi.org/10.4018/9781591409991.ch005.

[40] Romier, Geneviève, and Eric Fede. 'D6.8: Final EOSC Architecture | Eoscpilot.Eu'. EOSCPilot.eu, 20 May 2019. https://eoscpilot.eu/content/d68-final-eosc-architecture.

[41] Soman, Dilip. *The Last Mile: Creating Social and Economic Value from Behavioral Insights*. Toronto ; Buffalo ; London: University of Toronto Press, 2015.

[42] Tennant, Jonathan, Jennifer Elizabeth Beamer, Jeroen Bosman, Björn Brembs, Neo Christopher Chung, Gail Clement, Tom Crick, et al. 'Foundations for Open Scholarship Strategy Development'. Preprint. BITSS, 30 January 2019. https://doi.org/10.31222/osf.io/b4v8p.

[43] The Directorate-General for Communications Networks, Content and Technology. 'Digital Skills at the Core of the New Skills Agenda for Europe'. Text. Digital Single Market - European Commission, 10 June 2016. https://ec.europa.eu/digital-single-market/en/news/digital-skills-core-new-skills-agenda-europe .

[44] Vicente-Saez, Ruben, and Clara Martinez-Fuentes. 'Open Science Now: A Systematic Literature Review for an Integrated Definition'. Journal of Business Research 88 (July 2018): 428–36. https://doi.org/10.1016/j.jbusres.2017.12.043.