SCIENCE|BUSINESS®

Why Open Science is the Future

(And how to make it happen)

A report of the Science|Business Network's Cloud Consultation Group

July 2019

In the first half of 2018, members of the Science|Business Network's Cloud Consultation Group met twice to discuss potential business models for the European Union's proposed "science cloud" project. This report is based partly on those discussions, but is ultimately a product of Science|Business. The views expressed herein do not necessarily reflect those of individual members.

This report is the fourth in a series that the group is working on, to gather private and public sector expertise on topics of importance to the development of the European Open Science Cloud. The other reports are available on www.sciencebusiness.net:

The case for the cloud, May 2017
Governing the Open Science Cloud, October 2017
The European Open Science Cloud: Who pays for what?, February 2018

Cloud Consultation Group members:

Amazon
Association for Computing Machinery – Europe Policy Committee
Association of Commonwealth Universities
Barcelona Supercomputing Centre
CERN
ETH-Zurich
European University Association (EUA)
European Space Agency
GÉANT
Google
Huawei
Microsoft
Research Data Alliance Europe
University of Eastern Finland
University of Twente

Rapporteur: David Pringle

Avenue des Nerviens 79, Brussels, Belgium 1040
info@sciencebusiness.net
+322-304-7577

Executive Summary

Around the world, researchers are increasingly aware of the value and importance of open science. As scientific research becomes highly data-driven and dependent on computing, scientists are conscious of the growing need to share data, software and infrastructure to reduce wasteful duplication and increase economies of scale. In an ideal world, every step of the research process would be public and transparent – the full methodology and all the tools used, as well as the data, would be accessible to the public and all groups without restriction, enabling reproducibility and refinement by other scientists.

Given the growing body of evidence that collaboration between scientists across the world can achieve breakthroughs far quicker than individual research teams, there should be strong momentum behind fully open science. But, in reality, progress is patchy. Worryingly, one online survey[1] in the autumn of 2018 found that only 11% of researchers shared data from their last project with people they don't know personally, down from 14% in 2016.

With some notable exceptions, the sharing of research data and tools occurs on a piecemeal basis, facilitate by trusted bilateral relationships, rather than through fully open platforms. Moreover, openness is not yet fully designed and embedded into the scientific process. Too often, research tools and data are opened up as an afterthought through a time-consuming retrofit. Another major obstacle is the need to support the long-term preservation of research data and the software code, tools and operating environments required to make sense of the data. Depending on the discipline, important datasets can still yield scientific breakthroughs decades after they were first generated.

The European Open Science Cloud (EOSC) is seeking to make it much easier for researchers to pursue fully open science by federating the myriad of research infrastructures, tools and datasets being employed by the 1.7 million researchers across the EU. If it can pool Europe's research resources far better than today, the EOSC could make the long-term storage of scientific data and tools both practical and economic. Moreover, research funders are increasingly making open data mandatory, while requiring data management is built into project budgets.

But solving the financial, technical and interoperability issues won't be enough: The EOSC also needs to reduce the uncertainty around the regulatory and legal frameworks relating to open science. Today, researchers face many sovereignty and legal issues, encompassing questions about copyright and ownership and whether they will receive appropriate recognition for their work. To fully realise the benefits of open science, the EOSC will need to bring much greater coherence to the incentive, legal and regulatory frameworks governing research data and tools. Today, researchers need to navigate a variety of EU directives, regulations and national laws, as well as multinational initiatives, such as the Research Data Alliance and Plan S. By playing a coordinating role, the EOSC could help researchers cut through this complexity, so they can systematically pursue open science, safe in the knowledge that they are furthering their careers and enhancing their reputations.

1 Source: "Open data : The researcher perspective" by Elsevier and the Centre for Science and Technology Studies.

About this report

This report explores the case for open science, illustrated by use cases and case studies, spanning academia, start-ups and enterprises. It seeks to explain why public sector and private sector researchers should support open science by highlighting concrete examples of disciplinary breakthroughs that have been fuelled by the sharing of research data, tools and infrastructure.

The report draws on discussions and input from an independent consultation group representing research, industry and policy, which is coordinated by Science|Business. It begins by considering how to define open science, before then exploring the progress being made to date and the various obstacles standing in the way of more openness. The next section considers the work of the European Commission and the role of the European Open Science Cloud in making it easier for scientists to share their data, their tools and their methodologies. Finally, the paper outlines some examples of how open science has driven significant steps forward in various fields of research from astrophysics to the development of drugs.

Defining Open Science

How open does scientific research need to be? Some stakeholders focus on the importance of "open data" - the sharing of the full results of each experiment with the broader scientific community. This concept is encapsulated in the widely supported FAIR principles (Findability, Accessibility, Interoperability, and Reusability) of data stewardship.

"Successful implementation of open science will be characterized by researchers freely sharing data, publishing by default in open access journals and avoiding the use of restrictive IPRs." – Open Letter on Defining Success in Open Science But there is a compelling case for fully open science in which every step of the research process is public and transparent – the full methodology and all the tools used, as well as the data, would be accessible to the public and all groups without restriction. Some experts argue that open science also needs to encompass open infrastructure to ensure that all researchers can access the computing resources they need to further their work.

"Open science is about extending the principles of openness to the whole research cycle fostering sharing and collaboration as early as possible thus entailing a systemic change to the way science and research is done" - Foster Open Science's Introduction to Open Science[1]

As science becomes increasingly data-driven, researchers are developing highly specialist algorithms and software to enable them to process and analyse the data generated by their research and experiments. In a fully open scientific ecosystem, these computational tools need to be readily available along with the underlying data to enable other researchers to reproduce and extend the original research.

Although scientists' growing reliance on data has thrown a spotlight on open science, the underlying concept has been around for decades: some research organisations have made openness their modus operandi by insisting that researchers publish the full methodology of their project, as well as the results.

For example, the Structural Genomic Consortium (SGC), a worldwide partnership between universities and pharmaceutical companies, founded in 2004 requires scientists to share data and support reproducibility (e.g., by using electronic lab notebooks) in return for predictable funding. [1]

Moreover, the formal agreement to encourage the free distribution of research data, technology and resources that underpins the Human Genome Project has had a major positive impact on research across the life sciences. Similarly, the Allen Institute founded in 2003, has become a valuable resource for brain scientists worldwide by freely sharing gene-expression maps for human and mouse brains.[2]

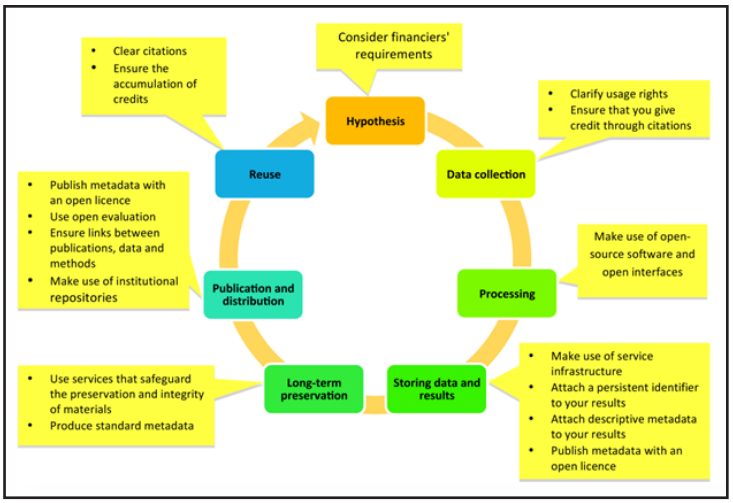

Figure 1: Promoting openness at different stages of the research process
Source: Open Science and Research Initiative, 2014

Experts note that in many scientific disciplines (DNA sequencing being a good example), the development of standardised software tools and algorithms is critical to progress – researchers need to be able to easily access a reference implementation without worrying about intellectual property. Collaborative software development hosting services, such as GitHub (owned by Microsoft), are increasingly being employed by researchers for this purpose.

1 https://mniopenresearch.org/articles/2-2

1 https://www.frontiersin.org/articles/10.3389/fnins.2018.00047/full#B16

2 https://www.fosteropenscience.eu/content/what-open-science-introduction

Why do we need open science?

From a big picture perspective, the arguments in favour of fully open science are persuasive. For society as a whole, open science offers numerous benefits: In general, openness and transparency allow new ideas to spread faster, avoid wasteful duplication, increase trust, and make it easier for experts to find each other and co-operate.

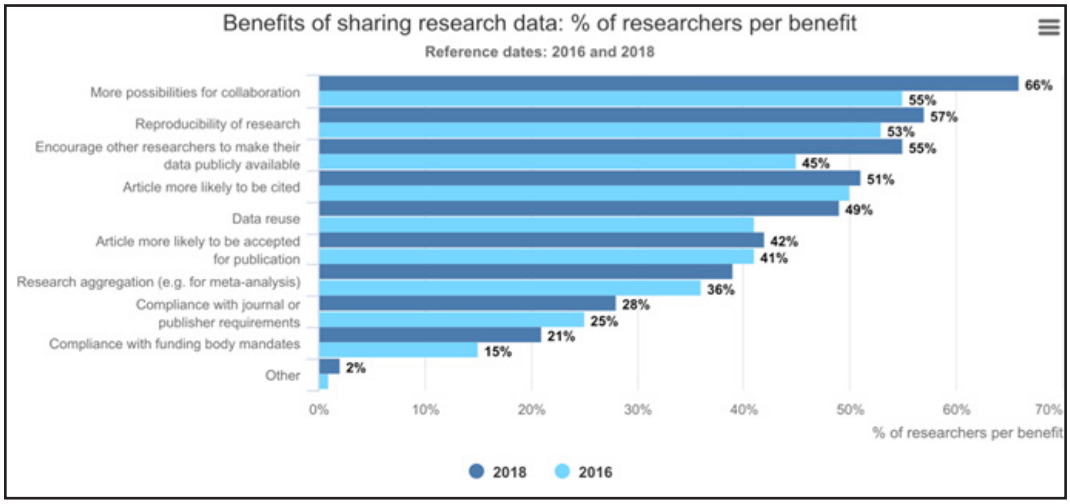

Figure 2: Researchers see multiple benefits to sharing their data
Sources: A survey of 1,200 researchers for the 2017 report "Open data : The researcher perspective" by Elsevier and the Centre for Science and Technology Studies (CWTS), part of Leiden University, and a follow-up online survey carried out between October/November 2018.

For researchers, open science is also becoming compelling. Sourcing and analysing the vast datasets that can be required to conduct cutting-edge research is an expensive and time-consuming exercise. It is far more efficient to reuse other scientists' data and software tools than to build them from scratch for each research project. In other words, open science is fast becoming an economic and practical necessity for the vast majority of researchers working on tight budgets.

Indeed, their need for data is driving a cultural and attitudinal shift among researchers. They are increasingly looking to partner with their peers in other institutions, while scientists in academia look to work with businesses that have the data sets they need for their research. As the importance of data and software tools rises, scientists that openly share their work should increasingly be rewarded through recognition and more opportunities for collaboration.

Indeed, individual researchers say they share their data primarily to open up more possibilities for collaboration, to help ensure reproducibility of research and to encourage other researchers to share their data. Global surveys suggest that support for open science is growing as researchers become increasingly aware of these benefits (see Figure 2).

In fact, stakeholders from across the scientific ecosystem are now calling for open science. For

- Generate more, and more diverse, high-quality datasets, together with the meta-data necessary to use them (including descriptions of methods, reagents, protocols and workflows, the instruments or platforms used in their generation, how and why data were collected).
- Curb waste within R&D, lowering otherwise rising costs and providing a better return on investment. Increased reuse of data, as fewer datasets are discarded or rendered inaccessible.
- Improve reproducibility, thereby increasing trust in scientific research and reducing the need for lengthy validations. A survey of researchers by Nature found that half of the respondents can not reproduce their own experimental results (see Figure 3).
- Build open detailed knowledge of scientific fundamentals, such as the basic biology and biochemistry of drug targets and pathways. In the pharmaceutical industry, for example, this would speed up the identification of the most promising drug targets.

• Give rise to a greater diversity of research, penetrating research 'white space' and seeding novel research domains, including new interdisciplinary fields. For example, the "Pathogen Box" – an open-access collection comprised of 400 compounds with demonstrated biological activity against specific pathogenic organisms that cause tropical and neglected diseases – has helped progress the discovery of drugs to combat neglected diseases. The Pathogen Box has been credited with the identification of Candida albicans biofilm inhibitor .

• Foster democratization of the research enterprise, resulting in a greater diversity of people meaningfully involved and gleaning value from the research process.

• Pave the way for increased scientific capacity in lower income, marginalized and developing communities.

• Lead innovation actors to concentrate their efforts where they can excel, reducing the redundancy of roles and activities.

• Decrease barriers to students moving between academia and industry, by increasing collaboration and knowledge flow between the two settings.

• Prompt the private sector and venture capital to invest in research, where otherwise they would not.

Note, the cloud services shown in the pink boxes in Figure 5 will involve a mixture of IaaS, PaaS and SaaS.

EOSC services

In its March 2018 staff working paper on the EOSC, the European Commission suggested that the EOSC will provide the following services:

- Identification and authentication to enable access to EOSC resources (free of charge).
- A protected and personalised work environment/space (e.g. logbook, settings, compliance record and pending issues) (free of charge).
- Relevant service information (status of the EOSC, list of federated data infrastructures, policy-related information, description of the compliance framework) and relevant guidelines. (free of charge).
- Catalogues of datasets (free of charge).
- Data mining, analytics, fusion and processing services, which enable users to find, access, re-use and analyse research data generated by others (may incur a fee).
- Services to make data FAIR, to store it and ensure long-term preservation (costs to be covered by EOSC business model).

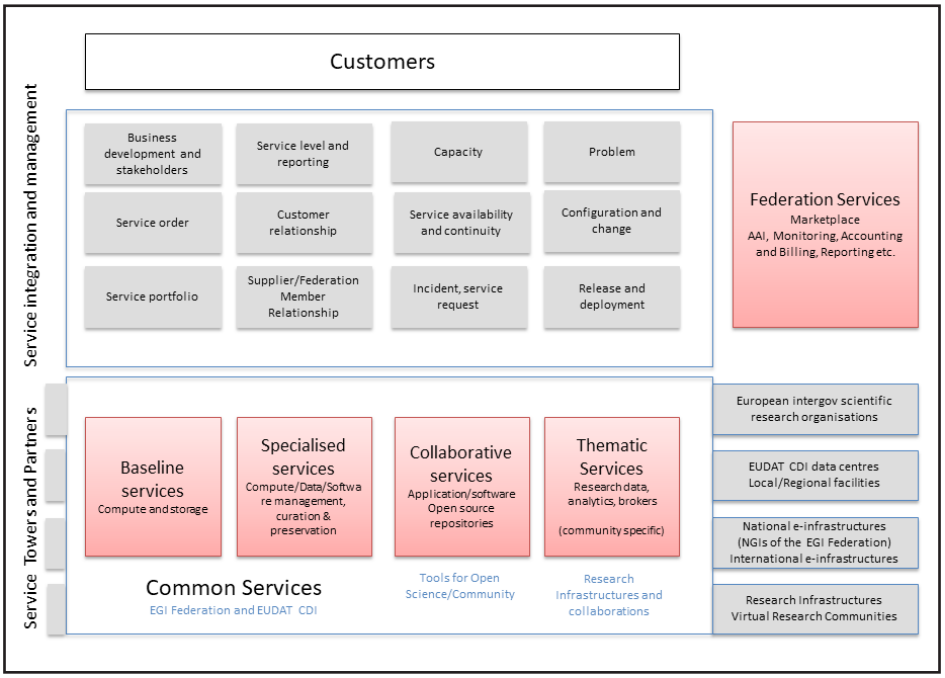

Figure 5: How the EOSC will bring together various services and tools for researchers
Source: The EOSCpilot

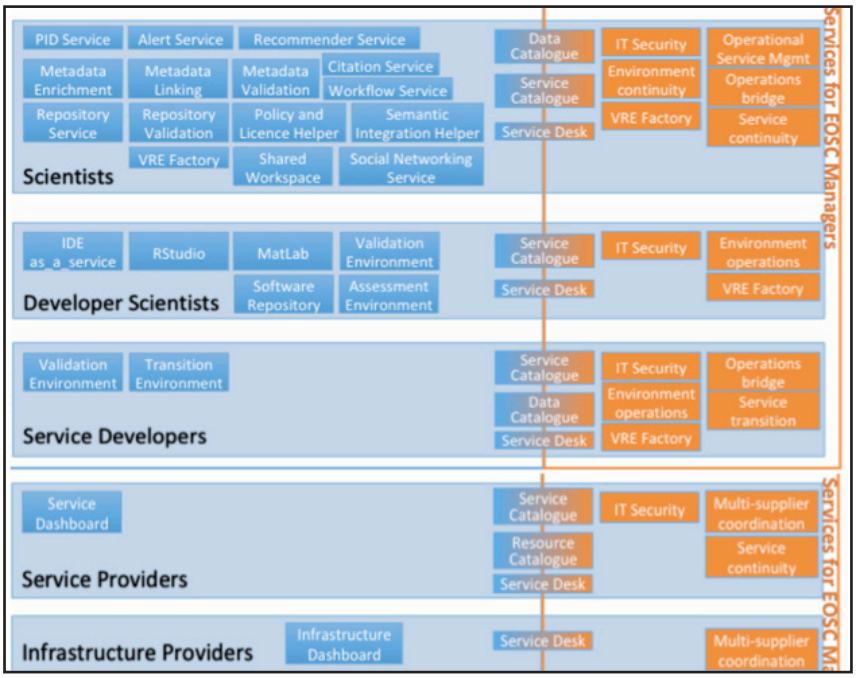

Figure 6: Overview of the services that the EOSC could provide to each group of users
Source: The EOSCpilot

The EOSCpilot has further segmented these services, as shown in Figure 6. It defines the roles of the various actors as:

- Scientists: use functions and tools to perform their domain-specific research activities and to collaborate.
- Developer scientists: develop new analytical models, new processes and tools to analyse data and derive knowledge
- Service developers: develop services for scientists by understanding their requirements and the Open Science vision.
- Service providers: operate and provide access to their own portfolio of services according to declared service level agreements .[1]
- Infrastructure service providers: provide access to computational, storage and network resources according to declared service level agreements .[2]
- EOSC managers: facilitate the operation, assistance and quality assurance of the EOSC system and coordination among different stakeholders.

Funding these services

In phase 1 (up until 2020), the Commission has said it will invest €300 million to support the core functions of the EOSC, while it has asked Member States to flag the national initiatives that they want to federate into the EOSC and the resources they are willing to provide in kind.

Moreover, funding agencies will start making research grants conditional on the recipients making the data generated by their research projects open and accessible. For example, EU grants awarded through the forthcoming Horizon Europe framework programme will require researchers to store their data in cloud-compatible repositories.

In phase 2, after 2020, the Commission envisages the activities of the EOSC could be financed by a mix of funding, including possibly deposit fees from national funders, as well as revenues generated from users. During Phase 1, the EOSC Governance Framework will be tasked with producing a full cost estimate for the running of the EOSC, which will be the basis of a financial plan for Phase 2, addressing both scalability and legacy.

1 The nature of such agreements and how they are enforced will depend on whether the provider is in the public or private sectors.
2 The nature of such agreements and how they are enforced will depend on whether the provider is in the public or private sectors.

The initial service charging model

To help the EOSC achieve critical mass, it would be prudent not to charge registered researchers in higher education to access EOSC-certified data and related research tools provided by their peers in higher education.

Under this "freemium" model, usage of the data and tools would be monitored to prevent abuse (only limited extracts could be downloaded), while enabling the generators of the data to receive recognition. This monitoring would enable the EOSC executive body to track usage patterns and refine the overall proposition. This freemium model could be applied indefinitely or it could only apply up until the end of 2020, after which researchers working in higher education could pay a low subscription fee to access EOSC-certified data or research tools.

In any case, the EOSC will have to compensate the providers of the authentication, search and discovery tools and catalogues researchers will employ as they use the EOSC. Moreover, researchers themselves will have to pay for their usage of commercial IaaS, PaaS and SaaS services provided by EOSC-accredited service providers[1] at commercial rates.

Note, these service providers would be competing with each other within the EOSC – the research team and/or its IT department would select the service provider that offers the most attractive proposition.

In time, they may also benefit from a volume discount negotiated between the EOSC executive body and the service provider once the EOSC is widely used. These services should, of course, be available from multiple providers to ensure that users of the EOSC have both choice and access to competitive pricing. (Note, some providers of PaaS and SaaS may source IaaS from another entity).

1 The EOSC High Level Expert Group has suggested a funding and payment model that is based on a certification programme for commercial and non-commercial providers of computing services that can provide scientifically useful services, that will meet EOSC-defined standards to ensure minimum levels of access and interoperability and that will accept specific, EOSC-defined financial transactions in payment for these services.

In the case of publicly funded cloud service providers, the transactions will need to be compliant with procurement regulations, which can restrict their ability to participate in commercial tenders, offer service level agreements and penalties, and receive more than 20% of their total revenue from outside their main funding source.

To illustrate how this charging and service model would work in practice, here is a high-level description of how researchers in higher education could engage with the EOSC in the start-up phase between 2018 and 2020. Note, the subsidies inherent in this process could be tapered off once the EOSC has achieved critical mass.

- **Step 1:** A university research team would register for access to EOSC resources and would then be authenticated online by an EOSC-accredited service provider. Registration would be free of charge. The research team will then have free access to a protected and personalised work environment/space and relevant service information provided by an EOSC-accredited service provider.
- **Step 2:** Before undertaking a new project, the research team would consult the EOSC catalogue to help identify any relevant data and research tools that already exist. Provided by an EOSC-accredited service provider, this facility is available for free.
- **Step 3:** During their project, the team would spend a portion of their research grant on IaaS and PaaS or SaaS services to access and analyse data generated by relevant projects they identified via the EOSC catalogue. In most cases, the research team would be prohibited from downloading or replicating the data, thereby enabling the EOSC to monitor usage.
- **Step 4:** During their project, the team may also spend a portion of their research grant on the IaaS and PaaS services required to access research tools and models developed by scientists elsewhere and made available through the EOSC.
- **Step 5:** During their project, the team would also spend a portion of their research grant on IaaS and PaaS or SaaS services made available via the EOSC to manage and analyse data generated by their own research.

- Note, these services could include collaborative tools that enable researchers to work remotely with other teams of scientists.
- **Step 6:** During their project, the team may also spend a portion of their research grant to use IaaS and PaaS services made available via the EOSC to develop new research tools and models.
- **Step 7:** At the end of their project, the team spend a portion of their research grant to employ IaaS and PaaS or SaaS services to organise the resulting data in line with FAIR principles, and store it where it can be accessed via the EOSC for a duration appropriate to the discipline in which the research is conducted. Note, where data needs to be stored long-term, the cost could be significant. They could also use IaaS and PaaS services to add any research tools they developed for the project to the EOSC, where they can be accessed by other researchers.

The role of EOSC credits/cloud coins

Once their research has been peer-reviewed and the resulting data and research tools certified as EOSC-compliant, the research team and its higher education institution should be given public recognition. For example, their contribution could be published in an "open science index", which ranks researchers by their contributions. This index could be supplemented with credits (or cloud coins)[1] that can be spent on EOSC-accredited services.

These credits would be shared in a transparent and proportionate manner between research teams and institutions.

Some kind of assessment process would be required to determine the value of the credits awarded. This process could be based on usage of the dataset.

The EOSC High Level Expert Group says: "Ideally, the marketplace would keep track of how frequently a data set is used and the provider of that data set compensated accordingly, similarly to how YouTube pays people who upload videos based on how many times they are viewed," it adds.

However, such a performance-based approach would require the EOSC executive body to contract service providers to monitor usage of individual data sets. While monitoring usage adds complexity and additional cost to the EOSC, publicly-funded researchers can justify their research budgets by showing to their funding agencies the impact and re-use of their results. Such monitoring will provide an incentive for data providers and service providers to participate in the EOSC. However, this model implies that researchers will not be able to download the data, which then creates a form of exclusivity for EOSC, which may encourage participation, but has to be managed carefully to ensure that the open science cloud is a genuinely open and competitive marketplace, rather than a de facto cartel.

Note, these credits may not be redeemable until after 2020, once the EOSC executive body has a legal status that enables it to manage funds and is beginning to generate revenues. As well as funding these credits, the EOSC would also need to finance the management of the marketplace itself. How much this relies on fixed subsidies, how much from grant conditions and how much on actual value will need to be carefully monitored.

From 2020 (once the EOSC Executive Body has a legal status that enables it to manage funds), private companies and commercial spinouts from universities and SMEs could pay the EOSC an annual subscription to access the services that are available for free to public-sector researchers.

1 The EOSC High Level Expert Group envisages the EOSC and member states would certify one or more brokers to manage the acquisition, distribution and payment for EOSC vouchers. These brokers could be government agencies in member states, entities within member states, transnational governments or private firms.

Other potential source of revenues (as identified by the EOSC High Level Expert Group) could be an annual subscription fee paid by conformant commercial service providers to offer their services in the marketplace, a transaction tax on EOSC credit vouchers (i.e. a service provider is taxed a small percentage on each voucher that is redeemed against their services), perhaps combined with a transaction tax on private sector users. A transaction tax has the advantage of removing upfront charges for participation by a service provider.

The High Level Expert Group says these revenues could contribute to funding the operational costs of the EOSC gateway/marketplace and the provision of key open data sets.

Although researchers working in the public sector could, at their own discretion, use some of their research grant to access EOSC services and make their data and tools available via the EOSC, the introduction of EOSC credits could allow for alternative mechanisms, such as:

- Researchers could apply for a specific data management budget when they apply for a research grant. This budget would be paid in EOSC credits.
- Researchers are paid in EOSC credits each time their data set, analytical model or research tool is accessed. However, as discussed above, this would require the EOSC or nominated agents to monitor usage of each dataset, analytical model and research tool.
- EU and Member States' funding agencies insist that a specific proportion of a research grant is used to pay for the storage, curation and management of research data via the EOSC. In effect, this could mean top-slicing research grants so that the research team must spend 2-3% of its received funding on data stewardship services from its chosen (and qualified) provider. The funding agency could make these funds available as EOSC credits that can be spent on any accredited cloud service.
- The funding agencies could link research funding to the identification of a capable agency that will provide the data stewardship services and products that will carry and manage the data long term. In most cases, this would be the research institution itself (as an IT or librarianship function), but researchers should have the flexibility to employ other service providers that meet the minimum standards specified by the EOSC. Note, a rigid top-slicing mechanism may face opposition from some researchers who want to retain the flexibility to spend all of their grant on conducting their research, if that proves necessary.

The use of EOSC credits could help drive competition between cloud providers: the researcher's IT specialists would spend the credits with the provider offering the best value. Such a system would also allow for measurement of activity (since the EOSC/Commission would see where the credits were getting redeemed) and help ensure that European science isn't becoming overly dependent on a single cloud vendor. EOSC credits could only be redeemed for services from accredited cloud service providers. This would prevent money being spent on facilities that don't meet the required standards, while helping to ensure privacy and security are respected in line with the EU's regulations including the General Data Protection Regulation.

Accredited cloud service providers will also need to have the capacity and capabilities to serve a lot of groups suddenly taking a close interest in a particular "hot" dataset and wanting to run workloads against it when a significant discovery is announced, and thus requires and attracts heavy scrutiny.

Moreover, an accredited cloud service provider should be able to deliver a pricing model, which is transparent about availability levels and data durability, and - importantly - is a function of total factor costs. That would reduce the risk that a university-built facility, for example, under charges for a service, because it is unaware that it is building on implicit or explicit subsidies (e.g. they don't factor costs of power or cooling or land into their "sell" price).

Such subsidies could become unsustainable if a university's low-priced cloud services become wildly popular, since the hidden mechanism in the back paying for the subsidy won't be ready for the onslaught (e.g. the campus estates department gets a bill for electricity they simply couldn't predict).

This could result in the long-term safety and stewardship of the data being imperilled. Regardless of whether the service provider is a private company or a publicly funded institution, total cost/price calculations should not rely be distorted by sunk costs and subsidies.

The EOSC should allow for business model innovations and dynamic competition in terms of pricing and proposition. For example, as they compete for EOSC-related usage, commercial cloud providers may wish to use hybrid models, involving so-called freemium models, in which the basic service is free, but users pay for premium features. Note, only a single copy of each dataset needs to be stored in the cloud (assuming the IaaS provider has appropriate back-up mechanisms to ensure data isn't lost in the event of a disaster): users pay for resources to process data and store outcomes, rather than paying to host their own copy.

Engaging the big data science factories

The success of the EOSC will depend on the active engagement of the major research infrastructures, such as CERN, EMBL, ESO, ESRF, and ILL, and e-infrastructures, such as GÉANT.

Some of the research infrastructures, known as the big data science factories, already make some data sets and analytical tools available in an open manner – in many cases, policy makers regard public usage of the data generated by research infrastructures as a key performance metric. For that reason, many of these organisations are likely to want to incorporate their existing propositions into the EOSC if the science cloud will make it easier for researchers to access their data.

For data that is already available on FAIR terms, the process of making this data accessible through the EOSC should not be too arduous or expensive. In short, the EOSC could help the big data science factories demonstrate how useful their work is to the wide scientific community and help them justify their public funding.

CERN, for example, already has two services that are designed to enable the wider world to access subsets of the data its experiments generate: the open data portal and Zenodo.

The data providers (LHC collaborations in the case of the open data portal or individual users in the case of Zenodo) have taken the effort to make their data understandable to others and agreed to make it open. However, in the case of the LHC experiments, only a fraction of their data is visible through the open data portal because of what is commonly known as an 'embargo period' and also because of the significant effort it takes to prepare it for re-use by people outside the experiment. Zenodo, is different in that it is open to everyone, including individual researchers, and hosts datasets from a multiple of research disciplines, not just high-energy physics. If CERN were to make these services compatible with the EOSC, it could extend their reach and help it further demonstrate the usefulness of its data to researchers.

However, some research infrastructures will need to be able to recover the cost of making their data and tools available through the EOSC. Once it is generating revenues, the EOSC could provide these research infrastructures with credits they can spend on the cost of using SaaS, PaaS and IaaS services to make their data and tools available through the EOSC.

Note, the EOSC High Level Expert Group suggests the Commission could, through the European Data Infrastructure (EDI) initiative, provide economic incentives for research infrastructure providers to use and co-develop shared facilities and data repositories through the EOSC.

Although these economic incentives would need to be temporary, they could be sufficient to persuade the big data science factories to make their existing open data services compatible with the EOSC.

Conclusions

Although it will ultimately need to break even, the initial priorities for the EOSC have to be driving participation and usage. Like most private businesses, the EOSC will probably need to operate at a loss (be subsidised) in its early years to ensure its proposition is appealing to both the data providers and the data users.

However, once it is benefitting from network effects, the EOSC will be able to implement a sustainable business model. To do that, the EOSC will need to be highly relevant for the private sector. As outlined by the European Investment Bank Advisory Services in a recent report[1] , access to relevant, commercially interesting data will clearly be needed to engage businesses in the EOSC.

The private sector has already shown strong interest in the data being created by some of Europe's leading research infrastructures, such as the ESA and CERN. To ensure these big science data factories make their open data services compatible with the open science cloud, the EOSC has to provide these players with the distribution reach and the monitoring tools they need to demonstrate to their funders the value of their experiments and research. It can do that by making it easy for the data factories to engage with the 725,000 researchers working in higher education across the EU, as well as those launching commercial spin-offs.

To ensure researchers in higher education engage with the EOSC, the freemium model outlined in this paper is designed to keep the barriers to usage as low as possible.

The aim is to enable researchers to explore the potential of the open science cloud without incurring significant financial expenditure. As the number of data providers and data users increases, the value of the EOSC to both groups should rise, attracting greater participation. Once the open science cloud is established as the preferred means of making scientific data open and accessible, the EOSC will have the scope to modify the freemium model to ensure it can become self-sustaining.

1 Financing the future of supercomputing: How to increase investment in high performance computing in Europe. Prepared for: DG Research and Innovation and DG Connect European Commission By: Innovation Finance Advisory European Investment Bank Advisory Services

SCIENCE|BUSINESS®
NETWORK

The Science|Business Cloud Consultation Group

In November 2018, the EU will launch a bold initiative to pull European science into the 'cloud.' The €6.7 billion European Open Science Cloud promises to change the future of research, cloud services and data technologies in Europe - changing the research game for industry, academia and governments. But how?

To track this important initiative, Science|Business runs a platform for companies, universities and research agencies. The group meets regularly in private and public to share intelligence and facilitate dialogue in Brussels. Topics include:

'Rules of Engagement'

Under what terms can private vendors, and public users, access the science cloud and the huge market – upwards of €50 billion – it can create?

Finance

Should public research grants include budgets to buy cloud services? Will labs and universities pay centrally? When corporate labs use it, will they be charged?

Infrastructure

How will broadband networks, data centres and existing research infrastructures be affected by the fast-growing new services?

Science data

How can total quality management of data be enabled throughout its lifecycle?

New technologies

How should the science cloud harness ground-breaking advances in computing and connectivity, such as AI, quantum computing and 5G?

Global strategy

How will Europe's science cloud be integrated into global research and on what terms?

C L O U D

Avenue des Nerviens 79, Brussels, Belgium 1040
info@sciencebusiness.net
+322-304-7577

SCIENCE|BUSINESS® NETWORK

Bringing together industry, research and policy

Industry

Amazon
Amgen
Dow Europe
Frontiers
GE
Google
Huawei
Microsoft
Nickel Institute
Novartis
Pfizer
Sanofi
Total
Toyota

Academia

Aalto University
ESADE Business & Law School
ETH-Zurich
Karolinska Institutet
KTH Royal Institute of Technology
KU Leuven
Norwegian University of Science and Technology (NTNU)
Politecnico di Milano
Polytechnique Montréal
Sorbonne University
Tallinn University of Technology (TalTech)
Trinity College Dublin
TU Berlin
University College London
University of Amsterdam
University of Birmingham
University of Bologna
University of Eastern Finland
University of Luxembourg
University of Pisa
University of Twente
University of Warwick

Public organisations

Barcelona Supercomputing Center
Business Finland
CERN
Centre National de la Recherche Scientifique (CNRS)
The COST Association
Eureka
Innovate UK
Innovation Norway
Republic of South Africa - Department for Science and Technology
European Investment Bank
European Space Agency
Fraunhofer

Associations

ACM Europe Policy Committee
ATTRACT
The Association of Commonwealth Universities (ACU)
European University Association (EUA)
Photonics21
The Guild

Group members

Trimbos Instituut
Simmons & Simmons
Deusto International Research School (DIRS)
Sant Joan de Deu - Barcelona Hospital
Tataj Innovation
GEANT
ICHOM - International Consortium for Health Outcomes Measurements