Recommendations for a FAIR EOSC

White Paper of the FAIRsFAIR Synchronisation Force

FAIRsFAIR - Fostering FAIR Data Practices in Europe

Authors

Ingrid Dillo (DANS), Simon Hodson (CODATA),

Sara Pittonet Gaiarin (Trust-IT),

Marjan Grootveld (DANS)

https://doi.org/10.5281/zenodo.5744786

Acknowledgements

Disclaimer

Abstract

This White Paper provides recommendations to the
EOSC Partnership and the EOSC Association Task
Forces in three areas related to the Strategic Research
and Innovation Agenda (SRIA) of the EOSC:

1) FAIR by Design,

2) sustainability of the components: the federated
data layer, and

3) a strategic and targeted approach to training.

Addressing the recommendations will require the
involvement of other stakeholders as well, such as
research infrastructures, research funders, policy
makers and projects, at national and international,
domain and cross-domain levels. Recommendations
for instruments to synchronise FAIR-related activities
conclude this White Paper.

Table of contents

Abbreviations and Acronyms

AAI	Authorisation and Authentication Infrastructure
API	Application Programming Interface
EOSC	European Open Science Cloud
EOSC-A	EOSC Association
FAIR	Findable, Accessible, Interoperable, Reusable
maDMP	machine-actionable Data Management Plan
MoU	Memorandum of Understanding
MVE	Minimal Viable EOSC
PID	Persistent Identifier
RDA	Research Data Alliance
RDM	Research Data Management
RI	Research Infrastructure
RPO	Research-Performing Organisation
SRIA	Strategic Research and Innovation Agenda
TFiR	Turning FAIR into Reality

Executive Summary

Using the 2018 Turning FAIR into Reality (TFiR) report by the European Commission expert group on FAIR data, the Synchronisation Force workshops identified a number of areas for which more activity is recommended.

In this White Paper the FAIRsFAIR main recommendations are taken a step further, and mapped against three priorities identified in the Strategic Research and Innovation Agenda of the EOSC (SRIA) and the relevant Task Forces under the EOSC Association, to facilitate their impact and uptake.

Figure 1: Eight recommendations for a FAIR EOSC, mapped to SRIA priorities and EOSC Task Forces

1. Synchronising our steps on the journey towards FAIR data

In 2014, the FAIR data guiding principles were introduced, and since then their uptake has been impressive. In 2018, the European Commission expert group on FAIR data published the Turning FAIR into Reality (TFiR)[1] recommendations and action plan. And from 2019 to 2021 the FAIRsFAIR Synchronisation Force (SF) organised yearly workshops[2]. The goal of these workshops was to chart the amount of activity in important initiatives in the EOSC and FAIR ecosystems to implement those recommendations, and to provide indicators of this amount of activity, per TFiR recommendation.

The three FAIRsFAIR Synchronisation Force workshop reports summarise the activity input from all participating initiatives, activity scores, as well as findings and recommendations for next steps.[3] As mentioned in the workshop reports, the findings and conclusions are a snapshot in time and depend on the initiatives and stakeholders represented. Furthermore, we recognise there are differences between domains in terms of the pace and amount of FAIR-related activities.

This year's workshop confirmed that there is broadly a good level of activity towards the TFiR recommendations, across the participating projects and initiatives. This is particularly strongly the case in the area of defining and communicating the definition of FAIR (Rec.1) and preparing metrics for FAIR Digital Objects and Outputs (Rec.12). Other areas are also receiving a lot of attention and commendable activities: certification for trustworthy data repositories, adoption of data management plans, interoperability frameworks and semantics, and training activities, to list a few examples. However, other areas recommended in the TFiR report require more attention or more activity; sometimes across the board, sometimes from specific stakeholders such as research communities, research infrastructures, or research funders.

In this White Paper our main recommendations are taken a step further, in that they are mapped against three priority topics in the Strategic Research and Innovation Agenda of the EOSC (SRIA[4]) and the relevant Task Forces[5] under the EOSC Association. By mapping our findings and recommendations against the core agenda of the EOSC Partnership and the charters of the main bodies of the EOSC Association, we hope to facilitate their impact and uptake. We conclude the White Paper in Section 3 by recommending four instruments for synchronisation and alignment of the various FAIR-related activities.

1   European Commission Expert Group on FAIR data (2018). Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR data. https://doi.org/10.2777/1524

2   FAIRsFAIR Synchronisation Force: https://www.fairsfair.eu/advisory-board/synchronisation-force

3   Coen, G., Mokrane, M., Pittonet, S., Hodson, S., & van Kessel-Hagesteijn, R. (2020). D5.3 Report on the First Synchronisation Force Workshop (1.0). FAIRsFAIR. https://doi.org/10.5281/zenodo.5361052 ; Dillo, I., Grootveld, M., Hodson, S., & Pittonet Gaiarin, S. (2020). Second Report of the FAIRsFAIR Synchronisation Force (D5.5) (1.0). Zenodo. https://doi.org/10.5281/zenodo.5361417; Grootveld, M., Hodson, S., Pittonet Gaiarin, S., Davidson, J., & Dillo, I. (2021). D5.6 Report 3 of the Synchronisation Force (V1.0_DRAFT). Zenodo. https://doi.org/10.5281/zenodo.5336658

4   EOSC Executive Board (2021). Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC). https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

5   EOSC Association Task Forces 2021-...: https://www.eosc.eu/task-force-faq

2. Recommendations for a FAIR EOSC

Using the TFiR report, the Synchronisation Force workshops identified a number of areas for which we recommend more activity. In this White Paper, we have distilled the ones we believe to be most important. Below, we show how they align with three priority action lines in the SRIA and can be mapped against the charters of the EOSC Association (EOSC-A) Task Forces.

The link with the SRIA

The SRIA is primarily concerned with the steps needed, over a six-year timescale from 2021, to build the EOSC. A lot of attention is paid to a staged approach for EOSC infrastructures and services, built on three levels: 1. the EOSC-Core; 2. the federation of research data services; 3. the EOSC-Exchange for data and research services. Some of the priorities and recommendations can be characterised as preconditions for a functioning Open Science and FAIR infrastructure at an international scale. This includes the network, Authentication and Authorisation Infrastructure (AAI), and rules of engagement which fall somewhat outside the scope of TFiR and the discussions of the FAIRsFAIR Synchronisation Force. Some other near-term SRIA priorities, such as PID infrastructure and FAIR assessment, fall into the scope of our discussions.

In this White Paper, the recommendations from the FAIRsFAIR Synchronisation Force workshops are mapped against the following three priorities identified in the SRIA and the EOSC Task Forces.

1 - FAIR by Design: the SRIA uses this phrase to underline the need for concerted and targeted effort to ensure that research data in the EOSC can be accessed and used computationally, and at scale. In other words, that EOSC research data should be machine-actionable and available for computational analysis such as machine learning and Artificial Intelligence. To this end, the SRIA reiterates firmly the keystone

TFiR Recommendation 4 to 'develop interoperability frameworks for FAIR sharing within disciplines and for interdisciplinary research'.[6]

There is little doubt that this is one of the most challenging areas described in the TFiR recommendations, and it is unsurprising that a number of the recommendations from our workshops fall under this category.

2 - Sustainable and Federated Data Layer: the SRIA presents a strategy to develop the EOSC-Core, on which can be built the EOSC-Exchange, a rich set of services, some rivalrous, in a form of marketplace. Between these, as the purpose of the former and the precondition of the latter, is a layer of federated data. Discussions of sustainability in the SRIA are focussed on EOSC-Core and this federated data layer. Discussions of sustainability in the EOSC Working Groups[7], as presented in the SRIA, have largely focused on the EOSC-Core or the Minimal Viable EOSC. This is understandable given the need for a stepwise approach. Nevertheless, the recommendations from our workshops indicate a pressing need to pay attention to the development and sustainability of a federated layer of data and the functionality to exploit it.

3- Strategic and Targeted Approach to Training: skills and training are identified in the SRIA as one of the 'boundary conditions' for the EOSC. If Europe is not to lose ground and miss out on the benefits of Open Science, it needs 'to ensure the availability of highly and appropriately skilled people'.[8] A welcome emphasis in the SRIA is the need for targeted training and professionalisation of key roles. Professional roles in Open Science and Data Stewardship are needed with clearly defined skills profiles and career paths. Certification systems must rest on accepted definitions of skills and curriculum frameworks. This chimes with the findings of the FAIRsFAIR Synchronisation Force workshops: there is a lot of training activity, but more focus is needed on the development of an agreed EU data steward curriculum framework, supported by the

---

6   SRIA pp.61-2, referencing TFiR Rec.4. See also 'Key strands of work were identified in the 'Turning FAIR into reality' report which set priorities for the implementation of a Web of FAIR Data that should be pursued on an EU level under Horizon Europe and in national and institutional funding cycles. These have been further validated by the EOSC FAIR Working Group and activities it has undertaken to assess FAIR practices across research communities and propose the EOSC Interoperability Framework.' p.71.

7   EOSC Working Groups 2019-2020: https://www.eoscsecretariat.eu/eosc-working-groups

8   SRIA, p.125.

establishment of professional accreditation bodies and certification processes.

In the next three sections, we make eight recommendations and relate each of them to the relevant EOSC Association Task Force(s).

2.1 Core recommendations: FAIR by Design

A central objective of the FAIR principles, that is sometimes neglected, is to better facilitate the machine-assisted analysis and reuse of data, at scale and using Artificial Intelligence. To achieve this the data need to be programmatically accessible, and described by machine-readable ontologies, terminologies and metadata. The SRIA puts this neatly: '*the machine must be able to find data ('know where it is'), then to be able to access and identify ('know what it is'); in order to operate on the data the machine needs to know what can be done with this object ('know how it can be handled') and for reusing the digital object the machine needs to know what it is allowed to do with it ('know which actions are allowed')*.'[9] Achieving this level of rich, machine-actionable metadata requires agreements in research domains, what TFiR describes as 'interoperability frameworks'.

It was evident from the FAIRsFAIR Synchronisation Force activities that these issues, highlighted in TFiR Recommendations and reiterated in the SRIA, were being little addressed by the participating projects, resulting in the following recommendations:

Rec.1: Develop domain and cross-domain interoperability frameworks at the level of vocabularies, ontologies, and metadata schema.

TFiR and the SRIA both emphasise the need for interoperability frameworks, rich metadata and semantics, in order for data to be interoperable and reusable. It is clearly recognised that 'interoperability frameworks' need to be further developed at the domain level. Increasingly, however, to address pressing interdisciplinary research topics, effort also needs to be targeted at 'interoperability frameworks' that work across domains. The SRIA makes this point by referencing and paraphrasing TFiR Rec.4: '*Research communities need to be encouraged to develop and maintain interoperability frameworks that define their practices for archiving, referencing and describing research artefacts of all types and formats. To support interdisciplinary research, these interoperability*

frameworks should be articulated in common ways and adopt global standards where relevant. Intelligent crosswalks, brokering mechanisms and semantic and other technologies such as artificial intelligence, should all be explored to break down silos and allow cross-disciplinary exploration, analysis and visualisation.'[10]

Addressing this recommendation will require projects (or other initiatives) that enable European Research Infrastructures (RIs) to work closely with a number of pertinent cross-domain case studies. In this process, it will be essential to engage with international domain and cross-domain activities and the organisations that maintain metadata specifications and semantics.

Relevant EOSC-A Task Forces: EOSC Semantic Interoperability Task Force. To a secondary degree, the EOSC FAIR Metrics and Data Quality Task Force.

Rec.2: Further develop and implement semantic technologies, particularly in domains where their use is less advanced.

In metadata and in semantics, just as in other aspects of good data practice, some research disciplines have made considerable advances. Some parts of the life sciences, astronomy, biodiversity and crystallography are often cited in this respect, though we should not infer by this that no further work is needed. Many other research disciplines have made considerably less progress with good RDM practices generally and the development and adoption of community-accepted metadata specifications and semantics. This is an issue for many disciplines, but also for certain institutions (RPOs and data archives) in some European countries. Understanding the drivers and context which has led some disciplines to advance their use of semantic technologies, while others have made less progress, will also be important.

Addressing this recommendation will require the participation of European RIs and projects. It will also require targeted activities for the domains and countries/institutions that need extra attention and resources. As will all matters of normalisation

---

9   SRIA, p.60.

10   SRIA, p.61.

in metadata and semantics, it will be essential to engage with international domain organisations that maintain metadata specifications and semantics. The appropriate level of resources should also be provided for the maintenance of established disciplinary interoperability frameworks.

Relevant EOSC-A Task Forces: Semantic Interoperability. To a secondary degree, FAIR Metrics and Data Quality.

Rec.3: Further develop machine-actionable DMPs, taking the Science Europe domain protocol approach as a start.

Data Management Plans, and the development of good practice around their use, have received a lot of attention. There is interest in making DMPs machine-actionable for a variety of reasons: e.g. to make them more effective as 'living' plans, to allow them to be active components in rules-based RDM data stewardship systems, and to better facilitate the use of DMP information in other areas of the research ecosystem, strategic planning for example. Initial work has started in RDA on this topic[11] and should be extended. The approach signposted by Science Europe to domain DMPs[12] should be followed and taken up by domain representative bodies, including European RIs and international organisations. It would be instructive also to extend this approach to domains which have made less progress in RDM and to emerging interdisciplinary research areas.

Addressing this recommendation will require extension of the existing exploratory work in RDA and the development of robust machine-actionable DMPs (maDMPs) as a service, which could ultimately be delivered through EOSC-Exchange, with appropriate user support from institutional/RPO and RI data stewardship competence centres.

Relevant EOSC-A Task Forces: Overseeing the further development of maDMPs (for example, by brokering the coordination necessary to make sure they take adequate account of different disciplinary needs) does not seem to fall under the remit of any of the current EOSC-A TFs and we recommend

that this is addressed. Funded project effort will be needed to build on the RDA Recommendation 'RDA DMP Common Standard for Machine-actionable Data Management Plans'.[13] Institutional ownership of some nature will be necessary for development to product and the sustainability of such services: such a service could, for example, be provided within EOSC-Exchange by an appropriate service provider.

Rec.4: Facilitate the Machine Processing of Data.

As noted above, one of the objectives of the FAIR principles, reiterated in TFiR and the SRIA, is to facilitate the machine processing of data in the research process. The PwC cost-benefit and opportunity cost analysis conducted for the European Commission[14] estimated that 80% of research project effort was expended on data wrangling of various sorts, upstream of analysis. One promise of the web in this digital age is to better mobilise computation to extract knowledge from complex data. The FAIR principles may be interpreted as a high level check list of the things that need to be done to the data in order to facilitate this. This recommendation is, therefore, at the heart of the Open Science, FAIR and the EOSC enterprise, but also extremely challenging.

A key objective of EOSC is to enable machine processing of FAIR data. This means programmatic access (under conditions where necessary). It also means enriching metadata in a number of ways, for example, with information about the 'type' of digital object (including what can be done to the FAIR Digital Object, by whom and how). It is also necessary to provide metadata about the structure of the data; about the relationships between measurements, variables and concepts; and, about the provenance of the data and the processing it has undergone. All this needs to be machine readable and actionable. The challenge in realising the scale of ambition in the FAIR principles is largely 'upstream': meaning that investment is required in data stewardship and in services that provide FAIR, research-ready data. There are also tasks required downstream in supporting research projects to demonstrate the potential of machine processing

---

11 Research Data Alliance: https://www.rd-alliance.org/groups/dmp-common-standards-wg and https://www.rd-alliance.org/groups/exposing-data-management-plans-wg

12 Science Europe (2018). Guidance Document Presenting a Framework for Discipline-specific Research Data Management. https://doi.org/10.5281/zenodo.4925907

13 https://doi.org/10.15497/rda00039

14 European Commission, PwC EU Services (2019). Cost of not having FAIR research data: Cost-benefit analysis for FAIR research data. https://doi.org/10.2777/02999

of FAIR data, and in further developing the necessary technologies.

Addressing this recommendation will require addressing the challenge of making data as machine actionable as possible. Ultimately this requires significant 'metadata uplift', as detailed above, and therefore greater investment in data stewardship and in data repositories and RIs that can provide research-ready, FAIR data. Input is required from the developers of software tools to facilitate the way semantics, FAIR vocabularies and ontologies are applied in data services. Funded projects also have a role in helping achieve the necessary 'metadata uplift' and to apply this in RIs, and to refine the APIs and downstream tools needed to undertake the machine processing of data.

Relevant EOSC-A Task Forces: Semantic Interoperability; to some extent FAIR Metrics and Quality.

2.2 Sustainability of the components: federated data layer

The SRIA presents a vision of EOSC in which the EOSC-Core enables FAIR access to a federated data layer and the EOSC-Exchange provides a marketplace for services operating on that data layer. This notion of the data layer, of data as infrastructure, even of data as a service, is longstanding and was one of the key elements of the Riding the Wave report[15], for example. The federation, interoperability, maintenance and sustainability of that data layer requires attention. A stepwise process to build capacity in data providers will be important. Sustainability is an important topic in the SRIA, but primarily concern the EOSC-Core (understandably enough given the need for a stepwise approach). Nevertheless, the data layer is complex, involving many services and institutions, at various national and European levels, and will require detailed attention from the perspective of ensuring sustainability. One approach is to include 'gate-keeper' requirements as part of the Rules of Participation, but in order not to exclude important but less well-funded services supporting smaller communities, this should be reinforced with investment and assistance in developing a range of business models and sustainability plans appropriate for a range of circumstances. Similarly, the key recommendations in the SRIA relating to cost assessments concern the EOSC-Core and the MVE.[16] It will be essential not to neglect costing and funding models, and other activities pertaining to the sustainability of the federated data layer. Task 3 of the EOSC-A TF on Defining Funding Models for EOSC is to 'Develop scenarios for financial sustainability of the federation of data and data services'. In summary: the scope of this needs to include the entities which will provide the data in that federated layer, not just the (considerable) task of alignment, interoperability and federation itself.

Rec 5. Provide continuous guidance and assistance to small repositories to engage with certification processes.

There has been substantial and laudable progress in the extension of certification and good practices for data repositories, particularly through the use of the CoreTrustSeal[17]. Nevertheless, 'meeting the CoreTrustSeal requirements remains a challenging task for smaller repositories and for institutional repositories'[18]. If the federated data layer is to include small repositories, which are important in a substantial range of domains and geographies, guidance, support and capacity building for these repositories is also required.

Addressing this recommendation will require capacity building and competence centres, explicitly tasked with assisting smaller, less-well funded research domains, institutional repositories and other services, and geographies/regions where research infrastructures are less advanced. Such support could be provided by European Research Infrastructures and Coordination and Support Actions. Likewise, a network of trusted repositories could play a key role in providing advice and peer to peer support.

Relevant EOSC-A Task Forces: Data stewardship, skills, roles and infrastructures, need to be developed and sustained in less well-funded research domains, through national and European investment. The TF with the most obvious role is Upskilling Countries to

---

15 European Union (2010). Riding the wave: How Europe can gain from the rising tide of scientific data.

16 SRIA, p.124.

17 https://www.coretrustseal.org/

18 Grootveld, M., Hodson, S., Pittonet Gaiarin, S., Davidson, J., & Dillo, I.. (2021). D5.6 Report 3 of the Synchronisation Force (V1.0_DRAFT). Zenodo. https://doi.org/10.5281/zenodo.5336658. p.16.

Engage in EOSC. Additionally, the TF on Long Term Data Preservation will produce recommendations on the creation of a European network of trustworthy digital repositories following FAIR-enabling principles with disciplinary and geographical spread, as well as recommendations for EOSC data services to connect to this network and a roadmap to further mature the long-term preservation aspects of these repository services.

Rec 6. Define costs for FAIR data stewardship (and how much it would cost not to have FAIR data stewardship).

Notwithstanding efforts to cost research data stewardship and to estimate the opportunity cost of not having FAIR data stewardship, there remains a need to further develop and operationalise three important practices in relation to data stewardship costs as better processes and support structures are put in place. Firstly, RIs, repositories and institutions need practical guides to costing and resourcing data stewardship; secondly, real-world cost information needs to be gathered throughout the data lifecycle, data preservation included; and thirdly, this cost information needs to be presented to senior management and to funders to help substantiate and quantify the necessary investment.

Addressing this recommendation will require funded projects to combine and further refine existing costing methodologies and to present these as practical guidance (notwithstanding regional and institutional variations). In this context, it will be essential to understand the costs of data stewardship in RPOs (which remains little understood) and to determine the level of investment and support needed to improve capacity. Furthermore, it requires processes for comparing costs across different institutions and stakeholders and appropriate sharing of costing information with senior management bodies, stakeholder groups and funders to assist

better planning and more targeted investment.

Relevant EOSC-A Task Forces: These issues are primarily a consideration for the EOSC-A TF on Financial Sustainability for EOSC and the EOSC-A TF Long Term Data Preservation.

Rec 7. Organise joint workshops involving research funders, policy makers at various levels, and other stakeholders to survey activity in the areas of funding, sustainability, governance and metrics.

This recommendation relates to the process undertaken by the FAIRsFAIR Synchronisation Force, through the workshops and in preparing the series of reports. As mentioned in Section 1, findings from the workshops are based on the projects and initiatives present; some stakeholder categories were underrepresented. Some TFiR recommendations rest more with or need the involvement of funders and policy makers. This is most obviously the case in those recommendations relating to funding, sustainability and governance, but also to those concerning metrics and incentives. Future synchronisation activities need to include funders and policy makers, as well as other stakeholders.

Addressing this recommendation will require commitment on the part of funders and policy makers to engage with the process. Synchronisation activities require input from all stakeholders, including funders and policy makers. More specific targeted activities may be needed to ensure that the workshops are useful for all stakeholders and lead to further implementation steps.

Relevant EOSC-A Task Forces: This recommendation relates to the work of the EOSC-A TF Financial Sustainability for EOSC, and to some degree to the TFs Rules of Participation Compliance Monitoring and Research Careers, Recognition, and Credit.

2.3 Strategic and targeted approach to training

Building capacity for FAIR data, for Open Science practices, remains an urgent priority for EOSC and for the European research space more generally. The SRIA underlines the need for ongoing culture change, assisted by improved rewards and recognition, incentives and strong governance. Most of all there is a need for very targeted training and the development of specific skills in a professional cohort realising EOSC.

Rec 8. Develop an EU data steward curriculum framework, including the establishment of professional accreditation bodies and certification processes.

It was evident from the FAIRsFAIR Synchronisation Force workshops that there was considerable activity in training through various RIs and projects. At the same time, there was a clearly stated recognition that

this activity needed to be accompanied by a number of targeted initiatives. These include, developing, normalising and implementing curriculum frameworks for data stewards and other data professionals needed by EOSC, building on existing work including by FAIRsFAIR and LIBER. The same process should also be undertaken to formalise certification and to encourage professional bodies for data stewardship as a profession. Finally, there needs to be clearly defined career profiles such that institutions can more easily identify their workforce requirements, create posts and recruit the necessary staff, and such that young professionals have a clear understanding of the skills needed in their portfolio to make a career as a data professional in an EOSC related institution. These same requirements are stated very forcefully in the SRIA.[19]

Addressing this recommendation will require significant coordinated effort through project activities, principally whichever project is funded through the HORIZON-INFRA-2021-EOSC-01-01 Call 'Supporting an EOSC-ready digitally skilled workforce'. The participation of policy makers and RPOs is essential to achieving these objectives.

Relevant EOSC-A Task Forces: This work relates to the remit of the EOSC-A TF Data Stewardship Curricula and Career Paths.

---

19 SRIA pp.125-30.

3. Instruments for synchronisation and alignment

The implementation of EOSC involves a huge and expanding number of organisations, projects, and individuals. The FAIR principles touch upon many aspects of EOSC and so are relevant, or even essential, for a large, diverse group of organisations, projects, and individuals. Implementing the FAIR principles requires conscious efforts towards synchronising and aligning both information and developments throughout EOSC, and towards avoiding unnecessary duplication of work. In this section we describe four instruments for synchronisation and aligning that have been used in FAIRsFAIR, which we believe could be usefully extended into other projects and initiatives.

3.1 Cross-project collaboration agreement

In the *INFRAEOSC-05-2018-2019 Support to the EOSC Governance* call seven projects were funded. All of them are party to a formal collaboration agreement[20]. The purpose of this agreement was to create synergy in the mutual activities related to the EOSC and to provide a framework for planning, implementing and monitoring the joint activities described in the agreement. To do so, the projects established a cross-project collaboration board plus six task forces focusing on FAIR data and infrastructures, Landscaping, Service onboarding, National policies and governance, Training and skills, and Dissemination and events. Each project is represented in the board and in each task force.

The bimonthly task force meetings serve to exchange information, compare and discuss developments, and identify collaboration opportunities. Output examples so far include a working report[21] and validation workshop[22] for indicators for EOSC readiness, a FAIR assessment workshop for representatives of the European Commission[23], a project providing webinars to another project, joint training events for researchers, and projects that co-organise conference workshops. An immediate and positive effect of this setup is also that individual project members get to know colleagues in 'sibling' projects. This makes it easier to reach out to them. Undoubtedly the FAIRsFAIR Synchronisation Force workshops (see 3.3) benefited from this, as the sibling projects were very engaged.

The European Commission mandated this collaboration and the type of agreement, after having funded the seven projects. Especially the initial stage of defining the scope, the activities that could be undertaken together, and the formal setup of the board and task forces, was intense because the extra effort didn't naturally fit the individual project plans and capacity, and the type of agreement was rather heavy-weight. However, without this requirement less cross-project output and collaboration may have emerged. In the case of the INFRAEOSC-05 projects it helps that EOSC Secretariat, one of the seven projects, provided logistical support for all board and task force meetings. The EOSC Secretariat also liaised between the collaborating projects and the EOSC Working Groups[24].

---

20 Collaboration Agreement: https://www.eosc-pillar.eu/infraeosc-5-collaboration

21 Arvola, M., Beckmann, V., Budroni, P., … Van Wezel, J.. (2021). Second Working Proposal for Living Indicators to Monitor MS Progresses Towards EOSC Readiness. Zenodo. https://doi.org/10.5281/zenodo.4452799

22 Indicator validation workshop: https://www.eoscsecretariat.eu/events/EOSC-validation-workshop-project

23 S. Pittonet Gaiarin (2020). FAIR assessment and certification in the EOSC region. Zenodo. https://doi.org/10.5281/zenodo.4486280

24 EOSC Working Groups 2019-2020: https://www.eoscsecretariat.eu/eosc-working-groups

3.2 Memorandum of Understanding

Projects or other initiatives can establish Memoranda of Understanding (MoU) to clarify the relationship between two initiatives and specify their joint objectives, ambitions, and agreements. An MoU can be strategic or rather focus on practical work. It can help avoid unnecessary duplication of effort, foster harmonisation in ways of working, and provide a mechanism for transfer of knowledge. The FAIRsFAIR project signed MoUs with the EOSC-hub project, the FAIRsharing initiative and the FAIRware project; all slightly different agreements with different goals. The MoU with EOSC-hub[25],

for example served as a mechanism to make sure that FAIRsFAIR could build on and carry forward the work of EOSC-Hub which was ending, in other words, to sustain project outputs. An MoU increases the partners' visibility in both their networks, while at the same time it can reduce the complexity for engagement with certain stakeholders. Finally, an MoU can also help to manage expectations of what each partner will contribute so that they enter into collaboration with a clear perception of the intended outcomes.

3.3 Synchronisation Force

A Synchronisation Force is a team dedicated to periodically bringing together initiatives that promote and work towards a common goal. The FAIRsFAIR Synchronisation Force was set up to maintain a dialogue across the complex EOSC and FAIR ecosystems so as to maximise collaboration, minimise duplication, and promote adherence to the Turning FAIR into Reality recommendations. The team reached out to the former EOSC Working

Groups, research infrastructures, e-infrastructures, and EOSC-building projects. The FAIRsFAIR's Synchronisation Force organised annual workshops as a series of online sessions. Here, each represented initiative reported on their progress against the 27 TFiR recommendations. These workshops were perceived as very useful, as can be seen in the text box.

What was the main value of the 2021 Synchronisation Force workshop for you?

The opportunity to give our say and get an overview of what is going on

Unique summary of inputs from a lot of people on these subjects

A reality check for the European Commission regarding the implementation of the EOSC

Find collaborators for realising project objectives together

Furthermore, the FAIRsFAIR Synchronisation Force organised 'FAIR National Roadshows'[26] with the aim to accelerate FAIR adoption in different countries whilst connecting to national initiatives. The Roadshow hosts are for instance national Open Science coordinators and national data service providers. Based on a simple

blueprint for engaging ninety-minute online sessions, the individual Roadshows are adapted to the audience - typically the research data community - and the country. National FAIR champions or data repositories can be involved, as fits the national interest and the FAIRsFAIR network.

---

25 https://www.fairsfair.eu/news/eosc-hub%C2%A0and-fairsfair-sign-memorandum-understanding; https://www.fairsfair.eu/news/rori-and-fairsfair-partner-fairware-initiative-funders-across-uk-europe-and-canada-join-force;https://www.fairsfair.eu/news/fairsharing-fairsfair-join-forces-support-repositories-all-around-europe-their-effort-towards

26 FAIRsFAIR National Roadshows: https://www.fairsfair.eu/events/fairsfair-national-roadshow-series

3.4 FAIR Champions

The FAIR Champions[27] are scientific experts and 'doers' in the field of FAIR data. The Champions work as FAIR ambassadors, sharing implementation stories, enhancing synergies, contributing to training activities and webinars, and encouraging cross-domain engagement with FAIR. Each Champion represents a distinct group of FAIRsFAIR stakeholders, for instance data centres, research councils, research-performing organisations, or academic libraries. They give advice on operational challenges and promote new policies to improve the adoption of FAIR data and support the EOSC.

3.5 Pointers for good practice

Our experience of the above instruments used in FAIRsFAIR, indicates that all activities that widen participation and foster greater collaboration across projects, are beneficial, or even essential, for an initiative that is as broad and far reaching as EOSC. Although the degree of formality varies, it is paramount to realise that in all cases the human component is an essential ingredient for the success of the instrument. We make the following observations with respect to the implementation of collaborative instruments such as the above. These pointers for good practice are targeted at the projects and initiatives themselves as well as the EOSC Partnership.

Be pro-active

Based on our experience in FAIRsFAIR we recommend organising collaboration across projects in the same funding call and/or with related remits.

Be on time and make it doable

The organisation of the collaboration mentioned above should be done before the projects start and preferably when there is still room to allocate capacity to the collaboration. Ideally the call for funding should require the allocation of resources for collaboration.

Be as lightweight as possible

Furthermore it is important to strive for lightweight MoUs which also address operational support for the implementation of the collaboration identified in the agreements.

Sustain the success of the Synchronisation Force and FAIR Champions

Finally, the two instruments that were created in the FAIRsFAIR project itself to promote further synchronisation and alignment among projects and initiatives in the FAIR ecosystems, the Synchronisation Force and the FAIR Champions, have proven to be very effective and successful. Their continuation beyond the FAIRsFAIR project is therefore recommended.

---

27 European Group of FAIR Champions: https://www.fairsfair.eu/advisory-board/egfc

To Conclude

The EOSC vision is wonderful, but challenging. EOSC is a multi-stakeholder initiative, drawing together European member states, public and private providers, and research communities. The EOSC initiative will implement a federated 'system of systems' architecture that brings domain-specific and national projects and initiatives together in a European-wide infrastructure.[28] Implementing a FAIR EOSC requires conscious efforts towards synchronising and aligning both information and developments throughout EOSC, and towards avoiding unnecessary duplication of work. In this White Paper we described four instruments that could be usefully extended into the future and we presented the results of one of them, the Synchronisation Force. The workshops of the Synchronisation Force resulted in eight recommendations for more activity, mapped against the three priorities identified in the SRIA and the EOSC Task Forces: FAIR by design, Sustainable and federated data layer, and Strategic and targeted approach to training.
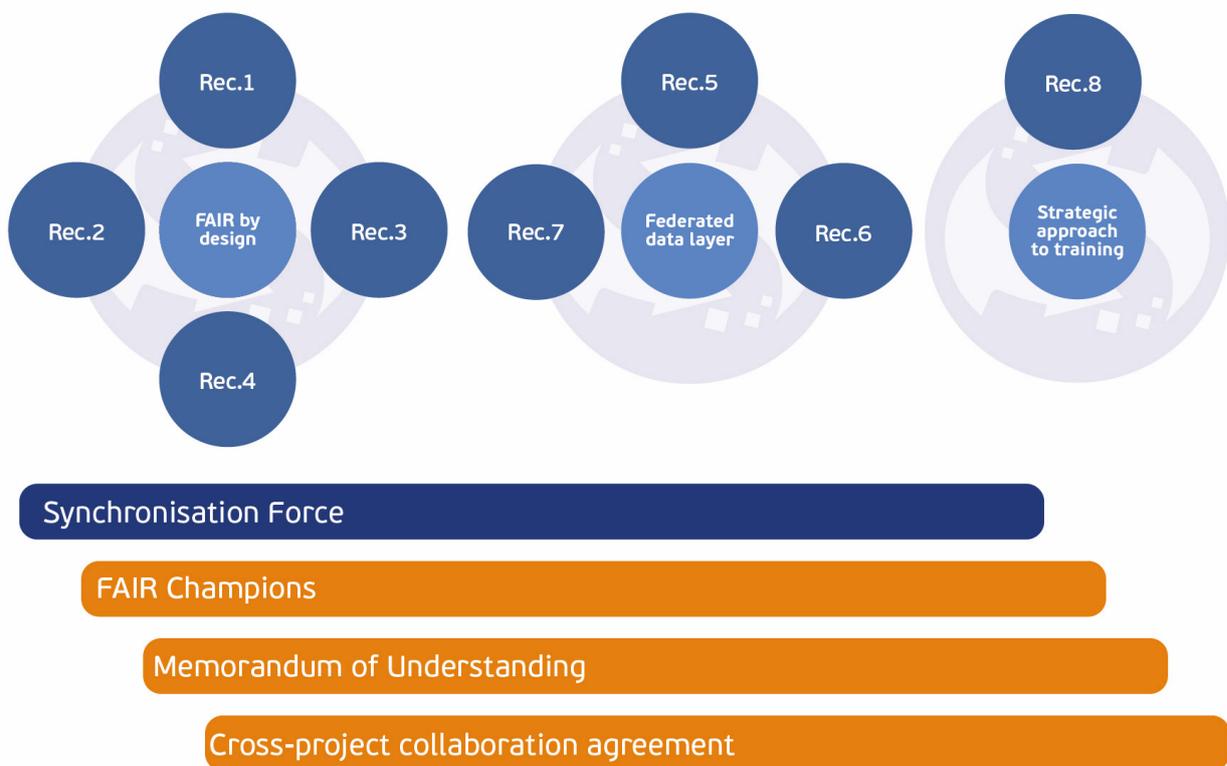

Figure 2: Four instruments to align activities and implement recommendations for a FAIR EOSC

---

28 Jones, S., Dimper, R., Dillo, I., Hanahoe, H., & Kurapati, S. (2021). Making the European Open Science Cloud (EOSC) work: where to go from here?. https://doi.org/10.5281/zenodo.5607692