

Project Title	Expanding FAIR solutions across EOSC
Project Acronym	FAIR-IMPACT
Grant Agreement No.	101057344
Start Date of Project	2022-06-01
Duration of Project	36 months
Project Website	https://fair-impact.eu/

M1.9 - Third synchronisation workshop

Work Package	WP1 - Project management, synchronisation and sustainability
Lead Author (Org)	Marita Everhardt (KNAW-DANS)
Contributing Author(s) (Org)	Anne Sofie Fink (DeiC), Clement Jonquet (INRAE), Esteban Gonzalez (UPM), Ingrid Dillo (KNAW-DANS), Josefine Nordling (CSC), Joy Davidson (DCC), Laurence Horton (DCC), Liisa Marjamaa-Mankinen (CSC), Maaïke Verburg (KNAW-DANS), Mike Priddy (KNAW-DANS), Nina Grau (INRAE), Olivier Rouchon (CNRS), Sara Pittonet Gaiarin (Trust-IT)
Due Date	2024-12-31
Date	2024-12-18
Version	V1.0
DOI	10.5281/zenodo.14537952

Dissemination Level

<input checked="" type="checkbox"/>	PU: Public
<input type="checkbox"/>	PP: Restricted to other programme participants (including the Commission)
<input type="checkbox"/>	RE: Restricted to a group specified by the consortium (including the Commission)
<input type="checkbox"/>	CO: Confidential, only for members of the consortium (including the Commission)

Versioning and contribution history

Version	Date	Author	Notes
0.1	2024.09.11	Marita Everhardt (KNAW-DANS)	TOC and V0.1
0.2	2024.11.07	Anne Sofie Fink (DeiC), Clement Jonquet (INRAE), Esteban Gonzalez (UPM), Ingrid Dillo (KNAW-DANS), Josefine Nordling (CSC), Joy Davidson (DCC), Laurence Horton (DCC), Liisa Marjamaa-Mankinen (CSC), Maaïke Verburg (KNAW-DANS), Marita Everhardt (KNAW-DANS), Mike Priddy (KNAW-DANS), Nina Grau (INRAE), Olivier Rouchon (CNRS), Sara Pittonet Gaiarin (Trust-IT)	Draft for concluding session
0.3	2024.11.27	Anne Sofie Fink (DeiC), Josefine Nordling (CSC), Maaïke Verburg (KNAW-DANS), Mike Priddy (KNAW-DANS), Nina Grau (INRAE)	Participants input processed
0.4	2024.12.13	Geta Mitrea (University Stefan cel Mare from Suceava), Fotis Mystakopoulos (OPERAS AISBL)	Proofreading
1.0	2024.12.18	Ingrid Dillo (KNAW-DANS), Sara Pittonet Gaiarin (Trust-IT)	Published version

Disclaimer

FAIR-IMPACT has received funding from the European Commission's Horizon Europe funding programme for research and innovation programme under the Grant Agreement no. 101057344. The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.



Table of Contents

1 Introduction	5
2 Description of the Milestone	7
2.1 Role of the Milestone	7
2.1.1 Means of verification	7
2.1.1.1 Proof of Milestone fulfilment as per the respective GA table	7
3 Highlights and recommendations from the Milestone	8
3.1 Metrics and assessing FAIRness	8
3.2 Sustainability of project outputs	12
3.3 Persistent identifiers	14
3.4 Legal & organisational interoperability	17
3.5 Trustworthy and FAIR-enabling repositories	20
3.6 Metadata, semantics and interoperability	23
4 Conclusions and next steps	28
5 Appendices	29
5.1 Underlying materials	29
5.2 Participant list	29

TERMINOLOGY

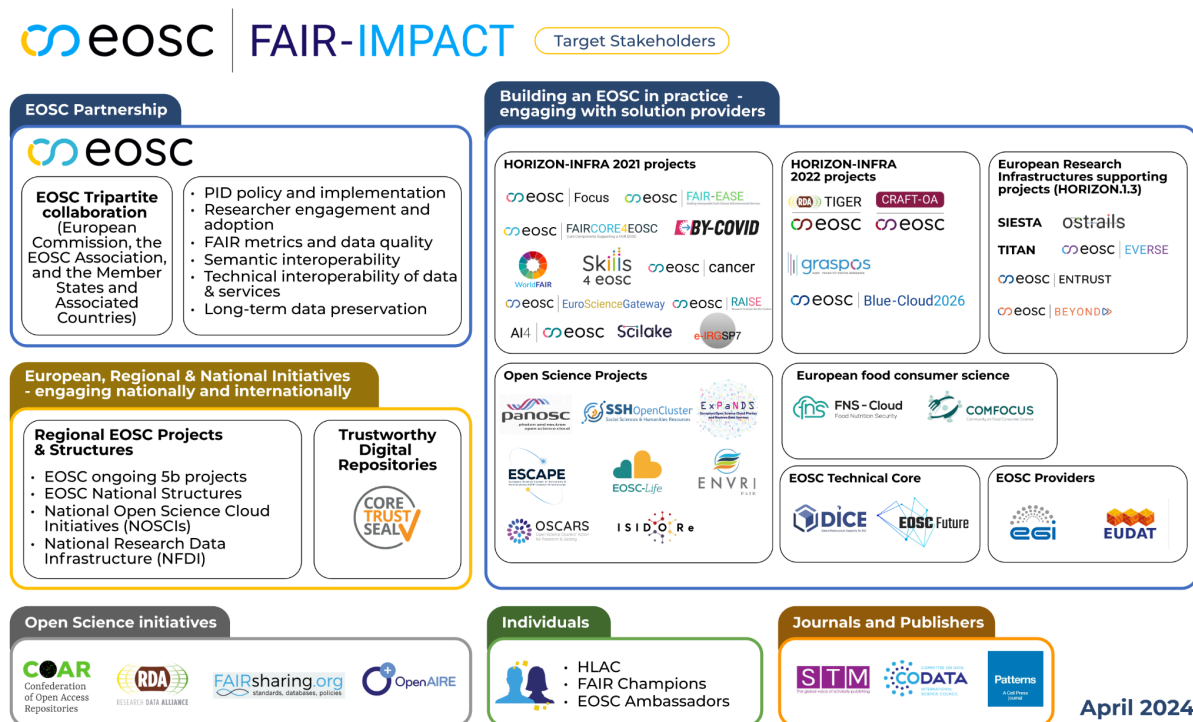
Terminology/Acronym	Description
API	Application Programming Interface
CAT	Compliance Assessment Toolkit
CoP	Community of Practice
EOSC	European Open Science Cloud
EOSC-IF	EOSC-Interoperability Framework
EOSC-OA	EOSC-Opportunity Area
EOSC-TF	EOSC-Task Force
FAIR	Findable, Accessible, Interoperable, Reusable
FIPs	FAIR Implementation Profiles
HE	Horizon Europe
KER	Key Exploitable Result
MAR	Multi-Annual Roadmap
MOD	Metadata for Ontology Description and Publication
MSCR	Metadata Schema and Crosswalk Registry
O'FAIRe	Ontology FAIRness Evaluator
PIDs	Persistent Identifiers
RDA	Research Data Alliance
SEP	Sustainable Exploitation Planning
SAC	Semantic Artefact Catalogues
SA	Semantic Artefact
WG	Working Group
WP	Work Package

1 Introduction

Extending the successful Synchronisation Force approach of the FAIRsFAIR¹ project, FAIR-IMPACT continues to foster dialogue to promote collaboration and harmonisation within the European Open Science Cloud (EOSC) and Findable, Accessible, Interoperable, Reusable (FAIR) ecosystems. This effort aims to minimise redundancy and to ensure that solutions are sustainable, widely accepted, and easily transferable to the EOSC Partnership, supporting current and future EOSC stakeholders in their uptake of FAIR-enabling practices.

To deal with these challenges, FAIR-IMPACT has established a Synchronisation Force with representatives from all work packages of the project. The main instrument is a series of three annual workshops organised between 2022 and 2024, each producing a brief report². Invited workshop participants include key stakeholders from the EOSC and FAIR ecosystems (see Image 1).

Image 1: FAIR-IMPACT's landscape of key stakeholders



April 2024

This landscape for synchronisation consists of the Board of Directors of the EOSC association and a selection of Task Forces under the EOSC association that are most relevant for the FAIR-IMPACT focus areas (top-left).

¹ FAIRsFAIR <https://fair-impact.eu/fairsfair-legacy>

² 2022 report <https://doi.org/10.5281/zenodo.7692063>

2023 report <https://doi.org/10.5281/zenodo.11082238>

FAIR is also in the remit of European projects, especially those in the HORIZON-INFRA-EOSC funding scheme (top, including 2021 and 2022 lines of funding). In addition to these, the most recently granted projects in support of European Research Infrastructures (Horizon 1.3 projects), as well as the Open Science projects (funded under the umbrella of the OSCARS project) and a few newcomers from the food sector, are also part of the FAIR landscape. The EOSC Technical Core projects and its discipline-independent providers complete the list of EOSC-related stakeholders (right-hand side).

Finally, representatives of regional and national initiatives and repositories (middle-left), Open Science initiatives (bottom-left), individuals and representatives of Journals and Publishers were invited to the 2024 workshop series.

Five topics that fit the FAIR-IMPACT core activity areas were defined to set the stage. Each topic focused on selected recommendations and ambitions from the Turning FAIR into Reality Report³ (2018), the EOSC interoperability framework⁴ (2021), FAIRsFAIR White Paper⁵ (2021), Strategic Research and Innovation Agenda⁶ (version 2022), EOSC Multi-Annual Roadmap⁷ (2023-2024), and the EOSC Multi-Annual Roadmap (MAR) 2025 and 2026-2027⁸. Additionally, with the project nearing its end, a sixth workshop was added to the series, covering the topic of sustainability.

Based on the workshop input and discussions, this report provides supporting recommendations for each topic.

³ Turning FAIR into Reality <https://doi.org/10.2777/1524>

⁴ EOSC Interoperability Framework <https://data.europa.eu/doi/10.2777/620649>

⁵ FAIRsFAIR White Paper <https://doi.org/10.5281/zenodo.5744786>

⁶ SRIA <https://data.europa.eu/doi/10.2777/935288>

⁷ EOSC Multi-Annual Roadmap https://eosc.eu/sites/default/files/2022-05/20220523_MAR_02_GL.pdf

⁸ EOSC MAR 2025 and 2026-2027 https://eosc.eu/sites/default/files/2023-01/MAR_2025-27_draft.pdf

2 Description of the Milestone

The Synchronisation Force workshops, annual series held between 2022 and 2024, brought together various stakeholders to periodically assess the progress in the FAIR-IMPACT focus areas. This Milestone concerns the 2024 workshop series.

2.1 Role of the Milestone

This Milestone demonstrates that FAIR-IMPACT, specifically Task 1.3 *Synchronisation with EOSC Partnership and relevant projects and initiatives*, organises an annual workshop series to monitor progress in FAIR practices across a wide range of EOSC initiatives and to disseminate such developments among the representatives of these initiatives.

2.1.1 Means of verification

Information about the 2024 Synchronisation Force workshop series is available via:

- The project website⁹, hosting the programme and slide decks;
- Zenodo (see Appendices for the links), containing supplementary materials such as the spreadsheet with input provided by the participants of the workshops.

The 2024 Synchronisation Force workshop series consisted of eight online sessions, held between 3 September 2024 and 7 November 2024:

- A kick-off session, introducing FAIR-IMPACT and the goal of the workshop as well as inviting participants to share details on their FAIR activities in preparation for the thematic sessions;
- A session on *Metrics and assessing FAIRness*;
- A session on *Sustainability of project outputs*;
- A session on *Persistent identifiers*;
- A session on *Legal & organisational interoperability*;
- A session on *Trustworthy and FAIR-enabling repositories*;
- A session on *Metadata, semantics and interoperability*;
- A concluding session, discussing highlights and recommendations from the thematic sessions with the participants of the workshops.

91 individuals registered for the workshop series; attendance in individual sessions ranged from 15 to 35 individuals.

2.1.1.1 Proof of Milestone fulfilment as per the respective GA table

This report serves as the verification method for achieving the Milestone.

⁹ FAIR-IMPACT

<https://fair-impact.eu/events/fair-impact-events/synchronisation-force-workshop-series-2024-edition>

3 Highlights and recommendations from the Milestone

3.1 Metrics and assessing FAIRness

Recommendations based on the Synchronisation Force workshop 2024 session

The recommendations discussed in the 2024 edition of this Synchronisation Force topical session were based on the results of the 2023 edition, reported in M1.8:

1. A broader community must be engaged and solicited for collaboration in developing a comprehensive **catalogue of methods**.
2. The development, sharing, discoverability, and reuse of **FAIR Implementation Profiles (FIPs)** necessitate enhanced community cooperation. This collaborative effort is crucial for supporting the creation of domain-specific assessments, ensuring that the principles of FAIR are effectively implemented and tailored to specific research needs.
3. Conducting **pre-assessment** could facilitate the gradual enhancement of the FAIRness of research outputs throughout the research data lifecycle.
4. Researchers require access to **dedicated local support** to assist the general and discipline-specific FAIRification processes of data, research software, and semantic artefacts.

The session was designed to assess the current status of the actively developing landscape on these points, and to critically discuss and reflect on the next steps that should be taken in the community. Over 40 participants joined the session and contributed to a lively discussion, representing different perspectives from various European projects and initiatives, scientific disciplines, and areas of expertise. This showed that FAIR metrics and assessment continue to be an important element of the scientific landscape to discuss and advance. This is also demonstrated in new projects that have started and now joined the Synchronisation Force workshops, for example OS Trails¹⁰, focused on advancing research assessment specifically.

Mari Kleemola (Finnish Social Science Data Archive, CESSDA) initiated the first discussion focused on the topic of ‘Supporting researchers in the FAIRification process’ (Related to recommendations #3 and #4). Sharing experiences from the institutional and community-based research data management (RDM) support that is provided, it was highlighted that the complexity of the landscape and the many involved stakeholders raises the question of division of labour when it comes to FAIR adoption and assessment. How much should researchers be expected to know about FAIR and the FAIRification of their digital objects? And how should they navigate the landscape of available support and resources on this topic? Audience discussion indicated agreement on the idea that **researchers are responsible for the good documentation of their resources**, but that it is important they receive **intensive professional support** on this. This support should be available on different levels, such as nationally and thematically, but it is deemed important

¹⁰ <https://ostrails.eu/>

that it starts “at home”, i.e. as close to the researcher as possible. It is important the expected effort from researchers is communicated in terms that mean the most to them. FAIR should not be presented as an end goal, but instead researchers could be motivated to document their data because it will ease future reuse by themselves or others, lead to increased citations, or indicate research integrity. Creating a **platform of publicly available success stories** on how implementing FAIR leads to advantages and successes were discussed as a useful means to increase importance and prioritisation of RDM in science. Such stories could include experiences shared by researchers on advantages of good data management, the perspective of institutions or communities that choose to invest in this, but also examples of data that follow good practices without being created to be FAIR, or stories outside of the EOSC or academic research context (e.g. OpenStreetMap¹¹), to highlight common advantages of this. Another way to prioritise research support would be **more mandates by funders or (inter)national bodies to invest in support staff**, ideally because it will lead to increased investments in, for example, hiring and professionalisation of support staff. The *Realising the European Open Science Cloud* report suggests that “well budgeted data stewardship plans should be made mandatory and we expect that on average about 5% of research expenditure should be spent on properly managing and stewarding data.”¹² Researchers can further be supported by the creation of **more automated systems** for the capturing of required information, so information only has to be provided in one place and can then be used in all the relevant other locations, instead of requesting the same information to be provided repeatedly (for example, machine actionable Data Management Plans (DMP)).

The second topic discussed during the session was “Discipline-specific metrics: the process and challenges” (related to recommendation #2), presented by conversation starter Robert Huber (University of Bremen, F-UJI lead developer, FAIR-IMPACT T5.1 lead). Community-specific metrics are challenging due to the diversity in FAIR practices and existing communities. Developing metric sets may lead to **further fragmentation of FAIR**. There are risks associated with diversifying FAIR metrics, not in the least the user frustration when different metrics and tools lead to widely different results. Discussion on the topic by the audience confirmed some of the experiences from FAIR-IMPACT and the challenges of discipline-specificity. One of the challenges stemming from this is that the implementation of FAIR and the interpretation of FAIR assessments is mostly done by humans, which brings along anthropological challenges. Discipline-specific standards may be agreed upon, but might not be actually adhered to in practice. Discipline-specific assessment tools may be more accurate, but also may lead to lower scores due to their strictness, causing frustration and possibly a user preference for the generic tools that present higher scores. Moreover, navigating a field of many different metrics and tools is currently already challenging for users. With further fragmentation, this will only increase, and multidisciplinary research will not easily find the right metrics for their assessment. Creating metrics is easy, but endorsement, uptake, and convergence is hard. And even with the creation of metrics, it is

¹¹ <https://www.openstreetmap.org/>

¹² Mons, B. *et al.* (2016) *Realising the European Open Science Cloud*, First report and recommendations of the Commission High Level Expert Group on the European Open Science Cloud, DOI 10.2777/940154, <https://op.europa.eu/en/publication-detail/-/publication/2ec2eced-9ac5-11e6-868c-01aa75ed71a1>, p17

important to **be aware about the biases** that are part of them. By clearly presenting biases in metrics, developers can reduce overconfidence and help shift focus to the assessment tool as an assistance tool.¹³ Similarly to trustworthiness and certification, FAIRness and assessment should be seen as a journey, not an all-or-nothing principle. By evaluating assessment results, or through the process of creating a FAIR Implementation Profile (FIP), users can already learn so much about FAIR and its implementation. FAIR should not be seen as the end goal but instead a means to reach other goals (as mentioned in the first discussion already), as the original principles were also deemed ‘guiding’ for this reason. The development of FIPs has been supported by projects such as WorldFAIR¹⁴, WorldFAIR+¹⁵, and initiatives like FAIRConnect¹⁶. New insights that have been gathered by organisations looking to create a FIP and those looking to evaluate it to identify community standards show that, again, the human component of this work presents challenges. The **quality of a FIP** depends greatly on the knowledge of the persons that create it. Moreover, organisational knowledge and resources are not always stable, as expertise leaves the organisation when employees leave. However, a FIP should be the outcome of the decision by a FAIR Implementation Community, be that within a discipline, research infrastructure, or perhaps a smaller community of users. There may be room for more education on how to create a useful FIP and how to care for it in the longer term. An advancement to FIPs based on this is the addition of the attribute ‘status’ to indicate whether or not a resource is approved by a specified community. Work is also being done to allow DMP creation to be based on existing FIPs, ensuring that the claims made in the FIP are reflected in the daily data management practices. A last part of the discussion on this topic related to **data quality** and the fact that FAIR assessment tools are often not equipped to evaluate this. Data quality, in the sense of ‘fitness for purpose’, is important for daily research practice and its value, but it is not a perspective addressed by FAIR. A dataset can have objectively poor quality but it may be essential to answering a particular research question, where the issues with the data object are known and managed. To consider data quality as part of FAIR assessment may be problematic as it can be subjective and biased. Therefore, it remains difficult to judge whether FAIR assessment tools should be expanded to include data quality, but some argue it is not the role of such tools to verify content.

The third and final topic of discussion, “Methods and tests: do we need a catalogue of these?” (related to recommendation #1) was introduced by conversation starter Daniel Garijo (Technical University Madrid (UPM), FOOPS!¹⁷ developer, FAIR-IMPACT T5.3 lead). Building on earlier discussions, this topic also brought up the topics of harmonisation, the **FAIRness of FAIR metrics** themselves, and how metrics and tests could be catalogued. As was brought up in the 2023 round of the Synchronisation Force workshop, there is a need to gain insight into FAIR metrics and tools. Efforts like FAIRassist¹⁸ are valuable, but due to the

¹³ Grootveld, M., Pittonet Gaiarin, S., Davidson, J., Dillo, I., O'Connor, R., Marjamaa-Mankinen, L., Verburg, M., & Jonquet, C. (2023). M1.7 - First synchronisation workshop. Zenodo. <https://doi.org/10.5281/zenodo.7692063>

¹⁴ <https://worldfair-project.eu/>

¹⁵ <https://worldfair-project.eu/worldfair-plus/>

¹⁶ <https://fairconnect.pro/>

¹⁷ <https://catalogue.fair-impact.eu/resources/foops>

¹⁸ <https://fairassist.org/#/>

manual entry and limited information, it does not offer a comprehensive insight. There is a great need to have metadata that is clear, transparent, and the same across different metrics and tools, making these objects themselves more FAIR. Having such information about metrics and tools can showcase the biases that are part of them (as also mentioned in the previous discussion topic), and can allow users to then make their own **informed decision about trusting and using the metrics and tools**. It can also help users navigate which tool(s) to use and how to interpret the results, based on what exactly is being tested and prioritised. This convergence of the topics of FAIR and trustworthiness through transparency aligns with the work being done in, for example, FAIR-IMPACT and OSTrills. The session's discussion then seemed to come full circle with a collection of considerations around the **division of labour** between researchers and repositories when it comes to FAIR assessment. Many automated FAIR assessment tools have the possibility of integration into a repository, leading to the question whether repositories should then not take more responsibility away from researchers in making digital objects more FAIR. The most realistic solution will be a co-responsibility between both stakeholders, as researchers remain vital in providing the documentation of their work and repositories have a position to offer tools and services to check compliance and return feedback. However, realistically, many repositories currently do not have the resources to take on additional responsibilities and would need increased investment to facilitate this.

Building upon the insights gained from the discussions in this topical session, new recommendations have been distilled to focus the next steps in the field of FAIR metrics and assessment:

- A platform of publicly available stories, experiences, and lessons learned related to the implementation of FAIR and associated benefits could improve the awareness and uptake of FAIR in the research community.
- Researchers should be supported throughout the research lifecycle by a cohesive network of supporters making use of increasingly automated systems, where possible, for capturing the necessary metadata and documentation about their digital objects to lessen the burden of this responsibility.
- Community-specificity is an essential part of FAIR, but also brings along increased fragmentation in the field of metrics and tools, as well as the associated risks. Therefore, community-lead recommendations and governance for tests, metrics, and tools would aid trust and transparency in the process of improvement and assessment of FAIRness.
- The development of a comprehensive catalogue of methods and tests should be considered by communities large and small, presenting metrics and tools with their associated metadata to give insight into underlying biases (linked to the tools), foster a sense of trust, and emphasise assistance as the main goal¹⁹.

¹⁹ Furthermore, sharing tests through a catalogue would aid and facilitate the creation of discipline-specific metrics for communities who wish to adapt or create FAIR assessment tools to meet their particular requirements.

3.2 Sustainability of project outputs

This topic was a new addition to the lineup for the Synchronisation Force this year. Planning for the long-term sustainability of project outcomes in an ever-changing landscape is a common challenge for many projects. The exploitation of outputs after a project ends depends partly on their uptake and adoption by the various stakeholders of a project, yet there is currently no consensus on how to best approach this. The Synchronisation Force as a mechanism in itself will also face these challenges of sustainability, running the risk of not being continued in the upcoming years. Given the shared nature of this topic and the associated challenges, it was decided to host a dedicated topical session to bring together different projects to discuss this together with some important stakeholders in the landscape.

Around 40 participants joined the session to discuss the topic of sustainability. Short presentations were given by Blagovesta Cholova (FAIR-IMPACT Project Officer), Kathrin Winkler (European Commission), and Illaria Nardello (EOSC-Association). These presentations discussed the different types of sustainability, the requirements and support mechanisms available towards sustainability, and next steps already on the radar to improve this.

Some initiatives for support towards (topics related to) sustainability were presented and discussed. The EOSC Implementation Macro Roadmap²⁰ catalogues all EU projects' and EOSC Association Task Forces' outputs to increase their findability. In a similar effort, the EU Open Research Repository pilot aims to provide a catalogue of research outputs from Horizon Europe (HE), Euratom and earlier Framework Programmes. Attention to cascading grants has also been increased with sustainability in mind, to facilitate that solutions coming from a project are used and improved by the community. Through the EOSC Partnership and Tripartite Governance, networking activities like workshops, conferences, winter schools, and other coordination meetings have been organised, which aim to enhance sustainability opportunities between projects. The HE Project Impact Working Group from the EOSC Association has worked on the Sustainable Exploitation Planning (SEP) tool to capture Key Exploitable Results (KERs). After initial testing with some projects, next steps will include a better contextualisation of the methodology steps and application rules, a more visual design, and the introduction of automation to simplify use and adoption. The Horizon Booster²¹, an initiative by the European Commission, provides tailored services to boost the dissemination and exploitation of research results. RDA TIGER reported being in the first stages of applying for this type of support, but could not yet share any specific experiences.

There is general agreement on the idea that the planning towards sustainability should start early in a project's lifetime, involve all partners, and be dynamic and flexible as the project advances. Each project also presents a unique collection of outputs (services, tools, policies, recommendations, guidelines, use cases, etc.) in a unique context and therefore requires a

²⁰ <https://eosc.eu/eosc-macro-roadmap/>

²¹ <https://www.horizonresultsbooster.eu/>

tailored sustainability plan. However, there are common challenges and opportunities that all projects should be supported in. Assigning responsibilities, discussing data sovereignty, and connecting to the right communities and stakeholders are big challenges to overcome when considering sustainability, and a project is not a legal entity where such things can be more easily agreed.

The EOSC federation has a crucial role for the HE INFRAEOSC projects' sustainability pathways. However, while many projects work to create solutions for EOSC or advance EOSC in some way, EOSC still remains to a certain extent a moving target and is currently not available for the integration, endorsement or uptake of these solutions. This forces the projects to consider other (temporary) sustainability solutions, and runs the risk of work not being sustained at all and never contributing to EOSC in its intended way. Shifts in the environment are very challenging for projects to respond to within the limits of their agreed work. For example, the fact there is no longer a dedicated EOSC Association Task Force on the topic of Persistent Identifiers (PIDs), as a body to endorse PID-related policies, was not expected and causes sustainability issues for the work developed in this field. The more security and clarity projects can get on the (near) future from the steering bodies like the European Commission and EOSC Association, the more they can navigate and try to be flexible with their solutions. It is also important to note that community uptake is not easily achieved without European Union and/or EOSC endorsement. This should be kept in mind when considering alternative sustainability options. Work is ongoing to create an EOSC Federation Handbook, which documents the decisions taken by the EOSC Tripartite Governance on the EOSC Federation and will serve as the reference document for the operation of the EOSC Federation²². Current unclarities about the progression and opportunities for open consultation remain, but this handbook could also include sustainability as a topic.

Based on the experiences of a collection of previous projects, the idea was raised to capture recurring challenges, possible solutions, and other recommendations or processes in a sustainability handbook. Networking events like the EOSC Winter School²³ could be places to discuss and collaboratively create such an output.

Taking together the insights from this topical session, several recommendations have been created on the topic of sustainability of project outputs:

- The EOSC Association should consider incorporating and/or referencing relevant outputs from EOSC-related projects in (future versions of) the EOSC Federation Handbook;
- The EOSC Association and the EOSC-related projects should start a dialogue to identify potential sustainability pathways for project outputs (possibly during the upcoming Winter School 2025);

²² <https://eosc.eu/eosc-federation-handbook/>

²³ <https://eosc.eu/events/eosc-winter-school-2025/>

- A Sustainability Handbook should be created to capture the common challenges, possible solutions, and experiences from previous projects in a way that provides useful references for new projects considering sustainability.

3.3 Persistent identifiers

This session took the approach of looking forward in time, considering the crucial activities related to persistent identifiers (PIDs) set out for the years 2025-2027 within EOSC as outlined in the Multi-Annual Roadmap (MAR)²⁴:

- Promote and sustain the use of Persistent Identifiers (PIDs) that are already common practice. Support activities where PID usage is not yet a common practice.
- Integrate widely used and adopted PIDs into institutional services and incentivise usage of PID technologies being developed in EOSC (like **PID Meta Resolver**²⁵, **Data Type Registry**²⁶, **PID graph**²⁷, **PID Policy Compliance Assessment Toolkit**²⁸).

Another topic for discussion was collaboration opportunities in a changing EOSC environment.

The main aim of this session was to understand what is needed to realise an inter-connected coordination among PID actors and to ensure sustainability of PIDs. A key building block of sustainable PIDs is the implementation of comprehensive and well-planned PID policies, so the session also delved into characteristics of PID policies and focused particularly on some exit strategies.

The session was attended by approximately 30 participants and was kicked-off by four presentations to spur the discussions; Tibor Kalman from GWDG on PID collaboration opportunities and sustainability of PIDs, Wim Hugo from DANS on creating EOSC compliant PID policies, Josefine Nordling from CSC - IT Center for Science on the future of the EOSC PID Policy, and Gabriela Mejias from DataCite on the needs of a collaboration mechanism for EOSC PID Service Providers.

Building sustainability and collaboration structures around PIDs is a multi-layer issue that requires the inclusion of several stakeholders. Currently, there is an uncertainty within EOSC, due to discontinuation of the Task Force on PIDs which was replaced by an Opportunity Area Expert Group on PIDs (OA-1) consisting of project representatives, but the mandate in terms of governance is based on loose terms. Addressing this issue is a work in progress and the OA-1 is aiming for a more prominent role by pointing to particular tasks of urgency when collaborating on PID efforts, with a concrete current assignment on updating

²⁴ Multi-Annual Roadmap (MAR) 1.2, which constitutes Section 8 of the EOSC Strategic Research and Innovation Agenda (SRIA). <https://eosc.eu/sria-mar/>

²⁵ <https://faircore4eosc.eu/eosc-core-components/eosc-pid-meta-resolver-pidmr>

²⁶ <https://faircore4eosc.eu/eosc-core-components/eosc-data-type-registry-dtr>

²⁷ <https://faircore4eosc.eu/eosc-core-components/eosc-pid-graph-pid-graph>

²⁸ <https://faircore4eosc.eu/eosc-core-components/compliance-assessment-toolkit-cat>

the EOSC PID Policy²⁹. Collaboration between the PID service providers within the EOSC Federation and related Nodes is a complex task where the PID services need to be interoperable with clearly defined responsibilities of service maintenance and provision. Also collaboration with actors outside of EOSC (e.g., RDA³⁰, FDO Forum³¹, ePIC³²) is crucial in order to contribute successfully to the PID ecosystem. Currently, PID services rely on varying technologies and associated policies, leading to diverse contractual agreements that create significant challenges for overall interoperability. The key outcome of this discussion topic is to **seek collaboration instead of sustaining the problems**.

- Promote and sustain the use of Persistent Identifiers (PIDs) that are already common practice. Support activities where PID usage is not yet a common practice.

Successfully implementing PIDs across the expanding EOSC community requires a well-defined coordination mechanism between EOSC and PID service providers. This will help to effectively convey the needs of the EOSC community, such as promoting existing PID practices and raising awareness of emerging PIDs for future consideration. **A lack of coordination on sustainability and cost models related to PIDs was predicted to have undesirable consequences on the broader European research community.** The probable scenario is continued inequalities across countries and research disciplines when it comes to access to operational and trustable identifier services, which would build barriers to the FAIRification of European research assets.

- Integrate widely used and adopted PIDs into institutional services and incentivise usage of PID technologies being developed in EOSC (like PID Meta Resolver, Data Type Registry, PID graph, PID Policy Compliance Assessment Toolkit).

There should be a means to enable interoperability across both existing PID services and those in development, such as the Data Type Registry, PID Meta Resolver, and PID Policy Compliance Assessment Toolkit. This will require ongoing collaboration across service providers, building on the work already undertaken within EOSC and beyond. A key focus for the near future, as highlighted in the EOSC Interoperability Framework report³³, is the need for greater access to highly scalable PID services to support the provisioning of FAIR Digital Objects (FDOs)^{34,35}.

²⁹ European Commission: Directorate-General for Research and Innovation, Hellström, M., Heughebaert, A., Kotarski, R., Manghi, P. et al., A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC), Publications Office, 2020, <https://data.europa.eu/doi/10.2777/926037>

³⁰ <https://www.rd-alliance.org/>

³¹ <https://fairdo.org/>

³² <https://www.pidconsortium.net/>

³³ New European Interoperability Framework at the website of the European Commission, ISA2 programme". 16 February 2017. Retrieved 14 April 2017. https://ec.europa.eu/isa2/eif_en/

³⁴ Ivonne, A., Christophe, B., Daan, B., Maggie, H., Sharif, I., Thomas, J., Larry, L., Karsten, P.-. von G., Robert, Q., Alexander, S., Ulrich, S., Stian, S.-R., Strawn, G., Dieter, . van U., Claus, W., Peter, W., & Carlo, Z. (2023). FDO Forum FDO Requirement Specifications. <https://doi.org/10.5281/zenodo.7782262>

³⁵ FAIR Digital Object Framework Documentation: <https://fairdigitalobjectframework.org/>

The PID ecosystem consists of the supply and demand sides. Developing a PID policy from the end user perspective, requires consideration of the supply side of the stack. 16 guidelines for PID Managers on creating a user friendly EOSC compliant PID policy³⁶ have been identified to date within the FAIR-IMPACT project, which are linked to features, characteristics and attributes of PID Stacks, as well as to the EOSC PID Policy. This work was influenced by the Principles of Open Science Infrastructure (POSI)³⁷, various national, institutional and thematic PID/data policies, published use of PIDs in workflows and specific use cases, and RDA outputs. The Compliance Assessment Toolkit (CAT) assesses compliance with the EOSC PID Policy, which helps in harmonising the PID implementations and usages across EOSC. According to the EOSC PID Policy, the PID Service Provider is to support versioning and have clear versioning policies in place, and by implication that is to be made for both kernel metadata as well as for the object or concept being referenced by the PID. Furthermore, it is the responsibility of PID service providers to provide a feedback mechanism for users of the system to ensure a well-functioning PID infrastructure. It is crucial to take the **end user perspective** when identifying what is important to them in terms of **guaranteeing persistent resolvability** at various time scales.

The P in PIDs was further explored through a discussion on exit strategies where e.g. DataCite³⁸ shared their best practice recommendations on “tombstones” as part of their non-deletion policy for DOIs. In rare instances, DataCite DOIs may be made publicly unavailable, such as in cases of research retraction. In all cases where the object the DOI is pointing to ceases to exist, the best practice is to provide a so-called tombstone : a landing page describing the item that has been removed. The organisation responsible for maintaining the DOI (the PID Manager) is also responsible for maintaining the tombstone page. The topic of persistence alongside resolvability, particularly in the context of landing page archiving efforts (e.g., using the Internet Archive), requires further discussion.

This session generated the following recommendations:

- There is a risk of fragmenting the PID service landscape even further due to several different providers having slightly different approaches affiliated with different domains using different schemes. The current EOSC PID Policy does not address these kinds of issues. Building sustainability and collaboration structures around PIDs in the context of the EOSC Federation and beyond is a multi-layer issue that requires the inclusion of several stakeholders.
- There is a call for convergence, as the current PID services rely on varying technologies and associated policies, leading to diverse contractual agreements that create significant challenges for overall interoperability.
- The issues and risks within the PID landscape are increasingly to be addressed from a social perspective, rather than solely a technical one, with the support and endorsement of the research community.

³⁶ van Horik, R., & Hugo, W. (2024). D3.3 - Guidelines for creating a user tailored EOSC Compliant PID Policy (V1.0 DRAFT NOT YET APPROVED BY THE EUROPEAN COMMISSION). Zenodo. <https://doi.org/10.5281/zenodo.11354246>

³⁷ <https://openscholarlyinfrastructure.org/about/>

³⁸ <https://support.datacite.org/docs/tombstone-pages>

- The end user involvement in ensuring sustainability, persistence and resolvability of PIDs is crucial.

3.4 Legal & organisational interoperability

For 2023 the topic for the workshop was to address the status and adoption of the legal and organisational recommendations presented by the EOSC Interoperability Framework³⁹ in different scientific domains. Moving on from this topic to a topic for 2024 was to collect feedback from concrete implementation of legal and organisational policies in data repositories, projects, and organisations.

The recommendations discussed in the 2024 edition of this Synchronisation Force topical session were based on the results of the 2023 edition, reported in M1.8:

1. EOSC and other relevant entities should advocate for Creative Commons (CC) licences unless another licence or licence family is predominant within a specific research domain or community. This aligns with the EOSC Interoperability Framework's support for permissive licences. "A list of EOSC-recommended licences and their compatibility with Member States' recommended licences should be provided."⁴⁰
2. Data creators and users should be shielded from the complexities of licence impacts, necessitating both harmonisation and comprehensive guidance potentially provided by local or domain-specific data stewards. EOSC is encouraged to take an active role in this harmonisation effort.
3. An integrated support programme for managing, protecting and licensing data is recommended for research-performing organisations.

For the 2024 workshop session 26 participants were present of which 20 participants were involved in various EU projects related to EOSC, and four were members of EOSC task forces.

For the workshop a panel was set up to address legal and organisational interoperability from different perspectives.

The workshop panel:

- Jessica Parland-von Essen, CSC, Finland
- Ohad Graber-Soudry, X-officio
- Lise Schroder, AquaINFRA

³⁹ EOSC Interoperability Framework

<https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en>

⁴⁰ EOSC Interoperability Framework

<https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en> (p. 24).

- Renato Juaçaba EMBL/EBI, BY-COVID, EOSC-Entrust
- Simon Hodson, CODATA, WorldFAIR

As the national NREN and the main service provider in Finland, CSC actively coordinates Open Science efforts. Copyright & IPR issues are relatively simple to address (CC licences recommended). However, data ownership issues are more difficult to deal with (organisationally) such as copying, sharing, access rights, etc. At the moment work at national level on contracts (data ownership contract templates) is ongoing. Also, centralised metadata services are in place and could meet needs provided that the metadata quality is high. In Finland, rights management is debated a lot in relation to the [Open Science and Reference architecture](#). The OS coordination has also produced a lot of [guidelines](#) (unfortunately some only in Finnish) with relation to legal and interoperability.

Going back to the report [Legal Interoperability and the FAIR Data Principles \(2021\)](#) focusing on legal and governance consultancy incl. When working closely with research infrastructures and research performing organisations it is tried to promote the same understanding within (distributed) RIs. In principle, factual data is not subject to copyright, however if research objects contain a creative element (for example, drawings, pictures, voice recording, etc), they may be subject to copyright. IPR policies including licences are formulated to facilitate the 'R' from FAIR. Additionally, interoperability of licenses is also important.

From a view from an EOSC project which is combining two different research domains (marine, inland water) with very different restrictions (inside country and across borders). The project is in the early phases, objective to have seamless access to the data in a distributed environment, but also to the different layers of licensing.

The challenges on copyright, IPR, and licences have been addressed in WorldFAIR, with a focus on data protection & privacy. The approach has been to automate and maximise access to sensitive and potentially disclosure data while respecting the necessary rights and protections. Three use cases have been identified to address them:

- Federated data analysis / data visiting used in projects / case studies across 4 african population health research centres (requires very detailed data description required).
- Fine-grained granular access to less disclosive variables, i.e. datasets with only few columns which are problematic (enable access to a modified version of the dataset, careful of potential triangulations). Discussed in CDIF⁴¹ document and WorldFAIR policy brief.⁴²
- Automating access permissions and accreditation.

The technical requirements were as follow:

⁴¹ <https://zenodo.org/records/11236871>

⁴² <https://zenodo.org/records/14236140>

- Clear, machine actionable articulations of what can and cannot be done with the data (lots of manual labour assessing the researcher accreditation to access data).
- Combined with data description at the variable level, including estimates of discursiveness risk.

Challenges for legal interoperability are highly dependent on the resources or projects. Each of them may have different answers to address them: license, data protection description, etc. For example Uniprot (repository of proteins) has its own license - based on CC. BY-COVID gathers partners with different policies which is an issue when trying to improve discoverability of datasets.

The DCAT-standard is proposed by the BY-COVID project as the standard format to expose resources/data, in addition to a custom XML schema. There have also been investigations around ODRL; it requires a lot of effort to learn, as none of the partners have sufficient expertise. Having said that, it remains a promising option. XML is mainly used and mappings can be described with the FAIRCore4EOSC [MSCR](#) tool.

CDIF recommends DCAT (JSON-LD) for discoverability purposes, and ODRL for expression of rights. CDIF adds recommendations for a set of controlled vocabularies:

- DDI-CDI⁴³ is used for description of variables - necessary for the detailed variable level data description for data visiting and fine-grained access.
- ODRL⁴⁴ - current assessment is that it will need extensions, additional vocabularies.
- DPV⁴⁵ - ditto
- schema.org/JSON-LD - recommended approach from a number of the WF case studies.

Multiple standards address particular challenges, functions, use cases. CDIF looks at those multiple functional requirements, and tries to identify where a standard is already being used, and recommend it (framework of multiple standards). In Finland, national guidelines were published in 2017 which aimed at machine-actionable expression of basis for all restrictions, see [metadata hub](#) for Finland (DCAT centred).

We are facing a huge challenge in legal interoperability as legal activity is interpretation of texts/laws (i.e. what something means). Lawyers don't always agree on what is the correct interpretation and in cases such as vocabularies or ontologies they are always ephemera and must be continuously reviewed (e.g. PIDs) due to the drifting of the concepts.

⁴³ <https://ddialliance.org/Specification/ddi-cdi>

⁴⁴ <https://www.w3.org/TR/odrl-model/>

⁴⁵ <https://w3c.github.io/dpv/2.0/dpv/>

There are different approaches to data interoperability in general e.g. for FAIRCORE4EOSC and CDIF approaches. 1-1 mapping is obviously useful. Additional to this, CDIF enables the use of multiple schemas/vocabularies together, which makes it easier and more powerful.

It can be difficult to articulate in a clear human-readable way what the legal rights are. A clear understanding of the constraints is needed, as well as collaboration with technicians for the encoding of these for machine actionability. Very good examples of interoperability general guidelines and the activation of this articulation may be found in the [linux master source code](#) and the example of [SDPX](#).

EU level governance and guidelines, especially vocabularies with interoperable and clear definitions (e.g. what is research use or who is a researcher etc) would help to overcome the challenges for legal interoperability. Additionally, guidelines on how to expose data (from a technical point of view) were missed as well as assessment tools needed to verify that the checklist/guideline have been implemented correctly.

Research Infrastructures, specifically those that distribute data, face particular challenges due to their fragmented nature cross cutting e.g. institutional setups, legal frameworks and/or national borders. They have limited capacity and are very much dependent on their nodes, service providers, etc. who are separate legal entities struggling to agree on principles and policies, weaknesses inherent to the distributed organisation structure.

To challenge this line of thought it was put out that methodology is more important than specific support. We need to be more specific, precise in the challenges we are trying to address and then step back and make sure we don't solve niche problems - how can we maximise access and usability to data?

Achieving a proper description of legal constraints is a big challenge to move interoperability forward - data creators (and curators!) use arguments like "it is unclear what license to apply" or "what if someone will use 'my' data for non-intended purposes" as an excuse to not invest in making their data (or even related metadata) FAIR or open, even though there may be no legal hindrance.

Recommendations based on the Synchronisation Force workshop 2024 session

- Training for those involved in RIs and service/data providers with (EU) central hubs to bring everyone to the same level of understanding - even basic level - is required e.g. about knowledge of IPR and GDPR.
- EOSC-related platform/contact point for advice on legal (organisational) interoperability-related documents - knowledge base, reference resources, good examples, etc. - for basic guidelines. Easy access to expertise and expert contact.

- Continued collaboration on legal and organisational interoperability for future projects such as FIDELIS⁴⁶, WorldFAIR+⁴⁷, and Climate-ADAPT4EOSC⁴⁸.
- Capacity building for RIs to have them leverage the expertise needed to develop the machine-actionable descriptions of what can/cannot be done - Use Case approach.

3.5 Trustworthy and FAIR-enabling repositories

Recommendations based on the Synchronisation Force workshop 2024 session

The recommendations discussed in the 2024 edition of this Synchronisation Force session were based on the results of the 2023 edition, reported in M1.8:

1. Increasing process transparency is vital to effectively assess repository trustworthiness, as encapsulated in the principle of **"Trust through transparency."**
2. Repository support should include **generic solutions** for widespread applicability and **detailed, customised support** for specific local or individual needs.
3. Due to their evident advantages, **support networks** for repositories at all levels (thematic, national, and international) should be established.

Last year's edition of this session discussed how transparency and trustworthiness relate, and how these qualities can be improved within data repositories through certification but also in other ways. This year's session focused on the support needed to improve transparency. Around 35 participants joined the discussion, including those offering such support and those who received it, to bring those two perspectives closer together.

To begin with, Socrates Varakliotis (UCL) shared his experience with the support received during the FAIR-IMPACT support action *'Recommendations for trustworthy and FAIR-enabling data repositories'*⁴⁹. He highlighted that **building a Community of Practice (CoP)**⁵⁰ around CoreTrustSeal, or more widely, certification or trustworthiness, is an essential type of support that is needed to advance in this area. Some experiences on building CoPs was shared, for example the initiative of RDA France⁵¹, but it was emphasised that such networks should be not only local and national, but more importantly European or even

⁴⁶

<https://www.horizon-europe.gouv.fr/enabling-network-eosc-federated-and-trustworthy-repositories-and-enhancing-framework-generic-and>

⁴⁷ <https://worldfair-project.eu/worldfair-plus/>

⁴⁸ <https://climate-adapt.eea.europa.eu/en>

⁴⁹ <https://fair-impact.eu/support-offer-3-recommendations-trustworthy-and-fair-enabling-data-repositories>

⁵⁰ Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511803932>

⁵¹ <https://www.ouvrirlascience.fr/certification-of-data-repositories-and-services/>

global. There are already some existing RDA working groups related to this topic^{52, 53, 54}, and it is important that these align with other projects and initiatives working towards the same goals. A challenge with CoPs is that they usually rely solely on volunteer effort, leading to risks with regards to sustainability over time and potentially an unfair division of time and effort. It is important that a community of practice is established carefully, and that roles, responsibilities, and common processes are developed with input and support from all involved members. Defining the area(s) of focus and the initial members of the community is also important to ensure that such groups are inclusive in nature. For example, for a community of trustworthy repositories, it must be clear whether the CoP includes certified repositories or also repositories that display trustworthiness in a different way, and whether it also welcomes those interested but not yet actively involved in the topic.

The second recommendation was introduced by Marine Vernet (IFREMER), who spoke about the mix of generic and customised support they received from the RDA France Certification of data repositories and services working group⁵⁵. The main lessons learned here were that certification, or more broadly working towards improved trustworthiness, takes a lot of time and effort. Engaging with peers and using available resources can be helpful, but in itself requires additional time and effort. However, participants of the support programme felt that the time and effort needed to work on trustworthiness did lead to direct benefits and rewards- even when the end goal is not to submit an application for certification. Tangible benefits included clarification of data workflows, consideration of all repository processes, and highlighting areas where improvements are needed. In addition, participation in the support programme fostered collaboration, both internally with the different areas of their own organisations and externally with other data centres. Taken together, the large efforts that are put in are met with valuable benefits and rewards.

When it comes to certification and trustworthiness schemes, there is a difference between discipline-agnostic and discipline-specific approaches. The advantage of the former is that it makes it broadly applicable and suitable for inclusion in national plans, which has happened in France. The advantage of the latter is it that it can be more tailored to reflect specific community needs. However, not all communities are organised in a way that allows them to develop and establish frameworks for discipline-specific solutions. This aligns with ongoing discussions in the topical session on metrics and assessing FAIRness, where it is also proven very challenging to move between harmonisation and fragmentation.

The last discussion was led by Kathleen Shearer (COAR), and Mari Kleemola (FSD). First, Kathleen Shearer (COAR) spoke about the challenges and experiences of repositories observed in a recent joint initiative undertaken with OpenAIRE, LIBER, and SPARC Europe to

⁵² <https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig/>

⁵³ <https://www.rd-alliance.org/groups/rdawds-trust-principles-outreach-and-adoption-working-group/>

⁵⁴ <https://www.rd-alliance.org/groups/community-based-catalogue-requirements-trustworthy-technical-repository-service-providers/>

⁵⁵ <https://www.ouvri.lascience.fr/certification-of-data-repositories-and-services/>

strengthen the European repository network⁵⁶. The initiative began with a survey of the European repository landscape which was undertaken in early 2023 and identified three main challenges: trying to maintain up-to-date, highly functional software platforms; applying consistent and comprehensive good practice in metadata, preservation and usage statistics; and achieving appropriate visibility of repositories in the scholarly ecosystem. Three areas of work needed in the coming years have been prioritised and include: advocating for the critical role of repositories; propagating good practices; and trying to strengthen national coordination networks as a mechanism to support communities of practice. These CoPs are thought to work best on the national level, where colleagues feel most comfortable to collaborate given they operate under the same funding and infrastructure jurisdictions. Such relationships are harder to establish at higher levels. Certification is a high bar to reach for many repositories, especially when considering understaffing and other challenges many repositories face to continue their services. It is important to develop models that allow repositories to incrementally adopt good practices and prioritise actions based on their own environment and capabilities. Forcing requirements too strictly and too quickly will only play into misuse and misunderstanding by those who enforce at the cost of those who should comply. Frameworks and criteria should be presented as supporting tools that can help advance in the right direction.

Next, Mari Kleemola (FSD) presented the new EU project FIDELIS that will start in January 2025. This project will focus on developing a European network of trustworthy repositories. FIDELIS will strive towards increased harmonisation and interoperability between repositories, as well as upskilling and knowledge exchange. The aforementioned desires and challenges regarding communities of practices will be explored and reported on as well. The project will not seek to replicate or create new criteria or requirements, but rather harmonise different existing initiatives and efforts. The consortium consists of a broad representation of repositories from different locations and domains in Europe, and will have cascading grants available to train and onboard other repositories. Apart from this, collaboration with existing organisations, networks, CoPs, and initiatives, like COAR and RDA^{57, 58, 59}, will be essential for the achievement of the goals that are set out.

From this session, several recommendations were extracted:

⁵⁶ Shearer, K., Nakano Koga, S. M., Rodrigues, E., Manola, N., Pronk, J. martine ., & Proudman, V. (2023). Current State and Future Directions for Open Repositories in Europe. Zenodo. <https://doi.org/10.5281/zenodo.10255559>

⁵⁷ <https://www.rd-alliance.org/groups/coordinating-earth-space-and-environmental-science-data-preservation-and-scholarly-publication-processes-wg/>

⁵⁸ <https://www.rd-alliance.org/news/take-part-in-the-rda-wds-trust-principles-outreach-and-adoption-working-group-survey/>

⁵⁹ <https://www.rd-alliance.org/groups/community-based-catalogue-requirements-trustworthy-technical-repository-service-providers/members/all-members/>

- Effort should be made to build **more communities of practice** around the topic of trustworthiness and certification, as an essential form of peer to peer support.
- When developing and providing support, it is important to realise that a **mixed approach of general and tailored support** is most effective. However, **financial support** is also needed.
- Assess the advantages and disadvantages to generic and discipline-specific trustworthiness schemes. Wherever possible, **harmonise trustworthiness schemes** to better enable the provision of support.
- As we move towards the development of a European network of trustworthy repositories, an **inclusive approach** focusing on **strong collaboration** between all involved stakeholders is essential.

3.6 *Metadata, semantics and interoperability*

Recommendations based on the Synchronisation Force workshop 2024 session

The recommendations discussed in the 2024 edition of this Synchronisation Force topical session were based on the results of the 2023 edition, reported in M1.8.

1. The EOSC-IF (Interoperability Framework) provides a set of recommendations that need to be implemented. The **use of indicators and metrics can assess** the framework's implementation. The use of a tool such as Compliance Assessment Toolkit (CAT) can facilitate this assessment and the recommendations'.
2. **Extend the set of semantic objects** described in the EOSC-IF to include artefacts such as mappings and crosswalk
3. Recognise the **semantic artefact catalogue as a critical part of the long-term viability** of any research data infrastructure.

The previous edition of this SyncForce session on Metadata, Semantics and Interoperability, highlighted the importance of a common understanding of semantic artefact (SA) definitions and the need of incentives to encourage their adoption across various disciplines. This year's discussion focused on potential indicators to support assessment of EOSC-IF implementation and thus within a broader framework.

This year's discussion attracted 44 online participants, with 27 registered in the session document. The organisational representation of the audience was diverse, with representatives from 18 EOSC projects, 5 EOSC Association Task Forces (EOSC-TF) and Opportunity Area (OA), 5 European research infrastructures, and about ten consortia and universities.

The session was structured in 5 parts, as follows: 1) after a short introduction of the EOSC-IF and its components, 2) an initiative to assess the implementation of an EOSC Interoperability Policy based on metrics extracted from the recommendations of the EOSC IF was outlined. This was followed by presentations on two semantic EOSC components: 3) the Semantic Artefact Catalogues (SACs) and 4) the Mapping repository. Due to intensive discussion during

the presentations, there was no time for the 5) planned collective activity to identify which EOSC component requires the most attention and which assessment methods are the most appropriate. Thus, the discussions are directly reported into each presentation description.

The session started with the presentation EOSC Interoperability Framework (IF)⁶⁰ as published in 2021, introduced by Esteban Gonzalez. M. He is an engineer in computer science at the Universidad Politécnica de Madrid (UPM)⁶¹ and aside from the FAIR-IMPACT project, he is also an active member EOSC Association Technical and Semantic Interoperability Task Force (TF)⁶². While the EOSC-IF is described as a set of policies and guidelines that enable interoperability of resources and services, and facilitate service composability, he emphasised that interoperability should be viewed holistically, encompassing both data and services. E. Gonzalez highlights the need for interfaces to facilitate interoperability that must interact technically and semantically. E. Gonzalez also shared with us the challenges and needs in supporting components where semantic interoperability is fundamental, pointing to the metadata catalogue, SAC and mapping repository.

This session brought out strong reactions regarding the definition to be allocated to semantic artefacts, terminologies, controlled vocabularies and ontologies, highlighting the confusion surrounding these terms and the need for standardisation. Although the “Semantic Spectrum”⁶³ of knowledge organisation systems was well-received, semantic artefact is reported to be used in an agnostic discipline, whereas the others terms seems to be mostly associated with a specific domain of discipline. This debate seems stormy and endless, as one of the participants said *“there is no good wording, we have to live with multiple explanations for the diverse audiences”*. Nevertheless, the audience acknowledged the importance of considering these nuances in mind for future discussions.

The second presentation, led by E. Gonzalez, outlined the development of an EOSC Interoperability Policy. This policy aims to transform EOSC-IF recommendations into actionable guidelines. This initiative addresses the need to evaluate interoperability within the EOSC ecosystem (components) and assess implementation levels. E. Gonzalez also provided examples on how to proceed for this transformation. Inspired by semantic interoperability profiles of the SIP Wizard questionnaire⁶⁴, this effort is a collaboration with FAIRCORE4EOSC⁶⁵, leveraging the Compliance Assessment Toolkit⁶⁶ (CAT) to assess interoperability policies. One policy could be an EOSC Interoperability policy based on the recommendations provided by the EOSC-IF and the EOSC-TFs. To be assessed, these recommendations should be transformed on metrics. .

⁶⁰

<https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en>

⁶¹ <https://www.upm.es/>

⁶² <https://eosc.eu/advisory-groups/technical-and-semantic-interoperability-task-force/>

⁶³ <https://datascience.cancer.gov/news-events/blog/semantics-primer>

⁶⁴ <https://sip-wizard.ds-wizard.org/wizard/>

⁶⁵ <https://faircore4eosc.eu/>

⁶⁶ <https://faircore4eosc.eu/eosc-core-components/compliance-assessment-toolkit-cat>

This session sparked intense discussion on the framework of interoperability, particularly regarding concerns over “political interoperability”. A clear political commitment is needed to break the persistent data silos within the EU itself, especially the EU Open Data Portal⁶⁷ and EOSC. These administrative barriers highlight the lack of concrete efforts to unify data publication, despite the 2017 publication of the “New European Interoperability framework”⁶⁸ which address interoperability efforts among European public administrations. The discussion then shifted to the challenges of defining and implementing an interoperability policy in complex contexts, particularly where the services come from different communities or political systems.

The third presentation focused on Semantic Artefact Catalogues, a key semantic functional component of the EOSC-IF by Clement Jonquet, who is a senior researcher at INRAE⁶⁹ and FAIR-IMPACT WP4 lead. After a brief introduction to the definitions of SA adopted in EOSC, C. Jonquet emphasised the critical role of SACs in hosting, aligning, enabling reuse, supporting FAIR principles and promoting semantic interoperability. SACs are even studied by the EOSC-TFs⁷⁰ as potential indicators of the effectiveness of semantic interoperability. Within WP4 of FAIR-IMPACT, SACs play a central role, with multiple aspects being developed, consolidated and/or transferred across various uses-cases. Examples of success-story were itemised, such as SACs in the agri-food (AgroPortal), ecology (EcoPortal), and earth sciences (EarthPortal). Those catalogues are built/consolidated within FAIR-IMPACT particularly with a mutualisation effort within OntoPortal Alliance⁷¹, a consortium of several research and infrastructure teams dedicated to promoting the development of SACs based on the open, collaboratively developed OntoPortal software. FAIR-IMPACT has demonstrated its capacity to transfer FAIR enabling tools/methods between scientific communities with respect to SAs, for instance, by deploying the O’FAIR tool (the Ontology FAIRness Evaluator⁷² that was originally available only on AgroPortal) to EcoPortal, EarthPortal and BiodivPortal. The work-package also presented the achievement of an exhaustive review of SACs and FAIR enabling dimension associated and finally the specification of Metadata for Ontology Description and Publication (MOD)^{73,74} and the corresponding standard API (Application Programming Interface) specification for SAC called MOD-API^{75,76}.

The audience appreciated all the accomplishments presented. The question addressed technical challenges related to using format like RDF for the APIs, and support SPARQL, highlighting the need to ensure semantic web objects are queryable across multiple catalogues without relying on new centralised APIs.

⁶⁷ <https://data.europa.eu/en>

⁶⁸

<https://op.europa.eu/fr/publication-detail/-/publication/bca40dde-deee-11e7-9749-01aa75ed71a1/language-en>

⁶⁹ <https://www.inrae.fr/>

⁷⁰ <https://zenodo.org/records/10518860>

⁷¹ <https://ontoportal.org/>

⁷² <https://github.com/agroportal/fairness>

⁷³ <https://github.com/FAIR-IMPACT/MOD>

⁷⁴ <https://zenodo.org/records/10725304>

⁷⁵ <https://github.com/FAIR-IMPACT/MOD-API>

⁷⁶ <https://zenodo.org/records/12579779>

The final topic dealt with the Mappings repository another key semantic component of the IF with an introduction of the topic and presentation of the FAIRCORE4EOSC's service MSCR (Metadata Schema and Crosswalk Registry)^{77,78} which were respectively presented by Yann Le Franc and Joonas Kesaniemi who respectively lead the SA crosswalk and mapping task in FAIR-IMPACT, and the MSCR development in FAIRCORE4EOSC and jointly co-lead the RDA⁷⁹ FAIR mappings Working Group (WG)⁸⁰. Y. Le Franc began by defining the terms Mapping and Crosswalks. Mapping refers to connections and relationships between different information elements by identifying similarities, while crosswalks are defined as sets of mappings. The FAIR-IMPACT project addresses various aspects of mappings and crosswalks, including developing guidelines, a governance framework, and a machine-actionable common exchange model to share mapping practices. These practices are informed by contributions identified through community-engagement workshops^{81,82,83,84} and the FAIR Mapping RDA WG activities. After this brief introduction of mappings and related initiatives, M. Kesaniemi carried on, by presenting the MSCR, enabling users to create, register and version schemas and crosswalks, making published content searchable and browsable. The platform allows projects and researchers to manage and share metadata schemas and crosswalks, supporting community reuse and extension with version control.

During this presentation, the conversation reflected a strong focus on addressing the practical barriers to using semantic resources, particularly the need for clearer incentives, better tools, and more accessible platforms to facilitate adoption and ensure quality. There was also recognition of the human aspect of interoperability, including how researchers engage with these tools and how to make the process more inclusive and user-friendly.

Before the launch of the SyncForce session, a questionnaire was shared with registered participants to gather their views regarding the EOSC-IF. Below is a summary of responses on implementing the EOSC-IF as an EOSC component. Key steps raised include: enhancing semantic interoperability through a repository of controlled vocabularies for consistent cross-domain use, expanding IF scope with a holistic approach focused on impactful solutions and practical guidelines, and fostering institutional and governmental support for interoperability. Regarding measurable adoption of the EOSC-IF, suggestions included collecting quantitative data on downloads, link-based metrics, and usage tracking at EOSC, national, and institutional levels, alongside compliance tools and national/institutional support. Finally, the question on future EOSC-IF priorities highlighted ongoing monitoring and support for interoperability and up-to-date semantic standards, as well as addressing cross-sectoral and global challenges.

⁷⁷ <https://cscfi.github.io/mscr-docs/>

⁷⁸ <https://faircore4eosc.eu/eosc-core-components/metadata-schema-and-crosswalk-registry-mscr>

⁷⁹ <https://www.rd-alliance.org/>

⁸⁰ <https://www.rd-alliance.org/groups/fair-mappings-wg/activity/>

⁸¹ <https://fair-impact.eu/events/fairimpact-events/why-mappings-matter-and-how-make-them-fair>

⁸² <https://fair-impact.eu/events/fair-impact-events/documenting-mapping-community-practices>

⁸³ <https://fair-impact.eu/events/fair-impact-events/developing-mapping-process-framework>

⁸⁴ <https://fair-impact.eu/news/collecting-ways-doing-mappings-take-survey>

To conclude, the topics discussed during this session touch on different facets of interoperability, semantic artefacts, governance and challenges in implementing these aspects across disciplines, highlighting the need of practical implementation and unified strategy for EOSC.

4 Conclusions and next steps

The 2024 workshop series was the third and final Synchronisation Force workshop series of FAIR-IMPACT. It was completed as planned and brought together a multitude of participants from various EOSC and FAIR initiatives. This included individuals who had participated in the previous FAIR-IMPACT (2022-2023) and FAIRsFAIR Synchronisation Force workshops (2019-2021). Information gathered during this workshop series is available on the project website⁸⁵ as well as on Zenodo (see Appendices).

It will help the FAIR-IMPACT partners identify and understand the current state of FAIR developments in a broader context. The recommendations from this report, together with those from the 2022 and 2023 reports, will form the basis of a White Paper, scheduled to be delivered by March 2025.

Finally, the Synchronisation Force as a mechanism will be part of the FAIR-IMPACT sustainability plan, with the aim to be taken up and maintained after the end of the project by other EOSC-related projects and/or the EOSC Association.

⁸⁵ FAIR-IMPACT Synchronisation Force <https://fair-impact.eu/synchronisation-force>

5 Appendices

5.1 Underlying materials

Available in the FAIR-IMPACT community in Zenodo⁸⁶:

- Data provided by workshop participants in the collaborative spreadsheet (separate spreadsheets per session)⁸⁷
- Slides from the opening session⁸⁸
- Slides from the session on *Metrics and assessing FAIRness*⁸⁹
- Slides from the session on *Sustainability of project outputs*⁹⁰
- Slides from the session on *Persistent identifiers*⁹¹
- Slides from the session on *Legal & organisational interoperability*⁹²
- Slides from the session on *Trustworthy and FAIR-enabling repositories*⁹³
- Slides from the session on *Metadata, semantics and interoperability*⁹⁴
- Slides from the concluding session⁹⁵

5.2 Participant list

The 91 workshop participants represent the following organisations:

Affiliation	Number of representatives
Aalborg University	2
Accurids	2
Alma Mater Studiorum - University of Bologna	1
ASREN	1
ATHENA RC / OpenAIRE	1
Bochum University of Applied Sciences	1
CESSDA ERIC	5
CLARIN ERIC	1
CNR	1
CNRS	2

⁸⁶ FAIR-IMPACT Zenodo community <https://zenodo.org/communities/fair-impact/>

⁸⁷ <https://doi.org/10.5281/zenodo.14507277>

⁸⁸ <https://zenodo.org/records/14509622>

⁸⁹ <https://zenodo.org/records/14509677>

⁹⁰ <https://zenodo.org/records/14509710>

⁹¹ <https://zenodo.org/records/14509722>

⁹² <https://zenodo.org/records/14509732>

⁹³ <https://zenodo.org/records/14509739>

⁹⁴ <https://zenodo.org/records/14509754>

⁹⁵ <https://zenodo.org/records/14509782>

Affiliation	Number of representatives
CNRS-IRD	2
COAR	1
CODATA	1
CSC - IT Center for Science	4
CSIC	1
DANS-KNAW	3
DataCite	2
DCC	3
DeiC	1
EC - DG RTD	1
EMBL	1
EOSC-A	1
ERINHA	1
ETT	1
Foundation for Research and Technology - Hellas (FORTH)	1
Fujitsu	1
GARR	1
GO FAIR Foundation	1
GWDG	2
HEAnet	1
IFCA-CSIC	1
Ifremer	2
INRAE	1
Lawyer	1
Leibniz Information Centre for Economics	1
Lund University	1
Lund University/ICOS Carbon Portal	1
Ministère de l'Enseignement Supérieur et de la Recherche (France)	1
National Forschungsdateninfrastruktur (NFDI)	2
National Oceanography Centre, British Oceanographic Data Centre	1
NFDI	1
Observatoire Astronomique de Strasbourg	1
OpenAIRE AMKE	3
OPENAIRE Non-Profit Civil Partnership	1
RDA Europe	4
Tampere University	1

Affiliation	Number of representatives
The University of Manchester	1
Trust-IT	1
UCL Advanced Research Computing	2
UK Data Service	2
UKRI	1
UL/FDV-ADP	1
UMC Utrecht	1
Universidad de Murcia	1
University of Copenhagen	1
University of Edinburgh	2
University of Limerick	1
University of Ljubljana	1
University of Oslo	1
University of Oxford	2
X-officio	2
ZBW	1
Grand Total	91