eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

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/

M1.8 - Second synchronisation workshop

Work Package	**WP 1 - Project management, synchronisation and sustainability**
Lead Author (Org)	**Marjan Grootveld (DANS)**
Contributing Author(s) (Org)	Sara Pittonet Gaiarin (Trust-IT), Joy Davidson (DCC), Ingrid Dillo (DANS), Maaike Verburg (DANS), Olivier Rouchon (CNRS), Mike Priddy (DANS), Josefine Nordling (CSC), Liisa Marjamaa-Mankinen (CSC), Esteban Gonzales (UOM), Anne-Sofie Fink (DeiC), Romain David (ERINHA), Richard Dennis (UCPH)
Due Date	**30-04-2024**
Date	**12-04-2024**
DOI	**https://zenodo.org/doi/10.5281/zenodo.10959527**
Version	**V1.0**

Dissemination Level	
X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Versioning and contribution history

Version	Date	Author	Notes
0.1	19.12.23	Marjan Grootveld	§1-2; structured the subsections of §3
0.2	31.01.24	Maaike Verburg, Olivier Rouchon, Mike Priddy, Josefine Nordling, Liisa Marjamaa-Mankinen, Marjan Grootveld, Esteban Gonzales, Sara Pittonet Gaiarin, Anne-Sofie Fink, Ingrid Dillo, Joy Davidson	Subsections of §3, appendices Draft for workshop 8th February.
0.3	20.01.24	Romain David, Richard Dennis	Feedback and resources added
0.4	24.03.24	Richard Dennis	Editing
0.5	2.04.24	Marjan Grootveld	Consistency check
1.0	12.04.24	Sara Pittonet Gaiarin	Final version

Table of Contents

TERMINOLOGY

| Terminology/Acronym	Description
CC	Creative Commons
CESSDA	Consortium of European Social Science Data Archives
DOI	Digital Object Identifier
EOSC	European Open Science Cloud
FIP	FAIR Implementation Profile
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
KPI	Key Performance Indicator
MAR	Multi-Annual Roadmap
ORCID	Open Researcher and Contributor ID (Identifier)
PID	Persistent Identifier
ROR	Research Organization Registry
SF	Synchronisation Force
SRIA	Strategic Research and Innovation Agenda

1.   Introduction

Expanding on the successful Synchronization Force approach (2019-2021) of the FAIRsFAIR[1] project, FAIR-IMPACT enhances dialogue for collaboration and harmonisation across EOSC and FAIR ecosystems. This effort aims to minimise redundancy and promote sustainable, widely accepted, and easily transferable solutions to the EOSC Partnership, facilitating the adopting of FAIR-enabling practices among current and future EOSC stakeholders.

To tackle these challenges, FAIR-IMPACT established a Synchronization Force comprising representatives from all project work packages. Between 2022 and 2024, the force organises three annual workshops. Each workshop produces a concise report[2], with crucial ecosystem representatives invited to contribute (see Image 1).

Image 1: FAIR-IMPACT's landscape of key stakeholders
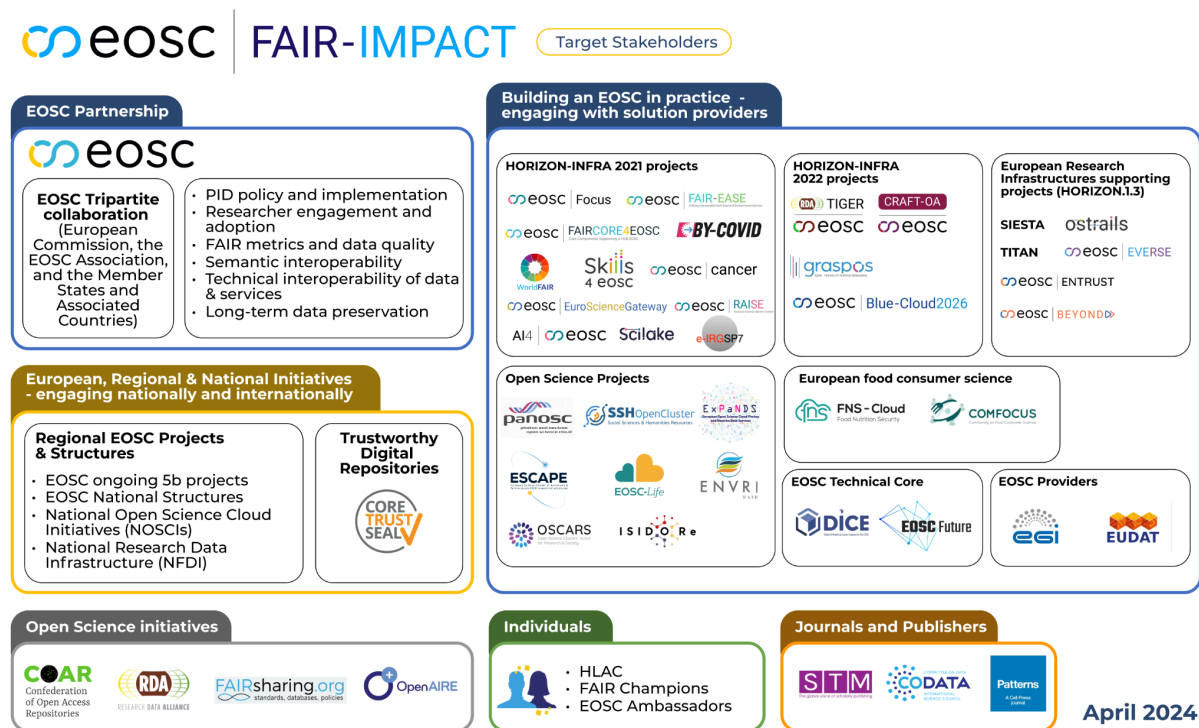
This landscape for synchronising consists of the Board of Directors of the **EOSC Association** and a selection of Task Forces under the EOSC Association that are most relevant for the FAIR-IMPACT focus areas (top-left).

**Funded by
the European Union**

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), as well as the most recent granted projects in support of European Research Infrastructures (Horizon 1.3 projects), as well as the Open Science projects (former ESFRI Cluster projects and those recently funded under the umbrella of the OSCARS project) and a few newcomers from the food sector. 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 workshop in 2023, which has been the second synchronisation force workshop organised by the FAIR-IMPACT project.

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 Report[3] (2018), the EOSC interoperability framework[4] (2021), FAIRsFAIR White Paper[5] (2021), Strategic Research and Innovation Agenda[6] (version 2022), EOSC Multi-Annual Roadmap[7] (2023-2024), and the EOSC Multi-Annual Roadmap (MAR) 2025 and 2026-2027[8].

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://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en
[5] FAIRsFAIR White Paper https://doi.org/10.5281/zenodo.5744786
[6] SRIA https://eosc.eu/sites/default/files/SRIA%201.1%20final.pdf
[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, which span 2022 to 2024, unite various stakeholders to review progress in the FAIR-IMPACT focus areas periodically. This particular milestone documents the activities and outcomes of the 2023 workshop.

2.1.   Role of the milestone

This milestone demonstrates how FAIR-IMPACT orchestrates an annual workshop, primarily through Task 1.3, *"Synchronisation with EOSC Partnership and relevant projects and initiatives."* This event monitors advancements in FAIR practices across a broad spectrum of EOSC initiatives and disseminates these developments among the initiatives' representatives.

2.2.   Means of verification

Details of the Synchronisation Force workshop 2023 are accessible through:

- The project website[9], which hosts the programme and slide decks;

- Zenodo contains supplementary materials such as a spreadsheet with participant contributions (refer to Appendices for links).

The 2023 workshop featured seven online sessions from 2 November 2023 to 8 February 2024, covering:

- An introduction to FAIR-IMPACT, outlining the workshop's objectives and 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 *Metadata, semantics and interoperability*

- A session on *Persistent Identifiers*

- A session on *Trustworthy and FAIR-enabling repositories*

- A session on *Legal and organisational interoperability*

- A concluding session summarising vital insights and recommendations from the thematic discussions for participant reflection.

---

[9] FAIR-IMPACT
https://fair-impact.eu/events/synchronisation-force-events/synchronisation-force-2nd-workshop-november-2023

Around 80 individuals registered for the workshop series, each attracting 30 to more than 40 attendees.

2.2.1. Proof of Milestone fulfilment as per the respective GA table

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

3.   Highlights and recommendations from the Milestone

3.1.   *Metrics and assessing FAIRness*

Underlying recommendations/questions: As outlined in the EOSC Multi-Annual Roadmap[10] (2023-2024), Operational Objective 6, page 15, emphasises the necessity to "Provide the metrics and tools to measure the adoption of the FAIR principles for research outputs."

Recommendations based on the Synchronisation Force workshop 2023 session

The *metrics and assessing FAIRness* theme remains a focal point for the Synchronisation Force workshops. The 2023 session aimed to evaluate advancements based on the objectives set during the 2022 session, focusing on:

- We need to work on a further convergence of metrics and tools, which requires further discussion, synchronisation and alignment;

- We need more domain-sensitive assessment methods to incorporate domain maturity and specific good practices and requirements.

- We need assessment tools for other research outputs, like software and semantic artefacts.

- The FAIR assessment and scoring instrument should be seen and used as the starting point for assistance and improvement.

The Synchronisation Force session was designed to tackle the critical challenge of enhancing convergence among FAIR assessment practices. It aimed to do so by facilitating a robust exchange of insights on existing and future FAIR assessment activities within the INFRA-EOSC project spectrum and beyond. This session drew approximately 40 participants from various European projects and initiatives, demonstrating a solid collective interest in advancing FAIR assessment methodologies.

Participants observed a **notable diversity in FAIR assessment tools**[11, 12] currently in use. This led to a consensus on the potential benefits of more systematically cataloguing FAIR assessment methods applicable across various research outputs. Current efforts to harmonise FAIR assessment tools, especially for data objects, were acknowledged. These efforts respond to the tools' divergent development trajectories. The EOSC Task Force on

---

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

[11] EOSC Association Task Force 'FAIR metrics and data quality' https://doi.org/10.5281/zenodo.7463421 and https://doi.org/10.5281/zenodo.10490289

[12] Community-driven Governance of FAIRness Assessment: An Open Issue, an Open Discussion https://doi.org/10.5281/zenodo.7390482

Semantic Interoperability is also developing a comprehensive catalogue of methods for assessing semantic artefacts[13], [14].

A suggestion was made to leverage the CodeMeta ontology# extension for **cataloguing FAIR assessment activities**, particularly for research software. The CodeMeta project works towards a schema.org extension for basic software descriptions. The potential use of FAIRConnect's **nanopublications**[15] to share and disseminate FAIR assessment methods was also discussed. This approach could facilitate greater accessibility and standardisation of assessment practices across the community.

In response to the second and third recommendations focusing on domain-sensitive assessment methods and metrics for **research outputs beyond data**, the discussion introduced two pivotal topics: the **FAIRness assessment of Research Software**[16] and the evaluation of semantic artefacts. The session showcased FAIR-IMPACT's efforts to adapt the FAIR Principles specifically for Research Software, developing a comprehensive set of metrics for its FAIRness assessment[17] alongside a customised F-UJI[18] tailored for this purpose. Emphasising the importance of domain-specific practices, this initiative includes collaboration with CESSDA[19] to align with the social science domain's unique requirements. The **FAIR Implementation Profiles (FIPs**[20]**)** generated from the WorldFAIR[21] project serve as crucial references for incorporating domain-specific considerations into the FAIR-IMPACT project. The forthcoming Open Science Trails project, launching in early 2024[22], aims to enrich this landscape by creating additional FIPs tailored to specific community needs. A call for enhanced collaboration on creation, accessibility, discoverability[23], and reuse of FIPs across the community was highlighted as a pathway to mutual benefit.

Furthermore, the session offered insights into the ongoing efforts to **assess semantic artefacts** within the FAIR-IMPACT project (see also Section 5.3), encompassing various resources such as vocabularies, ontologies, taxonomies, lexicons, and knowledge graphs. A

---

[13] Converging on a Semantic Interoperability Framework for the European Data Space for Science, Research and Innovation (EOSC) https://zenodo.org/records/8102786

[14] Proposal for the EOSC Semantic Interoperability Questionnaire https://zenodo.org/records/8028392

[15] For instance, search http://v2.fairconnect.pro/ for FIP = SeaDataNet.

[16] FAIR Principles for Research Software (FAIR4RS Principles) (1.0). Zenodo. https://doi.org/10.15497/RDA00068

[17] Metrics for automated FAIR software assessment in a disciplinary context (1.0 - DRAFT not yet approved by the European Commission). Zenodo. https://doi.org/10.5281/zenodo.10047401

[18] To keep track of the implementation of the research software metrics in F-UJI, you can watch this repository: https://github.com/softwaresaved/fuji. The development work in that fork will be integrated into the main F-UJI repository as we go (https://github.com/pangaea-data-publisher/fuji). This work has started in November 2023 and is expected to finish in Spring 2024.

[19] Consortium of Social Science Data Archives: https://www.cessda.eu

[20] FAIR Implementation Profiles: https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/

[21] WorldFAIR project https://codata.org/initiatives/decadal-programme2/worldfair/

[22] Open Science Trails project: https://ostrails.eu/

[23] FIPs can be found in FAIR Connect: https://fairconnect.pro/

significant challenge identified was determining the appropriate level of granularity for FAIRness assessment, suggesting the necessity to evaluate not only the artefacts in their entirety but also their components, such as concepts. The discussion acknowledged the importance of FAIRness in semantic artefacts while also bringing to light the critical issue of sustainability. This includes the need for careful management of the evolution of terms to ensure their continued relevance and FAIR compliance over time, as well as addressing the challenges associated with funding and hosting these resources. Therefore, the sustainability of semantic resources emerges as a complex challenge **requiring enhanced governance and community support** to ensure their long-term viability and effectiveness.

The concluding recommendation from the 2022 Synchronisation Force session emphasised that assessments should catalyse enhancement rather than merely achieving a FAIRness score. In alignment with this philosophy, participants concurred that the primary objective should **shift towards facilitating the practical application of FAIR principles** rather than solely focusing on their measurement. To further this aim, the concept of **'pre-assessments'** was introduced, advocating for evaluating research outputs at multiple stages throughout their lifecycle rather than limiting assessments to the point of publication. This approach acknowledges the limitations of current automated tools, which predominantly utilise persistent identifiers assigned at the deposit or publication stages.

Moreover, it was recognised that researchers would benefit significantly from the **specialised, local support of Data Stewards**, who play a crucial role in guiding the assessment and enhancement of research outputs' FAIRness throughout the lifecycle. This need for expert support is also echoed in the revised Multi-Annual Roadmap (MAR), which underscores the importance of ensuring the availability of Data Stewards and Software Engineers. These professionals are envisioned to possess expertise in the curation and long-term preservation of semantic artefacts and research software, furthering the mission to embed FAIR principles deeply within research practices.

Building upon the insights gained from the discussions, several key recommendations are proposed to augment those identified in the 2022 Synchronisation Force session.

- A broader community must be engaged and solicited for collaboration in developing a comprehensive **catalogue of methods.**

- 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.

- There is a pressing requirement to **improve governance** practices surrounding semantic artefacts.

- Conducting **pre-assessment** could facilitate the gradual enhancement of the FAIRness of research outputs throughout the research data lifecycle.

- Researchers require access to **dedicated local support** to assist the general and discipline-specific FAIRification processes of data, research software, and semantic artefacts.

3.2.   Metadata, semantics and interoperability

Interoperability will often be the more challenging of the four FAIR principles - especially interoperability across scientific domains. The session focused on exploring semantic and technical interoperability within and across scientific domains and the compliance of research institutes with the EOSC Interoperability Framework's recommendations.

Three recommendations were given in the 2022 SF Workshop[24]:

- More cross-disciplinary work is needed to align semantic artefacts with the same terms or concepts.

- Maintenance, sustainability, and governance of semantic artefacts deserve attention and agreement across disciplinary communities.

- The FAIR-at-large community should intensify the work on crosswalks and mapping.

Underlying recommendations/questions

"Develop domain and cross-domain interoperability frameworks at the level of vocabularies, ontologies, and metadata schema" (FAIRsFAIR White Paper[25], Recommendation 1)

"Develop interoperability frameworks for FAIR sharing within disciplines and for interdisciplinary research" (Turning FAIR into Reality[26], Recommendation 4)

Semantic and technical recommendations in EOSC Interoperability Framework[27]

---

[24] FAIR-IMPACT report https://doi.org/10.5281/zenodo.7692062
[25] FAIRsFAIR White Paper https://doi.org/10.5281/zenodo.5744786
[26] Turning FAIR into Reality https://doi.org/10.2777/1524
[27] EOSC Interoperability Framework
https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en

Recommendations based on the Synchronisation Force workshop 2023 session

A short questionnaire distributed to participants before the session revealed that 70% utilise semantic artefacts, primarily for metadata description. Moreover, 40% of the participants engage with data from disparate domains, underscoring the necessity for mechanisms to synchronise data across various domains and sources. The discussion highlighted the significance of developing crosswalks for semantic artefacts and creating mappings for standards, databases, and policies to facilitate this alignment.

The discourse also sheds light on multiple initiatives, often within projects, that leverage existing ontologies and semantic artefacts to tackle **challenges at the cross-disciplinary level.** For instance, the FAIR-EASE project conducts analyses across numerous environmental disciplines to ascertain the types of semantic artefacts employed, uncovering several semantic repositories and vocabulary services dedicated to earth science and marine communities. Notably, the marine sector frequently utilises the NERC Vocabulary Server[28], as reported by the Blue-Cloud2026 project[29], which amalgamates federated data from various Blue Data Infrastructures to enhance dataset search harmonisation.

Additionally, the ELIXIR infrastructure integrates over 50 community ontologies for utilisation within FAIRsharing[30] cataloguing nearly 800 terminologies and their relationships with databases, other standards, and data policies within research data management. In biology, EMBL-EBI also incorporates external reference resources/ontologies for their datasets, albeit grappling with challenges related to the incompleteness and upkeep of these external reference systems.

FAIR-IMPACT actively creates **ontology repositories and semantic artefact catalogues catering to many disciplines and communities.** This initiative aims to standardise these resources and, ultimately, federate them to foster a more interconnected and FAIR-compliant research ecosystem. Despite these examples, the foundational elements essential for achieving semantic interoperability, such as catalogues of semantic artefacts and their metadata or crosswalks bridging various metadata models, are still in the developmental phase and need stability.

Concurrently, numerous projects and research organisations produce and share their semantic artefacts. A highlighted good practice is employing the I-ADOPT Framework[31] as a

---

[28] NERC Vocabulary Server https://vocab.nerc.ac.uk/
[29] Blue-Cloud2026 https://blue-cloud.org/
[30] FAIRsharing https://fairsharing.org/
[31] I-Adopt Framework
https://www.rd-alliance.org/group/interoperable-descriptions-observable-property-terminology-wg-i-adopt-wg/wiki/i-adopt

semantic intermediary among models, such as for variable description. FAIRCore4EOSC[32] is creating **a registry for developing, registering, and versioning schemas and crosswalks**, which will be made available to the community—approximately two-thirds of the participants who shared detailed information adhered to guidelines for publishing their semantic artefacts. Additionally, various institutions have released guidelines and tools to promote adopting FAIR principles.

Despite thorough examination, there remains **a notable absence of universally accepted, explicit definitions for terms utilised within and across different communities**, alongside a pressing requirement for shared semantic artefacts. This lack of consensus was starkly illuminated during a session poll that posed two questions to the attendees: "What do you understand by metadata?" and "What do you understand by semantic artefact?" The responses ignited a vigorous debate among participants concerning the distinctions between semantics and metadata and between vocabularies and ontologies.

A further point of contention and significance is the challenge of harmonising semantic artefacts across disciplinary boundaries, mainly when researchers and data practitioners are unaware of their existence. Tools and components for crosswalking may facilitate alignment at the level of domain-specific ontologies or semantic artefacts. Nevertheless, the imperative for collaborative engagement across communities to address these challenges still needs to be improved.

Building on the insights from these discussions, several recommendations emerge, complementing those identified during the 2022 Synchronisation Force session:

- Establishing a **common understanding of semantic artefact definitions** to minimise ambiguity is paramount. This common ground will facilitate more precise communication and interoperability across various domains.

- It is imperative to **align semantic artefacts across different disciplines.** While creating domain-specific ontologies and semantic artefacts is a prevalent practice, the methodologies for their development and management ought to be harmonised and shared to the greatest extent feasible. Enhancing access to and across registries could support this objective.

- The community requires **expanding services that support the adoption and practical implementation of semantic artefacts**. Such services are essential to ensure that the potential benefits of semantic interoperability can be fully realised across different research and data management contexts.

3.3. Persistent Identifiers

Underlying recommendations/questions:

"Implement the EOSC PID policy and architecture" (Operational Objective 11 *Strategic Research and Innovation Agenda[33]* p. 166)

"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." and "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)" (European level priority 2.1.A/ national level priority 2.2.H, resp., national level priority 2.2.I/ institutional level priority 2.3.K EOSC Multi-Annual Roadmap (MAR) 2025 and 2026-2027[34]).

Recommendations based on the Synchronisation Force workshop 2023 session

Several presenters discussed the latest advancements in the broader domain of Persistent Identifiers (PIDs). The FAIRCORE4EOSC project has unveiled new opportunities for the different roles within the PID ecosystem, as defined by the EOSC PID Policy[35] — namely PID Authorities, Providers, Managers, and Owners. One significant advancement is introducing a Compliance Assessment Toolkit[36] (CAT) to evaluate a PID service's adherence to the EOSC PID Policy and other pertinent standards, such as those related to FAIR principles and the GDPR. Additionally, the toolkit enables PID Owners and Users to align PID services with their specific use cases by illustrating the potential benefits. A public beta version of the CAT is expected to be released in the summer of 2024.

The initial demonstration of the CAT was conducted during a FAIR-IMPACT workshop[37], where attendees discussed integrating Persistent Identifiers (PIDs) within research lifecycle workflows. This workshop highlighted the critical importance of achieving interoperability between PIDs and other research tools, and it brought to light the complexities involved in managing data workflows and maintaining PIDs within these processes. These observations were echoed during the Synchronisation Force workshop, where participants delved into the intricate challenge of consistently and accurately identifying successive versions of data sets,

---

[33] SRIA https://eosc.eu/sites/default/files/SRIA%201.1%20final.pdf
[34] EOSC MAR 2025 and 2026-2027 https://eosc.eu/sites/default/files/2023-01/MAR_2025-27_draft.pdf
[35] EOSC PID Policy https://data.europa.eu/doi/10.2777/926037
[36] CAT https://faircore4eosc.eu/eosc-core-components/compliance-assessment-toolkit-cat
[37] FAIR-IMPACT
https://fair-impact.eu/events/fair-impact-events/eosc-compliant-pid-implementations-practical-guidelines-implementing-best

underscoring the essential role of PIDs in ensuring data traceability and accessibility throughout the research lifecycle.

A recent Knowledge Exchange[38] report was introduced to spur discussion on the trust and sustainability of PID systems. This report explores PID trust indicators and factors that enable trust, spanning from individual contributors to organisational and technological levels. Despite minimal scepticism regarding the proficiency of PID providers, the report emphasises the significance of positive initial experiences with PIDs in establishing trust. Workshop discussions elicited views on what underpins trust and sustainability in PID services, highlighting the importance of openness, transparency, effective governance, the maturity and reliability of resolution processes, broad adoption, community endorsement, cost-effectiveness, and reasonable expenses. Notably, a well-articulated PID policy and adequate funding were underscored as crucial for the sustainability of PID systems.

Workshop participants also shared insights on their selection process for PID services, often influenced by the practices of significant research communities such as the RDA[39] or NASA[40] or their evaluations of sustainability, security, reliability, and compatibility with federated systems. This approach is guided by community-wide recommendations and the necessity for domain-agnostic and community-specific persistent identifiers, reflecting a collective stride towards a more interconnected and resilient research infrastructure.

The discussion regarding the management of sensitive data within repositories unveiled **a variety of potential data protection strategies**. These strategies ranged from decision-making regarding whether a PID should resolve to a descriptive landing page or directly to the data stream, data encryption and the differentiation between publicly accessible and restricted metadata. This diversity in approaches indicates that the complete automation of PID-related workflows might be an overly ambitious goal. However, it was also highlighted that not all challenges discussed are inherent issues with PIDs themselves, suggesting that some of these concerns might be more effectively addressed through other means.

In the FAIR-IMPACT PID workshop mentioned above, the complexity of research workflows was underscored, particularly the necessity for human oversight in determining which information is essential for downstream processes. This is critical for guiding automated systems accurately, especially in contexts involving sensitive or protected data. One

---

[38] Knowledge Exchange report: De Castro, P., Herb, U., Rothfritz, L., & Schöpfel, J. (2023). Building the plane as we fly it: the promise of Persistent Identifiers. Zenodo. https://doi.org/10.5281/zenodo.7258286. See also https://www.knowledge-exchange.info/event/pids-risk-and-trust.
[39] RDA National PID strategies https://www.rd-alliance.org/rda-national-pid-strategies-guide-and-checklist-final-outputs-and-supporting-materials-available
[40] NASA data citation https://pds.nasa.gov/datastandards/citing/#data-providers

participant noted, "Assuming that both humans and machines can make errors, maybe a combination of machine action and humans will give us the best quality at the end of the day".

Several actionable recommendations were formulated during the session to bolster trust in Persistent Identifier (PID) systems:

- Research communities and organisations should **prioritise early and positive experiences** with PIDs. Data stewards are to be instrumental in facilitating this communication, ensuring researchers understand the value and functionality of PIDs from the outset.

- PID providers are urged to **ensure the long-term persistence** of PID registries and the mechanisms for PID resolution. Research funders are proposed to establish a supportive funding structure to underpin the sustainability of these systems.

- PID providers are recommended to **clarify the relationships between different PIDs**, including how PIDs relate across different versions within a workflow and how PIDs can interlink entities, such as connecting a researcher's ORCID with a dataset's DOI. This recommendation calls for enhanced interoperability among PID systems, with encouragement from EOSC and financial backing from research funders for interoperability efforts.

- There is a call for PID providers and managers to **clearly define the roles** as outlined in the EOSC PID Policy, specifying who is accountable for each aspect of a PID system's maintenance. This also involves **communication with research communities and their data stewards** about these roles and responsibilities.

3.4. *Trustworthy and FAIR-enabling repositories*

Underlying recommendations/questions from the Synchronisation Force Workshop 2022[41]:

- Measure the percentage of repositories in EOSC that will have a certification such as CoreTrustSeal or expose trustworthiness using other mechanisms;

- Provide incremental, continuous and sustainable guidance and assistance to repositories and certification processes;

- Contribute to cooperation across the current initiatives to build and sustain a network of trustworthy digital repositories.

---

[41] FAIR-IMPACT SF Workshop 2022: 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

Recommendations based on the Synchronisation Force workshop 2023 session

The session on Trustworthy and FAIR-enabling repositories was well attended, with almost 30 participants in the virtual room. The session aimed to provide an update on recent efforts to improve the availability and visibility of trustworthy and FAIR-enabling repositories and consider whether any of these can be leveraged to assist the ongoing monitoring of the repository landscape by the EOSC Partnership.

The session started with the central question: What makes a repository trustworthy? The audience's answers highlighted the following aspects: **sustainability, transparency, documentation, provenance, expertise,** and **connection** to the community.

The session began with three conversation starters, each leading to dynamic discussions on the respective topics.

The discussion highlighted the challenge of assessing a repository's trustworthiness or quality. The term 'quality' can be ambiguous, raising questions about whether it pertains to the repository's policies and procedures or the data quality itself. There are varied **interpretations of quality** and debates over responsibility for it. While researchers often consider scientific quality assessments, which repositories can endorse and relay, the focus of repositories and their guidelines typically leans towards "compliance" with established workflows, schemas, etc. It was also noted that **while a certification can indicate the trustworthiness of a service, endorsement by the community is also important**, and the two do not necessarily go hand in hand.

The discussion emphasised the value of support networks for repositories and their certification processes, highlighting the substantial benefits of **thematic**, **national, and broader international networks.** These networks are most effective when they embrace repositories at diverse stages of development and include those aspiring to certification. They facilitate valuable knowledge and expertise exchange through networking between repositories of **differing maturity levels**. Proven support methods within these networks include inter-community assistance and personalised mentoring. The duration of support plays a crucial role in achieving certification success. Equally important is fostering a trusting environment among a small group of colleagues, allowing for open discussions about the strengths and vulnerabilities of repositories.

From the session, several recommendations were highlighted:

- Due to their evident advantages, **support networks** for repositories at all levels (thematic, national, and international) should be established.

- Increasing process transparency is vital to effectively assess repository trustworthiness, as encapsulated in the principle of "**Trust through transparency**."

- Repository support should include **generic solutions** for widespread applicability and **detailed, customised support** for specific local or individual needs.

3.5. *Legal and operational interoperability*

For the first time, the annual SF workshop featured a session on legal and organisational interoperability. This session focused on facilitating collaboration among organisations governed by varied legal and organisational frameworks, policies, and strategies. It addressed the need to ensure that differing legislations do not hinder establishing European public services within and across Member States.

Underlying recommendations/questions

Key question: What is the status and adoption of the legal and organisational recommendations presented by the EOSC Interoperability Framework[42] in different scientific domains?

Selected recommendations in the EOSC Interoperability Framework are:

- A clear management of permanent organisation names and functions needs to be provided.

- Standardised human and machine-readable licences, with a centralised source of knowledge and support on copyright and licences.

- A clear list of EOSC-recommended licences and their compatibility with Member States' recommended licences.

- GDPR compliance for personal data.

Recommendations based on the Synchronisation Force workshop 2023 session

Attendees provided concrete examples highlighting the **need for license harmonisation** across different scenarios: within project consortia spanning multiple countries, legally independent research centres affiliated with a single umbrella organisation, and over time, particularly with historical astronomy datasets.

---

[42] EOSC Interoperability Framework
https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en

Creative Commons[43] licenses are globally recognised and broadly applied, while the ROR[44] (Research Organization Registry) is noted for organisational identification but lacks uniform adoption. The widespread acceptance of CC licenses stems from their universal applicability. However, issues arise with CC-BY licenses when there's an unclear declaration of rights ownership. A CC-BY license, which permits the distribution, modification, and utilisation of material with proper attribution to the creator, gains its effectiveness if the creator's identity is clearly stated in the dataset.

Discussion and feedback revealed that Intellectual Property Rights (IPR) is not a significant concern for participants. Many indicated that datasets in their domain or organisation are openly available, making **IPR considerations** redundant. Yet, the concept of IPR intertwines with the complex issue of data ownership. In scenarios of public funding, the data producer, usually a researcher or team, typically does not hold ownership, which often lies with an institution such as a university. Consequently, the authority determining licensing and reuse conditions may differ significantly from the research contributors. Research institutions recognise these dilemmas but **"don't want to hassle with their researchers"** and do not view them as problematic.

Only a few attendees reported their datasets contained personal or sensitive data. One repository highlighted the 5 Safes Framework[45] as a strategy for securing confidential information. Additionally, a participant detailed her university's collaborative approach, involving data stewards, legal, and knowledge-transfer teams working alongside researchers to address data protection, sharing agreements, exploitation, end-user licenses, and IPR issues, underpinned by a blend of policy, support, and researcher obligations.

The lively discussion emphasised the session's relevance and that it is **hard to untie the respective aspects such as licensing, IPR, data-sharing agreements, and ownership**. The non-legal status of FAIR principles exacerbates these difficulties. The insufficient formalisation of these aspects hampers legal and organisational interoperability, restricts adoption, and complicates the machine-actionability of licensing constraints.

The session yielded several actionable recommendations:

- EOSC and other relevant entities should advocate for Creative Commons (CC) licenses unless another license or license family is predominant within a specific research domain or community. This aligns with the EOSC Interoperability Framework's

---

support for permissive licenses. "A list of EOSC-recommended licences and their compatibility with Member States' recommended licences should be provided."[46]

- Data creators and users should be shielded from the complexities of license 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.

- An **integrated support programme for managing, protecting and licensing data** is recommended for research-performing organisations.

---

[46] EOSC Interoperability Framework
https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1/language-en (p. 24).

4. Conclusions and next steps

The workshop proceeded as scheduled and effectively convened numerous participants from various EOSC and FAIR initiatives, including those who had previously engaged in the FAIR-IMPACT (2022) and FAIRsFAIR Synchronisation Force workshops (2019-2021). The compiled information is accessible on the project website[47] and disseminated on Zenodo (refer to Appendices). This information will assist work packages and project partners in comprehensively understanding the current landscape of FAIR development. The recommendations are intended for consideration and review during the final Synchronisation Force session slated for late 2024.

---

[47] FAIR-IMPACT Synchronisation Force https://fair-impact.eu/synchronisation-force

5.   Appendices

5.1.   Underlying materials

Available in the FAIR-IMPACT community in Zenodo[48]:

- Data provided by workshop participants in the collaborative spreadsheet (separate spreadsheets per session)[49]

- Slides from the opening session[50]

- Slides from 'Metrics and assessing FAIRness'[51]

- Slides from 'Metadata, semantics and interoperability'[52]

- Slides from 'PIDs' [53]

- Slides from 'Trustworthy and FAIR-enabling repositories'[54]

- Slides from 'Legal and organisational interoperability'[55]

- Slides from concluding session[56]

5.2.   Participant list

The 78 workshop participants represent the following organisations:

Affiliation	Organisation type	Country	Number
Alma Mater Studiorum - University of Bologna	Research Performing Organisations	Italy	1
CNR-ISTI	Research Communities & Infrastructures, Research Performing Organisations	Italy	1
CNRS	Research Performing Organisations	France	4
CNRS, Observatoire astronomique de Strasbourg	Service providers, Research Performing Organisations	France	1
CSC - IT Center for Science	Service providers, Data Infrastructures	Finland	3

---

[48] FAIR-IMPACT Zenodo community https://zenodo.org/communities/fair-impact/
[49] https://zenodo.org/doi/10.5281/zenodo.10931604
[50] https://zenodo.org/doi/10.5281/zenodo.10931705
[51] https://zenodo.org/doi/10.5281/zenodo.10931750
[52] https://zenodo.org/doi/10.5281/zenodo.10931788
[53] https://zenodo.org/doi/10.5281/zenodo.10931798
[54] https://zenodo.org/doi/10.5281/zenodo.10931820
[55] https://zenodo.org/doi/10.5281/zenodo.10931835
[56] https://zenodo.org/doi/10.5281/zenodo.7446826

Affiliation	Organisation type	Country	Number
KNAW-DANS	Research Communities & Infrastructures, Data Infrastructures	Netherlands	4
DCC	Service providers, Research Communities & Infrastructures	United Kingdom	1
Eastern Switzerland University of Applied Sciences	Research Communities & Infrastructures	Switzerland	1
EMBL-EBI	Research Performing Organisations	United Kingdom	1
ERINHA	Research Communities & Infrastructures	Belgium	1
ETT	Service providers	Italy	2
Finnish Meteorological Institute	Research Performing Organisations	Finland	1
Foundation for Research and Technology - Hellas (FORTH)	Research Performing Organisations	Greece	1
GÉANT	Service providers	Netherlands	1
GO FAIR Foundation	Research Performing Organisations	Austria	1
Grenoble Alpes University (UGA)	Scientific Societies & Academies	France	1
GRNET	Research Communities & Infrastructures, Data Infrastructures	Greece	1
Hasselt University	Individuals in Science, Publishers, Scientific Societies & Academies, Other	Belgium	1
Helmholtz Association, Helmholtz Open Science Office	Research Communities & Infrastructures, Research Performing Organisations, Policy Making Organisations, Data Infrastructures	Germany	1
ICOS Carbon Portal (hosted by Lund University)	Research Communities & Infrastructures, Research Performing Organisations, Data Infrastructures, Other	Sweden	2
INAF - Italian National Institute for Astrophysics	Service providers, Research Communities & Infrastructures, Research Performing Organisations, Data Infrastructures, Scientific Societies & Academies	Italy	1
Independent Consultant	Other	Greece	1
INRAE	Research Communities & Infrastructures	France	3
Institute of Applied Biosciences, Centre for Research and Technology Hellas	Research Performing Organisations	Greece	1
IT4Innovations National Supercomputing Center	Research Communities & Infrastructures	Czech Republic	1

Affiliation	Organisation type	Country	Number
KU Leuven	Research Performing Organisations	Belgium	1
Laboratoire d'Océanographie de Villefranche	Research Performing Organisations	France	1
Lawyer	Individuals in Science	Spain	1
Leibniz Universität Hannover	Research Communities & Infrastructures, Research Performing Organisations	Germany	1
LifeWatch	Research Communities & Infrastructures	Italy	1
MARIS	Data Infrastructures	Netherlands	2
NOC-BODC, Blue Cloud	Research Communities & Infrastructures	UK	1
Observatoire Astronomique de Strasbourg	National Level Initiatives, Research Performing Organisations, Policy Making Organisations, Data Infrastructures	France	1
OpenAIRE AMKE	Research Communities & Infrastructures	Greece	5
OpenAIRE AMKE	Service providers	France	1
OPERAS Research Infrastructure	Research Communities & Infrastructures	Greece	1
Premotec GmbH	Service providers, Research Communities & Infrastructures, Research Performing Organisations	Switzerland	1
Radboud University Nijmegen	Research Performing Organisations	Netherlands	1
Research Data Alliance	Service providers	Germany	1
Research Data Alliance	Research Communities & Infrastructures, Scientific Societies & Academies	Belgium	1
Research Software Alliance	Research Communities & Infrastructures	Australia	1
Sikt - Norwegian Agency for Shared Services in Education and Research	Service providers, National Level Initiatives, Data Infrastructures	Norway	1
SOCIB	Service providers, Data Infrastructures	Spain	1
SURF	Service providers	Netherlands	1
Tampere University	Research Performing Organisations, Data Infrastructures	Finland	1
The University of Manchester	Research Communities & Infrastructures	UK	1
Trust-IT	Other	Italy	2
TU Graz	Research Performing Organisations	Austria	1
TU Wien Library	Research Performing Organisations	Austria	1
UKRI	Service providers	UK	3

Affiliation	Organisation type	Country	Number
Uni-Freiburg	Service providers, Research Performing Organisations, Data Infrastructures	Germany	1
Universidad Politécnica de Madrid	Research Performing Organisations	Spain	1
Universitat Politècnica de València	Research Performing Organisations	Spain	1
Université Côte d'Azur, CNRS, Inria, I3S	Research Communities & Infrastructures, Scientific Societies & Academies	France	1
University of Copenhagen	Research Performing Organisations	Denmark	1
University of Notre Dame	Research Performing Organisations	United States	1
University of Trento	Scientific Societies & Academies	Italy	1
University Stefan cel Mare of Suceava	Research Performing Organisations, Research Funding Organisations	Romania	1
		Overall total	**78**