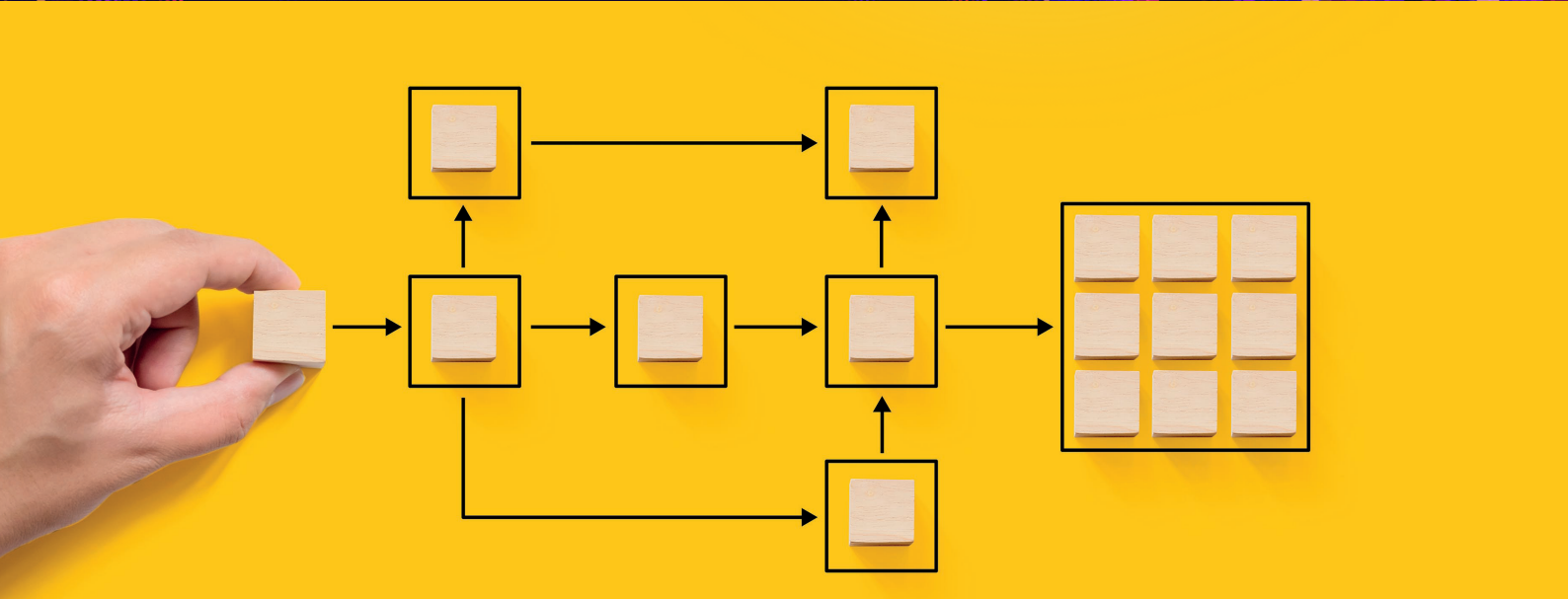
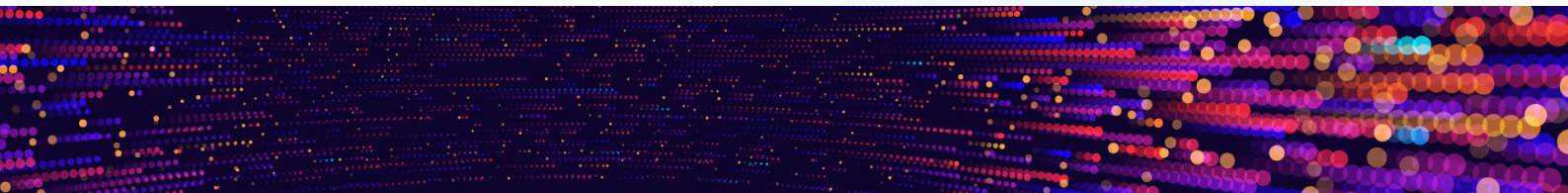




# Towards making the Research Analysis Identifier (RAI ID) EOSC PID policy compliant



## Authors names, affiliations and ORCIDs:

- Evdokimos Konstantinidis, EOSC RAISE Coordinator, Laboratory of Medical Physics and Digital Innovation, School of Medicine, Aristotle University of Thessaloniki <https://orcid.org/0000-0002-5522-9553>
- Gorka Epelde Unanue, Technical Lead EOSC RAISE, Digital and Biomedical Technologies, Vicomtech Foundation, Basque Research and Technology Alliance (BRTA), <https://orcid.org/0000-0002-5179-415X>
- Ilektra Makridou, EOSC RAISE Project Manager, Laboratory of Medical Physics and Digital Innovation, School of Medicine, Aristotle University of Thessaloniki, <https://orcid.org/0009-0004-0563-8959>

## Editors:

- Lassi Lager, CSC and FAIR-IMPACT, <https://orcid.org/0000-0002-5421-8246>
- Rene van Horik, DANS and FAIR-IMPACT, <https://orcid.org/0000-0001-6899-760X>

## Summary:

The persistent identification of research outputs is part of good research data management practice and is central to the FAIR Principles and the vision of the European Open Science Cloud (EOSC). There are many types of persistent identifiers (PIDs) currently being used to identify data and other kinds of research outputs but also different actors involved in the creation of outputs and the organisations that employ them or fund their work. To foster harmonisation on the use of different persistent identifiers, there is a need to define and implement research data and/or PID policies. FAIR-IMPACT's Creating EOSC compliant Persistent Identifier (PID) policies support action<sup>1</sup> aimed to help successful applicants to complete self-assessments with regard to their PID policy readiness through the use of FAIRCORE4EOSC's Compliance Assessment Toolkit (CAT) service<sup>2</sup> which strives to encode, record, and query compliance with the EOSC PID policy. The support action did not focus on any specific PID type but rather provided general best practice guidelines on the creation and assessment of PID policies. This FAIR Implementation Story outlines the experience of the EOSC-RAISE project, coordinated by Aristotle University of Thessaloniki in relation to their self-assessment of the application of the RAI ID, a PID for research analysis objects..

---

1 <https://fair-impact.eu/support-offer-2-creating-eosc-compliant-persistent-identifier-pid-policies>

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



## ■ Introduction:

---

The EOSC-RAISE project aims to provide the mechanisms for a distributed crowdsourced data processing system, moving from open data to data open for processing. The vision of the project is the EOSC Web of FAIR Data and Services for Science, an open, fair and reliable Research Community where every researcher will be accredited for their work and all research data will be equally accessible for processing without violating data protection regulations.

To make the research analysis (RAI ID) PID being developed in EOSC-RAISE EOSC PID policy compliant requires experts' support and the CAT tool's progressive assessment. Research Analysis Identifier (RAI), aims to guarantee the reproducibility of software analysis, reflecting authorship, source dataset (open dataset or protected dataset), algorithm and results. From a technological perspective, RAI ID was previously using EOSC-RAISE generated local identifiers (UUIDs) for the research analysis components (author, script, dataset and results), and governance policy and sustainability plans were still on an early status. So, we expected the action would help us update the PID technological implementation, governance, policies and other relevant aspects covered in the policy, to define globally unique and persistent identifiers for (software) research analysis, based on community best practices.

We had started a collaboration with the EOSC PID community during the EOSC Winter School 2024 within the EOSC Opportunity Area 1 activities. We presented and contrasted an initial analysis of the RAI compliance of EOSC PID Policy according to FAIRCore4EOSC D2.1 Annex D3.1. The aim was to further develop and refine our PID from different perspectives.

Additionally, we would like to make RAI aligned with PID ecosystem services (e.g., PID Metaresolver, PID Graph, DTR) being built to exploit the potential of PIDs and advanced uses.

## ■ Approach taken:

---

EOSC-RAISE had an initial PID concept defined in a previous non-EOSC project, and partners had limited knowledge of the EOSC PID Policy. We used the proposed tool, the CAT service, to evaluate our initial design for the PID. The initial results had a low score.

Based on the knowledge gained from the workshop and support action, the EOSC-RAISE project PID task force had several meetings, discussed the insights from the workshop and the initial assessment to re-engineer the RAI ID, bringing our open questions to PID programme office hours for clarification.

## ■ Challenges encountered and addressed:

---

Challenges encountered before participating in this action were to identify PID-related core sources of information and practical examples and decisions on how best to implement requirements from different areas of the EOSC PID policy. We have used this opportunity to learn how to plan and configure our services to support different policy areas. We are currently advancing our PID metadata redesign, as well as its core services re-engineering.

In the following months, we will rely on collected documentation, our notes, rolling notes and the FAIR-IMPACT planned knowledge base to include certain functionalities in our services and improve the description of our PID handbook to improve the score of our PID's CAT assessment. Some questions might arise in this phase which we will try to answer using documentation and identified forums.

Regarding specific issues with specific policy areas, updating and versioning of the core PID metadata doesn't make much sense in our case, as the PID records performed experiments (software research analysis) and results



information, being its objective to serve as the immutable proof of the performed research (including the necessary reproducibility information). We have interpreted this aspect within our PID, as parts of the non-core metadata that could be updated and could be interesting to be versioned.

## ■ ■ Impact:

---

The workshops and the discussions during office hours have helped us better interpret what should be in a PID policy from a PID (RAI ID) Manager perspective. It has helped us understand and mature aspects that go beyond the pure technical definition of the PID and metadata, considering other aspects such as the reliability or governance of the PID.

We have learned about the challenge of implementing a PID in production. Additionally, we have gained insights into assessing a PID implementation's progress toward fulfilling EOSC policy requirements and periodically evaluating the service's performance to advance the maturity of policy requirements and improve the quality of the service offered.

We still see it crucial to address governance and sustainability within our PID policy, aspects that were not identified before we participated in this programme.

## ■ ■ Key messages:

---

We recommend using the CAT service for guiding your PIDs iteratively to improve policy compliance. Check the examples provided by the service and the assessments made by others. Also, follow the hints provided by the service per each criterion to understand what the policy is asking for. Overall, refer to the FAIR-Impact PID knowledge base to identify existing PIDs for potential reuse, including the possibility of building your PIDs based on existing ones, and to find practical examples and guidance.





