eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

eosc | FAIRCORE4EOSC
Core Components Supporting a FAIR EOSC

Supporting FAIR Implementation: Panel with participants in the FAIR-IMPACT support actions

20 February 2025

FAIRfest 20 February 2025
The Hague, The Netherlands (Madurodam)
Celebrating the advancements in FAIR solutions for EOSC

Three routes of support provided

- Route 1: via dedicated guidance and one-to-one support
  - Research Performing Organisations
  - Repositories and Data Service Providers
  - National Level Initiatives

- Route 2: via support to test specific tools, approaches and solutions

- Via our series of public FAIR Implementation workshops

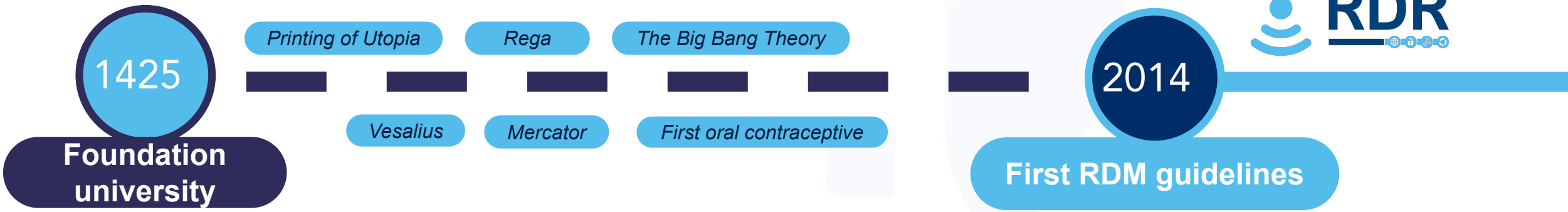
Speakers

- Dieuwertje Bloemen, KU Leuven
- Clara Boavida, ISCTE - University Institute of Lisbon
- Vaidas Morkevičius, Lithuanian Data Archive for Social Sciences and Humanities
- Rory Macneil, ResearchSpace
- Mateusz Pawlik, Paris Lodron Universität Salzburg
- Beth Knazook, Digital Repository of Ireland

Dieuwertje Bloemen, KU Leuven

- KU Leuven: the oldest & largest university of the Low Countries
- Open Science @ KU Leuven

- 2014: RDM working group
- 2018 & 2022: RISE analyses
- 2019: start two major infrastructure projects



Route 2: via cascading grants and support to test specific tools, approaches and solutions

#1.2 Enabling FAIR Signposting and RO-Crate for content/metadata discovery and consumption

#2.3 Recommendations for trustworthy and FAIR-enabling data repositories

#3.3 Testing the trustworthy and FAIR-enabling repositories prototype

#1.2 Enabling FAIR Signposting and RO-Crate for content/metadata discovery and consumption

- Aim: learn more about RO-Crate & integration RO-Crate in Dataverse software
  - Not the only Dataverse team that took part in the call
- Support from experts:
  - Signposting: Herbert Van de Sompel (DANS)
  - RO-Crate: Stian Soiland-Reyes (University of Manchester)
- Result
  - Implementation of Dataverse Signposting in our instance
  - Development of an RO-Crate metadata exporter for Dataverse
  - Creation of an RO-Crate previewer for Dataverse

#2.3 Recommendations for trustworthy and FAIR-enabling data repositories

#3.3 Testing the trustworthy and FAIR-enabling repositories prototype

- Context: building on the work from the RDA DRAWG (Data Repository Attributes Working Group)
- Took part in the RDA working group and the two related FAIR-IMPACT support actions
- #3.3 is still ongoing

#2.3 Recommendations for trustworthy and FAIR-enabling data repositories

#3.3 Testing the trustworthy and FAIR-enabling repositories prototype

- Aim: make repository information more transparent & machine readable at the source.
  - #2.3 test out the proposed set of repository properties
  - #3.3 test a technical prototype for the exposure of these repository properties

2.3 & 3.3: Transparency of our repository's information

- Before:
  - Our repository's characteristics weren't all as easy to find as we thought
  - some attributes weren't in parallel across platforms (re3data.org vs homepage vs documentation)
  - the information is difficult to expose in a machine readable way
- After:
  - addition of a all information on our repository's about page (great exercise to do)
  - exploration on how to choose one value per attribute and plan in work to make it parallel
  - work ongoing to make a prototype for Dataverse to make it machine readable

Challenges along the way

- our first support action took a lot longer than expected
  - FAIR-IMPACT support action was already finished, but we were still working on it much later
  - good thing for progress, but not so much for the project management side of things
  - in general, it was sometimes difficult to predict the time investments necessary up front.
- in the second support call, the differences in interpretation of certain terms in the metadata model was interesting
  - it also sparked interesting conversation on why it can be interpreted in multiple ways, what the intended way was, and how this ambiguity could perhaps be prevented in future versions.

Continued efforts

- RO-Crate
  - continued work on integrating RO-Crate in Dataverse more extensively (though a lot of community consultation necessary for this)
  - exploration of RO-Crate in our ManGO set-up and future cold storage solution
- Repository attributes
  - the support action is still in progress
  - might explore how to visualize the machine readable format in some way and see if a UI to edit and keep it up-to-date is necessary
  - submitted as a topic at the Dataverse community meeting to discuss further work on this in Dataverse

Clara Boavida, ISCTE - University Institute of Lisbon

Route 2: via cascading grants to enable testing of specific tools, approaches and methods

Support offer #4: Improving the availability and machine readability of data policies with FAIRsharing

FAIRsharing.org

Overview of my organisation

- The Instituto Universitário de Lisboa (ISCTE) is a public university in Portugal.

- It has approximately 15 000 students, more than 50% of whom attending postgraduate, masters and doctoral courses. 328 teaching staff, 467 resident researchers and 327 non-teaching staff.

- The main research areas of the Iscte are the social sciences, the humanities, management and information technologies. Iscte has 8 research units.

Overview of my organisation

2023

[Research Data Sharing and Management Policy]

(approved in November 2023)

2013

[CRIS System]

(since 2013)

2009

[Open Access Policy]

(since 2009, updated in 2015)

2006

[Institutional Repository]

(since 2006)

eosc

(November 2024)

Re.Data
Rede para a Gestão de Dados de Investigação

(January 2025)

Context

- The support action was carried out in Portugal from May to September 2024 and it aimed to support the registration of research data policies on the FAIRsharing platform.
- FAIRsharing - is a curated, informative and educational resource on data and metadata standards, databases and policies.
- It updated its own policy metadata in 2023 to reflect the fields covered by the FAIRsFAIR policy checklist.
- As a result, policies registered with FAIRsharing can now make the content of their policies more explicit and comparable by both humans and machines.

FAIRsharing.org
standards, databases, policies

Goals

1. Coordinate the participation of at least 20 stakeholders in the country
2. Engage stakeholders from research performing organisations, funding bodies and publishers
3. Organise and host a virtual event with participants from the country

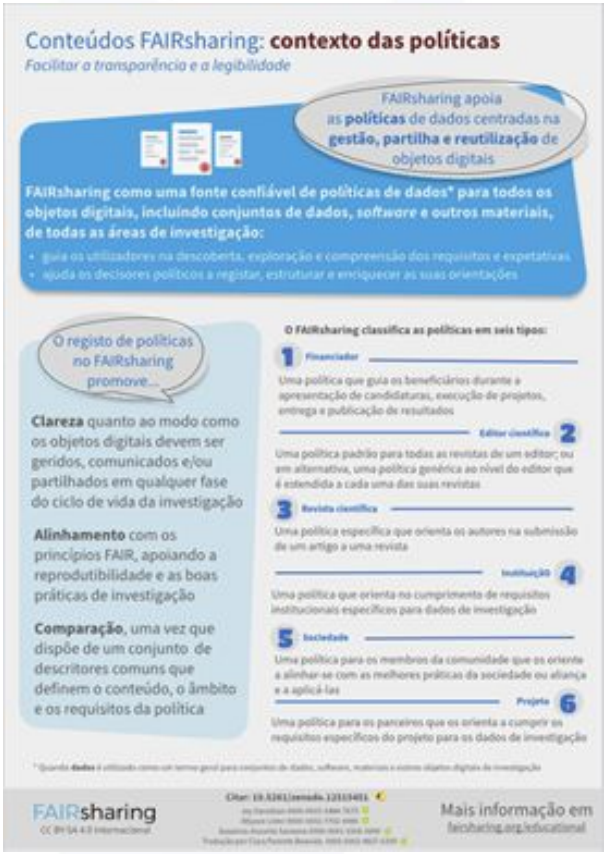
Approach

- Contact with funders, institutions, publishers and scientific journals
  - By email using a standard email about the initiative. List of contacts
  - One-to-one meetings (about 20 online meetings)

- Data Talk on the 27 June 2024
  - Public session on the support action
  - It gathered 32 participants

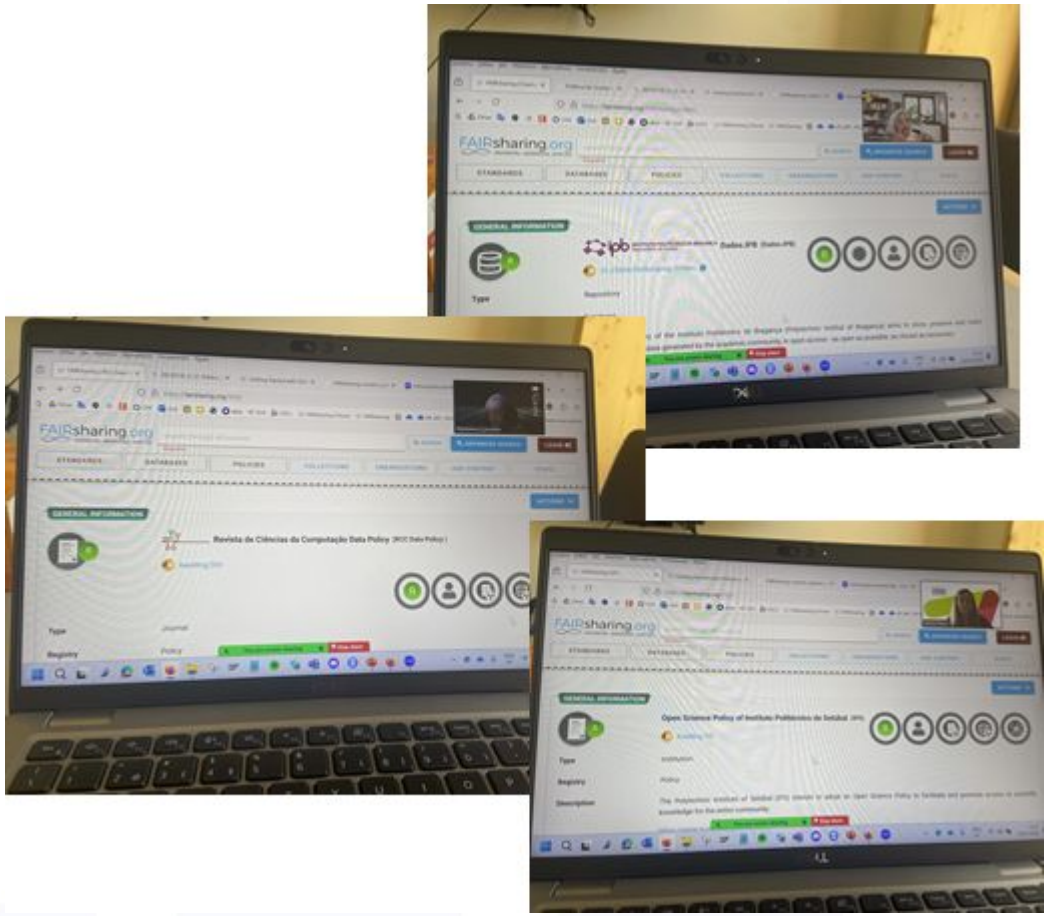
Approach

- Portuguese translation of the Policies overview factsheet
- Application to become a FAIRsharing Community Champion to facilitate the registration of policies



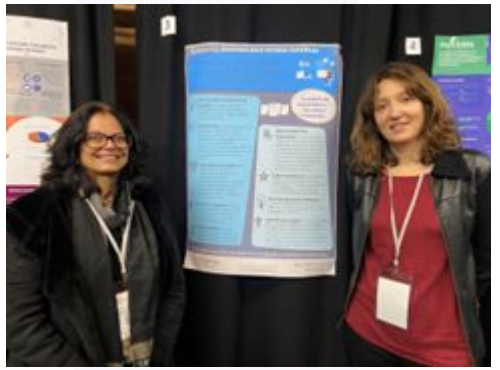
Approach

- National Workshop on the 10 July 2024
  - Online format
  - By invitation
  - 34 participants
  - Testimonials



Results

- More than 20 research data policies were registered increasing the availability of data policies
- Mostly data policies were from scientific journals
- Several institutional policies are still in the approval process
- On 21 and 22 November this year, the results of this initiative were presented at the GDI Forum conference



Challenges

On the one hand, the interactions with representatives of RPO and publishers were very productive. The one-to-one meetings allowed us to share knowledge and explore how they could get involved in and benefit from the FAIR-IMPACT Support Action.

On the other hand, we found it somewhat challenging to convince representatives of RPO and publishers of the value of registering their policies on the FAIRsharing platform and of the importance of having clear policies for sharing research data. They were unsure about the benefits of registration.

Vaidas Morkevcius, Lithuanian Data Archive for Social Sciences and Humanities (LIDA)

Route 2: via support to test specific tools, approaches and solutions
#2: Creating EOSC compliant Persistent Identifier (PID) policies (for LiDA)

Lithuanian Data Archive for Social Sciences and Humanities

Lithuanian Data Archive for SSH (LiDA) Lietuvos HSM duomenų archyvas (LiDA)

(LiDA)

Metrics 34,412 Downloads Contact Share

Lithuanian Data Archive for Social Sciences and Humanities (LiDA) is a virtual digital infrastructure for SSH data and research resources acquisition, long-term preservation and dissemination. It provides access to more than 600 data and research resources. All the data and research resources are documented in both Lithuanian and English according to international standards. Access to the resources is provided via this **Dataverse repository** (not all the resources are available, as in 2020-2029 a migration project from the old infrastructure is being implemented). LiDA curates different types of resources and they are published into catalogues according to the type: Survey Data, Aggregated Data (including Historical Statistics), Encoded Data (including News Media Studies), and Textual Data. Also, LiDA holds collections of social sciences and humanities data deposited by Lithuanian science and higher education institutions and Lithuanian state institutions (Data of Other Institutions).

Depositors interested in deposit of their data into the LiDA Dataverse repository should consult this page.

LiDA is hosted by the **Centre for Data Analysis and Archiving (DAtA) of Kaunas University of Technology** (data.ktu.edu).

Lithuanian Data Archive for Social Sciences and Humanities

LiDA curates different types of SSH research data (mostly, survey data and historical statistics) and employs Dataverse software for this purpose

- Dataverse software allows multiple and rather advanced opportunities to curate data according to FAIR principles

Before the action LiDA did not have a clearly defined PID policy (rather small and not very mature organisation in terms of policy development)

- It simply issued the EUDAT B2HANDLE supported Handles for datasets and all files contained within them

Creating EOSC compliant Persistent Identifier (PID) policies

Datasets to LiDA Dataverse repository are deposited by rather heterogenous data providers

- Primarily, individual researchers and research performing organisation
- Increasingly, governmental agencies and non-governmental organisations

Increasing need for FAIR-enabling data curation

- Participation in this support action was seen as an opportunity to create organisational level PID policy that would best serve interests of our depositors and at the same time become compliant with EOSC and other standard practices

Creating EOSC compliant Persistent Identifier (PID) policies

Support action aimed to help with the definition of EOSC compliant PID policies by completing self-assessments with regard to PID policy readiness through the use of FAIRCORE4EOSC's Compliance Assessment Toolkit (CAT) service, which strives to encode, record, and query compliance with the EOSC PID policy and more (including TRUST, FAIR, Reproducibility, GDPR, and Licences)

- The persistent identification of research outputs is part of good research data management practice and are central to the FAIR Principles and the vision of the EOSC

Creating EOSC compliant PID policy for LiDA

Redefined main goal

- Creating LiDA's PID policy for data sets

The policy should comply to…

- EOSC PID Policy
  - Criteria, Metrics, and Benchmarks:
    12 of the 35 criteria can be mapped to PID Managers

#	Criterion	Description	Imperative
C5	Update Functionality	The PID Manager **MUST** provide the functionality required to maintain PID attributes. (Also mapped to the Actor: Service)	MUST
C6	Ownership Transfer	The PID Manager **SHOULD** provide policies and contractual arrangements for transfer of ownership should the owner no longer be able to assume responsibilities in compliance with the policy.	SHOULD
C7	Resolution Integrity	The PID Manager **MUST** maintain the integrity of the relationship between entities and their PIDs, in conformance to a PID Scheme defined by a PID Authority.	MUST
C11	Versioning - Procedure	PID Services and PID Managers **SHOULD** have clear versioning policies. (Also mapped to the Actor: Service)	SHOULD
C14	Resolution Authenticity	PID Manager **MUST** ensure that the entity remains linked to the PID. In case that the entity being identified is deleted or ceases to exist, tombstone information needs to be included in the PID attribute set.	MUST
C16	Digital Representation	Physical and conceptual entities **MUST** be represented via a digital representation (e.g. landing page, metadata, attribute set, database index) to have a presence in the digital landscape.	MUST
C19	Accurate Entity Metadata	The PID Service **MUST** maintain entity metadata as accurately as possible in collaboration with the PID Owner. This copy is the authoritative version.	MUST
C22	No End User Cost	The basic services of PID registration and resolution **SHALL** have no cost to end users.	SHALL
C28	Certification	PID Authorities and Services **MUST** agree to be certified with a mutually agreed frequency in respect of policy compliance. (Also mapped to the Actors: Scheme, Authority, Service, Owner)	MUST
C29	Agreed Responsibilities	PID Services **SHOULD** agree with PID Managers the responsibilities for Kernel Information maintenance, preferably via contract (Also mapped to the Actor: Service)	SHOULD
C34	Persistent Median	PID Services **SHOULD** aim for a persistence median that is acceptable to and aligns with community and dependency	SHOULD

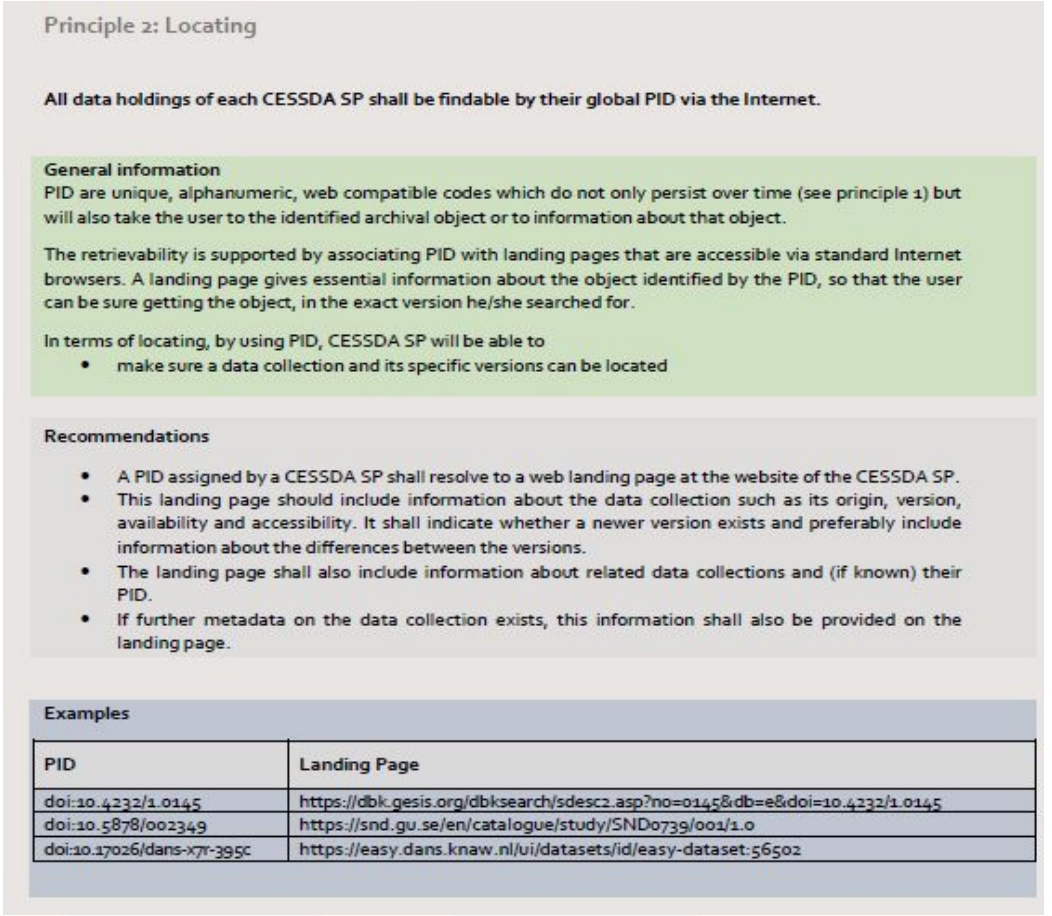
Creating EOSC compliant PID policy for LiDA

Redefined main goal

- Creating LiDA's PID policy for data sets

The policy should comply to…

- EOSC PID Policy
- CESSDA ERIC PID Policy (and Best Practice Guidelines)
  - Background and 5 principles: Identifying, Locating, Resolving, Referencing and Citation, Visibility



Principle 2: Locating

All data holdings of each CESSDA SP shall be findable by their global PID via the Internet.

General information

PID are unique, alphanumeric, web compatible codes which do not only persist over time (see principle 1) but will also take the user to the identified archival object or to information about that object.

The retrievability is supported by associating PID with landing pages that are accessible via standard Internet browsers. A landing page gives essential information about the object identified by the PID, so that the user can be sure getting the object, in the exact version he/she searched for.

In terms of locating, by using PID, CESSDA SP will be able to

- make sure a data collection and its specific versions can be located

Recommendations

- A PID assigned by a CESSDA SP shall resolve to a web landing page at the website of the CESSDA SP.
- This landing page should include information about the data collection such as its origin, version, availability and accessibility. It shall indicate whether a newer version exists and preferably include information about the differences between the versions.
- The landing page shall also include information about related data collections and (if known) their PID.
- If further metadata on the data collection exists, this information shall also be provided on the landing page.

Examples

PID	Landing Page
doi:10.4232/1.0145	https://dbk.gesis.org/dbksearch/sdesc2.asp?no=0145&db=e&doi=10.4232/1.0145
doi:10.5878/002349	https://snd.gu.se/en/catalogue/study/SND0739/001/1.0
doi:10.17026/dans-x7r-395c	https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:56502

Creating EOSC compliant PID policy for LiDA

Redefined main goal

- Creating LiDA's PID policy for data sets

The policy should be congruent with (take into account) opportunities and limitations of the software used (Dataverse)

- Currently, can issue PIDs for data sets and files
- Issuing PIDs for data catalogues?

Creating EOSC compliant PID policy for LiDA

Results

- LiDA Persistent Identifier Policy for Data Sets developed
  - https://zenodo.org/records/13827904
- Action plan for implementing and monitoring the policy
  - Document of alignment with EOSC/CESSDA ERIC policies and Dataverse software created

9	LiDA curates data following the international standards (references to FAIR etc.), which stipulate the use of standard and easy ways to identify and locate data sets. PIDs are employed as instruments for sustainable and reliable discovery and reuse of data sets, collections of data sets and files within data sets stored in the LiDA Dataverse Repository (https://lida.dataverse.lt). PIDs allow to find data sets and files within data sets (providing information if they changed and how; and in case they were removed/relocated, information is provided about the purposes and procedures in the tombstone page) and serve as means for referencing and citing them.	P7, P9	C7, C14	Background	PIDs are issued via APIs to all data sets that are published. No publishing of data set without PID is possible. Removal of data sets is possible, but the usual practice is to deaccession them providing necessary information in the tombstone page (to which PID resolves).	Care needed (data sets are not deleted, but deaccessioned)
10	PIDs shall be registered (issued) for all data sets curated by the LiDA. The data set PID shall not refer to a specific version of a data set, but metadata of the data set shall include information on a specific version of the data set that allows tracking of which version is disseminated and allows users to properly cite the data creator, title of the data set and which version was used.	P6	C11	Background; Principle 5: Visibility	PID is issued to every data set. Version information is indicated in the title and citation information. Different versions of the dataset can be accessed directly with reference to the PID (Talečiene, Audronė; Balžekiene, Aista; Budžyte, Agnė; Zolubiene, Eimante, 2023, "ISSP 2020: Environment IV, Lithuania, November 2022", https://hdl.handle.net/21.12137/RINOXY, Lithuanian Data Archive for SSH (LiDA), V2, UNF:6:Ycckhd9gATlwVPtP79Z1zg== [fileUNF] > https://lida.dataverse.lt/dataset.xhtml?persistentId=hdl:21.12137/RINOXY&version=1.0)	No

Creating EOSC compliant PID policy for LiDA

Results

- LiDA Persistent Identifier Policy for Data Sets developed
  - https://zenodo.org/records/13827904
- Action plan for implementing and monitoring the policy
  - Document of alignment with EOSC/CESSDA ERIC policies and Dataverse software created
    - 4 types of actions identified
      - No action needed
      - Attention needed (every two years)
      - Care needed (certain workflows for publishing)
      - Immediate action needed (in order for practice to align with the policy
  - Workflow (chart) for implementing the policy in practice is being developed

Rory Macneil, ResearchSpace

RO-Crate

- Open source electronic lab notebooks have adopted RO-Crate in the 'ELN exchange format' to facilitate exchange of data between ELNs in a common format
- Export in RO-Crate could also be used for data exchange with other kinds of tools that have adopted RO-Crate, e.g. repositories
- RSpace provides the option to export data in a variety of formats
- **We added the option to export an 'ELN archive' (a zip file containing XML).This is identical to a zip file with XML contents but also contains an `ro-crate-metadata.json` file describing the content**

PID-Policy

- We used the Compliance Assessment Toolkit to assess the RSpace PID Policy
- **Main conclusion**: RSpace acts as more as a broker between PID owners, PID managers, and PID services, and we is not well placed to decide on the details of a PID policy established by administrators of an RSpace instance
- **Benefit**: The CAT assessment enabled us to identify what supporting documentation to our solutions might be helpful for ensuring an educated use of our solutions, both for end users and administrators of RSpace.

- https://doi.org/10.5281/zenodo.8284206

Mateusz Pawlik, Paris Lodron Universität Salzburg

Route 1:

Research Performing Organisations

Route 1:

Repositories and Data Service Providers

Route 2:

#2 EOSC compliant interoperability policies

#3 Trustworthy and FAIR-enabling repositories

PARIS LODRON UNIVERSITY SALZBURG

ANC

Austrian NeuroCloud

FAIR-IMPACT impact:

- Establish Open Science Taskforce
- Implement FAIR-oriented RDM Policy
- Focus on Data Management Plans

PARIS LODRON UNIVERSITY SALZBURG

Domain-specific repository FAIR by design:

- Understand FAIR in the context of neuroscience
- Make FAIR transparent to the depositors
- Incrementally increase FAIRness of our datasets

ANC
Austrian NeuroCloud

FAIR-IMPACT impact:

- Use CoreTrustSeal as a guiding tool
- Focus on terms and policies, especially legal aspects of neuro RDM
- Improve cross-domain data interoperability (EOSC, OpenAIRE, CoreTrustSeal)

Beth Knazook, Digital Repository of Ireland

- Approx 58 member institutions contributing data from the Arts, Social Sciences and Humanities
- CoreTrustSeal Certified Repository
- Longstanding advocacy for Open Science and FAIR through the Research Data Alliance, the OECD, ALLEA, Europeana, the European Commission and the European Open Science Cloud (EOSC), and Irish National Open Research Forum

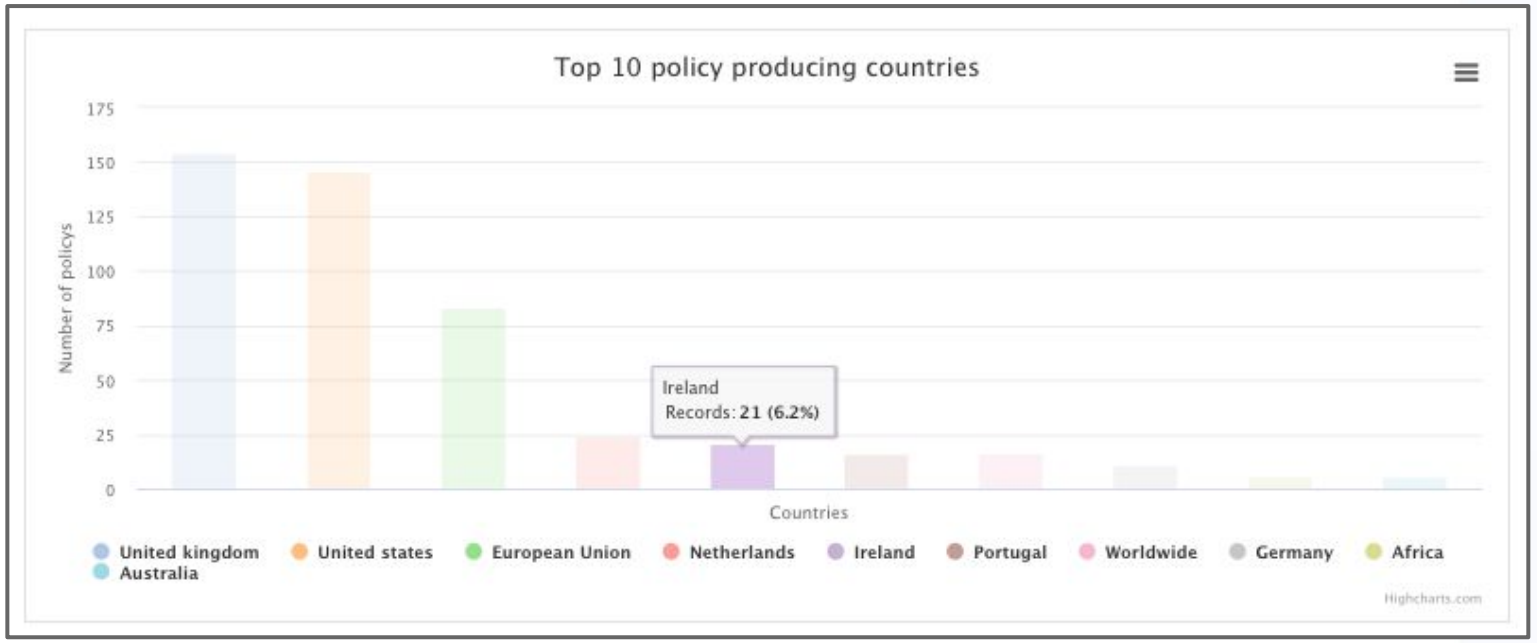
Participation:

FAIR for National Level Initiatives

Creating EOSC compliant interoperability policies

Testing the trustworthy and FAIR-enabling repositories prototype

Support action #2: improving the machine-actionability of data, policies and services



14th February 2025

What's love got to do with... data policies?

Panel questions

In what way was your participation in your specific support action useful?

Did you encounter any challenges?

Did your participation help you to progress some aspect of FAIR-enabling activity in your day to day practices?

Are there any tips that you would give to others who want to become more FAIR-enabling?

Any questions?

Join us at menti.com!

Use code 6927 7068

Funded by
the European Union


eosc
FAIR-IMPACT
Expanding FAIR solutions across EOSC
eosc
FAIRCORE4EOSC
Core Components Supporting a FAIR EOSC