Open Science Trails

Improving Digital Research Data Management: The OSTrails Project

14.11.2024

Daniel Spichtinger, University of Vienna

Funded by the European Union

Supporting eosc


OpenAIRE
ATHENA
Research & Innovation
Information Technologies
TU WIEN
POLITÉCNICA
Consiglio Nazionale
delle Ricerche
LifeWatch
ERIC
CTU
CZECH TECHNICAL
UNIVERSITY
IN PRAGUE
CITE
FECYT
FUNDACIÓN ESPAÑOLA
PARA LA CIENCIA
Y LA TECNOLOGÍA
universität
wien
CODEVENCE
TU
Graz
Universidade do
Minho
fct
Fundação
para a Ciência
e a Tecnologia
Sikt
ÖAW
ÖSTERREICHISCHE
AKADEMIE DER
WISSENSCHAFTEN
RWTH AACHEN
UNIVERSITY
UK Data Service
icm
FINNISH METEOROLOGICAL INSTITUTE
CSC
CLARIN
Common Language Resources and
Technology Infrastructure
UNIVERSITY OF
OXFORD
ESRF
SURF
SOCIB
Balearic Islands
Coastal Observing
and Forecasting System
ARISTOTLE
UNIVERSITY OF
THESSALONIKI
Tampere University
SOLEIL
SYNCHROTRON
OLLSCOIL TEICNEOLAÍOCHTA
BHAILE ÁTHA CLIATH
DUBLIN
TECHNOLOGICAL
UNIVERSITY DUBLIN
cessda
UNIVERSITY OF
BELGRADE
l'Observatoire
de Paris
PSNC
KNAW
Inria
CWTS
Meaningful metrics

Point of depature

- the **FAIR Principles** (Findability, Accessibility Interoperability and Reusability), have become *de facto* the global norm for good Research Data Management (RDM) practices, a prerequisite for replicability and reproducibility.
- However, the current fragmentation and inefficiency in managing data hinders the progress of research practices
- Against this background, the EU-funded OSTrails project (Grant agreement ID: 101130187) will work collaboratively with 40 partners on three pillars

ostrails
Funded by the European Union

DMP Platforms

Create and manage DMPs that support FAIR data practices.

- Argos (OpenDMP)
- DAMAP
- Data Stewardship Wizard (DSW)
- DMPTuuli
- Research Data Management Organiser (RDMO)

Scientific Knowledge Graphs

Enhance data interoperability and knowledge discovery.

- CESSDA Data Catalogue (CDC)
- CLARIN Virtual Language Observatory (CLARIN VLO)
- EURO-VO Registry
- FAIRsharing
- JERICO KG
- LifeBlock
- OpenAIRE Graph
- PANOSC Data Catalogue
- PUMA
- SOCIB KG
- SSH Open Marketplace

FAIR Assessment Tools

Evaluate and ensure the FAIRness of data and DMPs.

- CAT
- CESSDA Metadata Validator (CMV)
- CLARIN Curation Dashboard
- FAIR Aware
- FAIR Evaluator
- FAIR Research Objects Assessment (FAIROS)
- FAIR Validator
- FOOPS!
- IVOA services Validation
- PADC Validator
- SARA

Source: https://ostrails.eu/tools

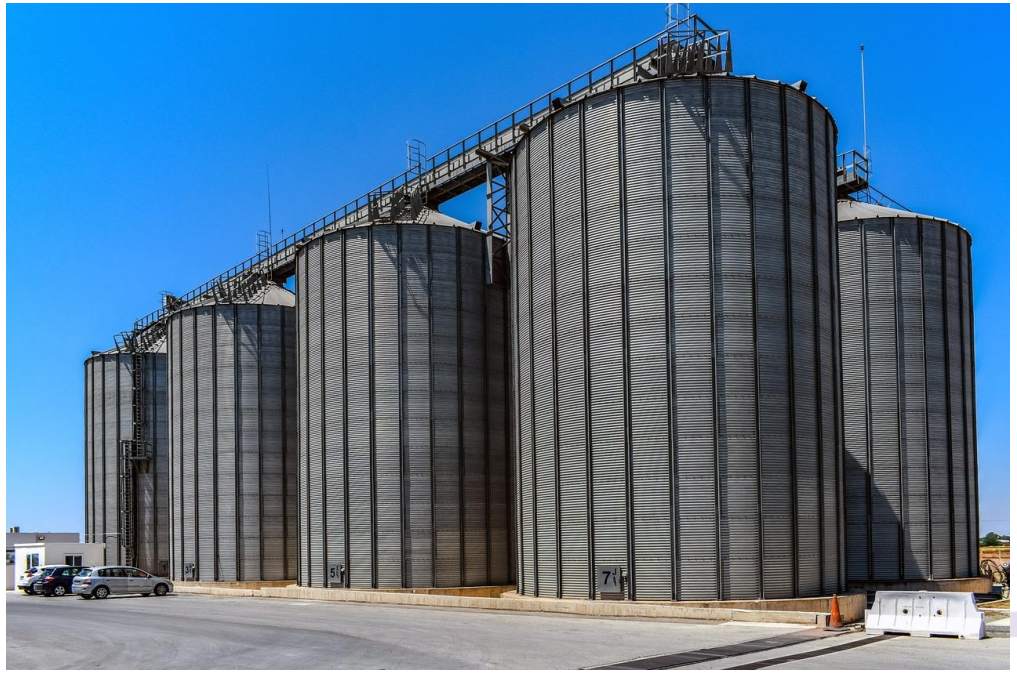
State of Play

- DMPs
  - DMPs themselves are usually not shared as FAIR outputs
  - Researchers feel extra burden to write the DMP
- SKGs are the talk of the town
  - Most useful in extracting knowledge, complementing AI/LLMs, and bibliometrics
  - Most SKGs are in their infancy with (i)quality issues related to metadata and relationships (ii) limited coverage for data, software, etc. ○ Isolated
- FAIR Assessment
  - funders, institutions, and infrastructures provide limited funding, support and guidance
  - Results are inconsistent among tools and/or assessments
  - Lack of a commonly agreed minimum set of guidance and direction to assisting researchers

ostrails
Funded by the European Union

Ambition: beyond silos

- **Data Management Plans (DMPs):** Must evolve into live, actionable tools to guide FAIR-by-design practices throughout the research lifecycle, ensuring continual updates and assessments of FAIRness.
- **Scientific Knowledge Graphs (SKGs):** To fully realize open science, SKGs need to transition from isolated units to interconnected networks, enabling more comprehensive research assessments beyond bibliometrics.
- **FAIR and FAIRness Assessments:** Transparent, consistent mechanisms are needed to evaluate FAIR compliance, shifting from judgement-based assessments to guidancedriven suppo



ostrails
Funded by the European Union

OSTrails: P-T-A framework

PLAN

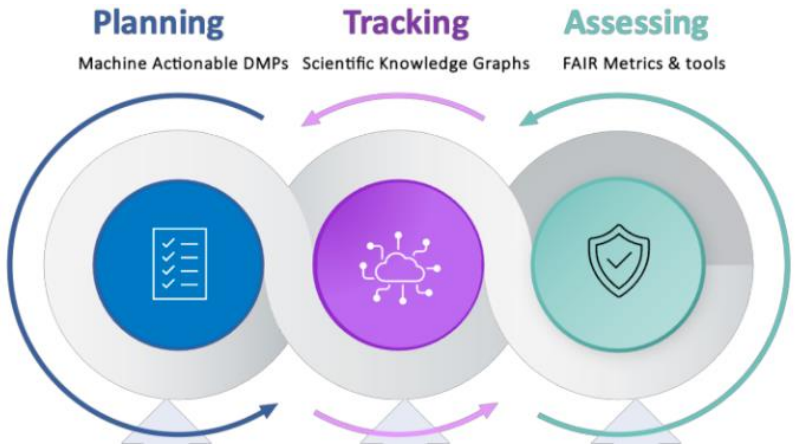
Increase the efficacy of DMPs and reach more researcher-centric, educative, and integrated “machine actionable” DMPs (maDMPs).

TRACK

Establish an open, interoperable and high quality SKG ecosystem of different types of research products, their relationships and metrics for evaluation.

ASSESS

Deliver modular and extendable FAIR tests, to make metrics “machine actionable”, complemented by user guidance.



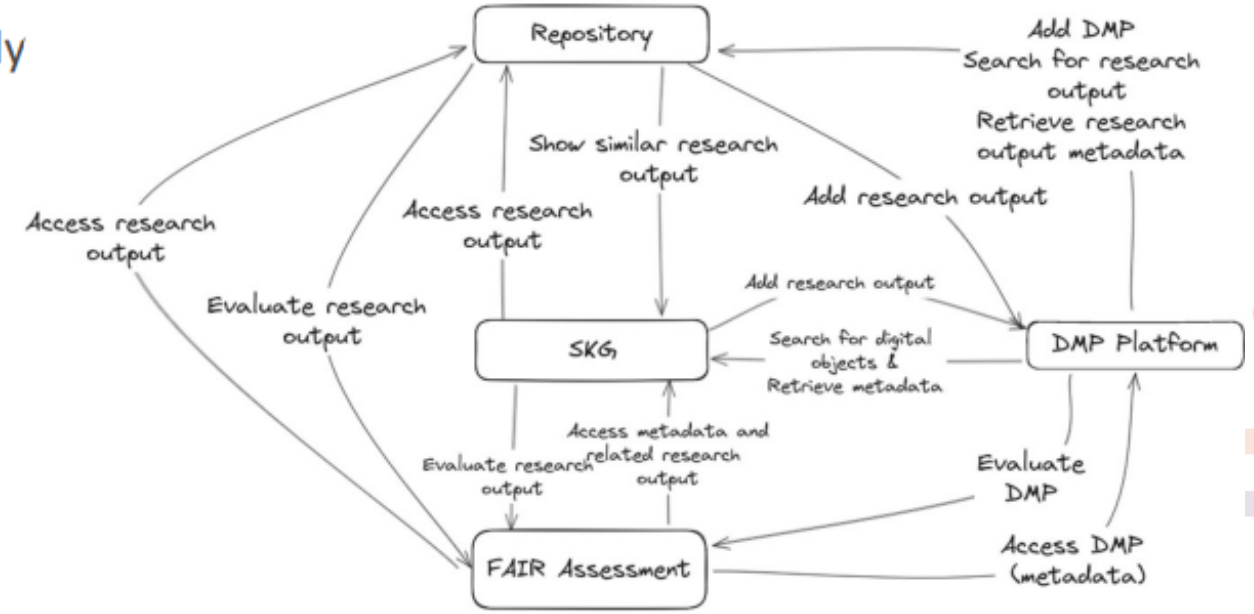
ostrails
Funded by the European Union

Only connect

OSTrails Pathways: User Actions

Generic diagram – User Actions

- Generic components are represented by the boxes
- Actions a user wants to perform actively are represented by the arrows
- Specific actions are represented by a label added to the arrows
- User groups include researchers, research managers, and research funders
- This diagram does not depict the processes in the background (i.e. what information is exchanged between the tools)



Source: D.1.1.

ostrails
Funded by the European Union

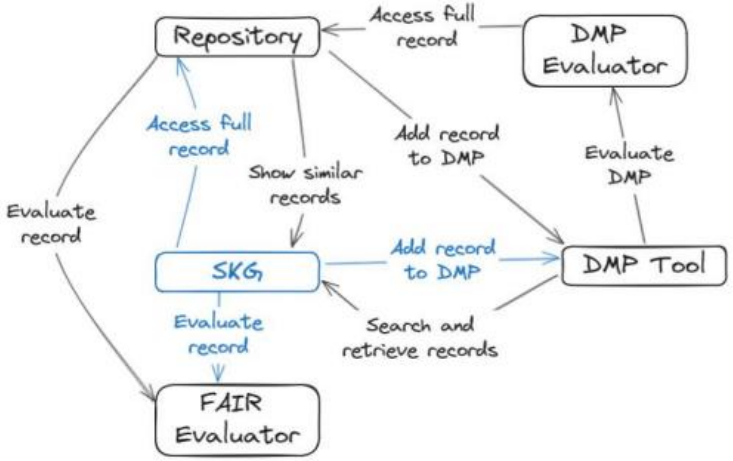


Figure 5: "Track" pathway

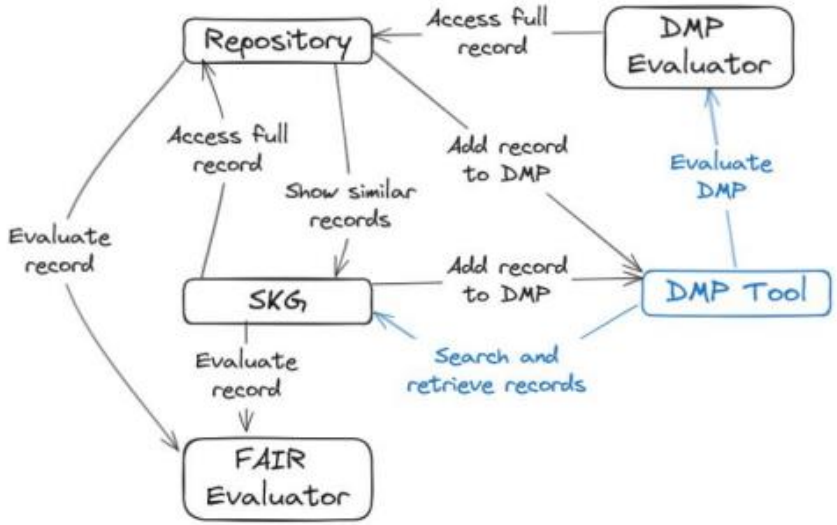


Figure 4: "Plan" Pathway

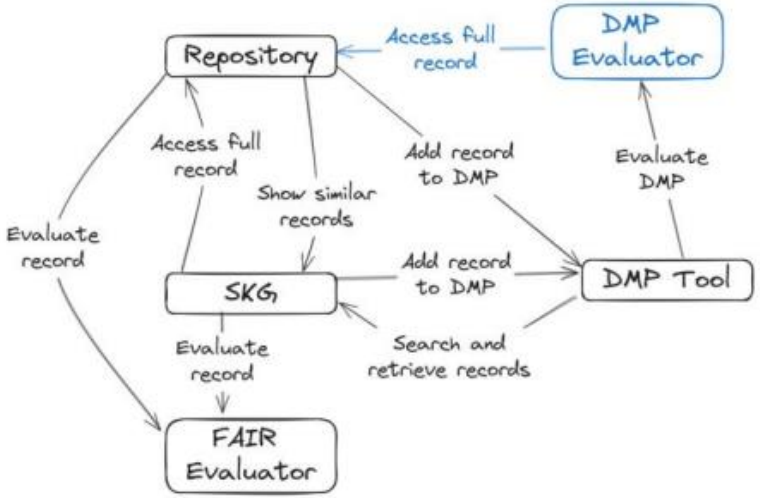


Figure 6: "Assess" pathway

Reichmann, S., Rey Mazón, M., Hasani-Mavriqi, I., Thaci, L., Eckhard, D., 2024. D1.1: Plan-Track-Assess Pathways. https://doi.org/10.5281/ZENODO.13145787

ostrails
Funded by the European Union

Key part: pilots

OSTrails project

PLAN: Increase the efficacy of DMPs

TRACK: Establish an open, interoperable and high quality SKG ecosystem

ASSESS: Deliver modular and extendable FAIR tests

National pilots #15

GR, PL, HR, NO, IR, RS, AT, ES, CZ, PT, NL, FR, FI, SE, DE

Thematic pilots #8

2 Multidisciplinary / Cross domain, 2 Social Sciences & Humanities, Language Resources, Biodiversity, Marine/Coastal, Physics, Astronomy and particle physics

Horizon Europe pilot

15+9+1 pilots

5 ESFRI clusters

Key part: pilots

Testing and Validation in Real-World Contexts

Diverse Contexts for Comprehensive Testing:

Case Studies and Proof-of-Concept:

Co-defining Metrics and Standards:

Community Engagement and Training:

Scaling and Sustainability:

Implementation

ostrails
Funded by the European Union

National pilots overview

1-Croatia Rudjer Boskovic Institute	2-Greece Hellenic Academic Libraries Link members, Athena RC	3-Netherlands SURF	4- Norway Sikt	5-Serbia U Belgrade
6- Austria- Uni Wien, TU Wien, Tu Graz	7-Poland PSNC, U Warsaw/ICM	8-Ireland TU Dublin	9-Portugal, UMinho & FCT	10-France INRIA
11-Finland CSC	12-Germany RWTH Aachen	13-Spain FECYT	14-Czech Republic, CTU	15-Sweden U Göteborg

& 9 thematic pilots (coordinated by LIFEWATCH)

ostrails
Funded by the European Union

National pilot use cases

I. develop maDMP templates to address national funders and local infrastructure (PIDs, CRIS, repositories)

II. extend repositories for archiving maDMPs and create links and relations (qualified references) between archived DMPs and publications;

III. interoperate with the OpenAIRE Graph or other SKGs to enhance qualified references with publications (incl. Pre-prints, open peer reviews), DMPs, maFAIRTests, data, workflows, instruments, software;

IV. co-define FAIR metrics for funders, institutions and research communities and assess deposited DO's and repositories;

V. co-define DMP evaluation criteria and assess DMPs quality according to national policy and criteria;

VI. extend national monitoring systems and funder's reporting systems with maDMPs.

Key messages from National pilot survey

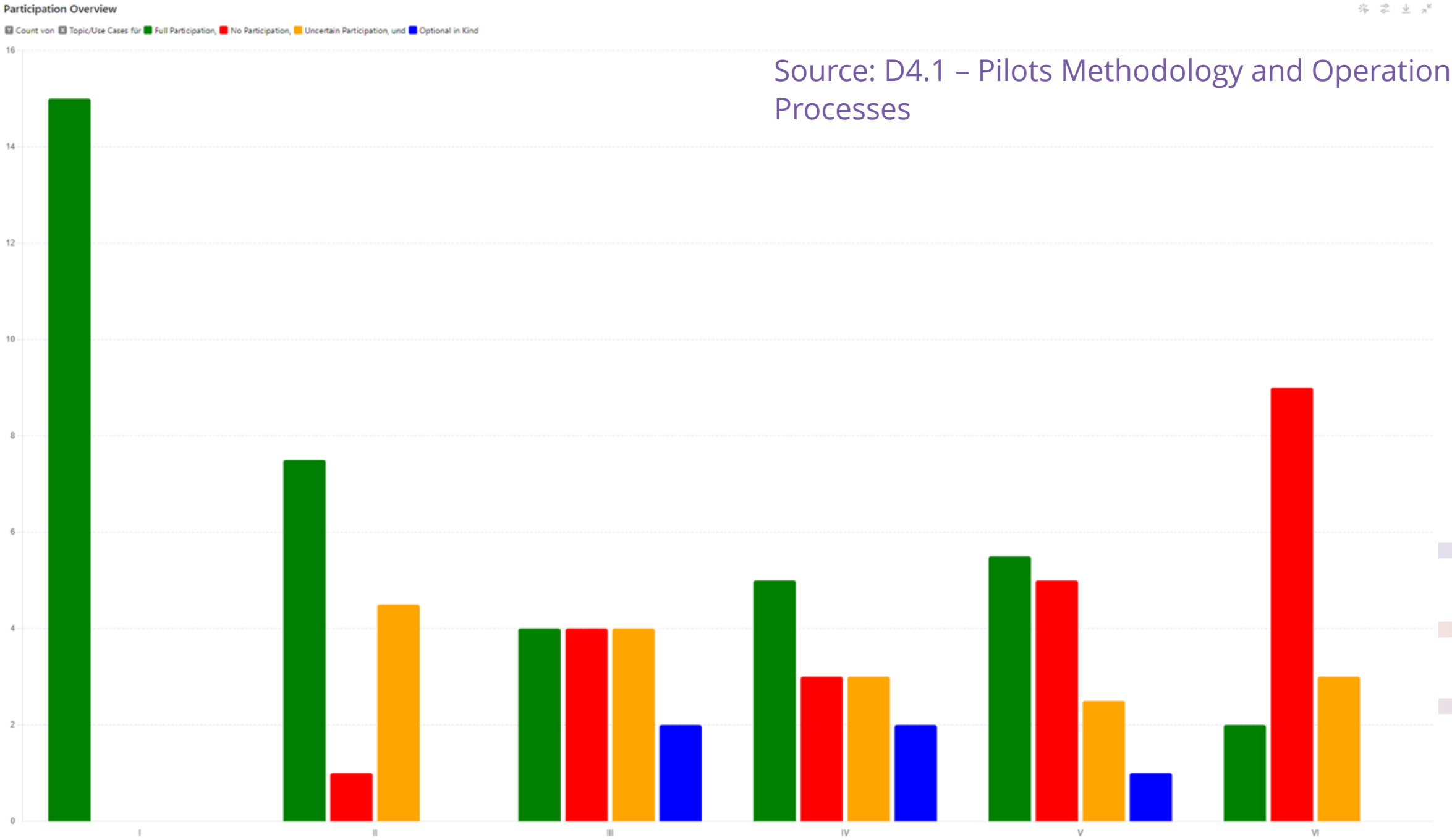
- There is substantial participation from national pilots, with multiple entries showing strong involvement (x)
- No participation is present but less frequent compared to full participation (o)
- Uncertain participation (?) is less frequent but present in some columns – this can still be refined
- There are some “optional in kind” contribution in use cases

National pilot survey

Table 1. National Pilot Participation Per Use Case.

Topic/Use Cases	Full Participation (x)	No Participation (o)	Uncertain Participation (?)	Optional in Kind	Other comments
I	15	0	0	0	0
II	7 to 8	1	4 to 5	0	1
III	4	4	4	2	1
IV	5	3	3	2	0
V	5 to 6	5	2 to 3	1	1
VI	2	9	3	0	2 for ?

Source: D4.1 – Pilots Methodology and Operational Processes



Source: D4.1 – Pilots Methodology and Operational Processes

Further results (summary)

- All national pilots will be developing maDMP templates tailored to national funder requirements and local infrastructure.
- Several pilots are working to integrate DMP tools with national infrastructures, including CRIS (Current Research Information Systems), repositories, and other data services.
- A core focus for pilots in Spain, Poland, France, and Germany is improving the FAIRness of research data. These pilots are developing or testing FAIR evaluation tools, establishing FAIR metrics, and assessing how to ensure data is accessible, interoperable, and reusable across different platforms and systems.
- Many pilots aim to increase collaboration between national stakeholders, such as universities, funders, and research infrastructures. By fostering better interoperability between systems, the pilots help to reduce the administrative burden on researchers and improve the quality and accessibility of research outputs.
- Several countries, including Serbia and Greece, are focused on enhancing repository and CRIS systems by linking DMPs with other research outputs, such as datasets and publications.
- Pilots across the project are actively testing and evaluating RDM tools and services. These include the implementation of maDMP tools like ARGOS, integration with lab notebooks and the FAIR assessment of digital objects in repositories.

Key part: pilots

OSTrails project

PLAN: Increase the efficacy of DMPs

TRACK: Establish an open, interoperable and high quality SKG ecosystem

ASSESS: Deliver modular and extendable FAIR tests

National pilots #15

GR, PL, HR, NO, IR, RS, AT, ES, CZ, PT, NL, FR, FI, SE, DE

Thematic pilots #8

2 Multidisciplinary / Cross domain, 2 Social Sciences & Humanities, Language Resources, Biodiversity, Marine/Coastal, Physics, Astronomy and particle physics

Horizon Europe pilot

15+9+1 pilots

5 ESFRI clusters

Thematic Pilots

The thematic pilots within the OSTrails project are performed by research infrastructures representing all **five ESFRI Science Clusters**.

- **T1 - FAIRsharing for ESFRI Science Clusters. -** Focuses on enhancing FAIRsharing as a collaborative resource for community standards, databases, and policies. The pilot aims to integrate FAIRsharing content with OSTrails components for improved interoperability across science clusters.

- **T2 - PaNOSC / ESRF, SOLEIL, ILL.-** Works on automating the generation of machine-actionable DMPs (maDMPs) within the Photon and Neutron communities. The pilot will contribute to the transformation of DMPs into dynamic resources aligned with FAIR principles.

- **T3 - ENVRI / JERICO. -** Extends the JERICO e-infrastructure by integrating DMPs and digital objects. It focuses on developing an interoperable schema for SKGs and enhancing data monitoring and reporting within marine research.

- **T5 - SSHOC/CESSDA. -** Aims to improve the interoperability and FAIRness of social science datasets by adapting the CESSDA Vocabulary Service and European Language Social Science Thesaurus to the DMP Interoperability Framework. The pilot will also integrate FAIR assessment tools into CESSDA services.

- **T6 - SSHOC / CLARIN / OEAW. -** Focuses on expanding the CLARIN central catalogue into a Scientific Knowledge Graph (SKG) and integrating it with FAIR assessment tools. This pilot aligns with OSTrails' objectives of enhancing metadata enrichment and data interoperability within the humanities and social sciences.

- **T7 - ENVRI / LifeWatch ERIC. -** Integrates multiple environmental data sources into federated searches, emphasizing semantic integration with LifeBlock and the development of SKGs. The pilot also incorporates maDMPs to automate and improve data management practices in the environmental domain.

- **T8 - SSHOC / CLARIN. -** Extends existing data catalogues into SKGs and enhances interoperability within OSTrails. The pilot works on aligning data models and vocabularies, contributing to the improvement of data management and FAIR assessment in the social sciences and humanities.

- **T9 - ESCAPE / ObsParis. -** Focuses on astronomy data management, specifically with the MASER portal and Cherenkov Telescope Array Observatory (CTAO). The pilot aims to adopt maDMP tools, integrate with SKGs, and implement FAIR assessment criteria tailored to astronomy datasets.

Horizon Europe Pilot

The Horizon Europe pilot positions the adoption of OSTrails results at larger scale to support project officers in their duties with respect to evaluating DMPs and to provide better tools and workflows for beneficiaries to implement data management practices according to Horizon Europe conditions, aiming to

- Develop a machine actionable template for Horizon Europe DMPs.
- Embed FAIR metrics in the machine actionable template to assess the FAIRness of their described outputs and activities.
- Embed DMP metrics in the machine actionable template for DMPs to be evaluated in a semi-automated way.
- Enhance the machine actionable template with semantics so that qualified references of DMPs with other research outputs can be created.
- Provide maDMPs (.json) along with the .pdf version of the project's DMP deliverables.

ostrails
Funded by the European Union

Conclusions/Summary

- OSTrails aims to overcome fragmentation through the PTA framework, working on maDMPs, SKG and FAIRness assessment
- The national, thematic, and Horizon Europe pilots play a critical role in the overall success of OSTrails by showcasing a diverse array of research environments, each with distinct challenges, goals, and infrastructure capabilities. This reflects the broad spectrum of Research Data Management (RDM) practices across Europe and various scientific disciplines.
- National and Thematic Pilots are focusing on 6 distinct use-cases/topics The Horizon Europe pilot scales adoption of OSTrails results to support the EC RDM policy in Horizon Europe.
- The pilots thus contribute to building a more integrated and FAIR-compliant research ecosystem.

The Uni Wien Team

- Gerda McNeill
- Daniel Spichtinger
- Susanne Blumesberger (PHAIDRA)
- Dominik Denk (PHAIDRA)

- Lars Kaczmirek (AUSSDA)
- Lisa Hirsch (AUSSDA)

ostrails
Funded by the European Union

Contact and Information

- Uni Wien Team (WP 4, National Pilot)
  daniel.spichtinger@univie.ac.at
- Website: https://ostrails.eu/
  - https://ostrails.eu/deliverables
  - https://ostrails.eu/tools (DMP

ostrails Funded by the European Union