



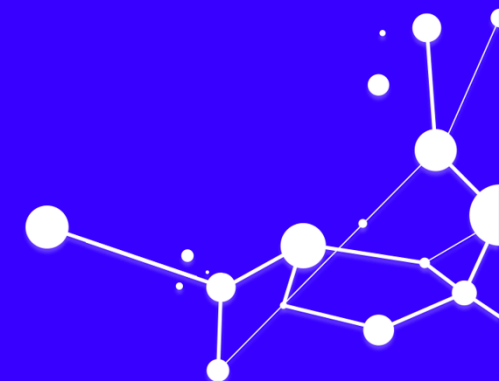
Elli Papadopoulou
ATHENA Research Center



Connecting research data workflows with Argos:
Embedding services and scientific outputs in data
management plans



About Argos



Overview



an **open source**, **configurable** and **extensible** tool for **planning Research Data Management (RDM)** activities according to **Open Access & FAIR** data policies.

- **Website:** <https://argos.openaire.eu/>
~2000 users (growing)

✓ Online or ad-hoc installation
✓ Free to use for researchers

Title: H2020 NEANIAS WP4 Space Astrophysics Datasets

Template: Horizon 2020

External References

Data Repositories

GAIA Catalogues, VIALACTEA Knowledge Base, SKA Science Data Challenges

External Datasets

Galactic Legacy Infrared Mid-Plane Survey Extraordinaire, 24 and 70 Micron Survey of the Inner Galactic Disk with MIPS, MIPS GAL Catalog, Co-Ordinated Radio 'N' Infrared Survey for High-mass star formation, Wide-field Infrared Survey Explorer, SKA Science Data Challenge #1, Canadian Galactic Plane Survey, BGPS Catalog, Southern Galactic Plane Survey, WISE Catalog, H2O southern Galactic Plane Survey, Three-mm Ultimate Mopra Milky Way Survey, GAIA Archive, VLA Sky Survey, ATLASGAL Catalog, The VLA Galactic Plane Survey, The Galactic Census of High and Medium-mass Protostars, James Clerk Maxwell Telescope, Herschel Infrared Galactic Plane Survey, The Multi-Array Galactic Plane Imaging Survey, Millimetre Astronomy Legacy Team 90GHz

Registries

VLKB Registries, GAIA Registries

Services

GAIA Services, VLKB Services

The H2020 NEANIAS Project aims to design, deliver, and integrate into the European Open Science Cloud innovative thematic services, derived from state-of-the-art research assets and practices in three major sectors: underwater research, atmospheric research and space research. In particular the Work Package 4 is focused on the thematic services related to the Space environment. The Space environment comprises astrophysicists and planetary scientists that will handle data within three services: S1 - FAIR Data Management and visualization for complex data and metadata service; S2- Map making and mosaicking for multidimensional images service; S3 - Structure detection on large map images with machine learning techniques service. This document describes datasets employed from astrophysics surveys mainly covering the Galactic Plane. Data holdings come partly from the ViaLactea Knowledge Base (VLKB) hosted at the Italian center for Astronomical Archives (IA2, hosted at INAF - Astronomical Observatory of Trieste) and partly from new data resources such as the GAIA Archive and SKA Data Challenges.

Dataset Description

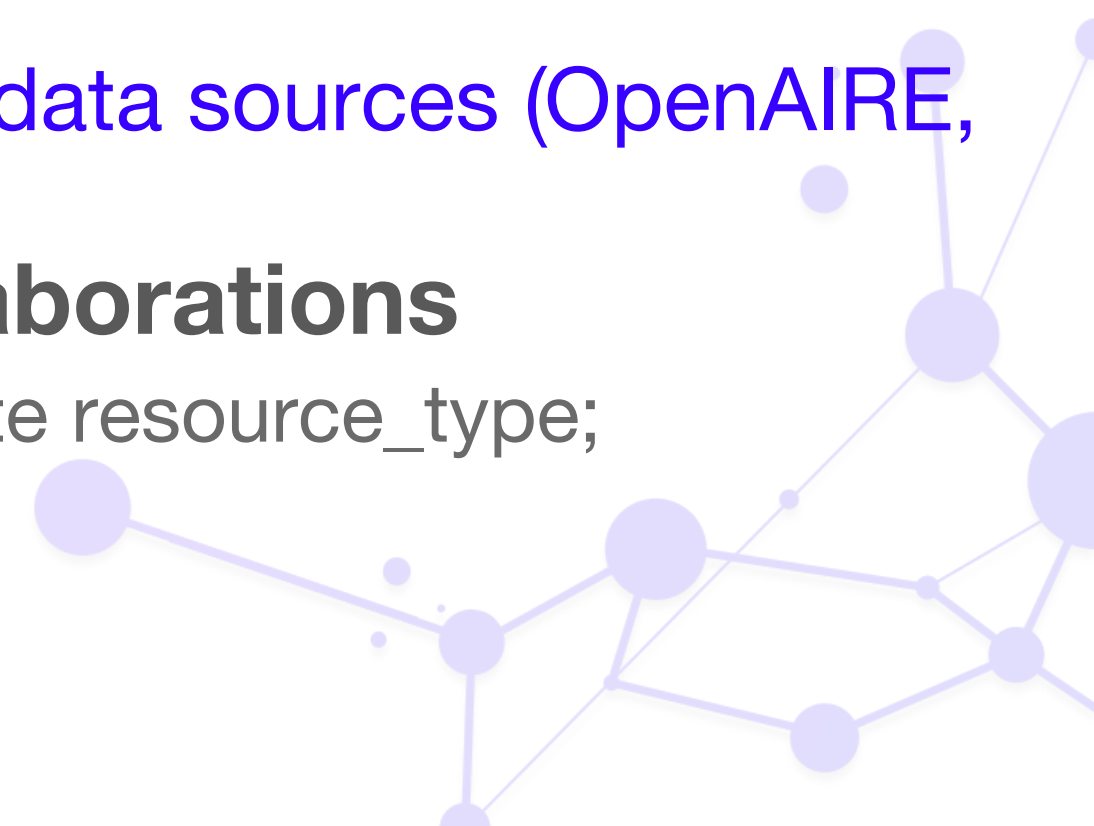
1 Data Summary

1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

The H2020 NEANIAS Project aims to design, deliver, and integrate into the European Open Science Cloud innovative thematic services, derived from state-of-the-art research assets and practices

– not just a tool!

- **Full DMP Lifecycle**
 - generate & publish DMPs according to Open and FAIR principles
- **Machine actionable DMP (ma-DMP) outputs**
- **Data Domain Protocols**
 - create many dataset profiles in a single DMP
 - -> e.g. new vs re-used vs sensitive vs discipline specific
- **Contextualized and exploitable DMP data**
 - connect with and enhance reference services and data sources (OpenAIRE, EOSC, etc).
- **Standardization of global practices and collaborations**
 - -> e.g. DMPs exposed in repositories with appropriate resource_type; connecting with data availability statements

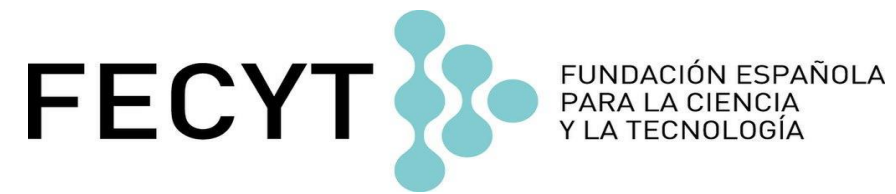


ARGOS for

Researchers



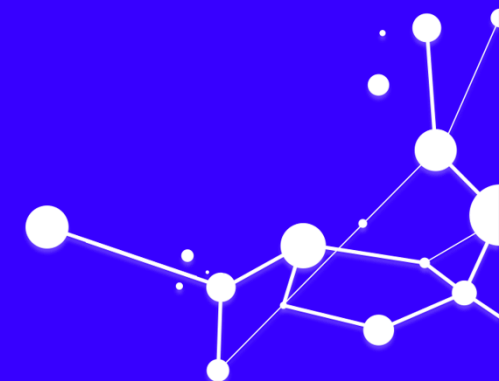
Funders & institutions



Research communities

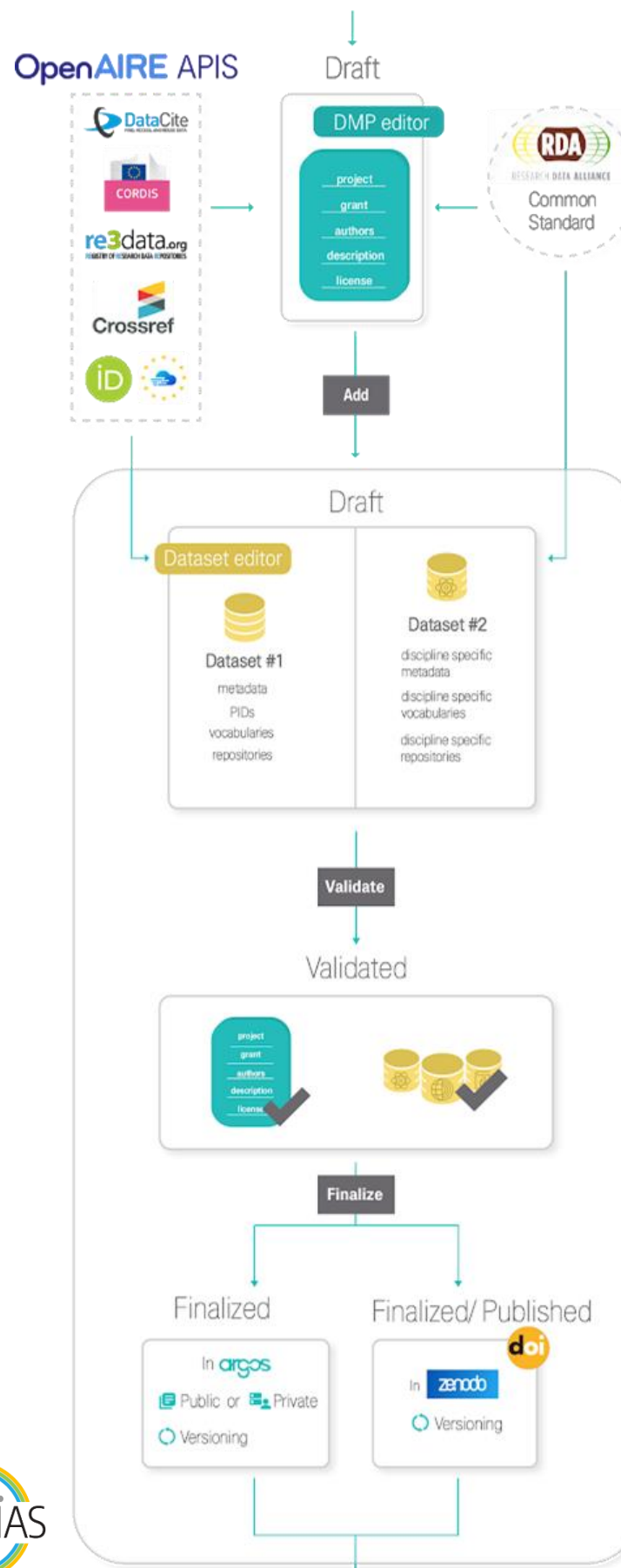


Why Argos



Write & publish

- Full DMP publication
- Two editors to create DMPs
 - Full DMP editor
 - Dataset wizard
- First validation
- Finalization
- ma-DMP outputs



- Rich documents
- Discoverable through OpenAIRE
- Versioned (provenance)
- Accessible: PIDs (ORCID & DOIs)
- Reusable: Licenses
- Preserved in Zenodo

Create dataset profiles

- Different data questions
- Configurable APIs
- Tailored instructions



The screenshot shows a DMP (Data Management Plan) interface. At the top, there is a teal 'DMP' button and the title 'Merge all datasets in a DMP'. Below this, the 'Owner' is listed as 'Version 0' and the 'Edited' date is '15 September 2021'. There are three teal circular icons: a pencil (edit), a document (copy), and a trash can (delete). The 'Grant' section is titled '4th European Symposium on Aerobiology 12.-16.8.2008'. Below the grant title, there are sections for 'Researchers' (indicated by a dash), 'Description' (indicated by a dash), and 'Datasets used'. The 'Datasets used' section contains three yellow buttons: 'First Template: Reused Data', 'third Dataset: Other Research Outputs', and 'Second Dataset: New Data', each with a right-pointing arrow icon. At the bottom, there is a '+ Add Dataset' button.

Design machine-actionable templates

- Many inputs to create a question
- Long list of input types
 - Boolean, Multiple choice, Free text, custom APIs...
- Collection of static APIs
- Conditional questions
- Multiplicity
 - # the question can be answered with different input
- RDA compatibility

4.1.4 What is the origin / provenance of the dataset?

Description

Required

Select RDA Common Standards

Default Value

Make Conditional Question

Conditional Questions

1 If Value is then show Question

Other If other, please specify 8695addc-0f4f-9a92-c90...

Word List Data

Multiple Selection

Input Placeholder Text

Select

Label: Primary data <input type="text"/>	Value: Primary data <input type="text"/>
Label: Secondary data <input type="text"/>	Value: Secondary data <input type="text"/>
Label: Other <input type="text"/>	Value: Other <input type="text"/>

+

Required

Free Text RDA Common Standards

Default Value

Make Conditional Question

Free Text Data

Input Placeholder Text

If other, please specify

Preview

What is the origin / provenance of the dataset?

Select *

If other, please specify

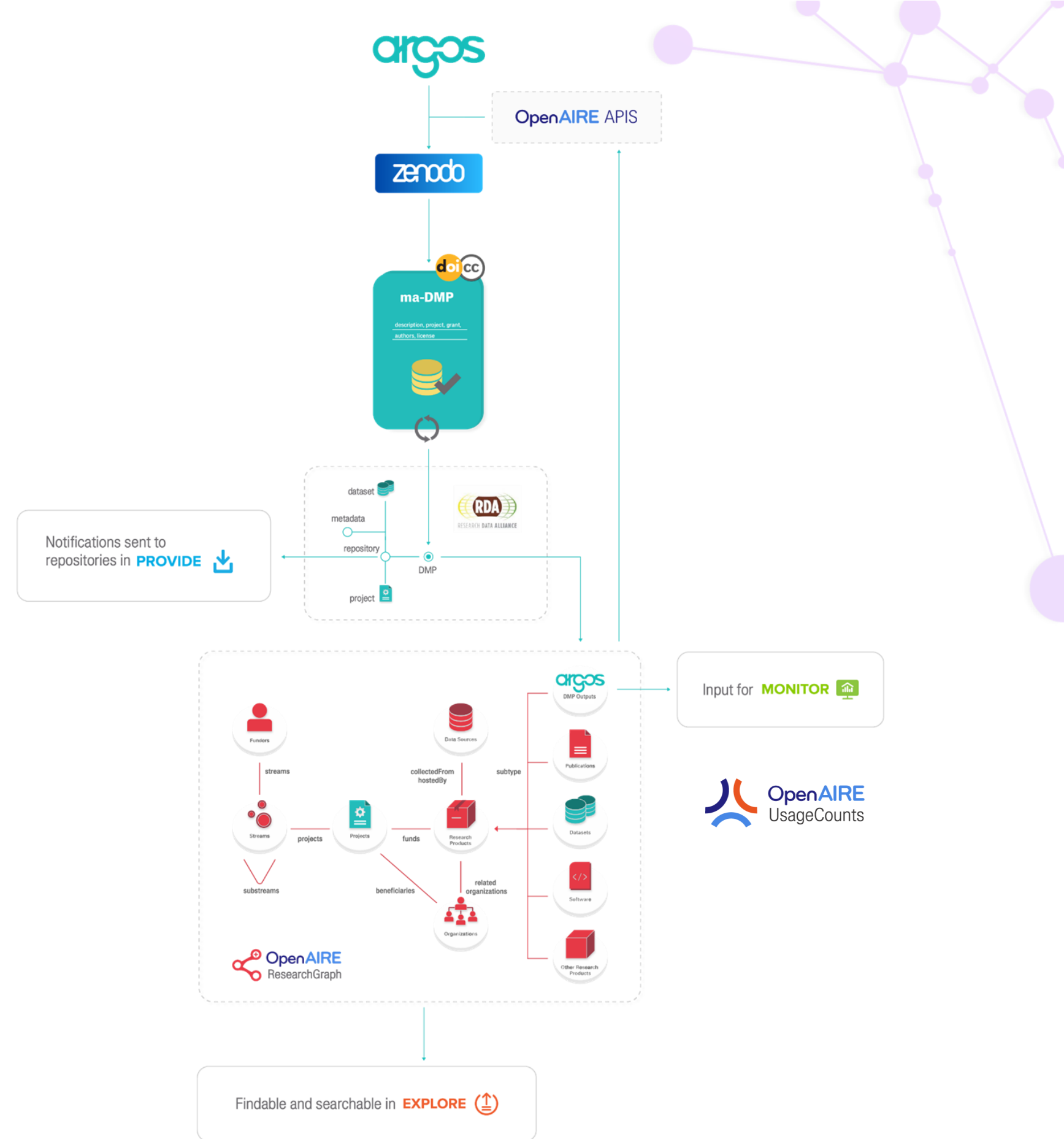
Please Specify

Provide additional information or justification about your selection



Connect workflows

- Benefit from OpenAIRE pool of data
- Connect to OpenAIRE services
- Connect to own services
- Personalized instances



Connect workflows

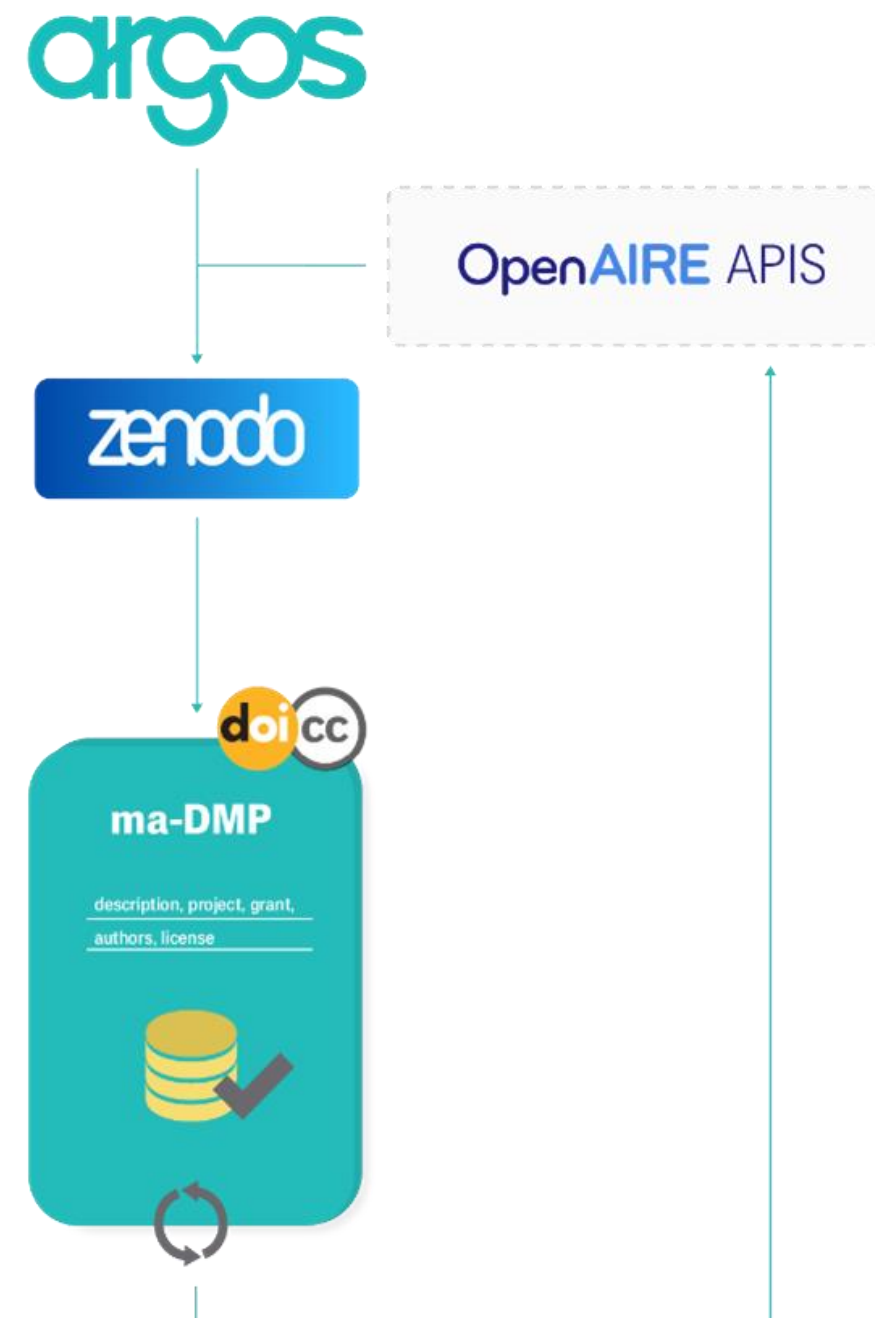
- Create subsets of Argos in OpenAIRE that meet your needs.
 - Exploit OpenAIRE and EOSC APIs in templates.
 - Configure own APIs

argos



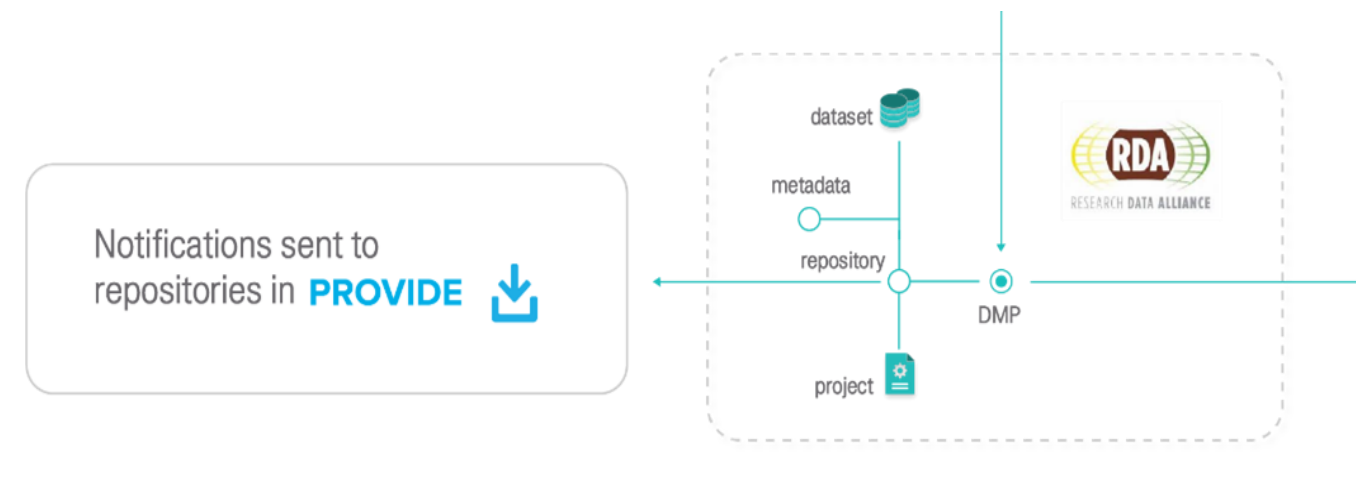
Connect workflows

- **Publish in Zenodo.**
 - Or integrate with own repositories
 - Integrations to enable pre-filling of DMP templates for re-used data



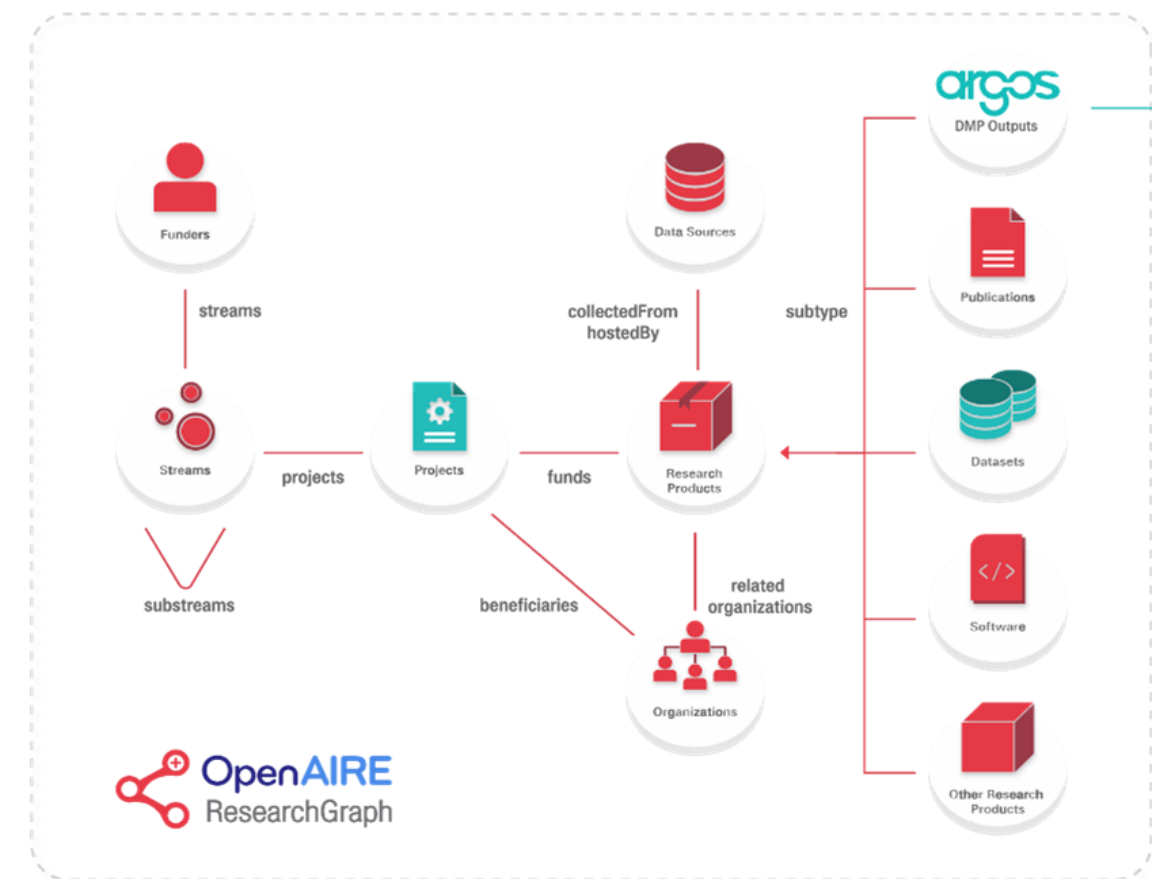
Connect workflows

- Notify repository managers for new datasets.



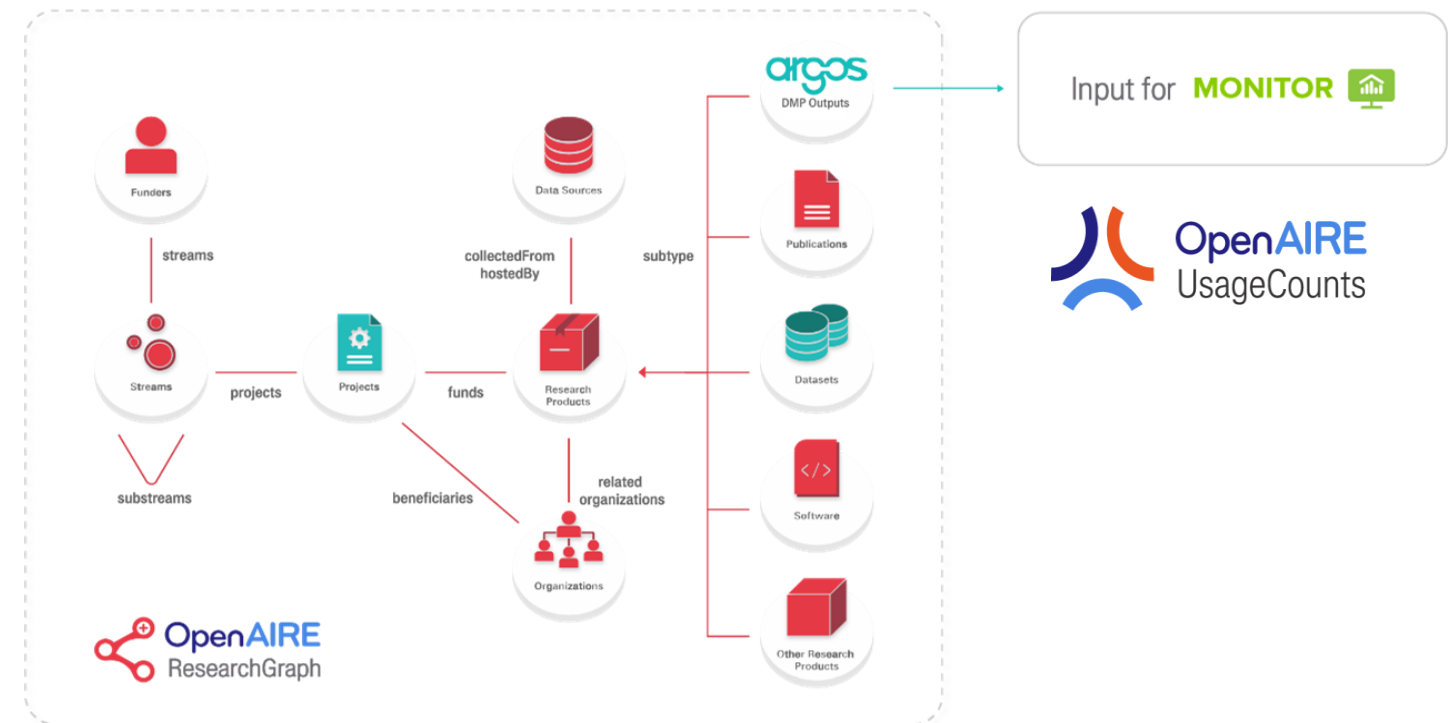
Connect workflows

- **Exploit DMP entities in the Research Graph.**
 - Create links between outputs and entities. -> Now: DMPs, data and projects.



Connect workflows

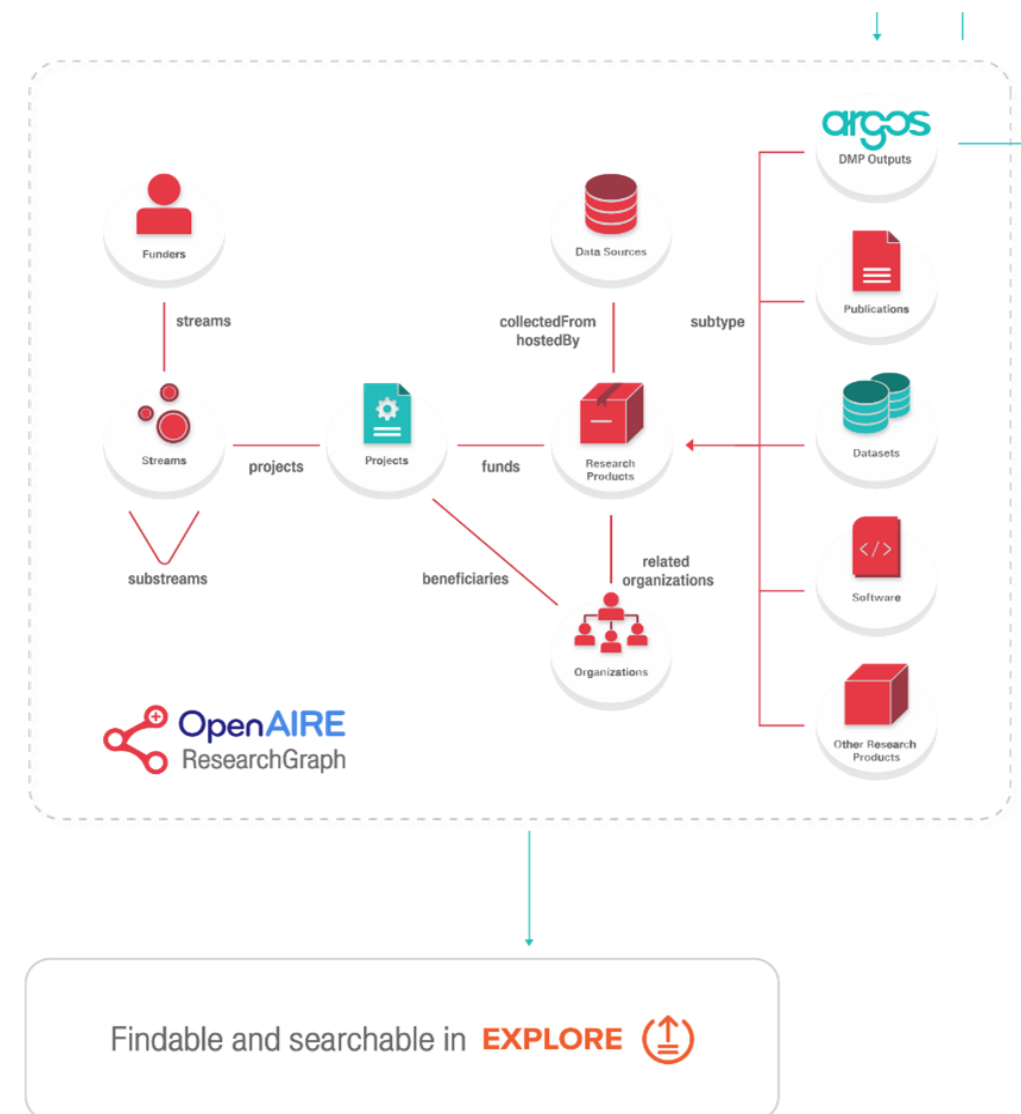
- Combine with validated OpenAIRE data and provide statistics.
 - Define indicators. Add to dashboards.



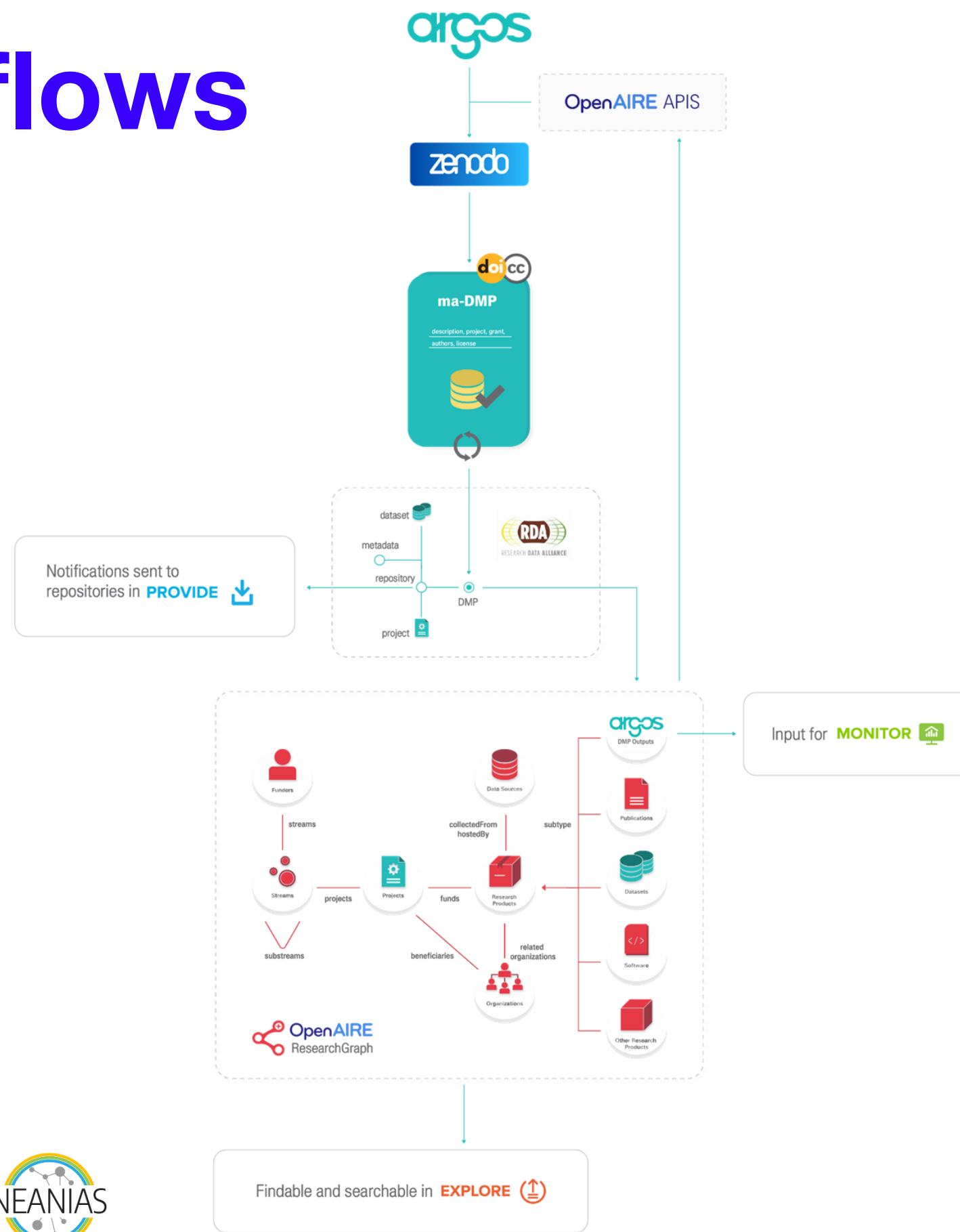
OpenAIRE
UsageCounts

Connect workflows

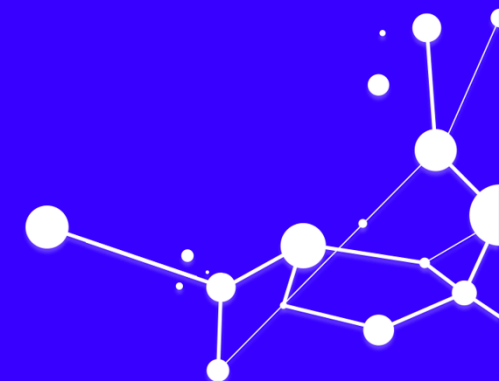
- Discover DMPs
- DMPs under the project's page.



Connect workflows



The way forward



Pre-filling of dataset information – Open invitation

Publishing

E.g. Re-used data

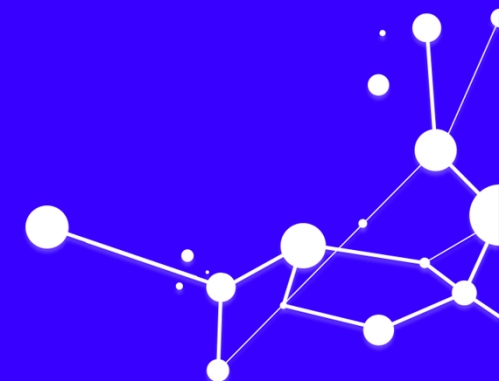
- Related ids with other outputs: publications, data, software, etc

Processing

E.g. New data, other research outputs (methods, workflows)

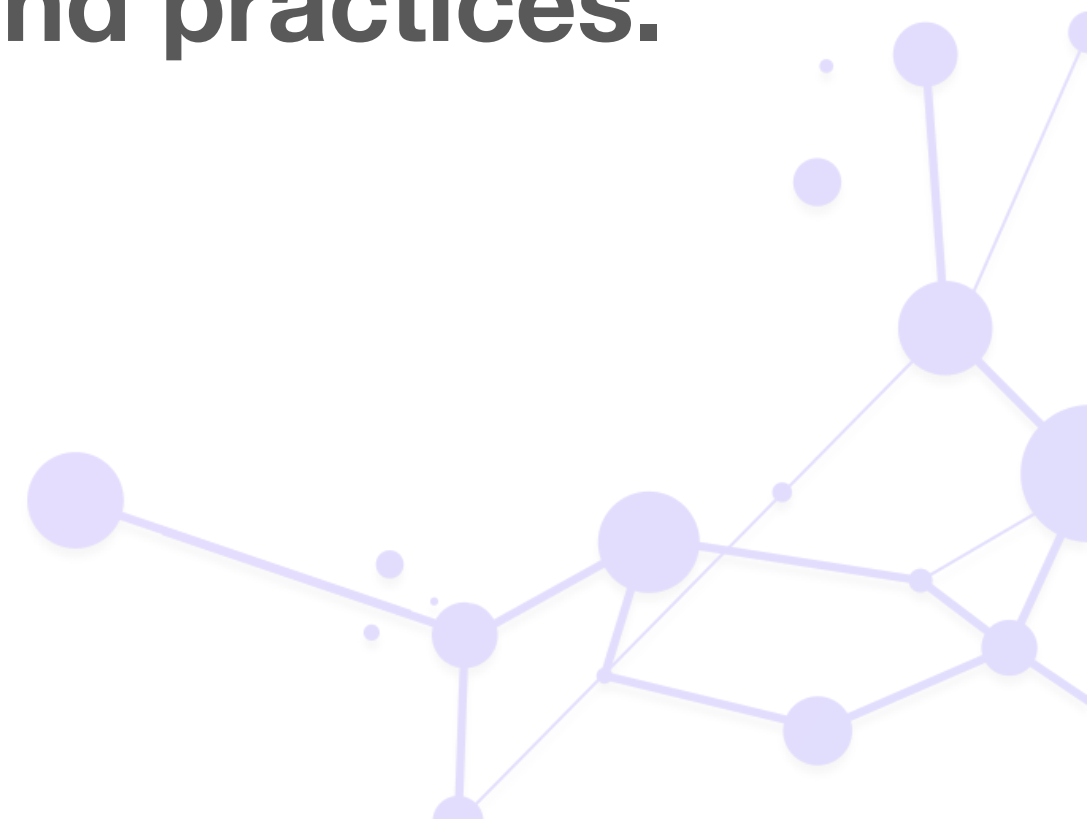
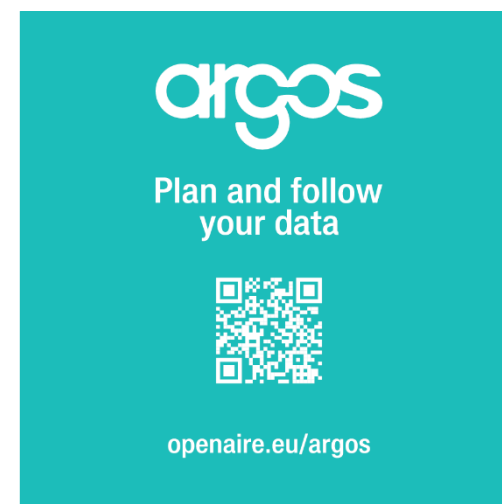
- Links with services and instruments: online storage, notebooks, etc

Take away



Take away

- ARGOS **prepares** all stakeholders for the next Horizon Europe DMPs requirements.
- ARGOS **simplifies** administrative processes and **connects** with University / Institutional workflows.
- ARGOS **enables** research communities to create templates tailored to domain standards and practices.





@openaire_eu
#OpenAIRE-Nexus

THANK YOU

Eli Papadopoulou | Georgios Kakaletris
elli.p@athenarc.gr | gkakas@athenarc.gr

