

www.openaire.eu



ARGOS

FAIR Data Management and DMP publications











www.openaire.eu

CONTENTS



OA and FAIR data in DMPs

Why RDM & DMPs: components and misconceptions



The OpenAIRE service for writing and publishing DMPs

ARGOS features and functionalities, data model and architecture



Added value

Why choosing ARGOS









DEPENDENCE OF A LOCAL DESCRIPTION OF A LOCAL







RDM in relation to policies

Funding Programmes conditions



2013: Open Data Pilot

2017: OA data by default (mandate)

Policy rules and recommendations

34/12

EN

Official Journal of the European Union

31 5 2018

RECOMMENDATIONS

of 25 April 2018

on access to and preservation of scientific information



European Open Science Cloud







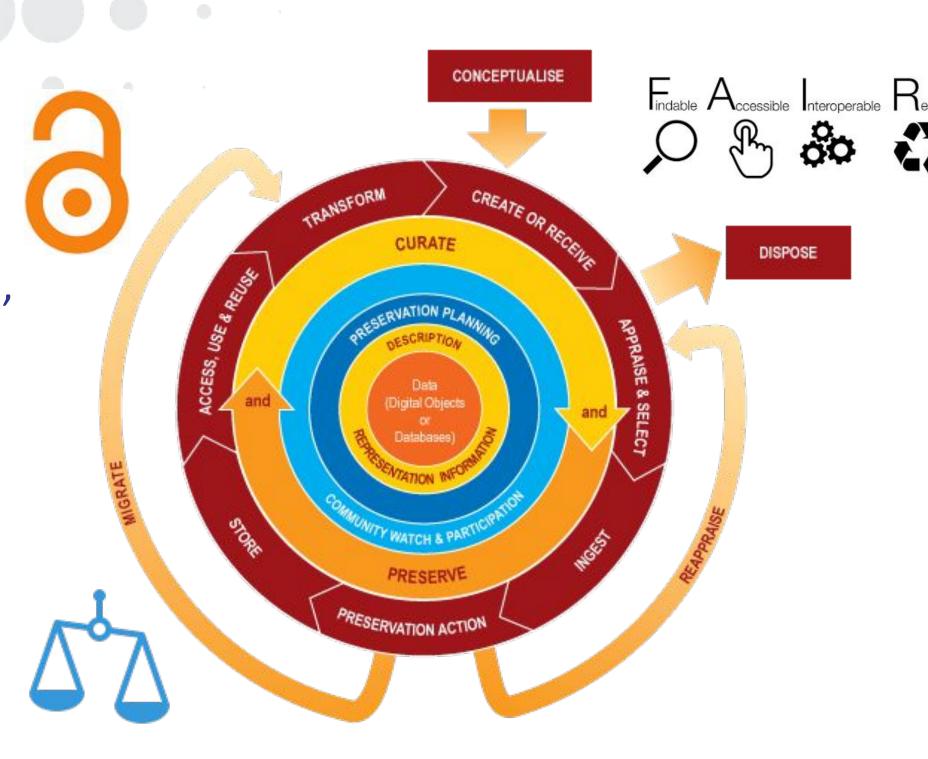
RDM elements

- Data collection, data processing, data preservation
- Indicative processes: ingestion, aggregation, analysis, classification, metadata enrichment, organisation, validation, storing, etc.
- There might be re-processing of data (e.g. data migration)
- Data disposal
- Data sharing/ publication









Data Management Plan (DMP)

- A formal declaration on how an activity's datasets are to be handled throughout activity's lifetime and on what terms those datasets will survive the activity.
 - The activity may be a project that may span several grants or a project that coincides a grant issued by a funder.
 - Datasets addressed by a DMP may be interim results,
 outputs or even inputs of the activity.
 - DMP addressed topics such as data description, data depositing and preservation, data identification, data licensing, data access etc



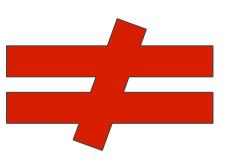




What is a DMP?

Deliverable and "living" document

- documents processes undertaken throughout data management lifecycle, including costs



What is not a DMP?

Research assessment method







What data does it cover?

- Scientific data and metadata
 - ✓ Useful for verifying / validating conclusions

- ✓ That does not infringe copyright
- ✓ Non-personal and / or sensitive content

Access restrictions

- Other raw or structured data and metadata
 - ✓ Guidelines for data reproducibility
 - ✓ Including an analysis tool / software

Restrictions in access are described in the Data Management Plan in advance







When is it delivered?

Proposal stage (not necessarily)

- What kind of data will the project create / collect?
- Which standards will be used?
- How will this data be shared / made available? If not, why not?
- How will this data be curated and preserved?

Project stage

- first version of the DMP within the first semester
- revisions







Who needs DMPs?









Organisations

- To track R&D outputs and any incurring bindings to them
- To identify the consumed data implications
- To facilitate R&D via data discovery and sharing

Funders

- To track direct and indirect products and impact of funding
- To identify or refine strategies w.r.t. research and data production/sharing/r euse

Researchers

 To facilitate and even enforce data referencing

Open Science

- To promote FAIRness
- To promote interdisciplinary research







Who is involved in DMPs

- Funders (offices)
 - Define rules for DMP
- Organisations (policy offices)
 - Define policies
- Projects (managers)
 - Apply project policy aligned to Organisational and Funding policies

- Data Managers
 - Manage the DMP
- Researchers
 - Describe datasets
 - Are attributed by datasets







H2020 FAIR DMP template 1/2

SUMMARY TABLE 1

FAIR Data Management at a glance: issues to cover in your Horizon 2020 DMP

This table provides a summary of the Data Management Plan (DMP) issues to be addressed, as outlined above.

DMP component	Issues to be addressed
1. Data summary	State the purpose of the data collection/generation
	Explain the relation to the objectives of the project.
	Specify the types and formats of data generated/collected
	Specify if existing data is being re-used (if any)
	Specify the origin of the data
	State the expected size of the data (if known)
	Outline the data utility: to whom will it be useful
2. FAIR Data	Outline the discoverability of data (metadata provision)
2.1. Making data findable, including provisions for metadata	 Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?
	Outline naming conventions used
	Outline the approach towards search keyword
	Outline the approach for clear versioning
	 Specify standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how.







H2020 FAIR DMP template 2/2

2.2 Making data openly accessible	 Specify which data will be made openly available? If some data is kept closed provide rationale for doing so
	Specify how the data will be made available
	 Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?
	 Specify where the data and associated metadata, documentation and code are deposited
	 Specify how access will be provided in case there are any restrictions
2.3. Making data interoperable	 Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.
	 Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?
2.4. Increase data re-use (through clarifying licences)	Specify how the data will be licenced to permit the widest reuse possible.
	 Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed
	 Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why
	Describe data quality assurance processes
	Specify the length of time for which the data will remain re-usable
3. Allocation of resources	 Estimate the costs for making your data FAIR. Describe how you intend to cover these costs
	Clearly identify responsibilities for data management in your project
	Describe costs and potential value of long term preservation
4. Data security	Address data recovery as well as secure storage and transfer of sensitive data
5. Ethical aspects	 To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former.
6. Other	Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)









Write and publish your DIMPs







Context



ARGOS: A service for machine actionable data management planning

Jointly designed and developed by OpenAIRE and EUDAT

- Launched implementation in 2017.
 - Currently at Release Candidate version with long testing history
 - Updates roll out on a monthly basis







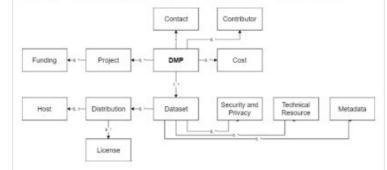
RDA DMP Common Standard for machine-actionable Data Management Plans

About this document

This is a metadata application profile to provide basic interoperability between systems producing or consuming machine-actionable data management plans (maDMPS). Further fields can be added in specific deployments, but they do not guarantee interoperability. DMP tools can use any other fields in their internal data models.

This application profile is intended to cover a wide range of use cases and does not set any business (e.g. funder specific) requirements. I represents information over the whole DMP lifecycle.

For more information see examples, FAQ and useful links to consultations, documents, tools, prototypes, etc. developed by the working group.



DMP

Provides high level information about the DMP, e.g. its title, modification date, etc. It is the root of this application profile. The majority of its fields are mandatory.

Project

Describes the project associated with the DMP, if applicable. It can be used to describe any type of project: that is, not only funded projects, but also internal projects, PhD theses, etc.

Funding

For specifying details on funded projects, e.g. NSF of EC funded projects.

Contact

Specifies the party which can provide any information on the DMP. This is

Structure

- dmp
 - contact
 - contact_id
 - identifier
 - type
 - mbox
 - name
 - contributor
 - contributor_id
 - identifier
 - type
 - mbox
 - name
 - role
 - o cost
 - currency_code
 - description
 - title
 - value
 - created
 - dataset
 - data_quality_assurance
 - dataset_id
 - identifier
 - type
 - description
 - distribution
 - access_url
 - available until
 - byte_size
 - data_access
 - description
 - download_url
 - formathost
 - availability

https://bit.ly/2t5jr3e





RDA DMP Common Standard

Incorporates global standards



Concepts

Data Management Plan

- is a collection of dataset descriptions, associated with an activity ("project" or "grant")
- may be versioned
- is exportable to various forms, (.pdf, .xml, .json)

Dataset description

is the description of a dataset according to a set of rules that is the dataset profile (aka template)

- Dataset profile (aka Template)
 is the set of rules that describe what dataset descriptions contain and how they are handled
- contains attributes/fields, behavioral rules (e.g. toggling visibility of fields) and validation rules
- are linked to DMPs for focusing users on specific formats of descriptions

Project & Grant

logical entities that define the context where one or more data management plans may be formed in

Users

- associated to DMPs and other entities for authorization
- may map to researchers

Secondary entities

- **Repository**: repositories where dataset resides on
- **Dataset**: external linked datasets (e.g. reused for dataset generation)
- **Registry**: catalogues listing the dataset
- **Service**: services manipulating the dataset under various protocols.
- **Researcher**: researchers associated with datasets
- Funder: is associated to a grant
- etc







A few key features

- Customisable templates for dataset descriptions
- Full DMP workflow support
- Support of collaboration on DMP authoring
- Authentication via a variety of providers
- Integration with several terms / vocabulary providers
- Import / export in human and machine-readable forms

- Accessible via any web browser
- Free and open source
- DMP versioning
- DMP and Dataset description cloning
- DMP publication mechanism (with PID issuing)
- Search features







More new features

- Support -> <u>argos@openaire.eu</u>
- JSON exports
- 10 more funders templates

```
A Data Management Plan consists of a set of questions that you should answer with a level of detail appropriate to the grant or with
  Academy Of Finland
  Dataset description template for Academy Of Finland projects
  Netherlands Organisation For Scientific Research (NWO)
  Dataset description template for Netherlands Organisation For Scientific Research (NWO)
  The Research Council Of Norway
  Dataset description template for The Research Council Of Norway
  National Science Center Poland
  Dataset description template for National Science Center Poland
  Swedish Research Council
  Dataset description template for Swedish Research Council
  UK Research Innovation
Dataset Description Template *
Select a template to describe your Dataset Description(s), If you want to select multiple Dataset Description Templates use "New DMP (Expert)". You can also add new Dataset
Descriptions at any time after the DMP creation.
If you can't find a template or if you want to create a personalized template for your institution, research community or training needs, please contact us.
```





```
"dmp": {
    "contact": {
        "mbox": "elli.libd@gmail.com",
        "name": "Elli Papadopoulou",
        "contact id": {
            "identifier": "a7c440e8-4e29-4d84-812a-1b2d5cb605f5",
            "type": "other"
    "created": "18-09-2019",
    "dataset": [
            "data_quality_assurance": [
            "dataset id": {
                "identifier": "56ac2eba-ea09-49cf-bba3-87fe48a10827",
                "type": "other"
            "issued": "9/18/19",
            "keyword": [
            "language": "en",
            "metadata": [
            "personal_data": "unknown",
            "security and privacy": [
            "sensitive data": "unknown",
            "technical resource": [
```

ARGOS Workflow and roles

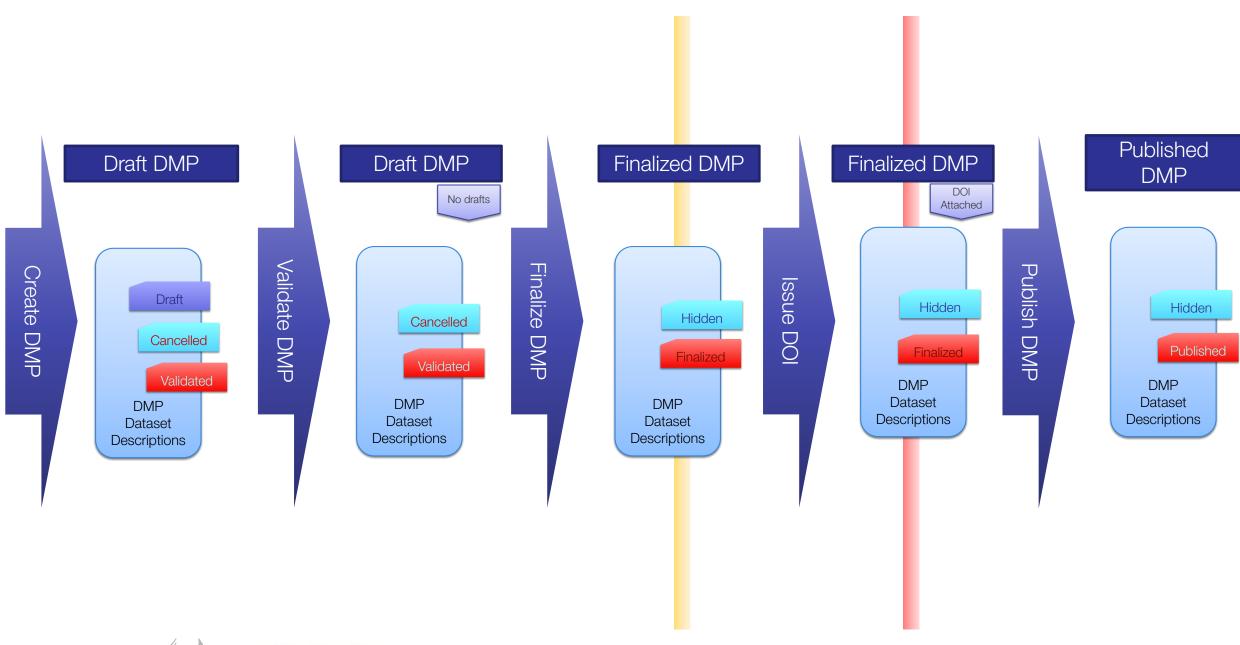
- Admins: Prepare & configure system
 - select and configure information sources
 - define dataset profiles (templates)
- DMP managers: launch DMPs
 - select appropriate Dataset profiles
 - invite researchers/dataset managers
- Researchers/Dataset managers
 - describe datasets according to fixed (systemic) and custom (profile-defined) attributes
- DMP Managers finalize DMPs
 - close the editing process
 - publish the DMPs (url, pdf, xml etc)
- System
 - Presents the DMP







DMP Lifecycle in ARGOS

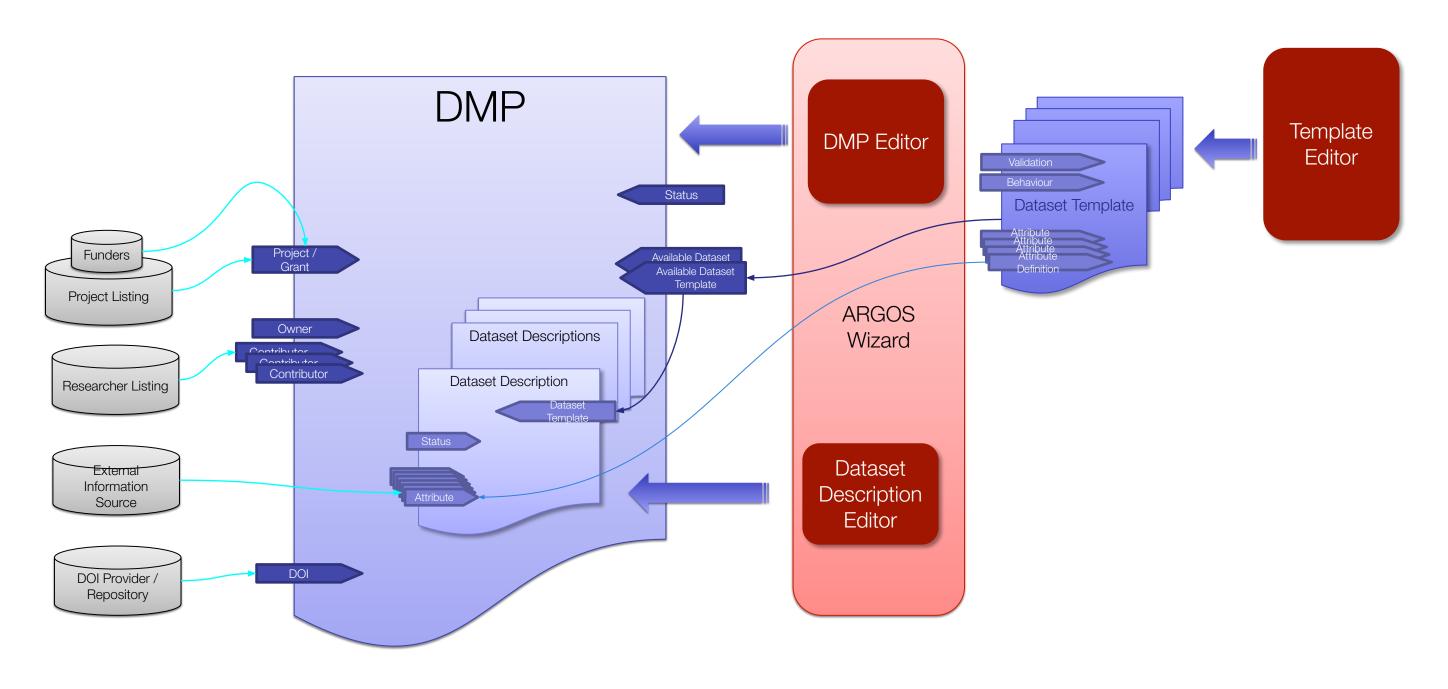








Data model overview







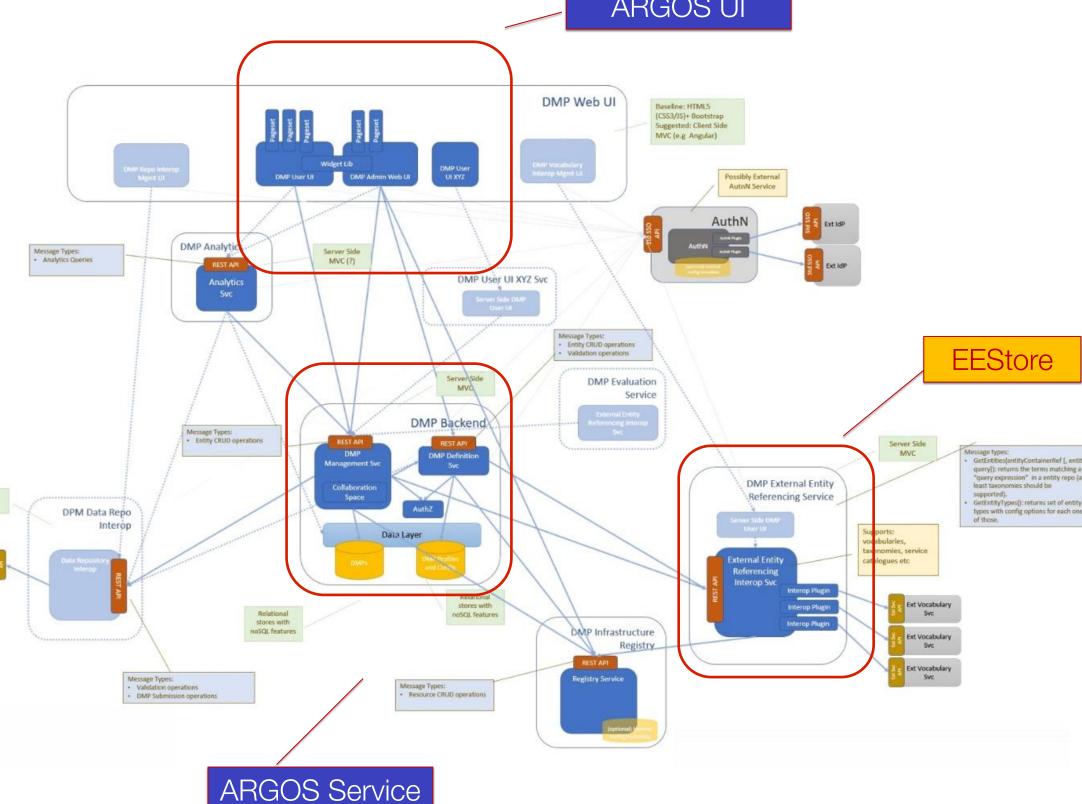


Architecture

ARGOS UI

ARGOS Core

- DMP Service: Core tools for editing, management and administration of DMPs/Dataset descriptions.
- DMP UI: End-user UI for editing **DMPs** and Dataset descriptions
- **DMP Admin UI: Administrative** UI and dataset profile definition UI
- EEStore: bridge to various providers of data for the ARGOS
- Core bridge component
- Several adapters for external providers









Technology & License

- Latest technology web application
- Frontend:
 - Single Page Application (SPA)
 - Angular / Typescript
 - Material Design / Responsive
- Backend:
 - REST backend services
 - Java/Spring Stack
 - PostgreSQL RDBMS
- License:
 - Unrestricted Open Source: Apache 2









ACCECIONE ACCIONE







Is ARGOS a questionnaire engine?

The reply is "Definitely No": The dataset profile (questionnaire) engine is a key, powerful, element of ARGOS but it is only a small part of it. Yet, even its questionnaire engine, offers essential facilities for DMP handling.

- ARGOS streamlines the handling of a DMP, not generic questionnaires:
 From the initiation to publication time, the tools and the underlying data model facilitate navigation, discovery and reuse of dataset descriptive information.
- ARGOS service is designed to handle workflows on actionable DMPs, where the questionnaire is a small step of the process. Metadata and tools that allow validation of information provided and interacting with external Data-oriented services are fundamental objectives behind the design of the system.
- ARGOS enables collaborative DMP composition that will surpass the capabilities of any generic tooling.







Why ARGOS?

- Is backed by strong collaboration that activates on several areas of EU Open Science paradigm
 - OpenAIRE and EUDAT
- · Can be exploited under different models of service provisioning
 - Software as a service
 - Installation software
- Closely follows standards and well-known specifications
 - RDA and several common IT standards
 - Common templates are provided
- Has been implemented from scratch taking into account the current trends on Actionable DMPs, Open Science and experiences from existing tools
- ... and for several technical and functional reasons!







Interoperability in ARGOS



REST API for utilization of core components by external services



Custom REST protocol for external source integration



Machine readable DMP export

- proprietary format
- RDA JSON experimental support

Open architecture

- for integrating external information providers to the DMP entities
- EEstore abstracts over the behavior and performance of external providers
- M2M communication data repositories, registries and other dataset handling services (work in progress)
 - E.g. depositing and verification of dataset(s) and DMP(s)

Flexible dataset and DMP model that can power different interoperability cases







Machine actionability



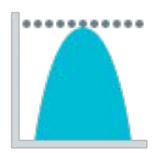


- ✓ Integrate DMPs with the workflows of all stakeholders in the research data ecosystem
- Use PIDs and controlled vocabularies
- Follow a common data model for maDMPs (Machine Actionable Data Management Plans)
- Make DMPs available for human and machine consumption
- Make DMPs updatable, living, versioned documents
- Make DMPs publicly available



Validation

- ✓ Allow automated systems to act on behalf of stakeholders
- ✓ Support data management evaluation and monitoring



Value

- Make policies (also) for machines, not just for people
- ✓ Describe for both machines and humans — the components of the data management ecosystem







Platform Walkthrough

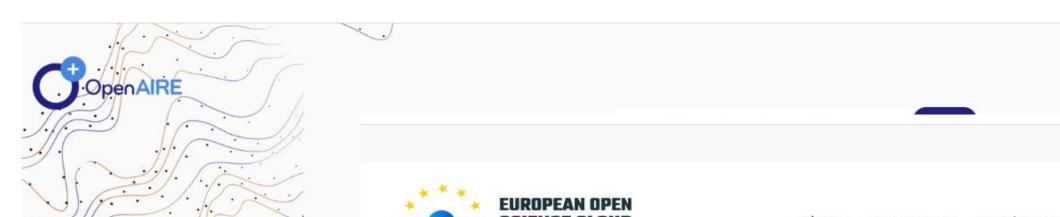
https://argos.openaire.eu/







Where to find ARGOS

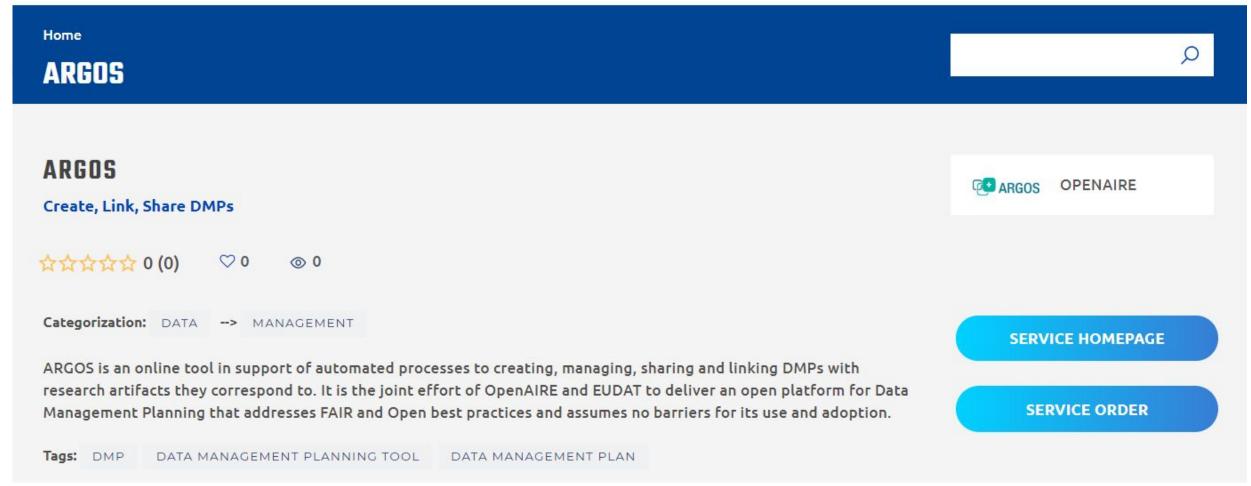


My Services Profile Contact Us Log In



About Governance Services & Resources Policy EOSC in practice Media For Providers

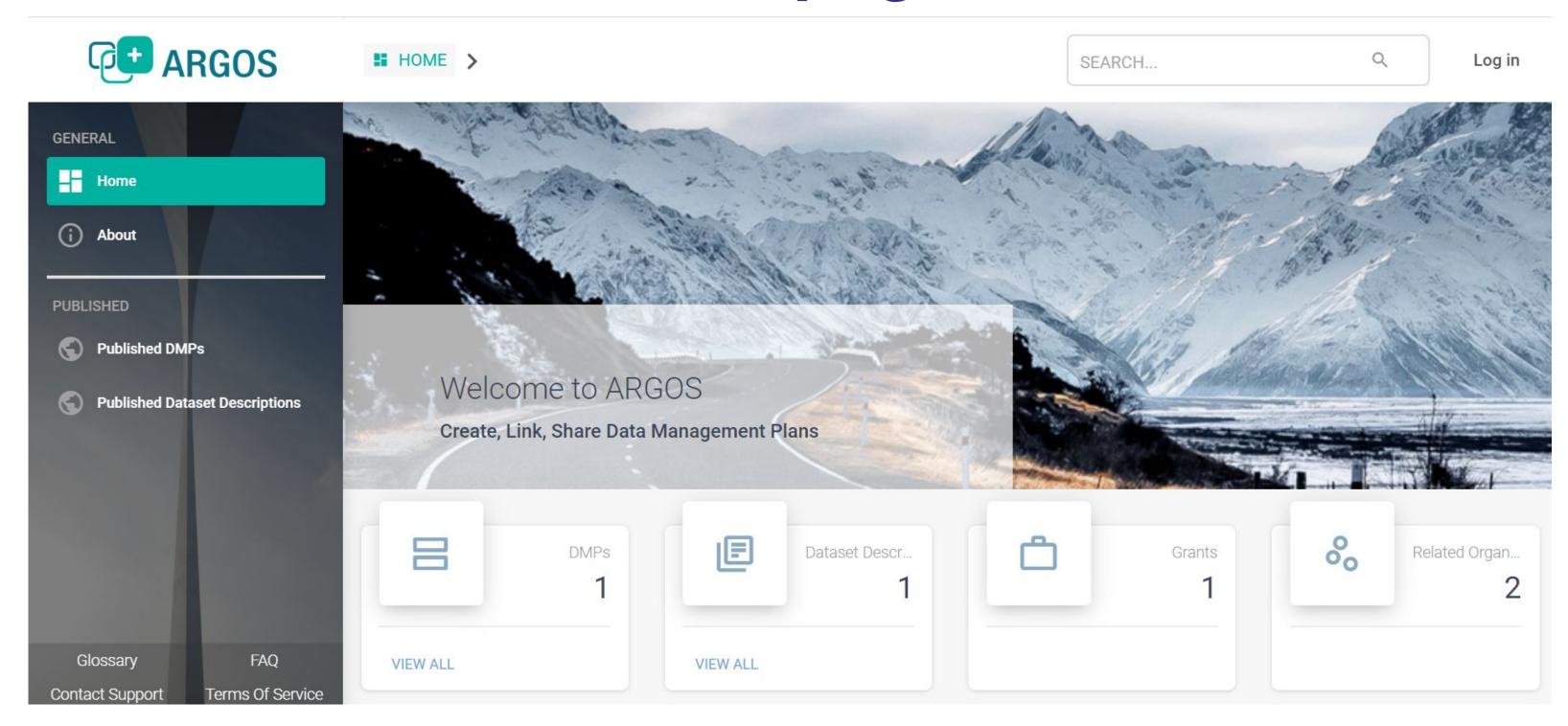
SERVICES





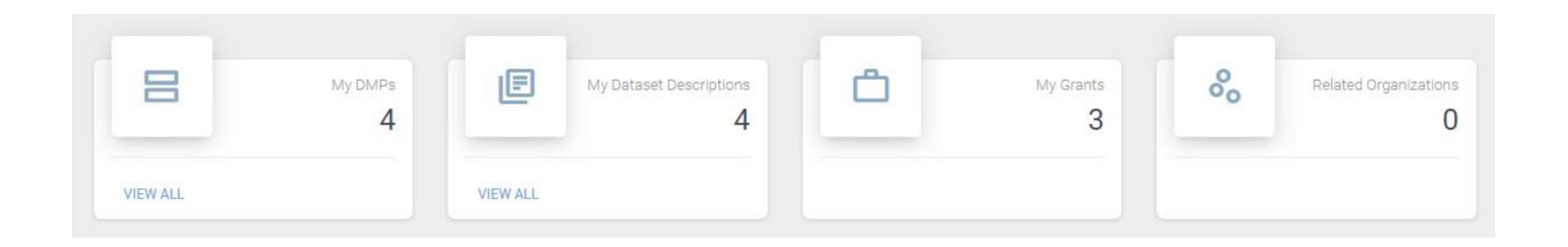


Homepage



Features (without log in) - Stats

 # of DMPs, Dataset Descriptions, Grants and Organizations included in ARGOS.

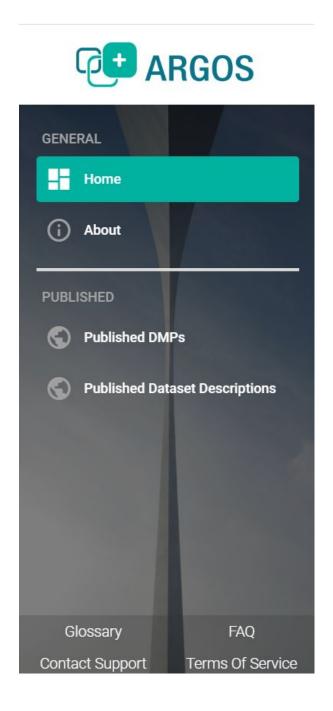


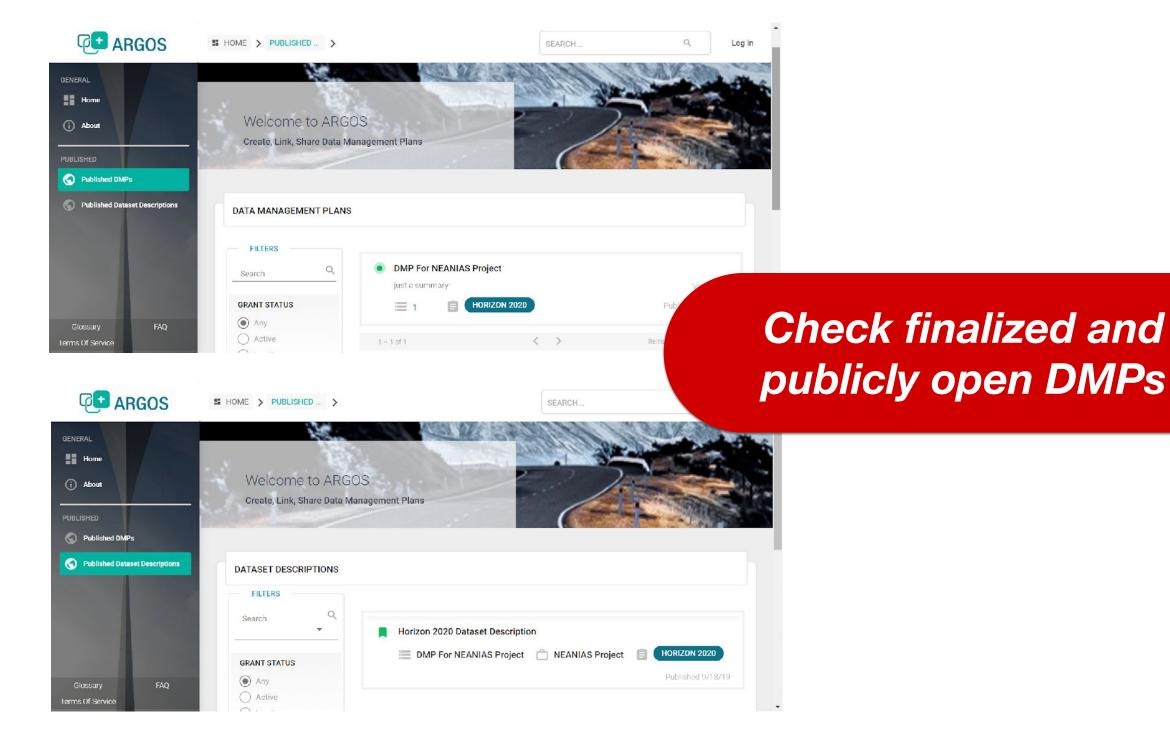
Note! When logged in to ARGOS the numbers change to correspond to information provided by each user's activity.





Features (without log in) – openly available outputs

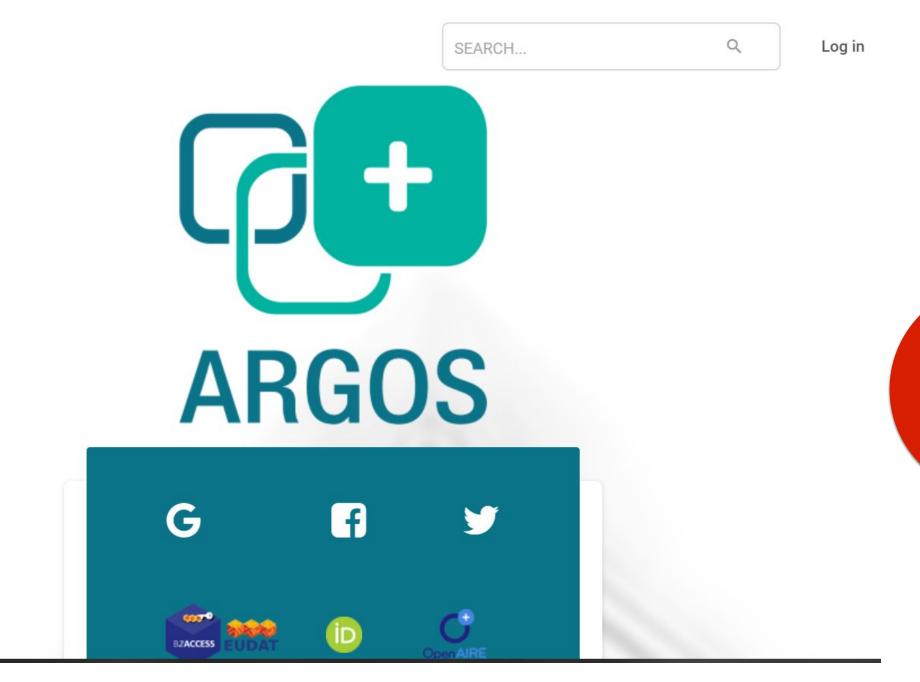








Login

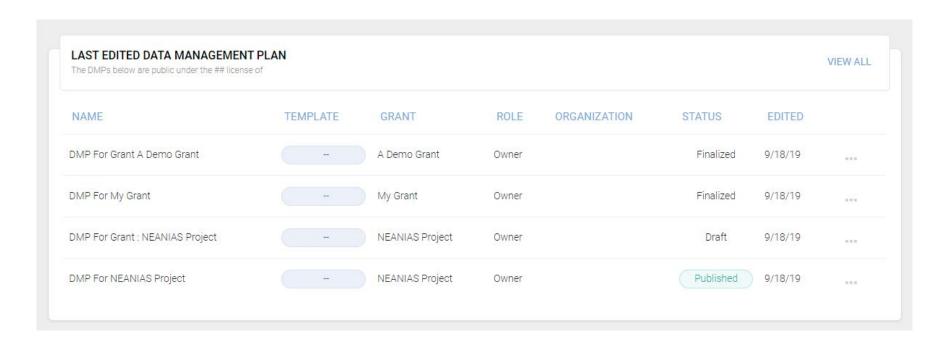


No user-account is required





Your dashboard - Recent Activity



Name - the name of the DMP
Template - the DMP template used
Grant - the number of the grant funded
the research under which the DMP is
conducted

Role - your role in the different DMPs that you are involved in

Organization - the organization responsible for the DMP development and delivery

Status – the status of the DMP development (Draft, Finalized, Published) **Edited** - date when the DMP was last edited





Two ways to create DMPs or Dataset Descriptions



- Provides an **easy** workflow
- Includes all essential elements
- Provides **guidance** through the required steps



- Integrates different workflow
- Includes more steps and elements
- Useful to **cross-check things** before finalizing work





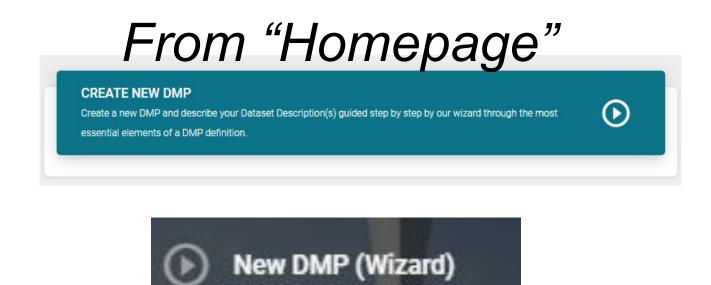
DMPs





Create DMPs - DMP Wizard

• **DMP Wizard** is an easy way to start writing a DMP. Wizard provides an interface that goes through the essential elements of a DMP and guides the process of its creation step by step.

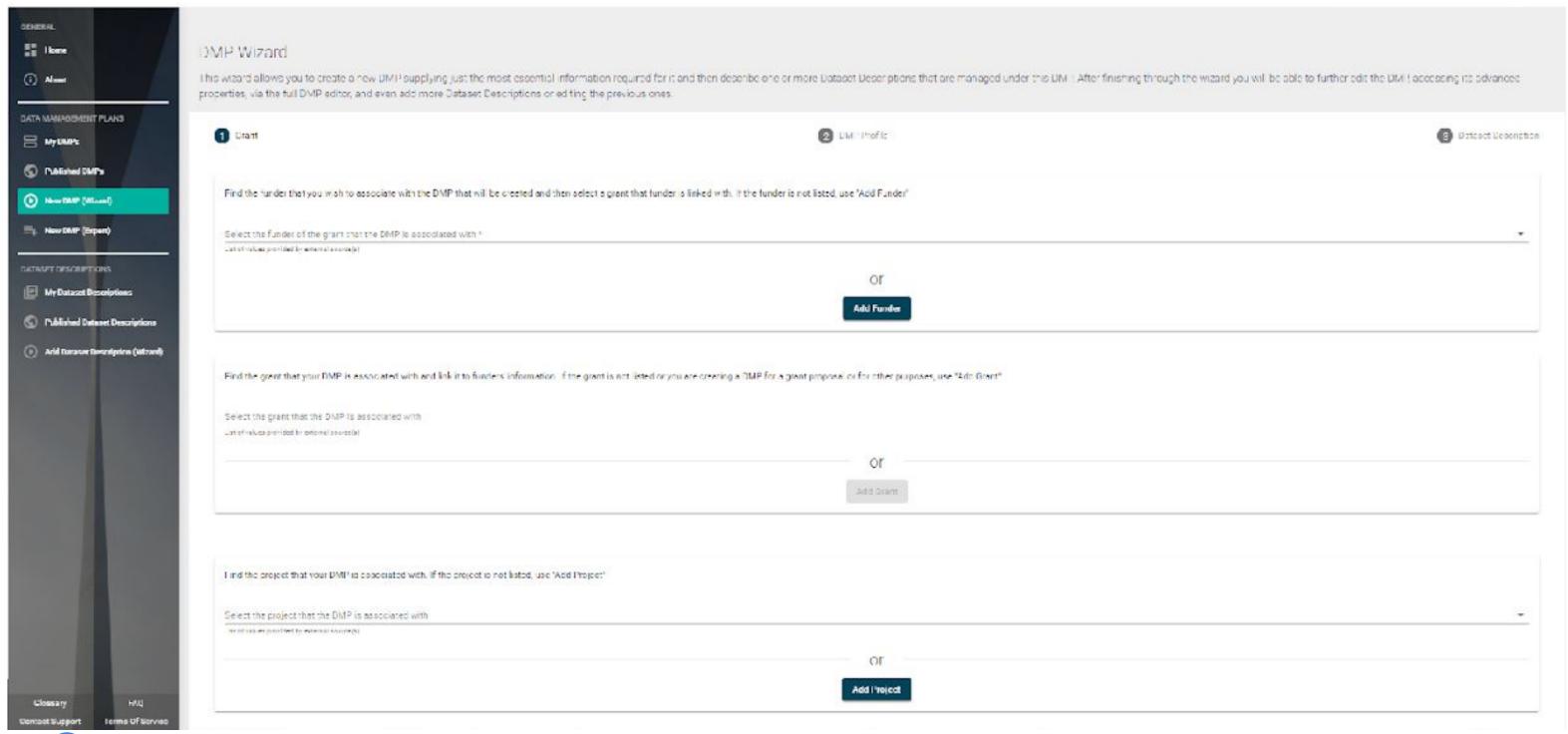








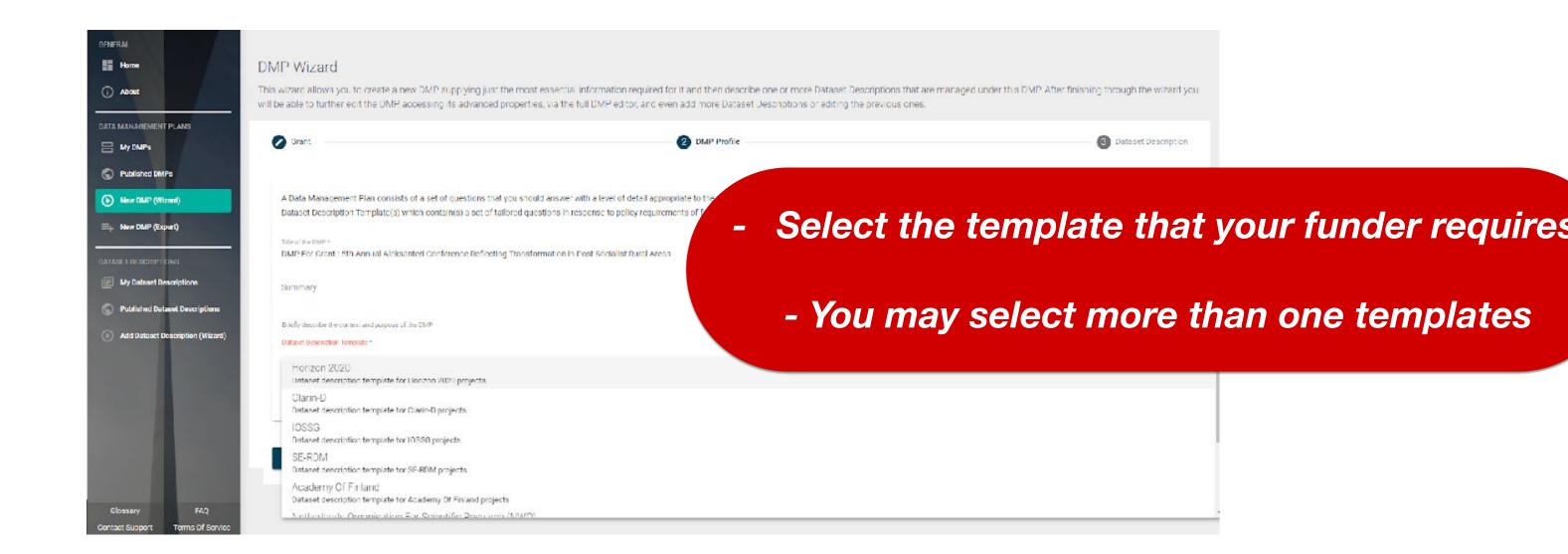
DMP Wizard – 1/3







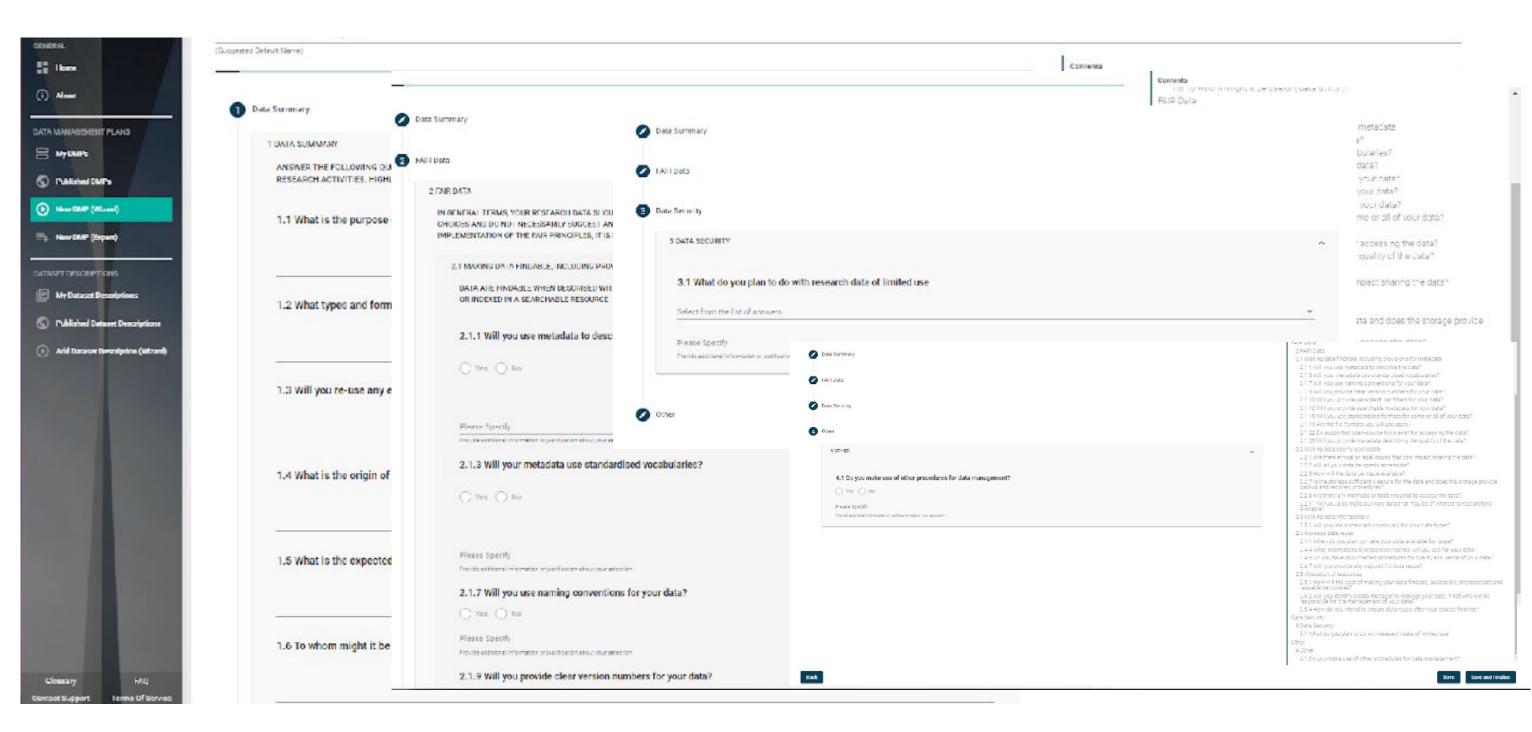
DMP Wizard – 2/3







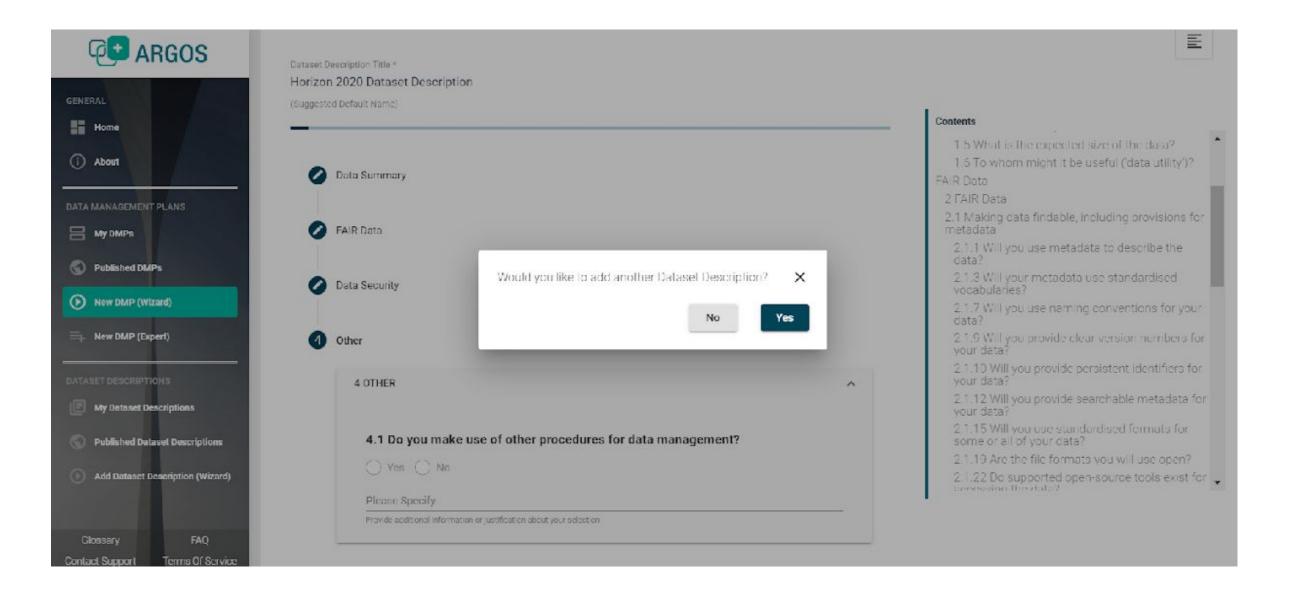
DMP Wizard – 3/3







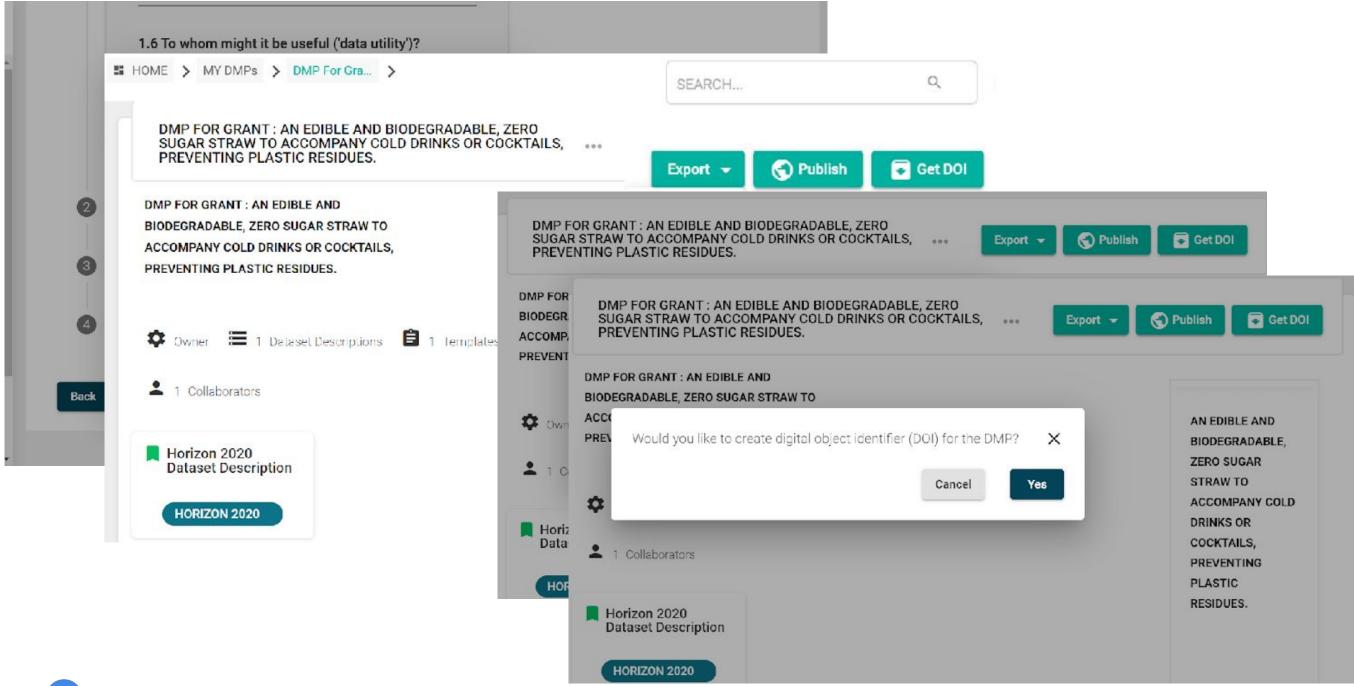
Save your DMP







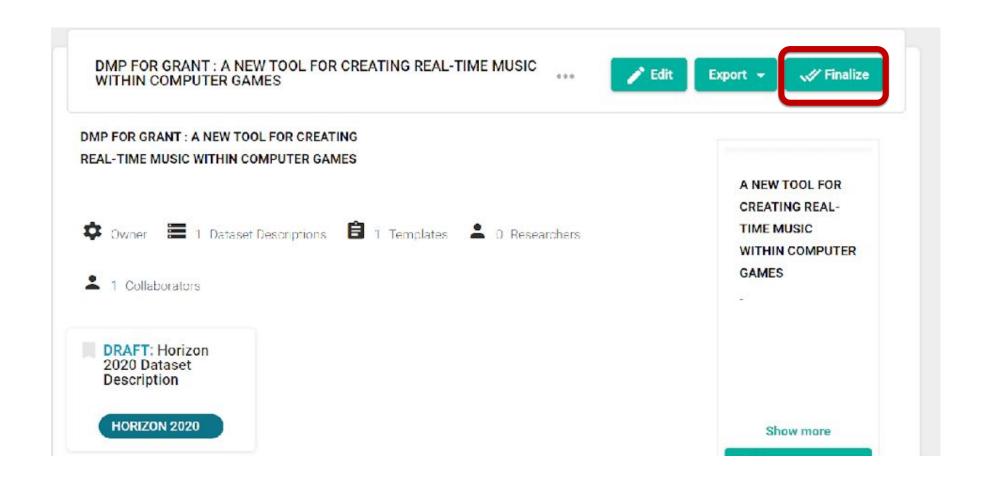
Save and finalize your DMP







Finalize saved DMPs







Create DMPs - DMP Expert

DMP Expert provides additional metadata to the DMP.

From "Homepage"



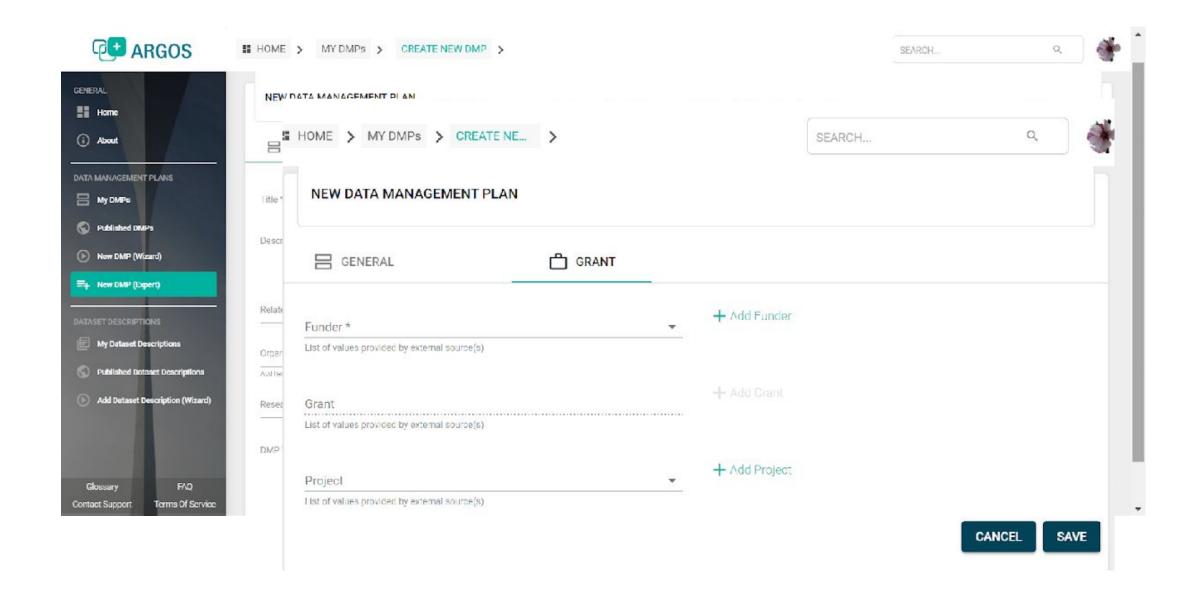
From "My DMPs"







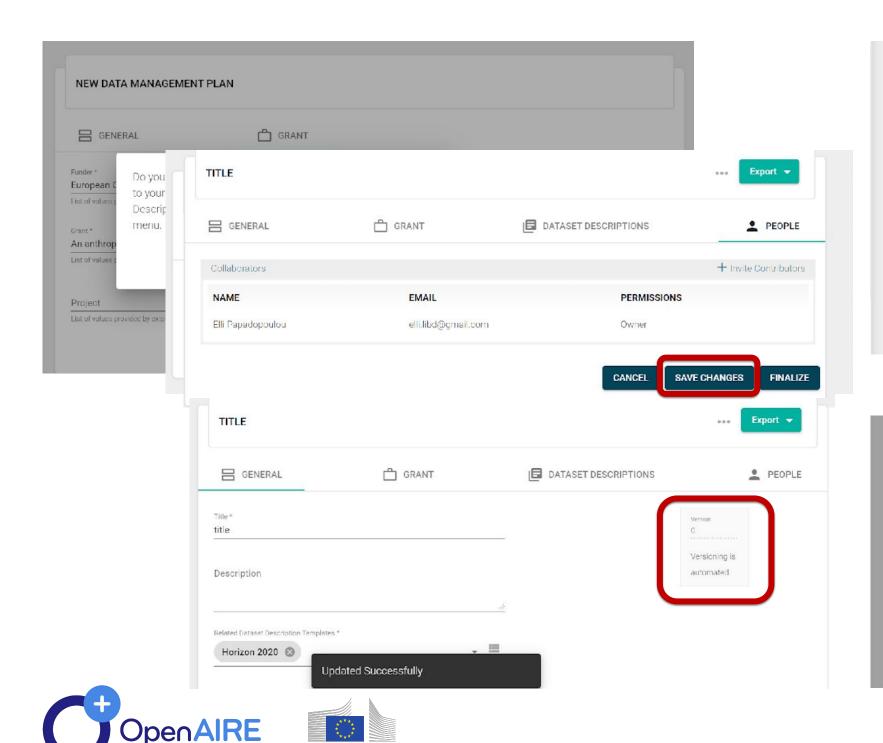
DMP Expert

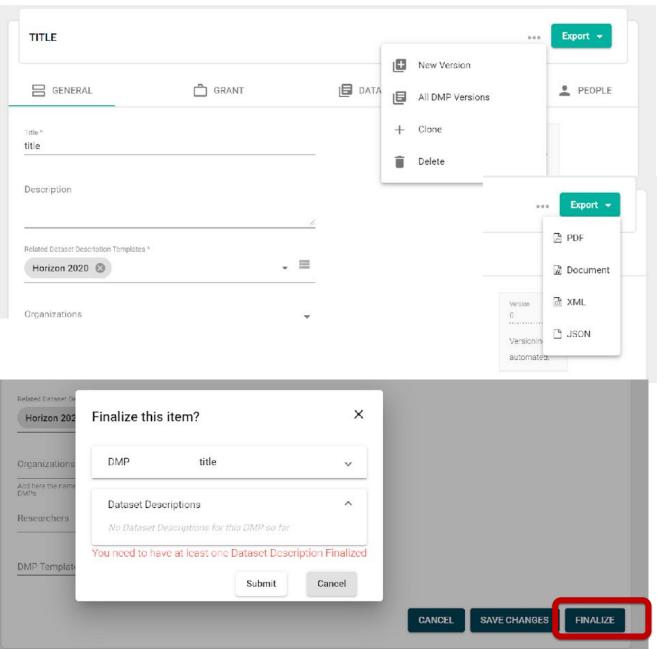






Save your DMP





Dataset Descriptions



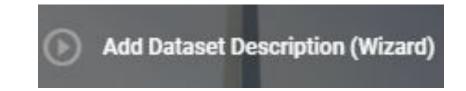


Describe your datasets - Wizard

 Dataset Description Wizard is an easy way to add new dataset descriptions to pre-existing DMPs.

From "Homepage"

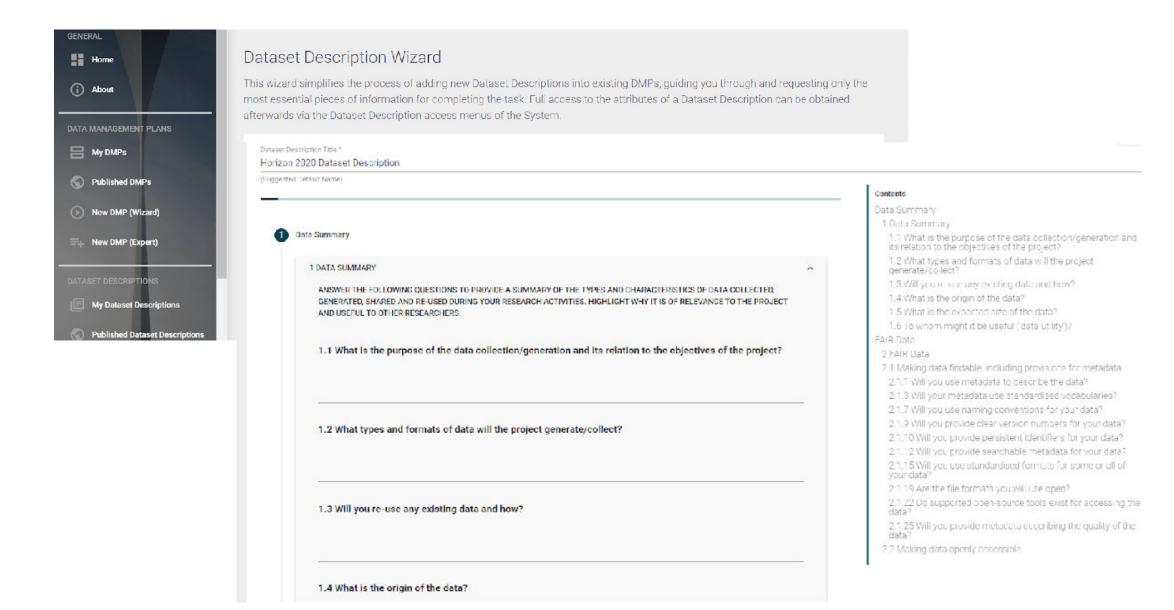








Describe NEW datasets with Wizard







Describe your datasets - Dataset Description Expert

 Dataset Description Expert provides additional metadata to the dataset description.

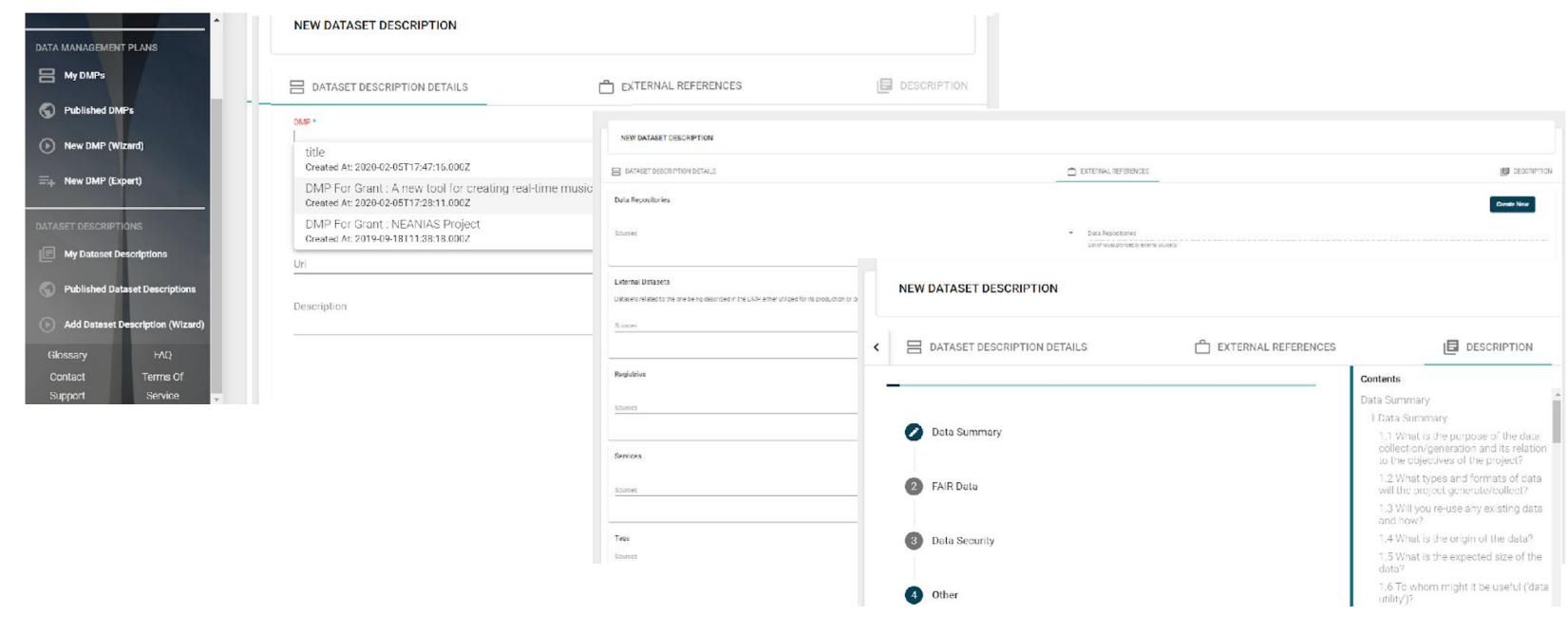
From "My Dataset Descriptions"







Describe NEW datasets with Experts







Next Steps





Next Steps

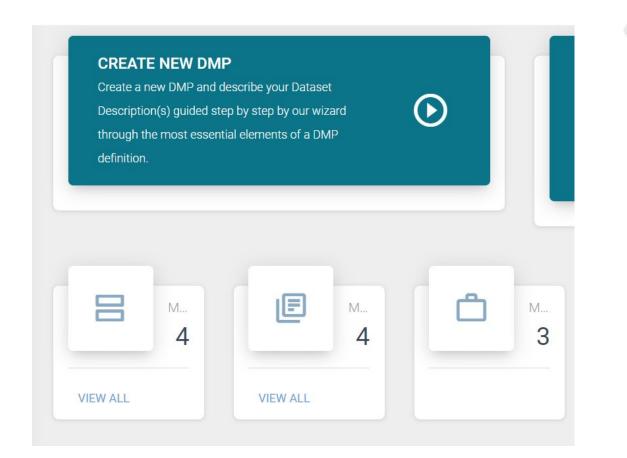
 Icons of the main page to become smaller & embedded to each other

 Integrate FundRef by Crossref



Integrate with





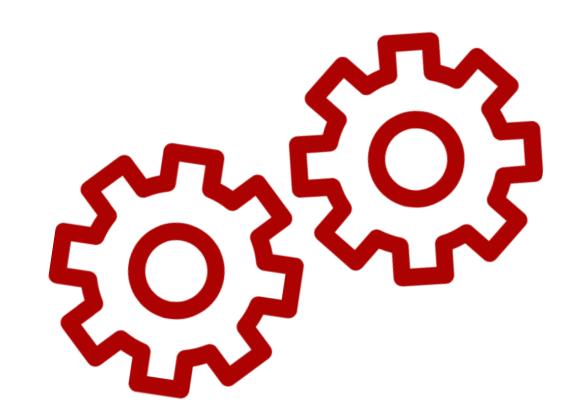


Next Steps – continuous processes

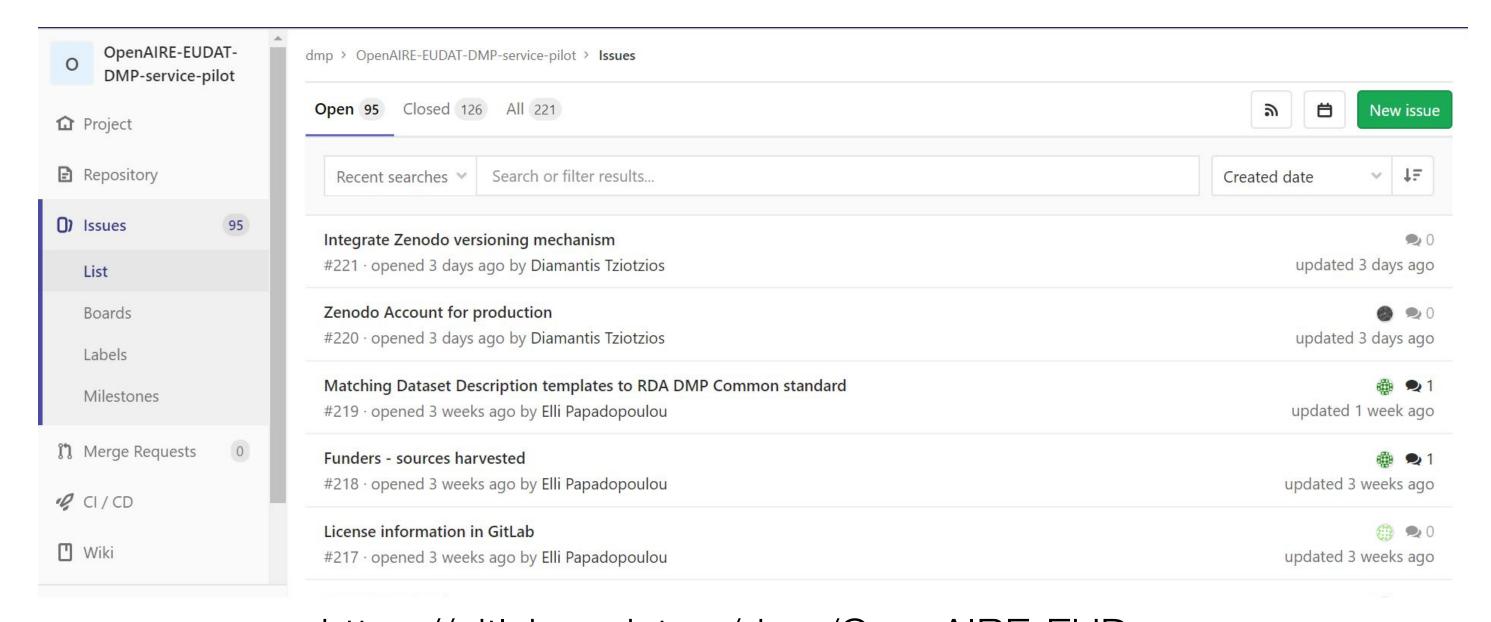
- Enrich ARGOS with more funders' templates
- Enrich resources inferred through the EEStore
- Usability tests and community feedback
- Debugging







GitLab tickets



https://gitlab.eudat.eu/dmp/OpenAIRE-EUD AT-DMP-service-pilot/issues



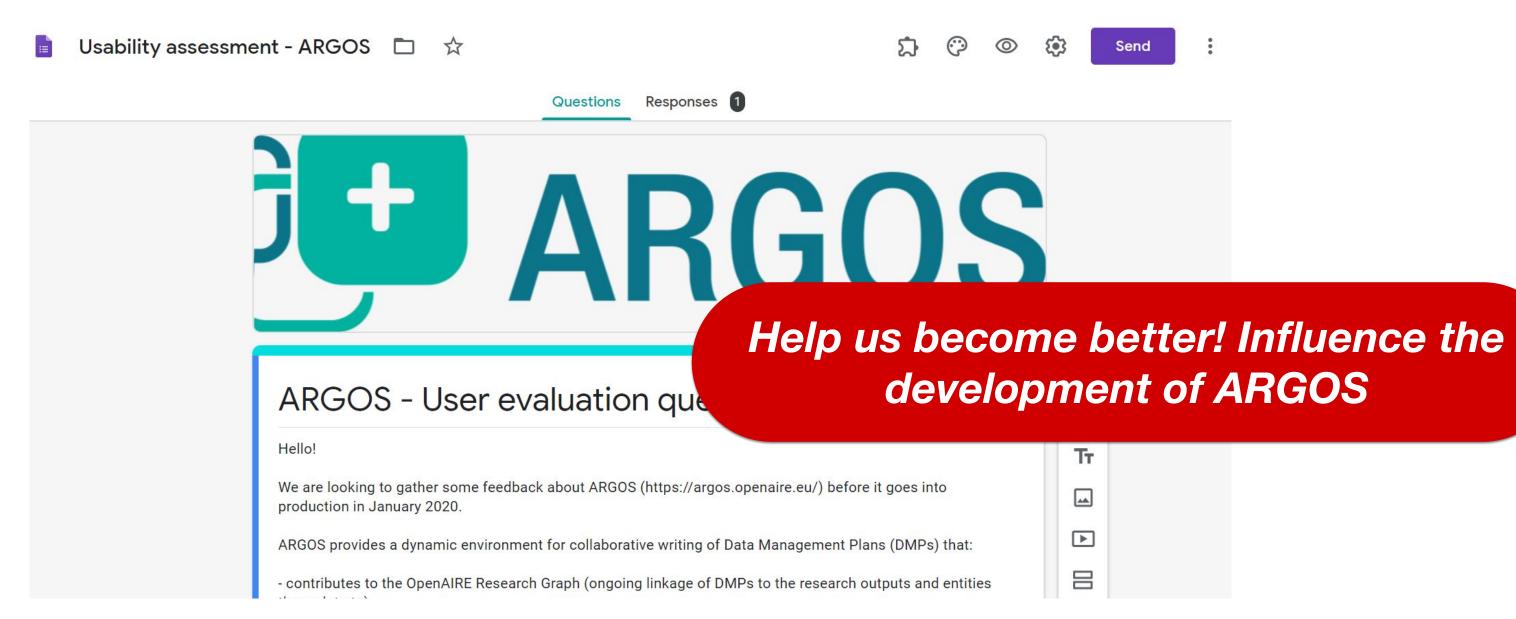


Feedback





Usability assessment – Users feedback



https://bit.ly/2tendao





Useful Resources

Blogpost:

https://www.openaire.eu/blogs/argos-the-newest-openaireservice-to-create-link-and-share-dmps-oaweek

Service Guide: https://www.openaire.eu/argos-guide

Trello Board: https://trello.com/b/x49lylnK/argos

Thank you!

Name

email

