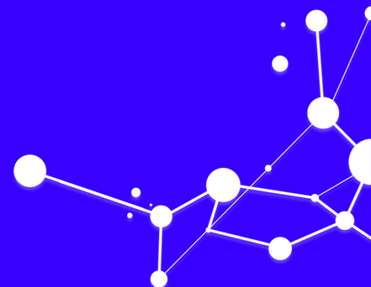
Elli Papadopoulou
ATHENA Research Center

Plan and follow your data

DMPs



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

Excluded

- preliminary analysis, drafts of scientific papers, plans for future research
- peer reviews or communications with colleagues
- physical objects (e.g., laboratory samples)
- trade secrets, commercial and confidential information
- personnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.

United States Office of Management and Budget, Circular 110

DMP Templates

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 in Annex I. You should refer to the annex and the main text of the guidelines for further guidance.

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 2.1. Making data findable, including provisions for metadata	• Outline the discoverability of data (metadata provision) • Outline the identifiability of data and refer to standard identification m you make use of persistent and unique identifiers such as Digital Object • 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 star discipline describe what type of metadata will be created and how

2. Data and Metadata Standards

The project will leverage existing metadata standards currently stored in Ecological Metadata Language (EML) format for the NutNet project. We will add additional metadata entries for the arthropod community composition and arthropod stoichiometry; field notes taken during the time of collection will be recorded. Morpho software will be used to generate the metadata file in EML. We chose EML format for our metadata since it allows integration with existing NutNet data housed in the Knowledge Network for Biocomplexity (KNB) data repository.

2 Example DMP – NutNet.
© DataONE 2011

3. Policies for Access and Sharing

After publication of manuscripts based on the data we collect, we will share our data and metadata with the NutNet community via data updates sent annually as .csv files from the existing central relational database. Other NutNet users will need to contact Lind for access to the data.

We will also submit both of our datasets (abundance and stoichiometry) to the U of M Digital Conservancy, an archive for digital preservation. Borer has access to this resource as a faculty member. This will occur within a year of publication. The data will be publicly available via the Digital Conservancy, which provides a permanent URL for digital documents.

DMPs content

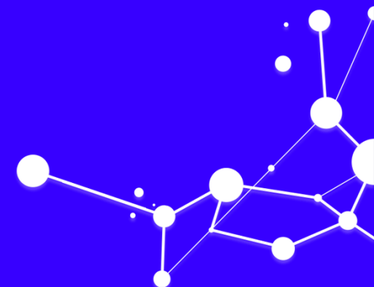
Major challenges (qualitative interviews only)

The following major challenges were raised by the interviewees in the qualitative interviews:

- reading and analyzing partner input and turning it into one understandable document, in particular at the beginning of the project, when there was little experience
- where to put the focus and how much details to give – internal procedures or output; also whether to tackle any data or data underlying publications (the latter strongly preferred)
- understanding the technicalities
- how to create the DMP from scratch with zero experience
- Understanding the requirements and convincing partners to submit thorough information (done through peer pressure). This is easier in newer projects since DMPs are more accepted
- Covering all partners, some of them in non-EU countries where different national policies apply (e.g. on protecting vulnerable groups)

https://phaidra.univie.ac.at/detail/o:1165751

Argos



Argos

EUROPEAN OPEN SCIENCE CLOUD

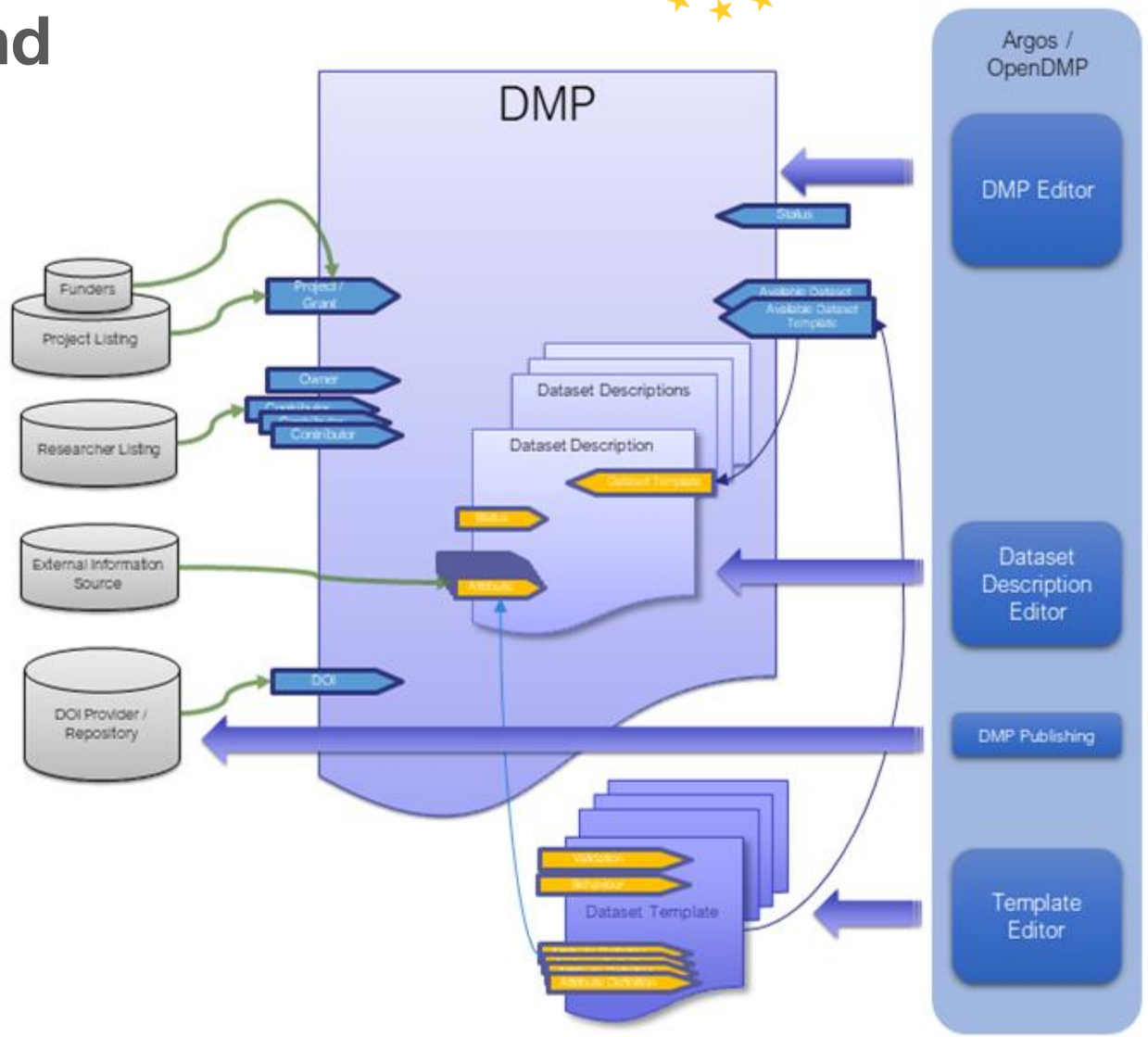
an **open source**, **configurable** and **extensible** tool for **planning Research Data Management** (RDM) activities according to Open Access & FAIR data policies.

- **Templating system**

Dynamic vs static parts

- Access points (APIs);

Import / Export; RDA standard



https://gitlab.eudat.eu/dmp/OpenAIRE-EUDAT-DMP-service-pilot

OpenAIRE NEXUS

CITE

ATHENA Research & Innovation Information Technologies

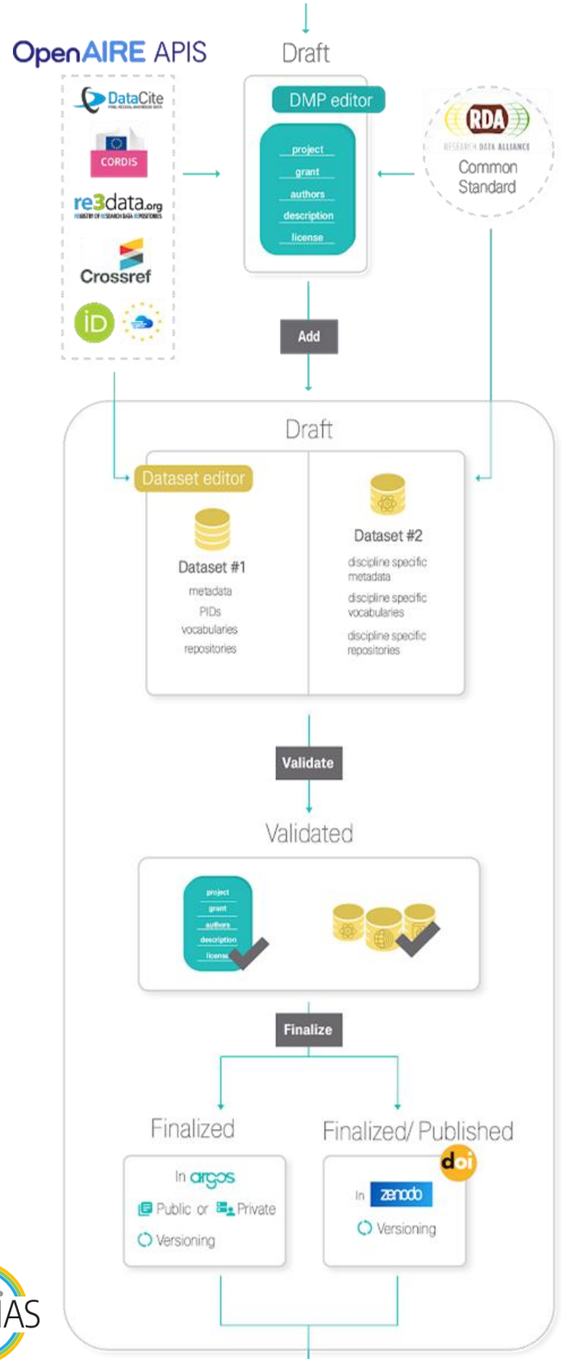
NEANIAS

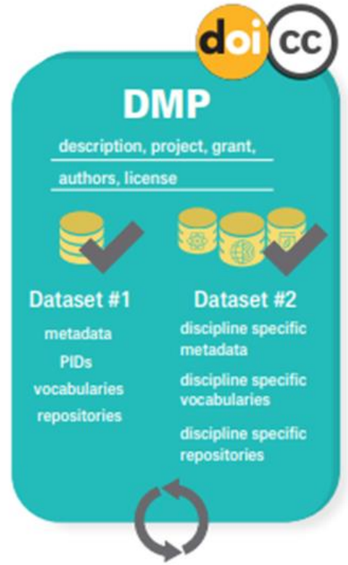
– not just a tool!

- **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. RDA DMP Common Standard; DMPs exposed in repositories with appropriate resource_type

Write & publish

- **Full DMP publication**
- **Two editors to create DMPs**
  - Full DMP editor
  - Quick wizard to add new datasets to DMPs
- **First validation**
- **Finalization**
- **ma-DMP outputs**

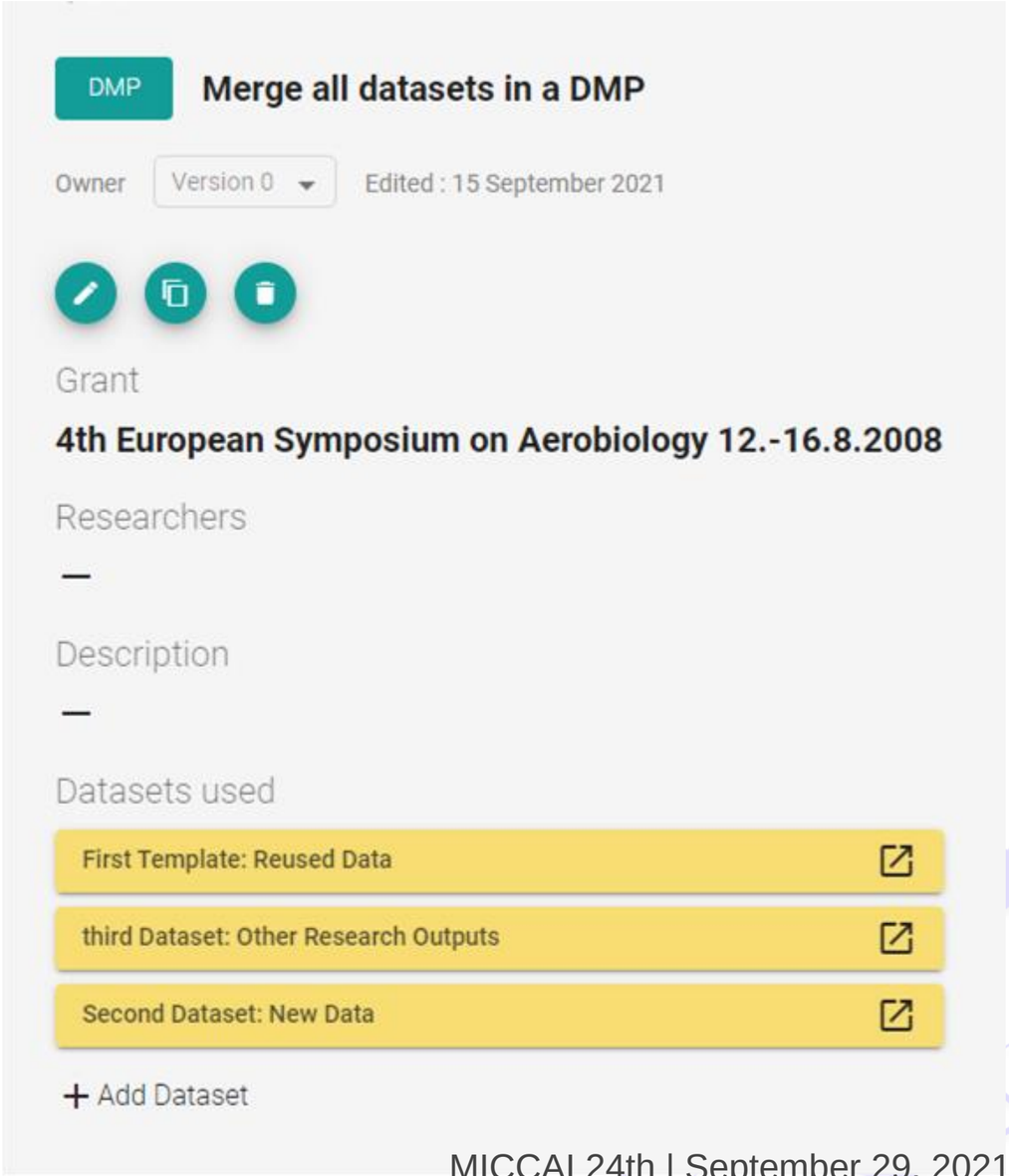




- Rich documents
- Discoverable through OpenAIRE
- Versioned (provenance)
- Accessible: PIDs (ORCIDs & DOIs)
- Reusable: Licenses
- Preserved in Zenodo

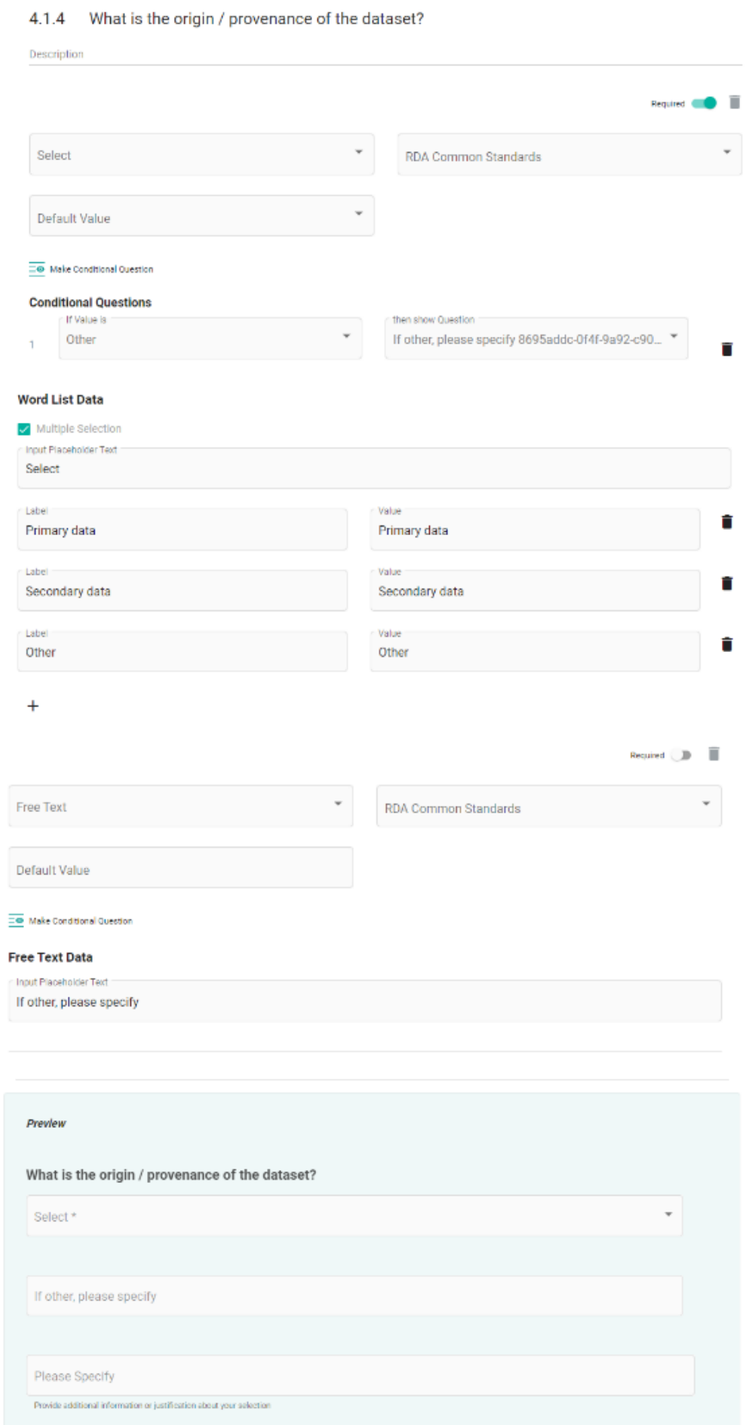
Create dataset profiles

- Different data questions
- Configurable APIs
- Tailored instructions



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
  - x the question can be answered with different input
- RDA compatibility



Connect workflows & benefit from OpenAIRE pool of data

- Create subsets of Argos in OpenAIRE that meet your needs.
- Publish in Zenodo.
  - Or integrate with own repositories
- Notify repository managers for new datasets.
  - Integrations to enable pre-filling of DMPs for re-used data
- Exploit DMP entities in the Research Graph.
  - Create links between outputs and entities.
- Combine with validated OpenAIRE data and provide statistics.
  - Define indicators; Add to dashboards.
- Add DMPs under the project's page.

Platform Walkthrough

https://argos.openaire.eu/

@openaire_eu

#OpenAIRE-Nexus

THANK YOU

Elli Papadopoulou | Georgios Kakaletris

elli.p@athenarc.gr | gkakas@athenarc.gr

OpenAIRE