

Machine-actionable DMPs - national & international alignments

Tomasz Miksa

TU Wien, OSTrails Technical Lead

The OSTrails logo is positioned on the right side of the slide, set against a dark purple background. It features the word "OSTrails" in a white, sans-serif font. The "O" is stylized with a white circle and a small orange dot at the top. The background of the slide is decorated with horizontal white and orange bars of varying lengths and colors, creating a modern, abstract design.

Data Management Plans (DMPs)

	Data Officer	<i>Who is responsible for the data management and the DMP of the project (name/email address)?</i>
I	Data Characteristics	
I.1	Description of the data	<i>What kinds of data/source code will be generated or reused (type, format, volume)? How will the research data be generated and which methods will be used? How will you structure the data and handle versioning? Who is the target audience?</i>
II	Documentation and Metadata	
II.1	Metadata standards	<i>What metadata standards (if any) will be in use and why? (see Digital Curation Centre)</i>
II.2	Documentation of data	<i>What information is needed for the data to be findable, accessible, interoperable and re-usable (FAIR) in the future? Is the data machine-readable? How are you planning to document this information?</i>
II.3	Data quality control	<i>What quality assurance processes will you adopt? How will the consistency and quality of data collection be controlled and documented? (This may include processes such as repeat samples or measurements, standardised data capture, peer review of data or representation with controlled vocabularies.)</i>
III	Data Availability and Storage	
III.1	Data sharing strategy	<i>How and when will the data be shared and made accessible? What repository will you be using? What persistent identifier will be used?</i>
III.2	Data storage strategy	<i>What data are to be preserved for the long-term, and what data will not be stored? How and where will the data be stored and backed up during the research? How and where will the data be stored after the project ends? For how long will the data be stored? Are there any costs that need to be covered for storage? At what point during or after the project will the data be stored? Are there any technical barriers to making the research data fully or partially accessible?</i>

DCC Checklist		DCC Guidance and questions to consider	
Administrative Data			
ID	A pertinent ID as determined by the funder and/or institution.		
Funder	State research funder if relevant		
Grant Reference Number	Enter grant reference number if applicable [POST-AWARD DMPs ONLY]		
Project Name	If applying for funding, state the name exactly as in the grant proposal.		
Project Description	Questions to consider: <ul style="list-style-type: none"> - What is the nature of your research project? - What research questions are you addressing? - For what purpose are the data being collected or created? Guidance: Briefly summarise the type of study (or studies) to help others understand the purposes for which the data are being collected or created.		
PI / Researcher	Name of Principal Investigator(s) or main researcher(s) on the project.		
PI / Researcher ID	E.g ORCID http://orcid.org/		
Project Data Contact	Name (if different to above), telephone and email contact details		
Date of First Version	Date the first version of the DMP was completed		
Date of Last Update	Date the DMP was last changed		
Related Policies	Questions to consider: <ul style="list-style-type: none"> - Are there any existing procedures that you will base your approach on? - Does your department/group have data management guidelines? - Does your institution have a data protection or security policy that you will follow? - Does your institution have a Research Data Management (RDM) policy? - Does your funder have a Research Data Management policy? - Are there any formal standards that you will adopt? Guidance: List any other relevant funder, institutional, departmental or group policies on data management, data sharing and data security. Some of the information you give in the remainder of the DMP will be determined by the content of other policies. If so, point/link to them here.		
Data Collection			
What data will you collect or create?	Questions to consider: <ul style="list-style-type: none"> - What type, format and volume of data? - Do your chosen formats and software enable sharing and long-term access to the data? - Are there any existing data that you can reuse? Guidance: Give a brief description of the data, including any existing data or third-party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.		
How will the data be collected or created?	Questions to Consider: <ul style="list-style-type: none"> - What standards or methodologies will you use? - How will you structure and name your folders and files? - How will you handle versioning? - What quality assurance processes will you adopt? Guidance: Outline how the data will be collected/created and which community data standards (if any) will be used. Consider how the data will be organised during the project, mentioning		

DMPs vs machine-actionable DMPs

Traditional DMP

```
<administrative_data>
```

```
  <question>Who is responsible for the DMP?</question>
```

```
  <answer>Moritz from our university.</answer>
```

```
</administrative_data>
```

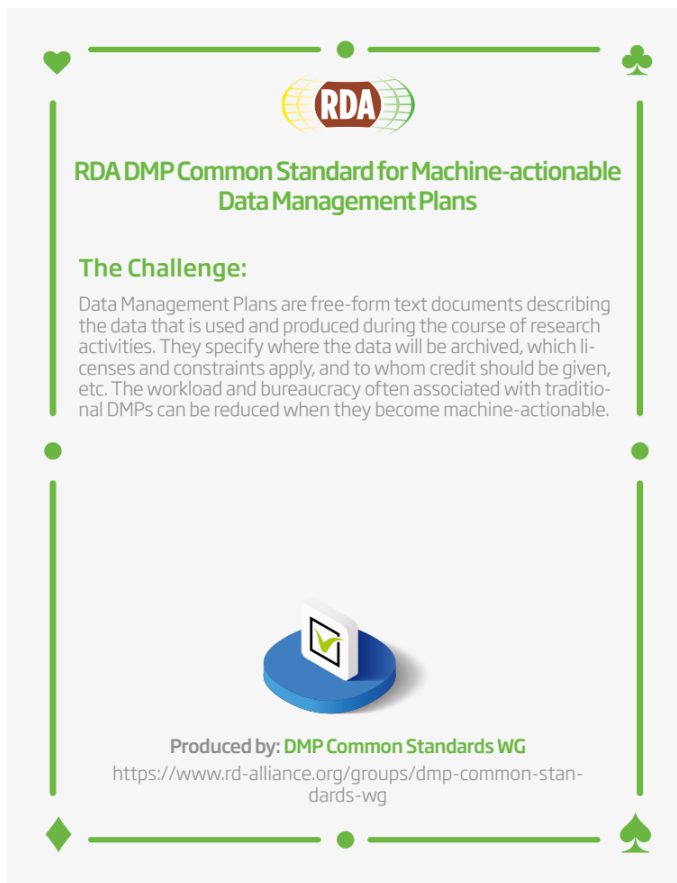
Machine-actionable DMP

```
"contributor" : [ {  
  "contributor_id" : {  
    "identifier" : "0000-0002-5164-2690",  
    "type" : "orcid"  
  },  
  "mbox" : "moritz.staudinger@tuwien.ac.at",  
  "name" : "Moritz Staudinger",  
  "role" : [ "Data Manager" ]  
}
```

maDMPs use PIDs and controlled vocabularies.

Example shows that Moritz is the one responsible for data management.

Official RDA Recommendation on maDMPs



RDA

RDA DMP Common Standard for Machine-actionable Data Management Plans

The Challenge:

Data Management Plans are free-form text documents describing the data that is used and produced during the course of research activities. They specify where the data will be archived, which licenses and constraints apply, and to whom credit should be given, etc. The workload and bureaucracy often associated with traditional DMPs can be reduced when they become machine-actionable.

Produced by: **DMP Common Standards WG**
<https://www.rd-alliance.org/groups/dmp-common-standards-wg>

RDA DMP Common Standard for Machine-actionable Data Management Plans

Recommendations of the RDA DMP Common Standards WG
Tamasz Miksa, Paul Walk, Peter Neish

Purpose

This application profile is meant for exchange of machine-actionable DMPs between systems. It is independent of any internal data organisation used by these systems. The application profile does not prescribe how information must be presented to the end user and does not enforce any specific logic on how this information must be collected or used. The application profile is an information carrier and the full machine-actionability can only be achieved when systems using the application profile implement appropriate logic.

This application profile is intended to cover a wide range of use cases and does not set any business (e.g. funder specific) requirements. It represents information over the whole DMP lifecycle, that is, it can express planned actions, as well as actions already performed.

The application profile is NOT intended to be a prescriptive template or a questionnaire, but to provide a re-usable way of representing machine-actionable information on themes covered by DMPs.

Overview

Figure 1 presents concepts used within the application profile. Each concept is further broken down into specific fields (not depicted). The full application profile specification can be found [online](#). Below we outline main concepts used within the application profile that are depicted in Figure 1.

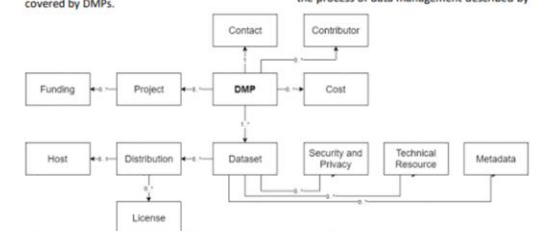
DMP - Provides high level information about the DMP, e.g. its title, modification date, etc. It is the root of this application profile.

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 or EC funded projects.

Contact - Specifies the party which can provide information on the DMP.

Contributor - For listing all parties involved in the process of data management described by



```
graph TD
    Funding --> Project
    Project --> DMP
    DMP --> Cost
    DMP --> Contact
    DMP --> Contributor
    Host --> Distribution
    Distribution --> Dataset
    Dataset --> License
    Dataset --> SecurityAndPrivacy[Security and Privacy]
    Dataset --> TechnicalResource[Technical Resource]
    Dataset --> Metadata
```

Figure 1: Overview of the application profile for the machine-actionable DMPs.

1

Machine-actionable DMPs in Austria



○ DAMAP

- Tool for DMPs and maDMPs
- RDA recommendation as a basis
- Implemented within *FAIR Data Austria*

○ DAMAP early adopters

- TU Wien, TU Graz

○ DAMAP is being rolled out within *Shared RDM* at other universities



www.damap.org



A...kademie der bildenden Künste Wien



DAMAP – Examples of automation

Martin Gruber ¹

Persons who are not members of TU Graz must be personally informed by you about the collection of their data in this tool. [297-7076](#)

Ilire Hasani-Mavriqi

ilire.hasani-mavriqi@tugraz.at

Laura Thaci

lthaci@tugraz.at

Search for person...

David Eck

Search service field **University**

David Eckhard

Show manual contributor input

Choose project from the project database Input project manually

Please select the project you want to create a DMP for:

Search project...

global

IEA IETS Task XV - Industrial excess heat recovery - Subtask 4

Dec 31, 2023 - Dec 30, 2025

Operation of the bio-geophysical variables systematic monitoring of the

Aug 11, 2023 - Aug 10, 2025

Im Zuge des angedachten Projektvorhabens zwischen der Firma Global-

May 31, 2023 - Dec 30, 2023

IMSIL-GlobalTCAD

May 31, 2023 - May 30, 2025

Repositories

Repositories are used for sustainable data management beyond the lifetime of the project and to make data discoverable and reusable to others. If available, use a domain specific repository for your data. The search filter below may help you to find one. Otherwise, you can always use TU Wien Research Data, TU Wiens institutional data repository.

Recommended Find repository

TU Wien Research Data

TU Wien Research Data is an institutional repository of TU Wien to enable storing, sharing and publishing of digital objects, in particular research data. It facilitates the funders' requirements for open access to research data and the FAIR principles by making research output findable, accessible, interoperable, and reusable. A DOI is assigned to each dataset published in TU Wien Research Data. This service is developed by the TU Wien Center for Research Data Management and hosted by TU.it.

Content types: [Raw data](#) [Scientific and statistical data formats](#) [Source code](#) [Standard office documents](#)


Metadata standards: [DataCite Metadata Schema](#) [Dublin Core](#)

Languages: [ENG](#)

<https://researchdata.tuwien.at/>

Use repository

Machine-actionable DMPs internationally



RDA DMP Common Standard for Machine-actionable Data Management Plans

The Challenge:
Data Management Plans are free-form text documents describing the data that is used and produced during the course of research activities. They specify where the data will be archived, which licenses and constraints apply, and to whom credit should be given, etc. The workload and bureaucracy often associated with traditional DMPs can be reduced when they become machine-actionable.



Produced by: **DMP Common Standards WG**
<https://www.rd-alliance.org/groups/dmp-common-standards-wg>

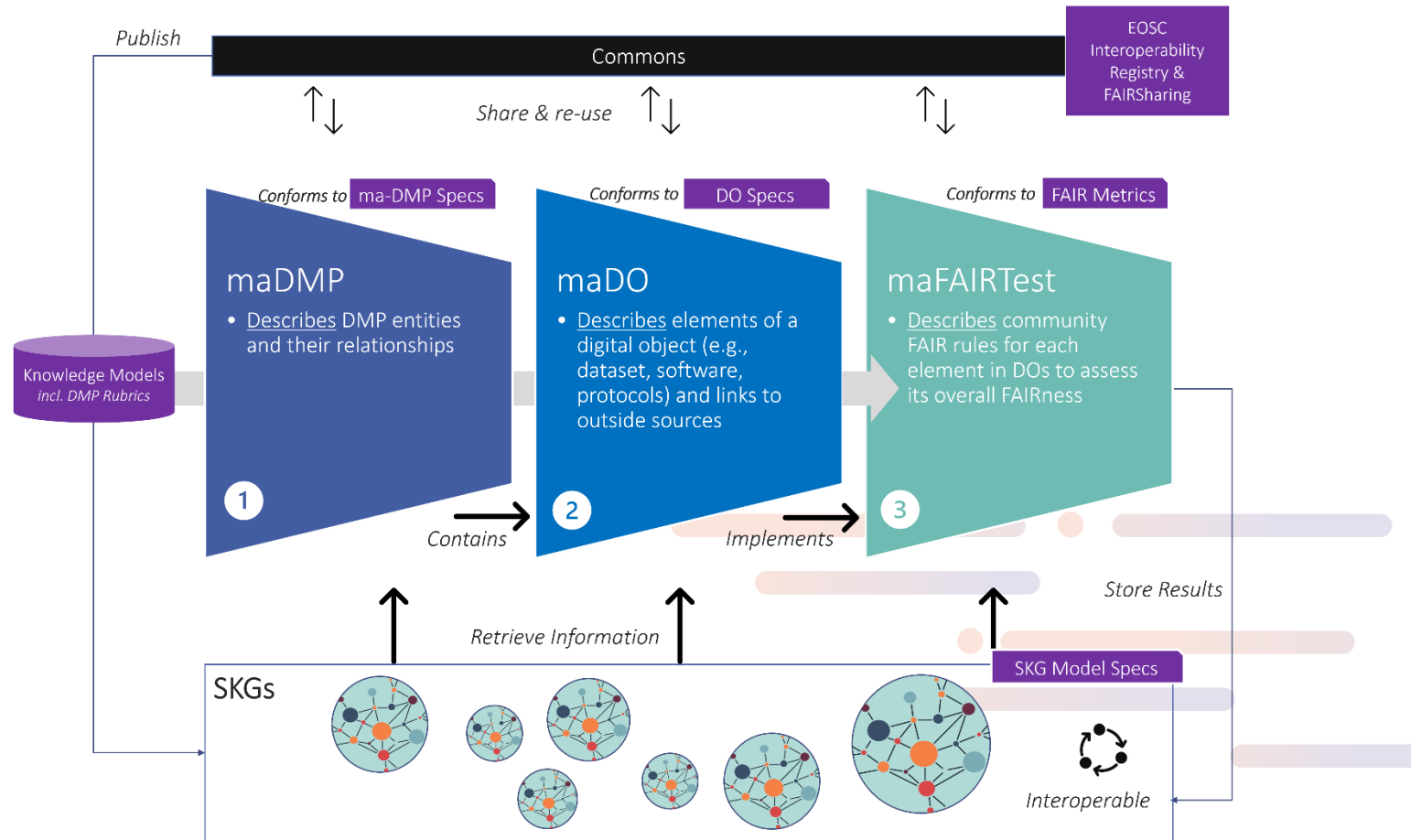


www.damap.org



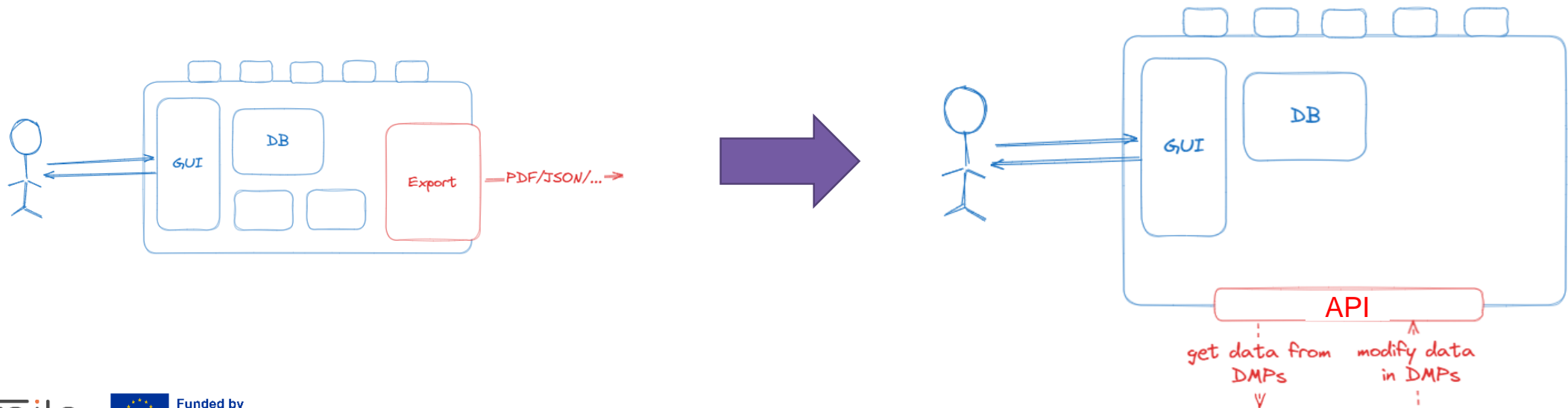
Open Science Trails (OSTrails)

- Horizon Europe Project
 - 2024 – 2027
 - Lead: OpenAIRE
 - Technical Lead: TU Wien
 - 8M EUR
 - 39 partners
 - 24 pilots
 - 17 national infrastructures
 - 5 science clusters



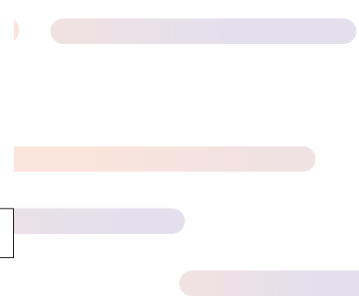
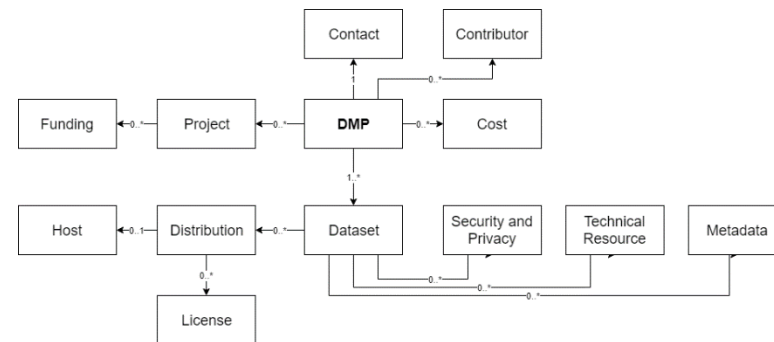
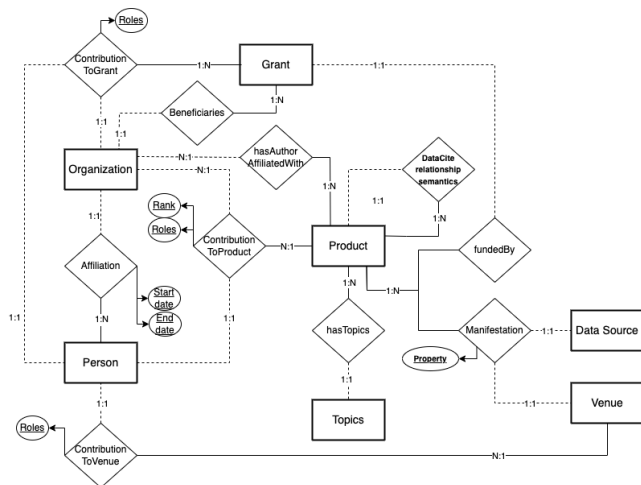
maDMPs in OStrails

- DMP Platforms come together
 - Common API to access/modify maDMPs
 - Alignment with Scientific Knowledge Graphs
 - Automated DMP Assessment and FAIRness indicators



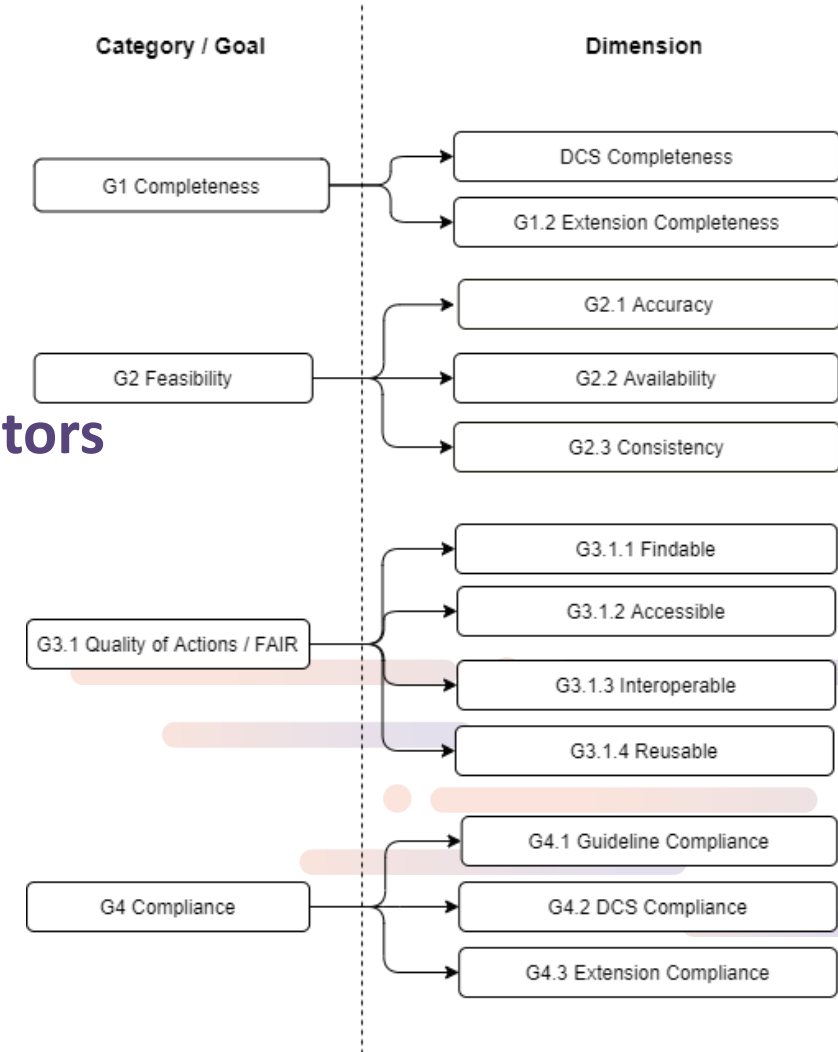
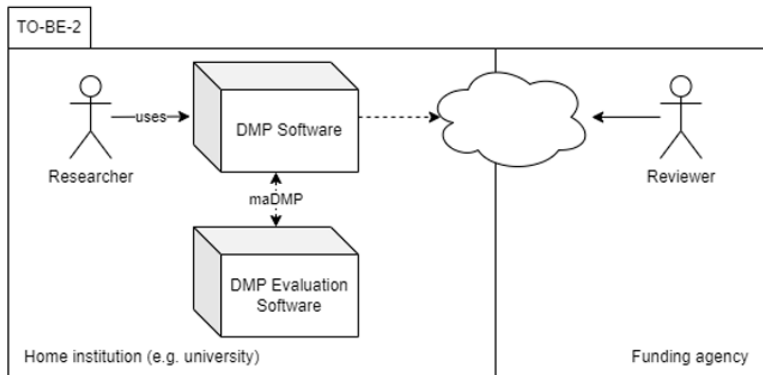
maDMPs in OStrails

- DMP Platforms come together
 - Common API to access/modify maDMPs
 - **Alignment with Scientific Knowledge Graphs**
 - Automated DMP Assessment and FAIRness indicators



maDMPs in OStrails

- DMP Platforms come together
 - Common API to access/modify maDMPs
 - Allignemnt with Scientific Knowledge Graphs
 - **Automated DMP Assessment and FAIRness indicators**



Summary

- The transformation in data management planning is happening now
- We already have many ingredients
 - DMP Platforms
 - RDA DMP Common Standard
 - SKGs, Registries, Repositories, ...
- We still have a way to go together
 - Further standardisation and alignment
 - DMP Assessment
 - Dropping PDFs...

