

2022 Webinar Series

Highlighting RDA Outputs

Hosted by RDA-US



DMP Common Standards WG Webinar

Integrating Research Data Management Services using
Machine-actionable Data Management Plans

November 22, 2022



rd-alliance.org



[@resdatall](https://twitter.com/resdatall) | [@RDA_US](https://twitter.com/RDA_US)



Individual Members

12,861 members
from 148 countries

69% Academic & Research
14% Public Administration
11% Enterprise & Industry
6% Other

RDA Groups

33 Working Groups
54 Interest Groups
1 Community of Practice



rd-alliance.org



[@resdatall](https://twitter.com/resdatall) | [@RDA_US](https://twitter.com/RDA_US)



30 National Groups

Asia	Czech Republic	Italy	Slovenia
Australia	Denmark	Lithuania	So Eastern Europe
Austria	Estonia	Hungary	Spain
Brazil	Europe	Ireland	Sweden
Bulgaria	Finland	Netherlands	Region of the Americas
Canada	France	Norway	United States
Costa Rica	Germany	Portugal	United Kingdom
Croatia	Greece	Romania	

5 Contributing Regions

Australia
Canada
EU
France
United States

www.rd-alliance.org/groups/rda-us

www.rd-alliance.org/groups/national-groups



rd-alliance.org



@resdata11 | @RDA_US



72 Organizations



rd-alliance.org



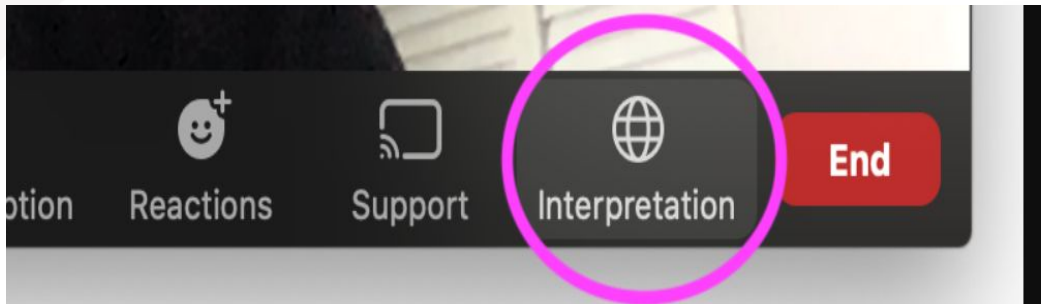
@resdata | @RDA_US





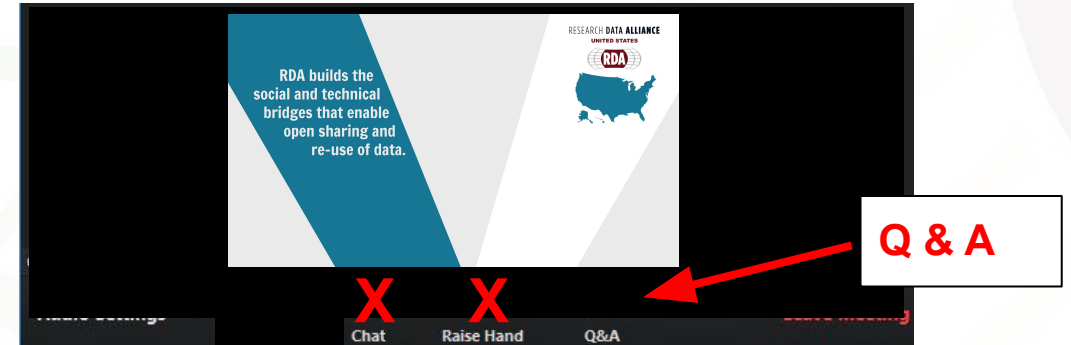
Translation

This **event is being live translated from English to French and Spanish**. To hear the translation select the Globe on the Zoom control area.



Questions and Answers

Please post to Q & A section of Zoom



Recording

Please note, this **event is being recorded**, and a links to the recording will be available on webinar page on the RDA website.



rd-alliance.org



[@resdatall](https://twitter.com/resdatall) | [@RDA_US](https://twitter.com/RDA_US)



2022 Webinar Series

Highlighting RDA Outputs

Hosted by RDA-US

Integrating Research Data Management Services using maDMPs

RESEARCH DATA ALLIANCE
UNITED STATES



Tomasz Miksa

TU Wien and SBA
Research

Austria



Elli Papadopoulou

ATHENA Research Center
and OpenAIRE Information
Management Systems
Institute

Greece



Maria Praetzellis

California Digital
Library / UC3 Curation
Center

United States



Marek Suchánek

CTU in Prague /
ELIXIR-CZ

Czech Republic

Register Here: bit.ly/NovemberDMP

#RDAwebinar

Nov. 22nd

17:00 UTC



rd-alliance.org



@resdataall | @RDA_US





Integrating RDM Services using maDMPs

Tomasz Miksa

DMP Common Standards WG

2022 Webinar Series Highlighting RDA Outputs

➤ Introduction

➤ Data Management Plans

➤ Machine-actionable DMPs

➤ Context

➤ RDA Recommendation

➤ What it is

➤ Example

➤ Integrating RDM Services

➤ RDA working group and further reading

Data Management Plans (DMPs)

Data Officer	Who is responsible for the data management and the DMP of the project (name/email address)?
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>

Directorate for Engineering Data Management Plans Guidance for Principal Investigators *updated: November 2018*

The Directorate for Engineering (ENG) supports research covering a broad spectrum of communities of investigators, and each community has its own best practices. ENG is aware of the need to provide flexibility to programs, principal investigators (PIs), and reviewers in assessing the quality of individual Data Management Plans (DMPs) from various communities. Therefore, guidance has evolved to accommodate changing community standards and expectations. ENG relies on the merit review process to determine which DMPs best serve each community.

The following guidance is to assist ENG investigators, reviewers and Program Officers in developing and evaluating effective, complete, and competitive DMPs. It is important to recognize that while all DMPs should address the five categories of information as specified in the PAPPG, they should not be generic. Each DMP should appropriately identify the data, metadata, samples, software, algorithms, curricula, documentation, publications, and other materials generated in the course of the proposed research. Moreover, the DMPs should describe how these materials will be disseminated, made accessible, and archived while incorporating the best practices and standards for the proposed research. DMPs are subject to peer review. Please contact your specific Program Officer if you have any questions related to DMPs in the program context.

PAPPG and NSF-WIDE REQUIREMENTS

All proposals must include a supplementary document of no more than two pages labeled "Data Management Plan," as described in [PAPPG Chapter II.C.2.j](#). The DMP is NOT part of the 15-page Project Description. *Proposals that do not include a Data Management Plan will be returned without review.*

You may request funds to cover costs of publication, page charges, or preparation of data as a direct cost in your budget proposal, which is evaluated as part of the merit review process. Any costs associated with implementing the DMP should be explained in the Budget Justification.

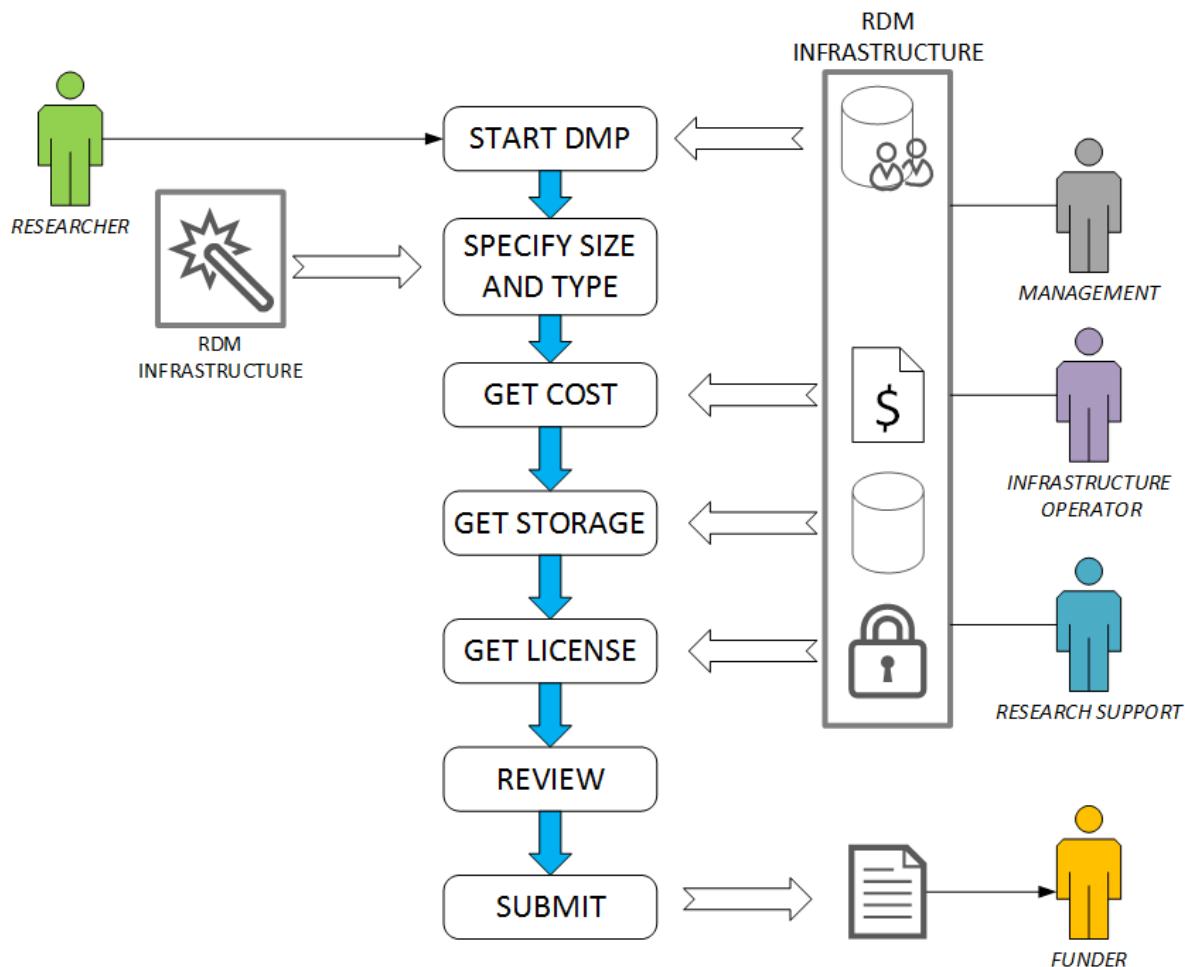
Some NSF Program Solicitations may contain specific and/or additional instructions that deviate from this guidance and/or provide exceptions to the two-page limit. Instructions in the solicitation take precedence over this guidance. Please check solicitations carefully for this information.

DATA MANAGEMENT PLAN (DMP) CONTENT


The DMP should clearly articulate how the investigators plan to manage and disseminate both the physical and digital data generated by the project, taking advantage of emerging information

Machine-actionable DMPs (maDMPs)

- Machine-actionable DMPs
 - Living documents
 - automate data management
 - collect information from systems
 - trigger actions in systems
 - facilitate validation
- This requires
 - well-defined RDM workflows
 - data management infrastructure
 - common standard to represent information




Official RDA Recommendation on maDMPs



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
Tomasz 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 of 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

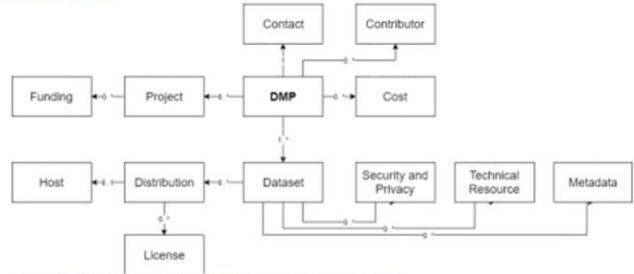


Figure 1: Overview of the application profile for the machine-actionable DMPs.

1

maDMPs - documentation

Properties in 'dmp'

Name	Description	Data Type	Cardinality	Example Value
contact	Contact person for a DMP	Nested Data Structure	1	
contributor	To list people that play role in data management related to this DMP, e.g. responsible for performing actions described in this DMP.	Nested Data Structure	0..n	
cost	To list costs related to data management. Providing multiple instances of a 'Cost' allows to break down costs into details. Providing one 'Cost' instance allows to provide one aggregated sum.	Nested Data Structure	0..n	
created	Date and time of the first version of a DMP. Must not be changed in subsequent DMPs.	DateTime	1	2019-03-13 13:13
dataset	To describe data on a non-technical level.	Nested Data Structure	1..n	

NOT a questionnaire!
NOT a template!

Most fields are optional!

Machine-actionable DMP

➤ Example: <https://doi.org/10.5281/zenodo.6467730>

```
"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.

Machine-actionable DMP

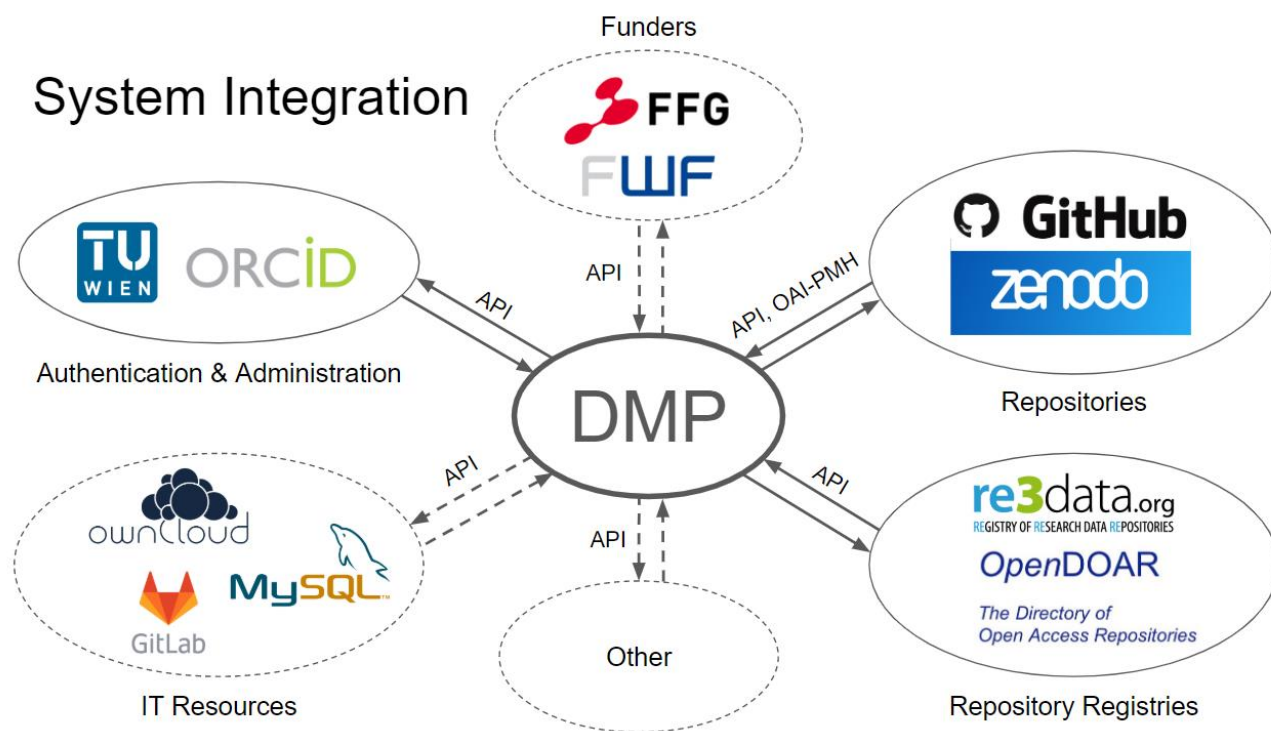
```
"dataset" : [ {  
  "description" : "For each dataset (fish and employee) the original dataset will be split into two subsets, one for training and one for testing the  
  performance.",  
  "distribution" : [ {  
    "access_url" : "https://zenodo.org/record/6467615",  
    "byte_size" : 2999302,  
    "data_access" : "open",  
    "description" : "For each dataset (fish and employee) the original dataset will be split into two subsets, one for training and one for testing the  
    performance.",  
    "format" : [ "STRUCTURED_TEXT" ],  
    "host" : {  
      "description" : "ZENODO builds and operates a simple and innovative service that enables researchers, scientists, EU projects and institutions to share  
      and showcase multidisciplinary research results (data and publications) that are not part of the existing institutional or subject-based repositories of  
      the research communities.\nZENODO enables researchers, scientists, EU projects and institutions to:\neasily share the long tail of small research results  
in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science.\ndisplay their research results and get  
credited by making the research results citable and integrate them into existing reporting lines to funding agencies like the European  
Commission.\neasily access and reuse shared research results.",  
      "pid_system" : [ "doi" ],  
      "storage_type" : "other",  
      "support_versioning" : "unknown",  
      "title" : "Zenodo",  
      "url" : "https://zenodo.org/"  
    },  
  },  
  "license" : [ {  
    "license_ref" : "https://creativecommons.org/licenses/by/4.0/",  
    "start_date" : "2022-05-01 22:00:00.0"  
  } ],  
  "title" : "Training and Test Subsets for Performance Comparison of kNN and GD"
```

Each **dataset** has a **title** and a human readable **description**.

It is also clear what the **format**, **size** and the **location** of the dataset are.

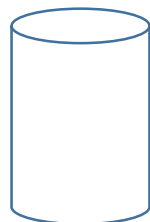
License and mode of **access**, including any exact **embargo** periods, are specified as well.

- maDMPs are the 'glue' between different systems
 - Automate getting information **in** and **out**

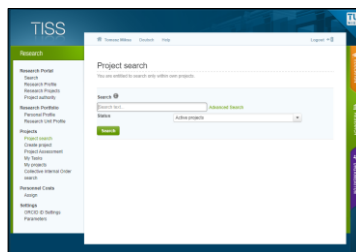


Example: TU Wien in Austria

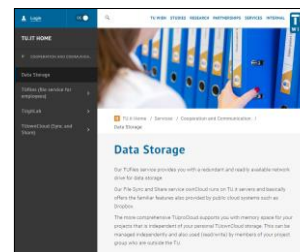
Knowledge Graph



CRIS



IT Storage



API

API

list of services

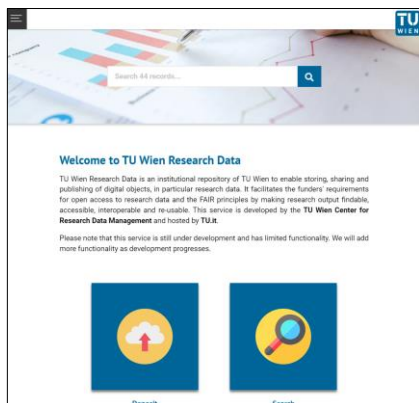
maDMP

maDMP

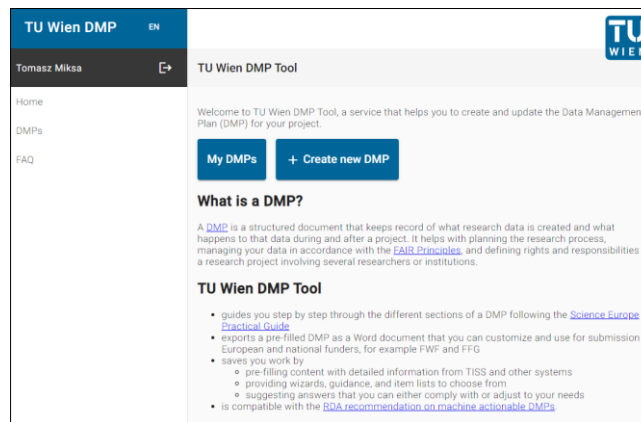
DMP

Funder integration?

Automated assessment and feedback?



Data Repository



www.damap.org

Adoptions (selected)



› Slides from all our sessions are in the repository

<https://www.rd-alliance.org/node/56938/file-repository>

› Join the group!

We have **250+** members!



WG

DMP Common Standards WG

Taxonomy:

Posts

Wiki

Events

Repository

Outputs

Case Statements

Plenaries

Members

create new content

Group Status:

WGs Maintaining deliverables (maintenance group)

You are the group manager

›

status: Recognised & Endorsed

Chair (s): Paul Walk, Peter Neish, Tomasz Miksa

Group Email: dmp-common@rda-groups.org

Secretariat Liaison: enquiries[at]rd-alliance.org

File Repository

VP17 Edinburgh

by Tomasz Miksa

22 April 2021

Attachment	Size
2021-RDA-DMP-VP17.pdf	3.71 MB
1-Zeno-FairDataAustria-DMAP.pdf	1.05 MB
3-Elli-Argos.pdf	1.66 MB
4-Fajar-DCSOntology.pdf	1.82 MB

VP16 Costa Rica Slides

by Tomasz Miksa

12 November 2020

Slides from the plenary session at the VP16: * 2020-RDA-DMP-VP16 - main deck of slides 1 - Claire Austin - maDMPs in a government context 2 - Maria Praetzellis - DMPHub 3 - Lucas Berent, Alexandre

Attachment	Size
1-madmps-goverment.pdf	962.26 KB
2-dmphub.pdf	1.34 MB
3-madmps-exposing.pdf	1.87 MB
4-madmps-repositories.pdf	794.44 KB
5-argos-knowledge-graph.pdf	2.09 MB
6-ro-crates-and-madpms.pdf	1.16 MB
2020-RDA-DMP-VP16.pdf	2.78 MB

Read more in...

- Describes the full story of developing the recommendation
- Example of a minimal maDMP
- Presents adoptions
 - Haplo
 - Open Research Publishing Platforms
 - DMP Tool
 - DMPonline
 - DMP OPIDoR
 - Data Stewardship Wizard
 - NSD DMP
 - Argos
 - Research infrastructure at TU Wien
 - Easy DMP

 **DATA SCIENCE JOURNAL**

Reading: Application Profile for Machine-Actionable Data Management Plans

Share: [f](#) [t](#) [g](#) [in](#)

Special Collection: Research Data Alliance Results

Research Papers

Application Profile for Machine-Actionable Data Management Plans

Authors: Tomasz Miksa , Paul Walk, Peter Neish, Simon Oblasser, Hollydawn Murray, Tom Renner, Marie-Christine Jacquemot-Perbal, João Cardoso, Trond Kvamme, Maria Praetzellis, Marek Suchánek, Rob Hooft, Benjamin Faure, Hanne Moa, Adil Hasan, Sarah Jones

Abstract

This paper presents the application profile for machine-actionable data management plans that allows information from traditional data management plans to be expressed in a machine-actionable way. We describe the methodology and research conducted to define the application profile. We also discuss design decisions made during its development and present systems which have adopted it. The application profile was developed in an open and consensus-driven manner within the DMP Common Standards Working Group of the Research Data Alliance and is its official recommendation.

Keywords: application profile, maDMPs, common standard, machine actionable, RDA

How to Cite: Miksa, T., Walk, P., Neish, P., Oblasser, S., Murray, H., Renner, T., Jacquemot-Perbal, M. -C., Cardoso, J., Kvamme, T., Praetzellis, M., Suchánek, M., Hooft, R., Faure, B., Moa, H., Hasan, A. and Jones, S., 2021. Application Profile for Machine-Actionable Data Management Plans. *Data Science Journal*, 20(1), p.32. DOI: <http://doi.org/10.5334/dsj-2021-032>

455	49	28
Views	Downloads	Twitter

 Published on 26 Oct 2021	 Peer Reviewed	 CC BY 4.0
 Accepted on 12 Oct 2021	 Submitted on 14 Jul 2020	

<http://doi.org/10.5334/dsj-2021-032>

- Key elements of the whole ecosystem needed to make DMPs machine-actionable

- Enterprise Architecture that uses maDMPs
- Examples of tasks automation at institutions using maDMPs

***ACM Transactions
on Management Information Systems***

<https://doi.org/10.1145/3490396>

Publications about maDMPs

- [Tomasz Miksa, Simon Oblasser, and Andreas Rauber. **Automating research data management using machine-actionable data management plans**. ACM Transactions on Management Information Systems, 13\(2\), dec 2021.](#)
- [Tomasz Miksa, Paul Walk, Peter Neish, Simon Oblasser, Hollydawn Murray, Tom Renner, Marie-Christine Jacquemot-Perbal, João Cardoso, Trond Kvamme, Maria Praetzellis, Marek Suchánek, Rob Hooft, Benjamin Faure, Hanne Moa, Adil Hasan, and Sarah Jones. **Application profile for machine-actionable data management plans**. CODATA Data Science Journal, 20\(1\):32, October 2021](#)
- [Raffael Foidl, Lea Salome Brugger, and Tomasz Miksa. **Automating Evaluation of Machine-Actionable Data Management Plans with Semantic Web Technologies**. In DaMaLOS - 2nd Workshop on Data and Research Objects Management for Linked Open Science : Co-located at the International Semantic Web Conference ISWC 2021. PUBLISSO, November 2021.](#)
- [Tomasz Miksa, Maroua Jaoua, and Ghaith Arfaoui. **Research Object Crates and Machine-actionable Data Management Plans**. In DaMaLOS - First Workshop on Data and Research Objects Management for Linked Open Science : Co-located at the International Semantic Web Conference ISWC 2020. PUBLISSO, November 2020.](#)
- [João Cardoso, Leyla Jael Garcia Castro, Fajar Ekaputra, Marie-Christine Jacquemot-Perbal, Tomasz Miksa, and José Borbinha. **Towards semantic representation of machine-actionable Data Management Plans**. In DaMaLOS - First Workshop on Data and Research Objects Management for Linked Open Science : Co-located at the International Semantic Web Conference ISWC 2020. PUBLISSO, 2020.](#)
- [Simon Oblasser, Tomasz Miksa, Asanobu Kitamoto: **Finding a Repository with the Help of Machine-Actionable DMPs: Opportunities and Challenges**. IDCC 2020](#)
- [Tomasz Miksa, Stephanie Simms, Daniel Mietchen, Sarah Jones \(2019\) **Ten principles for machine-actionable data management plans**. PLOS Computational Biology 15\(3\): e1006750.](#)
- [Tomasz Miksa, Peter Neish, Paul Walk, Andreas Rauber: **Defining requirements for machine-actionable Data Management Plans**. iPres 2018](#)
- [Tomasz Miksa, João Cardoso, José Luis Borbinha: **Framing the scope of the common data model for machine-actionable Data Management Plans**. BigData 2018: 2733-2742](#)
- [Asztrik Bakos, Tomasz Miksa, Andreas Rauber: **Research Data Preservation Using Process Engines and Machine-Actionable Data Management Plans**. TPDL 2018: 69-80](#)

NEXUS

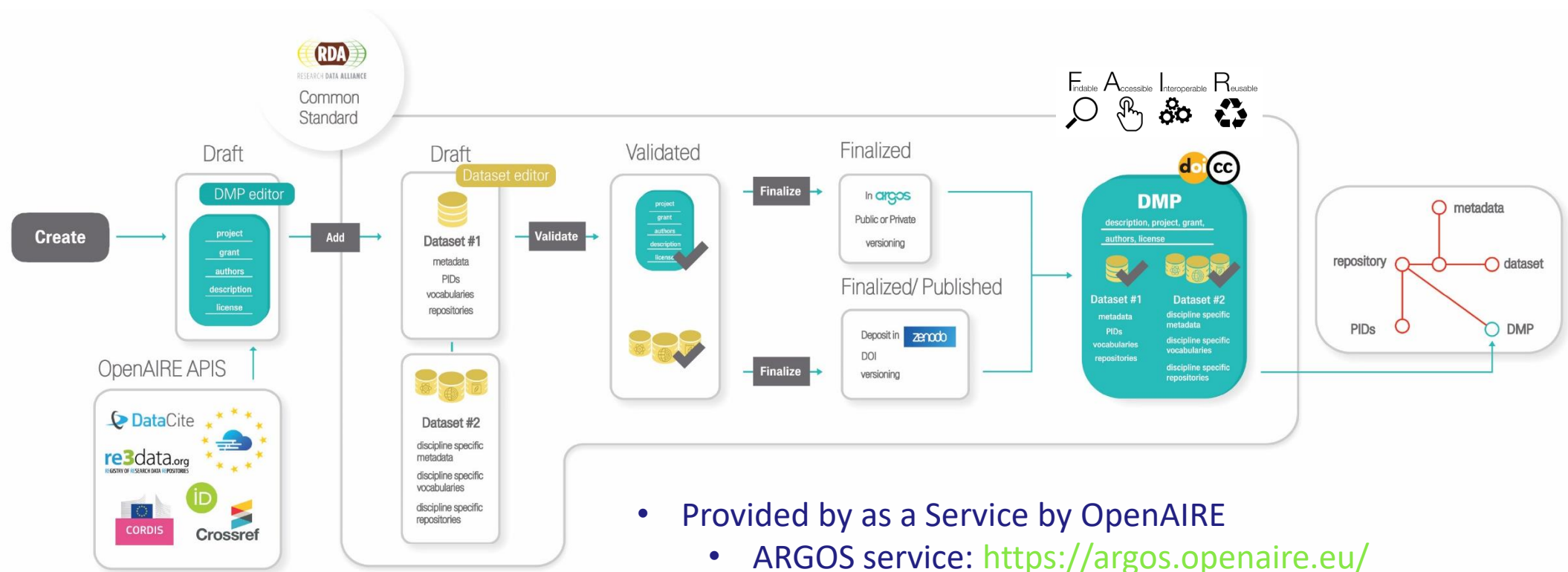
argos

DCS adoption

Elli Papadopoulou
ATHENA RC / OpenAIRE

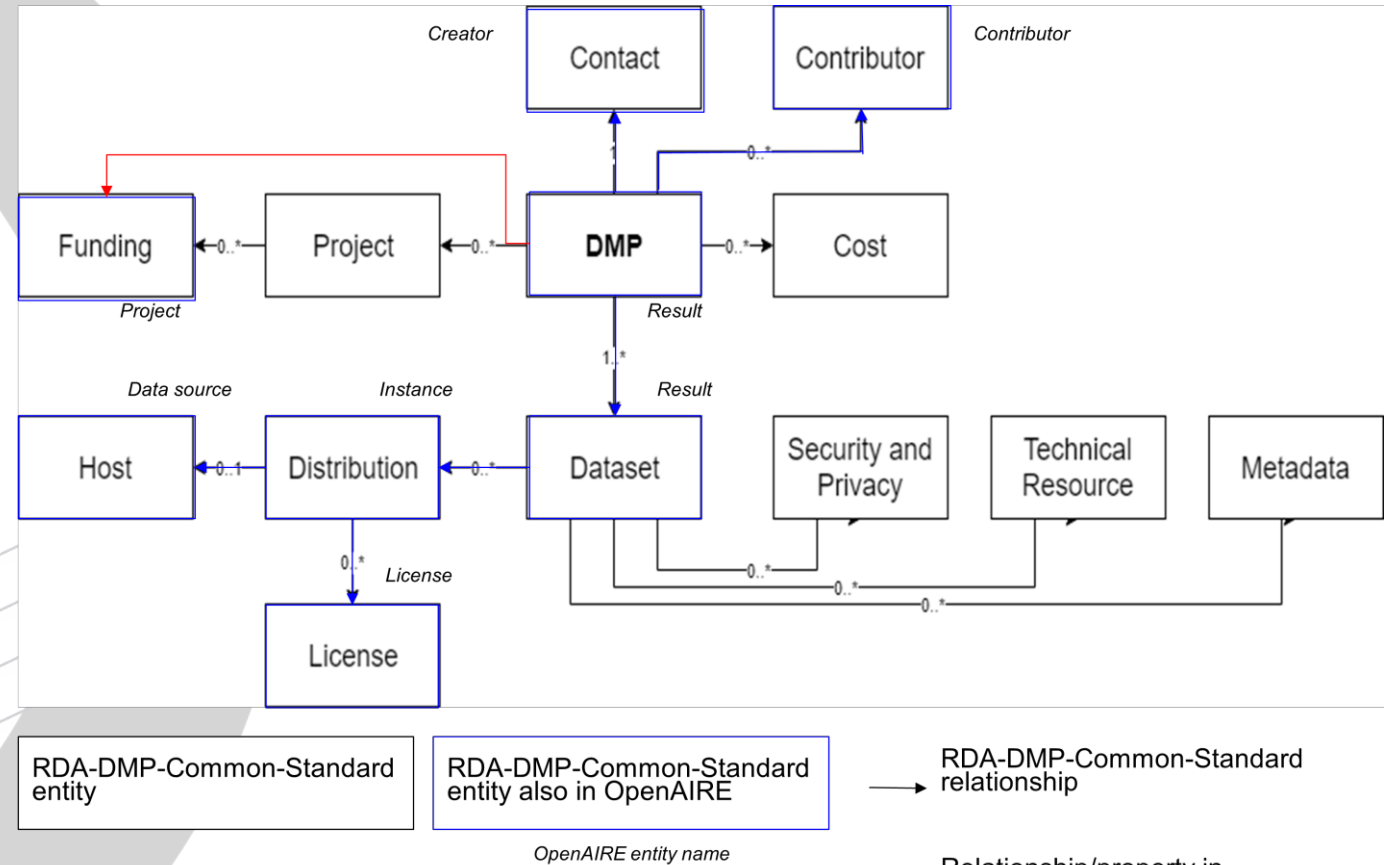
ARGOS – A GLIMPSE

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

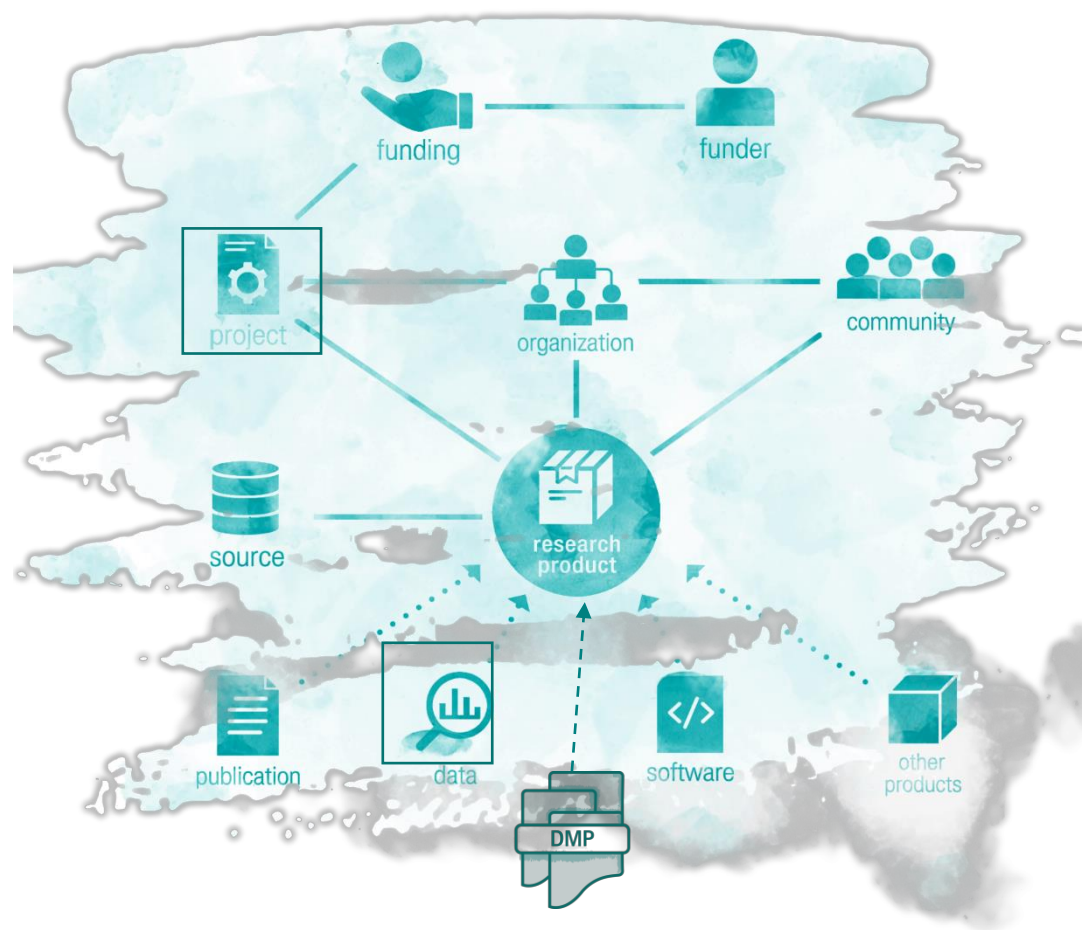


- Provided by as a Service by OpenAIRE
- ARGOS service: <https://argos.openaire.eu/>
[EOSC Resource \(eosc-portal.eu\)](https://eosc-portal.eu/)

MADMPs AND RESEARCH GRAPH



NEW - ENTITIES AND RELATIONSHIPS



OpenAIRE Research Graph

- “hasDMP”/”hasProject”
- “hasDataset”
- “hasDMP”

ARGOS

- Contact
- Cost
- Language
- License


(More in Argos ma-DMPs
ethics, security, quality, and preservation)

DISCOVERABILITY

- EXPLORE

- Adoption of term “data-management-plan” by repositories
 - disambiguation of DMPs
- Disambiguation of information contained in DMPs
 - E.g. [OpenOrgs](#) for Organizations

5

 OpenAIRE | EXPLORE

SEARCH

RESEARCH OUTCOMES (84) PROJECTS (2,812,592) CONTENT PROVIDERS (78,064) ORGANIZATIONS (147,114)

Filters

X Data Management Plan

Access Mode (1)

☐ Open Access (82)

Result Types (4)

☐ Publications
☐ Research data
☐ Software
☐ Other research products

Year range

e.g. 1800

 -

e.g. 2030

>

THIS YEAR | LAST 5 YEARS | LAST 10 YEARS

Funder (1)

☐ European Commission (23)

Type (4) Clear

☒ Data Management Plan (84)
☐ Other literature type (62)
☐ Article (13)
☐ Preprint (2)

Language (3)

☐ English (36)
☐ Portuguese (2)
☐ French (1)

Results per page: 10

Sort by: Relevance

Download Results

84 RESEARCH OUTCOMES, PAGE 4 OF 9

< 2 3 4 5 6 >

Publication . Data Management Plan . Article . 2017

Data Management Plan: Opening access to economic data to prevent tobacco related diseases in Africa

OPEN ACCESS ENGLISH

Authors: Woolfrey, Lynn;








Publisher: Pensoft Publishers



The purpose of this project is to demonstrate that tobacco-related data from selected Africa countries can be collected and distributed from an Open Data platform. The platform and data will improve the capacity for tobacco control research in key sub-Saharan African co...

Publication . Other Literature Type . Data Management Plan . 2019

The data management plan of Alien-CSI

OPEN ACCESS ENGLISH

Authors:  Groom, Quentin John;  Adriaens, Tim; Cardoso, Ana Cristina;  Essi, Franz; Martinou, Kelly; Moen, Toril Loennechen;  Jan Perg, Jan;  Pocock, Michael;  Reysenrove, Lien;  Schade, Sven; ...

DOI: 10.5281/zenodo.3265765  10.5281/zenodo.3265764 


Publisher: Zenodo




This document describes the data management principles and policies of the Alien CSI COST Action. This action aims to increase public awareness and levels of participation on issues related to invasive alien species and citizen science.

Publication . Data Management Plan . Other Literature Type . 2020

WorldBank_WDI_Visualization - DMP

OPEN ACCESS ENGLISH

Authors:  Lincoln, Craig N;

DOI: 10.5281/zenodo.3757220  10.5281/zenodo.3757219  10.5281/zenodo.3757520 

Initial DMP

DISCOVERABILITY - VOCABULARIES

↑ -- data management plan --

URI	http://purl.org/coar/resource/type/c_ah20
Definition (en)	A formal document describing the data management plan of a research project, which complies with the content policy of the community.
Preferred label (en)	data management plan
Preferred label (es)	plan de gestión de datos
Preferred label (ca)	pla de gestió de dades
Preferred label (fr)	plan de gestion des données
Preferred label (de)	Datenmanagementplan
Preferred label (ja)	データ管理計画
Preferred label (nl)	data management plan
Preferred label (ru)	план управления данными
Preferred label (it)	piano di gestione dei dati
Preferred label (sl)	načrt ravnanja s podatki
Alternative label (ca)	pla d'administració de dades
Alternative label (ru)	план осуществления
Alternative label (ru)	план регули
Alternative label (sl)	načrt upravljanja
Exact match	https://dictionary.com
Exact match	http://purl.org/coar/resource/type/c_ah20
Broader	project deliverable

Start typing a community name...



Argos DMP outputs

Upload type

- Book
- Book section
- Conference paper
- Data management plan**
- Journal article
- Patent
- Preprint
- Project deliverable
- Project milestone
- Proposal
- Report
- Software documentation
- Taxonomic treatment
- Technical note
- Thesis
- Working paper
- Other
- Journal article



Publication



Post

Publication type

OpenAIRE Guidelines

Search docs

CURRENT GUIDELINES

- OpenAIRE Guidelines for Literature Repositories
- OpenAIRE Guidelines for Data Archives
- OpenAIRE Guidelines for CRIS Managers
- Draft OpenAIRE Guidelines for Software Repository Managers
- Draft OpenAIRE Guidelines for Other Research Products

HOW TO CONTRIBUTE

Contributing

APPENDIX

How the Horizon 2020 Open Access requirements are met

<https://guidelines.openaire.eu/en/latest/>

MAJOR FINDINGS AND ADAPTATIONS

- **Import Tool and Export Tools enhanced for maDMPs exchange with other platforms**
- **Mappings:** When no direct mapping of entities, information ->
 - is still found in more abstract/ general fields of OpenAIRE or Argos, though diverged in cardinality and / or data type
 - might be covered by Argos ma-DMP outputs
 - can be tweaked to accommodate the needs of maDMPs documentation
 - rarely, it is omitted
- **Granularity**
 - Argos accommodates some entities on the level of datasets and not on the level of DMPs, e.g. ethical_issues_exist
- **Contextualisation of information** -> datasets in ma-DMPs vs datasets in DMPs

THANKS

Email

elli.p@athenarc.gr

Twitter

[@elli_lib](https://twitter.com/elli_lib)

ORCID

[0000-0002-0893-8509](https://orcid.org/0000-0002-0893-8509)

NETWORKING DATA MANAGEMENT PLANS

Maria Praetzelis, Product Manager, Research Data Management, CDL

RDA-US Webinar Series

Integrating Research Data Management Services Using maDMPs

November 22, 2022



University of California
CDL
California Digital Library

CDL was founded by the University of California in 1996

University of California Curation Center (UC3) is CDL's program concerned with maintaining, preserving, and adding value to digital research data throughout its lifecycle UC3 areas of focus:

- Research data management
- Data publication and data metrics
- Persistent identifiers
- Digital preservation
- Data/software skills training



AGENDA

1

Why are
maDMPs
important right
now?

2

How are we
building and
using maDMPs?

3

What's next?

WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY (OSTP) UPDATED U.S. POLICY



[Administration](#)

[Priorities](#)

AUGUST 25, 2022

OSTP Issues Guidance to Make Federally Funded Research Freely Available Without Delay



[OSTP](#) [BRIEFING ROOM](#) [PRESS RELEASES](#)

Today, the White House Office of Science and Technology Policy (OSTP) updated U.S. policy guidance to make the results of taxpayer-supported research immediately available to the American public at no cost. In a [memorandum](#) to federal departments and agencies, Dr. Alondra Nelson, the head of OSTP, delivered guidance for agencies to update their public access policies as soon as possible to make publications and research funded by taxpayers publicly accessible, without an embargo or cost. All agencies will fully implement updated policies, including ending the optional 12-month embargo, no later than December 31, 2025.

The new policy removes the optional 12-month embargo from public access to federally funded research.

This guidance covers all federal agencies with research and development expenditures. (Was only those >100M)

Recommends implementation no later than December 31, 2025.

WHAT'S NEW NIH DATA MANAGEMENT & SHARING POLICY?



Data Management and Sharing Plan

Policy is effective January 25, 2023.

Requires researchers seeking NIH funding to prospectively submit a plan outlining how scientific data from their research will be managed and shared.

Researchers should “maximize the appropriate sharing of scientific data.”

Data should be shared as soon as possible, and no later than the time of an associated publication or end of performance period (whichever comes first).

NSF CHIPS AND SCIENCE ACT OF 2022

Title III of the National Science Foundation for the Future, Section 10344 on research reproducibility and replicability establishes a requirement for...



...the inclusion of a machine-readable data management plan in award proposals. Requires the development of a set of criteria for trusted open repositories and provides support for the development of open data repositories to address any gaps. Requires the establishment of a single web-based point of access for data, software, and code resulting from Foundation-funded projects. Directs the Foundation to ensure that data resulting from Foundation-funded projects is made available in trusted open repositories. Supports research and development of tools and infrastructure to support research reproducibility.”



DUN22612 LM4

S.L.C.

AMENDMENT NO. _____ Calendar No. _____

Purpose: To improve the bill.

IN THE SENATE OF THE UNITED STATES—117th Cong., 2d Sess.

H. R. 4346

Making appropriations for Legislative Branch for the fiscal year ending September 30, 2022, and for other purposes.

Referred to the Committee on _____ and
ordered to be printed

Ordered to lie on the table and to be printed

AMENDMENT intended to be proposed by Mr. SCHUMER

Viz:

1 In lieu of the matter proposed to be inserted by the
2 amendment of the House to the amendment of the Senate,
3 insert the following:

4 **SECTION 1. TABLE OF CONTENTS.**

5 The table of contents for this Act is as follows:

Sec. 1. Table of contents.

Sec. 2. References.

DIVISION A—CHIPS ACT OF 2022

Sec. 101. Short title.

Sec. 102. Creating helpful incentives to produce semiconductors (CHIPS) for America fund.

Sec. 103. Semiconductor incentives.

Sec. 104. Opportunity and inclusion.

Sec. 105. Additional GAO reporting requirements.

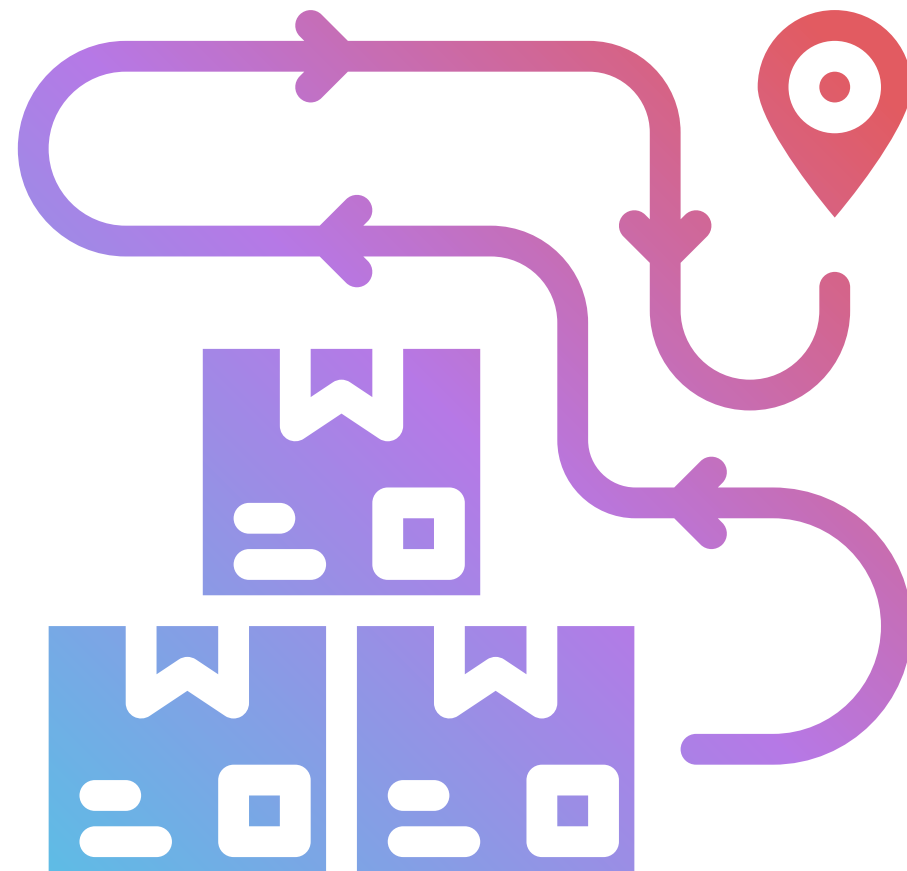
Sec. 106. Appropriations for wireless supply chain innovation.

Sec. 107. Advanced manufacturing investment credit.

DIVISION B—RESEARCH AND INNOVATION

Sec. 10000. Table of contents.

RESPONDING TO A CHANGING LANDSCAPE



With increasing mandates from federal funders, conversations are now turning toward implementation and compliance.

How can institutions demonstrate and track compliance with these new data-sharing requirements?

The next task in the drive toward building a robust research data infrastructure is to build on existing community-driven, open systems and workflows to track research outputs from project inception to long-term preservation.

DMPTOOL

A platform for DMP creation and guidance with templates for US funders (NSF, NIH, DOE, DOT, etc.) and many international funders.

Free, open-source community-supported tool working off of the shared DMPRoadmap codebase.

80K+ users from over 348 participating institutions around the world.



Sign in / Sign up

Email address *

For SSO, use institutional address.

Continue

Problems signing in? [Contact us.](#)



81,019 Users



348 Participating Institutions



78,554 Plans

Latest News from DMPTool

[Things to know about the updated DMPTool website](#)

[View all news](#)



DMPTOOL EDITORIAL BOARD



Heather L Barnes, PhD, Digital Curation Librarian, Wake Forest University

Raj Kumar Bhardwaj, PhD, Librarian, St Stephen's College, University of Delhi, India

Renata G. Curty, PhD, Social Sciences Data Curator, University of California, Santa Barbara

Jennifer Doty, Research Data Librarian, Emory University

Nina Exner, Research Data Librarian, Virginia Commonwealth University

Janice Hermer, Health Sciences Liaison Librarian, Arizona State University

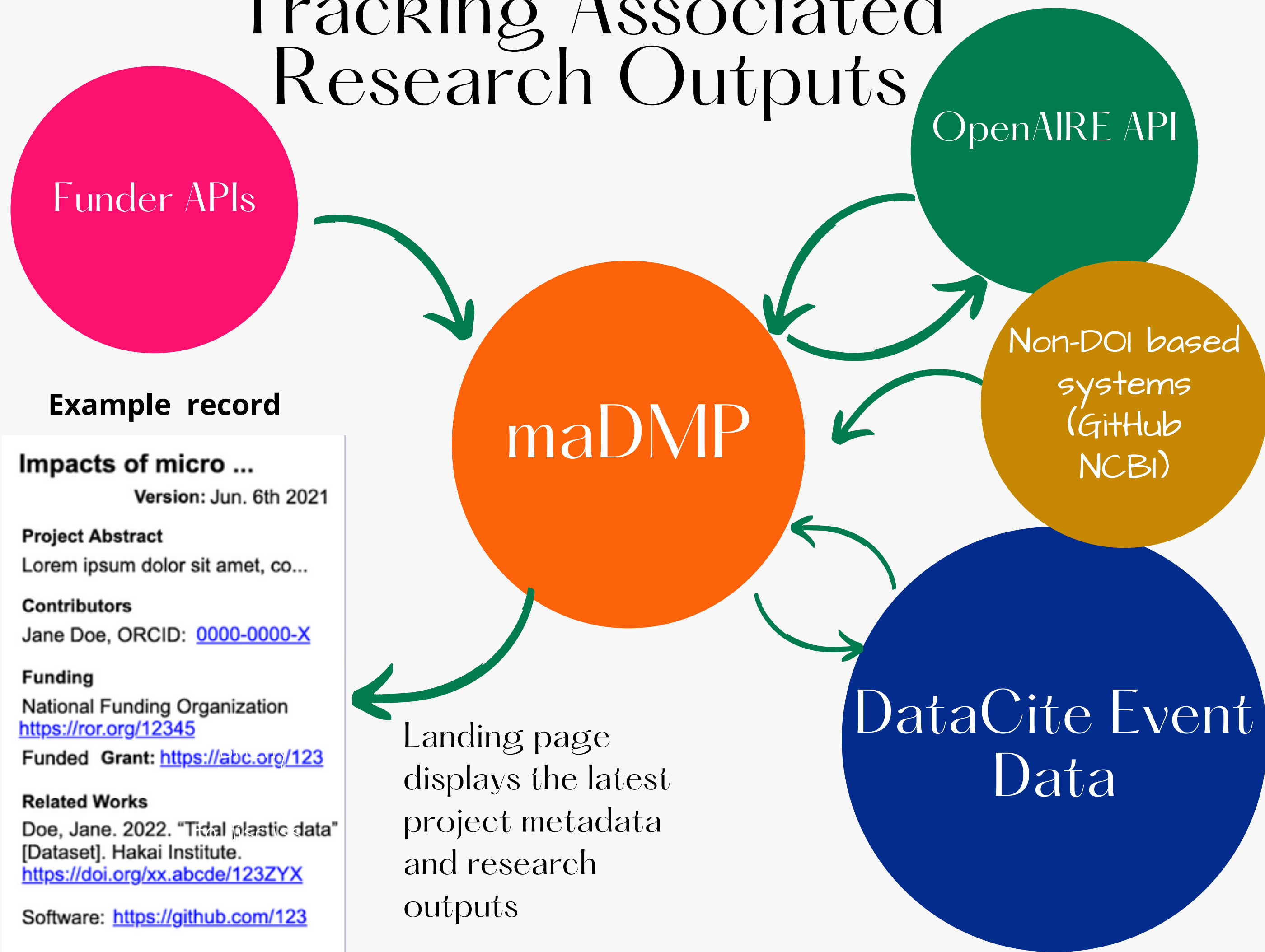
Megan O'Donnell, Data Services Librarian, Iowa State University

Reid Otsuji, Data Curation Specialist Librarian University of California, San Diego

Nick Ruhs, PhD, STEM Data & Research Librarian, Florida State University

Anna Sackmann, Data Services Librarian, University of California, Berkeley

Tracking Associated Research Outputs



Funder APIs

OpenAIRE API

Non-DOI based
systems
(GitHub
NCBI)

maDMP

DataCite Event
Data

Example record

Impacts of micro ...

Version: Jun. 6th 2021

Project Abstract

Lorem ipsum dolor sit amet, co...

Contributors

Jane Doe, ORCID: [0000-0000-X](#)

Funding

National Funding Organization

<https://ror.org/12345>

Funded Grant: <https://abc.org/123>

Related Works


Doe, Jane. 2022. "Tidal plastic data" [Dataset]. Hakai Institute.

<https://doi.org/xx.abcde/123ZYX>

Software: <https://github.com/123>

Landing page
displays the latest
project metadata
and research
outputs

Updated DMP-ID landing page with versioning



DMP ID: 10.48321/D1R316Version: 9 Nov 2022All Versions ▲


This page describes a data management plan written for the Tula Foundation using DMPTool.


2


3


Hakai Institute Juvenile Salmon Program Time Series

Contributors to this project

Brett Johnson: Data-curation, Investigation, Tula Foundation (tula.org),  <https://orcid.org/0000-0001-9317-0364>

Brian Hunt: Investigation, Tula Foundation (tula.org),  <https://orcid.org/0000-0003-4718-4962>

Krystal Bachen: Data-curation, Tula Foundation (tula.org),  <https://orcid.org/0000-0003-0886-157X>

Tim van der Stap: Data-curation, Tula Foundation (tula.org),  <https://orcid.org/0000-0002-0053-0795>

Project details

Research Domain: Earth and related environmental sciences

Project Start: May 12, 2015


Created: November 14, 2021

Modified: January 28, 2022

Ethical issues related to data that this DMP describes?: no

Citation

When citing this DMP use:

 Read
man

9 Nov 2022

30 Oct 2022

2 Oct 2022

15 Sep 2022

23 Dec 2021

maDMP Use Case: Field Station Directors



Problem: Field station directors have limited insight into the research outcomes and subsequent impact of research coming out of their field station/reserve.

Goal: As a field station director, I want to demonstrate the scientific impact of my field station so that I can demonstrate value and improve reporting for my administrators and Board.

Questions:

- What research outputs resulted from time spent on the station?
- Where were these outputs published?
- Where they cited/reused in subsequent publications?
- Where on the station was data gathered?
- Was the research we supported in compliance with university and government policies (including ethical, legal, and social aspects)?

maDMP Use Case: Research administrators

Problem: Research administrators need to demonstrate compliance with each award term. They are also interested in ensuring that all relevant data outputs are associated correctly with their affiliation to count towards citation counts.

Goal: To produce a report of data deposits and publications in case of an audit.

Questions:

- What outputs are planned for a given research project & what are the institutional responsibilities for data sharing?
- What outputs were generated during the research, and where were they deposited?
- Individuals and their contact info and which outputs they are responsible for.

The current state of maDMPs

Ability to link a plan to research outputs the project has generated– such as journal articles, preprints, datasets, data papers, and software. Example

Ability to access DMPs as structured JSON files complying with the RDA Common Standard Metadata Scheme for DMPs. Can be done via the UI or API.

The DMPTool currently supports the following identifiers within a DMP:

- IDs for DMPs
- RORs for research organizations
- Funder Registry IDs for funders
- ORCiDs for DMP creators and collaborators
- Registry of Research Data Repositories
- RRID (in development)

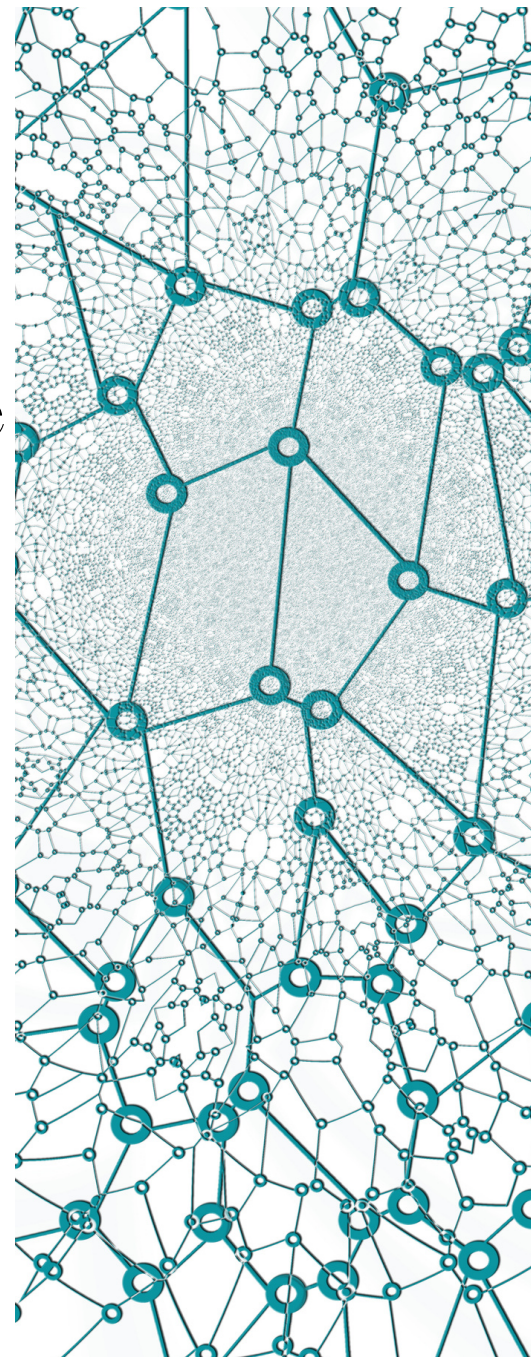


DMPTool

Scaling maDMPs

Our approach

The coupling of machine-actionable DMPs with PID-enabled infrastructure facilitates the effective tracking of projects over time, allowing these relationships to be shared with the larger research data ecosystem in an open and interoperable manner.



In development

We are working to PIDify DMPs at scale with the ability to upload externally created DMPs and generate DMP-IDs for them.

New output dashboards to track research outputs connected with DMPs

Maria Praetzellis

THANK YOU!

maria.praetzellis@ucop.edu



University of California

CDL

California Digital Library



UC Curation Center



DCS Adoption in Data Stewardship Wizard



**FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE**

Marek Suchánek
marek.suchanek@ds-wizard.org

- DSW allows to export DMPs (documents) by using a *document template*
- Document template uses Jinja2 language to specify the transformation
- In 2021, we implemented document template for RDA DMP Common Standard in both JSON and RDF formats (based on the ontology made by Cardoso et al.)
- We mapped our DSW Common Knowledge Model (questionnaire structure based on mind map of Rob Hooft) to maDMP specification and make some adjustments in the KM

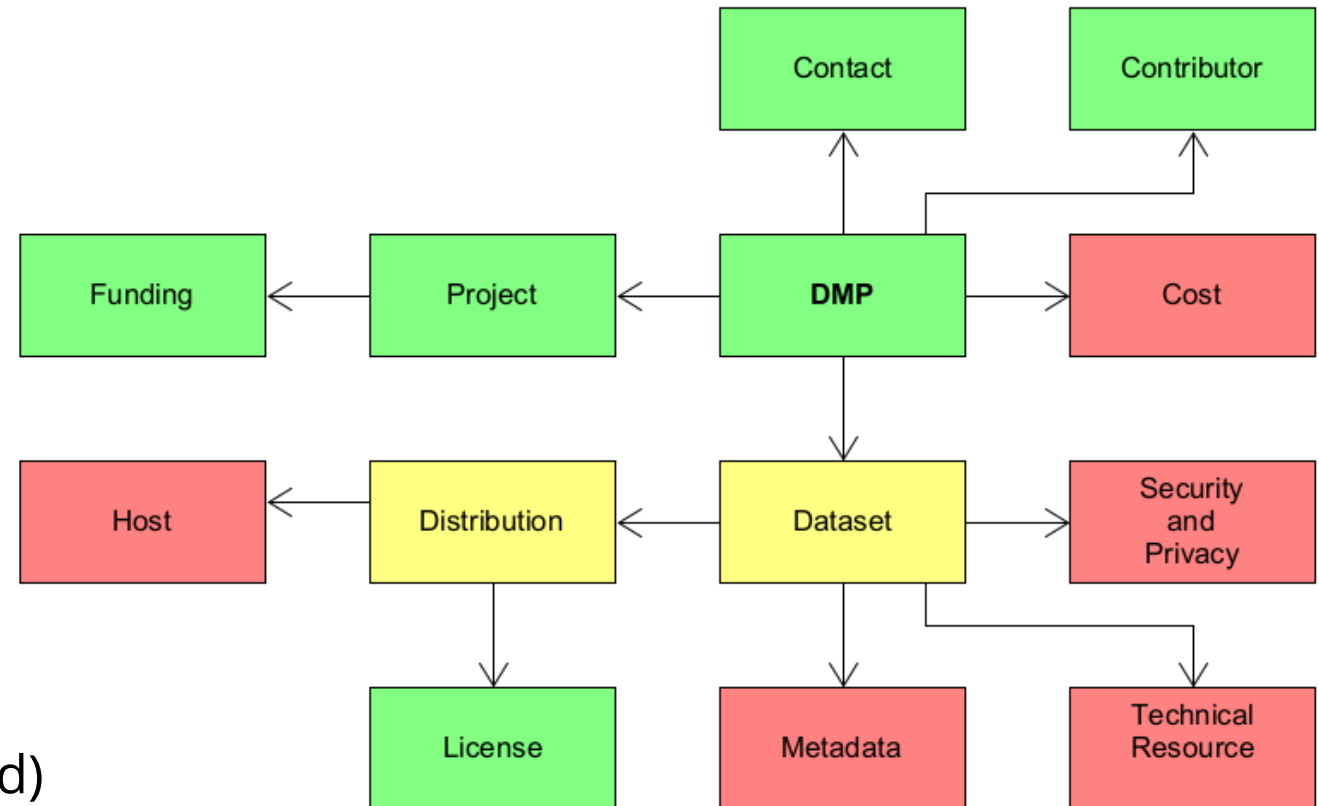
DSW maDMP Export (mapping)



- The coverage is not complete, there are many concepts in questionnaire that are not in maDMP and some also vice-versa

- Future plans:

- Improve dataset/distribution
- Include costs
- Include other missing info attached to datasets and distributions
(text synthesizing may be needed)



DSW maDMP Export (example)



New Document

Name

Example DMP

Answered (current phase): 0/20

Answered: 0/52

Document Template



maDMP (RDA DMP Common Standard) 1.13.1

Machine-actionable DMP according to RDA Common Standard

Format

☒ JSON

☐ Turtle (no blank nodes)

☐ Turtle

☐ N3

☐ RDF/XML

☐ JSON-LD

☐ N-Triples

☐ Trig

Cancel

Create

```
Questionnaire Metrics Preview Documents Settings

{
  "dmp": {
    "contributor": [],
    "created": "2022-11-14T20:24:36Z",
    "dataset": [
      {
        "dataset_id": {
          "identifier": "",
          "type": "other"
        },
        "description": "raw instrument data",
        "personal_data": "no",
        "sensitive_data": "no",
        "title": "HPLC ICP-MS analysis"
      }
    ],
    "description": "This maDMP has been created using Data Stewardship Wizard (DSW, ds-wizard.org) and is based on knowledge model Common DSW Knowledge Model (dsw:root:2.4.4). The questionnaire used for this DMP is identified by UUID \"601354fd-ac34-4f1a-b1ee-a45c7b9ac546\" within https://researchers.ds-wizard.org DSW instance.",
    "dmp_id": {
      "identifier": "https://researchers.ds-wizard.org/questionnaires/601354fd-ac34-4f1a-b1ee-a45c7b9ac546",
      "type": "url"
    },
    "ethical_issues_description": "Non-reference dataset \"Previous in-house Arsenic and Selenium analysis\" does not need an extension of consent because it does not contain personal data.",
    "ethical_issues_exist": "",
    "language": "eng",
    "modified": "2022-11-14T20:24:36Z",
    "project": [
      {
        "description": "The main goal of this study is to determine whether it is safe to use forage produced from the meadows in the vicinity of old metal mines in the area of P\u00159\u00ededbram, Kutn\u00e1 Hora, and Nal\u00edeovsk\u00e9 Hory (Czech Republic) to feed it to domestic herbivorous herds. Total and speciation analysis of As and Se in the plant biomass will be conducted using a hyphenated technique of HPLC and ICP-MS.",
        "end": "31.12.2021",
        "funding": [
          {
            "funder_id": {
              "identifier": "http://dx.doi.org/10.13039/501100001824",
              "type": "url"
            },
            "funding_status": "planned"
          }
        ],
        "start": "1.1.2021",
        "title": "Arsenic and Selenium Speciation Using Hyphenated Technique"
      }
    ],
    "title": "Science Europe Example DMP"
  }
}
```

- In 3.14 (Summer 2022) we introduced “project importer” feature that allows to flexibly import replies to a questionnaire in DSW
- We inversed the transformation used from export (document template)
- The only issue was with fields which we synthesize from multiple replies in the questionnaire (e.g. ethical issues text)
- Currently, importer supports only maDMPs in JSON
- Future work: improve coverage (together with export)

DSW maDMP Import (example)



Questionnaire

Metrics

Preview

Documents

Settings

View

Import answers

Warnings 2

Comments

TODOs

Version history

Current

DSW Replies (JSON)

RDA maDMP (JSON)

Before Submitting

Chapters

I. Administrative information 11

Contributors

Research Project(s)

To execute the DM...

Do you require har...

Describe national / ...

II. Re-using data 6

III. Creating a... 24

IV. Processing... 23

V. Interpreting... 9

VI. Preserving ... 11

VII. Giving acce... 2

1

Contributors

Horizon 2020 DMP

Horizon Europe DMP

Science Europe DMP

maDMP

Each person contributing to creating or executing the data management plan should be added as a contributor. A project probably should have a Contact Person, and a Data Curator.

☒ Desirable: Before Submitting the DMP

Jana Freeman

John Doe

Marek Suchánek

+ Add

2

Research Project(s)

Horizon 2020 DMP

Horizon Europe DMP

Science Europe DMP

maDMP

Add each of the research project(s) that you are (or will be) working on and for which the data and work are described in this DMP. Give each project a small identifying name for yourself.

Import answers

Example DMP

Follow the instructions in the importer window.

DSW maDMP importer - Personal - Microsoft Edge

https://s3.ds-wizard.org/project-importers/madmp/index.html

DSW maDMP importer

Select a machine-actionable DMP in JSON format according to the [RDA DMP Common Standard](#):

Choose File

No file chosen

Are we finished? What next?



- We plan to further improve coverage of the standard and support RDF format for import
- Do we want to work (just) with RDF/JSON files? No!
 - Direct DMP exchange with other tools (tools for DMPs but also others)
 - Development and integration of maDMP evaluators
 - Include RDF as part of RO-Crate
 - Enhance PID linking to maDMP
 - Add more DSW-specific data (in addition to standard “core”)

- We gained interoperability with other tools that support maDMPs (in terms of import and export of our questionnaire replies)
- Both document template and importer are open-source and can be adjusted for different knowledge models used in DSW:
 - <https://github.com/ds-wizard/dsw-madmp-importer>
 - <https://github.com/ds-wizard/madmp-template>
- There are plans for further improvements in terms of interoperability

Acknowledgements



The development and operation of DSW is supported by ELIXIR CZ research infrastructure (MŠMT Grant No.: LM2018131).

Contact information



- Data Stewardship Wizard

<https://ds-wizard.org>

 [@dswizard_org](https://twitter.com/dswizard_org)

info@ds-wizard.org



2022 Webinar Series

Highlighting RDA Outputs

Hosted by RDA-US



Thanks to....

Presenters and IG

Webinar Support - Meghan and Rebecca

French Interpreters - Digital Research Alliance in Canada

Spanish Interpreters - LA Referencia

www.rd-alliance.org/groups/rda-us



rd-alliance.org



@resdatall | @RDA_US





RDA PODCAST SERIES *Hosted by RDA-US*

Data Streams is a collection of conversations among members in the Research Data Alliance (RDA) community.

- | | | |
|------------|--|---|
| Episode #1 | The Data Citation WG | <i>Mark Parsons and Andreas Rauber</i> |
| Episode #2 | Data Repository Attributes WG | <i>Matthew Cannon and Michael Witt</i> |
| Episode #3 | Data Granularity WG | <i>Reyna Jenkyns, Brigitte Mathiak, Katy McNeill, Guangyuan Sun</i> |
| Episode #4 | Sensitive Data IG | <i>Nichola Burton, Aleks Michalewicz and Kristal Spreadborough</i> |
| Episode #5 | National PID Strategies WG | <i>Natasha Simons, Shawna Sadler and Melroy Almeida</i> |
| Episode #6 | FAIR Principles for Research Hardware IG | <i>Julien Colomb and Nadica Miljikovic</i> |

Subscribe: www.rd-alliance.org/podcast



rd-alliance.org



[@resdatall](#) | [@RDA_US](#)





Research Data Alliance 10th Anniversary Plenary meeting

Gothenburg, Sweden, March 21-23, 2023



CHALMERS



2022 Webinar Series

Highlighting RDA Outputs

Hosted by RDA-US



RDA Global

Web - www.rd-alliance.org

Email - enquiries@rd-alliance.org

Twitter - [@resdatall](https://twitter.com/resdatall)

LinkedIn - www.linkedin.com/in/ResearchDataAlliance

Slideshare - www.slideshare.net/ResearchDataAlliance

RDA US

Email – rdaus@rda-foundation.org

Twitter - [@RDA_US](https://twitter.com/RDA_US)

Web -

www.rd-alliance.org/groups/rda-us

Thank You For Joining Us!



rd-alliance.org



[@resdatall](https://twitter.com/resdatall) | [@RDA_US](https://twitter.com/RDA_US)

