RESEARCH DATA ALLIANCE

Data Management Planning: where are we and where do we want to be?

Joint session of Active Data Management Plans IG and DMP Common Standards WG and Discipline-specific Guidance for Data Management Plans WG

20th RDA Plenary

Agenda

- Part 1 – Getting to know each other and introduction for newcomers
- Part 2 – Updates and topics for discussion
  - Software Management Plans
  - Maintenance of maDMP specification
  - Santosh
  - Cross-fertilization workshop results
- Part 3 – Discussion in groups and reporting
- Part 4 – Wrap up

Participants

- Let’s get to know each other
  - Go to **slido.com**

#maDMPs

#activeDMPs

Active poll

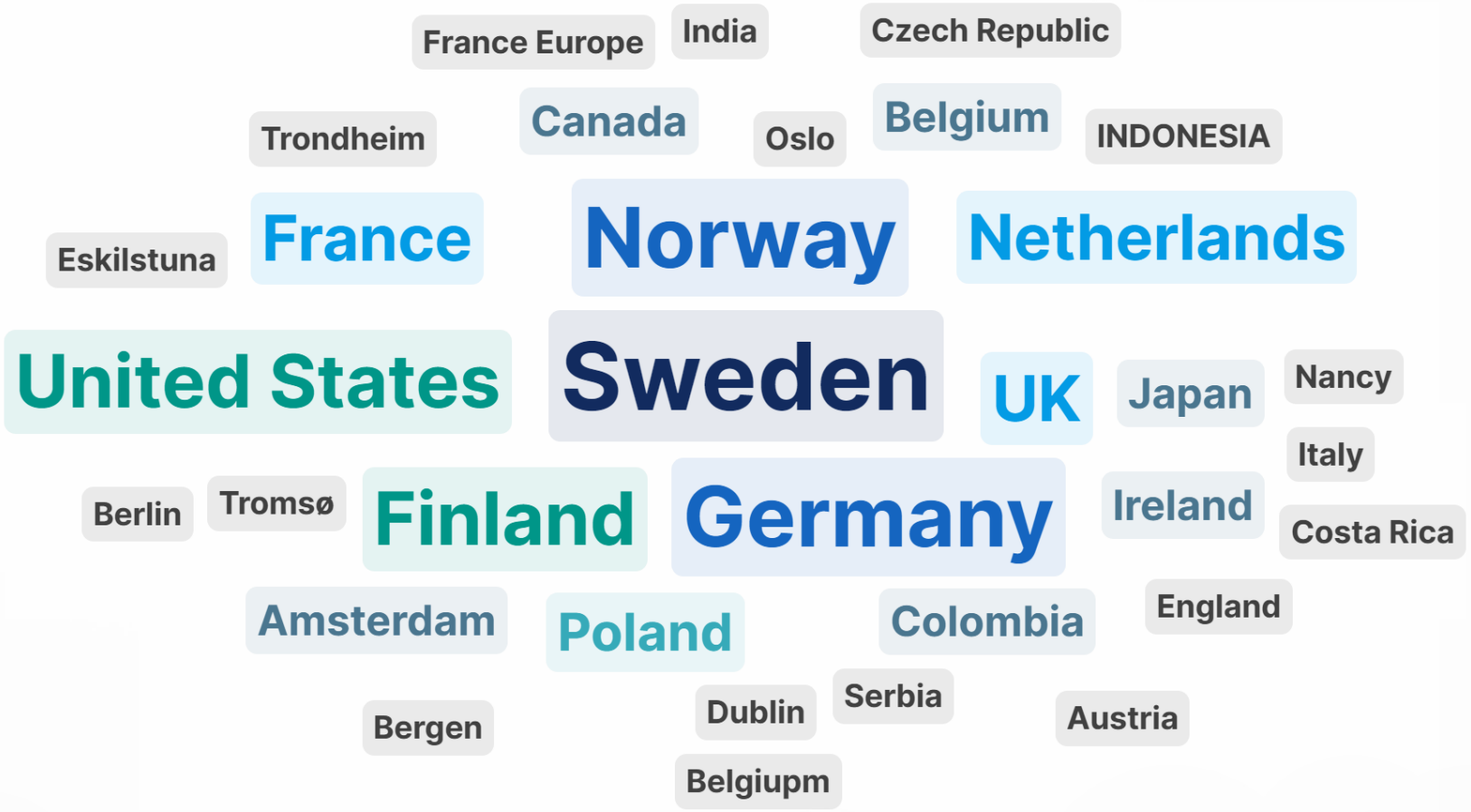
78

Join at

slido.com

#4225 137

Where are you from?



Active poll

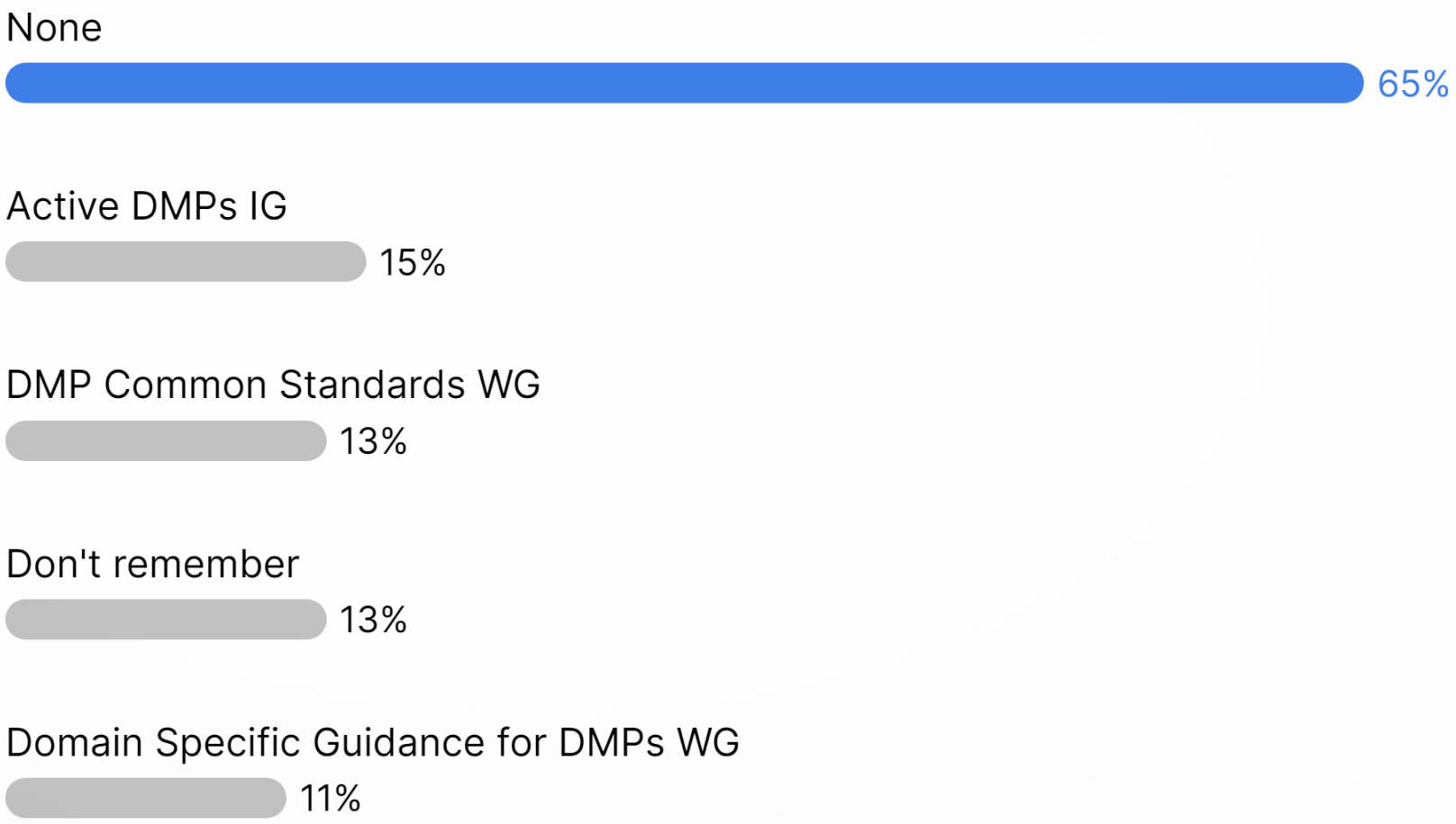
82

Join at

slido.com

#4225 137

Are you a member of the WG and/or the IG?



Active poll

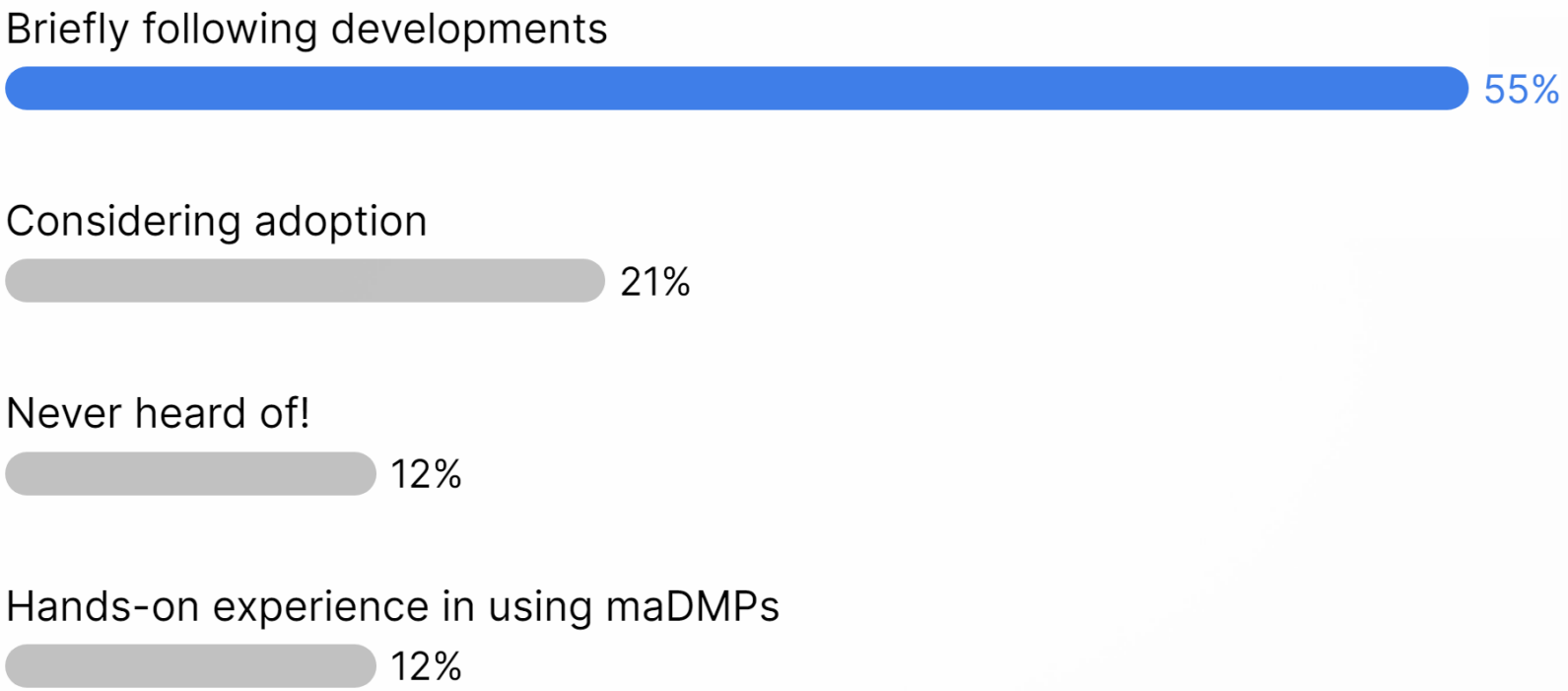
84

Join at

slido.com

#4225 137

How familiar are you with maDMPs?



Why joint session?

- Active DMPs *Interest Group*
  - Place for discussion on all topics related to DMPs
  - Can trigger new WGs
    - Like it happened in the past with the DMP Common Standards WG

400+ members!

- DMP Common Standards *Working Group*
  - Maintenance mode
    - Updates the recommendation when necessary
    - Supports adopters of the recommendation

240+ members!

- Domain-Specific Guidance for DMPs *Working Group*
  - Ends in May
    - Paper + Jupyter Notebook

110+ members!

Objective for today

Identify new hot topics relevant to this community

Introduction for newcomers on maDMPS

Part 1

Data Management Plans (DMPs)

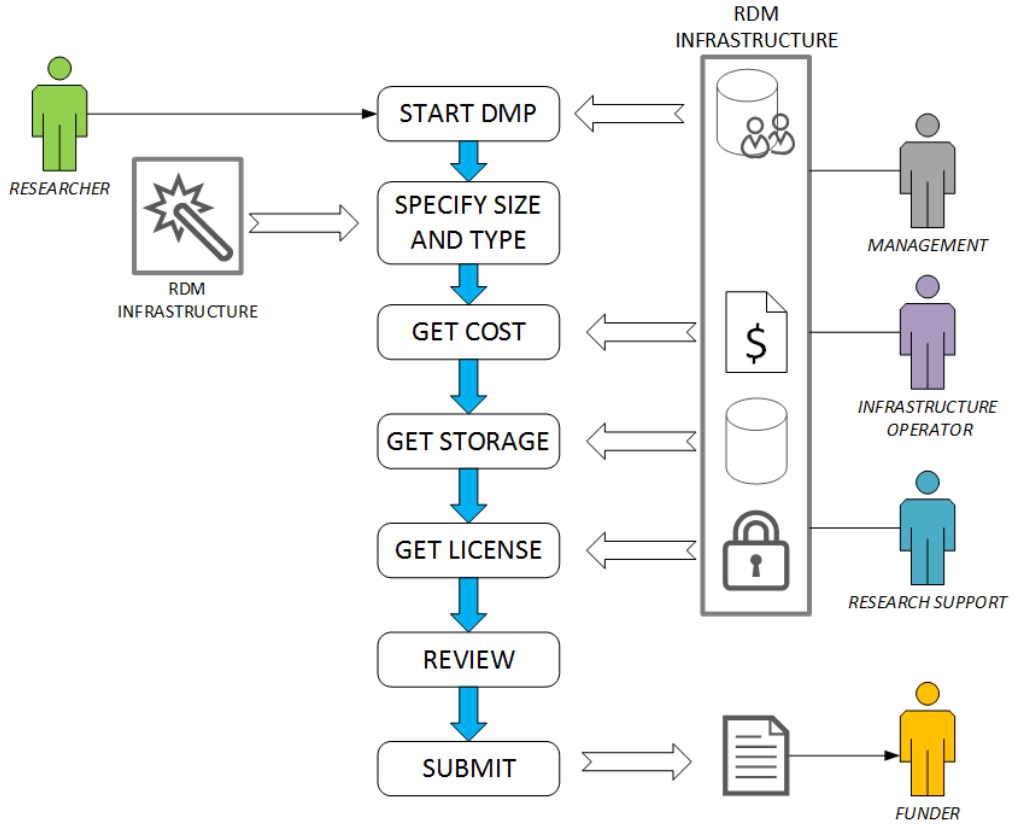
	Data Officer	*Who is responsible for the data management and the DMP of the project (name/email address)?*
I	**Data Characteristics**	
I.1	**Description of the data**	*What kinds of data/source code will be generated or reused (type, format, volume)?* *How will the research data be generated and which methods will be used?* *How will you structure the data and handle versioning?* *Who is the target audience?*
II	**Documentation and Metadata**	
II.1	**Metadata standards**	*What metadata standards (if any) will be in use and why? (see Digital Curation Centre)*
II.2	**Documentation of data**	*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?*
II.3	**Data quality control**	*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.)*
III	**Data Availability and Storage**	
III.1	**Data sharing strategy**	*How and when will the data be shared and made accessible?* *What repository will you be using?* *What persistent identifier will be used?*
III.2	**Data storage strategy**	*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?*

Table of Contents

4 For procedural elements of implementing DMPs, see the RDA DMP Common Standards Working Group: https://www.rd-alliance.org/groups/dmp-common-standards-wg

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

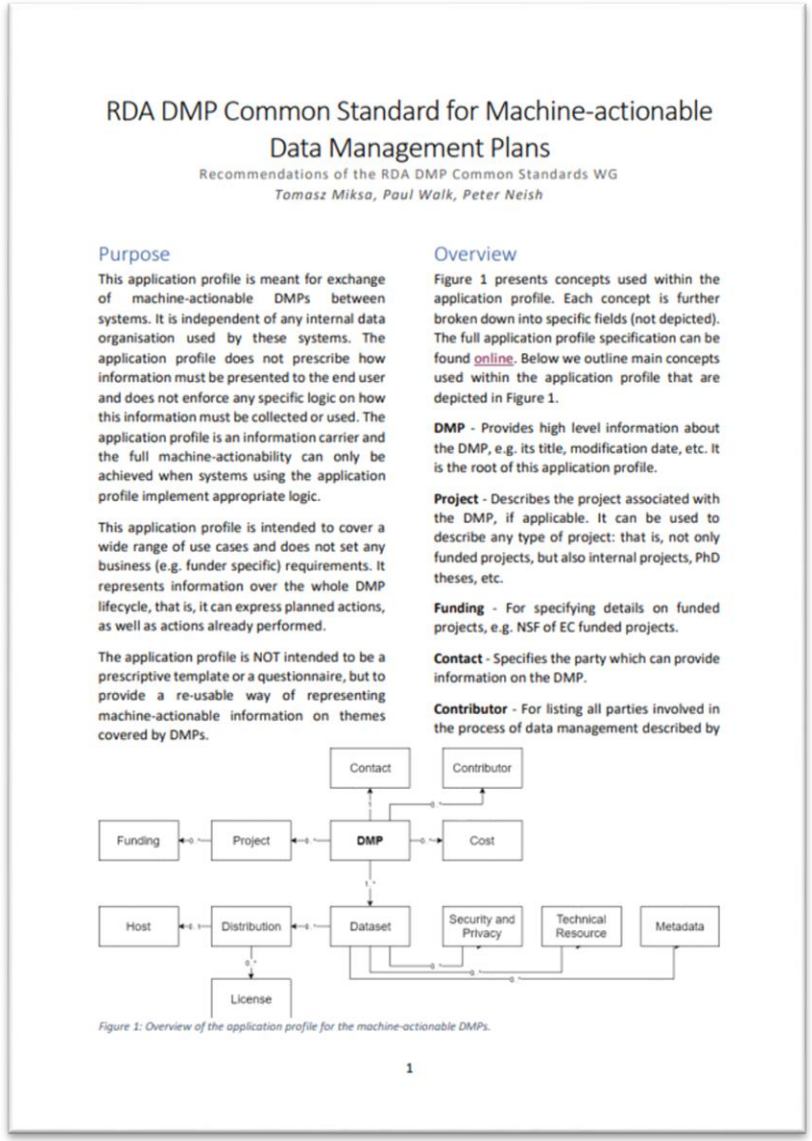
RDA DMP Common Standard for Machine-actionable Data Management Plans

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

Miksa, T., Walk, P., & Neish, P. (2019). RDA DMP Common Standard for Machine-actionable Data Management Plans. https://doi.org/10.15497/rda00039

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. resoponsible 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!

https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/index.md

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.

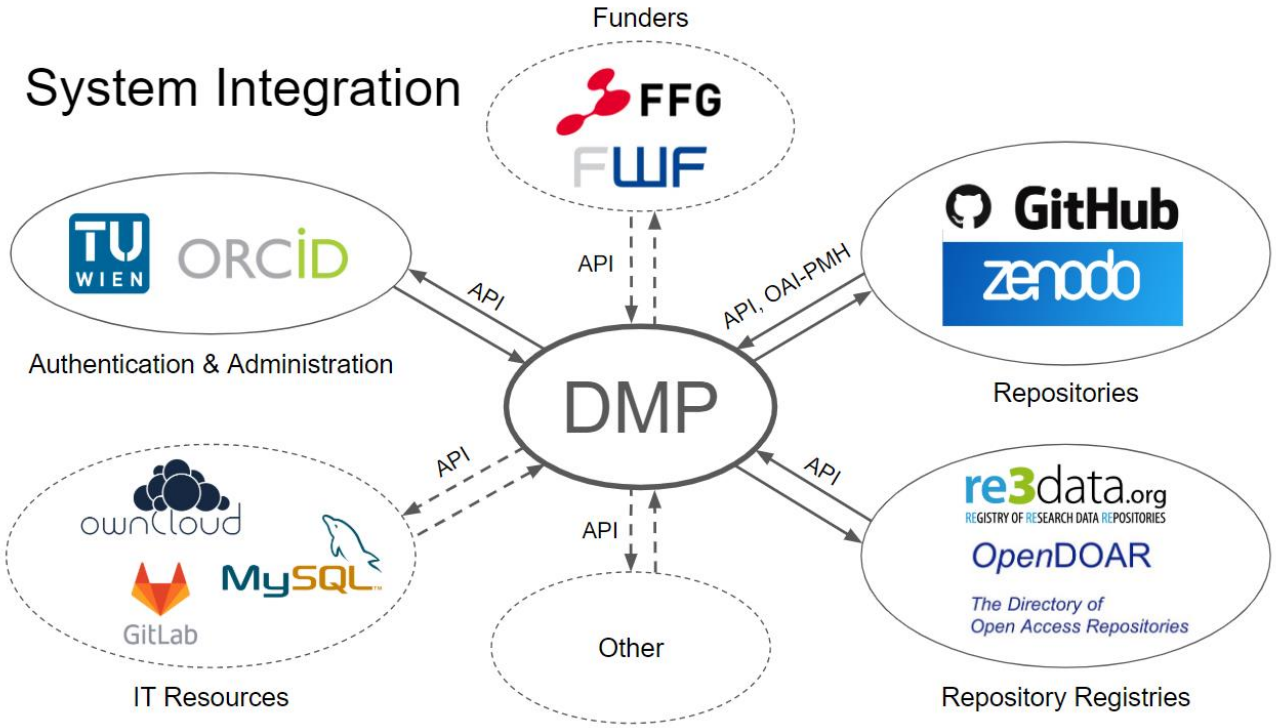
Machine-actionable DMP


```
"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.

RDM Infrastructure

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

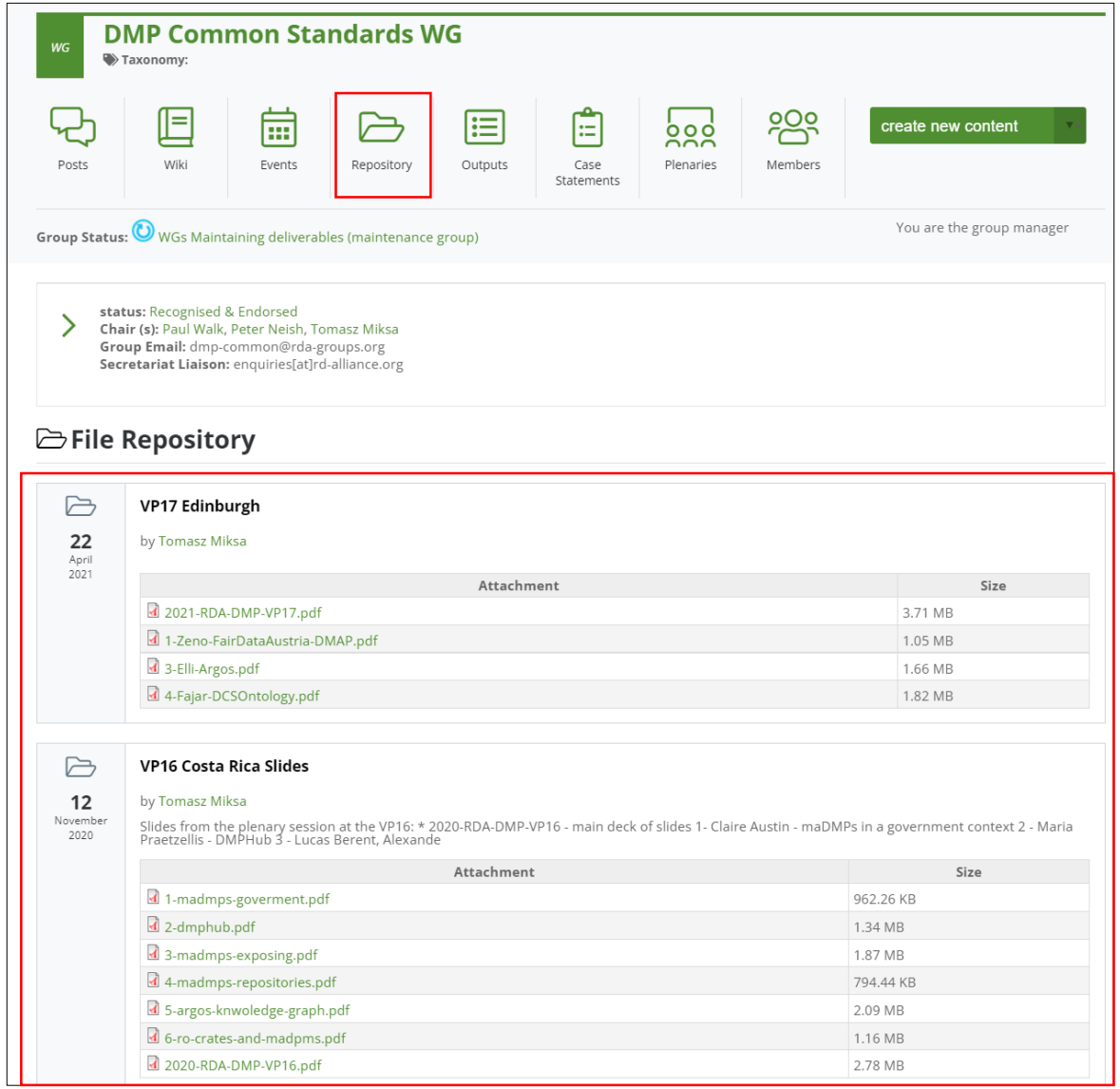


Adoptions (selected)

DMP Common Standards WG

- Slides from all our sessions are in the repository
- Today's presentations will also be there

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



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:

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 Views | 49 Downloads | 28 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

Read more in…

1

Automating Research Data Management Using Machine-actionable Data Management Plans

TOMASZ MIKSA, TU Wien & SBA Research, Austria
SIMON OBLASSER, TU Wien, Austria
ANDREAS RAUBER, TU Wien, Austria

Many research funders mandate researchers to create and maintain Data Management Plans (DMPs) for research projects that describe how research data is managed to ensure its reusability. A DMP being a static textual document is difficult to act upon and can quickly become obsolete and impractical to maintain. A new generation of machine-actionable DMPs was therefore proposed by the Research Data Alliance to enable automated integration of information and updates. Machine-actionable DMPs open up a variety of use cases enabling interoperability of research systems and automation of data management tasks.

In this paper we describe a system for machine-actionable data management planning in an institutional context. We identify common use cases within research that can be automated to benefit from machine-actionability of DMPs. We propose a reference architecture of a machine-actionable DMP support system that can be embedded into an institutional research data management infrastructure. The system semi-automates creation and maintenance of DMPs, and thus eases the burden for the stakeholders responsible for various DMP elements. We evaluate the proposed system in a case study conducted at the largest technical university in Austria and quantify to what extent the DMP templates provided by the European Commission and a national funding body can be pre-filled. The proof-of-concept implementation shows that machine-actionable DMP workflows can be semi-automated, thus workload on involved parties can be reduced and quality of information increased. The results are especially relevant to decision makers and infrastructure operators who want to design information systems in a systematic way that can utilise the full potential of machine-actionable DMPs.

CCS Concepts: • **Applied computing** → **Enterprise data management**; **Business process management**; **IT architectures**; • **Information systems** → **Digital libraries and archives**; • **Social and professional topics** → **Automation**.

Additional Key Words and Phrases: data management plan, machine-actionable, business processes, enterprise architecture, funder template, requirements engineering, automation, RDM, RDA, FAIR

ACM Reference Format:
Tomasz Miksa, Simon Oblasser, and Andreas Rauber. 2021. Automating Research Data Management Using Machine-actionable Data Management Plans . *ACM Trans. Manag. Inform. Syst.* 1, 1, Article 1 (January 2021), 22 pages. https://doi.org/10.1145/3490396

1 INTRODUCTION

The data revolution continues to transform every sector of science, industry, and government [AS19]. The economic and societal benefits and increased effectiveness of research funding by ensuring that data generated and (pre-) processed as part of research remains available for re-use,

Authors' addresses: Tomasz Miksa, tmiksa@sba-research.org, TU Wien & SBA Research, Vienna, Austria; Simon Oblasser, simon.oblasser@student.tuwien.ac.at, TU Wien, Vienna, Austria; Andreas Rauber, rauber@ifs.tuwien.ac.at, TU Wien, Vienna, Austria.

2158-656X/2021/1-ART1
https://doi.org/10.1145/3490396

Contains:
- 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

Practice Paper

Interconnecting systems using machine-actionable Data Management Plans - hackathon report

João Cardoso[1], **Leyla J. García**[2], **Tomasz Miksa**[3]
[1]Universidade de Lisboa, Instituto Superior Técnico & INESC-ID, Lisboa, Portugal
[2]ZB MED Information Centre for Life Sciences, Cologne, Germany
[3]SBA Research & TU Wien, Vienna, Austria

Corresponding author, joao.m.f.cardoso@tecnico.ulisboa.pt

This paper presents outputs of the Research Data Alliance Hackathon on Machine-actionable Data Management Plans, where participants proposed a series of topics looking forward to using and improving aspects related to this subject. The hackathon served three main purposes: broadening the community, improving the core supporting machine-actionable plans and exposing a growing endorsement on the adoption of the RDA DMP Common Standard application profile in a wide range of settings to enable exchange of DMP specific information in a machine-actionable way.

Keywords: Data management plans, machine-actionable data management plans, semantic web, community practice, open science.

1 Introduction

The Data Management Plan (DMP) was introduced to document and publish both data management practices and policies that are applied to data throughout its lifecycle. This implies describing the techniques, methods and policies on how data is to be created, collected, documented, processed, accessed, preserved, disseminated as well as the roles and responsibilities of associated actors (Michener, 2015).

The premise behind the concept of a machine-actionable DMP (maDMP) is that information contained within a DMP can be enacted both by humans and automated systems, thus addressing some of the limitations associated with traditional DMP documents. To that effect, data management workflows should integrate maDMPs and data management policies should take into account not only human agents but also machines. maDMPs should support both human and machine-processable representations so they act as an interchange format for dissemination and public access of the maDMP (Simms et al., 2017). In order to provide a machine-actionable representation of a maDMP, it becomes necessary to establish a standardised representation of the maDMP. The Research Data Alliance (RDA) (RDA, 2020) DMP Common Standards (DCS) working group (Miksa, Cardoso, and Borbinha, 2018; Miksa, Neish, et al., 2018; Miksa, Walk, and Neish, 2019) developed an application profile making it easier to express information from traditional DMP documents in a machine-actionable way. The DCS maDMP application profile allows for automatic exchange, integration, and validation of information provided in DMP documents. Thus, facilitating the exchange of information between systems acting on behalf of stakeholders involved in the research life cycle, such as researchers, funding bodies, repository managers, ICT providers, librarians, etc.

This paper reports on a hackathon organised by the DCS working group, which had as main motivation to promote the adoption of the maDMP concept by the research community, and, in particular, the usage of the DCS application profile for interchange of maDMPs. To that effect four main areas were identified: (1) serialisation, to encourage community development of serialisations of the DCS application profile; (2)

Contains:
- Summary of results from hackathon

DATA SCIENCE JOURNAL

http://doi.org/10.5334/dsj-2021-035

Introduction for newcomers on domain-specific guidance for DMPs

Part 1

Discipline-specific Aspects for DMP WG

Daniela Hausen Ivonne Anders, Santosh Ilamparuthi,
Yasemin Türkyilmaz-van der Velden, Shannon Sheridon, Briana Wham

Motivation

- Researchers should
  - Think about RDM in advance
  - Create a plan
  - Identify gaps

- Researchers want to
  - Have clear and unique concepts
  - Have direct support by structure, helping texts and examples
  - Support from the discpline

Create Guidance and Examples
adopt Terminology

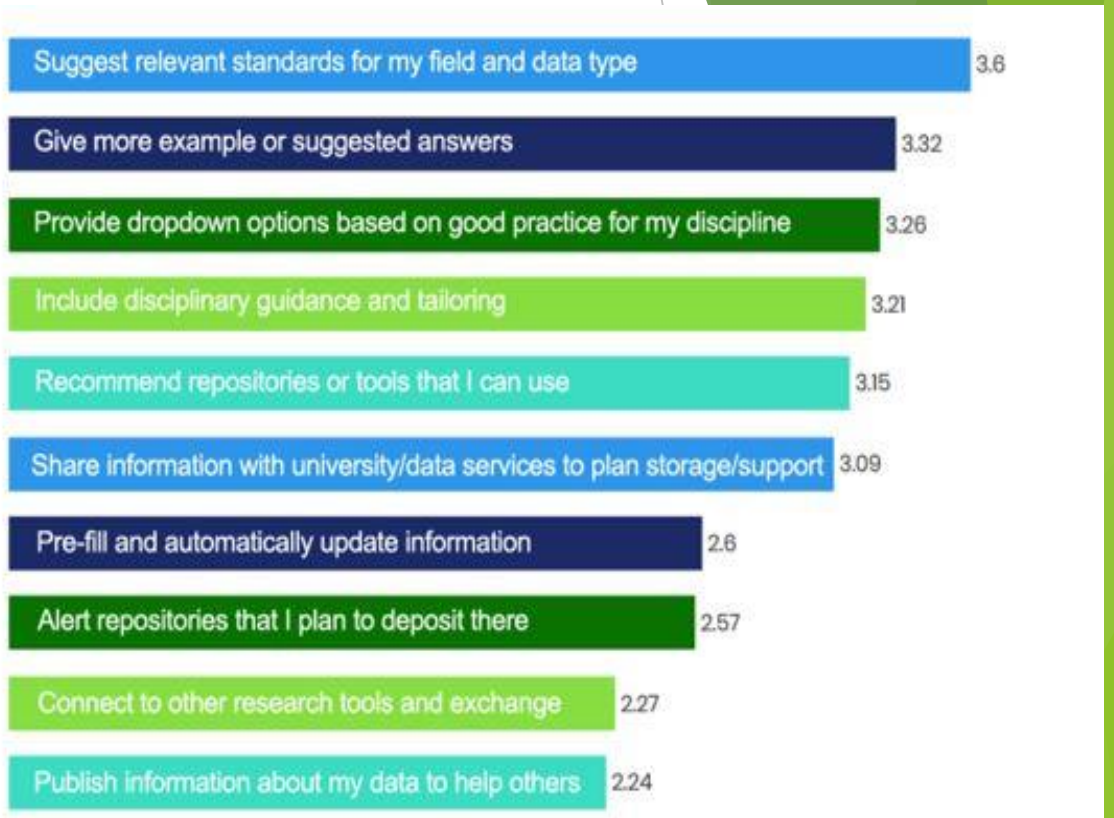


Figure 001 - Q11: priorities for a DMP template or tool

source: http://doi.org/10.5281/zenodo.1120245

RDA Online Survey on discipline-specific aspects for DMP templates


Daniela Hausen
RWTH Aachen


Ivonne Anders
DKRZ


Santosh Ilamparuthi
TU Delft


Yasemin Türkyilmaz-van der Velden
TU Delft

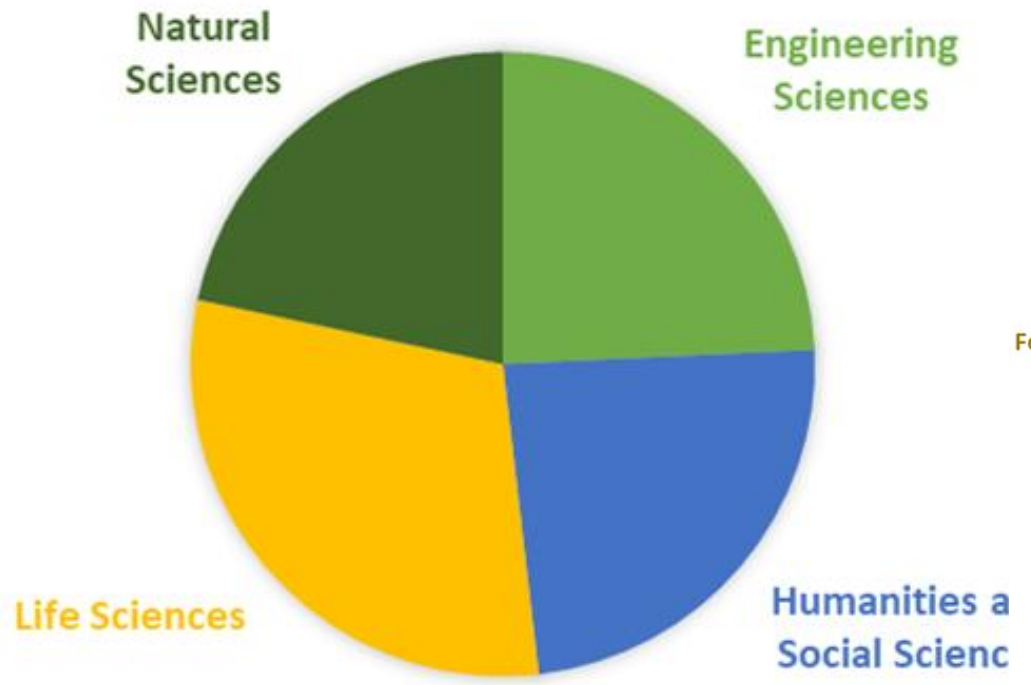

Shannon Sheridan
Pacific Northwest National Laboratory


Briana Wham
Penn State University

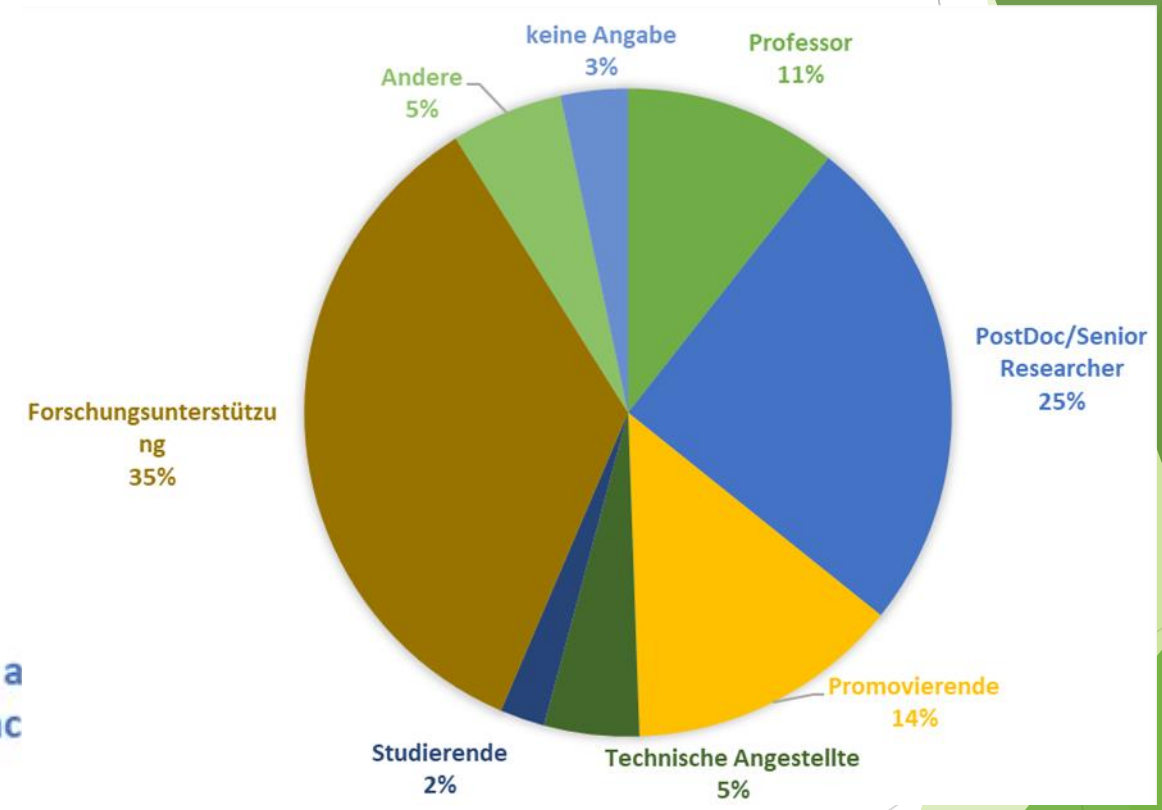
- survey 11. Okt. 2021 bis 14. Jan. 2022
- 358 participants
- 21 questions:
  - (1) Demographics,
  - (2) Data Description,
  - (3) Data Documentation & Quality,
  - (4) Data Archiving, Publishing & Sharing After the Project,
  - (5) Guidelines, Principles, & Best Practices

Results from the Online Survey

Disciplines of the participants



Position of the participants



Results from the Online Survey

Issues/areas with discipline-specific differences:

- Metadata and metadata standards
- Data types
- Data generation and data collection
- Post-use of data types
- Data documentation
- Quality control
- Framework conditions (legal, contractual and ethical)
- Types of data publication

Issues/areas with large overlaps:

- Discipline-specific guidelines
- Implementation of FAIR principles
- Data storage incl. storage media
- Use of naming conventions

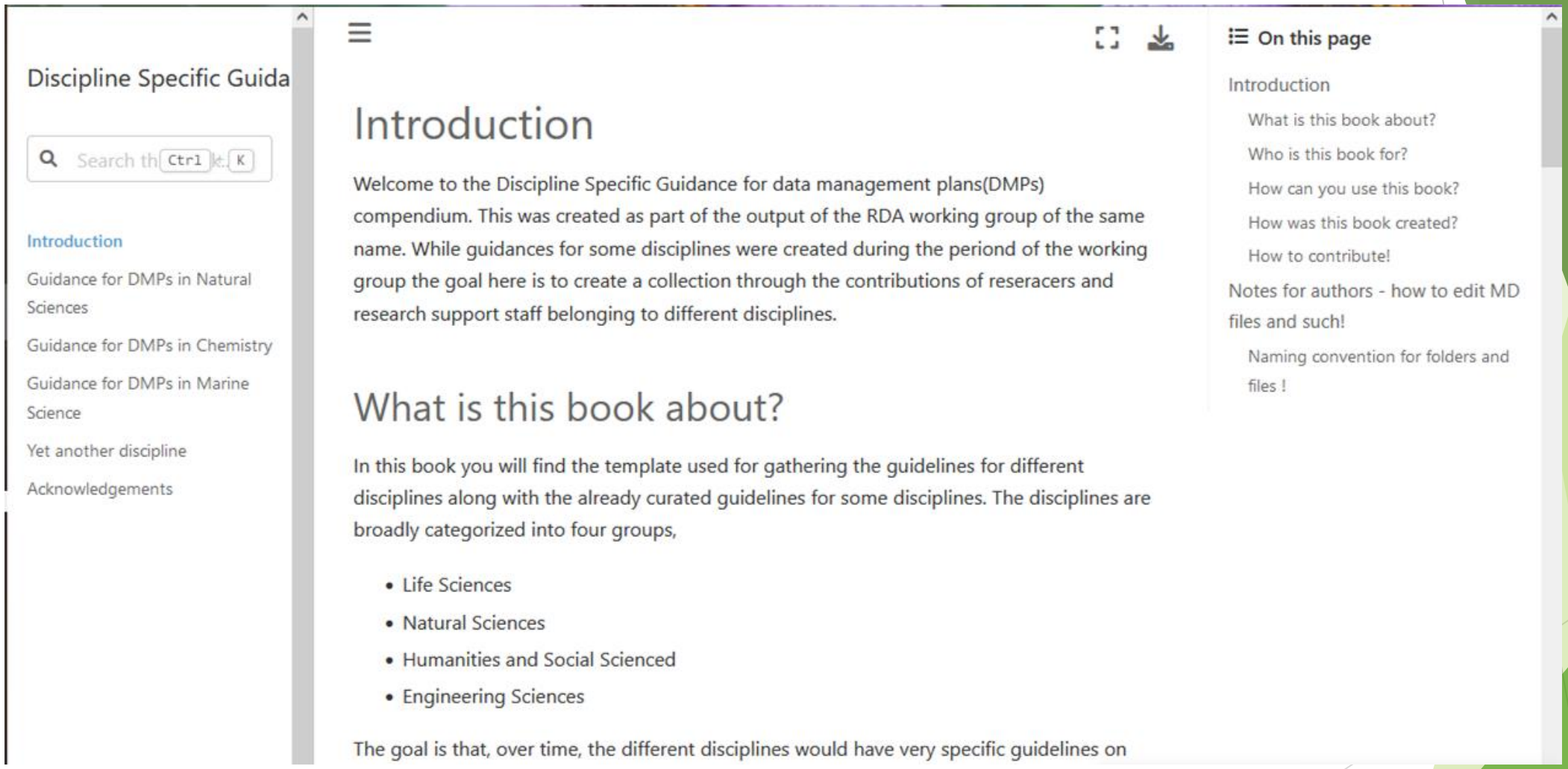
Issues/areas without discipline-specific focus:

- Allocation of persistent identifiers
- Allocation of licences
- Use of learning materials

Hausen, D.A., Wham, B., Anders, I., Ilamparuthi, S. and Sheridan, S., 2023. Discipline-specific Aspects in Data Management Planning. *Data Science Journal*, submitted

Wham, Briana, Hausen, D., Andres, I., Sheridan, S., Ilamparuthi, S., & Turkyilmaz-van der Velden, Y., 2022. Data Set from RDA WG Discipline-Specific Guidance on DMP - Online Survey. Zenodo. https://doi.org/10.5281/zenodo.7391669

Prototype



https://santoshilam.github.io/Discipline_Specific_Guidance_for_DMPs/intro.html
https://github.com/santoshilam/Discipline_Specific_Guidance_for_DMPs

Updates and topics for discussion

Part 2

Part 2

- “Jupyter Notebook” live demo - Santosh
- Software Management Plans - Kerstin (virtual)
- Maintenance of the maDMP specification - Marek
- Cross-fertilization workshop results - Elli

FAIRification of software development – (machine-actionable) software management plans

Discussion at RDA Plenary 20

suggested by Kerstin Helbig (Humboldt-Universität zu Berlin, Germany)

What is a software management plan?

Definition by DINI/nestor Working Group Research Data:

A software management plan (SMP) includes general and technical information about the software project, information about quality assurance, release and public availability, as well as legal and ethical aspects affecting the software.

The SMP summarizes information that sufficiently describes and documents the creation, documentation, storage, versioning, licensing, archiving and/or publication of the software generated or used in a project. Related hardware and necessary other resources, as well as related other software and software libraries, text and data publications, must also be described and are a feature of the SMP.

The purpose of an SMP is first of all to support the traceability and, if necessary, the long-term usability of the software (for direct application as well as for further processing) and to facilitate the support of the users in case of queries. The SMP therefore also serves the purpose of quality assurance (cf. FAIR4RS Principles).

The SMP can be linked to one or more data management plans (DMP) if the software is used for data generation or processing. SMP and DMP can be combined as output plans (cf. Software Sustainability Institute).

Source: https://forschungsdaten.info/praxis-kompakt/english-pages/glossary/#c499503

Funder requirements and integration of SMP in DMP tools

- Most funders treat research software as data -> no specialized requirements
- SMP is not a standard tool for software management planning
- With few exceptions, DMP tools do not offer up-to-date, comprehensive templates for software
- Machine actionable integration of SMP (maSMP) into GitHub could support FAIR4RS

Points for discussion

- Do we need SMPs or should we work towards more generalized output plans that combine data and software management?
- How can FAIRification of software be supported via SMPs?
- To what extent can maSMP help with this?
- …

Maintenance of the maDMP specification

Marek Suchánek

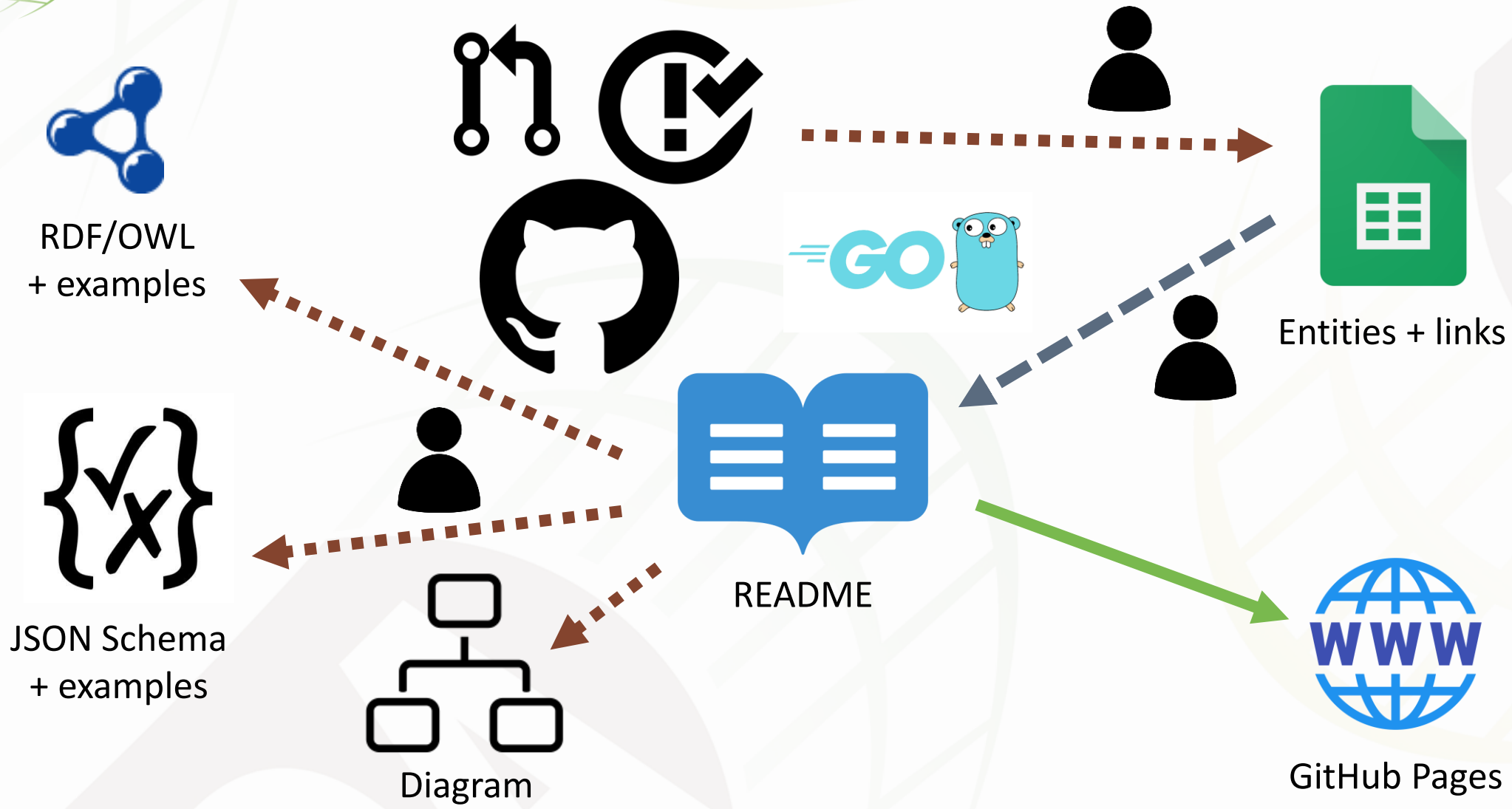
0000-0001-7525-9218

marek.suchanek@fit.cvut.cz

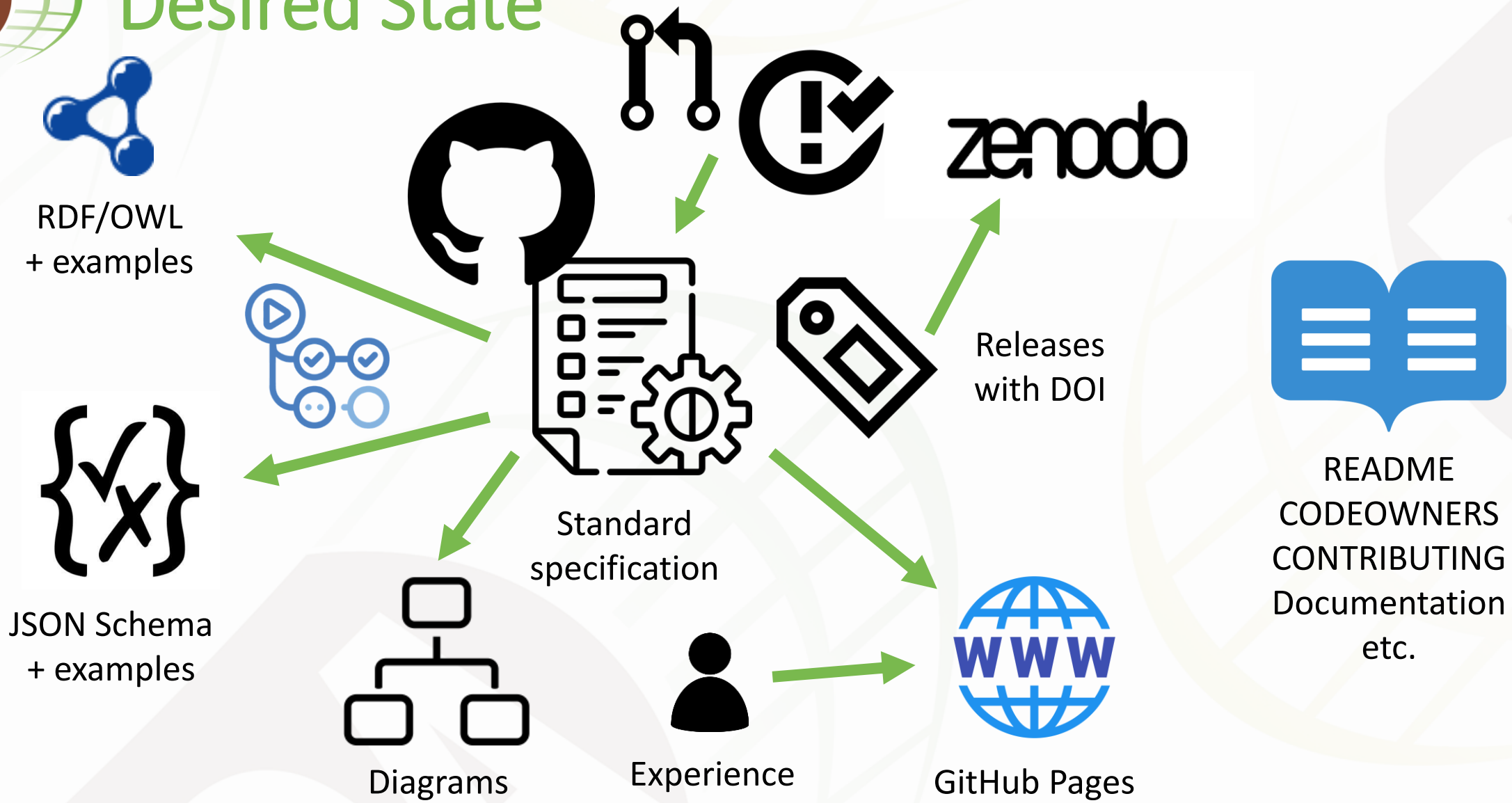
Current State and Issues

- GitHub repository, Google Spreadsheet, GitHub Pages (README-like)
- Not a single source of truth (figures, JSON schema, spreadsheet, generated README)
  - Harder to contribute (and adopt)
  - Possible inconsistencies
  - Issues with versioning
- No community content (adoption stories/experience)
- No automation (except the one with Google Spreadsheet)

Current State and Issues



Desired State



Work Plan

- EOSC Future support through *Optimising Frameworks and Guidelines in EOSC #2* grant
- Tasks and partial goals in the project:
  - Conduct user/adopter research, gather feedback, document issues
  - Develop tools and other resources supporting use and adoption as well as maintanence of the standard
  - Foster contributions and community development by providing proper way of crediting contributors and contributing guidelines
  - Enhance standard documentation (website, implementation recipes, sharing experience, clarify versions of standards, extensions, etc.)

Part 2

- Software Management Plans - Kerstin (virtual)
- Maintenance of the maDMP specification - Marek
- “Jupyter Notebook” - Santosh
- Cross-fertilization workshop results - Elli

RDA for Data Management Planning Community Cross-fertilisation Workshop Summary

RESEARCH DATA ALLIANCE

RDA for Data Management Planning

Celebrating A Decade of Data
RDA community cross-fertilisation workshop

doi. https://doi.org/10.15497/RDA00080

Version: December 2022

ABOUT THE WORKSHOP

The community cross-fertilisation workshop, 'RDA for Data Management Planning', brought chairs and members of RDA Working Groups (WGs) and Interest Groups (IGs) together, with members of the wider research data community, to share and discuss challenges, solutions and initiatives associated with data management plans (DMPs). The key findings of the workshop summarised herein will be used to direct the future strategy of the RDA community. Read more about the community cross-fertilisation workshop series in commemoration of the RDA's 10th Anniversary.

CHALLENGES TO BE ADDRESSED WITHIN THE THEME OF DATA MANAGEMENT PLANNING

Raising awareness & understanding about DMPs:

- Lack of awareness that DMPs are integral to good research data management (RDM).
- Creating a DMP is often considered an extra task.
- Creating a DMP requires prerequisite knowledge and skills related to RDM, tools and services.
- Insufficient incentives, reward and credit for RDM and the creation of DMPs.

Creation, adoption & implementation of DMPs:

- Lack of clarity about roles and responsibilities for supporting DMP creation and limited staff.
- Lack of automated and machine-actionable processes and workflows to harmonise DMP creation and implementation across stakeholders, tools, services and policies.
- Unclear definitions and language (vocabularies and terminologies) make DMPs challenging to understand and complete.
- Diversity of data means generic DMPs are unsuitable for specific research disciplines.
- As primarily text documents and parts of research proposals, DMPs are not FAIR research objects that enable their utility.

Review, evaluation & assessment of DMPs:

- No clear process or assessment criteria for DMP review and evaluation.
- Lack of accountability on following through on the implementation of DMPs that support funded research.

SOLUTIONS TO ADDRESS THE CHALLENGES

- Strengthen communication and collaboration between RDA groups and members (e.g., via submission of joint RDA Plenary sessions).
- Create new WGs and IGs, and collaborate with relevant organisations and communities.
- Expand RDA community work on DMPs to focus on funders as a key stakeholder.
- Design automated, machine-actionable and interoperable processes and workflows that enable different stakeholders to collaborate more easily on co-creating, adopting and implementing DMPs.
- Ensure research performing organisations offer tools and support services that aid creation of FAIR DMPs that can be reused.
- Encourage researchers to use tools, services and workflows and advise tool providers about what tools/infrastructure is needed.
- Employ Artificial Intelligence to create and review DMPs.
- Ensure that DMPs are considered as part of the research assessment reform.

PARTICIPATING GROUPS & WORKSHOP LEADS*

DMP Common Standards WG
Workshop lead: Tomasz Miksa
Outputs: RDA DMP Common Standard for Machine-actionable Data Management Plans & Publications

See community group card

Active Data Management Plans IG
Workshop lead: Elli Papadopoulou
Output: RDA DMP Common Standard for Machine-actionable Data Management Plans

See community group card

Discipline-specific guidance for Data Management Plans WG
Workshop lead: Santosh Ilamparuthi
Output: Online survey about the current state of discipline-specific DMPs and data/code management practices.

See community group card

*Workshop leads collected challenges, solutions and initiatives in preparation for the workshop and explained them during the workshop on behalf of their group.

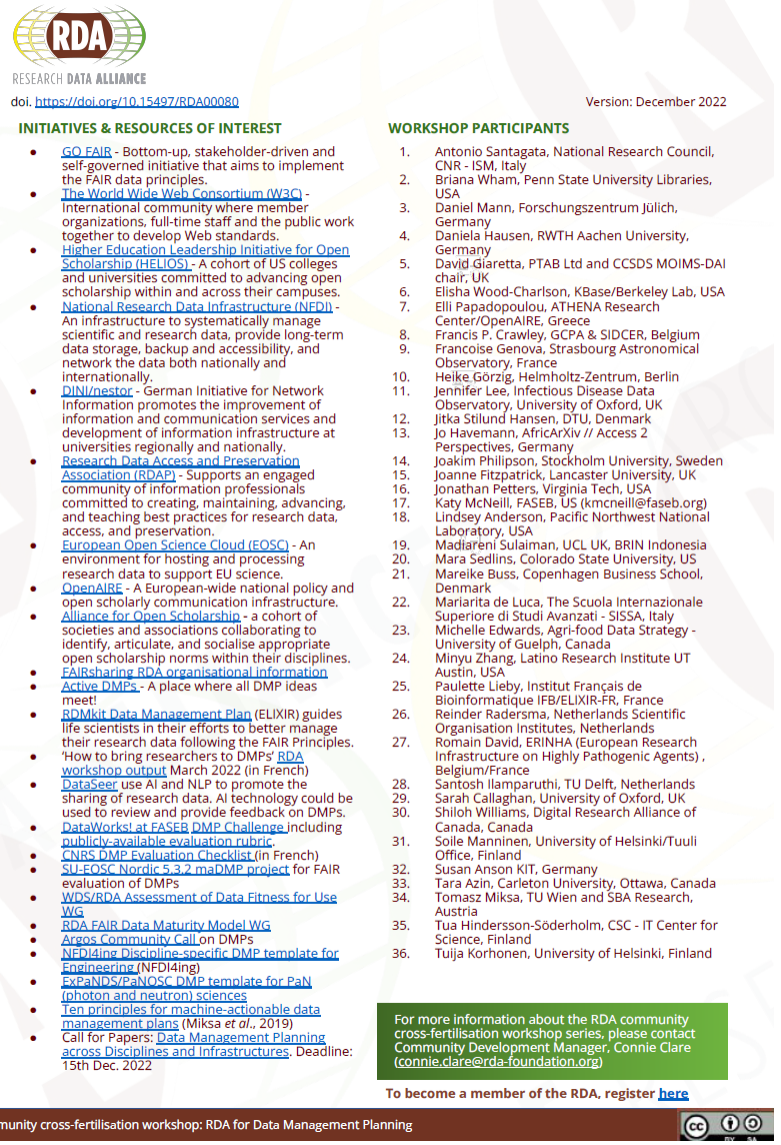
ACTIONS FOR THE RDA COMMUNITY

Collect 'gold star' DMP case studies. Real-world examples of DMPs from different institutions, disciplines and projects to demonstrate benefits of DMPs for various stakeholders. Case studies and infographics showcase best practices for how to create DMPs and integrate them into the research data lifecycle. This output supplements the 'Engaging Researchers with Data Management The Cookbook' by the RDA Engaging Researchers with Data IG.

Construction of DMP typologies. Define information, language and terminology to be included in DMPs that unambiguously describe RDM concepts and tasks. Typologies leverage existing ontologies and controlled vocabularies (E.g., T4FS, FIP ontology).

Create framework(s) for DMP evaluation and assessment. Collaborate with funding agencies (RDA Funders Forum and Research Funders and Stakeholders on Open Research and Data Management Policies and Practices IG) to create a scalable DMP evaluation framework that defines content to be assessed and evaluation criteria for DMPs dependent on institutional, disciplinary and project contexts.

Develop a toolkit for how to engage researchers in DMPs. DMP professionals, domain experts and RDM supporters co-create a step-by-step guide for how best to drive adoption and implementation of DMPs.



RESEARCH DATA ALLIANCE

doi. https://doi.org/10.15497/RDA00080

Version: December 2022

INITIATIVES & RESOURCES OF INTEREST

- GO FAIR - Bottom-up, stakeholder-driven and self-governed initiative that aims to implement the FAIR data principles.
- The World Wide Web Consortium (W3C) - International community where member organizations, full-time staff and the public work together to develop Web standards.
- Higher Education Leadership Initiative for Open Scholarship (HELIOS) - A cohort of US colleges and universities committed to advancing open scholarship within and across their campuses.
- National Research Data Infrastructure (NFDI) - An infrastructure to systematically manage scientific and research data, provide long-term data storage, backup and accessibility, and network the data both nationally and internationally.
- DINI/nestor - German Initiative for Network Information promotes the improvement of information and communication services and development of information infrastructure at universities regionally and nationally.
- Research Data Access and Preservation Association (RDAP) - Supports an engaged community of information professionals committed to creating, maintaining, advancing, and teaching best practices for research data, access, and preservation.
- European Open Science Cloud (EOSC) - An environment for hosting and processing research data to support EU science.
- OpenAIRE - A European-wide national policy and open scholarly communication infrastructure.
- Alliance for Open Scholarship - a cohort of societies and associations collaborating to identify, articulate, and socialise appropriate open scholarship norms within their disciplines.
- FAIRsharing RDA organisational information
- Active DMPs - A place where all DMP ideas meet!
- RDMkit Data Management Plan (ELIXIR) guides life scientists in their efforts to better manage their research data following the FAIR Principles.
- 'How to bring researchers to DMPs' RDA workshop output March 2022 (in French)
- DataSeer use AI and NLP to promote the sharing of research data. AI technology could be used to review and provide feedback on DMPs.
- DataWorks! at FASEB DMP Challenge including publicly-available evaluation rubric.
- CNRS DMP Evaluation Checklist (in French)
- SU-EOSC Nordic 5.3.2 maDMP project for FAIR evaluation of DMPs
- WDS/RDA Assessment of Data Fitness for Use WG
- RDA FAIR Data Maturity Model WG
- Argos Community Call on DMPs
- NFDI4ing Discipline-specific DMP template for Engineering (NFDI4ing)
- ExPaNDS/PaNOSC DMP template for PaN (photon and neutron) sciences
- Ten principles for machine-actionable data management plans (Miksa *et al*., 2019)
- Call for Papers: Data Management Planning across Disciplines and Infrastructures. Deadline: 15th Dec. 2022

WORKSHOP PARTICIPANTS

1. Antonio Santagata, National Research Council, CNR - ISM, Italy
2. Briana Wham, Penn State University Libraries, USA
3. Daniel Mann, Forschungszentrum Jülich, Germany
4. Daniela Hausen, RWTH Aachen University, Germany
5. David Giaretta, PTAB Ltd and CCSDS MOIMS-DAI chair, UK
6. Elisha Wood-Charlson, KBase/Berkeley Lab, USA
7. Elli Papadopoulou, ATHENA Research Center/OpenAIRE, Greece
8. Francis P. Crawley, GCPA & SIDCER, Belgium
9. Francoise Genova, Strasbourg Astronomical Observatory, France
10. Heike Görzig, Helmholtz-Zentrum, Berlin
11. Jennifer Lee, Infectious Disease Data Observatory, University of Oxford, UK
12. Jitka Stilund Hansen, DTU, Denmark
13. Jo Havemann, AfricArXiv // Access 2 Perspectives, Germany
14. Joakim Philipson, Stockholm University, Sweden
15. Joanne Fitzpatrick, Lancaster University, UK
16. Jonathan Petters, Virginia Tech, USA
17. Katy McNeill, FASEB, US (kmcneill@faseb.org)
18. Lindsey Anderson, Pacific Northwest National Laboratory, USA
19. Madiareni Sulaiman, UCL UK, BRIN Indonesia
20. Mara Sedlins, Colorado State University, US
21. Mareike Buss, Copenhagen Business School, Denmark
22. Mariarita de Luca, The Scuola Internazionale Superiore di Studi Avanzati - SISSA, Italy
23. Michelle Edwards, Agri-food Data Strategy - University of Guelph, Canada
24. Minyu Zhang, Latino Research Institute UT Austin, USA
25. Paulette Lieby, Institut Français de Bioinformatique IFB/ELIXIR-FR, France
26. Reinder Radersma, Netherlands Scientific Organisation Institutes, Netherlands
27. Romain David, ERINHA (European Research Infrastructure on Highly Pathogenic Agents) , Belgium/France
28. Santosh Ilamparuthi, TU Delft, Netherlands
29. Sarah Callaghan, University of Oxford, UK
30. Shiloh Williams, Digital Research Alliance of Canada, Canada
31. Soile Manninen, University of Helsinki/Tuuli Office, Finland
32. Susan Anson KIT, Germany
33. Tara Azin, Carleton University, Ottawa, Canada
34. Tomasz Miksa, TU Wien and SBA Research, Austria
35. Tua Hindersson-Söderholm, CSC - IT Center for Science, Finland
36. Tuija Korhonen, University of Helsinki, Finland

For more information about the RDA community cross-fertilisation workshop series, please contact Community Development Manager, Connie Clare (connie.clare@rda-foundation.org)

To become a member of the RDA, register here

https://doi.org/10.15497/RDA00080

CHALLENGES TO BE ADDRESSED WITHIN THE THEME OF DATA MANAGEMENT PLANNING

Raising awareness & understanding about DMPs:

- Lack of awareness that DMPs are integral to good research data management (RDM).
- Creating a DMP is often considered an extra task.
- Creating a DMP requires prerequisite knowledge and skills related to RDM, tools and services.
- Insufficient incentives, reward and credit for RDM and the creation of DMPs.

Creation, adoption & implementation of DMPs:

- Lack of clarity about roles and responsibilities for supporting DMP creation and limited staff.
- Lack of automated and machine-actionable processes and workflows to harmonise DMP creation and implementation across stakeholders, tools, services and policies.
- Unclear definitions and language (vocabularies and terminologies) make DMPs challenging to understand and complete.
- Diversity of data means generic DMPs are unsuitable for specific research disciplines.
- As primarily text documents and parts of research proposals, DMPs are not FAIR research objects that enable their utility.

Review, evaluation & assessment of DMPs:

- No clear process or assessment criteria for DMP review and evaluation.
- Lack of accountability on following through on the implementation of DMPs that support funded research.

ACTIONS FOR THE RDA COMMUNITY

Collect 'gold star' DMP case studies. Real-world examples of DMPs from different institutions, disciplines and projects to demonstrate benefits of DMPs for various stakeholders. Case studies and infographics showcase best practices for how to create DMPs and integrate them into the research data lifecycle. This output supplements the 'Engaging Researchers with Data Management The Cookbook' by the RDA Engaging Researchers with Data IG.

Construction of DMP typologies. Define information, language and terminology to be included in DMPs that unambiguously describe RDM concepts and tasks. Typologies leverage existing ontologies and controlled vocabularies (E.g., T4FS, FIP ontology).

Create framework(s) for DMP evaluation and assessment. Collaborate with funding agencies (RDA Funders Forum and Research Funders and Stakeholders on Open Research and Data Management Policies and Practices IG) to create a scalable DMP evaluation framework that defines content to be assessed and evaluation criteria for DMPs dependent on institutional, disciplinary and project contexts.

Develop a toolkit for how to engage researchers in DMPs. DMP professionals, domain experts and RDM supporters co-create a step-by-step guide for how best to drive adoption and implementation of DMPs.

Discussion

Part 3

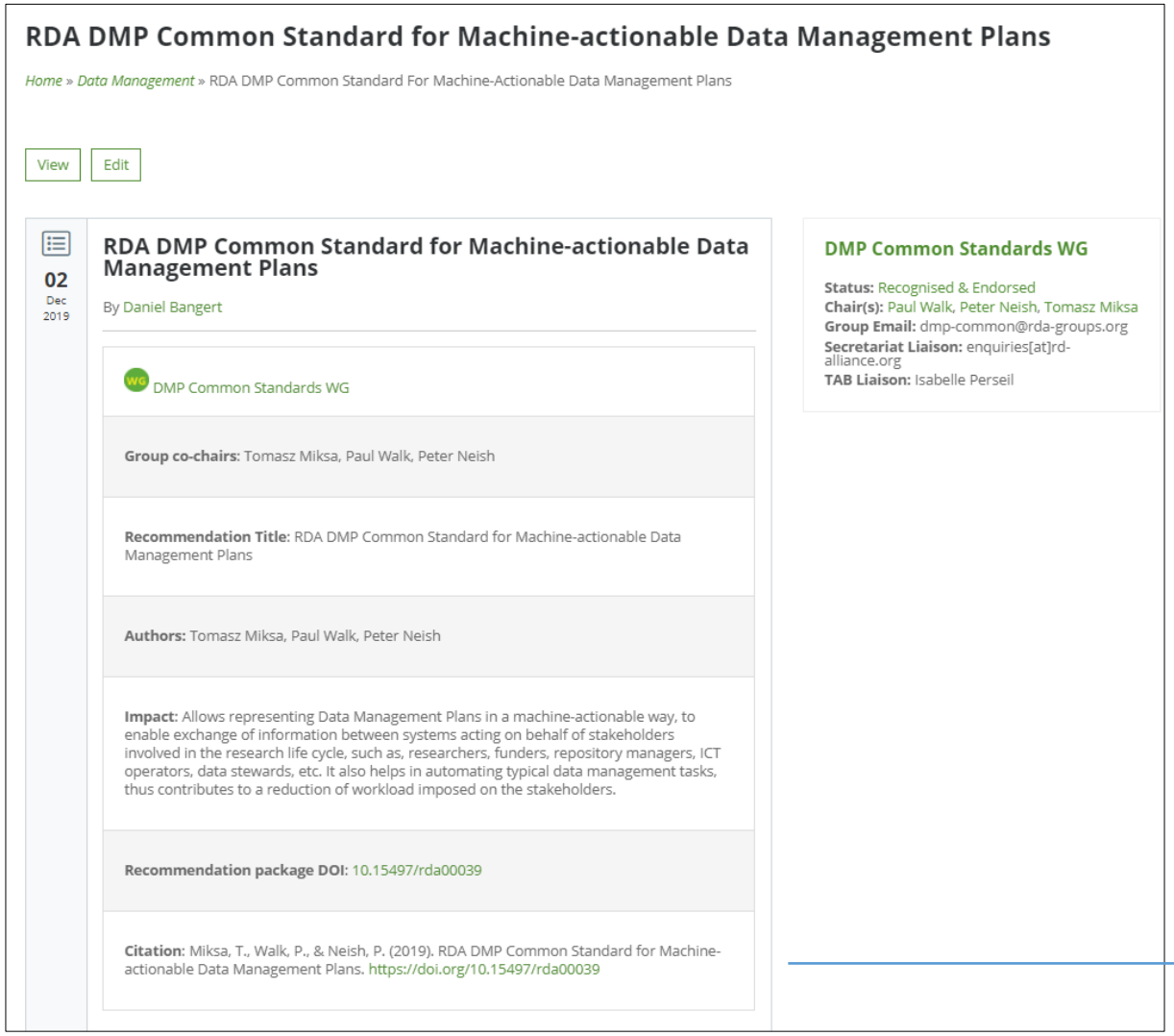
Wrap up

Part 4

Summary

- Bring topics that we could tackle together
- You can always
  - Contact group chairs
  - Write directly to the mailing list
- Further activities can be organised based on the needs
- We hope to see you all in person in Salzburg!

Please cite the recommendation



RDA DMP Common Standard for Machine-actionable Data Management Plans

Home » Data Management » RDA DMP Common Standard For Machine-Actionable Data Management Plans

View Edit

02 Dec 2019

RDA DMP Common Standard for Machine-actionable Data Management Plans

By Daniel Bangert

DMP Common Standards WG

Group co-chairs: Tomasz Miksa, Paul Walk, Peter Neish

Recommendation Title: RDA DMP Common Standard for Machine-actionable Data Management Plans

Authors: Tomasz Miksa, Paul Walk, Peter Neish

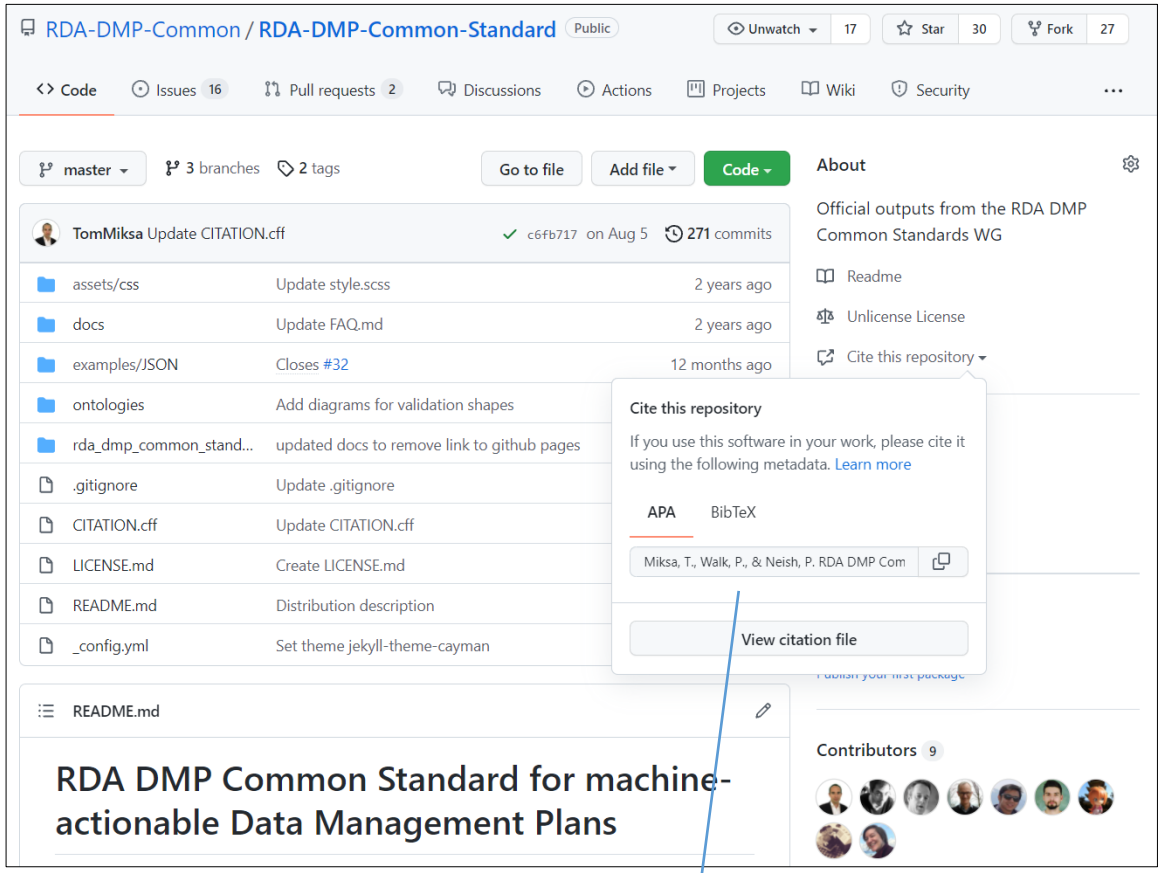
Impact: Allows representing Data Management Plans in a machine-actionable way, to enable exchange of information between systems acting on behalf of stakeholders involved in the research life cycle, such as, researchers, funders, repository managers, ICT operators, data stewards, etc. It also helps in automating typical data management tasks, thus contributes to a reduction of workload imposed on the stakeholders.

Recommendation package DOI: 10.15497/rda00039

Citation: Miksa, T., Walk, P., & Neish, P. (2019). RDA DMP Common Standard for Machine-actionable Data Management Plans. https://doi.org/10.15497/rda00039

DMP Common Standards WG

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
TAB Liaison: Isabelle Perseil



RDA-DMP-Common / RDA-DMP-Common-Standard Public

Unwatch 17 Star 30 Fork 27

Code Issues 16 Pull requests 2 Discussions Actions Projects Wiki Security

master 3 branches 2 tags Go to file Add file Code

TomMiksa Update CITATION.cff c6fb717 on Aug 5 271 commits

assets/css	Update style.scss	2 years ago
docs	Update FAQ.md	2 years ago
examples/JSON	Closes #32	12 months ago
ontologies	Add diagrams for validation shapes	
rda_dmp_common_stand...	updated docs to remove link to github pages	
.gitignore	Update .gitignore	
CITATION.cff	Update CITATION.cff	
LICENSE.md	Create LICENSE.md	
README.md	Distribution description	
_config.yml	Set theme jekyll-theme-cayman	

About

Official outputs from the RDA DMP Common Standards WG

Readme

Unlicense License

Cite this repository

Cite this repository

If you use this software in your work, please cite it using the following metadata. Learn more

APA BibTeX

Miksa, T., Walk, P., & Neish, P. RDA DMP Com

View citation file

README.md

RDA DMP Common Standard for machine-actionable Data Management Plans

Contributors 9

Miksa, T., Walk, P., & Neish, P. (2019). **RDA *DMP Common Standard for Machine-actionable Data Management Plans***. https://doi.org/10.15497/rda00039

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