



# DMP Common Standards WG at the 13<sup>th</sup> Plenary in Philadelphia

Tomasz Miksa, Paul Walk, Peter Neish

# Agenda

- › Part 1 – Introduction for newcomers
  - › goals of this WG
  - › summary of what we did so far
- › Part 2 – Common model for machine-actionable DMPs
  - › Presentation of the model
  - › Discussion
- › Part 3 – Ongoing and future pilot implementations
  - › Evaluations and prototypes developed to date
  - › New ideas for model adoption
  - › Discussion
- › Part 4 – Wrap up and Future of the WG
  - › Timeline
  - › Discussion

# Collaborative notes

› <https://www.rd-alliance.org/wg-dmp-common-standards-rda-13th-plenary-meeting>



Building the social and technical bridges to enable open data sharing

RDA EU RDA US CONTACT US LOGIN REGISTRATION

**O&A Members** 56 Active Organisational & Affiliate members

**MEMBERSHIP** Members: 8106 Becoming a member of RDA is simple and open to both individuals and organizations Register now

**RDA Groups** WG & IGs: 102 Discover what RDA Working and Interest Groups and all other Groups are up to and find out how to join them. [Explore Groups](#)

ABOUT RDA ▾ GET INVOLVED ▾ GROUPS ▾ RECOMMENDATIONS & OUTPUTS ▾ RDA FOR DISCIPLINES ▾ PLENARIES & EVENTS ▾ NEWS & MEDIA ▾

## WG DMP Common Standards - RDA 13th Plenary Meeting

10 JAN 2019 By Tomasz Miksa

**Meeting title**  
DMP Common Model and New Pilots (*Remote Access Instructions*)

**Collaborative session notes:** [https://docs.google.com/document/d/1-bySPMMy3KNb5Ok6Pit129x7\\_eW7kwbmW5ARly9ldgf/edit](https://docs.google.com/document/d/1-bySPMMy3KNb5Ok6Pit129x7_eW7kwbmW5ARly9ldgf/edit)

**Meeting Location:** Commonwealth B

**Short introduction describing the scope of the group and if any previous activities**  
Data Management Plans (DMPs) are semi-structured but largely free-form text documents describing data used and produced in research projects. The workload and bureaucracy often associated with traditional DMPs can be reduced when they become machine-actionable. However, there is no common definition of what machine-actionable DMPs really are. This hinders the communication between stakeholders and leads to scepticism, or conversely to exaggerated expectations.

The RDA working group on DMP Common Standards works to clarify what machine-actionable DMPs are. It develops a common data model for machine-actionable DMPs that will enable exchange of information between systems acting on behalf of stakeholders involved in the research life cycle, such as, researchers, funders, repository managers, ICT providers, librarians, etc. The group also develops standard workflows to demonstrate how the machine-actionable DMPs can be implemented by connecting them to various systems, such as, CRIS, repositories, or funder systems.

During the session we will present the common model for maDMPs and will discuss ongoing and future pilot projects implementing the model.

**Additional links to informative material related to the group**  
Slides from the last plenary summarizing group activity: <https://www.rd-alliance.org/system/files/documents/2018-RDA-DMP-Plenary...>  
Paper summarizing activities of the WG within the first 12 months (presented at IEEE CAS workshop): <http://doi.org/10.5281/zenodo.2161855>  
Paper describing results of the open stakeholder consultation (presented at IPRES 2018): <http://doi.org/10.5281/zenodo.1266211>  
Case Statement: <https://www.rd-alliance.org/sites/default/files/DMPCommonStandards-CaseS...>

Home

**Next Event**

**RDAP SUMMIT 2019** 

**RDAP Summit 2019**

The 2019 Research Data Access and Preservation Summit will be at the University of Miami in Coral Gables, FL from May 15-17, 2019. The theme is Building Communities: how different communities are impacted by our systems, technology, values, and practices; who our communities are by and for; and data services examined through a critical lens.

# Introduction for newcomers

Part 1

# Data Management Plans



How to discover these tools?

Which one do I need to use?

Why do I have to provide the same  
information again?

Why haven't they consulted us before?

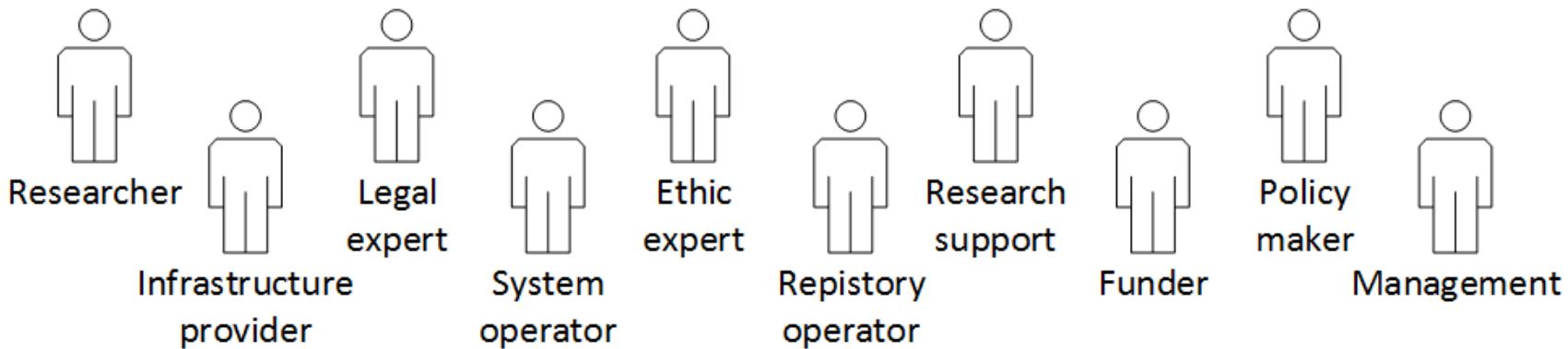
Who is going to pay for this?

We don't have enough people for that!

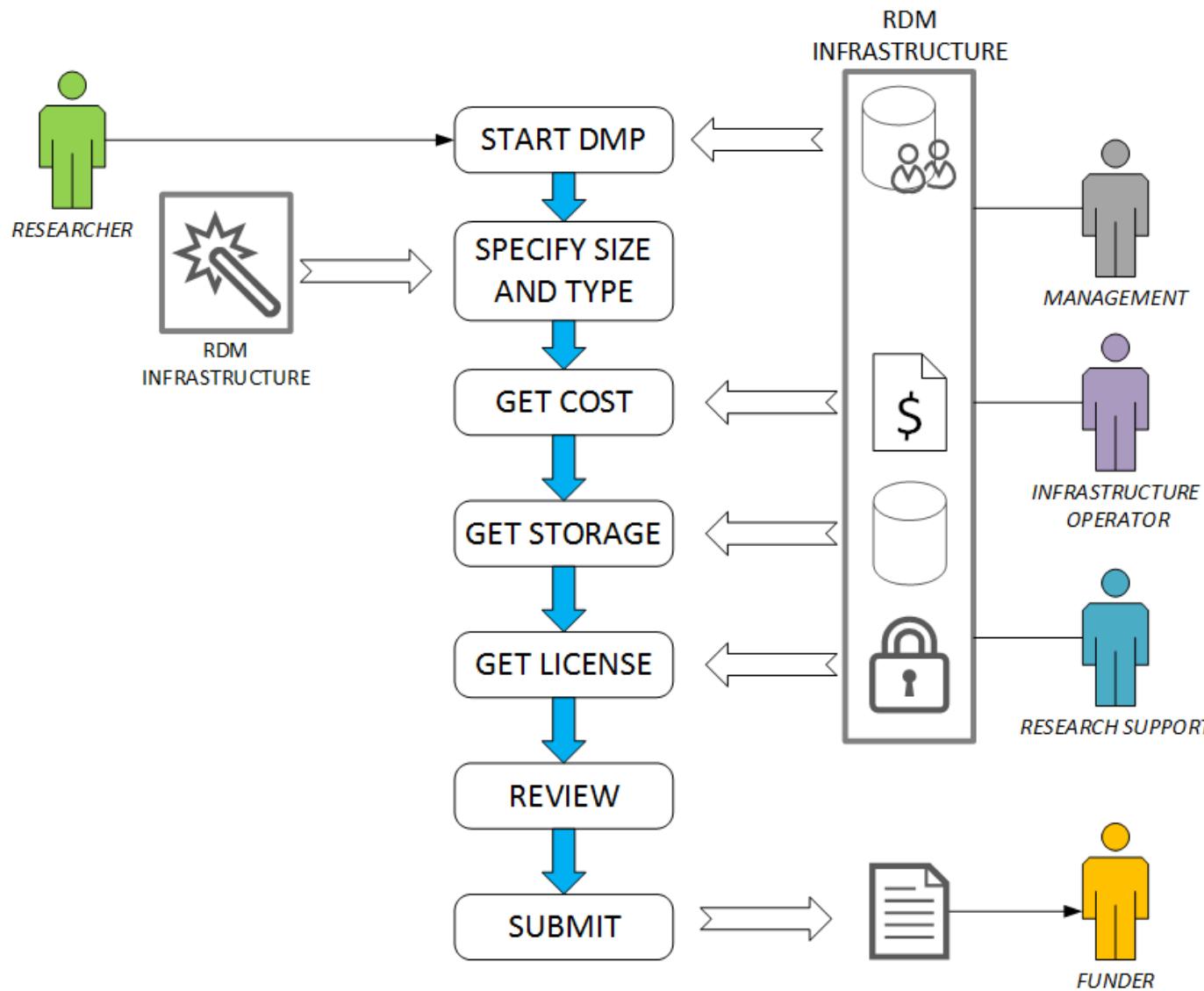


# Research data lifecycle

- › Stakeholders involved in research data management
  - › require information at certain stages
  - › can provide information if requested at a proper stage
- › Many problems can be avoided when
  - › timing is right
  - › information flow is ensured



# Automated Data Management Workflow



# Why do we need this WG?

- › Shortcomings of existing DMPs
  - › manually completed, vague, not updated, ...
- › 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 data model

- › **Common data model for machine-actionable DMPs**
  - › to model information from standard DMPs
  - › NOT a template
  - › NOT a questionnaire
  - › modular design
    - › core set of elements
    - › domain specific extensions
- › **Reference implementations**
  - › ready to use models
    - › JSON, XML, RDF, etc.
- › **Guidelines for adoption of the common data model**
  - › requirements for supporting systems
  - › pilot studies



# Example

- Current DMPs – model questionnaires

```
<administrative_data>
  <question>Who will be the Principle Investigator?</question>
  <answer>The PI will be John Smith from our university.</answer>
</administrative_data>
```

- Machine-actionable DMPs – model information

```
"dc:creator": [ {
    "foaf:name": "John Smith",
    "@id": "orcid.org/0000-1111-2222-3333",
    "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
    "madmp:institution": "AT-Vienna-University-of-Technology"
} ],
```

# Example

- Currently available – not very useful

```
<administrative_data>
  <question>Who will be the Principle Investigator?</question>
  <answer>The PI will be John Smith from our university.</answer>
```

Reuse existing  
standards, e.g. Dublin  
Core, PREMIS, etc.

&gt;

able DMP

```
"dc:creator": [ {
    "foaf:name": "John Smith",
    "@id": "orcid.org/0000-1111-2222-3333",
    "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
    "madmp:institution": "AT-Vienna-University-of-Technology"
} ],
```

# Example

- Currently available – not very useful

```
<administrative_data>
  <question>Who will be the Principle Investigator?</question>
  <answer>The PI will be John Smith from our university.</answer>
</administrative_data>
```

- Machine-actionable DMP

```
"dc:creator": [ {
    "foaf:name": "John Smith",
    "@id": "orcid.org/0000-1111-2222-3333",
    "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
    "madmp:institution": "AT-Vienna-University-of-Technology"
} ],
```

Use PIDs whenever possible, e.g. ORCID

# Example

- Currently available – not very useful

```
<administrative_data>
  <question>Who will be the Principle Investigator?</question>
  <answer>The PI will be John Smith from our university.</answer>
</administrative_data>
```

- Machine-actionable DMP

```
"dc:creator": [ {
    "foaf:name": "John Smith",
    "@id": "orcid.org/0000-1111-2222-3333",
    "foaf:mbox": "mailto:jsmith@tuwien.ac.at",
    "madmp:institution": "AT-Vienna-University-of-Technology"
} ],
```

Use controlled  
vocabularies

# Example

- Currently available – not very useful

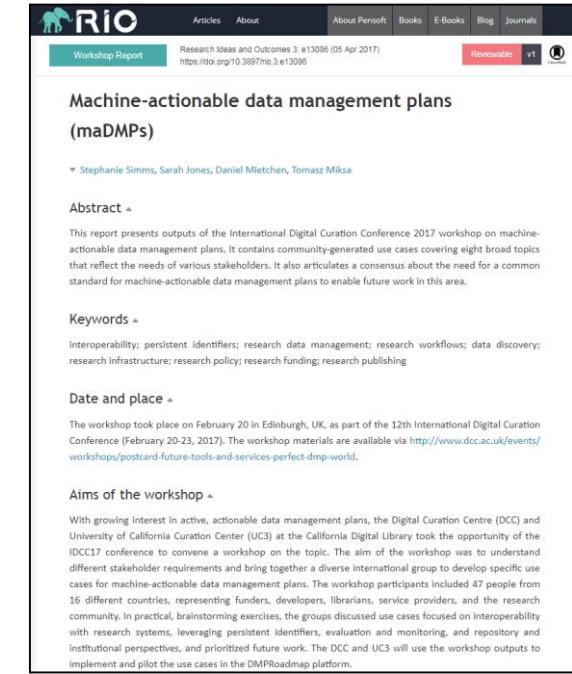
```
<administrative_data>
  <question>Who will be the Principle Investigator?</question>
  <answer>The PI will be John Smith from our university.</answer>
</administrative_data>
```

- Machine-actionable DMP

```
"dc:creator": [
    "foaf:name":"John Smith",
    "@id":"orcid.org/0000-1111-2222-3333",
    "foaf:mbox":"mailto:jsmith@tuwien.ac.at",
    "madmp:institution":"AT-Vienna-University-of-Technology"
  ],
```

Develop own  
concepts and  
vocabularies only  
when needed

- › Launched in October 2017
- › Result of a consultation made by Active DMPs IG
- › Focus on machine-actionable DMPs
- › 100+ members from all continents
- › DMP tool owners are part of it



A screenshot of the Research Ideas and Outcomes (RIO) website. The page title is "Machine-actionable data management plans (maDMPs)". The page includes a brief abstract, keywords (interoperability, persistent identifiers, research data management, research workflows, data discovery, research infrastructure, research policy, research funding, research publishing), and details about the workshop's date and place (February 20, 2017, Edinburgh, UK). At the bottom, there is a link to the workshop report: <https://doi.org/10.3897/rio.3.e13086>.

# Summary of actions till now

- › 1<sup>st</sup> consultation
- › 2<sup>nd</sup> consultation
- › Proof of concept tools
- › BPMN processes
- › Model development
  - › Part 3

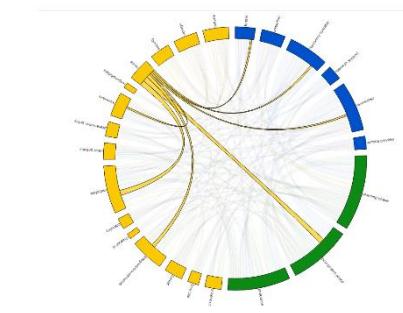
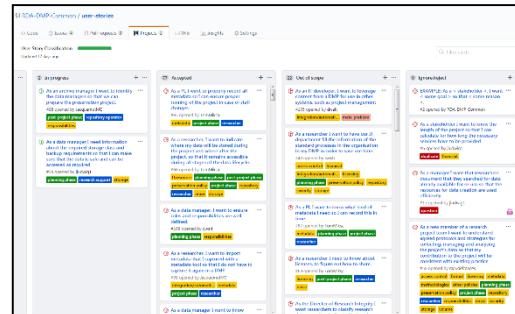
# 1<sup>st</sup> consultation – user stories

## › Goals

- › identify stakeholders at each lifecycle stage
- › define which information they **provide**
- › define which information they **expect**

*As a <stakeholder>, I want <goal> so that <reason >.*

*As a **researcher**, I want to **inform repository operator** on the amount of data in the planning phase, so that they provide **information on costs**.*



# 2<sup>nd</sup> consultation – existing models

## › 2<sup>nd</sup> consultation goes deep

### › how do we model specific requirements

- › which specific fields are needed?
- › which models exist?



### (Meta-) Data

#### Overview

This document is part of a consultation described here: [\[link\]](#)

From the previous consultation with [user stories](#) we have derived following high level requirements:

- Format
  - Format [80, 12, 99, 62, 67, 54, 80]
- Volume
  - Data size estimate [5, 77, 80, 100]
    - For specific type of data [62]
  - Data size real [54]
- Provenance [54]
- Metadata
  - taxonomy/classification [14, 11]
  - Links to metadata of the real data [89, 39]
  - Link publications to data [55]
  - Authorship [88]
  - Multilingual metadata [65]
  - Include raw metadata directly in the model [91, 85]
- Reuse
  - Links to (meta-)data location [89, 90, 56, 39, 60]
- Repository [42]
  - Persistent identifier for data [92]
  - Link publications to data [55, 88]
  - Link to License/Contract allowing data usage/storing [56]

Please help us:

- Break down existing requirements into more specific requirements,
- Add missing requirements,
- Provide examples of existing models, vocabularies, etc. that can be used to model these.

Please provide your suggestions below.

#### Requirements

[Quality](#) - `dqv:hasQualityAnnotation` (statement related to quality of the Dataset, including rating, quality certificate, feedback that can be associated to the Dataset.  
[Stat dimension](#), [stat:measure](#))

[Data Dimensions and units of measurement](#) ([stat:dimension](#), [stat:measure](#))

#### Models

[Format](#): `dct:format`  
[Volume](#): `dct:accrualPeriodicity`

[Provenance](#):  
`dct:creator`, `dcat:contactPoint`, `prov:generated`, [prov:qualifiedAttribution](#)

[Metadata](#):  
`Taxonomy/classification`: `dct:subject`, `dcat:theme`  
`Link publication to data`: `dct:relations` (link to Publications catalogue), `adms:identifier` (link to related publication-identifiers such as DOI, ISSN, ISBN)  
`Authorship`: `dct:publisher`, `prov:agent`, `foaf:name`  
`Conformity to data model`: `dct:conformsTo`  
`Multilingual metadata` - `dct:language`  
`Include raw data in the data model` - `adms:sample` (refers to a sample of data)

[Reuse](#):  
`Links to metadata location` - `dct:source`, `foaf:homePage` (documentation)  
`Repository`:  
`Persistent identifier for data` - `dct:identifier`  
`Link publications to data` - `dcat:distribution`  
`License/contract` - `dct:accessRights`, `dct:licence`

#### Other comments

<https://joinup.ec.europa.eu/release/statdcat-ap-v100>  
<https://joinup.ec.europa.eu/release/dcat-ap-v11>



CC BY-SA 4.0

# Proof of concept tools

## › Requirements

- › Provide minimum input
- › Import as much as possible from existing systems to help in creating maDMPs

## › Tools available as Docker containers on GitHub

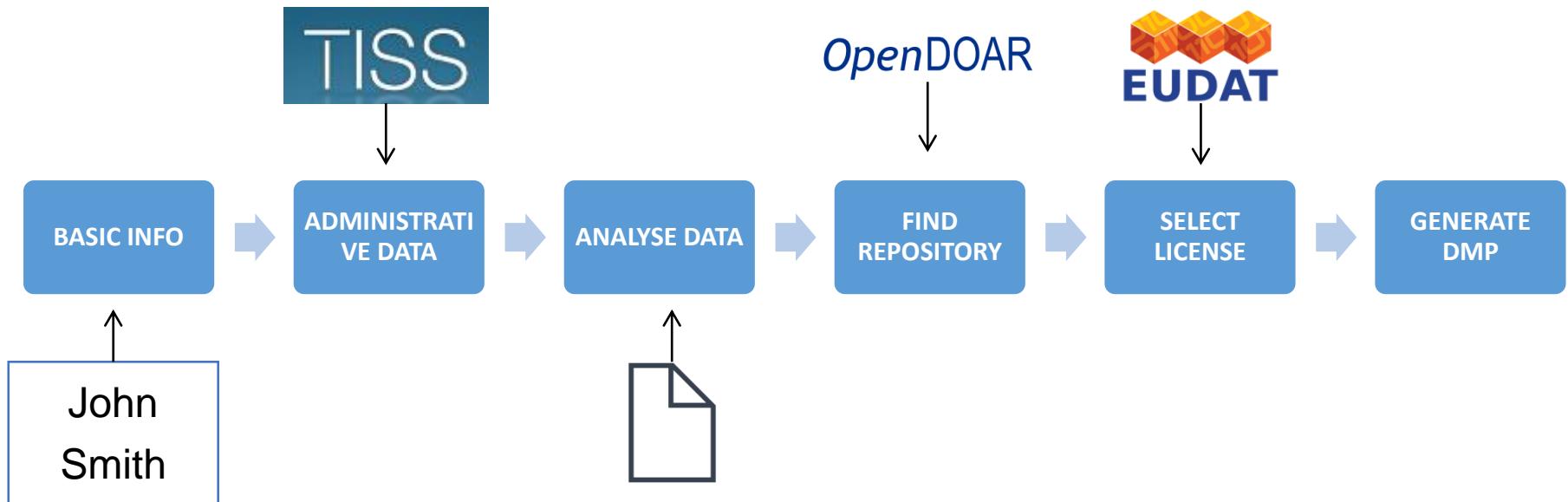
- › <https://github.com/TomMiksa/DMPGenerator>
- › [https://github.com/TomMiksa/digital\\_preservation\\_ex\\_1\\_2](https://github.com/TomMiksa/digital_preservation_ex_1_2)
- › <https://github.com/TomMiksa/tu-dpue-lab2-ss18>
- › [https://github.com/TomMiksa/DigitalPreservation\\_2](https://github.com/TomMiksa/DigitalPreservation_2)
- › <https://github.com/TomMiksa/digitalpreservation-dmp-generator>
- › <https://github.com/TomMiksa/DMPlanner>

## › Example of a landing page for maDMPs

- › <https://oblassers.github.io/fair-data-science/>
- › <https://github.com/oblassers/fair-data-science>

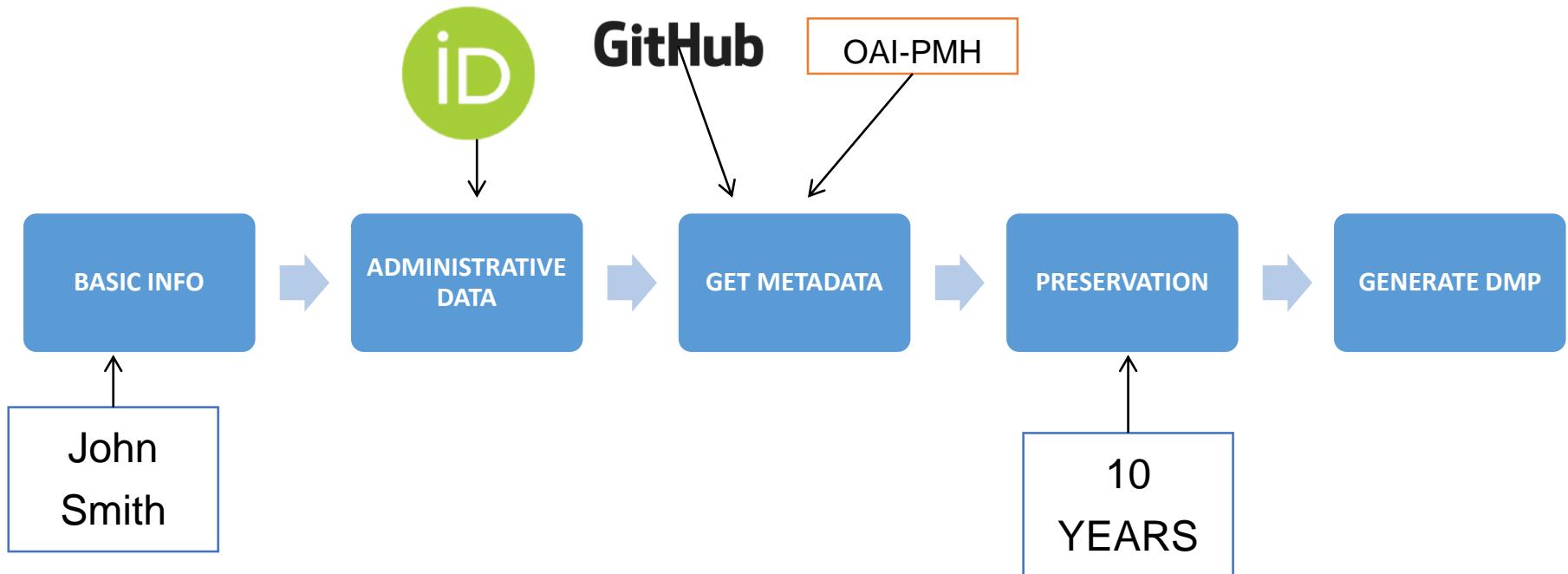
# Planning phase

- Goal: get **estimations** and **recommendations** (which are feasible to implement later)



# Project and Post-project phases

- Goal: **update DMP with real information by re-using (linking)** information provided elsewhere





## Name

Please provide your full name.

 Search ORCID

**full\_name** Tomasz Miksa

**orcid** 0000-0002-4929-7875

**current\_employment\_name** SBA Research



## Resources

Add as many Github repositories or OAI-PMH compliant DOIs as you like.\*

 Add Resource

**Zenodo** Ten Simple Rules For Machine-Actionable Data Management Plans (Preprint)

documentation ▾

Remove

**Github** TomMiksa/DMPlanner

software ▾

Remove



## Preservation Time

Choose how many years the data for each group should be kept.

Software 10 years ▾

Documentation 20 years ▾

# TUW DMP

A Data Management Plan created using DMPlanner.

## Creator

Name: Tomasz Miksa

ORCID: [0000-0002-4929-7875](#)

Current Work: SBA Research

## How will you manage copyright and Intellectual Property Rights (IPR) issues?

The software which was created in the course of the project has the license [restrictions "MIT License"](#).

## Which data are of long-term value and should be retained, shared, and/or preserved?

In this project especially the documentation, as well as the software has a long-term value and should at least be as long preserved as the targeted preservation time specifies. The targeted preservation time for the documentation is 20 years. The targeted preservation time for the software is 10 years.

## What is the long-term preservation plan for the dataset?

One of the main strategies of the long-term preservation plan is the use of public accessible repositories to save the components of the project. The documentation resource "Ten Simple Rules For Machine-Actionable Data Management Plans (Preprint)" is hosted on Zenodo. The software resource "DMPlanner" is hosted on Github.

## How will you share the data?

The data will be primarily shared through the public repositories listed above. This way the data is openly accessible and findable, as well as searchable. The data is available at the repositories as of this moment.

## Are any restrictions on data sharing required?

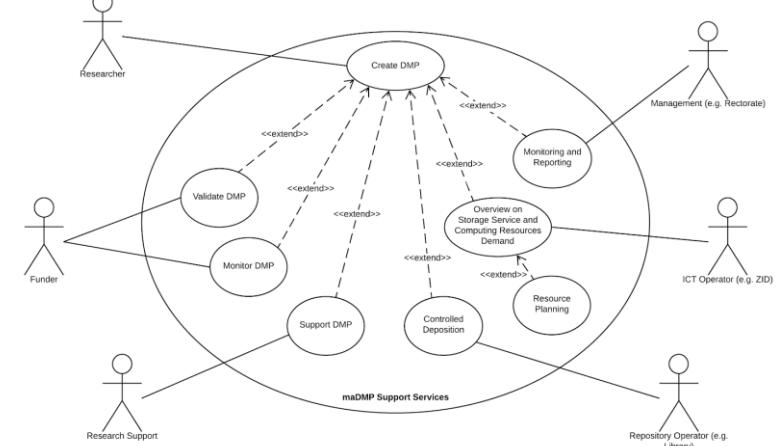
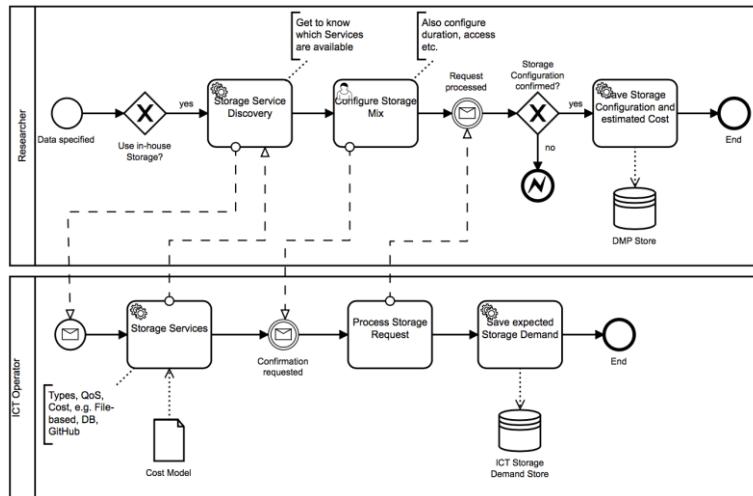
The restrictions on data sharing are composed of the used licenses together with the long-term preservation plan. With this in mind the following restrictions for the resources of the project apply. The documentation resource "Ten Simple Rules For Machine-Actionable Data Management Plans (Preprint)" will be hosted on Zenodo for at least 20 years. The software resource "DMPlanner" will be hosted on Github for at least 10 years.

## Who will be responsible for data management?

The creator of this data management plan is Tomasz Miksa. Therefore Tomasz Miksa is also the reference person for possible reviews and revisions regarding this data management plan in the future. Unless amended Tomasz Miksa is additionally responsible for the adherence to the plan.

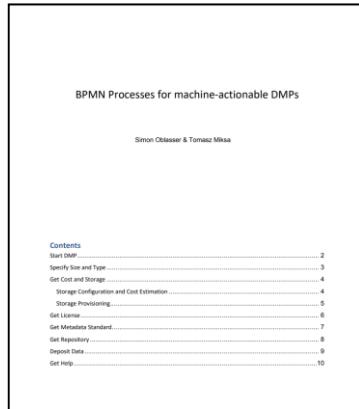
# Processes

- › Processes help identify
    - › **tasks** performed by stakeholders
      - › e.g. ICT operator provide costs of storage
    - › **systems** needed to be put in place
      - › e.g. maDMP repository or costing service
    - › **concepts** to be developed or agreed
      - › e.g. cost model for storage



# Processes

- › Useful in deploying maDMPs
- › Allow us to narrow down focus of this WG
  - › common model does not contain business logic
    - › e.g. cost estimation is done by a service that provides a value
  - › common model is an information carrier
    - › tools, services, processes make maDMPs *machine-actionable*



<http://doi.org/10.5281/zenodo.2607556>

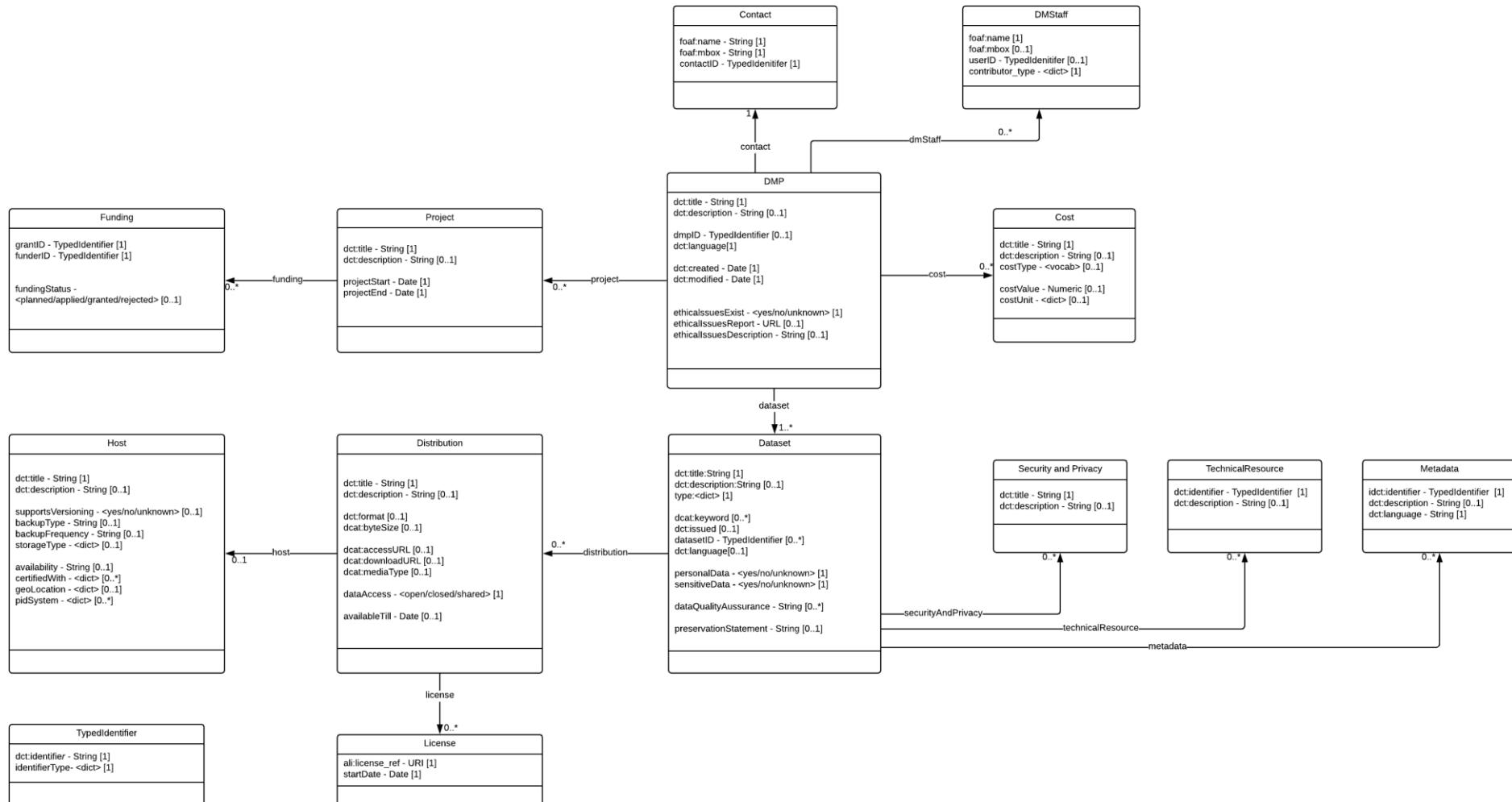
# Summary of actions till now

- › 1<sup>st</sup> consultation (user stories) went broad
  - › to define scope of maDMPs
- › 2<sup>nd</sup> consultation went deep
  - › to identify models for specific requirements
- › Proof of concept tools
  - › to demonstrate how model can be used to automate tasks
- › BPMN processes
  - › to identify systems and stakeholders involved
- › Model development

# Common model for maDMPs

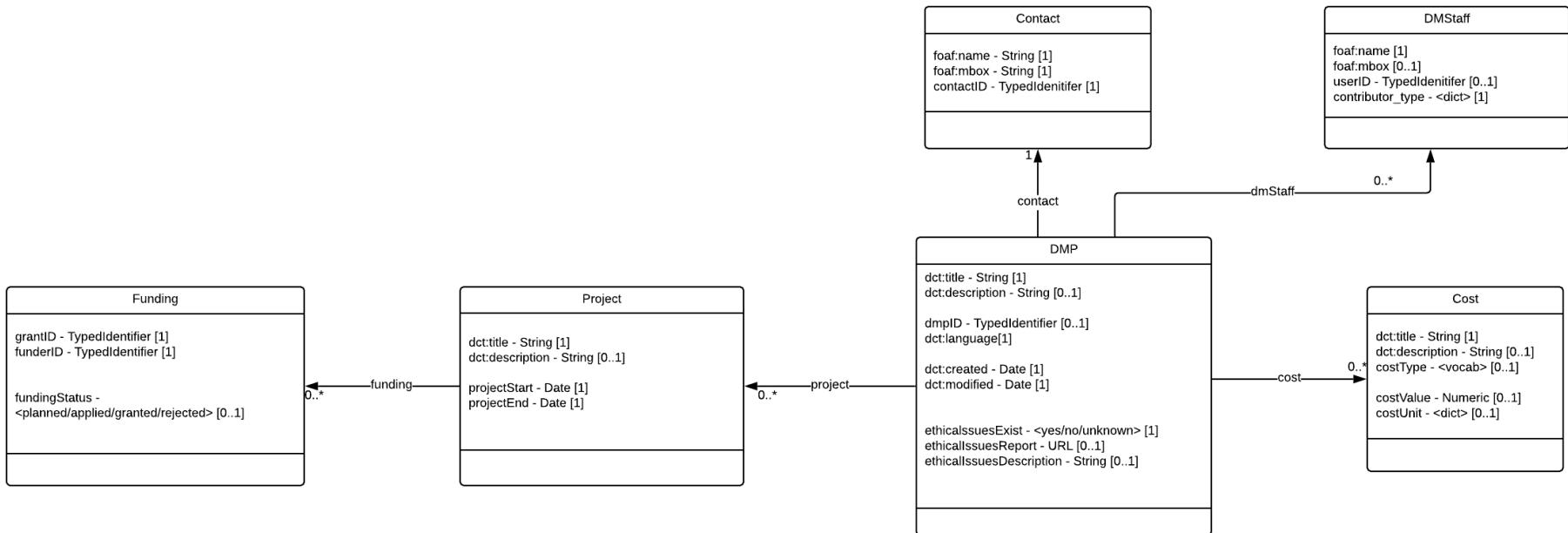
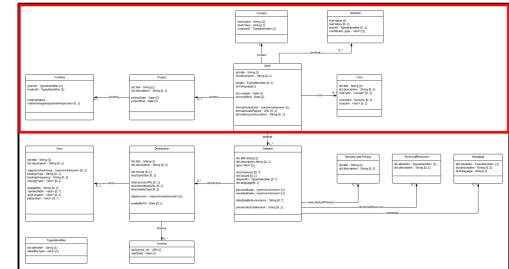
Part 2

# Common model for maDMPs



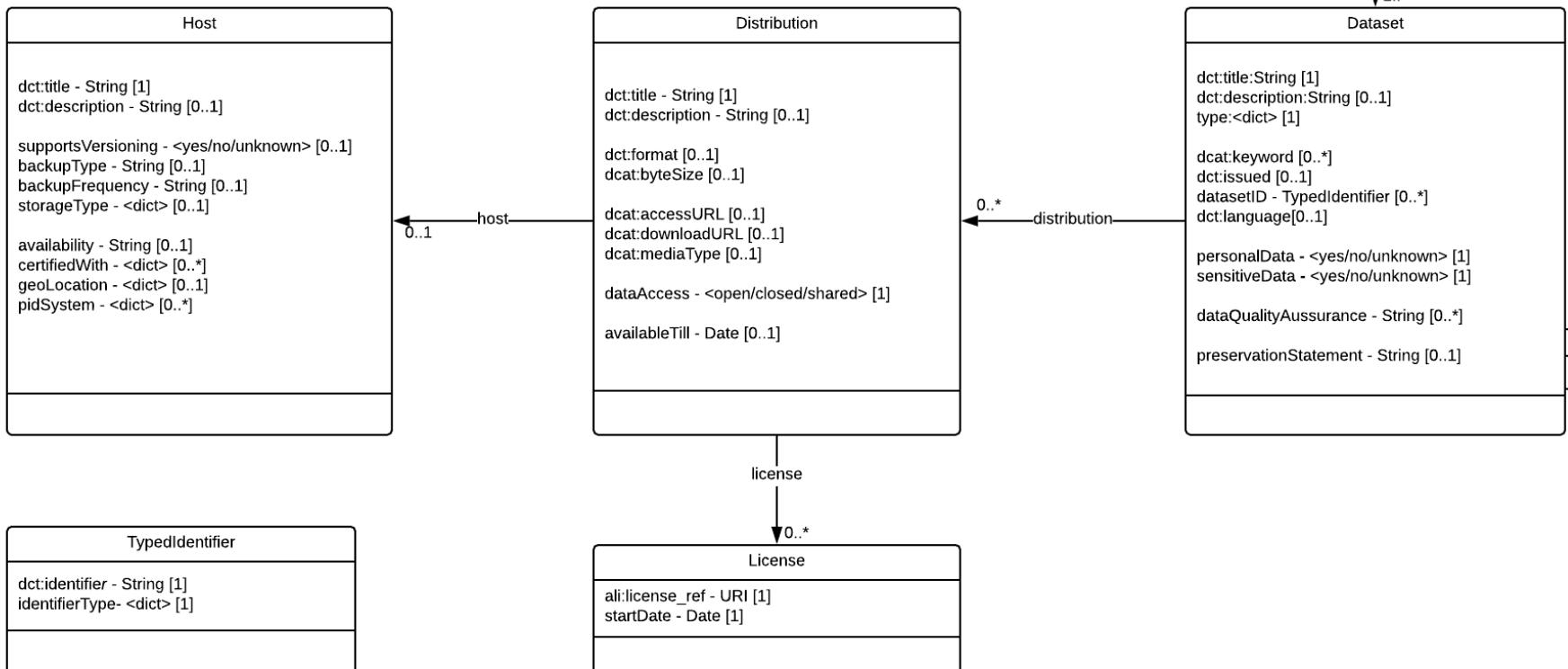
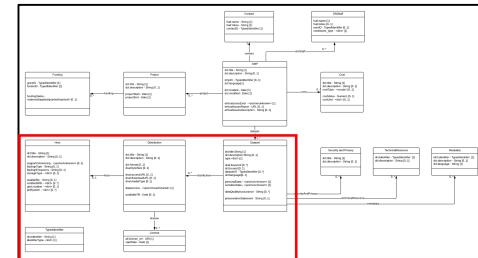
<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/diagrams/RDA-DMP-Common-Model-Diagram-190325.pdf>

# Common model for maDMPs



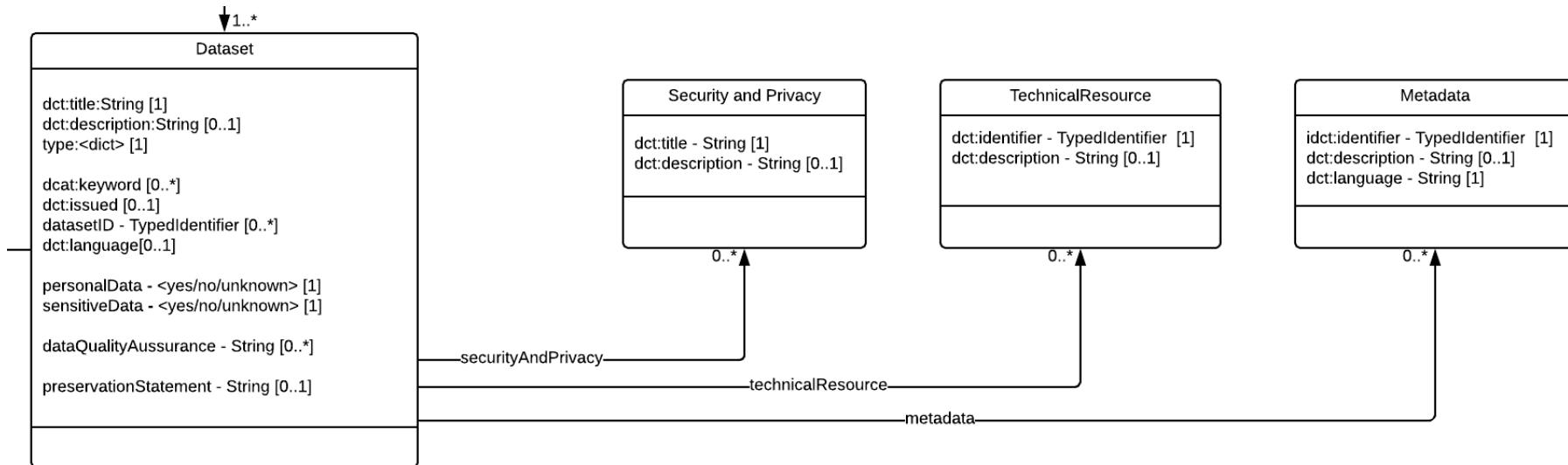
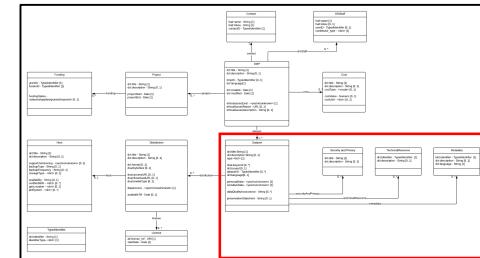
<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/diagrams/RDA-DMP-Common-Model-Diagram-190325.pdf>

# Common model for maDMPs



<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/diagrams/RDA-DMP-Common-Model-Diagram-190325.pdf>

# Common model for maDMPs



# Model - documentation

## Properties in 'contact'

Name	Description	Data Type	Cardinality	Example Value
contact_id	Identifier for a contact person	String	Exactly One	<a href="http://orcid.org/0000-0000-0000-0000">http://orcid.org/0000-0000-0000-0000</a>
mail	E-mail address	String	Exactly One	cc@example.com
name	Name of the contact person	String	Exactly One	Charlie Chaplin

## Properties in 'cost'

Name	Description	Data Type	Cardinality	Example Value
currency_code	Allowed values defined by ISO 4217.	Term from Controlled Vocabulary	Zero or One	EUR
description	Description	String	Zero or One	Costs for maintaining....
title	Title	String	Exactly One	Storage and backup
type	Type	Term from Controlled Vocabulary	Zero or One	
value	Value	Number	Zero or One	1000

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

# Model – FAQ

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

Unwatch 3 Star 0 Fork 5

Code Issues 1 Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master RDA-DMP-Common-Standard / docs / FAQ.md Find file Copy path

TomMiksa Update FAQ.md edd9820 21 hours ago  
1 contributor

85 lines (54 sloc) 8.34 KB Raw Blame History

## Frequently Asked Questions

Index:

- [When to use the model?](#)
- [Do I need to populate all fields?](#)
- [What is the granularity of a Dataset?](#)
- [What is a difference between Dataset and a Distribution?](#)
- [How versioning works?](#)
- [How to express something is planned?](#)
- [How to indicate actions that were performed?](#)
- [How to model embargoes?](#)
- [Why Metadata is referenced from a Dataset?](#)
- [Are there any other serialisations planned different than JSON?](#)
- [Is there a JSON Schema?](#)
- [Is there a model validator?](#)

### When to use the model?

The model is meant for exchange of machine-actionable DMPs between systems. The model is independent of any internal

<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/FAQ.md>

# Model – useful links

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

Unwatch 3 Star 0 Fork 5

Code Issues 1 Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master RDA-DMP-Common-Standard / docs / links.md Find file Copy path

TomMiksa Update links.md f846491 2 days ago

1 contributor

69 lines (45 sloc) 3.84 KB Raw Blame History

## Links

We have collected here links to all important resources created by the [RDA DMP Common Standards WG](#) (official website).

### 1st Consultation - scoping the maDMPs

Collection of user stories to identify scope of maDMPs.

- [Description of the consultation](#)
- [User stories organised on a project board](#)
- [Interactive visualisation of user stories](#)
- [Report from Vienna workshop for collecting user stories](#)
- [iPres conference paper summarising the consultation](#)

### 2nd Consultation - existing models

Collection of models that are relevant in view of requirements derived from the user stories

- [Description of the 2nd consultation \(includes further links\)](#)

<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/links.md>

# Model – JSON examples

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

Unwatch ▾ 3 Star 0 Fork 1

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master RDA-DMP-Common-Standard / examples / JSON / Create new file Upload files Find file History

TomMiksa missing ,	Latest commit ca8c7e6 12 days ago
..	
ex1-header-fundedProject.json	missing , 12 days ago
ex2-dataset-planned.json	JSON examples 12 days ago
ex3-dataset-finished.json	JSON examples 12 days ago
ex4-dataset-embargo.json	JSON examples 12 days ago
ex5-dataset-planned-host.json	JSON examples 12 days ago
ex6-dataset-closed.json	JSON examples 12 days ago
ex7-dataset-many.json	JSON examples 12 days ago

<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/tree/master/examples/JSON>

# DMP and Project – JSON example

40 lines (34 sloc) | 825 Bytes

[Raw](#) [Blame](#) [History](#)   

```
1  {
2      "DMP": {
3          "title": "Funded DMP",
4          "description": "Example of a DMP header for a funded project.",
5
6          "created": "2019-02-22T13:20:15.5",
7          "modified": "2019-02-22T15:10:56.9",
8          "contact": {
9              "name": "First Last",
10             "mbox": "test@test",
11             "contactID": {
12                 "identifier": "https://orcid.org/0000-0002-4929-7875",
13                 "identifierType": "HTTP-ORCID"
14             }
15         },
16         "ethicalIssuesExist": "false",
17
18         "project": {
19             "title": "Making maDMPs awesome",
20             "projectStart": "2017-01-01",
21             "projectEnd": "2020-12-31",
22
23             "funding": {
24                 "funderID": {
25                     "identifier": "501100002428",
26                     "identifierType": "FUNDREF"
27                 },
28                 "grantID": {
29                     "identifier": "1234567-AT",
30                     "identifierType": "custom"
31                 },
32                 "fundingStatus": "granted"
33             }
34         },
35
36         "dataset" : {}
37     }
38 }
```

<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/examples/JSON/ex1-header-fundedProject.json>

# DMP and Project – JSON example

40 lines (34 sloc) | 825 Bytes

```
1  {
2      "DMP": {
3          "title": "Funded DMP",
4          "description": "Example of a DMP header for a funded project.",
5
6          "created": "2019-02-22T13:20:15.5",
7          "modified": "2019-02-22T15:10:56.9",
8          "contact": {
9              "name": "First Last",
10             "mbox": "test@test",
11             "contactID": {
12                 "identifier": "https://orcid.org/0000-0002-4929-7875",
13                 "identifierType": "HTTP-ORCID"
14             }
15         },
16         "ethicalIssuesExist": "false",
```

<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/examples/JSON/ex1-header-fundedProject.json>

# DMP and Project – JSON example

```
18     "project": {  
19         "title": "Making maDMPs awesome",  
20         "projectStart": "2017-01-01",  
21         "projectEnd": "2020-12-31",  
22  
23         "funding": {  
24             "funderID": {  
25                 "identifier": "501100002428",  
26                 "identifierType": "FUNDREF"  
27             },  
28             "grantID": {  
29                 "identifier": "1234567-AT",  
30                 "identifierType": "custom"  
31             },  
32             "fundingStatus": "granted"  
33         }  
34     },  
35  
36     "dataset" : {}  
37  
38 }  
39 }
```

<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/examples/JSON/ex1-header-fundedProject.json>

# Controlled vocabularies

- › Broader problem
  - › identifier-type vocabulary
- › Can RDA help in solving this?

Name	Description	Data Type	Cardinality	Example Value
dataset_id	Dataset ID	String	Exactly One	<a href="http://doi.org/10.5281/zenodo.1172673">http://doi.org/10.5281/zenodo.1172673</a>
dataset_id_type	Identifier type	Term from Controlled Vocabulary	Exactly One	HTTP-DOI

Name	Description	Data Type	Cardinality	Example Value
funder_id	Funder ID, recommended to use CrossRef Funder Registry. See: <a href="https://www.crossref.org/services/funder-registry/">https://www.crossref.org/services/funder-registry/</a>	String	Exactly One	501100002428
funder_id_type	Identifier type	Term from Controlled Vocabulary	Exactly One	FUNDREF

# Model – reused standards

<b>id</b>	<b>label</b>	<b>uri</b>
ali	Access License and Indicators	<a href="http://www.niso.org/schemas/ali/1.0/">http://www.niso.org/schemas/ali/1.0/</a>
dces	Dublin Core Element Set	<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>
dct	DCMI Metadata Terms	<a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a>
foaf	Friend of a Friend (FOAF)	<a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>
dcat	DCAT	<a href="https://www.w3.org/TR/vocab-dcat/">https://www.w3.org/TR/vocab-dcat/</a>
datacite	Data Cite	<a href="https://schema.datacite.org">https://schema.datacite.org</a>
cerif	Cerif	<a href="https://www.eurocris.org/ontologies/cerif/1.3/index.html#currencyCode">https://www.eurocris.org/ontologies/cerif/1.3/index.html#currencyCode</a>
coar	COAR	<a href="http://vocabularies.coar-repositories.org/pubby/resource_type.html">http://vocabularies.coar-repositories.org/pubby/resource_type.html</a>
iso6391	ISO 6391-1	Two letter country code
iso4217	ISO 4217	Currency code

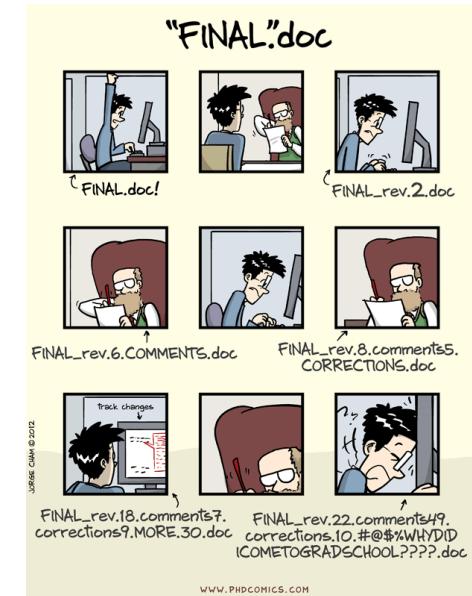
# Model assumptions – relaxed constraints

- › Model must be applicable in different settings
  - › relaxed constraints within the model
    - › e.g. DMP **can** relate to a project [0..\*]
  - › constraints introduced at the ‘business level’
    - › tool implementing the model
    - › e.g. DMP **must** relate to a project [1..\*]
  - › DMP instances are still compatible



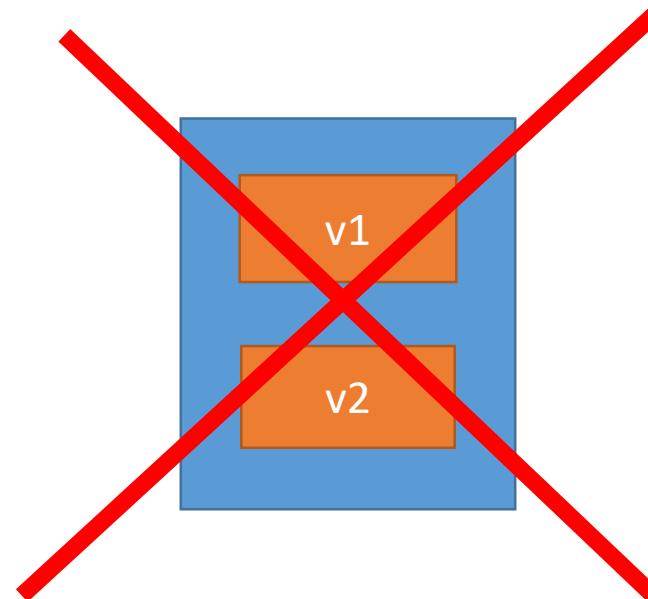
# Model assumptions - interoperability

- › Model will be pre-dominantly used to exchange information between systems
  - › internal representation of information in a DMP tool may differ (physical model)
    - › e.g. database may have a different schema
  - › No ‘meta-fields’ about DMP
    - › e.g. no DMP state field ‘final’



# Model assumptions - versioning

- › DMP versioning done by systems using the model
  - › model provides fields allowing to identify DMP version
  - › model does not track connections between versions



# Model assumptions – evolving information

- › Model expresses 'certainty' of provided information
  - › to support different phases of DMPs
- › Example
  - › Source code will be issued on 2019-06-30 (planned) in 'some-repo'. There will be an embargo period till 2019-12-31. Later on the source code will be available on a CC-BY license.

```
"DMP": {  
    "modified": "2019-02-22T13:20:15.5"  
    "dataset": {  
        "title": "Source Code",  
        "issued": "2019-06-30",  
        "distribution": {  
            "accessURL": "http://some-repo...",  
            "license": {  
                "license_ref": "https://creativecommons.org/licenses/by/4.0/",  
                "startDate": "2019-12-31  
            }  
        }  
    }  
}
```

# Ongoing and future pilot implementations

Part 3



TECHNISCHE  
UNIVERSITÄT  
WIEN

# RDM Infrastructure

Pilot project

# Mock-up for a tool + prototype

## › Goal

- › generate easily and quickly DMPs
- › not a training tool

## › Mock-ups

- › To define requirements of ALL stakeholders

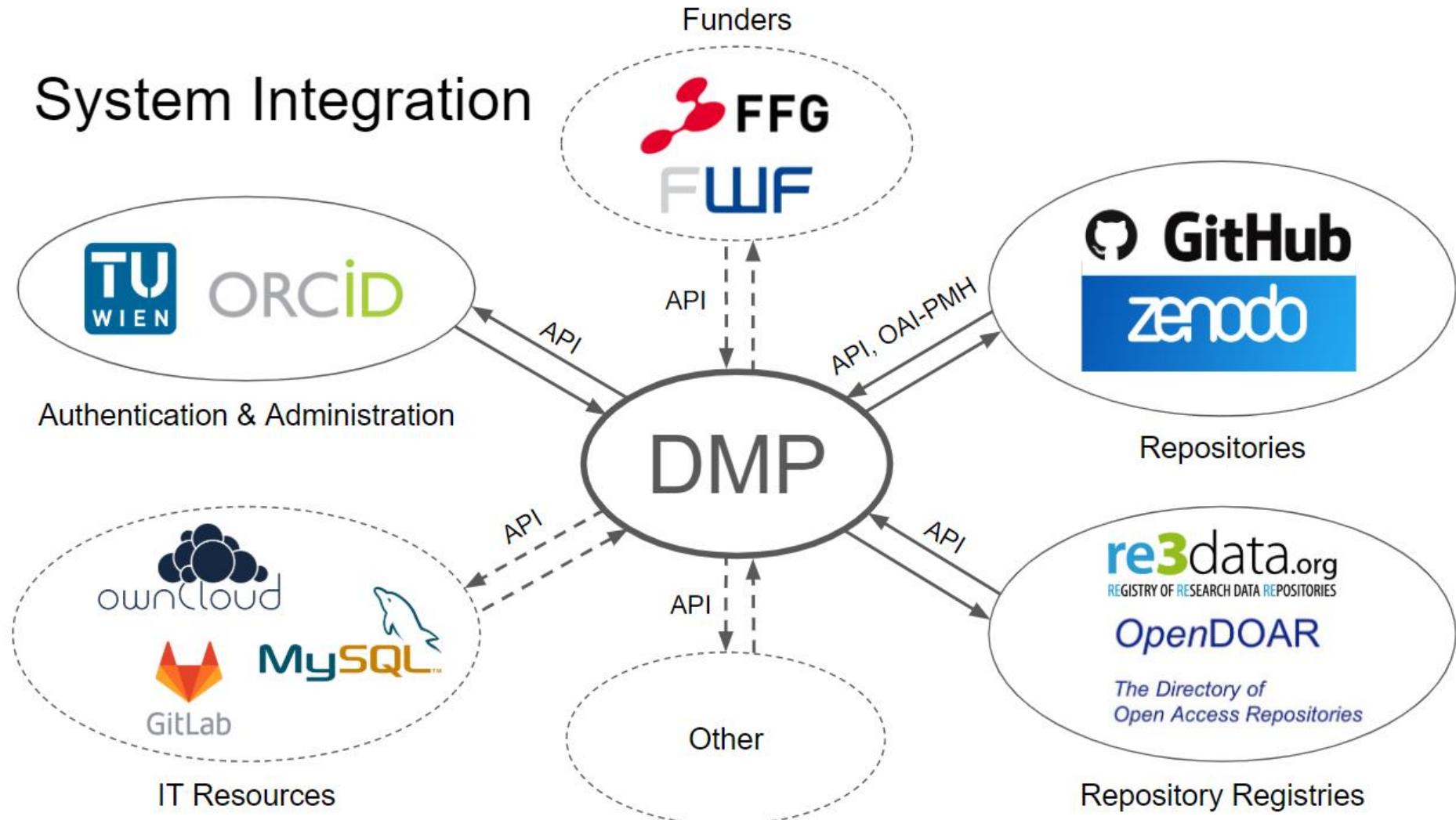
## › Deployment requires integration with university specific services

- › e.g. researchers database, research support ticket system, etc.

## › There are common services to be co-developed

- › e.g. repository recommendation service

# System Integration



# Mock-up of a funder view for maDMP

DMP Funder View

<https://dmpviewer.tuwien.ac.at/dmps/54365437012341>

## DMP Funder View

[Home](#) > [DMPs](#) > DMP#54365437012341

### Reuse of pre-existing data

Dataset title	Origin	License
Calculating Thermal Bremsstrahlung Emission from Stellar Winds	<a href="https://doi.org/10.5281/zenodo.1476587">doi:10.5281/zenodo.1476587</a>	MIT
Occurrence records download on 2018-11-05	<a href="https://doi.org/10.26197/5be00801ec357">doi:10.26197/5be00801ec357</a>	CC-BY

### FAIR Data

#### Metadata standards

- [Dublin Core](#)
- [DataCite Metadata Schema](#)
- [DDI - Data Documentation Initiative](#)
- [CIF \(Crystallographic Information Framework\)](#)
- [CSMD \(Core Scientific Metadata Model\)](#)

#### Metadata

Lore ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

#### Inferred FAIRness by repository selection

Selected repository	Dataset	Data access	PID system	AID system	Certificate	Quality Mgmt.	Versioning	Location	API
<a href="#">GitHub</a>	Source code for client application	open	none	none	none	no	yes	U.S.	<a href="#">other</a>
<a href="#">Zenodo</a>	Supplementary material	open	DOI	ORCID	none	yes	yes	EU	<a href="#">REST OAI-PMH</a>
<a href="#">GESIS Data Archive</a>	Raw data Processed data	open	DOI	none	CoreTrustSeal	no	-	Germany	<a href="#">OAI-PMH</a>

#### Licensing

Dataset	Sharing strategy	Selected license	License planned to be active from
Supplementary material	keep closed	-	-
Raw data	keep closed	-	-
Source code for client application	publish	<a href="#">Apache License 2</a>	2020-01-01
Processed data	publish	<a href="#">Creative Commons Attribute (CC-BY)</a>	2021-03-01

Please click on the scrollbar to see more.
 



RESEARCH D

# Mockups

Machine-actionable Data Management Planning Application

[View on GitHub](#)

## Introduction

Currently we are designing a system to make research data management planning machine-actionable. This involves the automation of workflows and exchange among information systems and services. If you are interested in machine-actionable DMPs or are a stakeholder of research data management (e.g. researcher), feel welcome to **try out our mockups** and **give us feedback**. Your help is very appreciated!

<https://oblassers.github.io/dmap-mockups/>

# John Chodacki

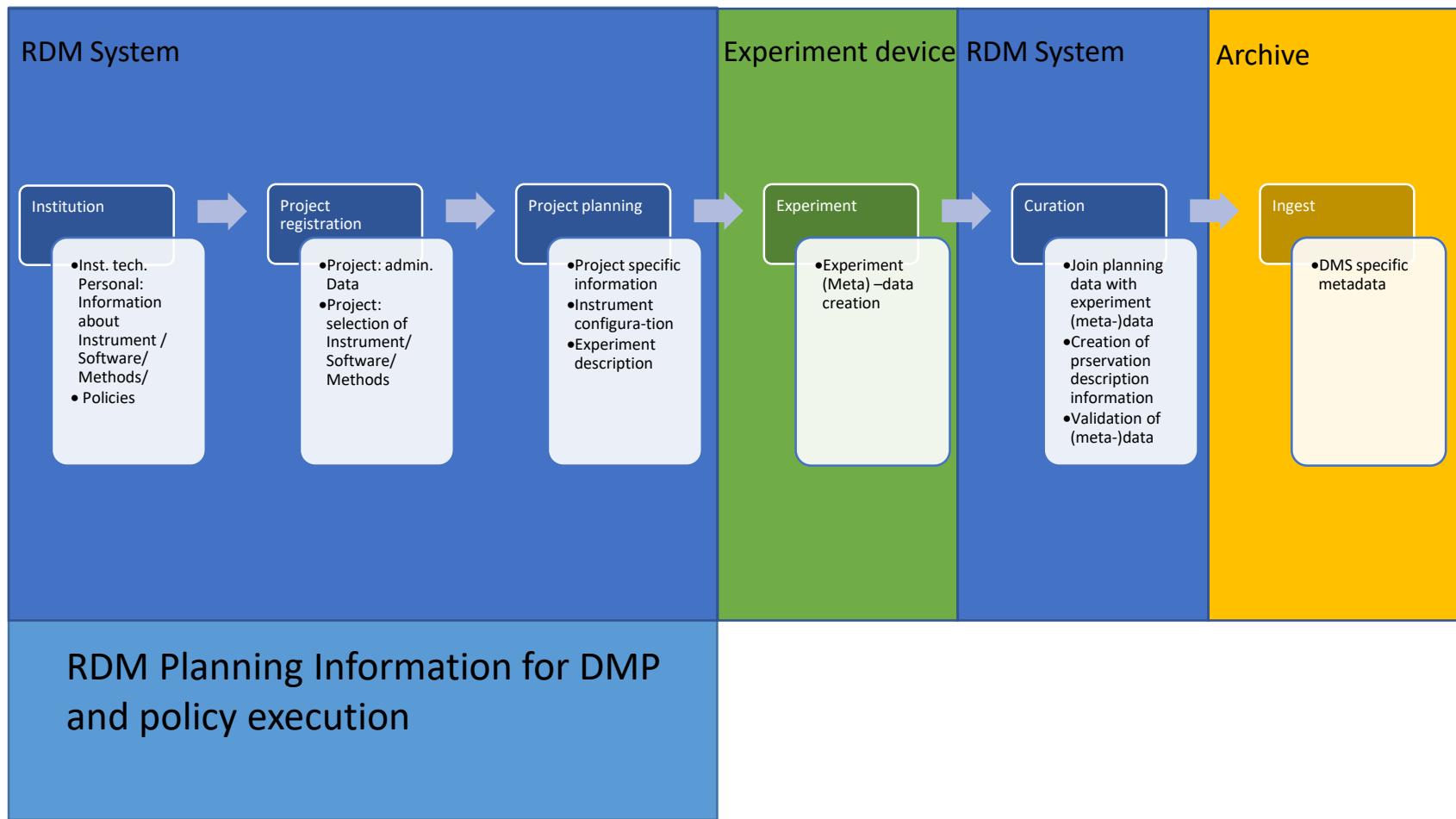
Idea 1

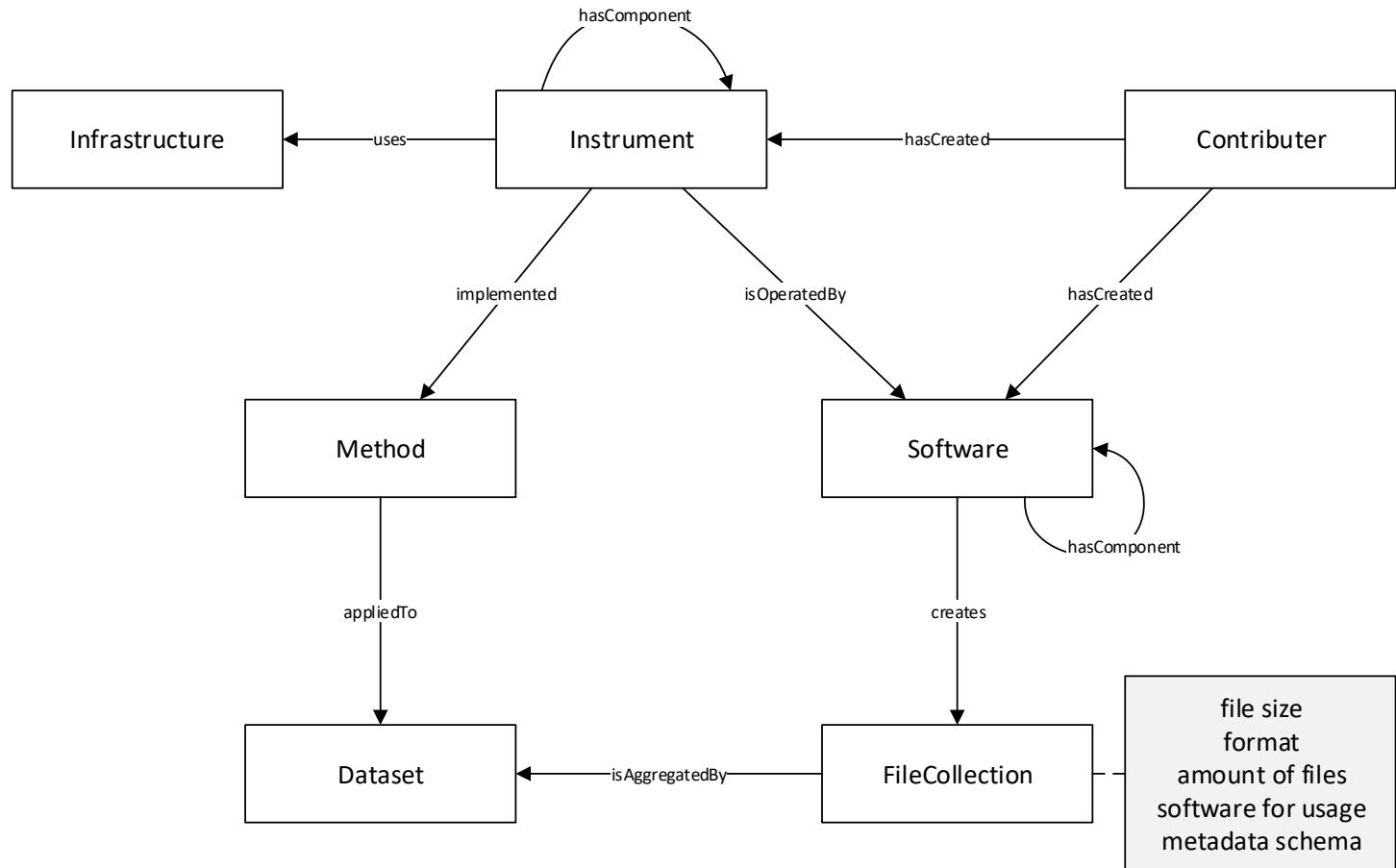
# Heike Görzig

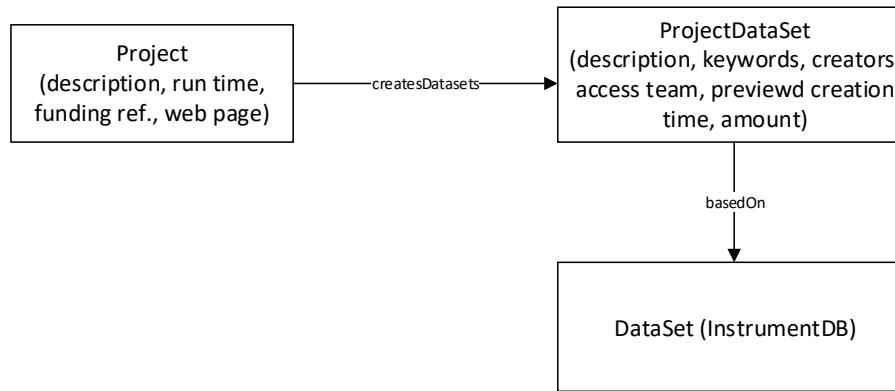
# Application of DMP-common-standard model on Horizon 2020 Template (FAIR)

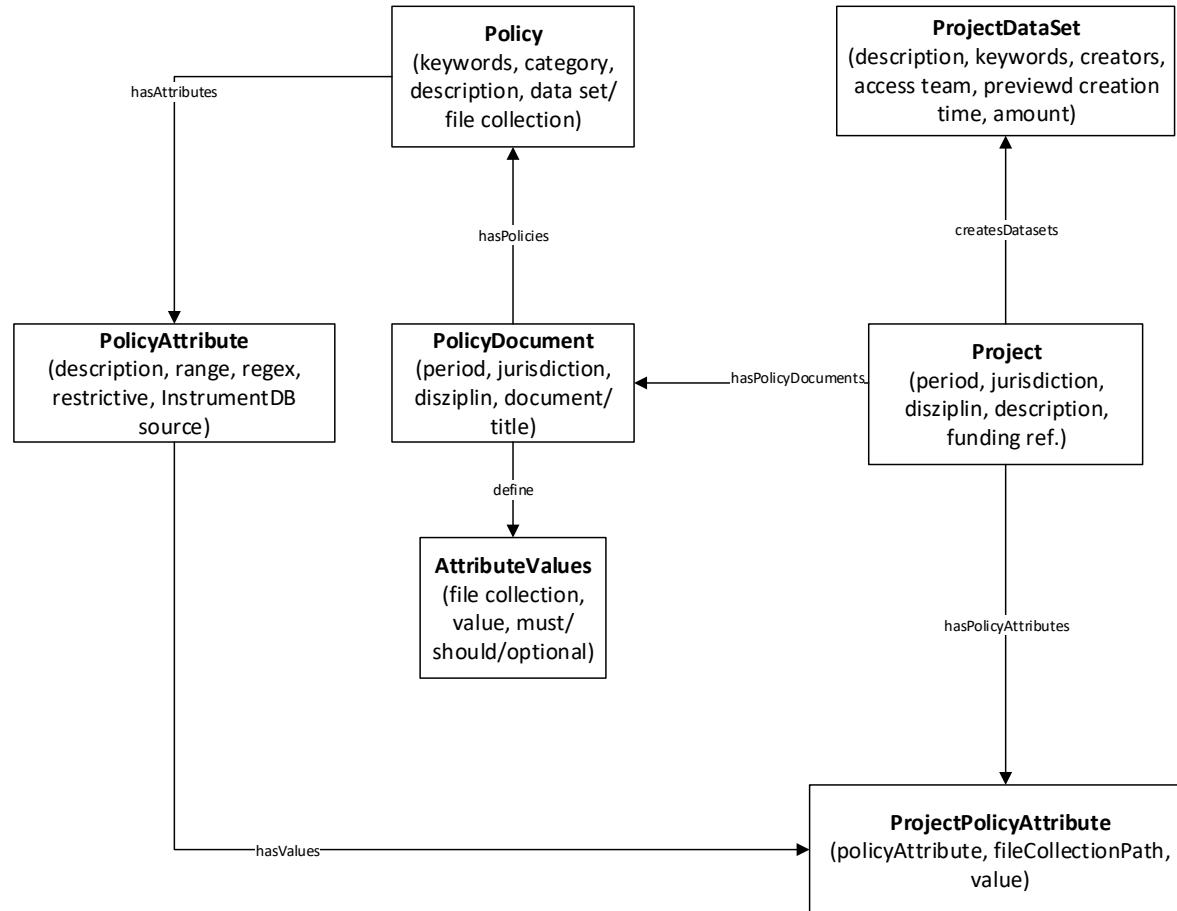
Mapping to InstrumentDB of RDMinfoPool

- Helmholtz Zentrum Berlin
- Photon and Neutron Science, Laboratories
- No personal data
- Data Policy
- Repository – ICAT









# FAIR

## What metadata will be created?

### DMP-common-standards

Metadata:  
identifier  
description  
language

### RDMinfoPool

Metadata\_schema:  
name  
description  
namespace  
metadata\_schema  
metadata\_catalog:yes

## Will search keywords be provided that optimize possibilities for re-use?

### DMP-common-standards

Dataset:dcat:keyword

### RDMinfoPool

Static keywords generated from instrument name, method name, software name, Dataset:keywords on data curation

# FAIR

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

DMP-common-standards	RDMinfoPool
Metadata: identifier, description, language	file: file_metadata_schema (referencing table Metadata_schema)
Dataset:description	method:name, description

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. ...?

DMP-common-standards	RDMinfoPool
Distribution:format	file: format
	file_collection: read_software
For describing provenance: Technical resource: description	For describing provenance: instrument:instrument_metadata software:name, version , configuration

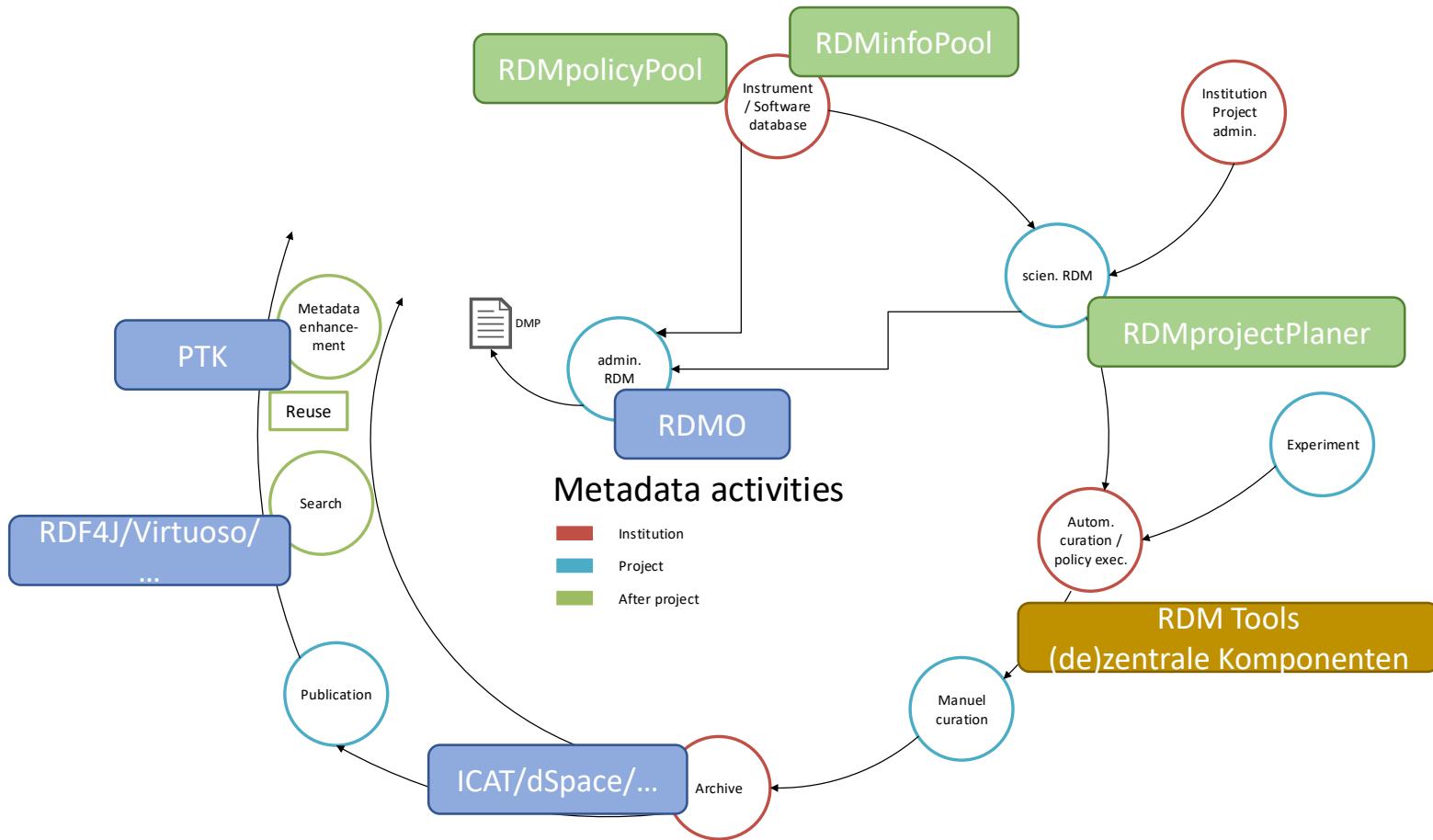
# FAIR

How long is it intended that the data remains re-usable? Are data quality assurance processes described?

DMP-common-standards	RDMinfoPool
Dataset:dataQualityAssurance	Policy:description where policy type is validation (metadata_schema:metadata_schema, file, format parameter for validations with PolicyDB (which needs to be implemented))

# Thank you!

Role	Before project / OPA / EPA	Project initiation	Project planning	Project execution	Project finalisation	After project / OPA
<b>Instrument scientist</b>	Instrument and software description, selection of applicable metadata standards, general description of datasets					Actualization of instruments and software
<b>Data manager</b>	Administration of controlled vocabularies and standards, mapping metadata to standards, general data policies, policy execution.				Automatic metadata extraction and validation	Open access of research data, validation of policy execution. Actualization of standards and policies
<b>User co-ordination</b>		Initializing proposal				
<b>Researchers</b>			Dataset description, metadata schema selection, amount of datasets produced	Experiment execution: parameter and configurations	Dataset selection, metadata completion and validation	
<b>Project administration</b>		Specific policies, DMP	DMP actualisation	DMP actualisation after experiments		



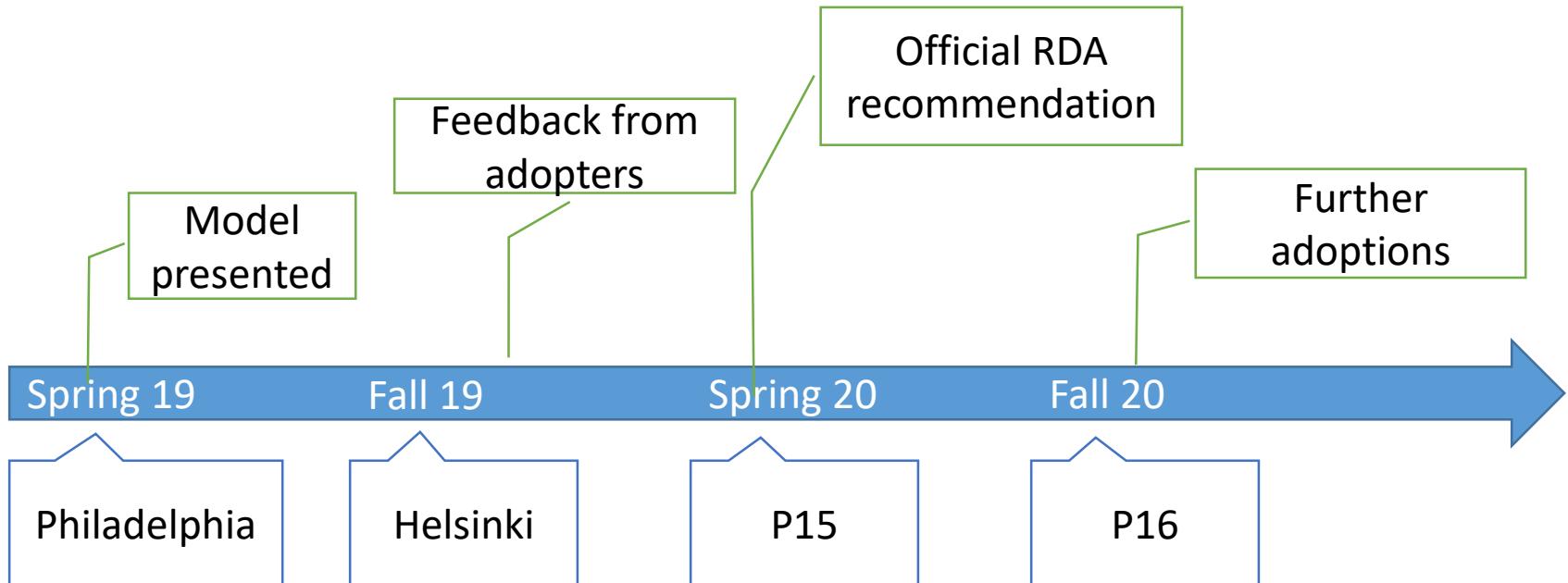
# Discussion

# Wrap-up and next steps

Part 4

# Next steps – long term goals

- › Focus on adoption
- › Standard is frozen
- › Create change requests on GitHub
  - › To request new fields or modify existing
- › Standard will be reviewed before the next plenary



# Next steps – mid-term goals

- › JSON Schema development
- › Validator for maDMP instances
- › Further serialisations
  - › XML
  - › OWL
  - › JSON-LD
- › Assistance in complying with the standard

# Visit Active DMPs IG session

- › More on machine-actionable DMPs
- › [Active Data Management Plans: Machine-actionable DMPs - Revisit Requirements, Review Outputs, Reflect on Next Steps](#)
- › Day 2
- › 12:00-13:30

# Staying in touch!

- › Sign up to the group
  - › <https://www.rd-alliance.org/groups/dmp-common-standards-wg>
- › Visit GitHub repository
  - › <https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard>
- › Participate in model adoption
- › Contact group chairs



Tomasz Miksa



Paul Walk



Peter Neish

# Publications

- › [Tomasz Miksa, Peter Neish, Paul Walk, Andreas Rauber: Defining requirements for machine-actionable Data Management Plans. iPres 2018](#)
- › [Tomasz Miksa, Stephanie Simms, Daniel Mietchen, Sarah Jones: Ten principles for machine-actionable data management plans. PLOS Computational Biology \(in press\)](#)
- › [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](#)