

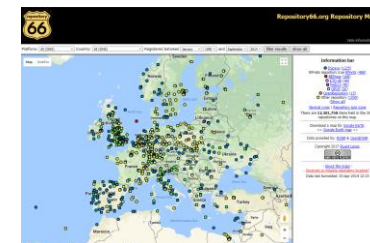
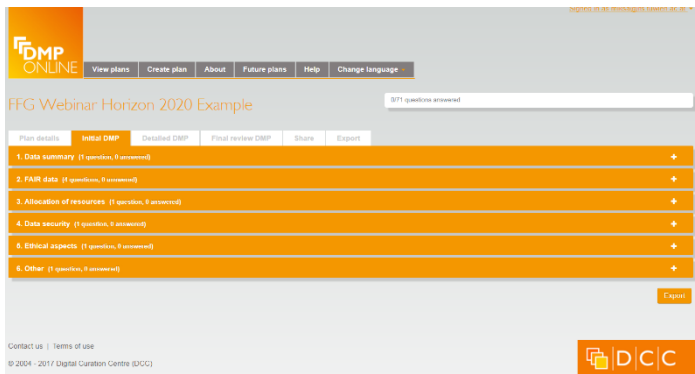
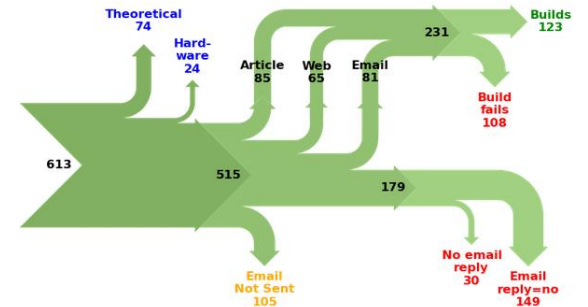
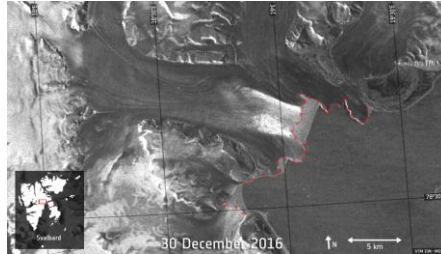
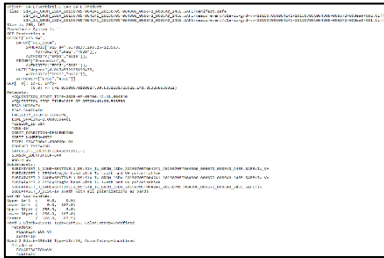
Machine-actionable Data Management Plans

Dr. Tomasz Miksa

SBA Research & TU Wien

tmiksa@sba-research.org

Previous lecture on DMPs



Data Management Plans (DMPs)

- manually created text documents
- considered as bureaucracy
- created too late
- vague
- depend on human factor
 - scrupulousness
 - awareness



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?

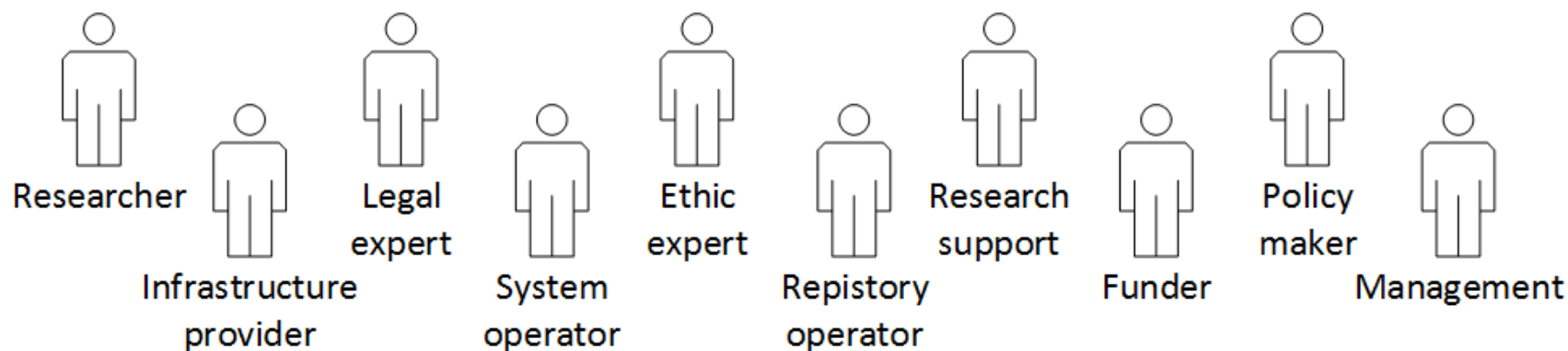


Why haven't they consulted us before?
Who is going to pay for this?
We don't have enough people for that!

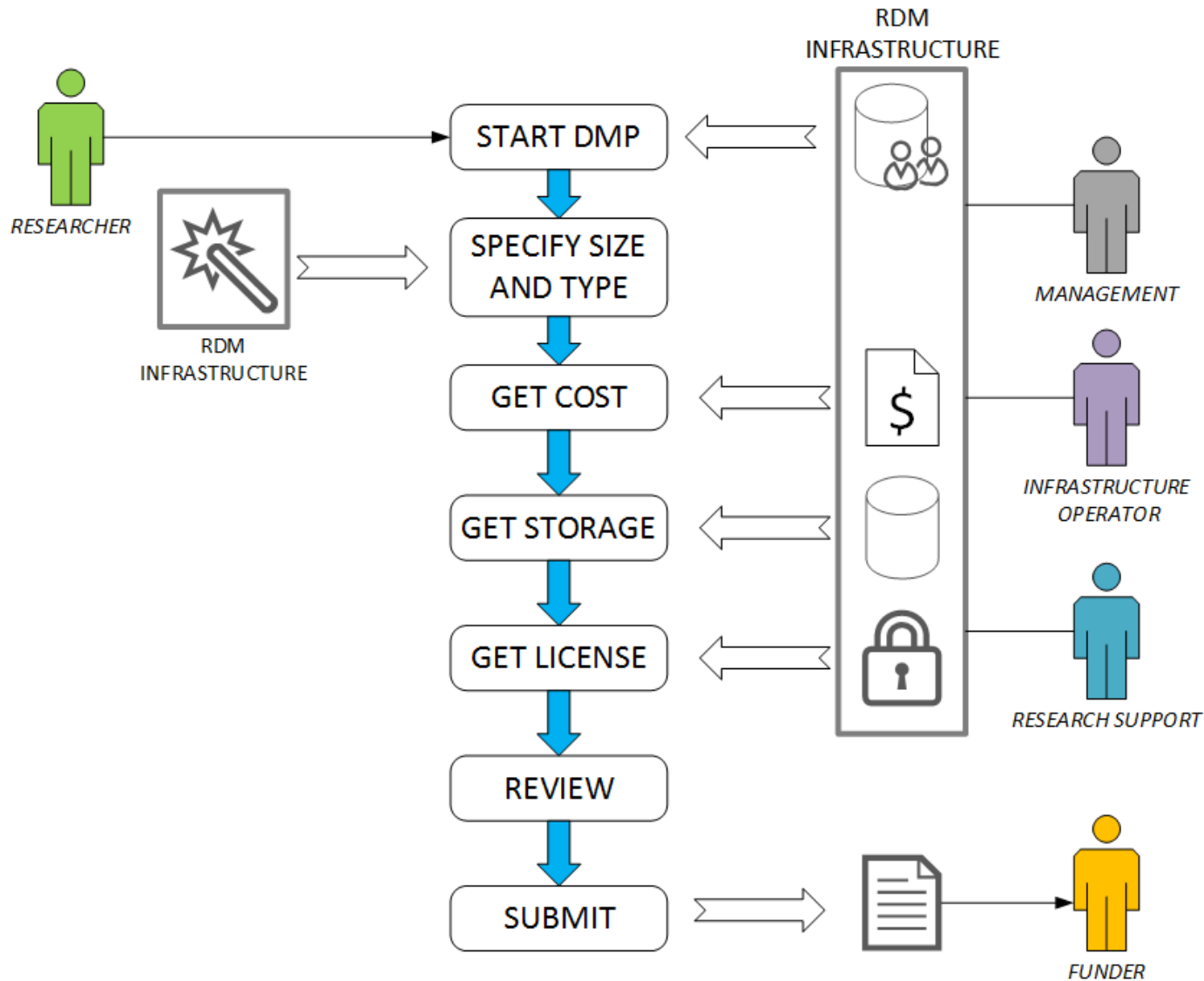


Research data lifecycle

- 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



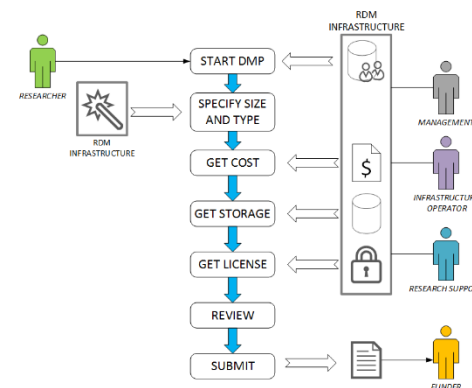
Automated Data Management Workflow



Automated Data Management Workflow

■ Requires

- common data model
 - to exchange information
- well-defined RDM workflows
 - Who? What? When? How?
- data management infrastructure
 - systems and services implementing workflows



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>

John Smith from our university.</answer>

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

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

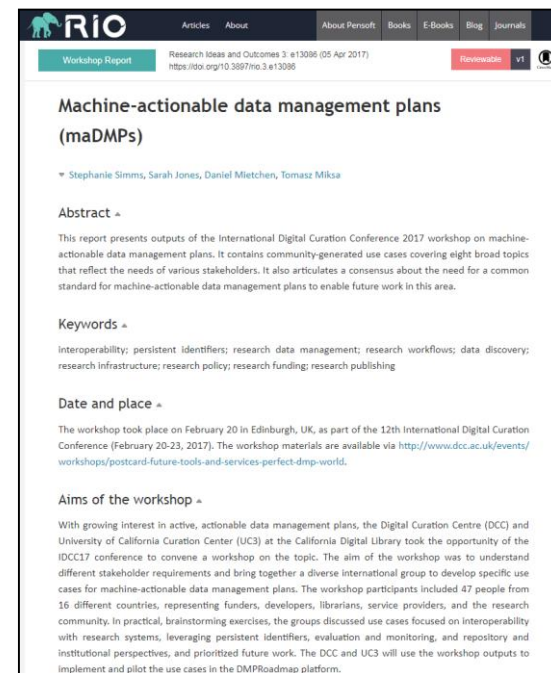
What is RDA

- Research Data Alliance
 - community-driven organization
 - 6,000 members from 130 countries
 - different stakeholders
- Plenary meetings
- Interest Groups (IGs)
 - Active DMPs
- Working Groups (WGs)
 - DMP Common Standards



<https://rd-alliance.org/>

- 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



<https://doi.org/10.3897/rio.3.e13086>

DMP Common Standards WG

Home » Working And Interest Groups » Working Group » DMP Common Standards WG

WG

Group details

Status: Recognised & Endorsed
Secretariat Liaison: Lynn Yarmey
TAB Liaison: Wenbo Chu

WG Getting started (~0-6 months after RDA endorsement)

History

The need for establishing this working group was articulated during the 9th plenary meeting in Barcelona during the Active DMPs IG session. The discussion was framed by a white paper by Simms et al. on machine-actionable data management plans (DMPs). The white paper is based on outputs from the IDCC workshop held in Edinburgh in 2017 that gathered almost 50 participants from Africa, America, Australia, and Europe. It describes eight community use cases which articulate consensus about the need for a common standard for machine-actionable DMPs (where machine actionable is defined as "information that is structured in a consistent way so that machines, or computers, can be programmed against the structure")

The specific focus of this working group is on developing common information model and specifying access mechanisms that make DMPs machine-actionable. The outputs of this working group will help in making systems interoperable and will allow for automatic exchange, integration, and validation of information provided in DMPs, for example, by checking whether a provided PID links to an existing dataset, if hashes of files match to their provenance traces, or whether a license was specified. The common information models are NOT intended to be prescriptive templates or questionnaires, but to provide re-usable ways of representing machine-actionable information on themes covered by DMPs.

The vision that this working group will work to realise is one where DMPs are developed and maintained in such a way that they are fully integrated into the systems and workflows of the wider research data management environment. To achieve this vision **we will develop a common data model with a core set of elements**. Its modular design will allow customisations and extensions using existing standards and vocabularies to follow best practices developed in various research communities. We will **provide reference implementations of the data model using popular formats**, such as JSON, XML, RDF, etc. This will enable tools and systems involved in processing research data to read and write information to/from DMPs. For example, a workflow engine can add provenance information to the DMP, a file format characterization tool can supplement it with identified file formats, and a repository system can automatically pick suitable content types for submission and later automatically identify applicable preservation strategies.

DMP Common Standards WG

Status: Recognised & Endorsed
Chair(s):
Group Email:
Secretariat Liaison: Lynn Yarmey
TAB Liaison: Wenbo Chu

Public - accessible to all site users

Join Group

Index

Add new content

Click here to create a wiki index for this group.
Group Mailing list Archive

Case Statement

RDA WG DMP Common Standards Case Statement
Comments 2

Outputs & Recommendations

Create new Outcomes

https://www.rd-alliance.org/groups/dmp-common-standards-wg

SCOPING MADMPS

Scoping maDMPs by DMP

Common Standards WG

- 1st consultation
- 2nd consultation
- Proof of concept tools
- BPMN processes
- Model development

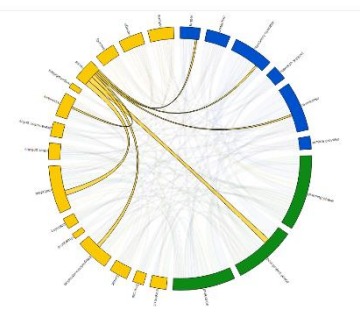
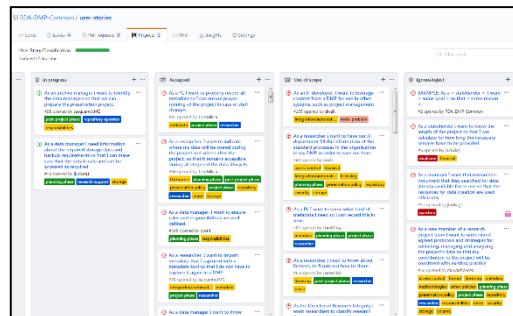
1st consultation – user stories

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**.*



2nd consultation – existing models

- 2nd consultation goes deep
 - how do we model specific requirements
 - which specific fields are needed?
 - which models exist?



(Meta-) Data

Overview

This documents 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

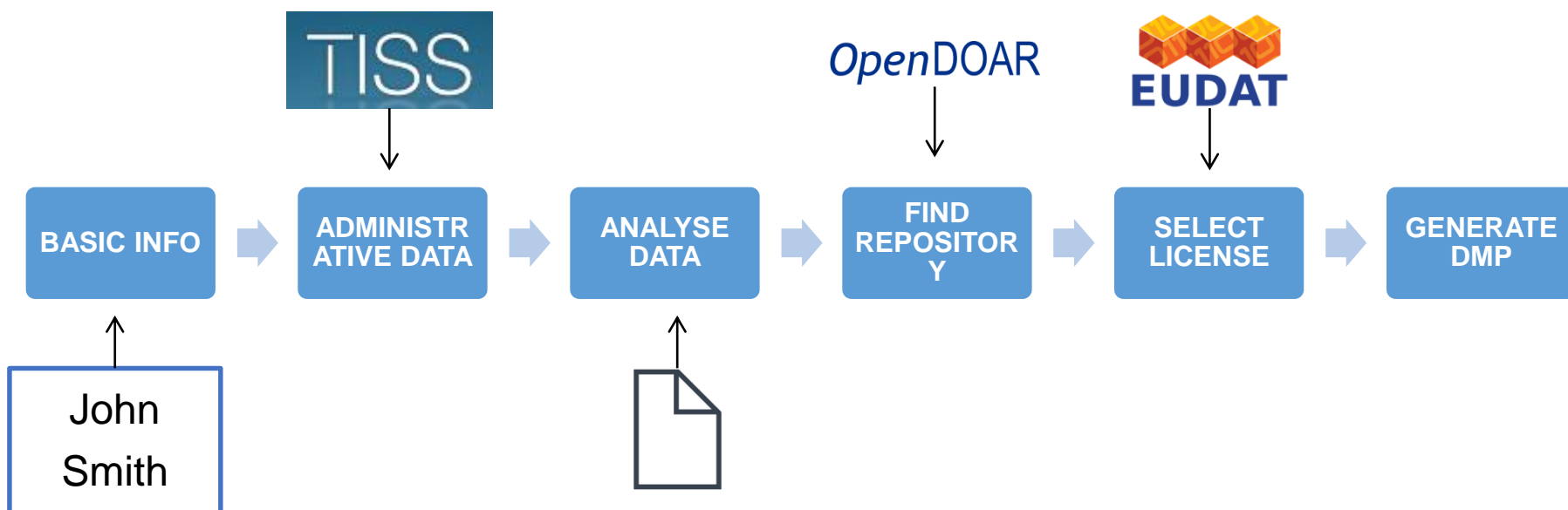
<https://joinup.ec.europa.eu/release/statdcat-ap-v100>
<https://joinup.ec.europa.eu/release/dcat-ap-v11>

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/tu-dpue-lab2-ss18>
 - https://github.com/TomMiksa/DigitalPreservation_2
 - <https://github.com/TomMiksa/digitalpreservation-dmp-generator>
 - <https://github.com/TomMiksa/DMPPlanner>
- Example of a landing page for maDMPs
 - <https://oblassadors.github.io/fair-data-science/>
 - <https://github.com/oblassadors/fair-data-science>

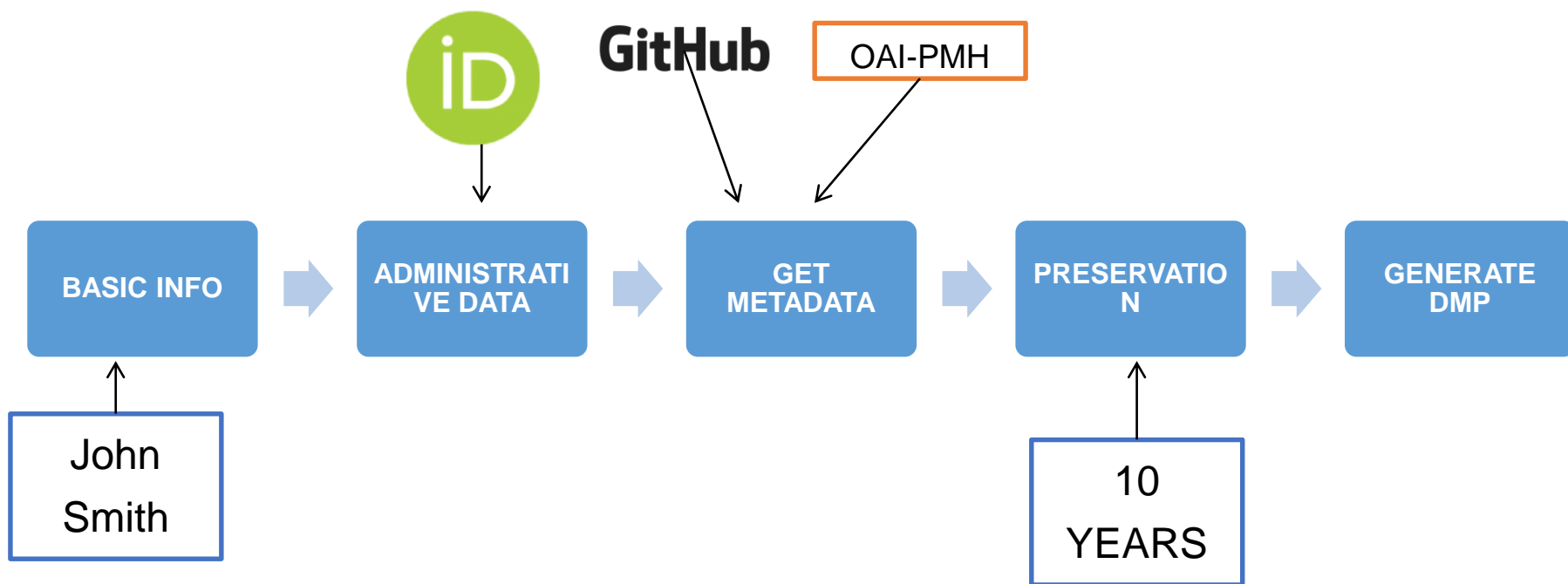
Planning phase

- 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.

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.*

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](https://orcid.org/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.

Machine-actionable DMPs: What can we automate?

DMPTool Blog
Guidance & resources for your data management plan




Search

About the DMPTool
Go to the Tool
DMPTool Webinars
Presentations and Press
Maintenance

Posted on August 20, 2018 by stephaniesimms

[← Previous](#) [Next →](#)

Machine-actionable DMPs: What can we automate?



— From Flickr by Alan Levine CCo <https://www.flickr.com/photos/cogdog/30766387175>

Following on some initial thoughts about [Scoping Machine-Actionable DMPs \(maDMPs\)](#), we're keen to dive into the substance. There are plenty of research questions we plan to explore here and over the course of our maDMP prototyping efforts. Let's begin with these:

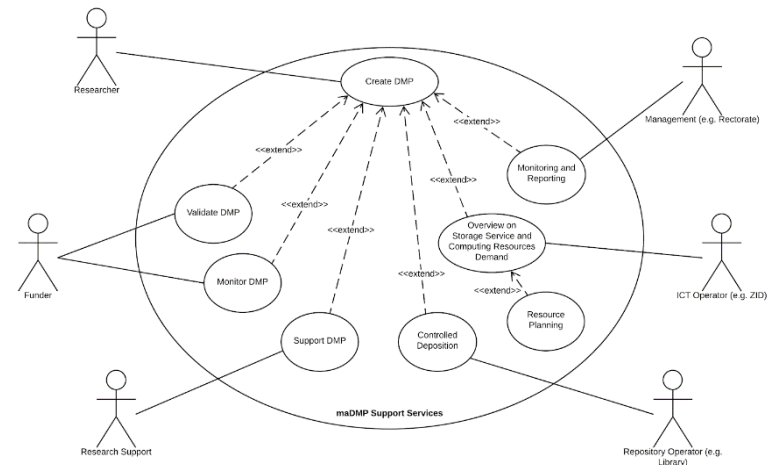
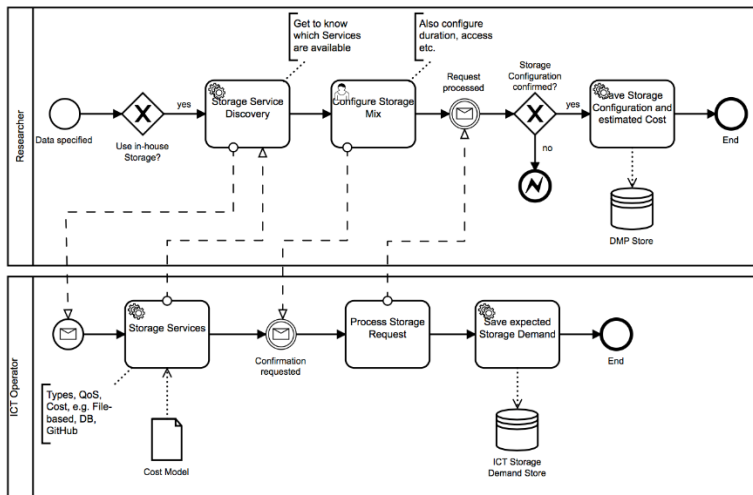
What can we automate?
What needs to be entered manually?

One of the major goals for maDMPs is to automate the creation and maintenance of some pieces of information.

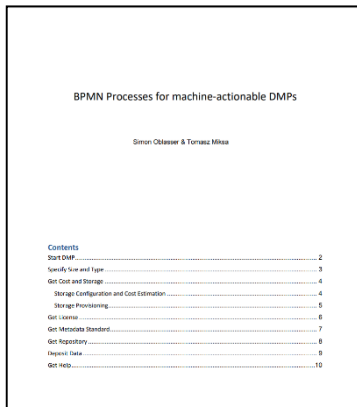
Automation stands to alleviate administrative burdens and improve the quality of information contained in a DMP.

<https://blog.dmptool.org/2018/08/20/machine-actionable-dmps-what-can-we-automate/>

- 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



- Useful in deploying maDMPs
- Allow us to narrow down focus
 - 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*



BPMN Processes for machine-actionable DMPs	
Simon Ottaviani & Torsten Mitzel	
Contents	
Start DMP	2
Specify Use and Type	3
Get Cost and Storage	4
Storage Configuration and Cost Estimation	4
Storage Provisioning	5
Get License	6
Get Metadata Standard	7
Get Repository	8
Request Data	9
Get Help	10

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

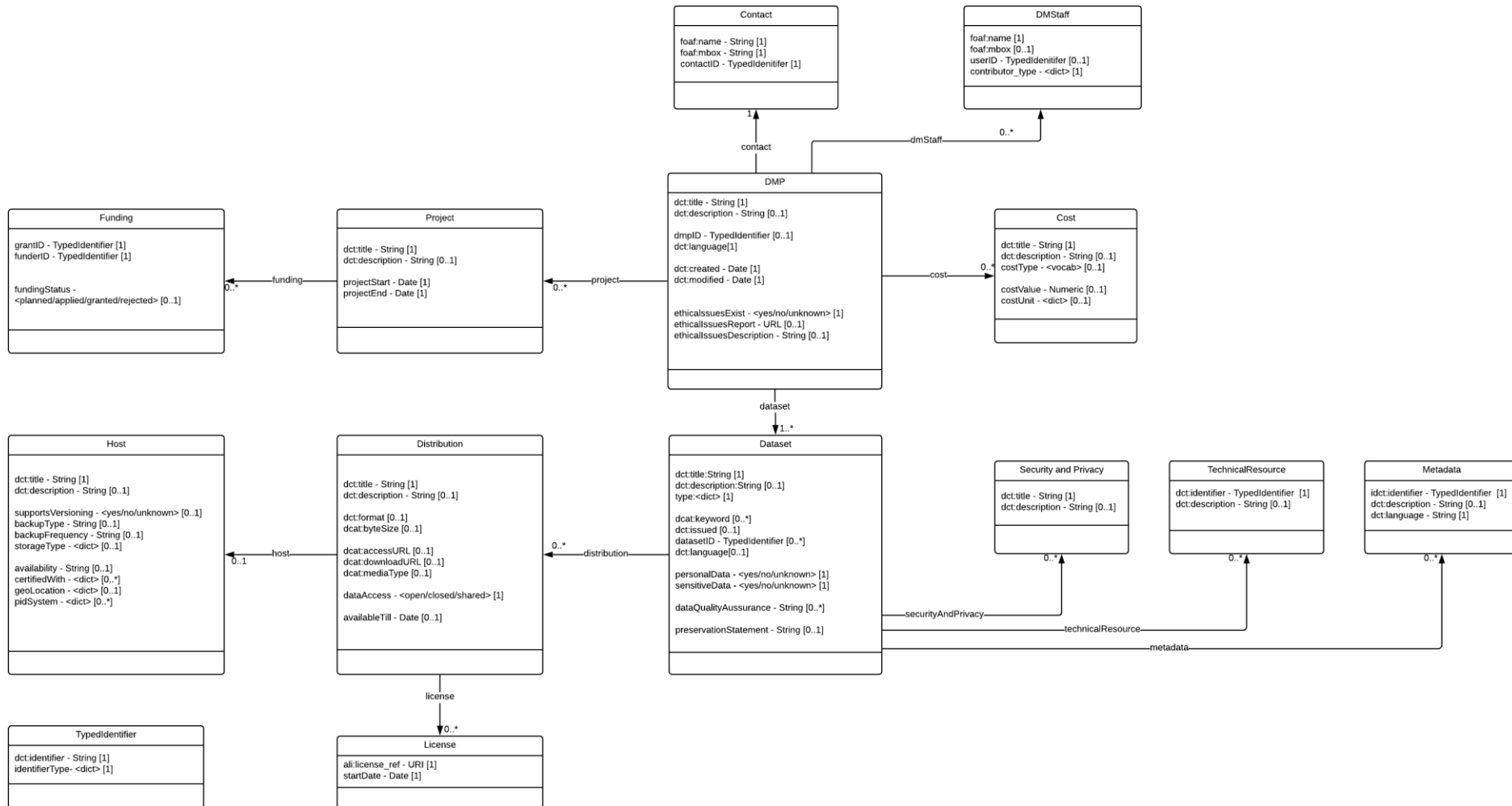
Scoping maDMPs - summary

- 1st consultation (user stories) went broad
 - to define scope of maDMPs
- 2nd 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

Part 2

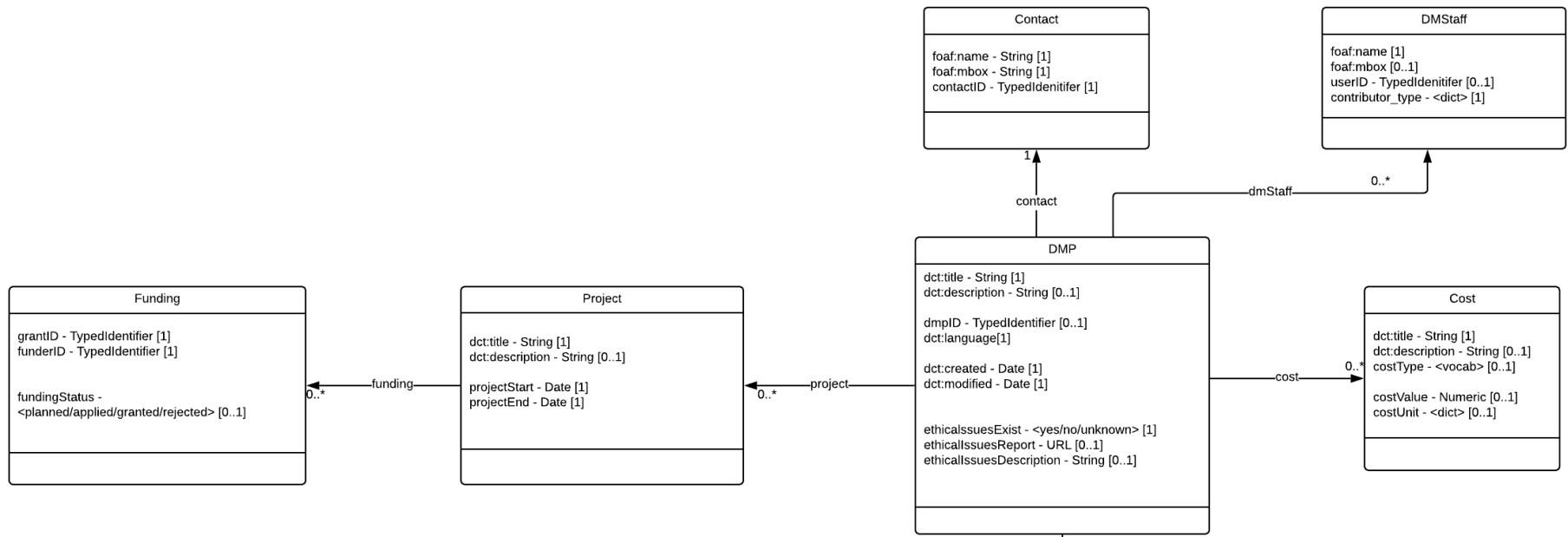
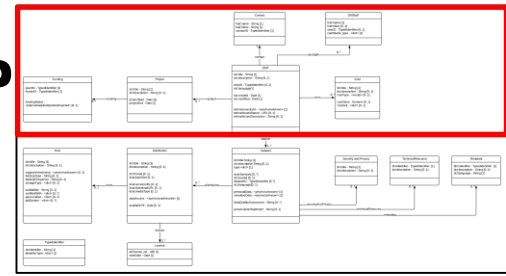
COMMON MODEL FOR MADMPs

Common model for maDMPs



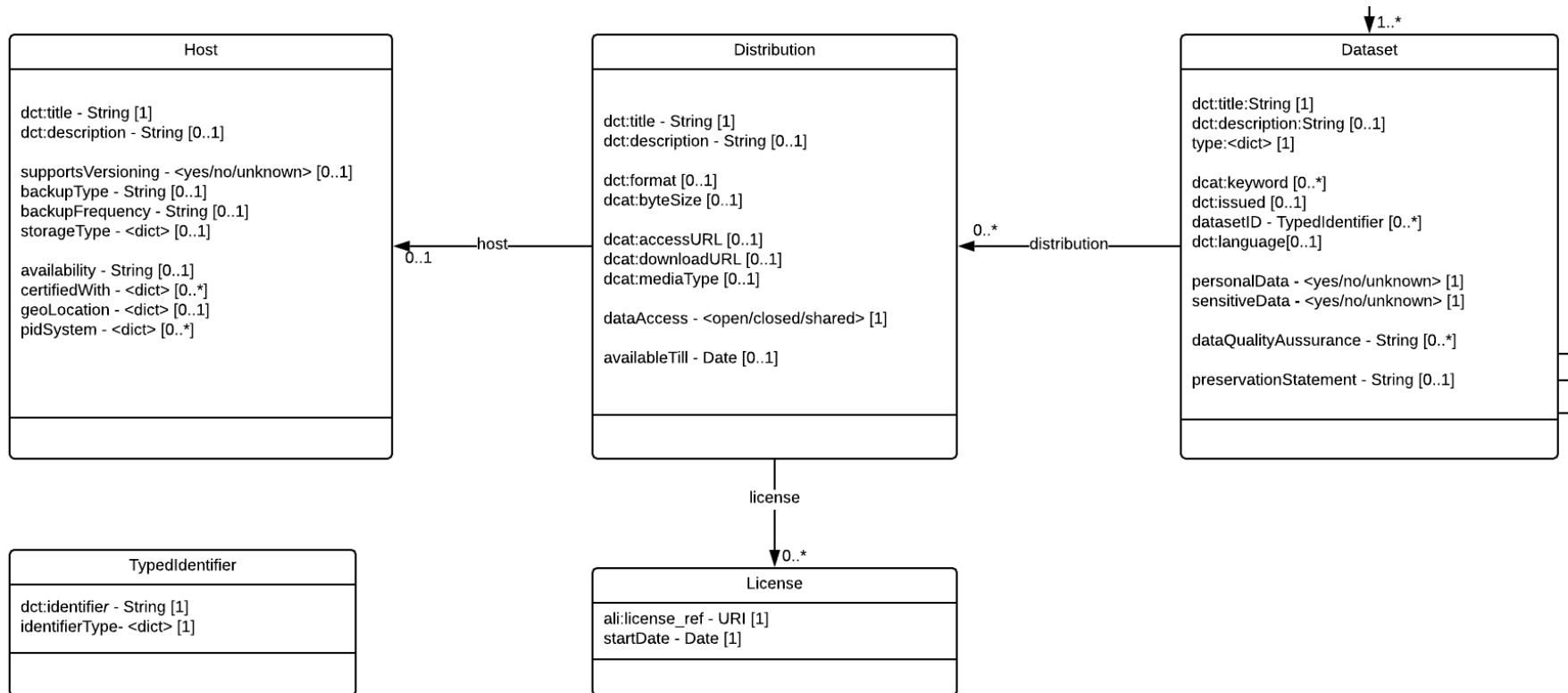
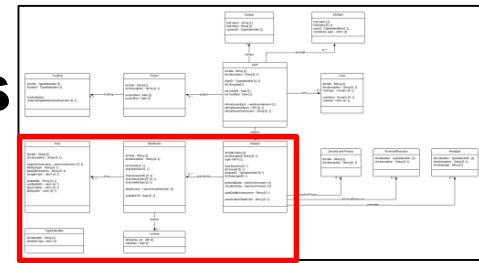
<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/diagrams/RDA-DMP-Common-Model-Diagram-190325.pdf>

Common model for maDMP



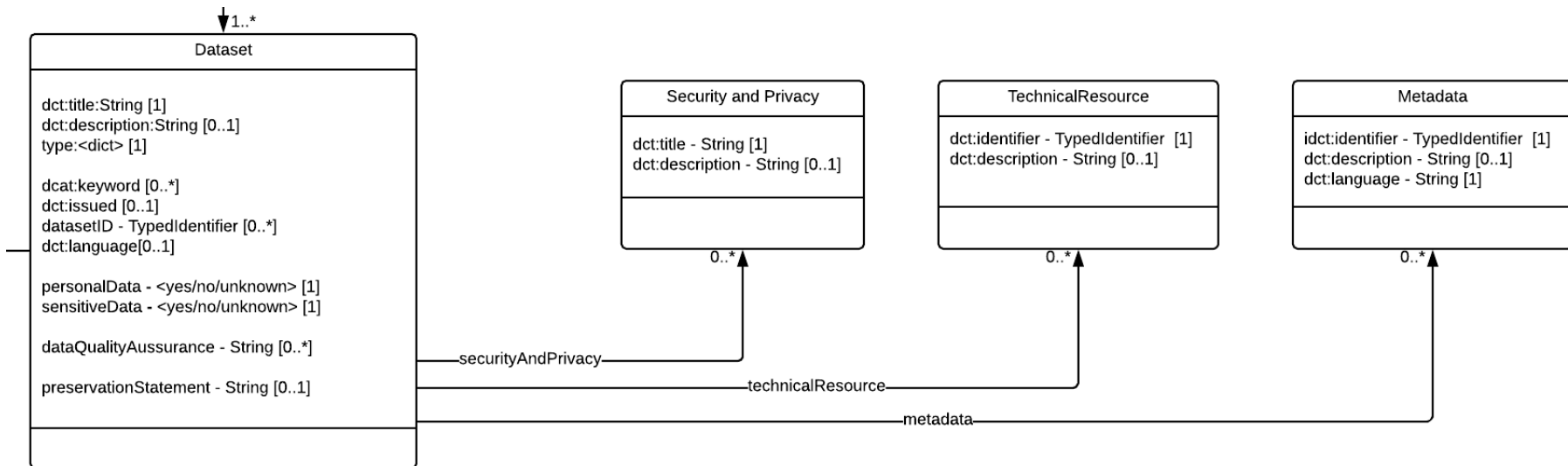
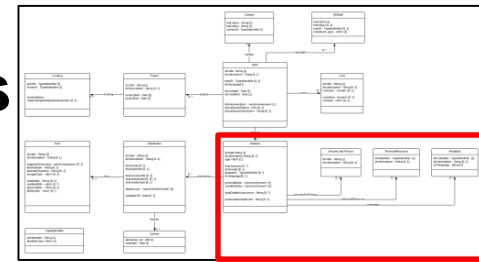
<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/blob/master/docs/diagrams/RDA-DMP-Common-Model-Diagram-190325.pdf>

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	http://orcid.org/0000-0000-0000-0000
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:

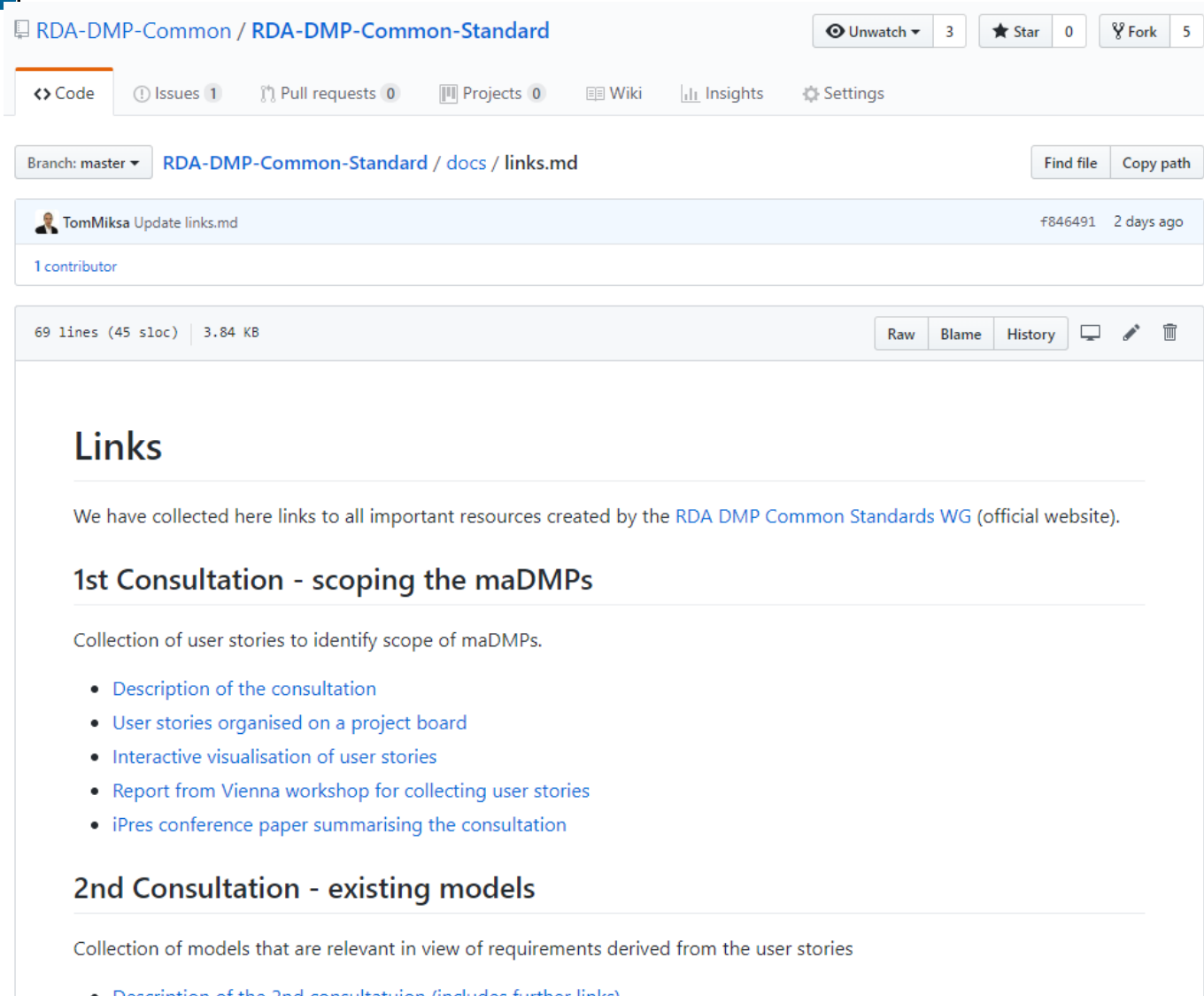
- [When to use the model?](#)
- [Do I need to populate all fields?](#)
- [What is the granularity of a Dataset?](#)
- [What is a difference between Dataset and a Distribution?](#)
- [How versioning works?](#)
- [How to express something is planned?](#)
- [How to indicate actions that were performed?](#)
- [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



The screenshot shows a GitHub repository page for `RDA-DMP-Common / RDA-DMP-Common-Standard`. The repository has 3 watchers, 0 stars, and 5 forks. The file `docs / links.md` is selected, showing a commit by TomMiksa 2 days ago. The file content is as follows:

69 lines (45 sloc) | 3.84 KB

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

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": "50110002428",
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>

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>

Model – reused standards

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

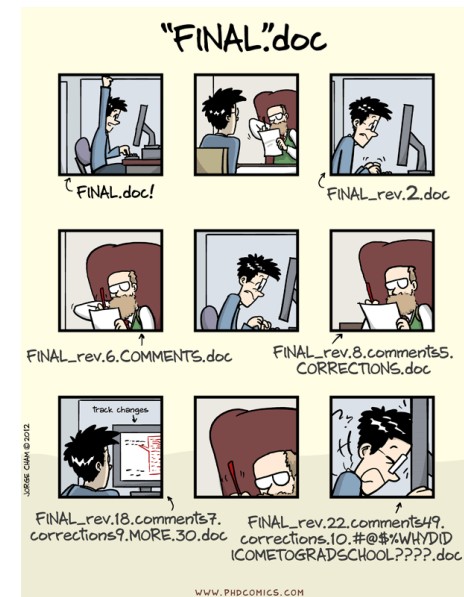
Model assumptions – relaxed constraints

- 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



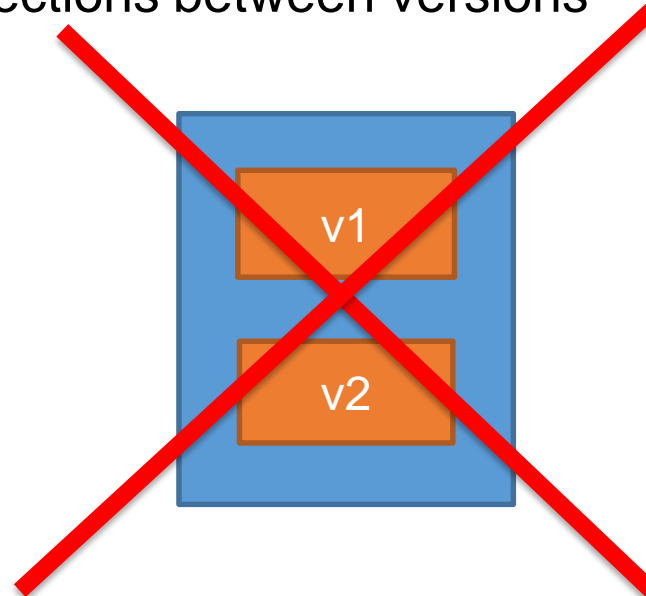
Model assumptions - interoperability

- 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"  
      }  
    }  
  }  
}
```



TECHNISCHE
UNIVERSITÄT
WIEN

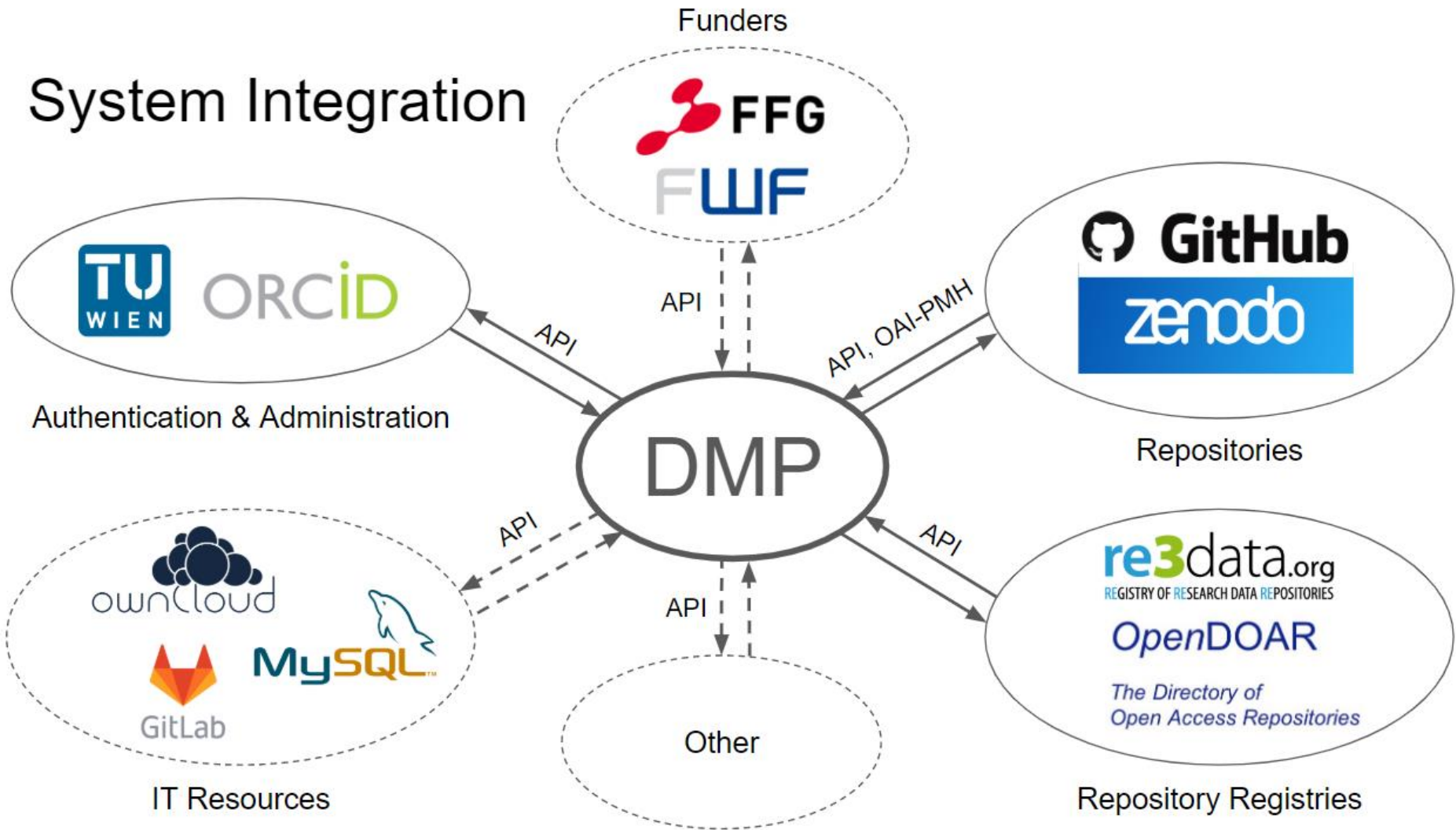
Pilot project

RDM INFRASTRUCTURE

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



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://oblassadors.github.io/dmap-mockups/>

Mock-up of a funder view for maDMP

DMP Funder View

[←](#) [→](#) [↻](#)

DMP Funder View

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

Reuse of pre-existing data

Dataset title	Origin	License
Calculating Thermal Bremsstrahlung Emission from Stellar Winds	doi:10.5281/zenodo.1476587	MIT
Occurrence records download on 2018-11-05	doi:10.26197/5be00801ec357	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

Lorem ipsum dolor sit amet, consectetur adipiscing 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
GitHub	Source code for client application	open	none	none	none	no	yes	U.S.	other
Zenodo	Supplementary material	open	DOI	ORCID	none	yes	yes	EU	REST OAI-PMH
GESIS Data Archive	Raw data Processed data	open	DOI	none	CoreTrustSeal	no	-	Germany	OAI-PMH

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	Apache License 2	2020-01-01
Processed data	publish	Creative Commons Attribute (CC-BY)	2021-03-01

Please click on the scrollbar to see more.

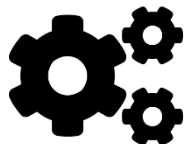


10 PRINCIPLES FOR MADMPS

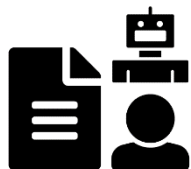
10 principles for maDMPs



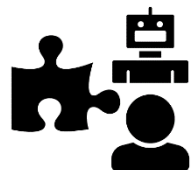
1 Integrate DMPs with the workflows of all stakeholders in the research data ecosystem



2 Allow automated systems to act on behalf of stakeholders



3 Make policies (also) for machines, not just for people



4 Describe—for both machines and humans—the components of the data management ecosystem



5 Use PIDs and controlled vocabularies

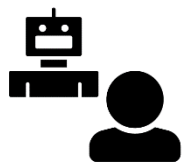
Miksa T, Simms S, Mietchen D, Jones S (2019) Ten principles for machine-actionable data management plans. PLoS Comput Biol 15(3): e1006750.

<https://doi.org/10.1371/journal.pcbi.1006750>

10 principles for maDMPs



6 Follow a common data model for maDMPs



7 Make DMPs available for human and machine consumption



8 Support data management evaluation and monitoring



9 Make DMPs updatable, living, versioned documents



10 Make DMPs publicly available

Miksa T, Simms S, Mietchen D, Jones S (2019) Ten principles for machine-actionable data management plans. PLoS Comput Biol 15(3): e1006750.

<https://doi.org/10.1371/journal.pcbi.1006750>

You should know

- Why traditional DMPs are not enough
- What is required to make DMPs machine-actionable
- How common model for maDMPs works
- 10 principles for maDMPs

- 6.5: Machine-actionable DMPs
- 13.5: Repositories and Certification
- 20.5: Institution-wide Data Management
- 27.5: Reproducibility
- 3.6: Exercise – mid-term presentation

Funding opportunity



- FEMTech Funding opportunity
 - Funding by FFG (Austrian Research Funding Agency)
 - Target group: **female students**
 - Duration: 2-6 months, **paid**
 - Can be used to work on a thesis or project (Praktikum)

- More info: <https://www.ffg.at/femtech-praktika>

- Interested?
 - ➔ Contact me (tmiksa@sba-research.org)
 - ➔ Deadline 13.05.2019
 - ➔ Also possible later: October 2019, and again in 2020