



RDA DMP Common Standards

Tomasz Miksa, Paul Walk, Peter Neish

Agenda

- Part 1 – Introduction
- Part 2 – Consultations
 - Overview of two consultations performed
- Part 3 - Tools for machine-actionable DMPs
 - Examples of what can be automated
- Part 4 – Use cases and processes to include all stakeholders
 - How to define landscape of maDMPs
 - Mock-ups
- Part 5 – Towards Common Data Model
 - Architecture and examples
- Part 6 – Wrap-up and next steps

Data Management Plans (DMPs)

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



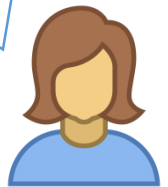
Data Management Plans



How to discover these tools?
Which one do I need to use?
Why do I have to provide the same
information again?

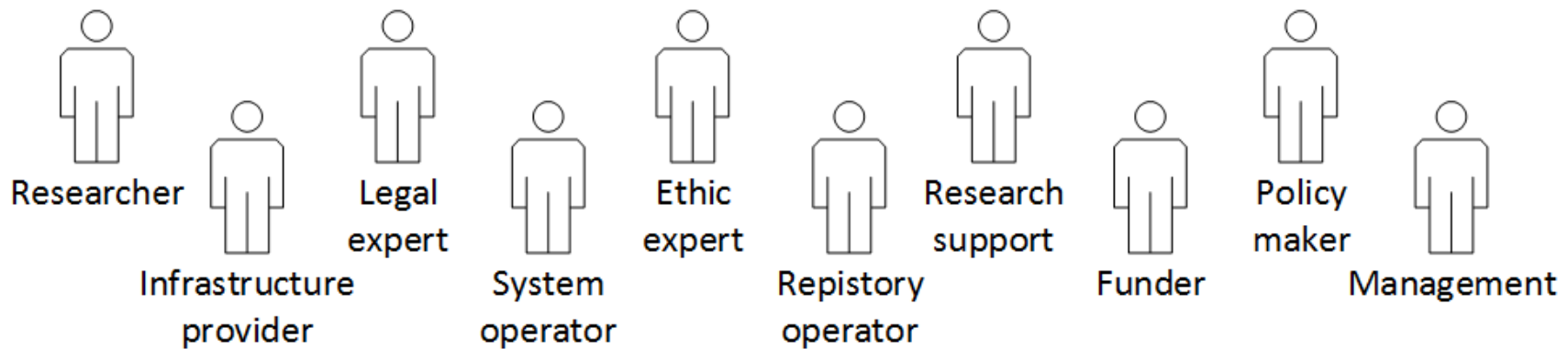


Why haven't they consulted us before?
Who is going to pay for this?
We don't have enough people for that!

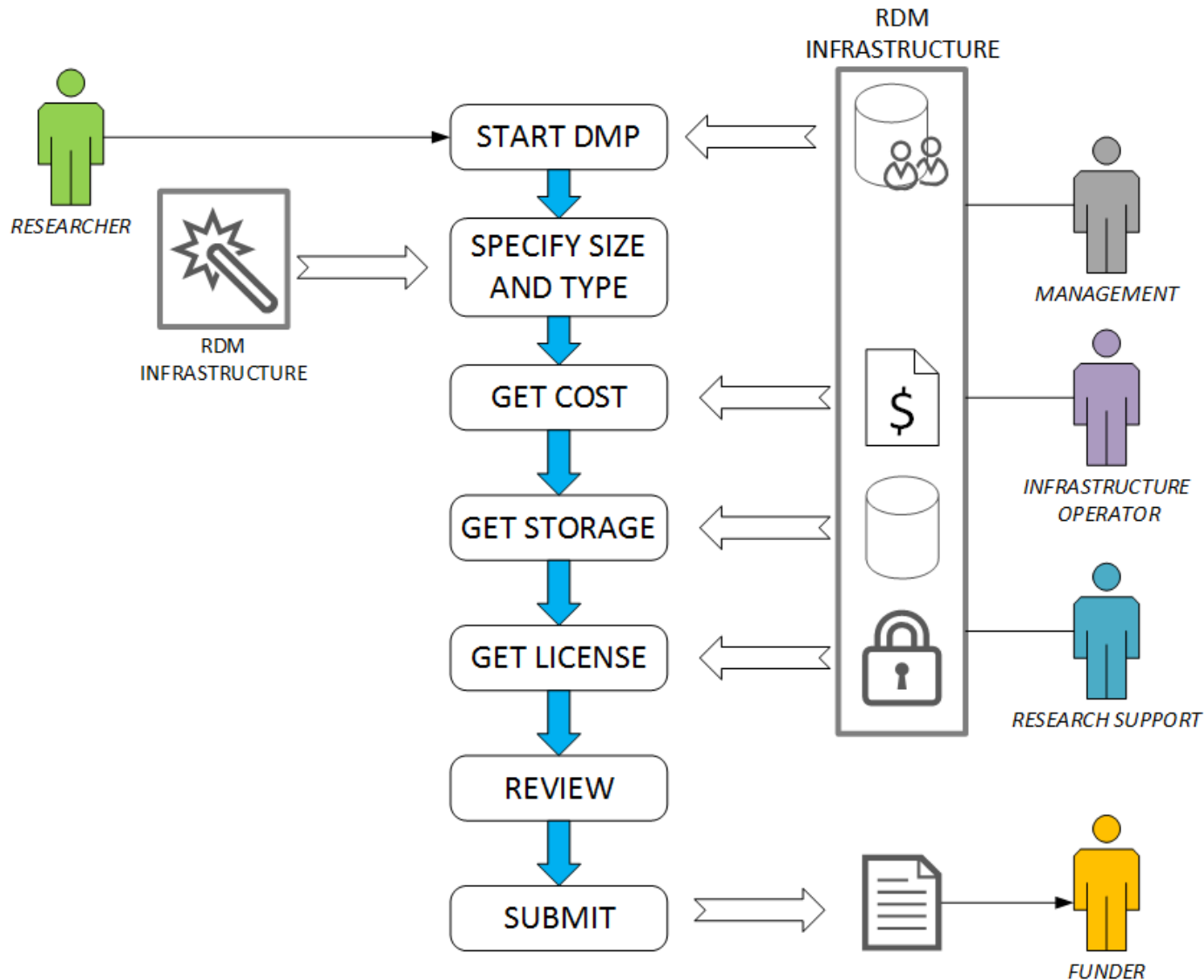


Research data lifecycle

- 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

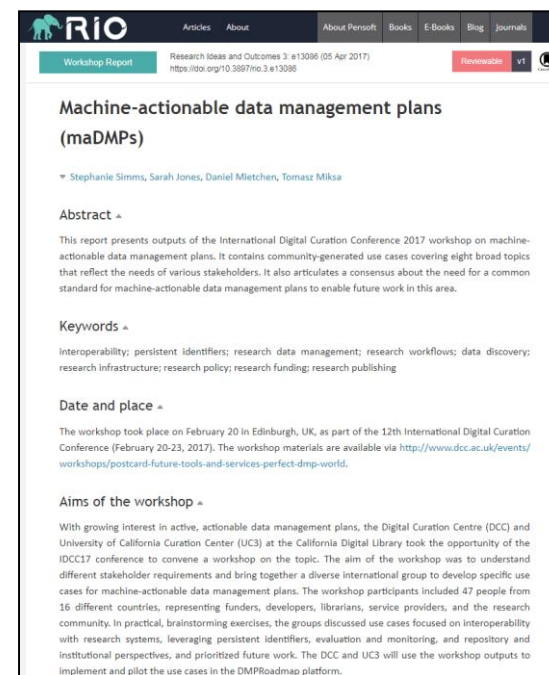


Automated Data Management Workflow



Why do we need this WG?

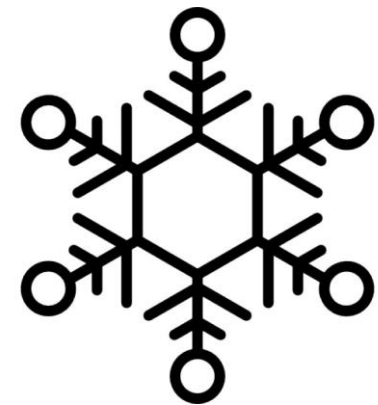
- Shortcomings of existing DMPs
 - manually completed, vague, not updated, ...
- Machine-actionable DMPs
 - living documents
 - automate data management
 - collect information from systems
 - trigger actions in systems
 - facilitate validation
- This requires
 - well-defined RDM workflows
 - data management infrastructure
 - common data model



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

DMP Common Standards - Outputs

- **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.

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

Use PIDs whenever possible, e.g. ORCID

```
"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 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>

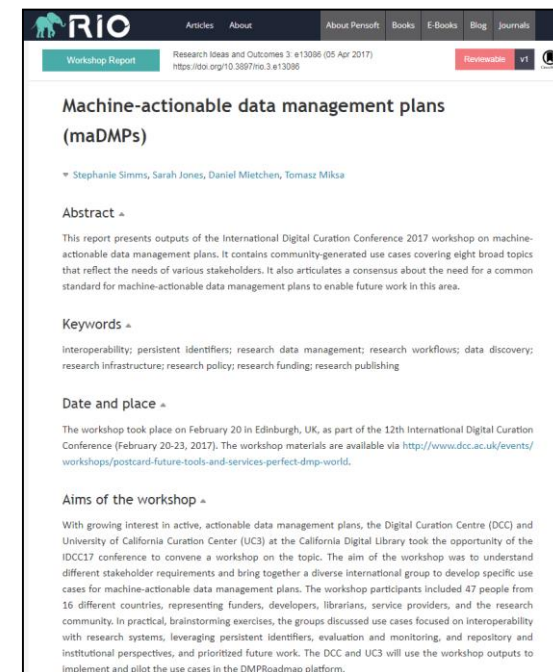
- Machine-actionable DMP

```
"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

DMP Common Standards WG

- 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

› Staying in touch

- › Regular plenary meetings
- › Workshops
- › Calls
- › Mailing list
- › GitHub

› <https://github.com/RDA-DMP-Common>



Tomasz Miksa



Paul Walk



Peter Neish

DMP Common Standards WG

[Home](#) » [Working And Interest Groups](#) » [Working Group](#) » [DMP Common Standards WG](#)

WG

i Group details

Status: Recognised & Endorsed
Chair (s): Tomasz Miksa, Paul Walk, Peter Neish
Secretariat Liaison: Lynn Yarmey
TAB Liaison: Isabelle Perseil

WG Wrapping up (from ~12 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

DMP Common Standards WG

Status: Recognised & Endorsed

Secretariat Liaison: Lynn Yarmey
TAB Liaison: Isabelle Perseil

Public - accessible to all site users

[Join Group](#)

Index	Add new content
<div> Click here to create a wiki index for this group. Group Mailing list Archive </div>	

Group sessions at RDA Plenaries

WG DMP Common Standards - RDA 13th Plenary Meeting

By Tomasz Miksa On 10, Jan 2019

Case Statement

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

Recent Activity

Type

Event
Group event
Plenary session page
Post to Group Mailinglist
Wiki Page

Surname

Search by keyword contained in title

07
MAR
2019



DMP Common Standards - Model Your Own MaDMP

By Tomasz Miksa

Dear group members,
We have created some JSON examples of maDMPs. You can find them here:
<https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard/tree/master/ex...>
They are based on the model that you have already seen (which continuously undergoes adaptations):
<https://www.lucidchart.com/invitations/accept/ee26bc71-01a6-442a-b946-5b...>

[Read more](#) [3 comments](#) [Log in or register to post comments](#)

01
MAR
2019



Re: [EXTERNAL] Re: [Dmp-Common] RDA DMP Common Standards - February 2019 Call

By Dennis Walworth

What about adopting ISO 19115 metadata domains? CI_RoleCd and MD_ScopeCd seem particularly adaptable for Roles and Dataset Types, respectively.
Dennis
On Fri, Mar 1, 2019 at 6:12 AM jacquemotmc <***@***.***> wrote:

[Read more](#) [Log in or register to post comments](#)

13
FEB
2019



RDA DMP Common Standards - February 2019 Call

By Tomasz Miksa

Dear group members,
Since our last call in January, we have introduced some changes in the model. The last call was very fruitful and we would like to organise another one next week.
Please indicate your availability:

File Repository

latest files uploaded

- Notes IDCC19 Unconference
- DMP Common Standards WG - Call January 2019
- Botswana plenary slides
- Oct18-Call: Overview of 12 months and next steps
- Berlin Plenary Slides
- 2018.03 Summary of activities and next steps
- 10th Plenary Montreal - Presentation of the WG

Latest Webconference

There are no Webconferences available yet for this group

March						
S	M	T	W	T	F	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

ical

DMP Common Standards WG members

Adil Hasan
Allen Dearth
Ana Slavec
Andre Filipe de Moraes Batista
Andreas Rauber
Andrew White
Angus Whyte
Antonio J. Sánchez-Padial
Antonio S. Cofiño
Bethania de Araujo Almeida

Consultations

Part 3 - Overview of two consultations performed

Summary of actions till now

- 1st consultation (user stories) went broad
 - to define scope of maDMPs
- 2nd consultation went deep
 - to identify models for specific requirements

1st consultation – user stories

› Goals

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

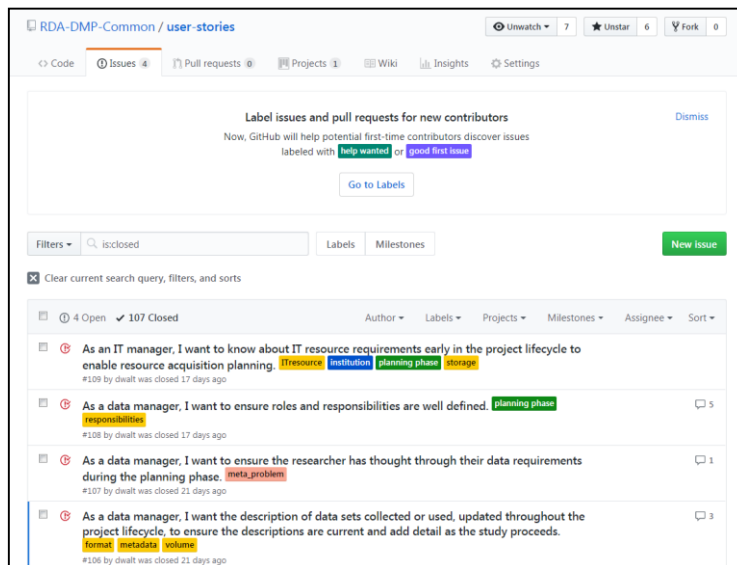
As a <stakeholder>, I want <goal> so that <reason >.

*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**.*

<https://github.com/RDA-DMP-Common/user-stories/>

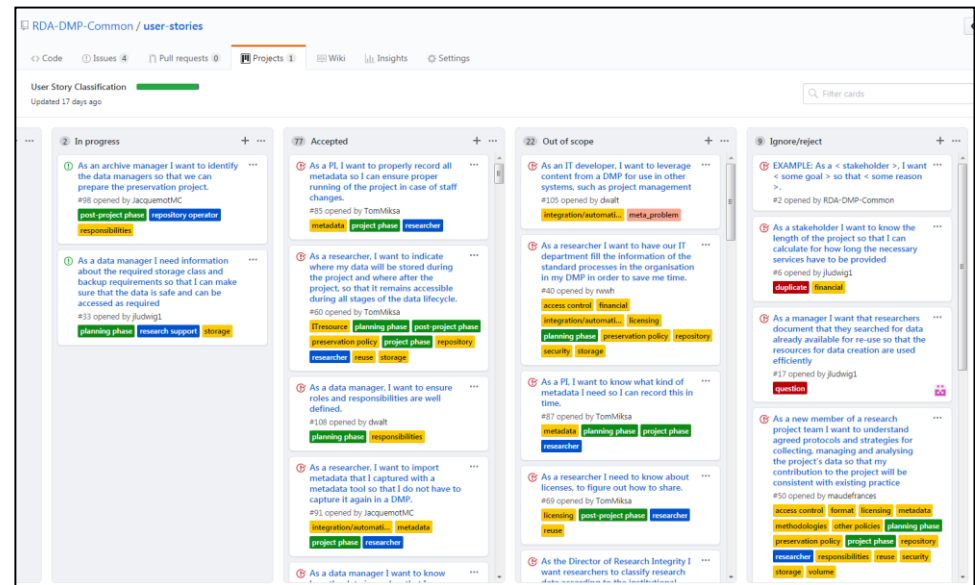
User story consultation

- <https://github.com/RDA-DMP-Common/user-stories/>
- 100+ issues defined
- inputs from Europe and Australia
- inputs from individuals and [workshops](#)

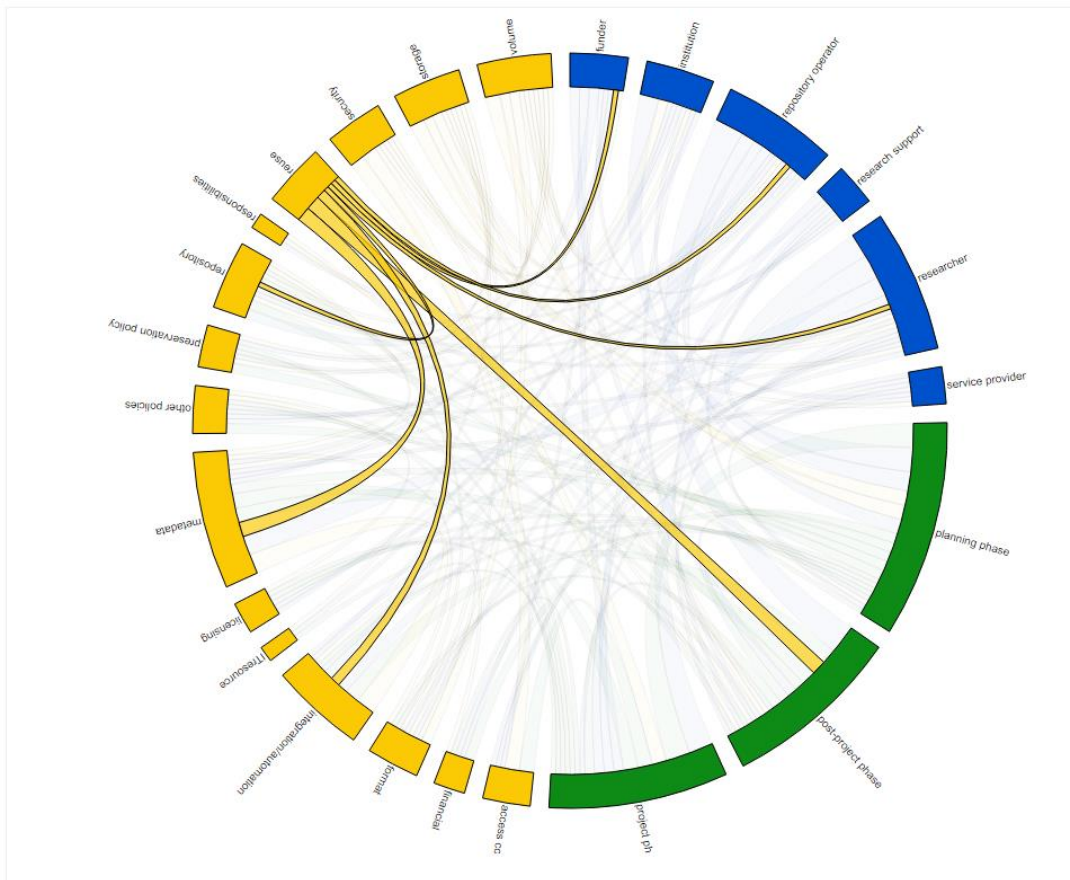


User story labelling

- <https://github.com/RDA-DMP-Common/user-stories/projects/2>
- <https://github.com/RDA-DMP-Common/user-stories/wiki>
- 3 major categories (colours)
 - stakeholders involved
 - project phase
 - subject of information conveyed
 - access control
 - volume
 - financial
 - licensing
 - metadata
 - repository
 - security
 - storage
 - etc.



User story visualisation



- <https://goo.gl/znBL3F>
- interactive visualisation - changes on GitHub are visible immediately
- shows relations between stakeholders, phases and information

Defining requirements for machine-actionable Data Management Plans

- › Defines machine-actionability
- › Describes results of user story consultation
- › <http://ifs.tuwien.ac.at/~miksa/papers/2018-iPres-maDMPs.pdf>



From user stories to requirements

- <https://docs.google.com/document/d/1sWVy0Rqj9fGsjs6GyFnBd3fH6XF2088zjK8U-1wLq4c/edit?usp=sharing>
- Refactoring of user stories
- Goal: finding overlaps, gaps, duplicates
- Example below

- Metadata
 - taxonomy/classification [14,11]
 - Links to metadata of the real data [89, 39]
 - Funder information [7]
 - Link publications to data [55]
 - Authorship [88]
 - Multilingual metadata [65]
 - Include raw metadata directly in the model [91, 85]

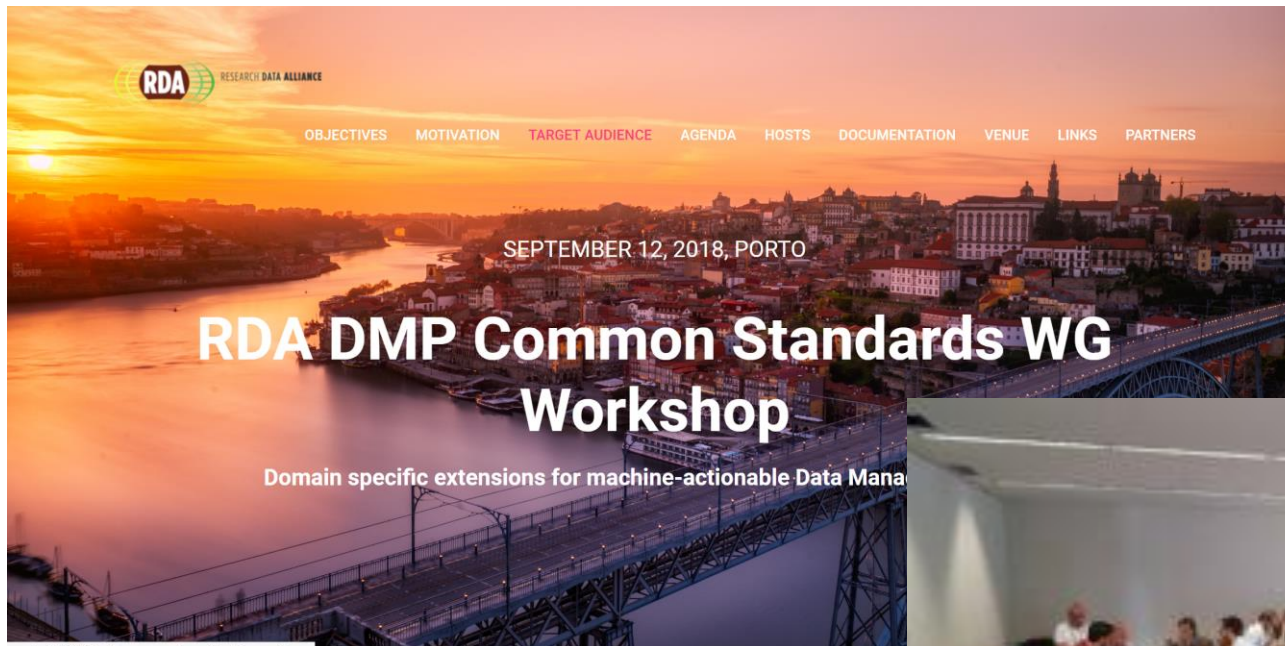
2nd consultation:

From user stories to requirements

- https://docs.google.com/document/d/1mMJqmvqEAkbEWbdV7rtFU9hiQMOuH0ESn4Up_TDn1Es/edit?usp=sharing
- 5 documents to collect requirements, models, specific fields, etc.
 - [Administrative, Roles and Responsibilities](#)
 - [Data](#)
 - [Infrastructure](#)
 - [Security, Privacy and Access Control](#)
 - [Policies, legal and ethical aspects](#)

RDA DMP Common Standards WG Workshop

➤ <http://rda-ws-tpdl2018.idsswh.sysresearch.org/#>



2nd consultation – feedback collected

(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.

<https://docs.google.com/document/d/1GRBxgOKf5VGfJ9YGzcQqID2qn6V5PKcwNUIAYGsJwj0/edit?usp=sharing>

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>

Consultations summarised

- 1st consultation (user stories) went broad
 - helped us defined the scope of the maDMPs
 - what information should a maDMP contain?
 - who provides and uses this information?
- 2nd consultation goes deep (ongoing)
 - how do we model specific requirements
 - which specific fields are needed?
 - which models exist?

Tools for maDMPs

Part 3 - Examples of what we can automate with maDMPs

Prototypes developed by TU Wien students

› Requirements

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

› Tools available as Docker containers on GitHub

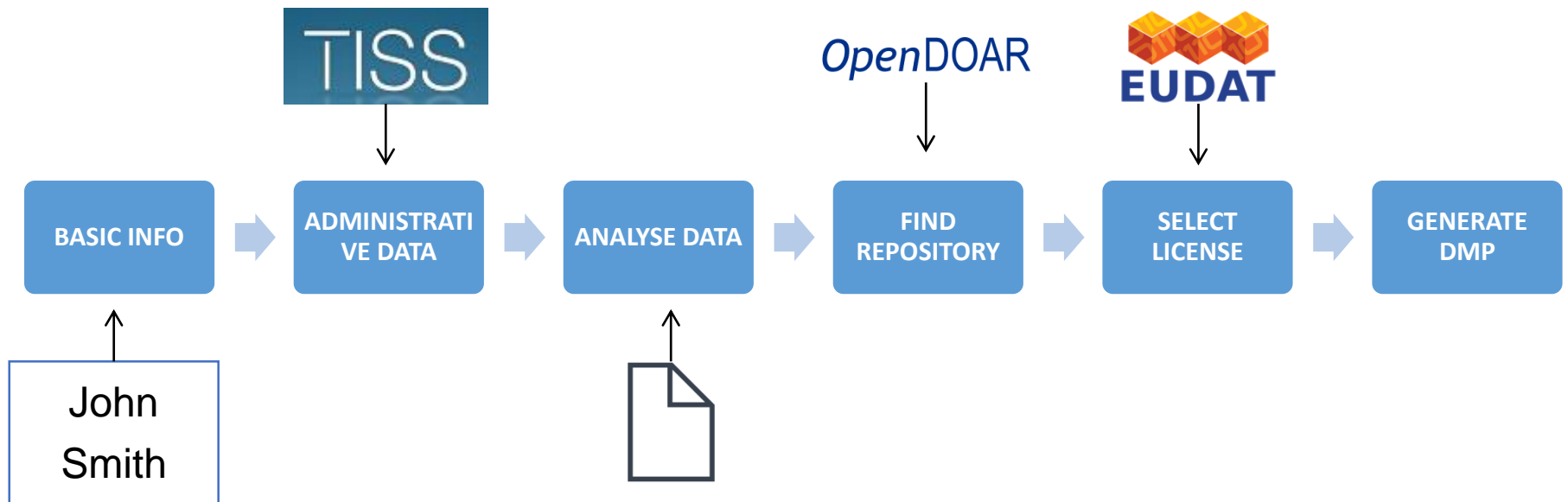
- › <https://github.com/TomMiksa/DMPGenerator>
- › https://github.com/TomMiksa/digital_preservation_ex_1_2
- › <https://github.com/TomMiksa/tu-dpue-lab2-ss18>
- › https://github.com/TomMiksa/DigitalPreservation_2
- › <https://github.com/TomMiksa/digitalpreservation-dmp-generator>
- › <https://github.com/TomMiksa/DMPlanner>

› Example of a landing page for maDMPs

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

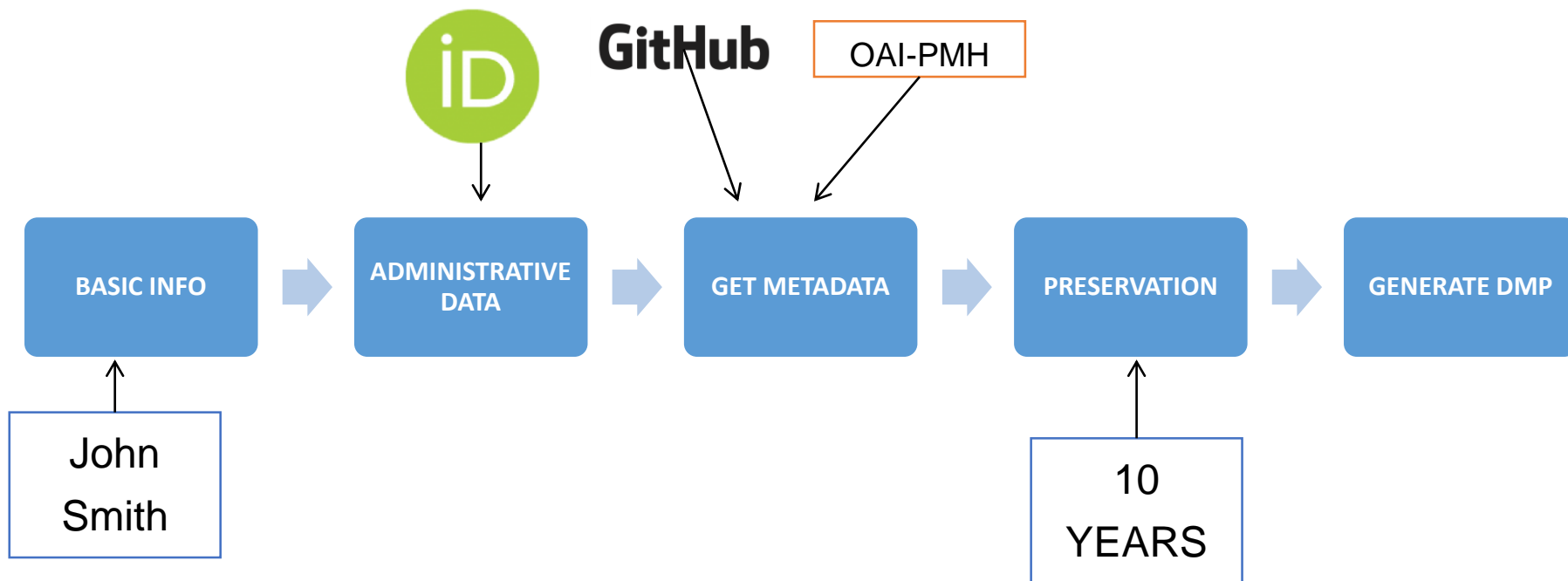
Planning phase

- 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



Planning phase - demo

<https://github.com/IrinaAvram/DMPGenerator>



Basic Information

Project Name*

TUW Report

Projektass.
Dr.techn.
Mag.

First Name

Tomasz

Last Name*

Miksa

Contact

Email: tomasz.miksa@tuwien.ac.at

Position

Projektassistent at Forschungsbereich Information und Software Engineering

NEXT



1. BASIC INFORMATION





2. DATA ANALYSIS



3. REPOSITORY AND LICENSING

Upload Sample files

Choose Files 2018-TenRulesMADMPS.pdf

Name	Mime Type	Size	Input/Output	Nr. of Files	
Fig1.tif	image/tiff	1009010 Bytes	input ▼	100	
2018-TenRulesMADMPS.pdf	application/pdf	178186 Bytes	output ▼	500	

Required Storage space

Input: 100901000 Bytes

Total: 189994000 Bytes

Output: 89093000 Bytes

COMPUTE REQUIREMENTS

Repositories

Architektur-Informatik <http://architektur-informatik.scix.net/cgi-bin/works/Home>



CumInCAD Digital Archive <http://cumincad.architecturez.net/>



Digitale Landesbibliothek Oberösterreichische <http://digi.landesbibliothek.at/>



Project and post-project phase - demo 1

<https://github.com/mdietrichstein/digitalpreservation-dmp>



Connecting Research
and Researchers

EDIT YOUR RECORD

ABOUT ORCID

CONTACT US

HELP

5,529,789 ORCID iDs and counting. [See more...](#)

Marc Dietrichstein

ORCID ID

 <https://orcid.org/0000-0003-4890-3498>

 Print view?

Country

Austria



Email

e0327606@student.tuwien.ac.at



Works (1 of 1)

Sort

Correlating Alcohol Consumption and UFO Sightings in the USA



TU Wien

2018-03-29 | supervised-student-publication

DOI: [10.5281/zenodo.1209833](https://doi.org/10.5281/zenodo.1209833)

Source: Marc Dietrichstein

 Preferred source

 Help

DMPGen - Mozilla Firefox

workflows x DMPGen x +

DATA IMPORT PRESERVATION DMP

Import Data

Requirements:

- Author has to be registered on orcid.org
- Profile information on orcid has to be **public**
- The respective project has to be listed first in the profile's *Works* section
- The project has to be hosted on zenodo.org
- The project must have a link to a Github repository

Search for an author and click on the corresponding orcid id to import data relevant to the DMP

0000-0003-4890-3498

0000-0001-8985-2979

0000-0002-0923-3084

Find valid Orcid Profile to continue

Search for an author and click on the corresponding orcid id to import data relevant to the DMP

ORCID Author Info

Orcid Id	0000-0003-4890-3498
Given Name	Marc
Family Name	Dietrichstein
E-mail	e0327606@student.tuwien.ac.at
Project Title	Correlating Alcohol Consumption and UFO Sightings in the USA
Project DOI	https://doi.org/10.5281/zenodo.1209833
Publication Date	2018/03/29

ORCID Author Info

Orcid Id	0000-0003-4890-3498
Given Name	Marc
Family Name	Dietrichstein
E-mail	e0327606@student.tuwien.ac.at
Project Title	Correlating Alcohol Consumption and UFO Sightings in the USA
Project DOI	https://doi.org/10.5281/zenodo.1209833
Publication Date	2018/03/29

Zenodo Project Info

Creators	Marc Dietrichstein, sorx
Rights	openAccess
Type	software
Github Url	https://github.com/mdietrichstein/digitalpreservation-dmp/tree/1.0.0

For each file below, select it's role in the context of preservation and the preservation duration if applicable

View DMP

Filename	Path	Tag	Preservation
.gitignore	.gitignore	Ignore ▾	Ignore ▾
Dockerfile	Dockerfile	Ignore ▾	Ignore ▾
LICENSE	LICENSE	Ignore ▾	Ignore ▾
README.md	README.md	Ignore ▾	Ignore ▾
README.pdf	README.pdf	Ignore ▾	Ignore ▾
ufo_alcohol.csv	data/processed/ufo_alcohol.csv	Ignore ▾	Ignore ▾
DP_LIVE_22032018202902423.csv	data/raw/DP_LIVE_22032018202902423.csv	Input Data ▾	20 Years ▾
ufo-scrubbed-geocoded-time-standardized.csv	data/raw/ufo-scrubbed-geocoded-time-standardized.csv	Input Data ▾	20 Years ▾
architecture.png	documentation/architecture.png	Documentation ▾	5 Years ▾
description.txt	documentation/description.txt	Documentation ▾	10 Years ▾
metadata.xml	documentation/metadata.xml	Documentation ▾	10 Years ▾
01_data-preprocessing.ipynb	notebooks/01_data-preprocessing.ipynb	Software ▾	5 Years ▾
02_visualization.ipynb	notebooks/02_visualization.ipynb	Software ▾	5 Years ▾
.keep	reports/.keep	Ignore ▾	Ignore ▾

DMPGen - Mozilla Firefox
arkflow x DMPGen x +

DATA IMPORT
PRESERVATION
DMP

Correlating Alcohol Consumption and UFO Sightings in the USA

Authors

Marc Dietrichstein

- Orcid Id 0000-0003-4890-3498
- e0327606@student.tuwien.ac.at

Document Version and Date

29.03.2018

Gathered Data

DP_LIVE_22032018202902423.csv	size: 114328 b	preserve: 20 years
checksum: b37dbbdf8927a3670a104d51e71bbdd67b7870c		
ufo-scrubbed-geocoded-time-standardized.csv	size: 13929415 b	preserve: 20 years
checksum: 5c6b7af9c458ffdb38301ad948a443cfa7a3719a		

Software

01_data-preprocessing.ipynb	size: 25788 b	preserve: 5 years
checksum: 194bee784a56d6030f99705c873d29d4ddee0a42		
02_visualization.ipynb	size: 108779 b	preserve: 5 years
checksum: 0f0d585f887c7b24ffd079acac225064a0621cfd		

Documentation

architecture.png	size: 87594 b	preserve: 5 years
checksum: ba6f8585ff107d55d42e00c127ffcb41f448d5ef		
description.txt	size: 341 b	preserve: 10 years
checksum: 069b7f8024120952a09c6f482cebbdde9505719d		
metadata.xml	size: 997 b	preserve: 10 years
checksum: 934ae1c38d721790e353a9dfdc498d3c1d5283e3		

Ethical Questions

DMPGen - Mozilla Firefox

kflow X DMPGen X +

Ethical Questions

<No information>

Licenses and Redistribution

Files are marked with their respective license. The license-information of input-files is not known.

Code Preservation

The created code will be stored on github. The repository can be found through the link given below under "Github Repository"

Data Preservation

The files that should be preserved are marked throughout the lists of files, which can be seen above. Each file states the duration that it should be preserved for. All github releases are stored on Zenodo as well.

The service provided by Zenodo is free and does not incur any costs - neither during the project nor afterwards.

Zenodo Infos:

CERN
European Organization for Nuclear Research
att: IT Department, Digital Repositories Section
1211 Geneve 23
Switzerland
<http://zenodo.org/>

Access and Security

Code and data are hosted on the given git repository on github.

Data Sharing

All code, data and documentation is available on Github, which is licensed under the MIT license. Each Github release then is published to the Zenodo repository where it also gets assigned a DOI

Github Repository

<https://github.com/mdietrichstein/digitalpreservation-dmp/tree/1.0.0>

```
{
  "@context": {
    "dmp": "http://purl.org/madmps#",
    "foaf": "http://xmlns.com/foaf/0.1/",
    "dc": "http://purl.org/dc/elements/1.1/",
    "dcterms": "http://purl.org/dc/terms/",
    "premis": "http://www.loc.gov/premis/rdf/v1#"
  },
  "@id": "http://example.org/dmps/mydmp",
  "@type": "dmp:DataManagementPlan",
  "dcterms:title": "mdietrichstein/digitalpreservation-dmp: Submission Release",
  "dcterms:description": " ☹ Exploring the connection between alcohol consumption and the number of ufo sightings in the USA ☹ ",
  "dc:creator": [
    {
      "@id": "0000-0003-4890-3498",
      "foaf:name": "Marc Dietrichstein",
      "foaf:mbox": "e0327606@student.tuwien.ac.at"
    }
  ],
  "dc:date": "29.03.2018",
  "dmp:hasDataObject": [
    {
      "@id": "https://doi.org/10.5281/zenodo.1209833",
      "@type": "dmp:SourceCode",
      "dmp:hasIntellectualPropertyRights": {
        "dcterms:license": "https://opensource.org/licenses/MIT"
      },
      "dmp:hasDataRepository": "https://github.com/mdietrichstein/digitalpreservation-dmp/tree/1.0.0",
      "dmp:hasPreservation": "All files that need preservation, are marked with their respective preservation duration. The files the",
      "dmp:hasDataSharing": "All code, data and documentation is available on Github and is licensed under the MIT license. To make t",
      "dmp:hasEthicsAndPrivacy": " <No information> ",
      "dmp:hasDocumentation": "The documentation can be found in all files that are marked as type documentation. These files can be",
      "dmp:hasDataCollection": "All files that are collected from external sources are marked as input-files.",
      "dmp:hasDataObject": [
        {
          "@type": "dmp:input_data",
          "dc:title": "DP_LIVE_22032018202902423.csv",
          "dmp:hasIntellectualPropertyRights": {
            "dcterms:license": "https://opensource.org/licenses/MIT"
          },
          "dmp:hasMetadata": {
            "premis:hasObjectCharacteristics": {
              "premis:fixity": {
                "premis:hasMessageDigestAlgorithm": "premis:Fixity:SHA",

```

Project and post-project phase - demo 2

<https://github.com/alexschwarzresearch/DMPlanner>



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/DMPPlanner

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.

```
{
  "@context": {
    "dc": "https://purl.org/dc/elements/1.1/",
    "dcterms": "https://purl.org/dc/terms/",
    "dmp": "https://purl.org/madmps#",
    "foaf": "http://xmlns.com/foaf/0.1/",
    "premis": "http://id.loc.gov/ontologies/premis.html#",
    "schema": "https://schema.org/",
    "time": "https://www.w3.org/2006/time#"
  },
  "@type": "dmp:DataManagementPlan",
  "dc:creator": {
    "@id": "https://orcid.org/0000-0002-4929-7875",
    "@type": "foaf:Person",
    "foaf:Organization": "SBA Research",
    "foaf:name": "Tomasz Miksa"
  },
  "dcterms:created": "2018-09-16",
  "dcterms:title": "TUW DMP",
  "dmp:hasDataObject": [
    {
      "@id": "https://doi.org/10.5281/zenodo.1172673",
      "@type": "dmp:Documentation",
      "dmp:hasDataObject": [],
      "dmp:hasIntellectualPropertyRights": [
        {
          "dcterms:license": "Creative Commons Attribution 4.0"
        },
        {
          "dcterms:license": "Open Access"
        }
      ]
    }
  ],
  "dmp:hasMetadata": {
    "dc:date": "2018-02-13T15:41:09Z",
    "dcterms:abstract": "

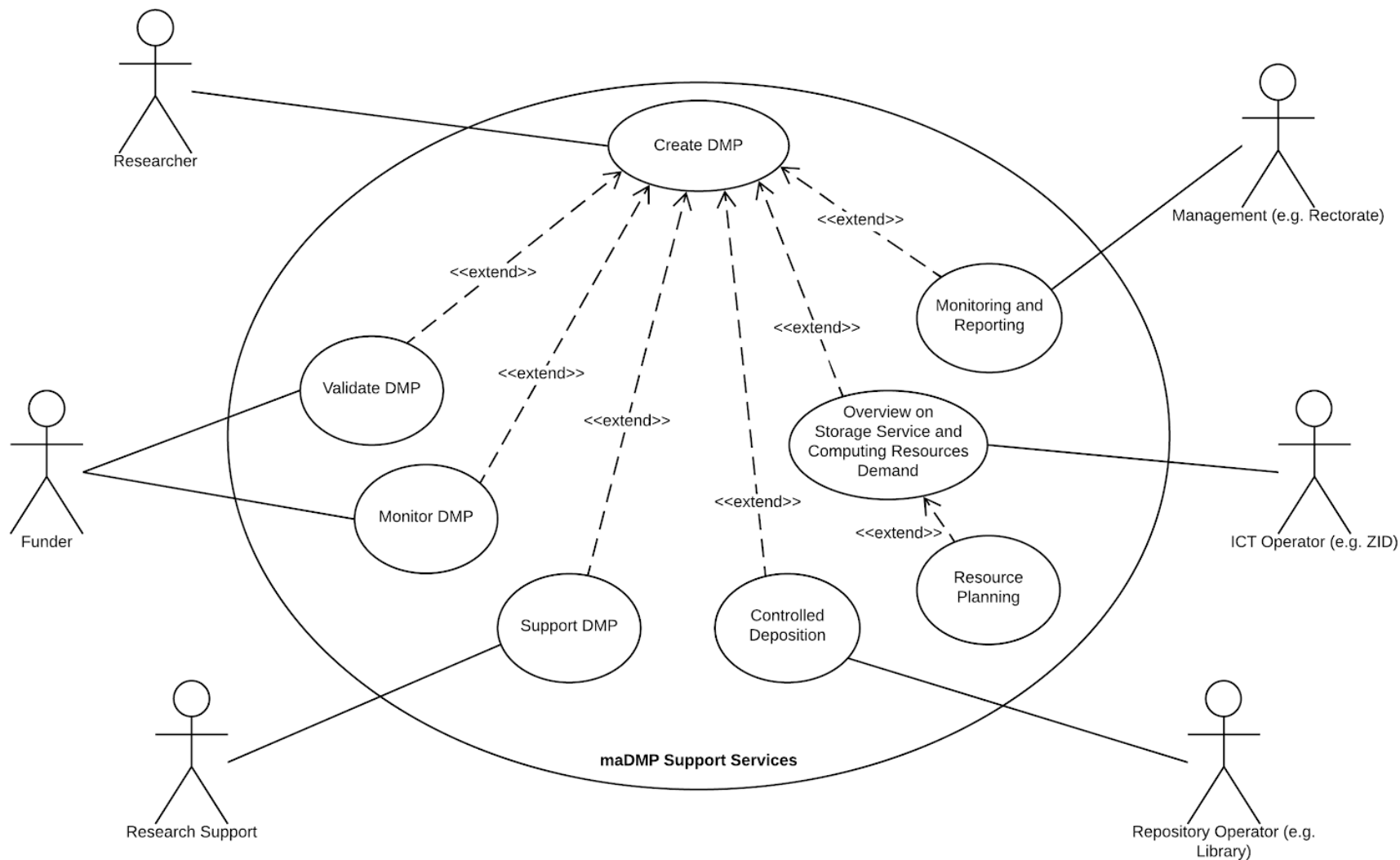
```

Data management plans (DMPs) are documents accompanying research proposals and project outputs. They describe data

Processes for maDMPs

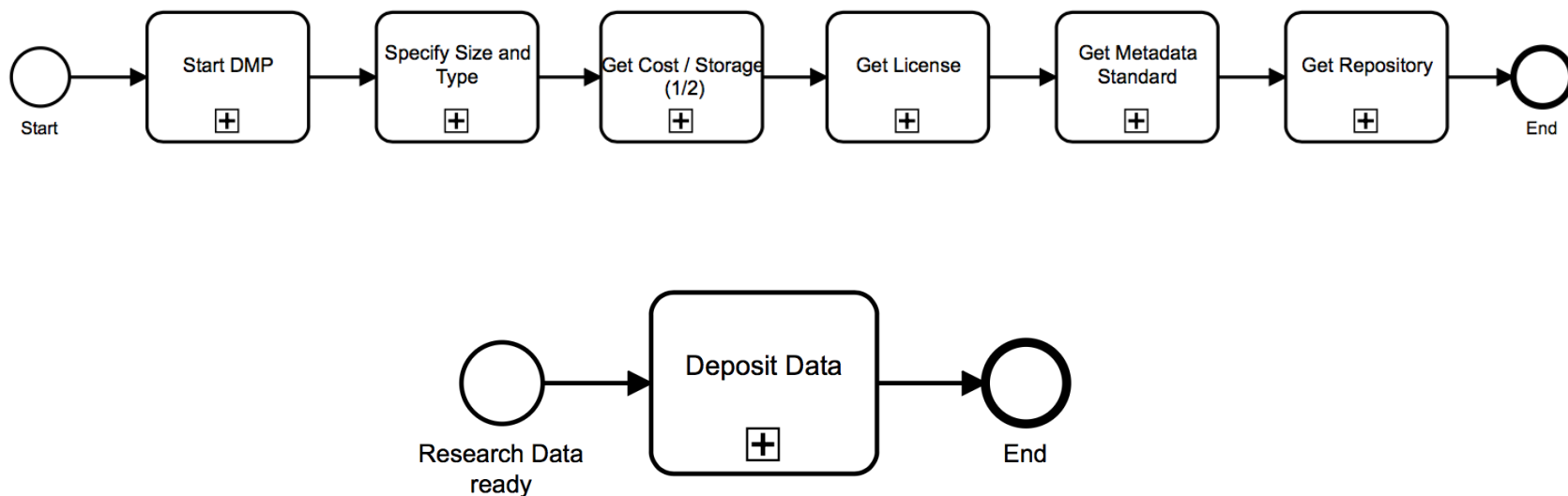
Part 4 – Identifying stakeholder interactions and services

maDMPs use cases

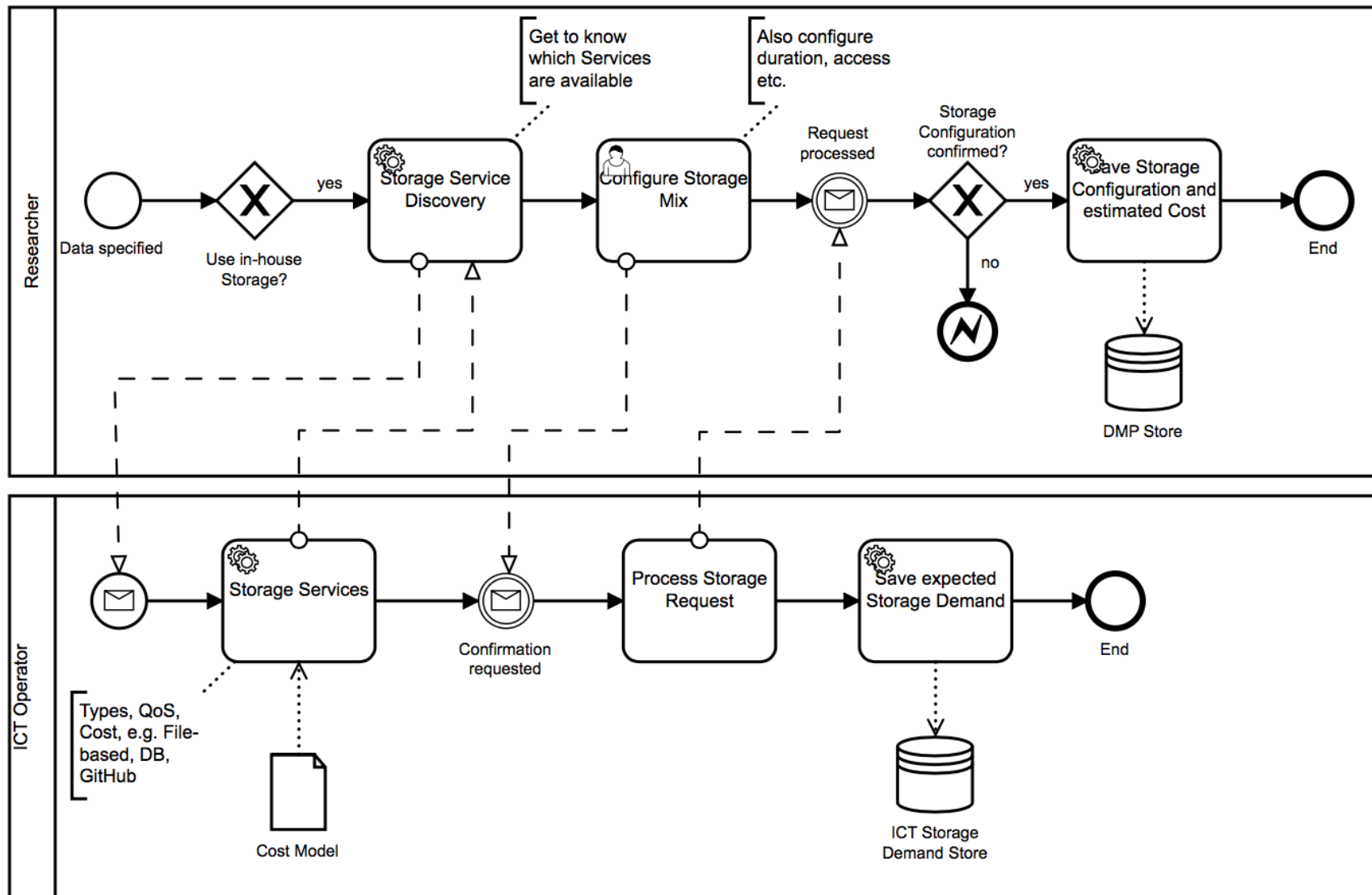


BPMN process - overview

- Business Process Modelling Notation (BPMN)
- Defined 10 workflows



Get Cost / Storage



BPMN Processes for machine-actionable DMPs

Simon Oblasser & Tomasz Miksa

Contents

Start DMP	2
Specify Size 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
Deposit Data	9
Get Help	10

<http://rda-ws-tpdl2018.sysresearch.org/documents/2018-TPDL-Porto-Handout-BPMN.pdf>

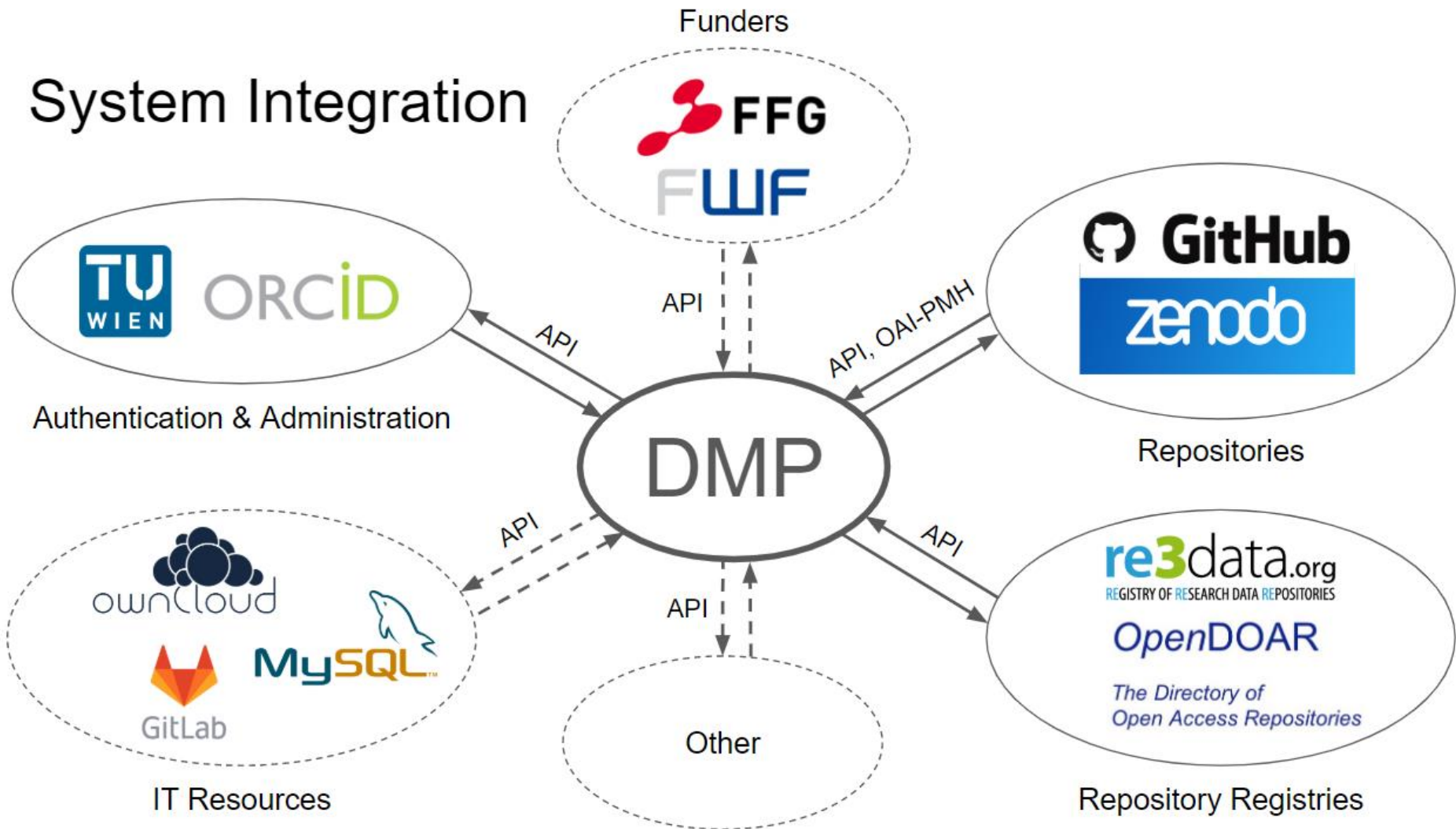
Processes - summary

- 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 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*

Mock-up for a tool + prototype

- Goal
 - generate easily and quickly DMPs
 - not a training tool
- Mock-ups
 - To define requirements of ALL stakeholders
- Prototype currently developed at the TU Wien
 - Spring Boot + Vue.js
- 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

← → ↻

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 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
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.

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/>

Summary of actions

- 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
- Mock-ups
 - Implement processes and create a proof of concept

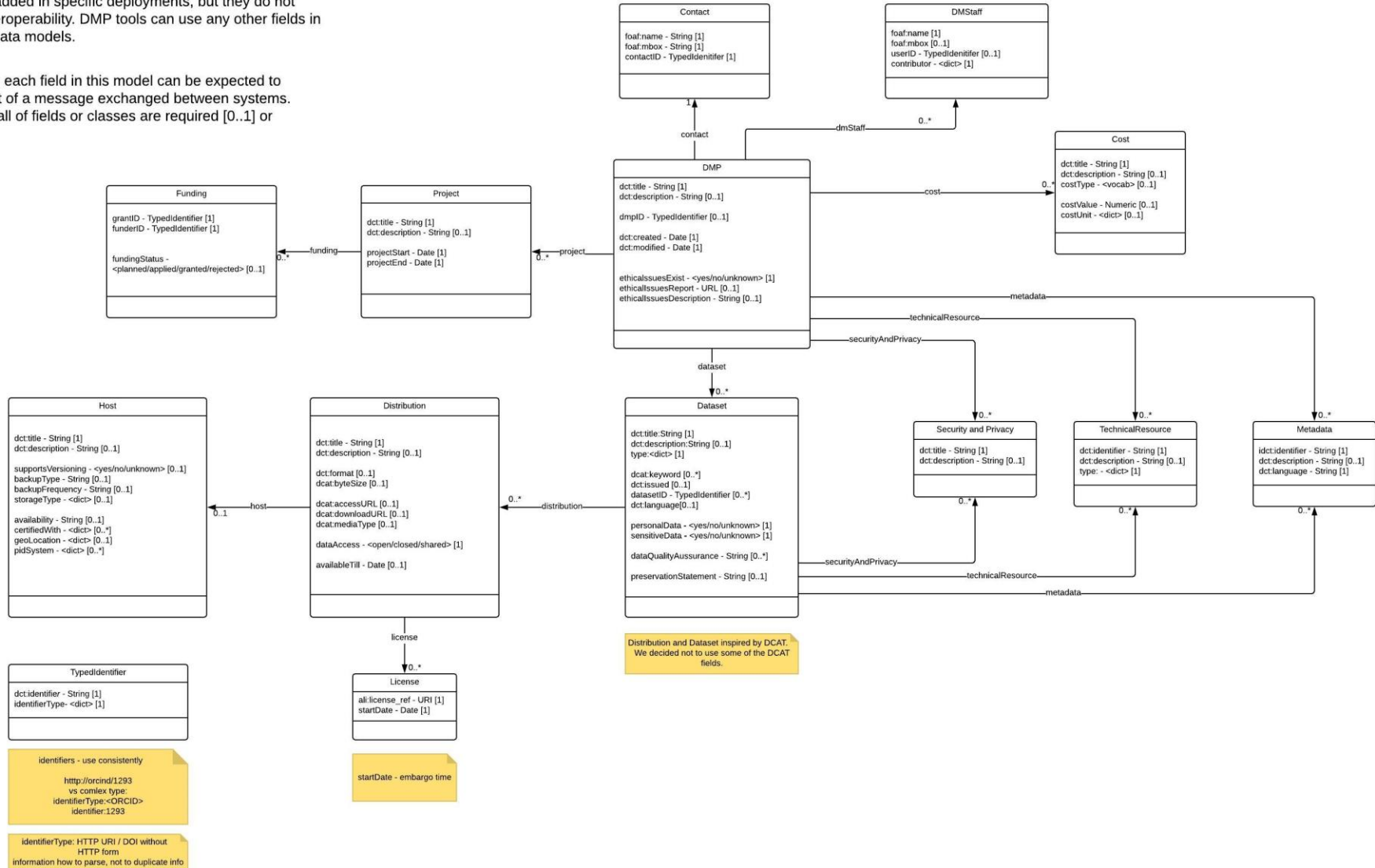
Common Data Model

Part 5 – work in progress!

Interchange Format for maDMPS

This is a minimum set of universal terms that we agree on which ensure basic interoperability of systems using maDMPS. Further fields can be added in specific deployments, but they do not guarantee interoperability. DMP tools can use any other fields in their internal data models.

Cardinalities: each field in this model can be expected to appear as part of a message exchanged between systems. However, not all of fields or classes are required [0..1] or [0..*].



<https://www.lucidchart.com/invitations/accept/ee26bc71-01a6-442a-b946-5b9c910fb926>

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 – 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

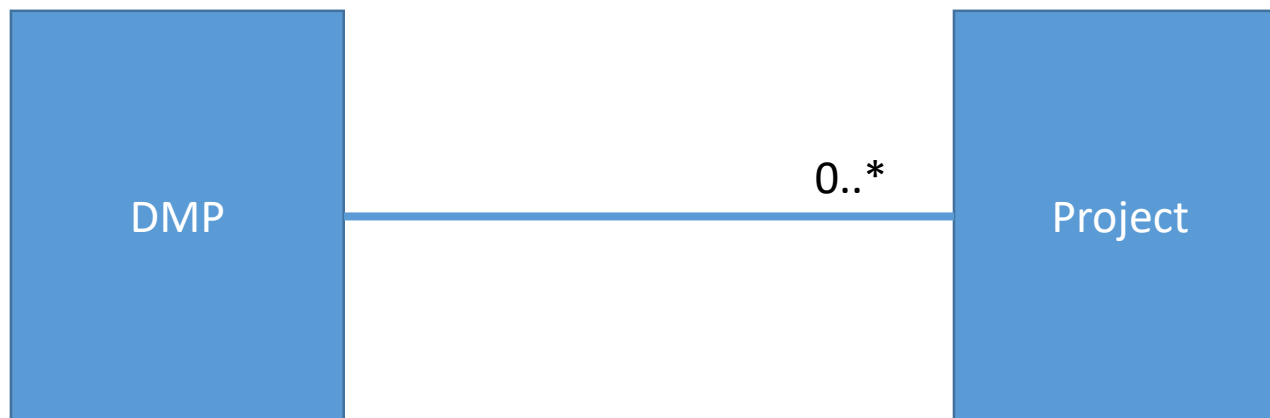
Raw Blame History

```
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 }
39 }
```

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

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

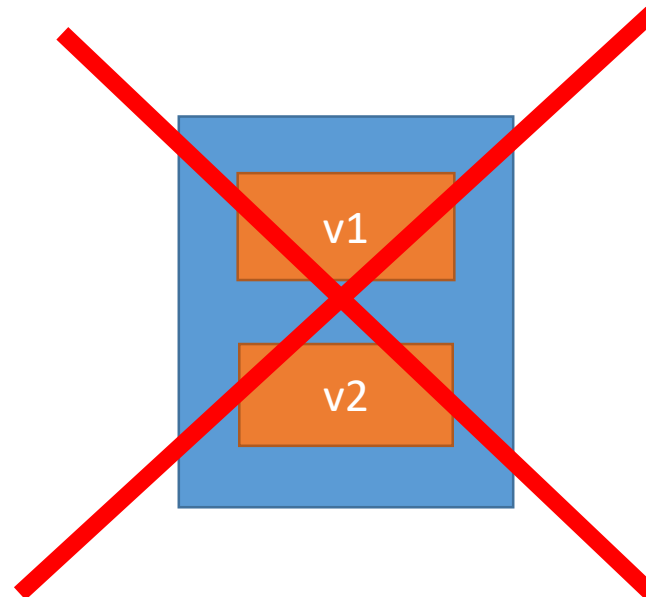
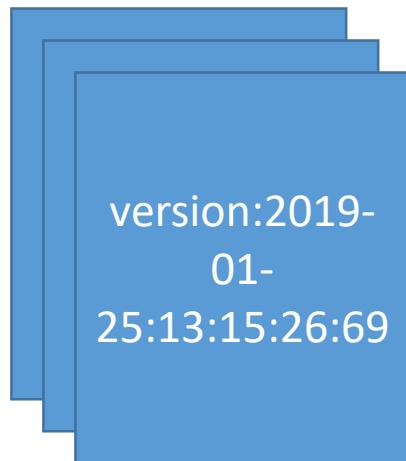


Model assumptions - interoperability

- 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

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"  
      }  
    }  
  }  
}
```

Wrap-up and next steps

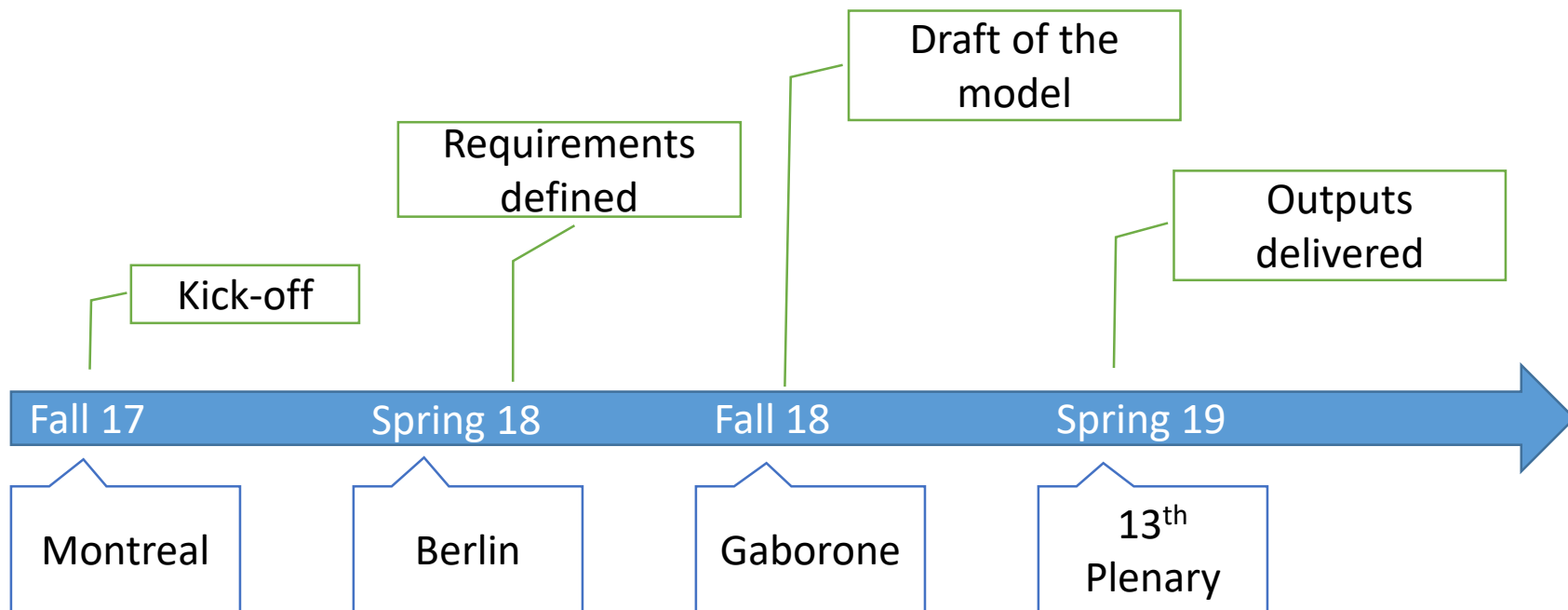
Part 5

RDA Plenary in Philadelphia - April

- Present the model
- Launch pilots
 - to use the model
 - to deploy tools using the model
- Become an official RDA recommendation
- In meantime
 - TU Delft to review mock-ups and identify its internal processes to include all stakeholders in data management
 - DMP Online to map their model into the common model

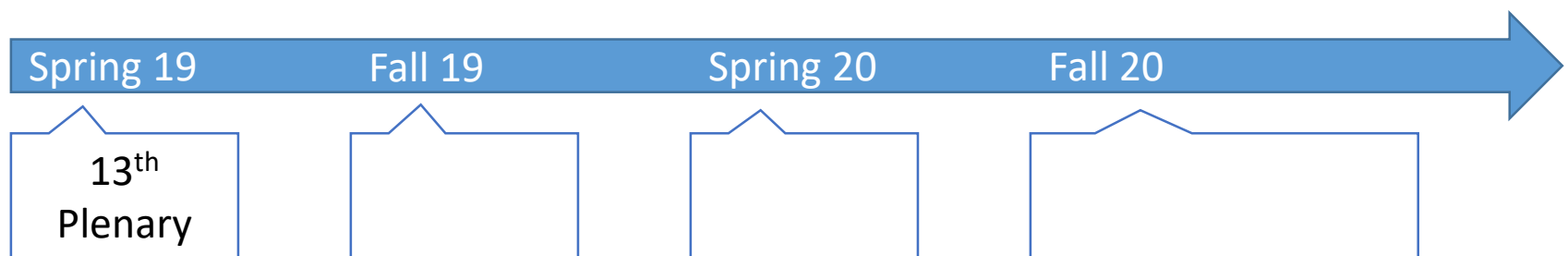
Timeline

- Slide taken from the Berlin plenary presentation
 - Seems we're on track



Future work

- Adoption of the model
 - By DMP tool providers
 - By system owners
 - Funders
 - Repositories
 - Universities
- Adoption can be gradual
- What is your use case for maDMPs?



Call For Europe Adoption Grants

[Home](#) » [Previous Calls](#) » [Call For Europe Adoption Grants](#)

03 December 2018 - 10:00 To 04 March 2019 - 17:00

[Frequently Asked Questions](#)

Status: Closed

RDA Europe has launched a Call for projects in European organisations that adopt existing RDA Recommendations and Outputs.

The intent of the call is to support and encourage examples of adoption which can benefit others, to promote these examples and to learn lessons about benefits and challenges that arise from making use of RDA recommendations. Since its inception in 2013, 17 Recommendations and 11 Outputs have been produced by RDA Working Groups and Interest Groups (see [the following link](#) for a complete list).

These are very wide ranging, addressing registries for persistent identifiers and data types, policy templates, repository audit methodologies, standards directories, curricula, wheat data interoperability, data/literature cross-linking and many other topics. Opportunities for implementation are equally wide-ranging and address a diverse set of stakeholder groups. A number of recommendations have also been endorsed as ICT technical specifications by the European Multi-Stakeholder Platform on ICT specifications (see [them here](#)). The RDA Europe Adoption Grants Call will support 8 projects up to a maximum of €15,000 each. Projects should be short and focused activities that run for up to 12 months that result in the use of one or more RDA Recommendations and Outputs in a context capable of providing lessons to other potential adopters. A number of conditions must be met as outlined below. The deadline for submitting applications is 04 March 2019.

- Projects must adopt one or more existing RDA recommendation(s) and/or Output(s) (focus on the RDA ICT Technical Specifications is a plus) and have sufficient knowledge to implement without significant assistance from the originating WG/IG members.
- The project must add value to the RDA output by providing a practical use case, a description of benefits of adoption, guidelines that can assist others in adoption, and/or constructive criticism and recommendations for improvement if the planned adoption was challenging for any reason.
- Projects can focus on particular research domains or cross-discipline use cases, but cross-disciplinary and multi-disciplinary projects will be favoured where this is appropriate to the recommendation being adopted.
- The project partners must demonstrate that there is appropriate co-funding (e.g. person effort, in-kind contributions, match funding etc)
- All projects are expected to present their adoption case in a webinar and/or plenary and provide written case studies and final reports to assist others to make decisions on output adoption and to carry out that adoption. The RDA has the right to publish outcomes and written summaries non-exclusively on RDA media.
- Outcomes / Adoption grant results must be available for at least 4 years after the end of the project.

<https://grants.rd-alliance.org/OpenCalls/call-europe-adoption-grants>

Staying in touch!

- Sign up to the group

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

- Participate in model deployment

- What is your use case for maDMPs?

- Contact group chairs



Tomasz Miksa



Paul Walk



Peter Neish