Machine-actionable DMP application (DMap)

RDA 14th plenary, Helsinki

23 Oct 2019 | Simon Oblasser | simon.oblasser@student.tuwien.ac.at
Automated DMP workflow

Research prototype & demo tool

https://github.com/oblassers/dmap

- automation & system integration
- developed using stakeholder feedback on graphical mockups
- first tool to support RDA DMP Common Standard
Welcome to DMap
The Machine-actionable Data Management planning application
DMap assists you in creating Data Management Plans (DMPs) for your research projects.

GET STARTED
<table>
<thead>
<tr>
<th>ID</th>
<th>Projects</th>
<th>Contact</th>
<th>Created</th>
<th>Last Update</th>
</tr>
</thead>
</table>
### Create a new DMP

**Select project(s)**

Please select the project(s) you want to create a DMP for.

Currently you have no projects selected.

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Start Date - End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explainable CPS (ECPS)</td>
<td>15.07.2019 - 30.09.2019</td>
</tr>
<tr>
<td>Setup and management of the EOSC Secretariat supporting the EOSC Governance (EOSCSecretariat.eu)</td>
<td>01.01.2019 - 30.06.2021</td>
</tr>
<tr>
<td>Artificial Researcher in Science: Efficient Scientific Publication Mining (AR-Science)</td>
<td>01.01.2019 - 31.03.2021</td>
</tr>
<tr>
<td><strong>openEO - a common, open source interface between Earth Observation data infrastructures and front-end applications (openEO)</strong></td>
<td>01.10.2017 - 30.09.2020</td>
</tr>
</tbody>
</table>

**Search projects...**

**NEXT STEP**
Select project(s)

Please select the project(s) you want to create a DMP for.

You selected 1 project(s):

- **openEO - a common, open source interface between Earth Observation data infrastructures and front-end applications (openEO)**
  - 01.10.2017 - 30.09.2020

- **Explainable CPS (ECPS)**

- **Setup and management of the EOSC Secretariat supporting the EOSC Governance (EOSCSecretariat.eu)**
  - 01.01.2019 - 30.06.2021

- **Artificial Researcher in Science: Efficient Scientific Publication Mining (AR-Science)**
  - 01.01.2019 - 31.03.2021

- **Innovationslehrgang Data Science und Deep Learning (iDSL)**

- **openEO - a common, open source interface between Earth Observation data infrastructures and front-end applications (openEO)**
  - 01.10.2017 - 30.09.2020

Project not in the list? Try to find it with the search function.
### Select project(s)

### People involved in data management

Please select the people involved in data management and assign roles.

Currently you have no people selected.

- **Projektass. Dr.-Ing. Matthias Schramm**
  - Sub project manager

- **Senior Scientist Dipl.-Ing. Alexandra von Beringe**
  - Project member

- ** Univ.Prof. Dipl.-Ing. Dr.techn. Wolfgang Wagner**
  - Project manager

- **Projektass. Dipl.-Ing. Alena Dostalova**
  - Project member

- **Dipl.-Ing. Bernhard Gößwein BSc**
  - Project member

**Search for person...**
Create a new DMP

Select project(s)

People involved in data management

Please select the people involved in data management and assign roles.

You selected 1 person(s):

1. **Projektass. Dr.-Ing. Matthias Schramm**
   - Contact Person

2. **Projektass. Dr.-Ing. Matthias Schramm**
   - Sub project manager

3. **Senior Scientist Dipl.-Ing. Alexandra von Beringe**
   - Project member

4. **Univ.Prof. Dipl.-Ing. Dr.techn. Wolfgang Wagner**
   - Project manager

5. **Projektass. Dipl.-Ing. Alena Dostalova**
   - Project member

6. **Dipl.-Ing. Bernhard Gößwein BSc**
   - Project member

Assign contribution roles

- Contact Person
- Data Manager

Person not in the list? Try the search function.
Create a new DMP

1. Select project(s)
2. People involved in data management
3. Specify your research data
   - What kinds of research data will you create?
     - Don’t know yet
     - No data will be created or analysed
     - Specify with assistance
4. Documentation and data quality
5. Legal and ethical aspects
6. Specify license(s)
7. Specify repository/repositories
8. End
Create a new DMP

1. Select project(s)

2. People involved in data management

3. Specify your research data

   What kinds of research data will you create?
   - Don’t know yet
   - No data will be created or analysed
   - Specify with assistance

   Please explain why you won’t create or analyse any research data.

   [Text box for explanation]

4. Documentation and data quality

5. Legal and ethical aspects
Specify your research data

What kinds of research data will you create?

- Don’t know yet
- No data will be created or analysed
- Specify with assistance

Currently you have no research data specified.

Create names for your datasets

- Sumatra Image Collection
- Processing App

Estimate the type, format and volume of your research data manually and/or by uploading sample data and group them into datasets.

MANUALLY UPLOADING SAMPLE DATA

Manual Data Estimation

PREVIOUS STEP NEXT STEP

Specify license(s)

Specify repository/repositories
Specify your research data

What kinds of research data will you create?

- Don’t know yet
- No data will be created or analysed
- Specify with assistance

Currently you have no research data specified.

Create names for your datasets

[Sumatra Image Collection] [Processing App]

Estimate the type, format and volume of your research data manually and/or upload sample data:

- MANUALLY
- UPLOADING SAMPLE DATA

Manual Data Estimation

- [PREVIOUS STEP] [NEXT STEP]

Documentation and data quality

Legal and ethical aspects

Specify licenses

Specify repository/repositories

New Data Estimation

Estimated type:

- Standard office documents (text documents, spreadsheets, presentations)
- Network-based data (websites, email, chat history, etc.)
- Databases (DBASE, MS Access, Oracle, MySQL, etc.)
- Images (JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.)
- Structured graphics (CAD, CAM, 3D, VRML, etc.)
- Audio/visual data (WAVE, MP3, MP4, Flash, etc.)
Specify your research data

What kinds of research data will you create?

- Don’t know yet
- No data will be created or analysed
- Specify with assistance

Currently you have no research data specified.

Create names for your datasets
- Sumatra Image Collection
- Processing App

Estimate the type, format and volume of your research data manually and / or online

MANUALLY UPLOADING SAMPLE DATA

Manual Data Estimation

Estimated type
Images (JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.)

Estimated size
- < 100 MB
- 100 - 1000 MB
- 1 - 5 GB
- 5 - 20 GB
- 20 - 50 GB
- 50 - 100 GB
Specify your research data

What kinds of research data will you create?
- Don’t know yet
- No data will be created or analysed
- Specify with assistance

Currently you have no research data specified.

Create names for your datasets
- Sumatra Image Collection
- Processing App

Estimate the type, format and volume of your research data manually and / or automatically.

MANUALLY UPLOADING SAMPLE DATA

Manual Data Estimation

- Estimated type: Images
- Estimated size: 20 - 50 GB
- Belongs to dataset: Sumatra Image Collection

Comment:
Satellite images of northwestern Sumatra

CANCEL SAVE
Specify your research data

What kinds of research data will you create?

- Don't know yet
- No data will be created or analysed
- Specify with assistance

You specified 1 dataset(s)

Sumatra Image Collection

20GB - 50GB

Images

(JPEG, JPEG2000, GIF, TIF, PNG, SV, etc.)

Satellite images of northwestern Sumatra

Create names for your datasets

Sumatra Image Collection

Processing App

- Estimate the type, format and volume of your research data manually and/or by uploading sample data and group them into datasets.

Manual Data Estimation

<table>
<thead>
<tr>
<th>Estimated type</th>
<th>Estimated size</th>
<th>Dataset</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Images</td>
<td>20 - 50 GB</td>
<td>Sumatra Image Collection</td>
<td>Satellite images of northwestern Sumatra</td>
</tr>
</tbody>
</table>
You specified 2 datasets:

**Sumatra Image Collection**
- 20GB - 50GB
- Images (JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.)
- Satellite images of northwestern Sumatra

**Processing App**
- 100MB - 1GB
- Source code:
  - (scripting, Java, C, C++, Fortran, etc.)
  - Python scripts
- Configuration data:
  - (parameter settings, logs, library files)
  - Configuration files to setup the runtime environment

Create names for your datasets:
- Sumatra Image Collection
- Processing App

Estimate the type, format and volume of your research data manually and/or by uploading sample data and group them into datasets.

**MANUALLY UPLOADING SAMPLE DATA**

### Manual Data Estimation

<table>
<thead>
<tr>
<th>Estimated type</th>
<th>Estimated size</th>
<th>Dataset</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Images</td>
<td>20 - 50 GB</td>
<td>Sumatra Image Collection</td>
<td>Satellite images of northwestern Sumatra</td>
</tr>
<tr>
<td>Source code</td>
<td>100 - 1000 MB</td>
<td>Processing App</td>
<td>Python scripts</td>
</tr>
<tr>
<td>Configuration data</td>
<td>&lt; 100 MB</td>
<td>Processing App</td>
<td>Configuration files to setup the runtime environment</td>
</tr>
</tbody>
</table>
You specified 2 dataset(s)

**Sumatra Image Collection**

20GB - 50GB

Images

- JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.
- Satellite images of northwestern Sumatra

**Processing App**

100MB - 1GB

Source code

- (scripting, Java, C, C++, Fortran, etc.)
- Python scripts

Configuration data

- (parameter settings, logs, library files)
- Configuration files to setup the runtime environment

Create names for your datasets

- Sumatra Image Collection
- Processing App

Estimate the type, format and volume of your research data manually and/or by uploading sample data and group them into datasets.

MANUALLY  UPLOADING SAMPLE DATA

- 9.6 MB
- 13 Creep (Radio Edit).mp3
Create names for your datasets

Sumatra Image Collection
Processing App

Estimate the type, format and volume of your research data manually and / or by uploading sample data and group them into datasets.

**MANUALLY UPLOADING SAMPLE DATA**

Drop files here to analyze

<table>
<thead>
<tr>
<th>Format</th>
<th>Format ID</th>
<th>Size</th>
<th>Amount</th>
<th>Total</th>
<th>Dataset</th>
<th>Comment</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEG 1/2 Audio Layer 3</td>
<td>lm/134</td>
<td>10MB</td>
<td>1</td>
<td>10MB</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rows per page: 5
1-1 of 1

**Specify license(s)**

Specify repository/repositories
Edit Sample Data Estimation

Format: MPEG 1/2 Audio Layer 3
Format ID: fmt/134
Size: 10MB

Amount: 500
Total size: 5GB

Belongs to dataset: Interviews

Comment: Recordings of interviews from X

CANCEL  SAVE
You specified 3 dataset(s):

**Sumatra Image Collection**
- **Size**: 20GB - 50GB
- **Images**: JPEG, JPEG2000, GIF, TIF, PNG, SVG, etc.
- Satellite images of northwestern Sumatra

**Processing App**
- **Size**: 100MB - 1GB
- **Source code**: Java, C, C++, Fortran, etc.
- Python scripts
- **Configuration data**: (parameter settings, logs, library files)
- Configuration files to setup the runtime environment

**Interviews**
- **Size**: 5GB
- **Source**: MPEG 1/2 Audio Layer 3
- Recordings of interviews from X

Create names for your datasets:
- Sumatra Image Collection
- Processing App
- Interviews

Estimate the type, format and volume of your research data manually and/or by uploading sample data and group them into datasets.

MANUALLY UPLOADING SAMPLE DATA

Drop files here to analyze

<table>
<thead>
<tr>
<th>Format</th>
<th>Format ID</th>
<th>Size</th>
<th>Amount</th>
<th>Total</th>
<th>Dataset</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPEG 1/2 Audio Layer 3</td>
<td>fmt/134</td>
<td>10MB</td>
<td>500</td>
<td>5GB</td>
<td>Interviews</td>
<td>Recordings of interviews from X</td>
</tr>
</tbody>
</table>
## Documentation and data quality

Please indicate which metadata will be provided to help others identify, discover and reuse the data.

**Metadata**

How will the research data be generated and which methods will be used?

**Data generation**

How will you structure the data and handle versioning?

**Structure and versioning**

Who is the target audience?

**Target audience**
Legal and ethical aspects

This step helps you to consider issues like personal data, intellectual property rights and ownership.

Does your data contain personal information?
- Yes
- No

Is your data sensitive?
- Yes
- No

Are there any other legal restrictions on how data is processed or shared?
- Yes
- No

Are there any ethical issues associated with your research data?
- Yes
- No

If so, was your plan to deal with these ethical issues approved by the ethics committee?
- Yes
- No

Link to ethics report or related document:
Legal and ethical aspects

This step helps you to consider issues like personal data, intellectual property rights and ownership.

Does your data contain personal information?

- Yes
- No

Please indicate which datasets contain personal information:

- Sumatra Image Collection
- Processing App
- Interviews

Is your data sensitive?

- Yes
- No

Are there any other legal restrictions on how data is processed or shared?

- Yes
- No

Are there any ethical issues associated with your research data?

- Yes
Legal and ethical aspects

This step helps you to consider issues like personal data, intellectual property rights and ownership.

Does your data contain personal information?

- [ ] Yes
- [X] No

Please indicate which datasets contain personal information:

- Interviews

Please describe:

Audio recordings contain personal information...

Is your data sensitive?

- [ ] Yes
- [X] No

Are there any other legal restrictions on how data is processed or shared?

- [ ] Yes
- [X] No

Are there any ethical issues associated with your research data?

- [ ] Yes
Create a new DMP

Select project(s)

People involved in data management

Specify your research data

Documentation and data quality

Legal and ethical aspects

Specify license(s)

Please indicate which datasets you plan to keep closed or publish. Select a license for the datasets you want to publish and estimate the release date.

Sumatra Image Collection

Processing App

Interviews
Please indicate which datasets you plan to keep closed or publish. Select a license for the datasets you want to publish and estimate the release date.

**Sumatra Image Collection**

CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

**Planned license**

<table>
<thead>
<tr>
<th>Public Domain Dedication (CC Zero)</th>
<th>Planned to be active from</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2020-09-30</td>
</tr>
</tbody>
</table>

See also [http://creativecommons.org/publicdomain/zero/1.0/](http://creativecommons.org/publicdomain/zero/1.0/)

**Processing App**

- Keep closed

**Interviews**

- Keep closed
Documentation and data quality

Legal and ethical aspects

Specify license(s)

Please indicate which datasets you plan to keep closed or publish. Select a license for the datasets you want to publish and estimate the release date.

Sumatra Image Collection

CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

Planned license:

- Public Domain Mark (PD)
- Public Domain Dedication (CC Zero)
- Creative Commons Attribution (CC-BY)
- Creative Commons Attribution-ShareAlike (CC-BY-SA)
- Creative Commons Attribution-NoDerivs (CC-BY-ND)
- Creative Commons Attribution-NonCommercial (CC-BY-NC)

Planned to be active from:

2020-09-30

Publish

Keep closed
Choose a License

Answer the questions or use the search to find the license you want

What do you want to deposit?
- Software
- Data

Search for a license...

Creative Commons Attribution (CC-BY)
This is the standard creative commons license that gives others maximum freedom to do what they want with your work.

Publicly Available

Creative Commons Attribution-NoDerivs (CC-BY-ND)
The no derivatives creative commons license is straightforward; you can take a work released under this license and re-distribute it but you cannot change it.

Publicly Available

Creative Commons Attribution-NonCommercial (CC-BY-NC)
A creative commons license that bans commercial use.

Publicly Available

Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND)
The most restrictive creative commons license. This only allows people to download and share your work for no commercial gain and for no other purposes.
Doc: 4. Documentation and data quality

- Legal and ethical aspects

- Specify license(s)

Please indicate which datasets you plan to keep closed or publish. Select a license for the datasets you want to publish and estimate the release date.

**Sumatra Image Collection**

The most restrictive creative commons license. This only allows people to download and share your work for no commercial gain and for no other purposes.

Planned license:
Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND)

See also [http://creativecommons.org/licenses/by-nc-nd/4.0/](http://creativecommons.org/licenses/by-nc-nd/4.0/)

**Processing App**

**Interviews**
Documentation and data quality

Legal and ethical aspects

Specify license(s)

Please indicate which datasets you plan to keep closed or publish. Select a license for the datasets you want to publish and estimate the release date.

**Sumatra Image Collection**

The most restrictive creative commons license. This only allows people to download and share your work for no commercial gain and for no other purposes.

 Planned license
 Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND)  
 See also [http://creativecommons.org/licenses/by-nc-nd/4.0/](http://creativecommons.org/licenses/by-nc-nd/4.0/)

**Processing App**

A short, permissive software license. Basically, you can do whatever you want as long as you include the original copyright and license.

 Planned license
 The MIT License (MIT)  
 See also [http://opensource.org/licenses/mit-license.php](http://opensource.org/licenses/mit-license.php)

**Interviews**

- keep closed

[PREVIOUS STEP] [NEXT STEP]
Please select the repository/repositories where you plan to deposit your data.

Currently you have no repositories selected.
Specify repository/repositories

Please select the repository/repositories where you plan to deposit your data.

Currently you have no repositories selected.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>CancerData.org</th>
<th>SELECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content types</td>
<td>Claremont Colleges Digital Library</td>
<td>SELECT</td>
</tr>
<tr>
<td>Countries</td>
<td>PRISM: University of Calgary Institutional Repository</td>
<td>SELECT</td>
</tr>
<tr>
<td>Certificates</td>
<td>Peptidome</td>
<td>SELECT</td>
</tr>
<tr>
<td>PID systems</td>
<td>Digital Lunar Orbiter Photographic Atlas of the Moon</td>
<td>SELECT</td>
</tr>
<tr>
<td>AID systems</td>
<td>International Satellite Cloud Climatology Project</td>
<td>SELECT</td>
</tr>
<tr>
<td>Repository access</td>
<td>EVIA Digital Archive Project</td>
<td>SELECT</td>
</tr>
<tr>
<td>Data access</td>
<td>Export Import Data Bank</td>
<td>SELECT</td>
</tr>
<tr>
<td>Data upload</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data licenses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repository types</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institution types</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2415 repositories found
figshare allows researchers to publish all of their research outputs in an easily citable, sharable and discoverable manner. All file formats can be published, including videos and datasets. Optional peer review process, figshare uses creative commons licensing.

Repository URL: http://figshare.com/
Repository types: other
Certificates: none
PID systems: DOI
AID systems: ORCID
Versioning: yes
Repository access: open
Data access: open
Data upload: restricted (registration)
Data licenses: CC, Apache License 2.0, CC, other, other, BSD
Metadata standards: DataCite Metadata Schema, Dublin Core
Content types: Standard office documents, images, Structured graphics, Audiovisual data, Raw data, Plain text, Archived data
Scientific and statistical data formats, Source code
Subjects: Humanities and Social Sciences, Life Sciences, Natural Sciences, Engineering Sciences

University of Adelaide figshare

University of Amsterdam / Amsterdam University of Applied Sciences figshare

Brunel figshare
Specify repository/repositories

Please select the repository/repositories where you plan to deposit your data.

Currently you have no repositories selected.

Filter

Subjects
Select subjects...

Content types
Select content types...

Countries
Select countries...

Certificates
Select certificates...

PID systems
Select PID systems...

AID systems
Select AID systems...

Repository access
Select repository access...

Data access
Select data access...

Data upload
Select data upload...

Data licenses
Select data licenses...

Repository types
Select repository types...

figshare

University of Adelaide figshare

figshare

University of Amsterdam / Amsterdam University of Applied Sciences figshare

Brunel figshare

DMU Figshare

monash.figshare

St. Edward's University institutional repository

University of Salford Data Repository

23 repositories found
Specify repository/repositories

Please select the repository/repositories where you plan to deposit your data.

You selected 1 repository/repositories:

figshare

- [x] Sumatra Image Collection
- [ ] Processing App
- [ ] Interviews

Available until: 2030-09-30

University of Adelaide figshare

University of Amsterdam / Amsterdam University of Applied Sciences figshare

Brunel figshare

DMU Figshare

monash.figshare

St. Edward's University institutional repository

University of Salford Data Repository
Congratulations, you successfully created an initial DMP! Now you can select an export type and download the DMP or send it directly to your funder.
RDA DMP Common Standard

This is a minimum set of universal terms that we agree on which ensure basic interoperability of systems using DMPs. 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..*].

Full documentation:
https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard
DMP info
Contact
DM staff & roles
Ethical issues
Project & funding
Datasets
<table>
<thead>
<tr>
<th>Datasets</th>
<th>Type</th>
<th>License</th>
<th>Repository</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal &amp; sensitive data</td>
<td>Image</td>
<td>Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND)</td>
<td>figshare</td>
<td>2020-09-30</td>
</tr>
<tr>
<td>Processing App</td>
<td>Source code</td>
<td>The MIT License (MIT)</td>
<td>rikodata</td>
<td>2020-09-30</td>
</tr>
</tbody>
</table>
Congratulations, you successfully created an initial DMP! Now you can select an export type and download the DMP or send it directly to your funder.

Select export type
RDA DMP Common Standard (JSON)

Latest submission: 20.10.2019 14:04:21
Thank you!

Useful links:

https://github.com/oblassers/dmap

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

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

https://github.com/RDA-DMP-Common/RDA-DMP-Common-Standard