

Tekniska möjligheter för ett digitalt DHP-verktyg

Johan Fihn Marberg, Olof Olsson
2019-03-12




SND

Svensk nationell datatjänst | Göteborgs universitet - Karolinska institutet - Lunds universitet - Stockholms universitet - Sveriges lantbruksuniversitet - Umeå universitet - Uppsala universitet



Nationellt verktyg för DHP

- Flerspråkighet?
 - Enbart engelska, enbart svenska eller bägge språken
 - Integreringar mot andra system
 - Lärosätesspecifika system
 - Ansökningssystem
 - Inlogg via SWAMID (och eventuellt andra aktörer)
 - Output
 - PDF-dokument
 - Övriga exportformat
 - Översikt för finansiärer?
 - Översikt för lärosäten och intitutioner?
 - "Machine actionability"
- 

Maskinläsningsbara DHP

Data management plans (DMPs) are documents accompanying research proposals and project outputs. They describe data and tools employed in scientific investigations, **mostly in free-form text**. DMPs are **often seen as an administrative exercise and not as an integral part of research practice**.

There is now widespread recognition that the DMP can have more thematic, machine-actionable richness with added value for all stakeholders: researchers, funders, repository managers, research administrators, data librarians, etc. The larger goal is to improve the experience for all involved by exchanging information across research tools and systems and **embedding DMPs in existing workflows**. This will enable parts of the DMP to be automatically generated and shared, thus reducing administrative burdens and improving the quality of information within a DMP.

This paper presents 10 simple rules outlining specific steps to put machine-actionable DMPs into practice and realize their benefits.

Miksa, Tomasz, Simms, Stephanie, Mietchen, Daniel, & Jones, Sarah. (2018). Ten simple rules for machine-actionable data management plans (preprint) (Version preprint).

<http://doi.org/10.5281/zenodo.1172673>

Ten simple rules for machine-actionable data management plans



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



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

Fig 2. 10 rules for machine-actionable DMPs at a glance. <http://doi.org/10.5281/zenodo.1172673>

Flera olika verktyg

Platform	Organisation(s)	Resource link(s)
DMPRoadmap	California Digital Library Digital Curation Centre Portage Network INIST CNRS	https://github.com/DMPRoadmap/roadmap
University of Queensland Research Data Manager (UQRDM)	University of Queensland	https://research.uq.edu.au/project/research-data-manager-uqrdm
ReDBox DLC	Queensland Cyber Infrastructure Foundation	https://www.redboxresearchdata.com.au/rbdlc.html
(new DMP tool under development)	University of Auckland	https://doi.org/10.17608/k6.auckland.c.3912652.v1
RDMOrganiser (RDMO)	Leibniz Institute for Astrophysics Potsdam (AIP), Potsdam University of Applied Sciences (FHP), Karlsruhe Institute of Technology Library (KIT)	http://rdmorganiser.github.io/en/
Data Stewardship Wizard	ELIXIR Europe, Dutch Techcentre for Life Sciences (DTL)	https://github.com/DataStewardshipPortal/ https://dmp.fairdata.solutions/
ezDMP	Interdisciplinary Earth Data Alliance (IEDA)	https://www.iedadata.org https://www.rd-alliance.org/system/files/documents/4-RDA10-Lehnert-ezDMP.pptx
Data planning tool	UNINETT Sigma2 / NorStor Research Data Archive	https://www.sigma2.no/content/data-planning-tool
DMP Service	OpenAIRE and EUDAT	Public beta due in February 2018



Existerande verktyg – DMP Roadmap

- Öppen källkod
 - Läggs upp på fristående portal
 - Utvecklad av DCC och University of California Curation Center
 - Används i:
 - Storbritannien - <https://dmponline.dcc.ac.uk>
 - Danmark - <https://dmponline.deic.dk>
 - Finland - <https://dmptuuli.fi>
 - Frankrike - <https://dmp.opidor.fr>
 - Kanada - <https://assistant.portagenetwork.ca>
 - Spanien - <https://dmp.csuc.cat>
- 

DMP Roadmap

sektion

2.2 Making data openly accessible [FAIR data] (0 / 5)

In general terms, your research data should be 'FAIR' that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard or implementation-solution.

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

B *I*

Save

kommentarer

Guidance

Comments (2)

text

olof.olsson@snd.gu.se at 11 Mar 10:04

Edit Remove

Add comments to share with collaborators

B *I*

Save

hjälptexter

Specify how the data will be made available

B *I*

Save

fritext

Guidance

Comments

EC Tuuli

For example by deposition in a repository. The Registry of Research Data Repositories provides a useful listing of repositories that you can search to find a place of deposit.



Existerande verktyg – EasyDMP

- Utvecklat av Sigma2 i samarbete med EUDAT
- Används i EUDAT



EasyDMP

easy.DMP

Create data management plans

Your plans

Current plan

Create a new plan

About

Support

Log out

Data about cats

[Go to plan summary](#)

Your answers are saved every time you press "prev" or "next"

1 Data Summary

2 FAIR data

3 Allocation of resources

4 Data Security

5 Ethical aspects

6 Other

2.4 Increase data reuse

Q2.4.2 When do you plan to make your data available for reuse?*

Even after your project completes your data may still have value to fellow researchers. You should consider providing access to other researchers once the data is no longer of primary value to you.

- The data will not be made available for reuse.
- The data will be made available for reuse once the article describing the research results has been published.
- The data will be made available for reuse immediately after the quality of the data have been verified and all metadata provided.
- The data will be made available for reuse later than the end of the project.
- The data will be made available for reuse at the end of the project.

More information

If you need to go more in depth, do so here. This will be shown in the generated text

The cat data will be purrfect for re-use!

Prev

Question 2/9

Next

sektion

Inledande flerval

fritext



Existerande verktyg – RDMO

- Öppen källkod
 - Finansierat av Deutsche Forschungsgemeinschaft (DFG)
- 

Questionnaire for project *test*

Metadata and referencing / Metadata

Please fill in the form for each dataset. The different datasets will be referred to in following questions. You can add a new dataset using the green button. Once created, you can edit or delete datasets using the buttons in the top right corner.

cat data Add dataset ✎ 🗑

Which information is necessary for other parties to understand the data (that is, to understand their collection or creation, analysis, and research results obtained on its basis) and to re-use it?

- Location
- Content
- Methodology
- Creation process
- Technology
- Documentation of the software necessary to use the data
- Time
- Sources
- Agents
- Identifiers
- Other:

Which standards, ontologies, classifications etc. are used to describe the data and context information?

- Discipline-specific standards, classifications etc. are used:
- A custom description system is used (please briefly outline and, if necessary, indicate where it is documented in more detail):

Progress



Navigation

Please note that using the navigation will discard any unused input.

[skip to previous page](#)

[skip to next page](#)

[Back to project overview](#)

Overview

[General](#)

[Content classification](#)

[Technical classification](#)

[Data usage](#)

[Metadata and referencing](#)

→ [Metadata](#)

[Metadata costs](#)

[Structure, granularity, and referencing](#)

[Persistent Identifiers \(PIDs\)](#)

[PIDs costs](#)

[Legal and ethics](#)


[Storage and long-term preservation](#)

sektion

Inledande flerval



Existerande verktyg – Elixir Data Stewardship Wizard (DS Wizard)

- Öppen källkod
 - Utvecklat av ElixirNL och ElixirCZ
 - Arbetar för aktiva DHP
 - Koppling till FAIR
- 

DS Wizard

hjälpstexter

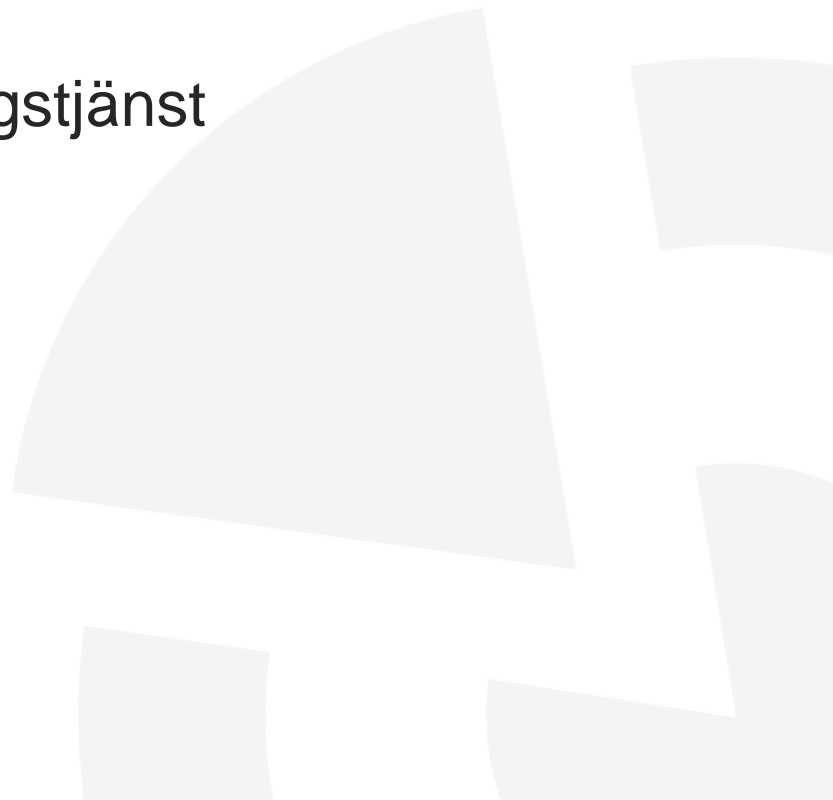
The screenshot shows the DS Wizard interface. On the left is a sidebar with navigation options: 'DS Wizard', 'Knowledge Models', 'Questionnaires', and 'KM Editor'. The main content area is titled 'Johan (Common ELIXIR Knowledge Model, 1.0.0)'. It features a 'Current Phase' dropdown menu set to 'Before Submitting the Proposal'. Below this is a list of sections: 'Design of experiment' (highlighted in orange), 'Data design and planning', 'Data Capture/Measurement' (highlighted with a red dashed box), 'Data processing and curation', 'Data integration', 'Data interpretation', 'Information and insight', and 'Summary Report'. The 'Design of experiment' section contains two questions: 'Is there any pre-existing data?' and 'Will reference data be created?'. The 'Data Capture/Measurement' section is highlighted with a red dashed box. The 'Will reference data be created?' question is also highlighted with a red dashed box. Red lines connect these highlighted elements to external labels: 'sektion' points to the sidebar, 'hjälpstexter' points to the 'Design of experiment' section, and 'Inledande flerval' points to the 'Will reference data be created?' question.

sektion

Inledande flerval



Existerande verktyg – NSD

- Egenutvecklat inom Norsk Senter For Forskningsdata
 - Stöd för flera discipliner
 - Rådgivning online via chat
 - Integrationer mot arkiveringstjänster och etikprövningstjänst
- 

seksjon

[← Back](#)

[Share](#) [Export](#) [Show history ▾](#) [Delete](#) [Save](#)

Technical information

What types of data will be collected / generated? **Primary**

- | | |
|--|---|
| <input type="checkbox"/> Dataset | <input type="checkbox"/> Image |
| <input type="checkbox"/> Video | <input checked="" type="checkbox"/> Sound |
| <input checked="" type="checkbox"/> Text | <input type="checkbox"/> Other |

Comment

If possible, estimate the presumed size / scale of the data. **Primary**

- Less than 1 GB
- 1 GB – 100 GB
- 100 GB – 1 TB
- 1 TB – 100 TB
- More than 100 TB
- Don't know

Comment


Inledande flerval

fritext

1. Project
2. Data description
3. Technical information
4. Ethical and legal issues
5. Data security, handling and storage
6. Preservation



Sammanfattning

- Hur användbart är endast fritext?
 - Ger väldigt begränsade möjligheter för översikt
 - Bör kombineras med kontrollerade vokabulär, en/flerval
 - Integrering
 - Federerad inloggning är ett måste, men kanske inte enda sättet
 - Möjligt att importera information från andra system?
 - Användargränssnittet
 - Viktigt att få det rätt
 - Svårt / omodernt = mindre användning
 - Exporter och rapporter
 - Lista över dokument som endast innehåller fritext ger ingen bra överblick över innehållet
 - Viktigt att relevanta organisationer kan få bra översikt baserad på ifylld information
- 



SND

Svensk nationell datatjänst

| Göteborgs universitet - Karolinska institutet - Lunds universitet - Stockholms universitet - Sveriges lantbruksuniversitet - Umeå universitet - Uppsala universitet

Johan Fihn Marberg
Olof Olsson