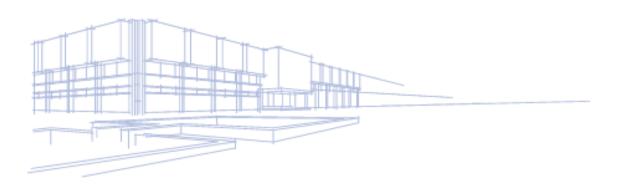


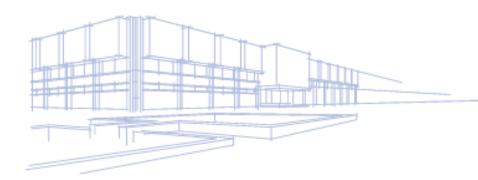
Open Research Data in H2020 Projects Data Management Plan

Tiiu Tarkpea
University of Tartu Library
Data Librarian
04.06.2020



Topics

- 1. EC requirements for H2020 research data
- 2. Research Data Lifecycle
- 3. FAIR data
- 4. Tool for DMP: DMPonline
- 5. DMP topics
- 6. Data Repositories
- 7. Tutorials and Services



Research Data

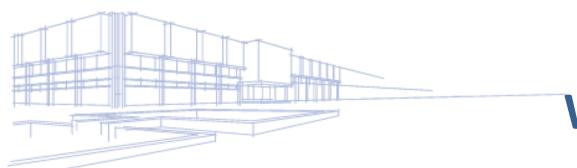
Research Data is any information collected and processed to produce original research results

H2020: Research Data refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation.

Open Research Data is the research data made accessible to others via Open Access

EC: As open as possible, as closed as necessary!

FAIRsFAIR*: Open if possible and FAIR at all times!



Why open?

Broader access to scientific publications and data helps to:

- build on previous research results (improved quality of results)
- encourage collaboration and avoid duplication of effort (greater efficiency)
- **speed up innovation** (faster progress to market means faster growth)
- involve citizens and society (improved transparency of the scientific process).

RESEARCH DATA - OPEN BY DEFAULT

Horizon 2020 grantees are required

take measures to ensure open access to the data underlying their scientific publications

provide open access to any other research data of their choice Horizon 2020 grantees are encouraged to also share datasets beyond publication



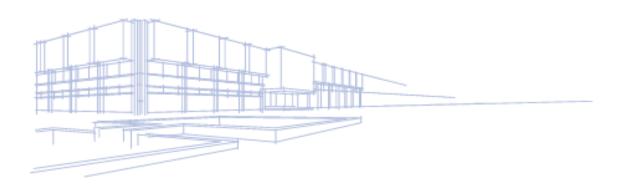












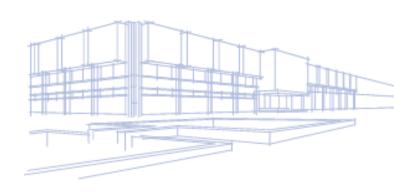
Opt-out

As open as possible, as closed as necessary!

Data cannot be shared:

- Personal data (+ pseudonymized data)
- National defence and security
- Trade secrets, patent applications
- Some environmental data (e.g. endangered species)
- Metadata of sensitive data (in some cases)

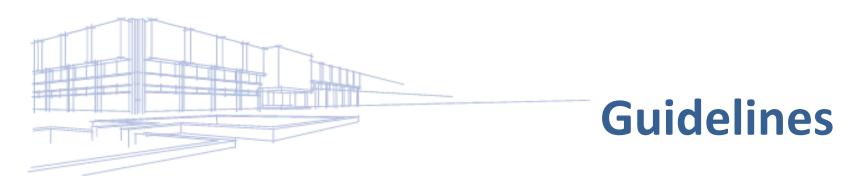
You can opt out at any stage (applying, GAP, contracting), giving the reason in DMP.



Requirements for RD in H2020 projects

- 1. Data underlying publications (raw data)
- 2. Open data, linked to publications
- 3. Create a Data Management Plan
- 4. Mandatory since 2017, possibility to opt-out
- 5. Earlier projects: read article 29.3
- 6. Costs to RDM are eligible
- 7. Read your agreement carefully!

 How to comply with H2020 mandate for research data



<u>Guidelines on FAIR Data Management in Horizon 2020</u> (Version 3.0, 26 July 2016)

Guidelines to the Rules on Open Access to Scientific

Publications and Open Access to Research Data in Horizon

2020 (Version 3.2 21 March 2017)

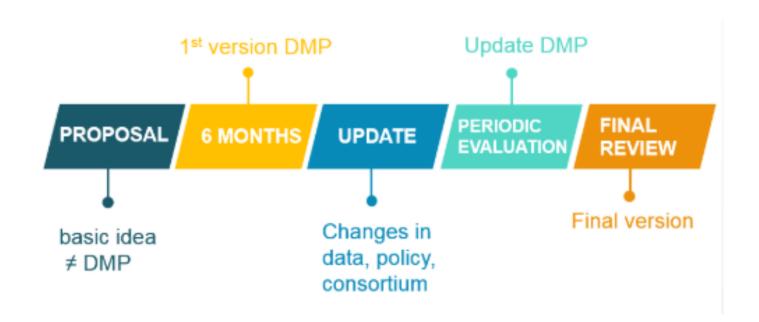
DMP template for H2020 projects

- Data summary
- FAIR Data
- RDM Costs
- Information Security
- Research Ethics

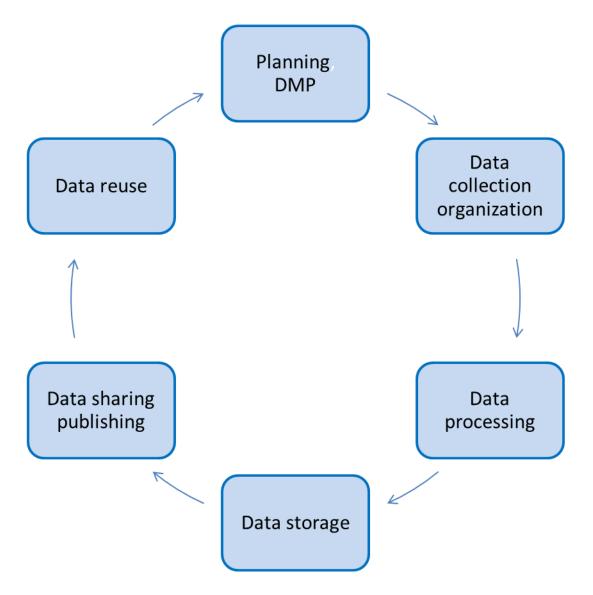
<u>DMPonline</u> – a tool for creating a DMP

Data Management Plan

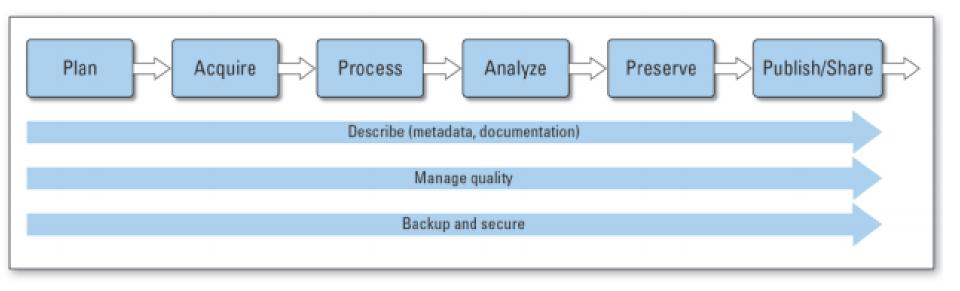
Data management plan (DMP) is an official document which describes how research data are managed during and after the completion of a research project.



Research Data Lifecycle



The United States Geological Survey Science Data Lifecycle Model



Faundeen, J.L., Burley, T.E., Carlino, J.A., Govoni, D.L., Henkel, H.S., Holl, S.L., Hutchison, V.B., Martín, Elizabeth, Montgomery, E.T., Ladino, C.C., Tessler, Steven, and Zolly, L.S., 2013, The United States Geological Survey Science Data Lifecycle Model: U.S. Geological Survey Open-File Report 2013–1265, 4 p., http://dx.doi.org/10.3133/ofr20131265

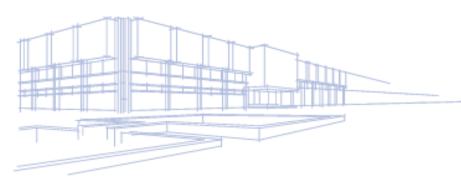
FAIR data for humans and machines

- Findable
- Digital Object Identifier (DOI) and metadata
- Accessible
- Open, free and universal software
- Interoperable
- Data exchange, relevant standards
- Reusable
- Metadata and licenses

Action plan

- 1. Create a DMP
- 2. Find a suitable data repository
- 3. Store your data
- 4. License CCO, CC-BY
- 5. Link your data and publication





DMPonline

- Developed by <u>Digital Curation Centre</u> (UK)
- Open source software
- Create DMP
- Share DMP
- Download formats
- Funders templates incl. H2020
- Guidelines
- Tutorials



https://dmponline.dcc.ac.uk/



Home

Public DMPs

Funder requirements

lelp

Welcome

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).

Join the growing international community that have adopted DMPonline:



17.622 Users



23.083 Plans

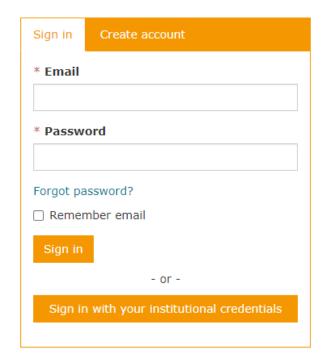


203 Organisations



89 Countries

Some funders mandate the use of DMPonline, while others point to it as a useful option. You can download funder templates without logging in, but the tool provides tailored guidance and example answers from the DCC and many research organisations. Why not sign up for an account and try it out?



Create a new plan

Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.

* What research project are you planning?		
		 mock project for testing, practice, or educational purposes
* Select the primary research organisation		
University of Tartu	8	 or -
* Select the primary funding organisation		
European Commission (Horizon 2020)	8	- or - $\ \square$ No funder associated with this plan or my funder is not listed
Create plan Cancel		

H2020 DMP



H2020 DMP

Project Details

Plan overview

Initial DMP

Detailed DMP

Final review DMP

Share

Downloa

expand all | collapse all

9/9 answered

1. Data summary (1 / 1)



2. FAIR data (4 / 4)

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.

2.1 Making data findable, including provisions for metadata:

- Outline the discoverability of data (metadata provision)
- Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?
- · Outline naming conventions used
- · Outline the approach towards search keyword
- · Outline the approach for clear versioning
- Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

Guidance

DCC

Comments

EC

The Research Data Alliance provides a Metadata Standards Directory that can be searched for discipline-specific standards and associated tools.

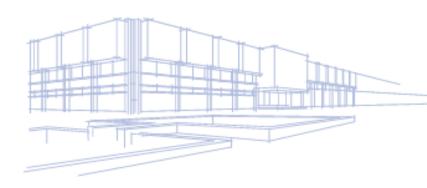
How to make data findable (F)

- The data have a permanent identifier DOI. See <u>DataCite Estonia</u>
- · metadata is in DataCite registry
- · administrative metadata (project cover sheet)
- descriptive metadata (authors, title, abstract, subjects)
- · technical metadata (hardware and software, instruments, tools)
- standard metadata

Metadata standards are standards about what fields should be filled in.

Directory of Metadata Standards

Controlled vocabularies and classifications of metadata tell you what to write in these fields, using standard terminology.

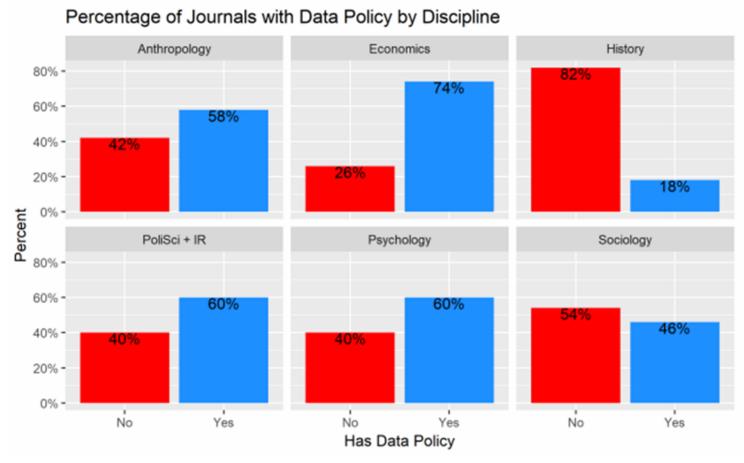


Data sharing and publication

- Data repositories
- Data Journals, peer-rev., e.g. Nature <u>Scientific</u>
 <u>Data</u>, <u>Open Health Data</u>
- Open Access peer-reviewed + open data, e.g. <u>PLoS One</u>. <u>An example</u>
- Check out the journal's data policy!

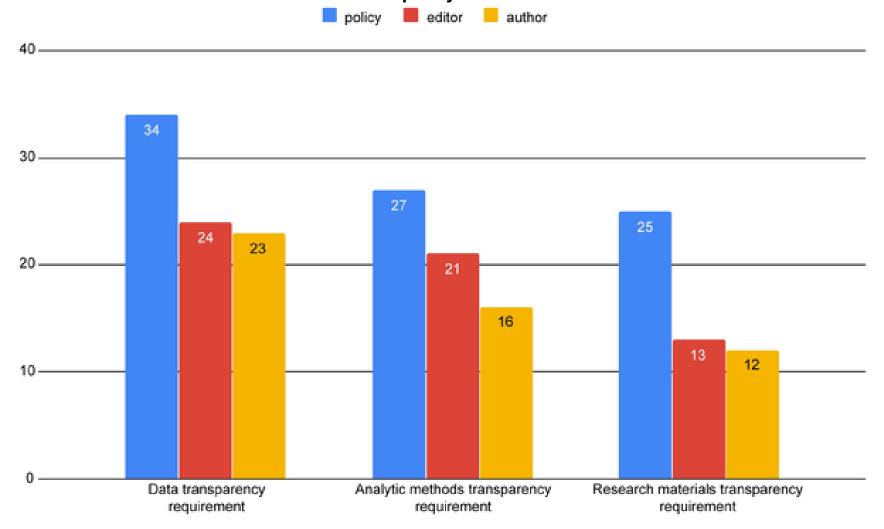
Data policies of highly-ranked social science journals

Figure 1: Data Policies by Discipline



Crosas, M., Gautier, J., Karcher, S., Kirilova, D., Otalora, G., & Schwartz, A. (2018, March 30). Data policies of highly-ranked social science journals.

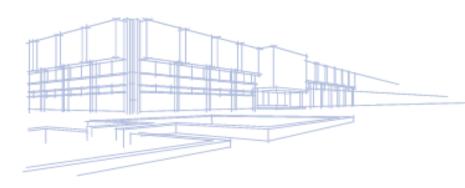
Fig 1. Comparison of editors' conception of the policy, authors' understanding of the policy, and the policy as written.



Christian TM, Gooch A, Vision T, Hull E (2020) Journal data policies: Exploring how the understanding of editors and authors corresponds to the policies themselves. PLOS ONE 15(3): e0220281. https://doi.org/10.1371/journal.pone.0220281

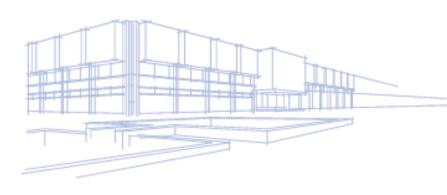
https://journals.plos.org/plosone/article?id=10.1371/journal.por .023 28

S ONE



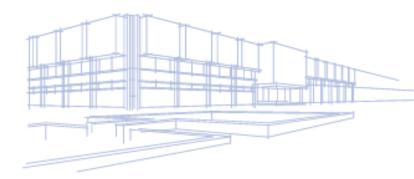
Journal policy requirements

- Submit data underlying article findings to a trusted repository
- Submit analytic methods (e.g., code, scripts, packages) to a trusted repository
- Submit research materials (e.g., codebook, readme files) to a trusted repository
- Explain access restrictions for data that cannot be shared due to legal or ethical reasons
- Describe the process for accessing data that cannot be shared due to legal or ethical reasons
- Other (Please specify) https://doi.org/10.1371/journal.pone.0230281



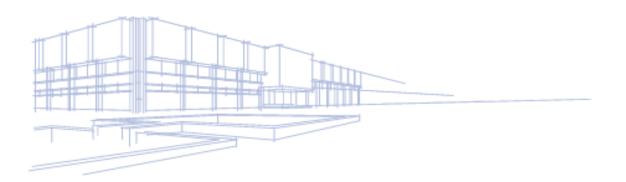
Research Data Repositories

- Goal: FAIR data
- DOI, linking (publications, <u>ORCID</u>)
- Long term storage ≈ 10 years
- Data Curation
- Embargo period, access management
- Metadata standards
- Licensing (CC 0, CC-BY)
- Statistics



Data Repositories

- Universal (Zenodo, Figshare, Dataverse)
- Disciplinary (ICPSR, Dryad, GenBank, GitHub)
- Institutional (UT DataDOI, Utrecht Yoda)
- National (UK Data Service, ANDS)
- Disciplinary repositories in Estonia, provide DOI:
 - PlutoF (biodiversity)
 - QsarDB (computational chemistry)
 - CELR (Center of Estonian Language Resources)
 - SARV (geocollections)



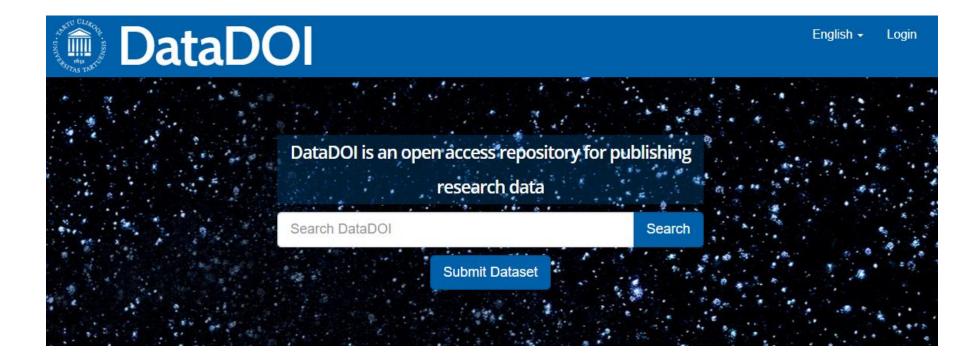
Re3data

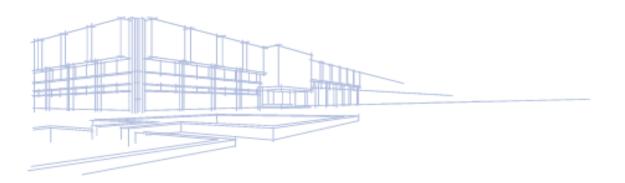
- Registry of research data repositories
- About 2500 repositories
- 41 filters to specify the search
- Find a repository → find the data



UT DataDOI

- Free for the University of Tartu researchers and Tallinn University
- DataCite DOI, indexed in DataCite
- Automatically indexed in OpenAIRE portal



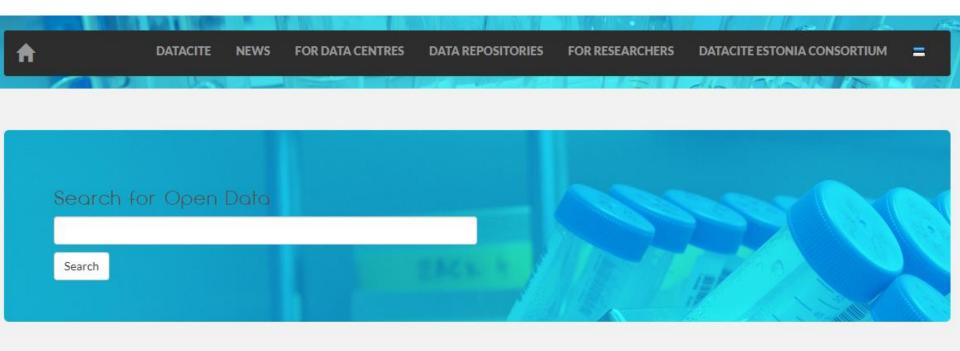


DataCite

- DataCite Estonia Consortium
- Provides DOIs for research data and registers the metadata
- Metadata schema (M; R; O)
- ≈ 18 million datasets
- Citation formatter
- Repository finder



DataCite Estonia









re3data tutorial

Citation Formatter



- Supports Open Access, Open Research Data and Open scholarly communication policy in Europe, building e-infrastructure and European Open Science Cloud (EOSC)
- Linked projects, publications, data, funders
- Horizon 2020 publications and data + FP9
- Zenodo repository











PARTICIPATE

SEARCH

MC



¬ CATCH-EyoU: Public Authorities engaging with youth: Survey Data: Estonia: Wavel



SHARE - BOOKMARK













Noack, Peter, Macek, Petr, Kalmus, Veronika; Rämmer, Andu; Beilmann, Mai; Kello,

Katrin; Kõuts, Ragne; Opermann, Signe (2017)

Publisher: Tartu Ülikool

Languages: English

Type: dataset

Identifiers: 3 doi:http://dx.doi.org/10.15155/repos-6

The data set was generated within the research project Constru European Youth: Policies, Practices, Challenges and Solutions Horizon 2020 Programme http://www.catcheyou.eu Grant Agreen

DOWNLOAD FROM

Tartu Ülikool

□ DataDOI



♠ DataDOI / Socialia / Eesti Sotsiaalteadustik Andmearhiiv (ESTA). Uuringute andmestikud / CATCH-EyoU / N

subjects: youth, European citizenship, surveys, identity, civic and p CATCH-EyoU: Public Authorities engaging with youth: Survey Data: Estonia: Wave1

Noack, Peter; Macek, Petr; Kalmus, Veronika; Rämmer, Andu; Beilmann, Mai; Kello, Katrin; Kõuts, Ragne; Opermann, Signe

≛ Export +

31

Nimi	Suurus	Kirjeldus
Abstract docx	11.95Kb	
WP7 survey methodology_Estonia.pdf	90.18Kb	
Estonian_Questionnaire_Older_Group_ET.pdf	4.279Mb	
Estonian_Questionnaire_Younger_Group_ET.pdf	4.281Mb	
CatchEyou_WP7_surveyWave1_dataset_Estonia.por	778.3Kb	CatchEyou WP7 survey Wave1 dataset Estonia
README_CatchEyou_WP7_surveyWave1_Estonia.pdf	296.9Kb	README file for CatchEyou WP7 survey Wave1 Estonia

A dataset in UT DataDOI can be found in OpenAIRE portal



Kuupäev 2017

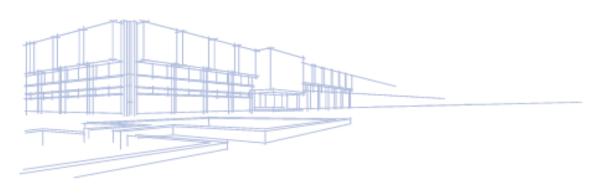
URI

http://dx.doi.org/10.15155/repos-6 http://datadoi.ut.ee/handle/33/57

Kokkuvöte

The data set was generated within the research project Constructing AcTive CitizensHip with European Youth: Policies, Practices, Challenges and Solutions (CATCH-EyoU) funded by European Union, Horizon 2020 Programme - Grant Agreement No 649538. The data set is a truncated version of the adolescents' and young adults' survey that was carried out in Estonia from October to December 2016. It merges results of two polls (15-19) and 20-30 year olds). Survey was conducted by University of Tartu (UT) wi_Rohkem

youth, European citizenship, surveys, identity, civic and political participation



EOSC

- European Open Science Cloud (EOSC)
- Free virtual environment for researchers
- For data storage, managemant and reuse
- Consolidate and connect existing research data e-infrastructures
- Federated search

The Backbone of EOSC



- UT guide: <u>Data Management Plan</u>
- UT PhD course materials: Research Data management and Publishing
- DataCite Estonia Consortium
- OpenAIRE Guides on RDM
- RDM services are provided by universities' libraries, repositories and archives

Service

Databases

Open Science

About us

Conference Centre

₩ E-bookstore

Home / Data Management Plan...

Open Science

Open Science at UT

Data Management Plan (DMP)

- > DMPonline
- › Examples of Data Management Plans
- Data Management support in UT

DataCite Estonia

OpenAIRE and Horizon 2020

RDA

EODOPEN

Events

Data Management Plan (DMP)

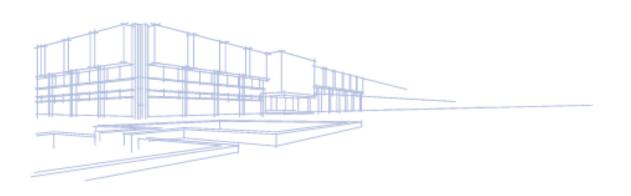
When planning research, it is important to consider carefully and document the ways of collecting and processing data during the research project, to specify who has access to these data and who is responsible for them, what will happen to the data after the closure of the project, etc. In order to do all this, it is necessary to create the data management plan and to follow it throughout the project.

DATA COLLECTION

- Acquiring data
- > Data description
- > Data formats
- > Data volume
- > How will the data be collected or created
- > Software
- > Organization of data

DOCUMENTATION AND METADATA

- Data documentation
 - · use this guide for data documentation:
 - Siiri Fuchs, & Mari Elisa Kuusniemi. (2018, December 4). Making a research project understandable - Guide for data documentation (Version 1.2). Zenodo. DOI: http://doi.org/10.5281/zenodo.1914401



Thank you!

Contact tiiu.tarkpea@ut.ee