DAMAP demo

Cluster Forschungsdaten Expo: January 16, 2025

www.damap.org

Agenda

- Intro *-10 minutes*
  - Background
  - Goals of the session

- Hands on demo of DAMAP *-30 minutes*

- Post-demo *-10 minutes*
  - Customizing DAMAP
  - Questions

Background

- DAMAP was created as part of the FAIR Data Austria project and a partnership between research data centers at TU Wien and TU Graz.

- We are funded as part of the Shared RDM Services and Infrastructure project through the research data cluster.

- Shared RDM goals include:
  - Developing a critical base infrastructure for research data management and also concepts for cross-university operating models for shared RDM services and infrastructures.
  - Ensuring a leading role of Austrian universities for shared RDM services in an international context.
  - Establishing networking with European and international initiatives, projects, and infrastructures.
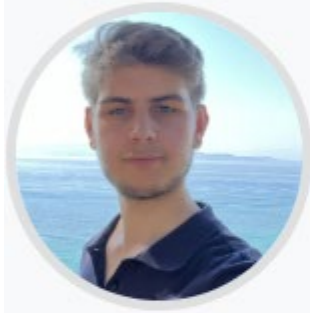
DAMAP
Team

Andres Tabima

Christiane Stork

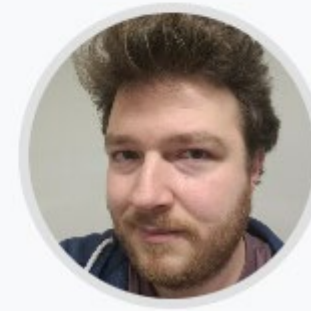Derek Molnar

Geoffrey Karnbach

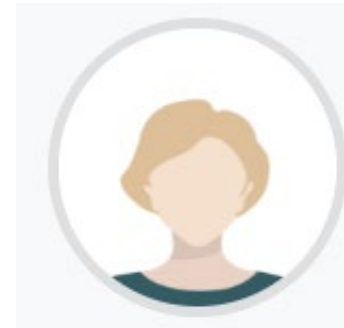Ilire Hasani-Mavriqi

Laura Thaci

Tomasz Miksa

Valentin Futterer

Zeno Casellato

SHARED RDM

Tell us about yourself

- What institution are you coming from?

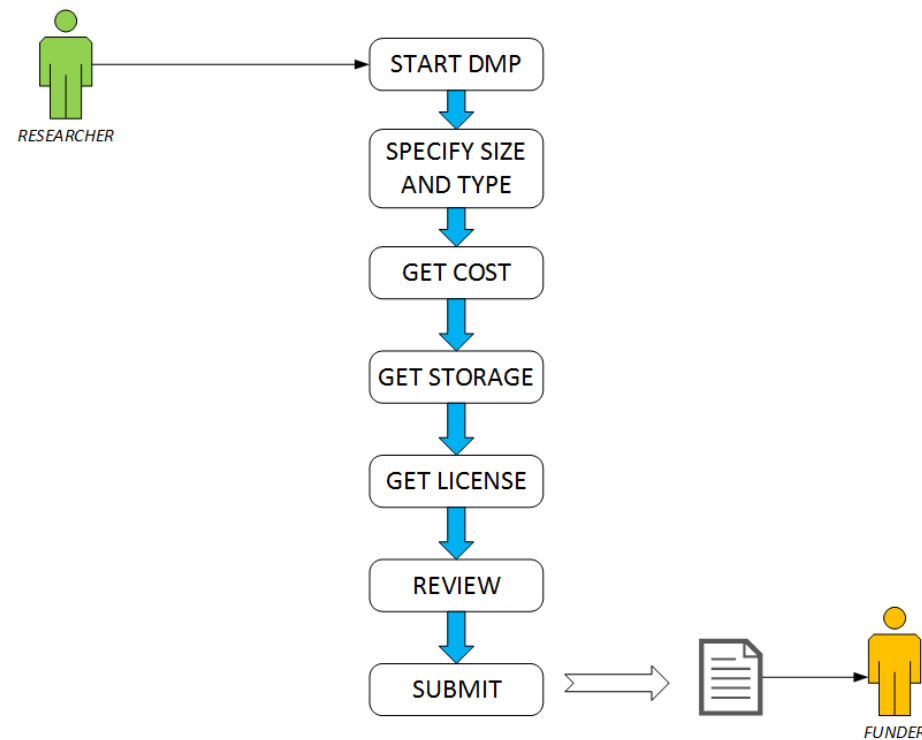- What is your experience with DMPs?

- Have you used other DMP tools?

Goals of the session

By the end of the session you will know:

- What a DMP tool is and what it should do

- What DAMAP is and why you would use it

- How to use DAMAP to begin creating a DMP

What a Data Management Plan (DMP) Tool is

- DMP tools guide researchers through the creation of a DMP

What a DMP tool should do

- Make it easier for researchers to create an accurate, high quality DMP

- Integrate and reuse information from internal and external systems (including Current Research Information Systems -CRIS)

- Be built on the principle of machine actionability to ensure future adaptability
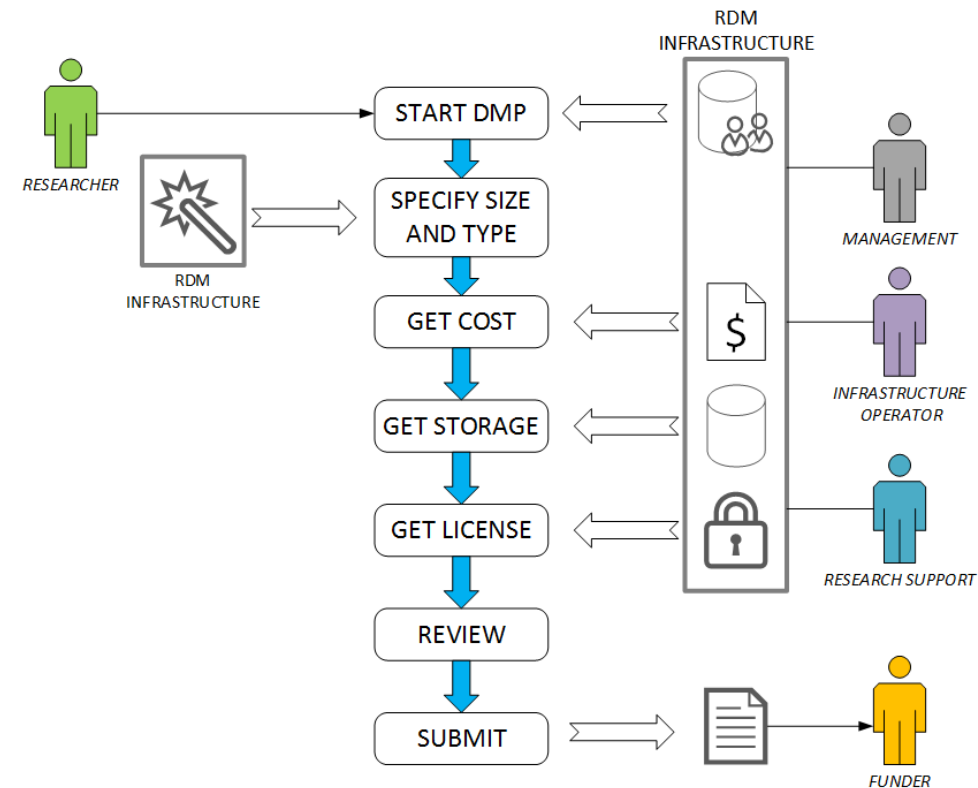
What DAMAP is

DAMAP is a community-based, open-source platform for creating machine-actionable DMPs.

www.damap.org

How we built it: maDMP

- DAMAP was developed with the RDA recommendations for maDMPs in mind, i.e., we focus on modelling information, not questionnaires.

How we built it: Science Europe practical guide

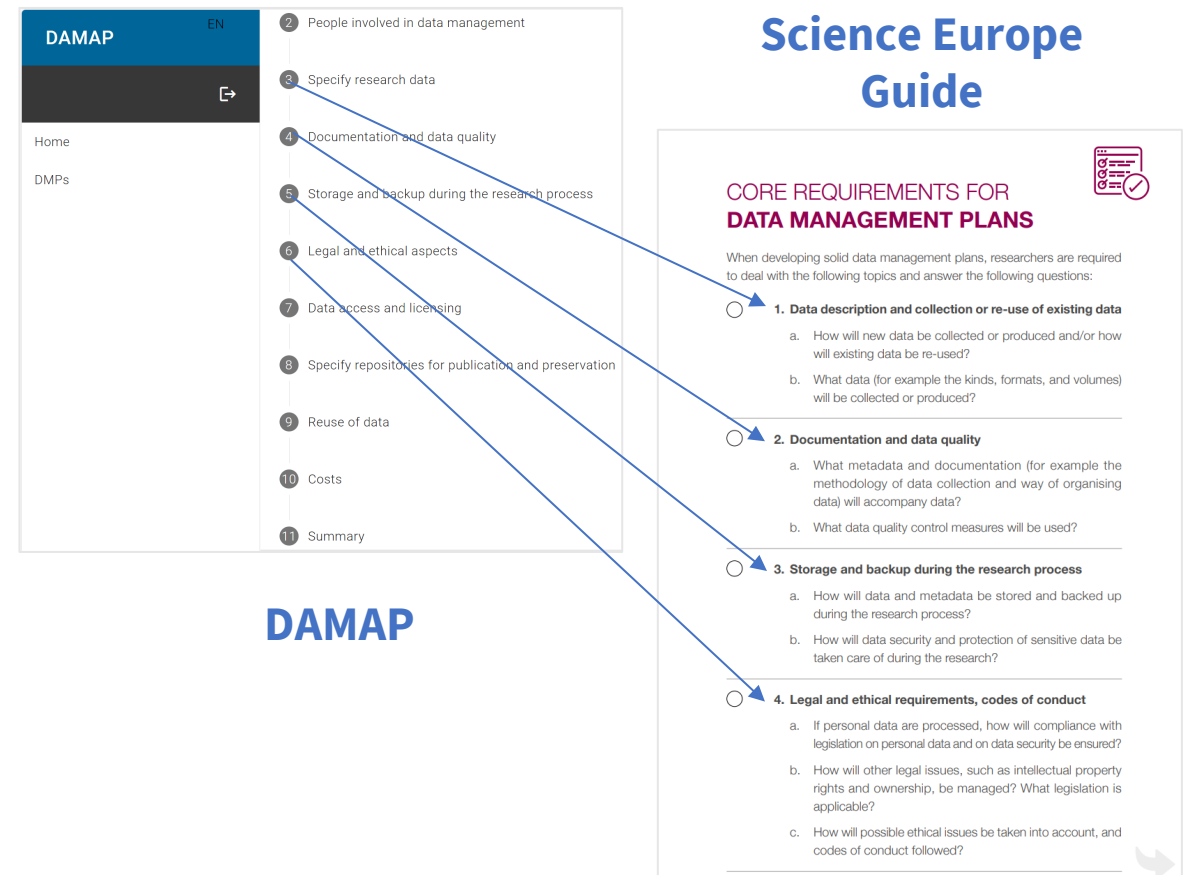- DAMAP aligns directly to an established standard for DMPs: the Science Europe practical guide.

GUIDANCE FOR ORGANISATIONS:
CORE REQUIREMENTS FOR
DATA MANAGEMENT PLANS

GUIDANCE FOR ORGANISATIONS:
CRITERIA FOR THE SELECTION OF
TRUSTWORTHY REPOSITORIES

GUIDANCE FOR RESEARCHERS:
Translating the Core Requirements into a DMP template
Guiding the Selection of Trustworthy Repositories

GUIDANCE FOR REVIEWERS:
Evaluation Rubric for Data Management Plans

SCIENCE EUROPE

PRACTICAL GUIDE TO
THE INTERNATIONAL ALIGNMENT OF
RESEARCH DATA MANAGEMENT

Extended Edition
with DMP Evaluation Rubric

SCIENCE
EUROPE

Science Europe. (2021). Practical Guide to the International Alignment of Research Data Management - Extended Edition. https://doi.org/10.5281/zenodo.4915862

How we built it: Science Europe practical guide (cont.)

DMP topics:

1. Data description and collection or re-use of existing data

2. Documentation and data quality

3. Storage and backup during the research process

4. Legal and ethical requirements, codes of conduct

5. Data sharing and long-term preservation

6. Data management responsibilities and resources

Includes:

- guiding questions and explanations for researchers

- evaluation rubric for funders: "comply or explain"

DAMAP

Science Europe Guide

Why would you use DAMAP?

- Requires less effort from researchers
  - Reuses information from **internal** and **external** systems
  - Uses "Comply or explain" principle
- Improves the quality of DMPs through:
  - Step-by-step guidance
  - Cross checking and validation
  - Providing recommended services
- Makes DMPs actionable
  - Integrated with other systems, e.g., your institution's CRIS
- Meets funders' goals
  - Accurate, high-quality DMPs
  - Science Europe and Horizon Europe compliant

Will personal data be collected/used as part of the project?
Option currently selected in the CRIS system: No.
⦿ Yes
◯ No
Please select the datasets containing personal data:
☐ A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration ☑ testdaten interview ☐ Test 2.pdf

How will compliance with data protection be ensured?

☐ by gaining informed consent for processing personal data
☐ by anonymisation of personal data for preservation and/or sharing (truly anonymous data are no longer considered personal data)
☐ by pseudonymisation of personal data (the main differences with anonymisation is that pseudonymisation is reversible)
☐ by encryption of personal data (the encryption key must be stored separately from the data, for instance by a trusted third party)
☐ other measures

Demo: How to use DAMAP to create a DMP

DAMAP - a tool for machine actionable DMPs

▶ Start now   📡 Introduction

GUIDANCE, EXPORT, AUTOMATION & COMPATIBILITY

https://DAMAP.org
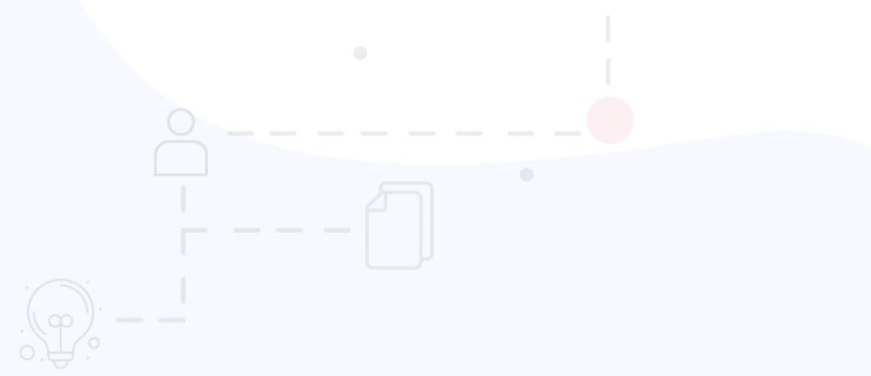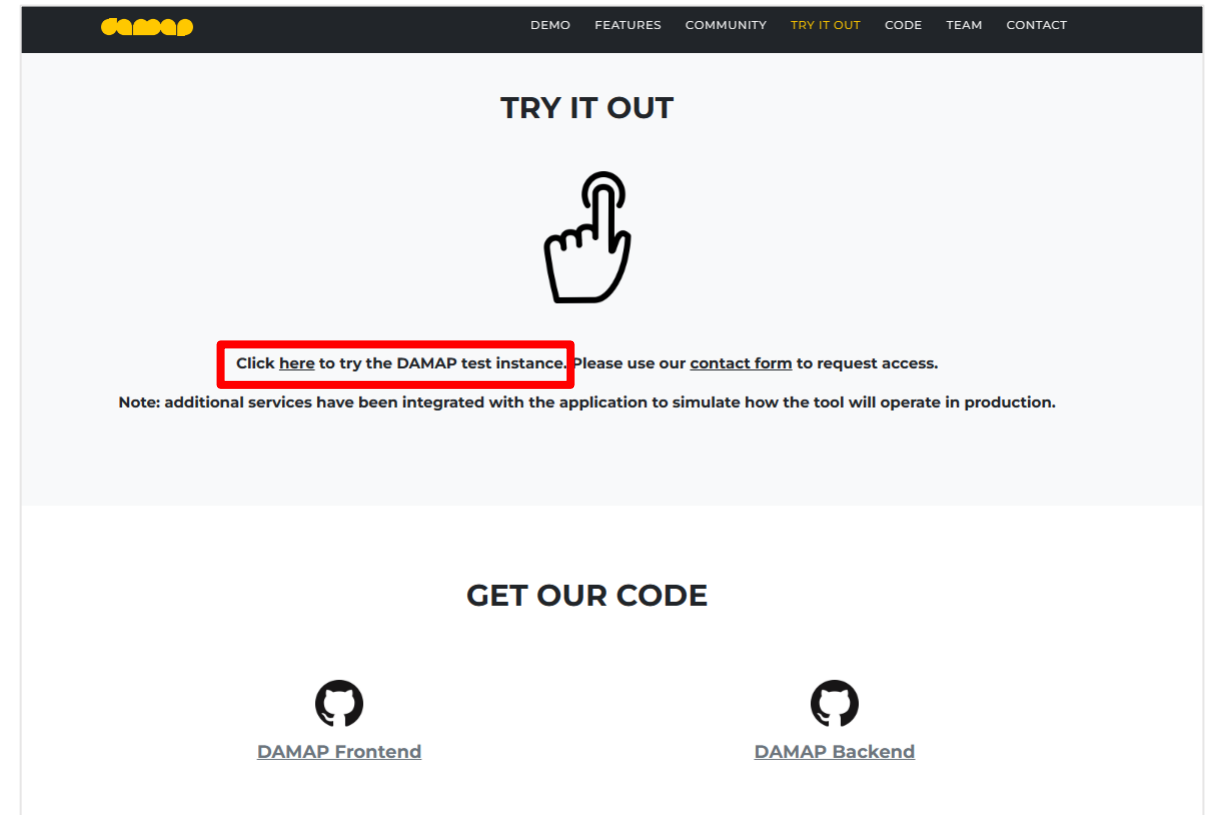
DEMO BEGIN
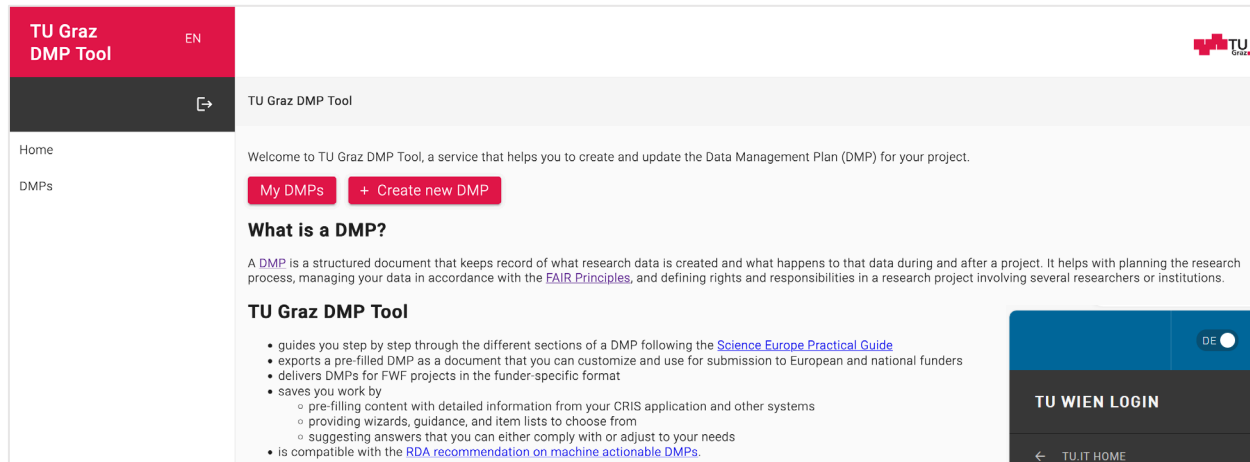
How to Login

- Go to DAMAP.org

- Scroll to "Try it out" section

- Click the "here" hyperlink

- Use the credentials on your card

Logging in to DAMAP -customization examples

Step 1: Choosing a project -instructions

- Create a new DMP

- Input project manually

- Enter project information:
  - Title: Forschungsdaten Demo Project
  - Start: 1st June 2025
  - Duration: 12 months
  - Description: be creative ☺

- Update project
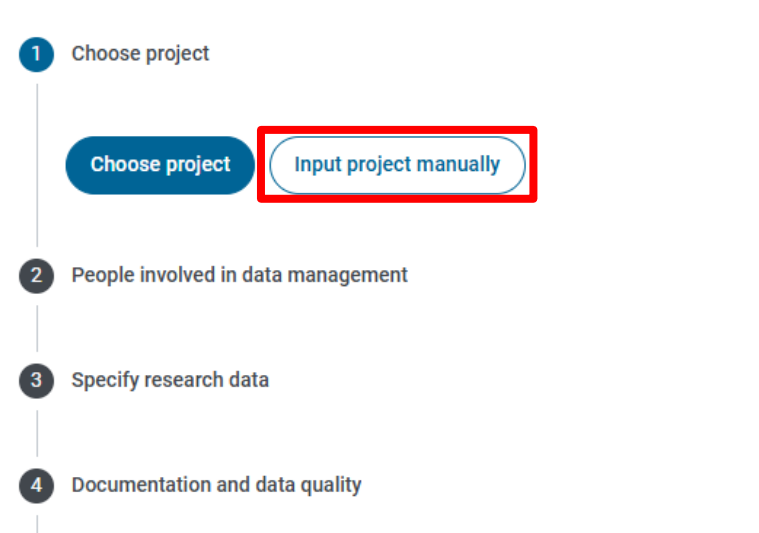
Step 1: Choosing a project -customization examples

- What you would see if you connected DAMAP to your project database

Step 2: People involved in data management -instructions

- Add person via ORCID
  - Name: Tomasz Miksa
  - Change role to Supervisor
- Add person via manual contributor
  - Name: yourself
  - Change your role to Data Manager
  - Add your email
  - Mark yourself as a contact
- Save changes
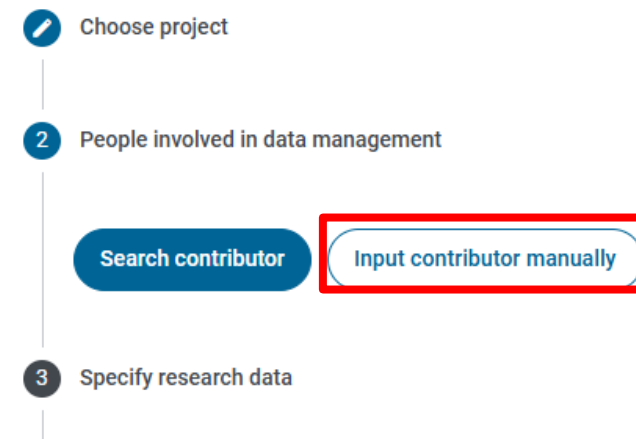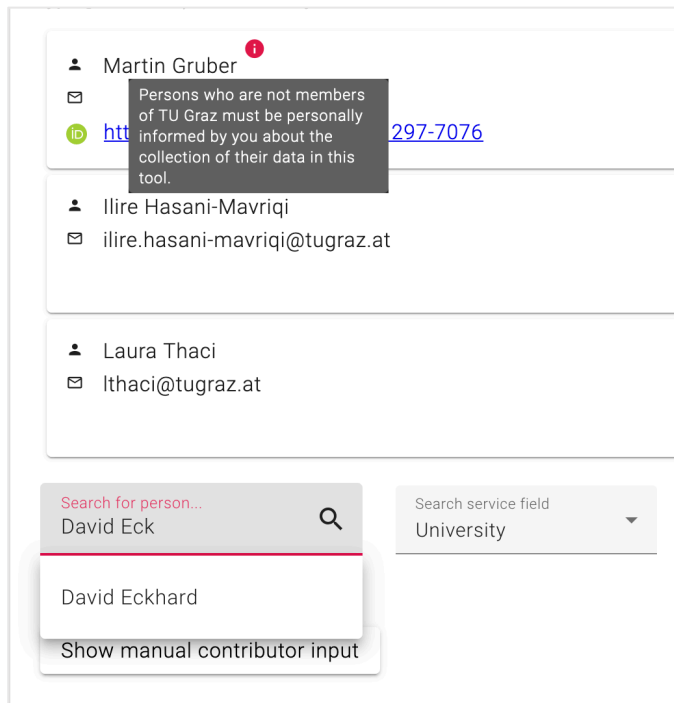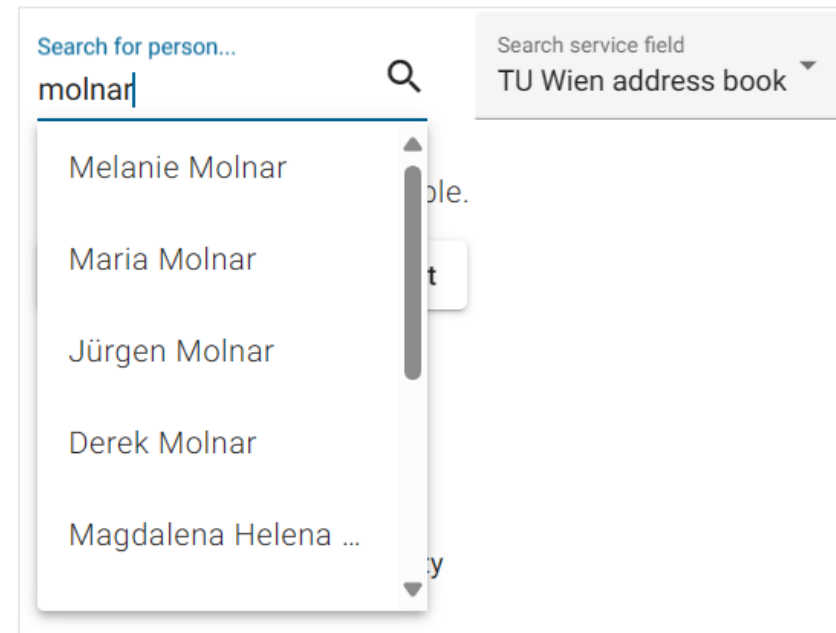
Step 2: People involved in data management -customization examples

- What you would see if you connected DAMAP to your HR database

Step 3: Specifying the research data -instructions

- Add <u>new datasets</u> according to the following information:

Name	Type	Size	Sensitive data	Deletion	Storage during research	Repository	License
Interview recordings	Audiovisual data	1-5 GB	yes	31.03.26 by Tomasz Miksa	Institutional Cloud	no	-
Anonymised interview transcripts	Standard office document	100-1000 MB	no	no	Institutional Cloud	Zenodo	CC BY 4.0
Analysis software	Software application	< 100 MB	no	no	Institutional GitLab	GitHub/ Zenodo	MIT
Interpreted results	Standard office document	< 100 MB	no	no	Institutional Cloud	Zenodo	CC BY 4.0

- Add <u>a reused dataset</u>: doi:10.5281/zenodo.6430412

Step 4: Documenting data organization and quality -instructions

- How will you structure the data and handle versioning?
  - The respective work package leader…

- Please indicate which metadata will be provided to help others identify, discover and reuse the data.
  - As there are no domain specific…at **project** level (remove other text)

- How will you provide documentation needed to validate data analysis and facilitate data reuse?
  - We will provide documentation needed to validate data analysis and facilitate data reuse in accompanying README files

- How will the consistency and quality of data collection be controlled?
  - Data entry validation
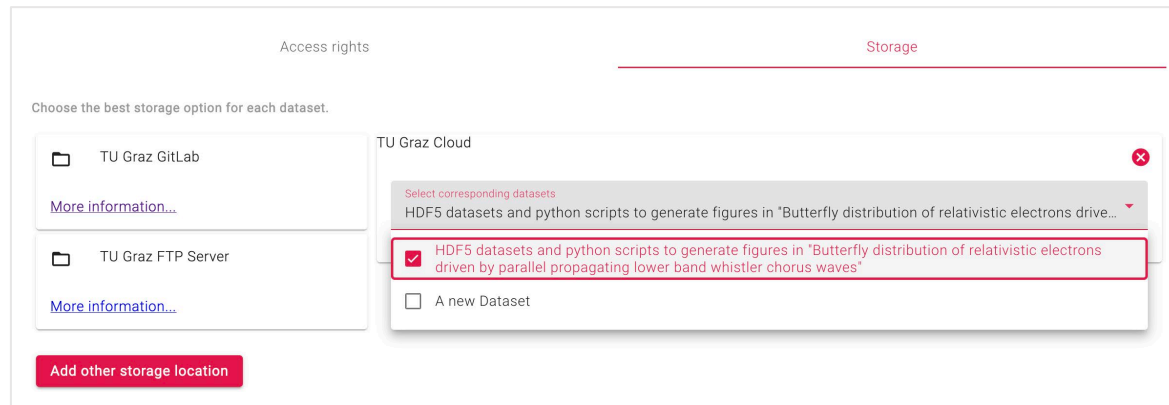 - Representation with controlled vocabularies

Step 5: Storage and back-up during the research process: -instructions

- Select Storage, add other storage location with the following info

Name	Type	Size	Sensitive data	Deletion	Storage during research	Repository	License
Interview recordings	Audiovisual **data**	< 5 GB	yes	31.03.25 by Tomasz Miksa	Institutional Cloud	no	-
Anonymised interview transcripts	Standard office document	< 1 GB	no	no	Institutional Cloud	Zenodo	CC BY 4.0
Analysis software	Software application	< 100 MB	no	no	Institutional GitLab	GitHub/ Zenodo	MIT
Interpreted results	Standard office document	< 100 MB	no	no	Institutional Cloud	Zenodo	CC BY 4.0

Step 5: Storage and back-up during the research process -customization examples

- Recommended storage locations can be configured to your institution's repositories

Step 6: Legal and ethical aspects -instructions

Legal Aspects

- Is your data or part of it sensitive…? **Yes**

- Please select the datasets containing sensitive data: "**interview recordings"**

- Will personal data be collected/used as part of the project? **Yes**

- Please select the datasets containing personal data: "**interview recordings"**

- How will compliance with data protection be ensured? **By gaining informed consent… and Anonymization for preservation**

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

Ethical Aspects

- Will you involve human participants in the project…? **Yes**

- Beyond the use of personal data, are there any other ethical issues associated…? **No**

- Was your research plan reviewed by an ethics committee…? **No**

Step 6: Legal and ethical aspects -integration example

- DAMAP shows what was selected in the project database for your institution

Will personal data be collected/used as part of the project?

Option currently selected in the CRIS system: No.

◉ Yes

◯ No

Step 7: Data access and licensing -instructions

- Assign the licenses to your datasets based on the following info:
  - Note: Data Access for the "Interview recordings" data set will be "Closed" and you will need to mark the dataset for deletion.

Name	Type	Size	Sensitive data	Deletion	Storage during research	Repository	License
Interview recordings	Audiovisual data	< 5 GB	yes	31.03.25 by Tomasz Miksa	Institutional Cloud	no	n/a
Anonymised interview transcripts	Standard office document	< 1 GB	no	no	Institutional Cloud	Zenodo	CC BY 4.0
Analysis software	Software application	< 100 MB	no	no	Institutional GitLab	GitHub/ Zenodo	MIT
Interpreted results	Standard office document	< 100 MB	no	no	Institutional Cloud	Zenodo	CC BY 4.0

Step 8: Specifying repositories for publication and preservation -instructions

- Select the relevant repositories

Name	Type	Size	Sensitive data	Deletion	Storage during research	Repository	License
Interview recordings	Audiovisual data	< 5 GB	yes	31.03.25 by Tomasz Miksa	Institutional Cloud	no	-
Anonymised interview transcripts	Standard office document	< 1 GB	no	no	Institutional Cloud	Zenodo	CC BY 4.0
Analysis software	Software application	< 100 MB	no	no	Institutional GitLab	GitHub	MIT
Interpreted results	Standard office document	< 100 MB	no	no	University Cloud	Zenodo	CC BY 4.0

Step 8: Specifying repositories for publication and preservation -customization examples

Repositories

Repositories are used for sustainable data management beyond the lifetime of the project and to make data discoverable and reusable to others. If available, use a domain specific repository for your data. The search filter below may help you to find one.

Recommended | Find repository

TU Graz Repository

An institutional repository at Graz University of Technology to enable storing, sharing and publishing research data, publications and open educational resources. It provides open access services and follows the FAIR principles.

Content types: Raw data | Software applications | Source code | Standard office documents
Metadata standards: DataCite Metadata Schema
Languages: DEU | ENG
https://repository.tugraz.at/

Repositories

Repositories are used for sustainable data management beyond the lifetime of the project and to make data discoverable and reusable to others. If available, use a domain specific repository for your data. The search filter below may help you to find one. Otherwise, you can always use TU Wien Research Data, TU Wiens institutional data repository.

Recommended | Find repository

TU Wien Research Data

TU Wien Research Data is an institutional repository of TU Wien to enable storing, sharing and publishing of digital objects, in particular research data. It facilitates the funders' requirements for open access to research data and the FAIR principles by making research output findable, accessible, interoperable, and reusable. A DOI is assigned to each dataset published in TU Wien Research Data. This service is developed by the TU Wien Center for Research Data Management and hosted by TU.it.

Content types: Raw data | Scientific and statistical data formats | Source code | Standard office documents
Metadata standards: DataCite Metadata Schema | Dublin Core
Languages: ENG
https://researchdata.tuwien.at/

Use repository

30

Step 9: Detailing your reuse of data -instructions

- Who is the target audience and who could be interested in reusing the data and why?
  - Students and general public

- Indicate whether potential users need specific tools or software to access and (re)use the data
  - The analysis software is provided in GitHub

Step 10: Costs -instructions

- Select: "as outlined below…"

- Add Cost

- Name: Change from "New cost" to "DAMAP Demonstration"

- Estimated cost: you decide ☺

- Cost type: Training

10 Costs

In this step you are asked to provide a realistic estimation of ne

Select the statement which best describes the situation in

○ There are no costs dedicated to data management.

⦿ As outlined below, there are costs dedicated to data

⊕ Add cost
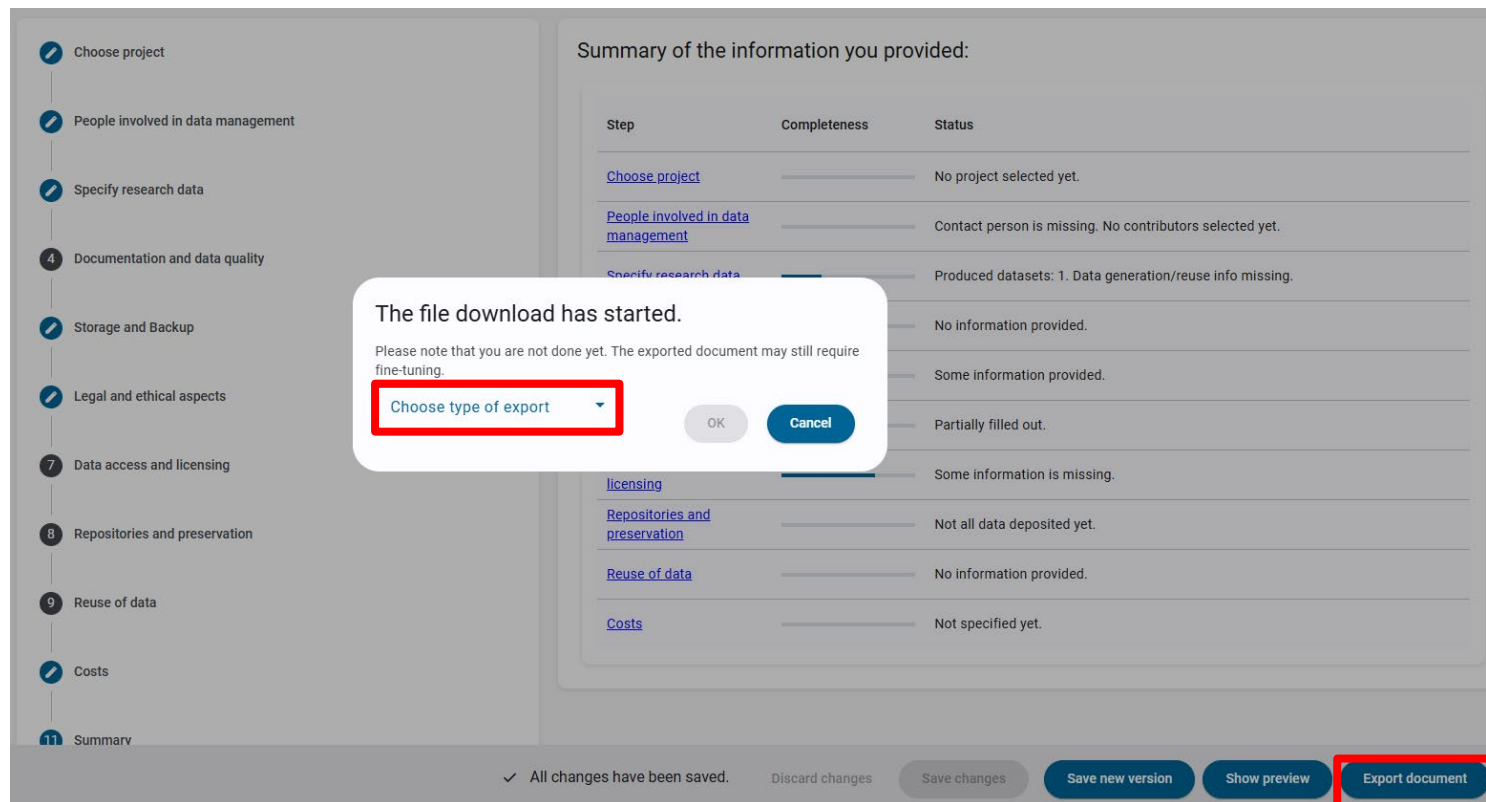
Step 11: Summary

- Review the ten steps to see if any information is missing

- As needed, click on any of the sections to provide additional info

- If you made any changes, click "Save changes"

33

Exporting the document -instructions

- Click on the "Export Document" button

- Select Science Europe

Comparing exports and machine actionability

```
"contributor" : [ {
  "contributor_id" : {
    "identifier" : "0000-0002-5164-2690",
    "type" : "orcid"
  },
  "mbox" : "moritz.staudinger@tuwien.ac.at",
  "name" : "Moritz Staudinger",
  "role" : [ "Data Manager" ]
```

DEMO END

Customizing DAMAP for your institution

Simple changes:

- Colors for look and feel of tool
- Default suggested answers
  - IT Storage options
  - Export document policy
- Default text in template
- ORCID, re3data (list of data repository), OpenAire (for reused data sets)
- FITS (use to upload a sample file, automatically tells you the format)

More complex changes:

- Integration with your institutions' CRIS and other systems, e.g. HR and project databases.
- Setting up Single Sign On (SSO)

Revisiting the goals of the session

By the end of the session you will know:

- What a DMP tool is and what it should do
  - 1) Make it easier for researchers to create an accurate, high quality DMP
  - 2) Integrate and reuse information from internal and external systems
  - 3) Be built on the principle of machine actionability to ensure future adaptability
- What DAMAP is and why you would use it
- How to use DAMAP to begin creating a DMP

Questions?

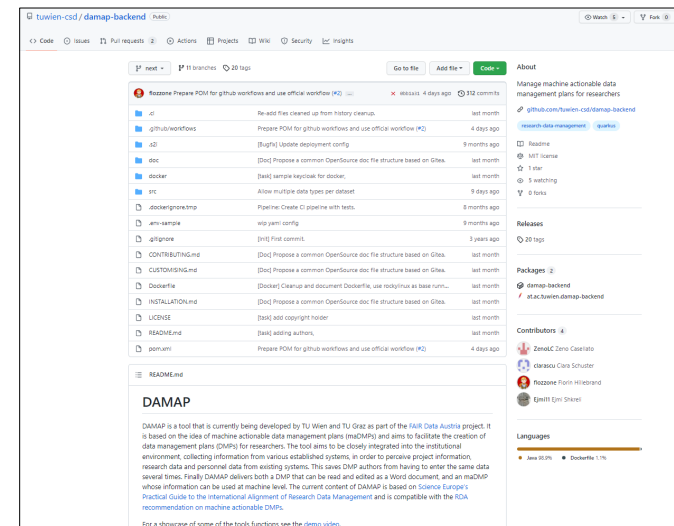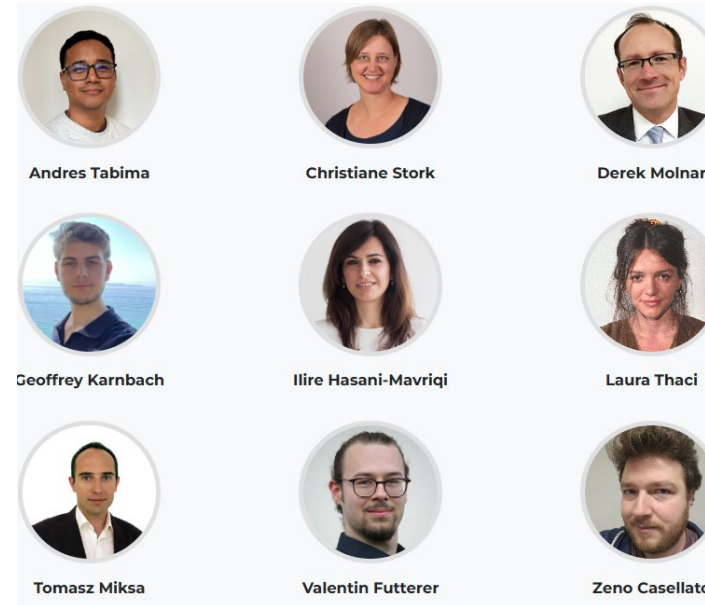# **Stay in Touch!**

Contact us at: info@DAMAP.org

Code:

tuwien-csd/damap-frontend (github.com)
tuwien-csd/damap-backend (github.com)

www.damap.org

END