Love Data Week

EPFL 2022 edition

Wed. Feb.16 | 11:30-12:00
Webinar "Data Storage vs. Backup vs. Archive: EPFL services"

Thu. Feb. 17 | 11:00-11:45
Online training "Which platforms to publish my research data/code?"

Fri. Feb. 18 | 10:00-12:00
Online speed-dating with EPFL Data Champions

go.epfl.ch/ldw22
Data Management Plan... Not only grant applications

- Data collection/reuse
- Data preprocessing
- Data processing/analysis
- Write paper
- Paper submitted to journal
- Data curation
- DMP update?
- DMP publication?
- Data REUSE (others / yourself)
- Discovery (others / yourself)
- Data / Code Preservation
- Paper & Data PUBLICATION

Source: “Research Data Management for junior researchers”, S.Dürr, J.Pidoux, F.Varrato, CC-BY-4.0
Many flavors

SNSF (various grant schemes + NCCR, …)
- Mandatory DMP
- Researchers must share (at least) the data underlying their publications

ERC (H2020 + MSCA + Horizon 2021-2027...)
- Mandatory DMP
- The research data needs to be open by default

U.S. Federal Grants + Wellcome Trust + …

EPFL
- Internal projects or compliancy
- Not mandatory across the board
Now the EPFL template exists

```
go.epfl.ch/epfl-dmp
```

+ find it at

```
go.epfl.ch/rdm-guide
```
Not in a vacuum ...

SNSF DMP template

ERC DMP template.s (H2020 + MSCA)

Research Office (REO)
Open Science Unit (OSU)
Data Protection Officer (DPO)
Scientific Information and Libraries (SISB = Library ☺)

+ EPFL Data Champions
# 1. PERSONAL INFORMATION (M)

1a. Fellow's HPI is a mandatory

1b. Fellow's HPI

1c. Fellow's HPI

# 2. PROJECT

2a. Project Title

2b. Type of F

2c. Institution

2d. Laboratory

2e. Supervisors

2f. Starting date

# 3. INTRODUCTION (M)

3a. Could any of the data/code that you create, produce, collect or process contain personal data? (Y)

3b. Could any of the data that you an

3c. Could any of the data that you collect contractually obliged?

3d. Will you put requirement, be

# 4. DATA CODE THAT WILL BE COLLECTED, PRODUCED AND PROCESSED

**PART 1: DATA CODE THAT WILL BE COLLECTED, PRODUCED AND PROCESSED**

Please use the guidelines to fill in the table. (M = Mandatory / MA = Mandatory if applicable / R = Recommended)

<table>
<thead>
<tr>
<th>P1a. # (M)</th>
<th>P1b. Centers personal data? (R)</th>
<th>P1c. Collects personal data? (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td>Y</td>
</tr>
</tbody>
</table>

**PART 2: UNDERSTANDABILITY OF COLLECTED, PRODUCED OR PROCESSED DATA**

This part is about how you make your datasets understandable for you and others.

Please use the guidelines:

| P2a. For whom are you e.g. for yourself, for a colleague |
| P2b. What documentation needs to be provided? (MA) |
| P2c. How you will ensure e.g. choosing processes, peer review |

**PART 3: DATA CODE MANAGEMENT INFRASTRUCTURE**

Please use the guidelines to fill in the table. (M = Mandatory / MA = Mandatory if applicable / R = Recommended)

<table>
<thead>
<tr>
<th>P3a. Data Code Management System (M)</th>
<th>P3b. Stores</th>
</tr>
</thead>
<tbody>
<tr>
<td>External SSO drive</td>
<td></td>
</tr>
<tr>
<td>DMP (repositories)</td>
<td></td>
</tr>
<tr>
<td>Data (Knowledge)</td>
<td></td>
</tr>
</tbody>
</table>

**PART 4: DATA CODE LONG-TERM PRESERVATION**

Please use the guidelines to fill in each field. (M = Mandatory / MA = Mandatory if applicable / R = Recommended)

<table>
<thead>
<tr>
<th>P4a. # (MA)</th>
<th>P4b. What will be preserved? (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Part of the dataset which is useful for future projects</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

**ANNEX 1: "AS OPEN AS POSSIBLE": RESEARCH DATA CODE (ORD)**

This part is intended for use when datasets are published to an ORD platform/repository. We are aware that not all datasets you create can be made available. You can at least strive for publication of datasets at the end of the project. For more details, please check guidelines. (M = Mandatory / MA = Mandatory if applicable / R = Recommended)

**ANNEX 2: FORM FOR A DATA PROTECTION ANALYSIS**

Need help? Please contact the EPFL DPO: dpo@epfl.ch or chiara.tannen@epfl.ch

**Part 1: Datasets collected/produced**

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1a. What is the intended purpose?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1b. How many participants will be included?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2b. Data Subjects: do you consider only Swiss residents? UE residents? Others?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2c. Have you got the consent (explicit consent in case of sensitive data)?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Part 2: The 9 criteria to evaluate if we face to "likely to result in a high risk" problem**

- The current FADP does not require a Data protection impact assessment (DPIA). For this reason, if you do a DPIA, you also need to evaluate if there is a high risk. If you do this, the DPO will have to sign off on it.

**ANNEXES**

Data Life Cycle planning
Some use-cases

No need
- If a funder already has a DMP template (ex. SNSF, ERC)

Expected use of the EPFL DMP template
- EPFL services, but are not forced to (ex. REO, HREC, TTO, etc.)
- Any EPFL funder (School, Institute, Platform, …)
- Any researcher willing to streamline data management for a project

REO already used it for a couple Interdisciplinary Seed Fund calls
Support by EPFL Library

To researchers

- Info & Suggestions & Guiding material
- Draft review (all necessary rounds)
- Answer to specific RDM questions
- Put in touch with other services (HREC, REO, TTO,...)
- Accompany & Updates

To other EPFL services

- Integrate, depending on needs
- To researchers (Info + Review + Questions + …)
- Provide feedback / Reporting

REO already used it for a couple Interdisciplinary Seed Fund calls
A file is nice, but …

Ongoing pilot

- EPFL DMP template version on digital platform
- STI + other faculties can join the pilot

dmponline.dcc.ac.uk
Ongoing collaborations

- Got precious input from early experiences
- Expecting increased use
- Will evolve during 2022 (no disruptive changes)
- *How-to guide* will be published during 2022

- An institutional DMP template might become a future norm (?)
- Planning and documenting the data management is / should be (?) the norm

Research Office (REO)
Open Science Unit (OSU)
Data Protection Officer (DPO)
Scientific Information and Libraries (SISB = Library 😊)

+ EPFL Data Champions
+ Researchers input
Questions & Answer...

... Discussion
Love Data Week

EPFL 2022 edition

Wed. Feb. 16 | 11:30-12:00
Webinar "Data Storage vs. Backup vs. Archive: EPFL services"

Thu. Feb. 17 | 11:00-11:45
Online training "Which platforms to publish my research data/code?"

Fri. Feb. 18 | 10:00-12:00
Online speed-dating with EPFL Data Champions

go.epfl.ch/ldw22