Keep calm and fill in your DMP
Lessons learnt from a Swiss DMP-template initiative

Eliane Blumer, EPFL Library
Fantin Reichler, EPFL Library

IDCC2018, Barcelona, 21st February 2018
2 Swiss Federal Institutes of Technology:

- Ecole Polytechnique Fédérale de Lausanne (EPFL)
- Eidgenössische Technische Hochschule Zürich (ETHZ)
Context
National initiative: Data Life-Cycle Management

WWW.DLCM.CH
Funding agencies requirements

Starting at submission date October 2017, the **Swiss National Science Foundation** (SNSF) requires to provide a full dynamic DMP

The SNSF DMP form includes four sections:

1. Data collection and documentation
2. Ethics, legal and security issues
3. Data storage and preservation
4. Data sharing and reuse

Data underlying a publication must be share ASAP
Concept and Design

EPFL Preparation work
Basic structure of template

May 2017

Finding internal and external partnerships
- Internal: working group established and led by the library
  - EPFL Library
  - Research Office (ReO)
  - Technology Transfer Office
  - Vice Presidency for Information Systems
- External: ETH Library

Feedback and input from all DLCM partners

Realisation of templates and supplementary tools

Sept 2017

Workshops@EPFL and ETH Library
+ promotion on respective websites

RESEARCHDATA.EPFL.CH

ETHZ.CH/RESEARCHDATA
DMP template

Main structure:

I. Question and guidelines from SNSF
II. Recommendations from the library
III. Examples taken from actual DMPs (anonymised)
IV. Contact points

+ supplementary materials (data publication decision tree & recommended file format document)
Supplementary tools

List of EPFL recommended file formats

<table>
<thead>
<tr>
<th>TYPE OF DATA</th>
<th>Sub type</th>
<th>Preferred file formats for sharing, reuse and preservation</th>
<th>Other Acceptable formats (up to 10 year)</th>
<th>Not suitable for preservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset</td>
<td>Tabular data with extensive metadata</td>
<td>comma-separated values (CSV) file, Unicode UTF-8 (.csv) with CSV on the Web descriptive metadata. Hierarchical Data Format version 5 (HDF5)</td>
<td>plain text data, ASCII (txt) HyperText Mark-up Language (HTML) (.html)</td>
<td>MS Excel (.xls, .xlsx)</td>
</tr>
<tr>
<td></td>
<td>Tabular data with minimal metadata</td>
<td>comma-separated values (CSV) file (.csv) tab-delimited file (.tab) OpenDocument Spreadsheet (.ods) Structured Query Language (SQL) dump, preferably from an open tool (PostgreSQL, MariaDB)</td>
<td>Extensible Mark-up Language (XML) text according to an appropriate Document Type Definition (DTD) or schema (.xml)</td>
<td>MS Excel (.xls, .xlsx)</td>
</tr>
<tr>
<td>Test</td>
<td>Textual data</td>
<td>PDF/A (.pdf) plain text, Unicode UTF-8 (.txt) Open Document Text (.odt, .odm) LaTeX (.tex)</td>
<td>widely-used proprietary formats, e.g. MS Word (.doc), PowerPoint (.pptx)</td>
<td>MS Word (.doc), PowerPoint (.ppt)</td>
</tr>
<tr>
<td></td>
<td>Code</td>
<td>Plain text formats (e.g. Matlab/Octave .m, R project .R, Python .py, and so on) Jupyter Notebook (.ipynb) Rstudio (.rstudio), Rmarkdown (.rmd) NeoCDF</td>
<td>Text files for S-plus (.sdd) Matlab .mat, should be saved in HDF format.</td>
<td>Matlab (.mat), R-project (.rdata)</td>
</tr>
</tbody>
</table>
Workshop formats

EPFL Library
• 1.5h on the lunch hour: general overview of the data management and the SNSF DMP (discipline specific)
+ tailored trainings

ETH Zurich
• ETH-wide + specific information events
• 1.5h comprehensive training
+ tailored training
Workshop results

- EPFL: 8 workshops in 2017 (68 participants) and 4 workshops in 2018 (2 given so far, 55 participants and counting)

- ETH Zurich: 434 part. since 2017 (all kind of RDM services)
Benefits & limitations

⇒ Improvement of libraries’ services visibility

⇒ Libraries reaffirm their role as centres of competences regarding data management

⇒ Institutional, National & International Collaborations:
  • Research Office, Technology Transfer Service, Vice Presidency for Information System
  • ETH Zürich, UNIGE, Inist-CNRS, Cambridge University

⇒ BUT difficult to find out the exact number of researchers benefitting from the service
Next steps

- Continue to adapt and further develop DMP-services
  - EPFL Library: specific trainings for selected labs
  - ETH Zurich: new webinars

- Interactive version of the template

- More discipline-oriented services