From planning to implementation: data management in real-time research

Montreal, March 31st, 2018

Alexandra Stam and Brian Kleiner
FORS, Swiss Centre of Expertise in the Social Sciences
**FORS**

Swiss Centre of Expertise in the Social Sciences

- Methodological research
- Large-scale surveys
- Data and research information services (DARIS)

**DARIS**

<table>
<thead>
<tr>
<th>Data service</th>
<th>Data enhancement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data archiving</td>
<td>Data management</td>
</tr>
<tr>
<td>New requirements</td>
<td>Training</td>
</tr>
<tr>
<td>Long-term preservation</td>
<td>Consulting</td>
</tr>
<tr>
<td>Enhance the value of</td>
<td></td>
</tr>
<tr>
<td>research projects</td>
<td></td>
</tr>
</tbody>
</table>

**Data access**
- Direct access to:
  - + 500 data sets
  - + 11’000 project descriptions
Data management at DARIS

Early days: Focus on DM from a data service point of view

- Early 2010’s: increasing awareness of the importance of data management
- Follow the flow

Current days: Focus on DM from the researcher’s point of view

- Need for more concrete guidelines and solutions
- Focus on ‘day-to-day’ data management
- It is important to think beyond the declaration of intent (DMP)
- It is important to think in practical terms (operationalisation)
- Data management represents an opportunity
2017: our DM story takes a new turn

- June: we start a pilot with the University of Lausanne to focus on day-to-day data management
- October: Our main funder, the Swiss National Science Foundation (SNSF) requires DMPs with all proposals

  Strong reactions from the research community (‘one more administrative burden’)

  A great opportunity for us

  *How to use the DMP (and forced-interest in data management) to pass on our messages?*
Our initial intention

- Develop one good DMP example
- Include post-funding guidance

What we did:

- Provided assistance with DMPs
- Organised and participated in DMP events

What we learned:

- There is no such thing as one good example
- There are various shortcomings with the current template
- Even reluctant researchers can get to appreciate data management (plans)

⇒ Instead of one good example, provide a reflexive model
Content of the SNSF DMP

1. Data collection and documentation
   - What data will you collect, observe, generate or re-use?
   - How will the data be collected, observed or generated?
   - What documentation and metadata will you provide with the data?

2. Ethics, legal and security issues
   - How will ethical issues be addressed and handled?
   - How will data access and security be managed?
   - How will you handle copyright and intellectual Property Rights issues?

3. Data storage and preservation
   - How will your data be stored and backed-up during the research?
   - What is your data preservation plan?

4. Data sharing and reuse
   - How and where will the data be shared?
   - Are there any necessary limitations to protect sensitive data?
   - [checkbox: I will choose digital repositories conform to the FAIR data principles]
   - [Yes/No button: I will choose digital repositories maintained by a non-profit organisation]
Shortcomings of the SNSF DMP

1. Data collection and documentation
   - What data will you collect, observe, generate or re-use?
   - How will the data be collected, observed or generated?
   - What documentation and metadata will you provide with the data?

2. Ethics, legal and security issues
   - How will ethical issues be addressed and handled?
   - How will data access and security be managed?
   - How will you handle copyright and intellectual property rights issues?

3. Data storage and preservation
   - How will your data be stored and backed-up during the research?
   - What is your data preservation plan?

4. Data sharing and reuse
   - How and where will the data be shared?
   - Are there any necessary limitations to protect sensitive data?
   - [Checkbox: I will choose digital repositories conform to the FAIR data principles]
   - [Yes/No button: I will choose digital repositories maintained by a non-profit organisation]

- It suggests a chronological order

Data collection → Data sharing
## Shortcomings of the SNSF DMP

1. **Data collection and documentation**
   - What data will you collect, observe, generate or re-use?
   - How will the data be collected, observed or generated?
   - What documentation and metadata will you provide with the data?

2. **Ethics, legal and security issues**
   - How will ethical issues be addressed and handled?
   - How will data access and security be managed?
   - How will you handle copyright and intellectual property rights issues?

3. **Data storage and preservation**
   - How will your data be stored and backed-up during the research?
   - What is your data preservation plan?

4. **Data sharing and reuse**
   - How and where will the data be shared?
   - Are there any necessary limitations to protect sensitive data?
   - [Checkbox: I will choose digital repositories conform to the FAIR data principles]
   - [Yes/No button: I will choose digital repositories maintained by a non-profit organisation]

- It suggests a chronological order
- Administrative feel
Shortcomings of the SNSF DMP

1. Data collection and documentation
   - What data will you collect, observe, generate or re-use?
   - How will the data be collected, observed or generated?
   - What documentation and metadata will you provide with the data?

2. Ethics, legal and security issues
   - How will ethical issues be addressed and handled?
   - How will data access and security be managed?
   - How will you handle copyright and intellectual property rights issues?

3. Data storage and preservation
   - How will your data be stored and backed-up during the research?
   - What is your data preservation plan?

4. Data sharing and reuse
   - How and where will the data be shared?
   - Are there any necessary limitations to protect sensitive data?
   - [Checkbox: I will choose digital repositories conform to the FAIR data principles]
   - [Yes/No button: I will choose digital repositories maintained by a non-profit organisation]

- It suggest a chronological order
- Administrative feel
- ‘One size fits all’
Shortcomings of the SNSF DMP

1. Data collection and documentation
   - What data will you collect, observe, generate or re-use?
   - How will the data be collected, observed or generated?
   - What documentation and metadata will you provide with the data?

2. Ethics, legal and security issues
   - How will ethical issues be addressed and handled?
   - How will data access and security be managed?
   - How will you handle copyright and intellectual property rights issues?

3. Data storage and preservation
   - How will your data be stored and backed-up during the research?
   - What is your data preservation plan?

4. Data sharing and reuse
   - How and where will the data be shared?
   - Are there any necessary limitations to protect sensitive data?
   - [Checkbox: I will choose digital repositories conform to the FAIR data principles]
   - [Yes/No button: I will choose digital repositories maintained by a non-profit organisation]

- It suggests a chronological order
- Administrative feel
- ‘One size fits all’
- Intention of good practices
Shortcomings of the SNSF DMP

1. Data collection and documentation
   - What data will you collect, observe, generate or re-use?
   - How will the data be collected, observed or generated?
   - What documentation and metadata will you provide with the data?

2. Ethics, legal and security issues
   - How will ethical issues be addressed and handled?
   - How will data access and security be managed?
   - How will you handle copyright and intellectual property rights issues?

3. Data storage and preservation
   - How will your data be stored and backed-up during the research?
   - What is your data preservation plan?

4. Data sharing and reuse
   - How and where will the data be shared?
   - Are there any necessary limitations to protect sensitive data?
   - [Checkbox: I will choose digital repositories conform to the FAIR data principles]
   - [Yes/No button: I will choose digital repositories maintained by a non-profit organisation]

- It suggest a chronological order
- Administrative feel
- ‘One size fits all’
- Intention of good practices
- Most important, it doesn’t provide true guidance
Towards a reflexive SNSF DMP

1 ← 2 ← 3 ← 4
Data collection  Data sharing

DMP

['?']
['?']
['?']

Social Sciences
How do we do that?

1. Data sharing and reuse
   - ....
   - ....

2. Data collection and documentation
   - ....
   - ....

3. Ethics, legal and security issues
   - ....
   - ....

4. Data storage and preservation
   - ....
   - ....

1. Bring sharing to the forefront
2. Revisit questions
3. Provide tips and guidance
4. Get people to think beyond the proposal
Data collection and documentation – an example

- What data will you collect observe, generate or re-use?
- How will the data be collected, observed or generated?
- What documentation and metadata will you provide with the data?

Reflexive DMP

- Will I / am I obliged to collect sensitive and personal data?
- Will I share (all or part of) my data?

- What type of re-use would be relevant for my data?
- What is my epistemological stand?
- What documentation is needed to re-use the data or reproduce findings?

Documentation: We will for example provide three case studies representing three disciplines: ethnography, qualitative sociology, and statistics

We will raise awareness about what needs to be done after the project gets funded, including project management
Conclusion and next steps

- It is possible to get people to engage seriously with DMPs and subsequent DM practices
- There is a need to approach DMPs in a reflexive way and address the right questions
- As next steps we will finalize our reflexive template
- We also aim to build a decision tree
Thanks for your attention!
alexandra.stam@fors.unil.ch
Natural sciences
Engineering
Mathematics
School of Law
Social Sciences
Humanities
...

I love Social Sciences
Content of the DMP

1. Data sharing and reuse
   - What data will you collect, observe, generate or re-use?
   - How will the data be collected, observed or generated?
   - What documentation and metadata will you provide with the data?

2. Ethics, legal and security issues
   - How will ethical issues be addressed and handled?
   - How will data access and security be managed?
   - How will you handle copyright and intellectual Property Rights issues?

3. Data storage and preservation
   - How will your data be stored and backed-up during the research?
   - What is your data preservation plan?

4. Data sharing and reuse
   - How and where will the data be shared?
   - Are there any necessary limitations to protect sensitive data?
   - [checkbox: I will choose digital repositories conform to the FAIR data principles]
   - [Yes/No button: I will choose digital repositories maintained by a non-profit organisation]