

# Social Science Research Data Management

---

2022-04-25 12:00

John D. Martin III ([john.d.martin.iii@unc.edu](mailto:john.d.martin.iii@unc.edu))

## OVERVIEW

### Decisions, decisions

#### Data lifecycle

- What is the lifecycle of your data?
- What constraints are your data subject to? Now? Future?
- What are the ethical, legal, and other policy considerations for your data?
- Who are we archiving, preserving, and curating the data for?
- What are the costs involved?

#### Sensitive data

- What tools are needed to protect personal identifying information?
- Do data managers need to be on your IRB?
- Can ingest of sensitive data be automated?

#### Documentation

- Metadata
- Policies
- Workflows
- Version control
- End-to-end reproducibility

## ACTION ITEMS and TOOLS

1. **Establish a repository of information for your team:** OSF <https://osf.io/>
2. **Develop a data management plan:** Data Management Plan Tool <https://dmptool.org/>
3. **Establish a repository for your tools:** GitHub <https://github.com>

## NOTES

- **Interview data** is difficult to deidentify because it always contains context. Consider this as you think about how to preserve data.
- **Social media data** is similarly difficult to deidentify due to context. It is also already available online, which makes it searchable.

## USEFUL LINKS and REFERENCES

Data Curation Lifecycle Model

<https://www.dcc.ac.uk/guidance/curation-lifecycle-model>

Open Science Research Training Handbook

<https://github.com/Open-Science-Training-Handbook>

Data management plans – Stanford Libraries

<https://library.stanford.edu/research/data-management-services/data-management-plans>

Write a data management plan – MIT Libraries

<https://libraries.mit.edu/data-management/plan/write/>

Write an effective data management plan – UCSD Library

<https://library.ucsd.edu/lpw-staging/research-and-collections/data-curation/data-management/data-management-plan.html>

Get started with GitHub <https://docs.github.com/en/get-started>

Getting started with git

<https://docs.github.com/en/get-started/getting-started-with-git>

Happy Git and GitHub for the user <https://happygitwithr.com/index.html>

Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059–1078.

<https://doi.org/10.1002/asi.22634>

Choudhury, S., Huang, C., & Palmer, C. L. (2020). Updating the DCC Curation Lifecycle Model. *International Journal of Digital Curation*, 15(1), 12.

<https://doi.org/10.2218/ijdc.v15i1.721>

Huang, C., Lee, J.-S., & Palmer, C. L. (2020). *DCC Curation Lifecycle Model 2.0: Literature Review and Comparative Analysis*. University of Washington.

<https://digital.lib.washington.edu/researchworks/handle/1773/45392>

Mayernik, M. S. (2016). Research data and metadata curation as institutional issues. *Journal of the Association for Information Science and Technology*,

67(4), 973–993. <https://doi.org/10.1002/asi.23425>

Poole, A. H. (2016). The conceptual landscape of digital curation. *Journal of Documentation*, 72(5), 961–986. <https://doi.org/10.1108/JD-10-2015-0123>

Shilton, K. (2012). Participatory personal data: An emerging research challenge for the information sciences. *Journal of the American Society for Information Science and Technology*, 63(10), 1905–1915.

<https://doi.org/10.1002/asi.22655>

Soderberg, C. K. (2018). Using OSF to Share Data: A Step-by-Step Guide.

*Advances in Methods and Practices in Psychological Science*, 1(1), 115–120.

<https://doi.org/10.1177/2515245918757689>