

# Identifying Opportunities for Inclusive Language in Research Data Management Support in Academic Libraries



Hannah C. Gunderman, Ph.D. (Twitter: @WhovianVeganGeo)

\*My pronouns are she/her\*

Research Data Management Consultant

Carnegie Mellon University Libraries

Empirical Librarians 2020: Knoxville, TN

# Why have this conversation?

Started with a Medium article

My positionality: a neurodiverse researcher with generalized anxiety disorder, working in the world of open science and RDM

# Open science is really scary y'all.



Darren L Dahly [Follow](#)

Apr 26, 2019 · 3 min read



I did something stupid last week. I publicly posted my data and code.

I've been living in a nightmare ever since, terrified by every email and notification, fearful it's someone pointing out how I screwed up.



A few thoughts:

1. So what if someone finds a mistake. It is not like nobody ever made a mistake before.
2. If you set up your pipeline to share from the start, then you will do it without thinking about it. We write our data every night to GitHub on a script. No big deal. Variable names are in the experiment script which sets up the MySQL table in which the data are stored.
3. So what if your code is not too legible. People will write their own if they have the data and want to check. Nobody reads anybody else's code.
4. I find OSF to be a hard tool because it is hard to automate things. Automation for me is the key to limiting mistakes.
5. Being open is not an all or none proposition. You can start slow, maybe just data, and advance as you get comfortable.
6. Reliability and openness go hand in hand. When your pipeline is reliable, it is easier to be open. Being open forces you to make your pipeline more reliable.

# Why have this conversation?

Serendipity on Twitter!

RDM education as condescending

Pushing unfit software onto researchers

My first consultations - what kind of language was effective?

# Intent is not to replace “best practices,” but unpack them

Federal mandates for data sharing and open data - certain standard practices are important!

Best practices can introduce more standardization


Use of the term “best practices” in multiple settings

# 3 areas of opportunity for reframing towards inclusivity

When to start implementing RDM techniques

Storage

Data sharing



“Best practices” may not truly  
capture all abilities,  
institutional access,  
socioeconomic statuses, etc.  
of our researchers we are  
supporting!

# When is the “best” time to implement RDM?

“Start as early as possible”

“Start from day one”

“Start before you collect data”

These are absolutely true - it *does* help to start early! But....

This language can frame it as though if a researcher *doesn't* start RDM early, it's too late to implement it later.

**Meet the researcher where they are at, and express that you are there to help them.**



# Storage (3-2-1 rule)

3 copies of your data, with 2 copies on different storage media, with one located offsite

Institutions, labs/research groups, and individual researchers may have uneven access to these resources!

This can pose a significant financial burden to researchers!

We can't assume all researchers can take on this financial burden, particularly considering the competitiveness of grants

In consultations, try to understand what the researcher has access to, supplementing any gaps with potential library services to help support their backup strategy

# Data Sharing

Data sharing is compulsory with many funding agencies and journals

But - it can be scary to put data out there!

What if I made a mistake?

What if someone misinterprets it?

What if someone finds issues in my code?

**Acknowledge nervousness or fears around data sharing! Frame a README file as a way to address some of those fears. Don't dismiss the researcher's emotions!**

# Conclusions

We suggest using “good practices” or “recommended practices”

Meet the researchers where they are at!

Inclusive, open language can create more trust, continued partnerships

Language for consultations at each stage of the research data lifecycle in Inclusive RDM Toolkit