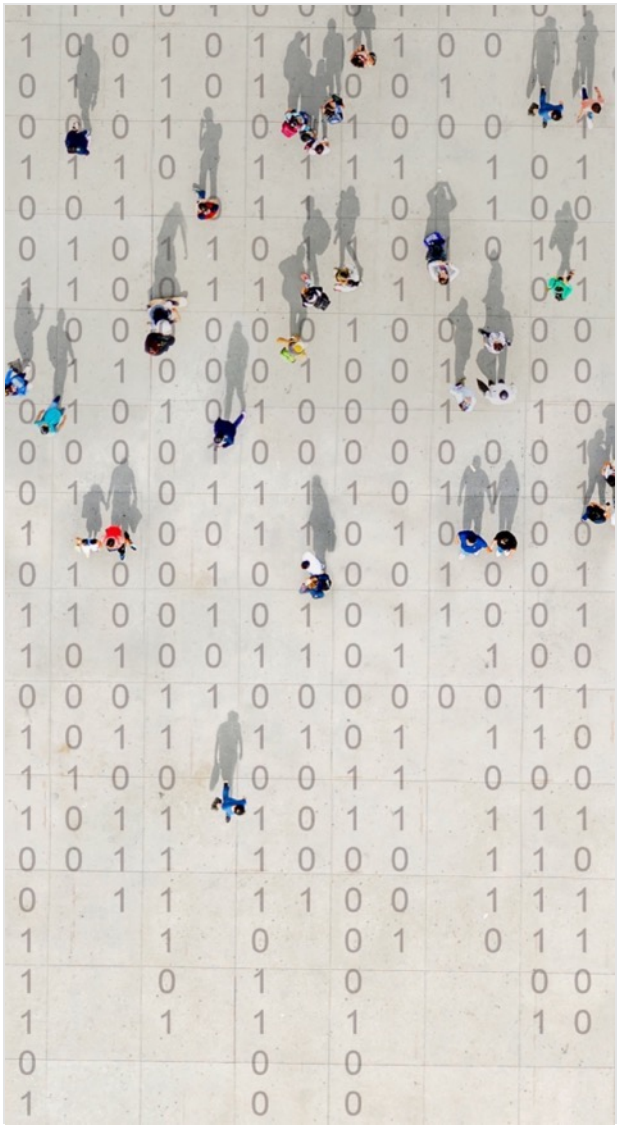
The value of open data

A meta-research(er) perspective

AGU/CHORUS Forum: How open is open data and software?
April 19, 2023

Kathleen Gregory
University of Vienna
Scholarly Communications Lab, University of Ottawa



Today’s talk

Focus data sharing/reuse

Three questions (‘provocations’)

Three responses

So what can be done?

Provocation 1:

Does ‘openness’ depend on context and demand flexibility?

No need to wait for an answer:

A resounding yes!

See for sharing data

Openness depends on community, norms, infrastructures, data, types research, legal requirements…

*We have to be very careful sharing the raw data, otherwise we **can be sued by the lab or by the university**…In terms of the clinical data (in Portugal), **the situation is different**…it is not that formal as in the United States. It is **quite frequent to share clinical data** with other researchers.*

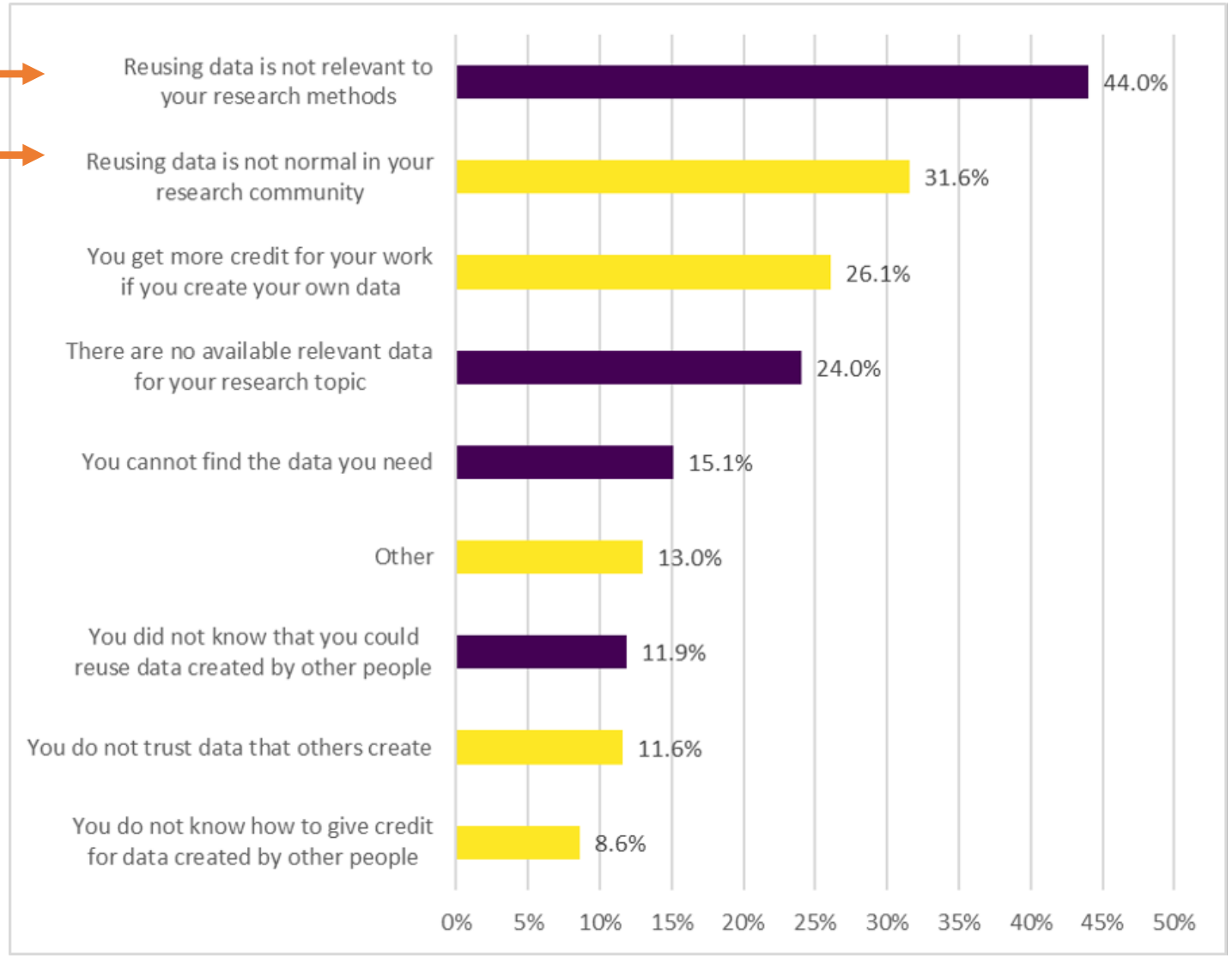
Gregory et al., 2019

See for reusing data

Re-using data is not relevant for all

And challenges exist

Why do you not reuse data?



Gregory, Ninkov, Ripp, Roblin, Peters, Haustein, 2023

Weighted percentages of respondents to this question (n=466). Items with a significant disciplinary difference in purple. Multiple responses possible.

So what?

A chance to get creative

Recognize the spectrum of research

- Flexible ‘snap-in’ policies
- Sharing ‘openly’ or reusing data are not always relevant or possible

Provocation 2:

What is our responsibility around attribution and credit?

Wait a minute…

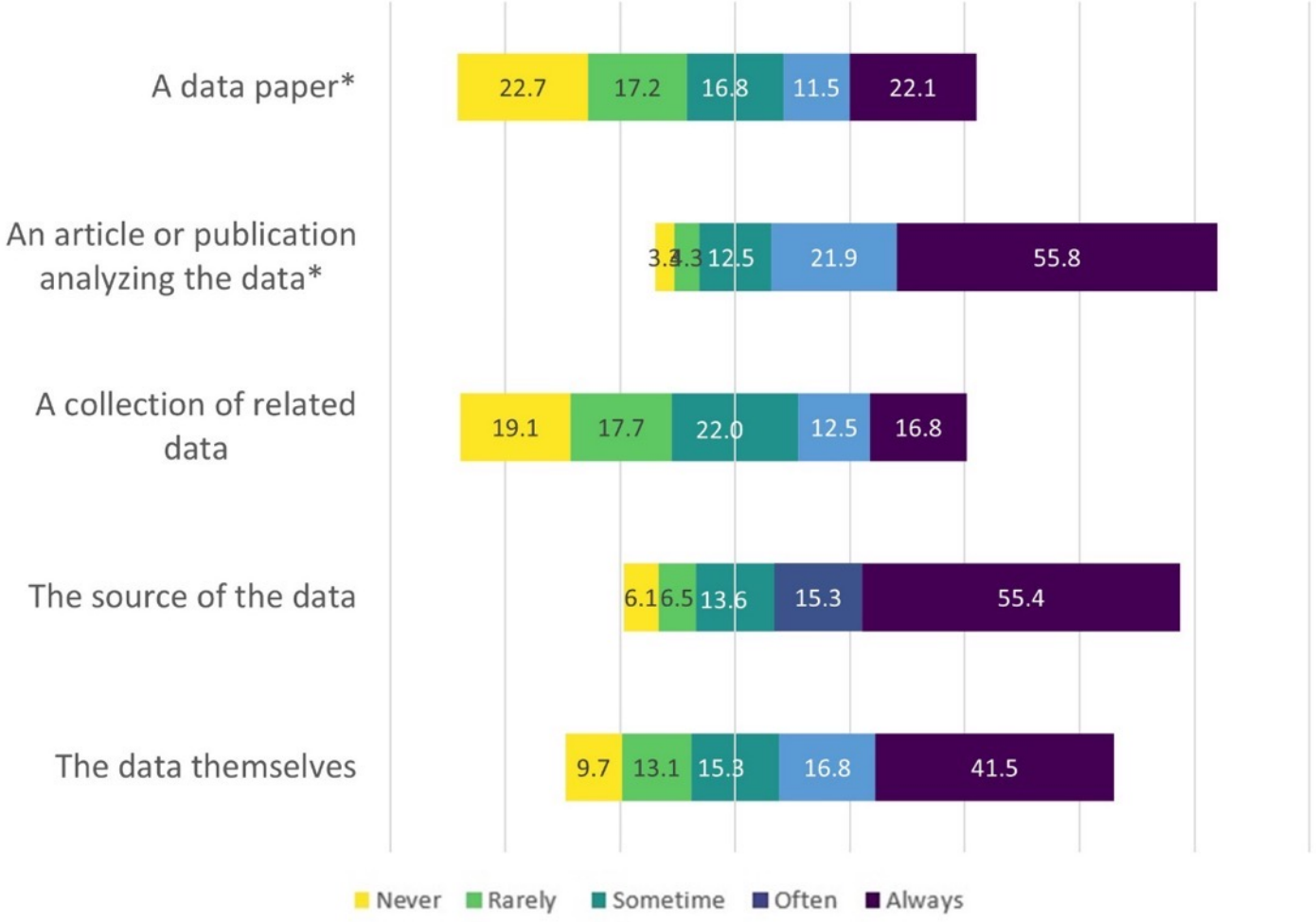
Are researchers already ’responsibly’ attributing data use?

Citing data

Majority of data re-users report citing!

Variety in practice

What data objects do you cite or reference?



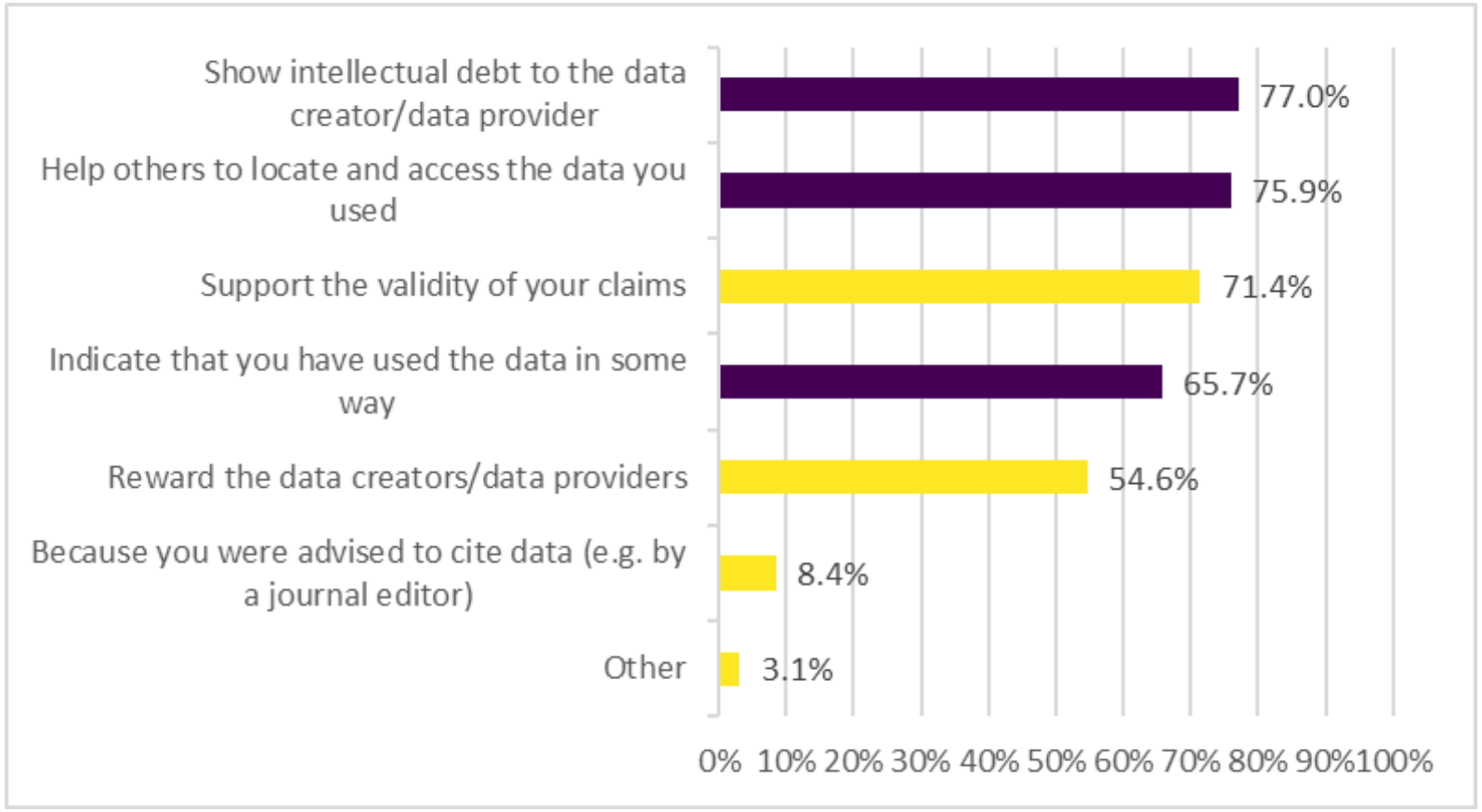
Gregory, Ninkov, Ripp, Roblin, Peters, Haustein, 2023

Weighted percentages of respondents to this question (n=2,026). Significant disciplinary differences marked with asterisk.

Motivations

Reasons reflect ideal (responsible?) research practice

Why do you *cite* data?



Gregory, Ninkov, Ripp, Roblin, Peters, Haustein, 2023

Weighted percentages of respondents to this question (n=2,026). Items with a significant disciplinary difference in purple. Multiple responses possible.

So what?

Build on existing practices

Data citation practices

- Encourage standardized data citation
- But meet researchers where they are

Provocation 3:

What is the value of

data management and data sharing ?

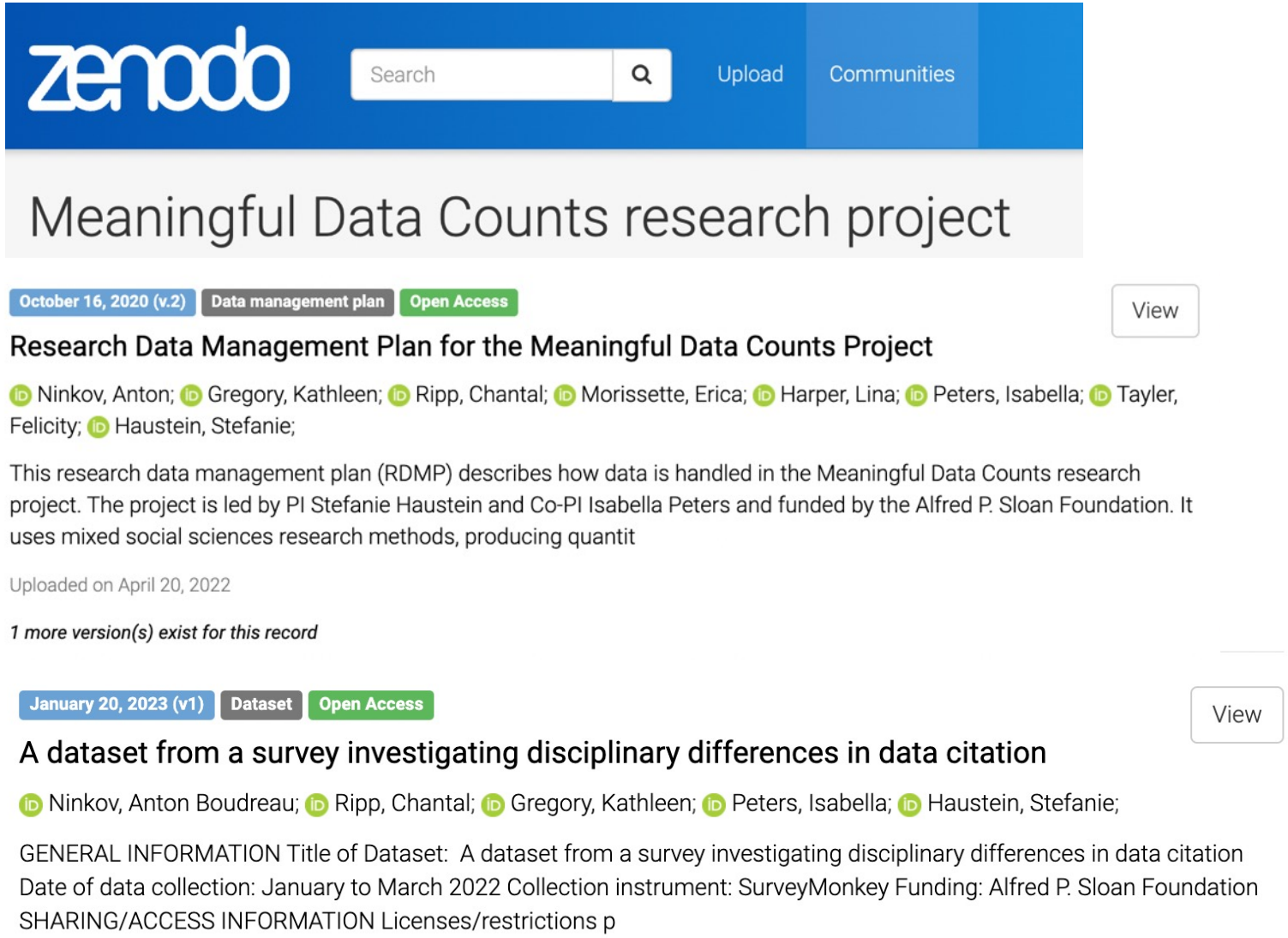
The value of **Slow (Reflexive) Science**

Wait a minute – isn’t faster science the goal?

This takes work.

Detailed, tedious

Lots of questions – also an opportunity



https://zenodo.org/communities/meaningfuldatacounts/

So what?

This is really tricky.

Data management takes time and work.

Existing steps (e.g. DMPs) as a check-in not a checkbox.

Thank you.

Questions/Discussion?

Kathleen Gregory
@gregory_km | kathleengregory@mastodon.social
kathleen.gregory@univie.ac.at

References

Gregory, K., Ninkov, A., Ripp, C., Roblin, E., Peters, I., & Haustein, S. (2023). Tracing data: A survey investigating disciplinary differences in data citation. Zenodo. https://doi.org/10.5281/zenodo.7555266

Gregory, K., Cousijn, H., Groth, P., Scharnhorst, A., Wyatt, S. (2019). Understanding data search as a socio-technical practice. *Journal of Information Science*. https://doi.org/10.1177/0165551519837182