

Vrije Universiteit Amsterdam
RENÉ BEKKERS

FACILITATING OPEN SCIENCE

What do Researchers Need?





23. Describe the data collected for and/or used in your research honestly, scrupulously and as transparently as possible.
24. Manage the collected data carefully and store both the raw and processed versions for a period appropriate for the discipline and methodology at issue.
25. Contribute, where appropriate, towards making data findable, accessible, interoperable and reusable in accordance with the FAIR principles.¹³
35. Be transparent about the method and working procedure followed and record them where relevant in research protocols, logs, lab journals or reports. The line of reasoning must be clear and the steps in the research process must be verifiable. This usually means that the research must be described in sufficient detail for it to be possible to replicate the data collection and its analysis.
36. Be explicit about any relevant unreported data that has been collected in accordance with the research design and could support conclusions different from those reported.

THE OPPOSITION

A provocative visualization

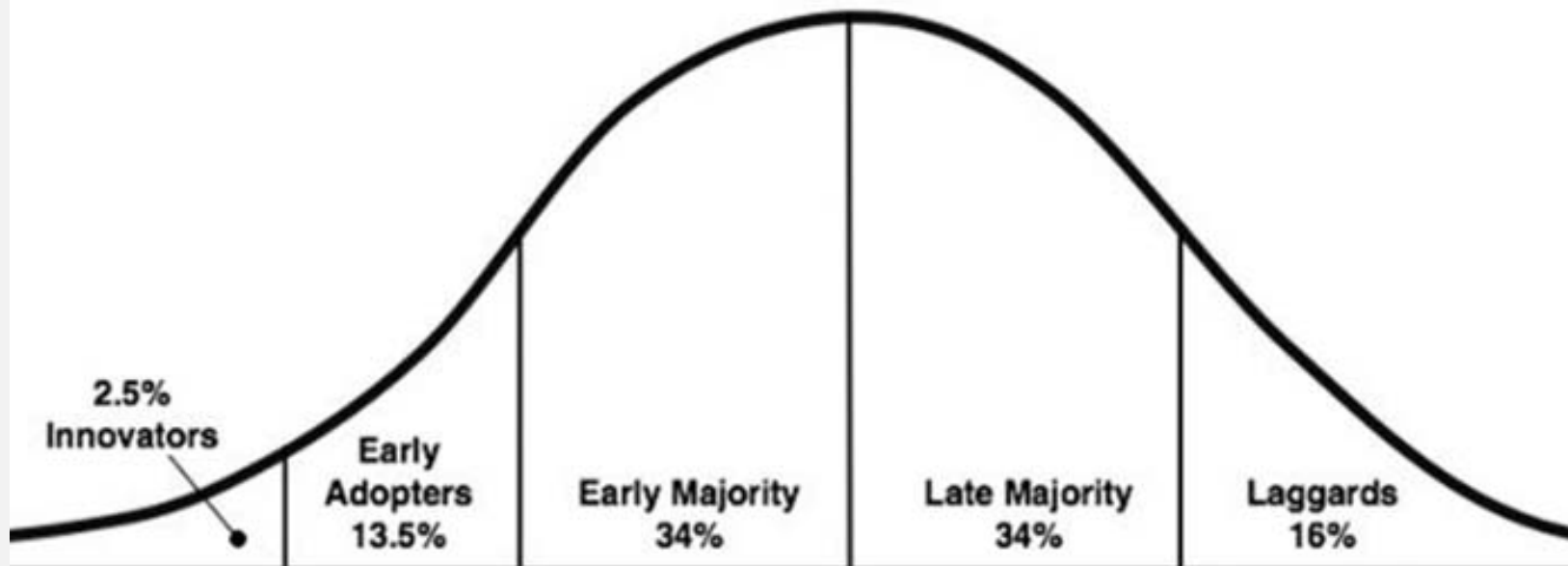
What we did previously



What we do now



Changing a Research Culture

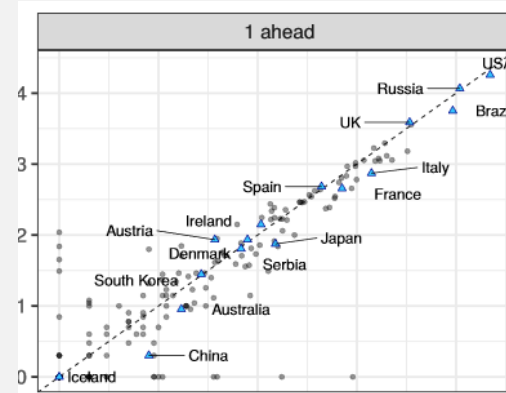


Source: Everett Rogers, *Diffusion of Innovations* (1962).

DATA? WHAT ARE DATA?



Field	Type	Nullable	Attributes	Null	Default	Extra
<input type="checkbox"/> patient_id	bigint(20)	NO	PRIMARY	NO	None	auto_increment
<input type="checkbox"/> hospital_id	int(10)	NO	FOREIGN	NO	None	
<input type="checkbox"/> title	varchar(255)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> first_name	varchar(255)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> middle_name	varchar(255)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> last_name	varchar(255)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> date_of_birth	date	NO	None	NO	None	
<input type="checkbox"/> age	int(10)	NO	FOREIGN	NO	None	
<input type="checkbox"/> sex	varchar(2)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> occupation	varchar(100)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> marital_status	varchar(10)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> religion	varchar(15)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> title	varchar(255)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> preferred_language	varchar(15)	YES	UNICODE_32	NO	None	
<input checked="" type="checkbox"/> language	varchar(15)	YES	UNICODE_32	NO	None	
<input type="checkbox"/> status	varchar(1)	NO	FOREIGN	NO	1	
<input type="checkbox"/> date_created	datetime	NO	None	NO	None	
<input type="checkbox"/> date_modified	datetime	NO	None	NO	None	



DATA FROM INTERVIEWS

RAW

PROCESSED

ANALYZED

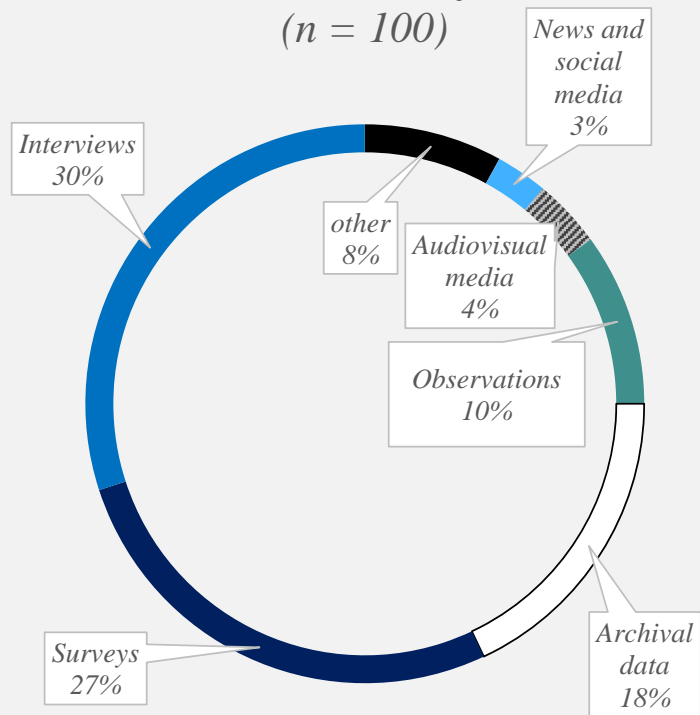


Patient 27 said:
“This was no ordinary fever. I knew that this was a much more serious disease.”

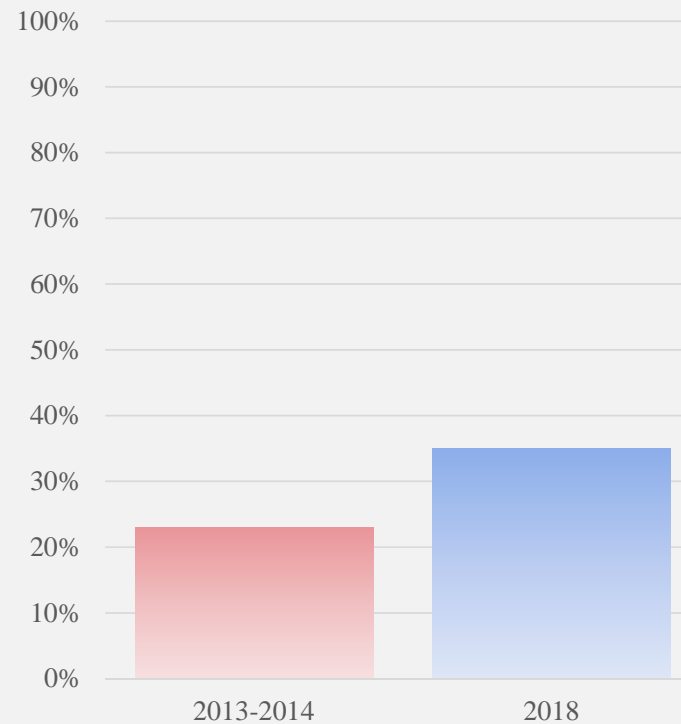
DATA REPORTED IN PUBLICATIONS

Interviews have become the major source of data in the Faculty of Social Sciences @ VU Amsterdam

*Data in Publications from 2018
(n = 100)*



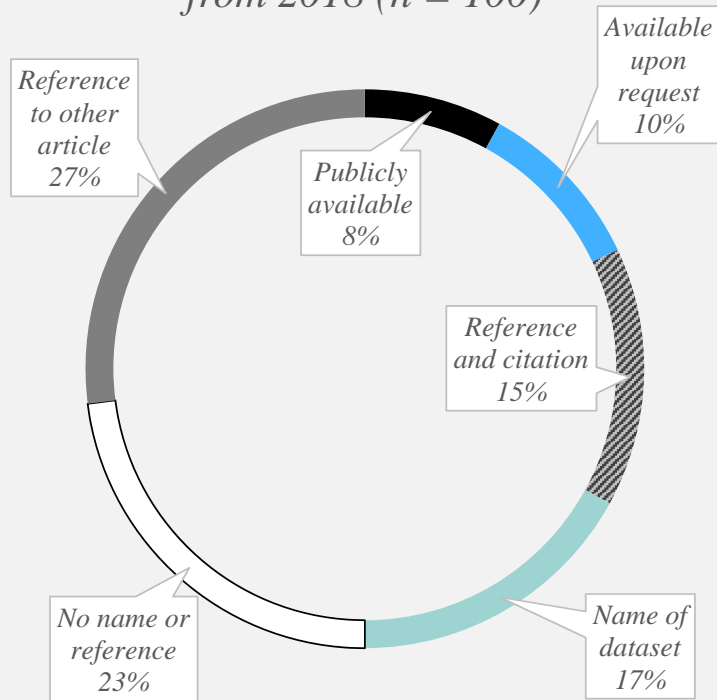
Primary data from interviews



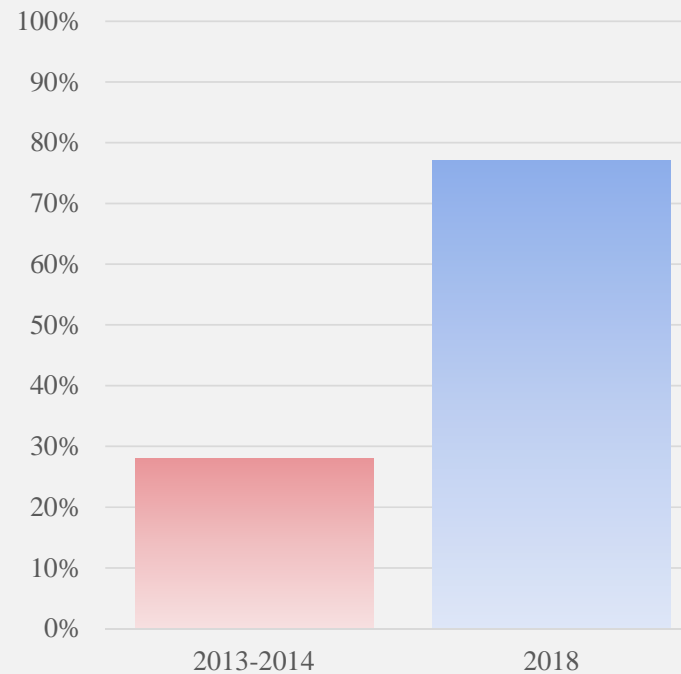
AVAILABILITY OF DATA IN PUBLICATIONS

We are very far from the ideal, but we are making progress

*Details on Data in Publications
from 2018 (n = 100)*



*At least some details on the
data analyzed*



REPRODUCIBLE RESEARCH

*Views from the
workfloor*

REPRODUCIBLE RESEARCH IS

Desirable

- When the goal of research is nomothetic: to develop knowledge that can be generalized to other samples



REPRODUCIBLE RESEARCH CAN BE

Desirable

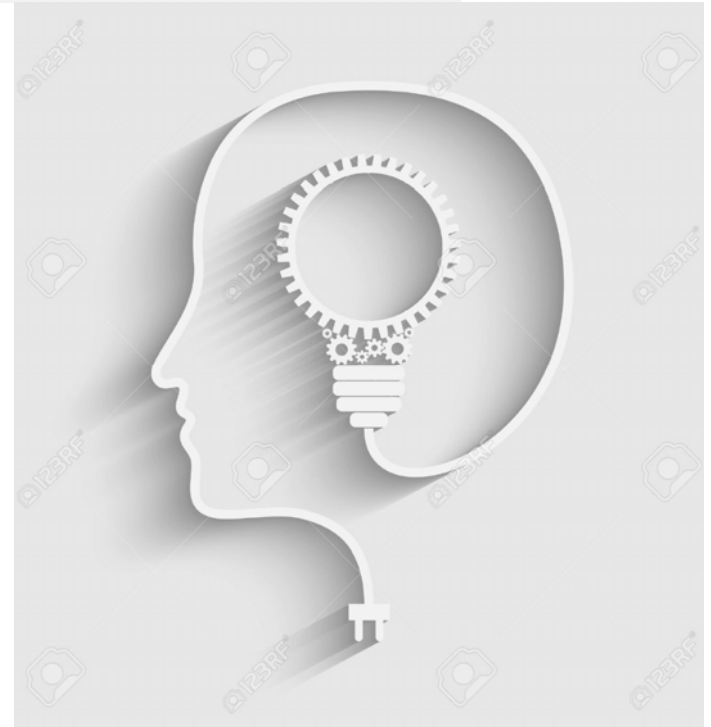
- When the goal of research is nomothetic: to develop knowledge that can be generalized to other samples
- ...but this is not necessarily the goal of idiographic research



MAKING RESEARCH REPRODUCIBLE CAN BE

Meaningless

- When the only data are memories of a researcher, as in participant observation without field notes



MAKING RESEARCH REPRODUCIBLE CAN BE

Less interesting

- When data are pseudonymized, the granularity of the data is reduced

MAKING RESEARCH REPRODUCIBLE CAN BE

Dangerous

- When identifying those observed brings them in danger
- E.g., interviews with rebels in war zones



MAKING RESEARCH REPRODUCIBLE CAN BE

A threat

- To your credibility as a researcher
- E.g., when you've promised anonymity to informants



MAKING RESEARCH REPRODUCIBLE CAN BE

A threat

- To the credibility of science
- E.g., when the results are not reproducible from the data.
- Donner's argument against transparency: you don't want to know how the 'sausage factory' works, it's ugly





SOLUTIONS

*Governance
instruments for science*

SOME P-VALUES

Let's p-hack our way to good science

What we did previously

- Prescribe a code of conduct, assume people know the rules, and follow them
- Perverse incentives – reward publishing in ‘high impact factor’ journals, assuming that peer review keeps out bad science

What we need to do

- Prevention by training, also of seniors
- Performance evaluation – reward transparency and research quality
- Provide resources: formats, templates, good examples, support, tools
- Pay for the time good data management takes

SOME P-VALUES

Let's p-hack our way to good science

What I think we should do as well

- Practice what we preach: lead by example
- Praise & prizes for exemplars
- Pick new talent carefully

MORE P-VALUES

Let's p-hack our way to good science

Questionable

- Patience & Persistence – leaving responsibility up to individuals
- Preaching without practicing
- Payments for good behaviors

Future options

- Plead guilty: come clean, confess mistakes
- Policing: audits and fines for violations
- Permits: introduce a research license that can be revoked

SEND YOUR THOUGHTS

René Bekkers

Research Ethics Review Committee ✉ [*r.bekkers@vu.nl*](mailto:r.bekkers@vu.nl)

Faculty of Social Sciences 🌐 [*renebekkers.wordpress.com/*](http://renebekkers.wordpress.com/)