Minor adjustments, major benefits:

Assessing reproducibility in hydrology research

Jim Stagge

The Ohio State University Department of Civil, Environmental and Geodetic Engineering

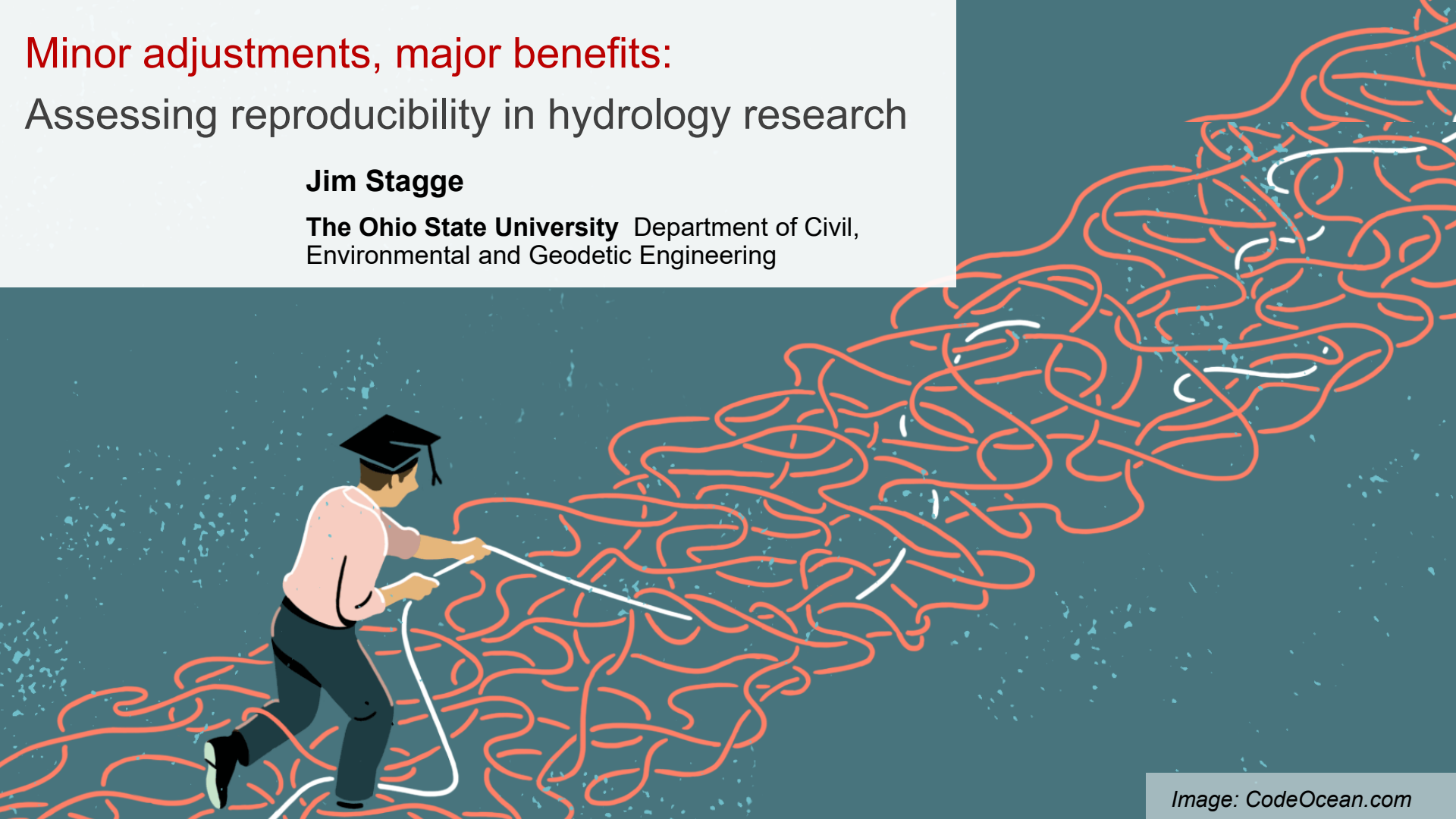
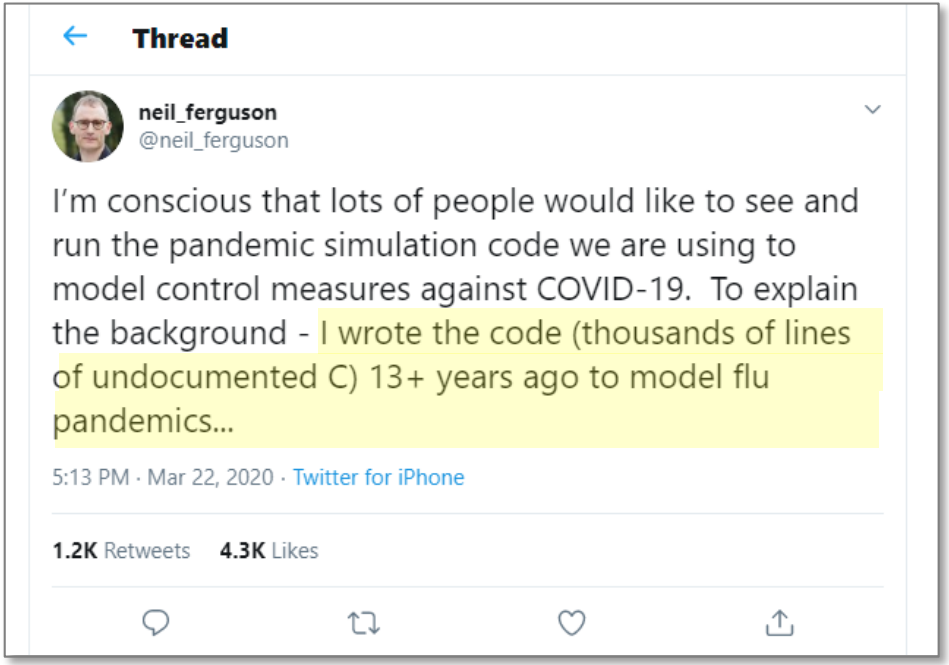
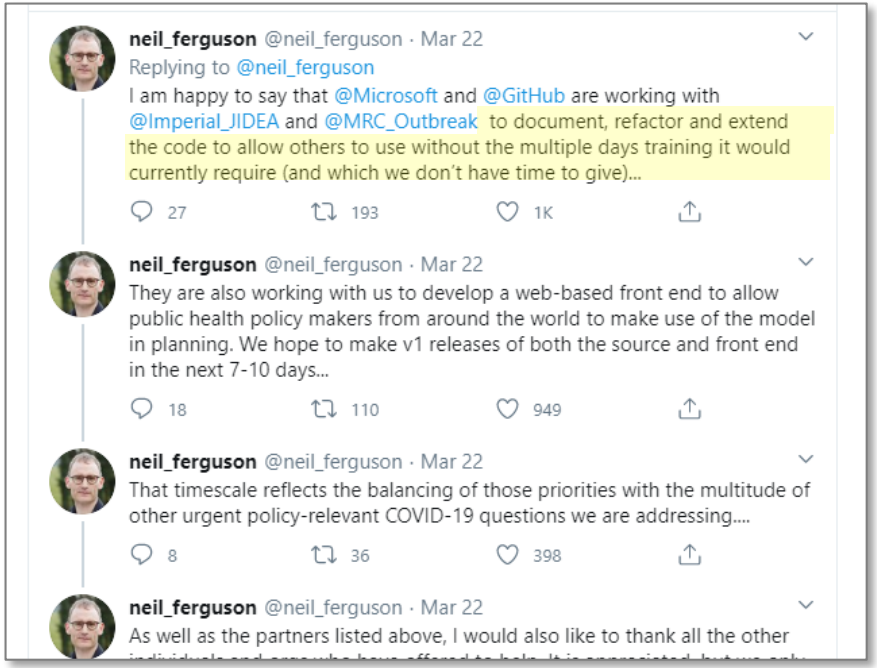


Image: CodeOcean.com

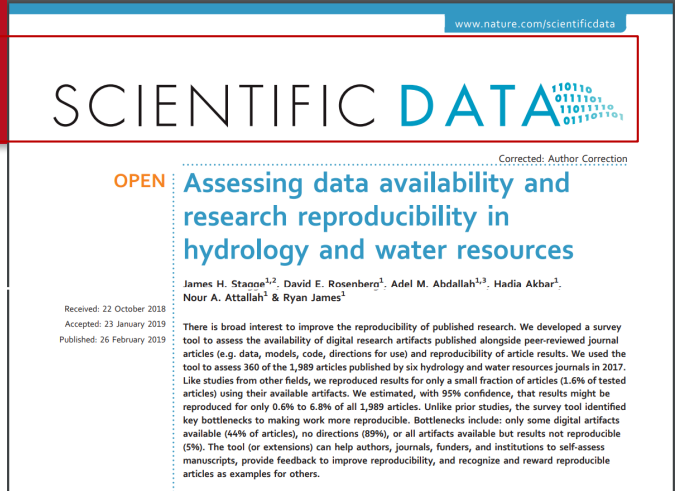
Some relevant motivation





Purpose of study

Significant effort has been put into improving reproducible research.



www.nature.com/scientificdata

SCIENTIFIC DATA

Corrected: Author Correction

OPEN

Assessing data availability and research reproducibility in hydrology and water resources

James H. Stagge[1,2], David E. Rosenberg[1], Adel M. Abdallah[1,3], Hadia Akbar[1], Nour A. Attallah[1] & Ryan James[1]

Received: 22 October 2018
Accepted: 23 January 2019
Published: 26 February 2019

There is broad interest to improve the reproducibility of published research. We developed a survey tool to assess the availability of digital research artifacts published alongside peer-reviewed journal articles (e.g. data, models, code, directions for use) and reproducibility of article results. We used the tool to assess 360 of the 1,989 articles published by six hydrology and water resources journals in 2017. Like studies from other fields, we reproduced results for only a small fraction of articles (1.6% of tested articles) using their available artifacts. We estimated, with 95% confidence, that results might be reproduced for only 0.6% to 6.8% of all 1,989 articles. Unlike prior studies, the survey tool identified key bottlenecks to making work more reproducible. Bottlenecks include: only some digital artifacts available (44% of articles), no directions (89%), or all artifacts available but results not reproducible (5%). The tool (or extensions) can help authors, journals, funders, and institutions to self-assess manuscripts, provide feedback to improve reproducibility, and recognize and reward reproducible articles as examples for others.

- How are we doing (hydrology and water resources)?
- How can we realistically assess and identify bottlenecks?

Stagge, J., Rosenberg, D., Abdallah, A. et al. Assessing data availability and research reproducibility in hydrology and water resources. Sci Data 6, 190030 (2019). https://doi.org/10.1038/sdata.2019.30

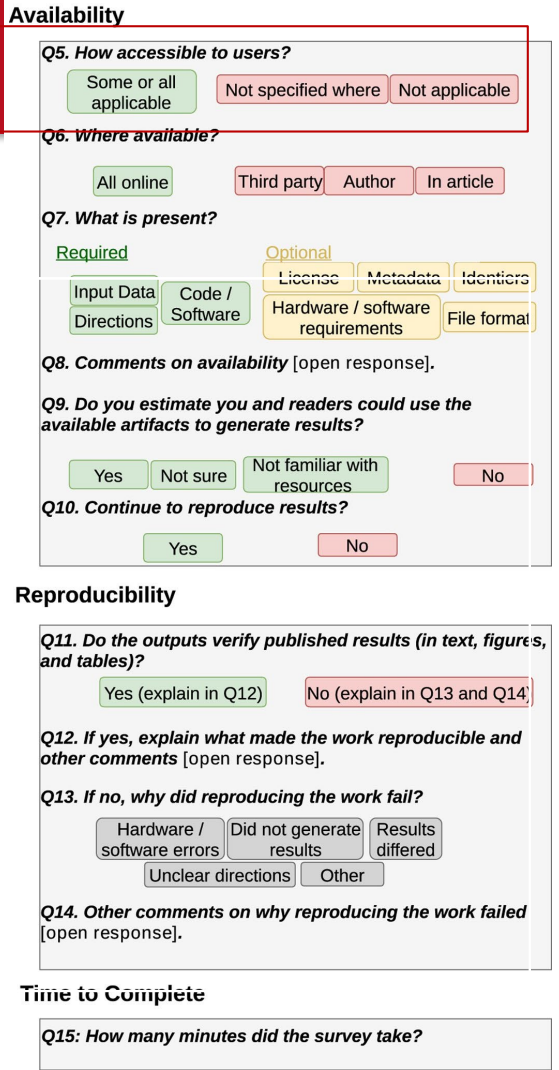
Reproducibility as a Continuum

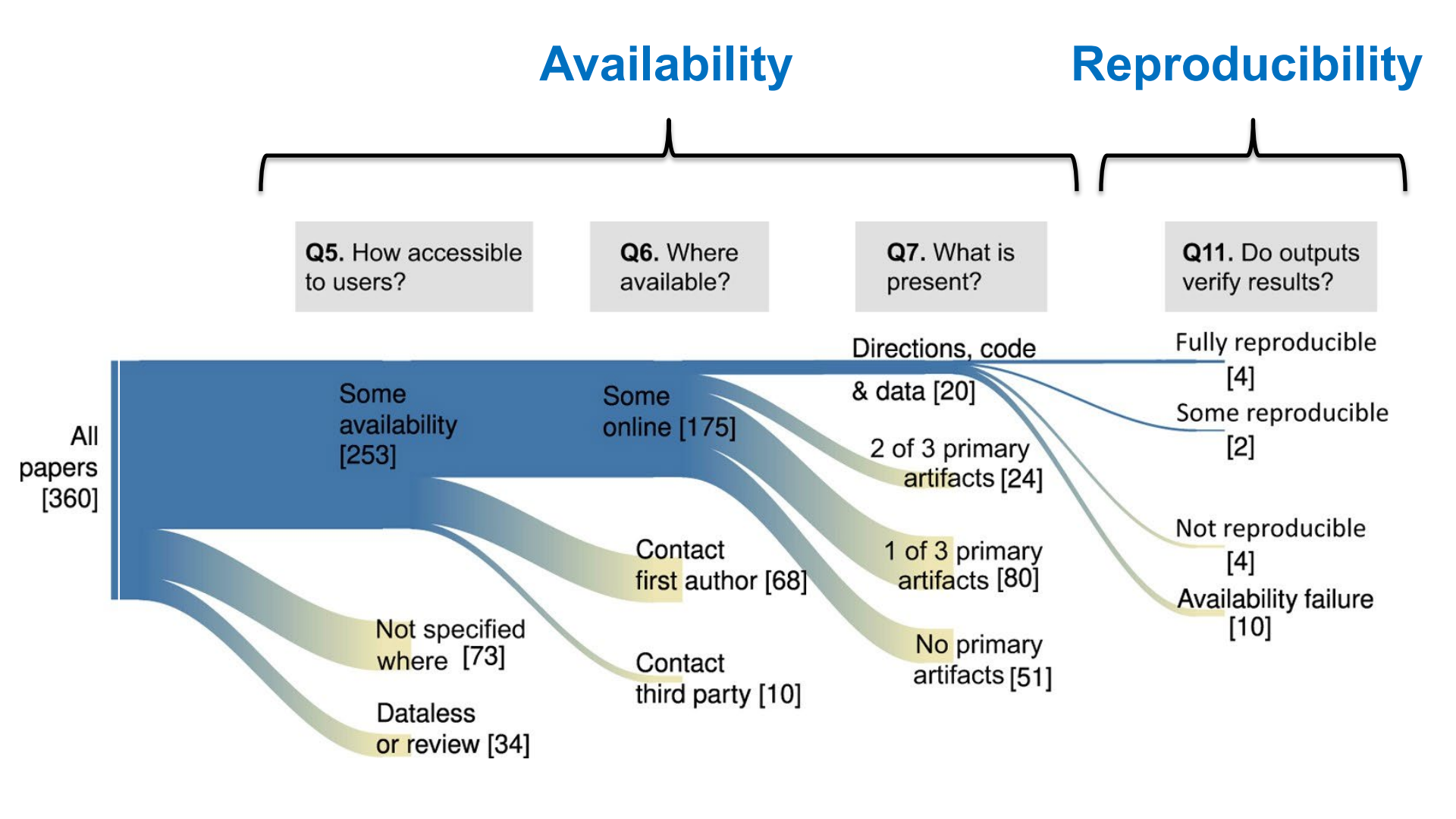
Institute of Education Sciences, US Dept of Education, and NSF. (2018) Companion Guidelines on Replication and Reproducibility in Education Research

Survey of 360 Articles

- Created 15 question survey instrument:
  - Availability - Self-evaluation & unfamiliar reviewer
  - Reproducibility: Subject reviewer

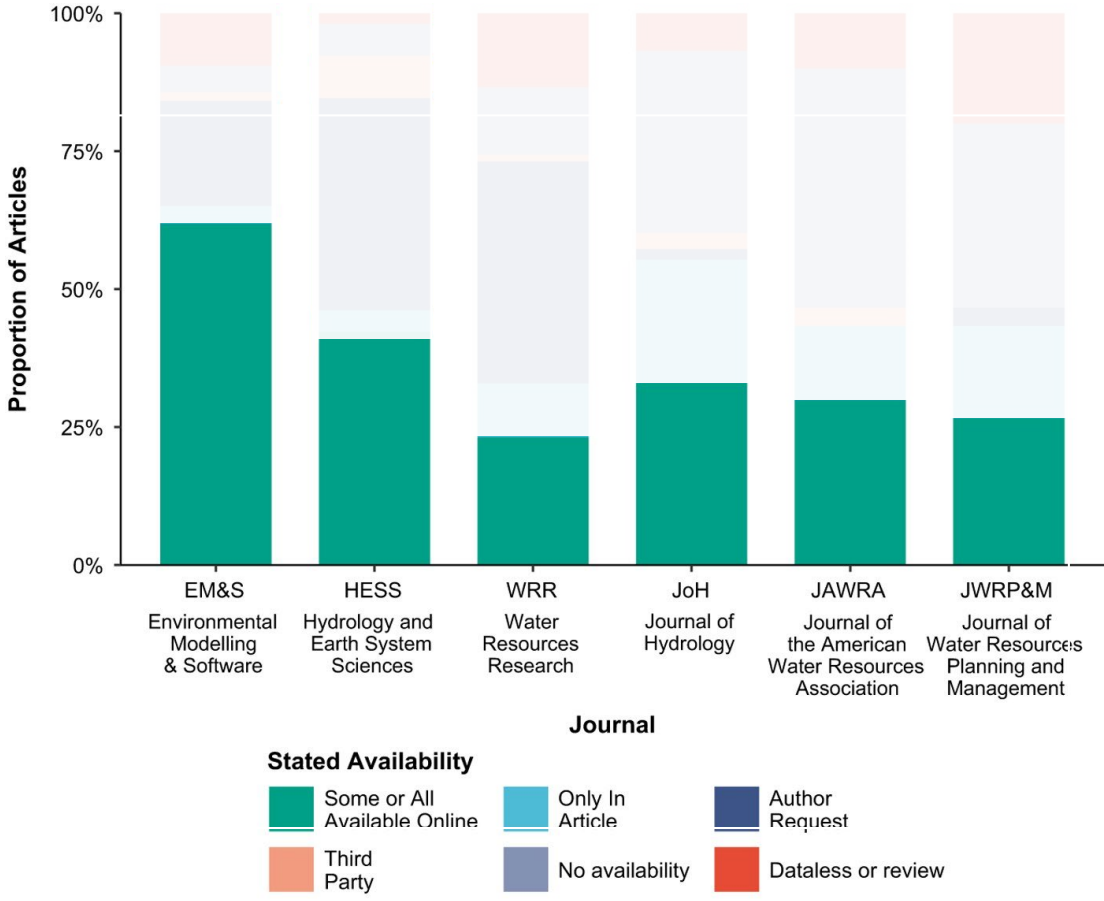
- 6 hydrology/water resources journals
- 360 of 1,989 articles published in 2017
- 60 articles per reviewer
- 5-20 min for avail, 1-3 hours for reproducibility




Availability
Reproducibility
Q5. How accessible to users?
Q6. Where available?
Q7. What is present?
Q11. Do outputs verify results?
All papers [360]
Some availability [253]
Not specified where [73]
Dataless or review [34]
Some online [175]
Contact first author [68]
Contact third party [10]
Directions, code & data [20]
2 of 3 primary artifacts [24]
1 of 3 primary artifacts [80]
No primary artifacts [51]
Fully reproducible [4]
Some reproducible [2]
Not reproducible [4]
Availability failure [10]

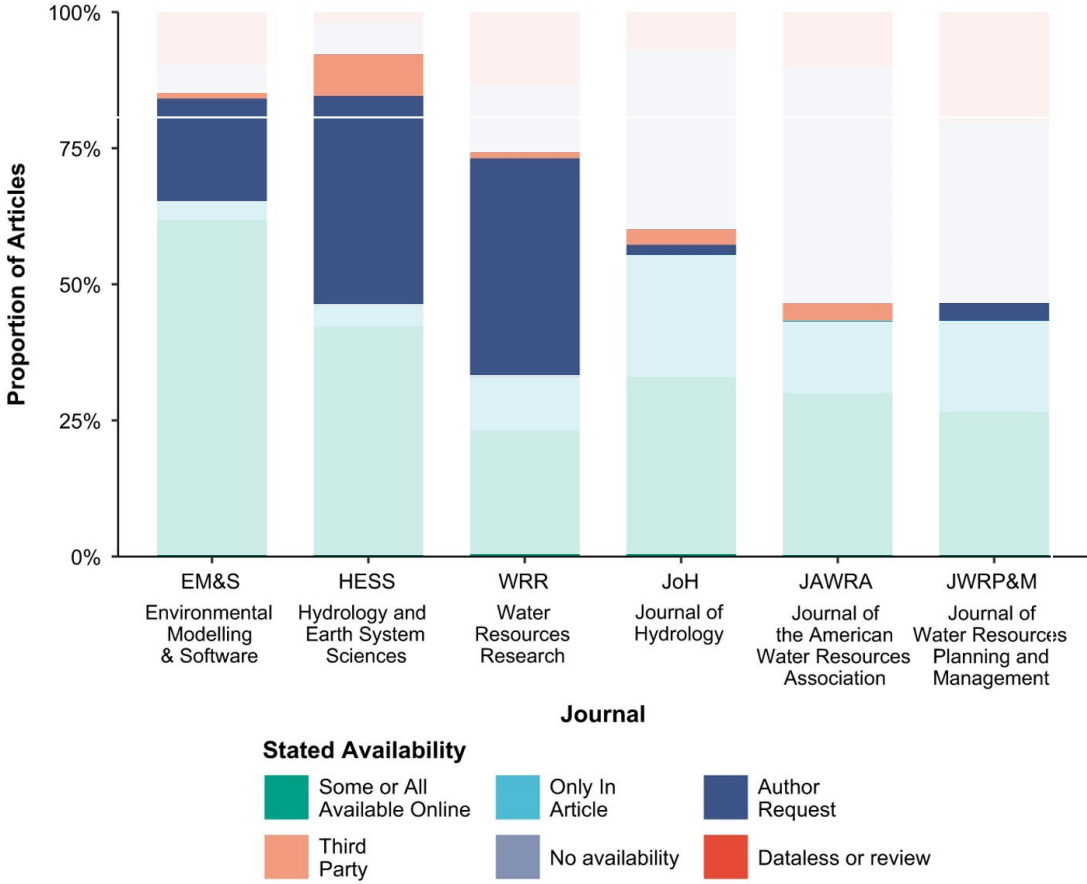
Stated Availability by Journal

- **70.3%** stated some materials online
- Data availability statement required vs encouraged

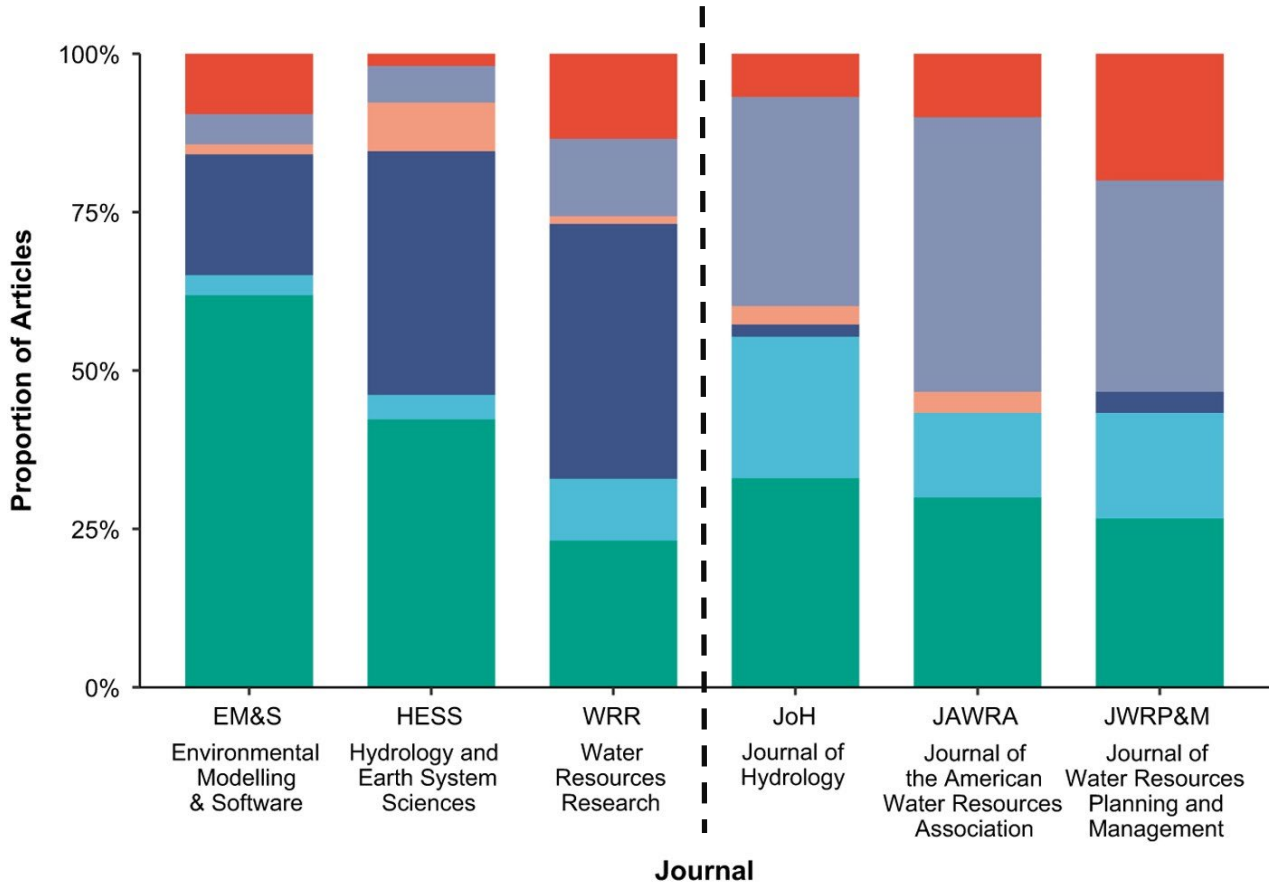


Stated Availability by Journal

- Data availability statement increased the author/third party requests
- Did not test

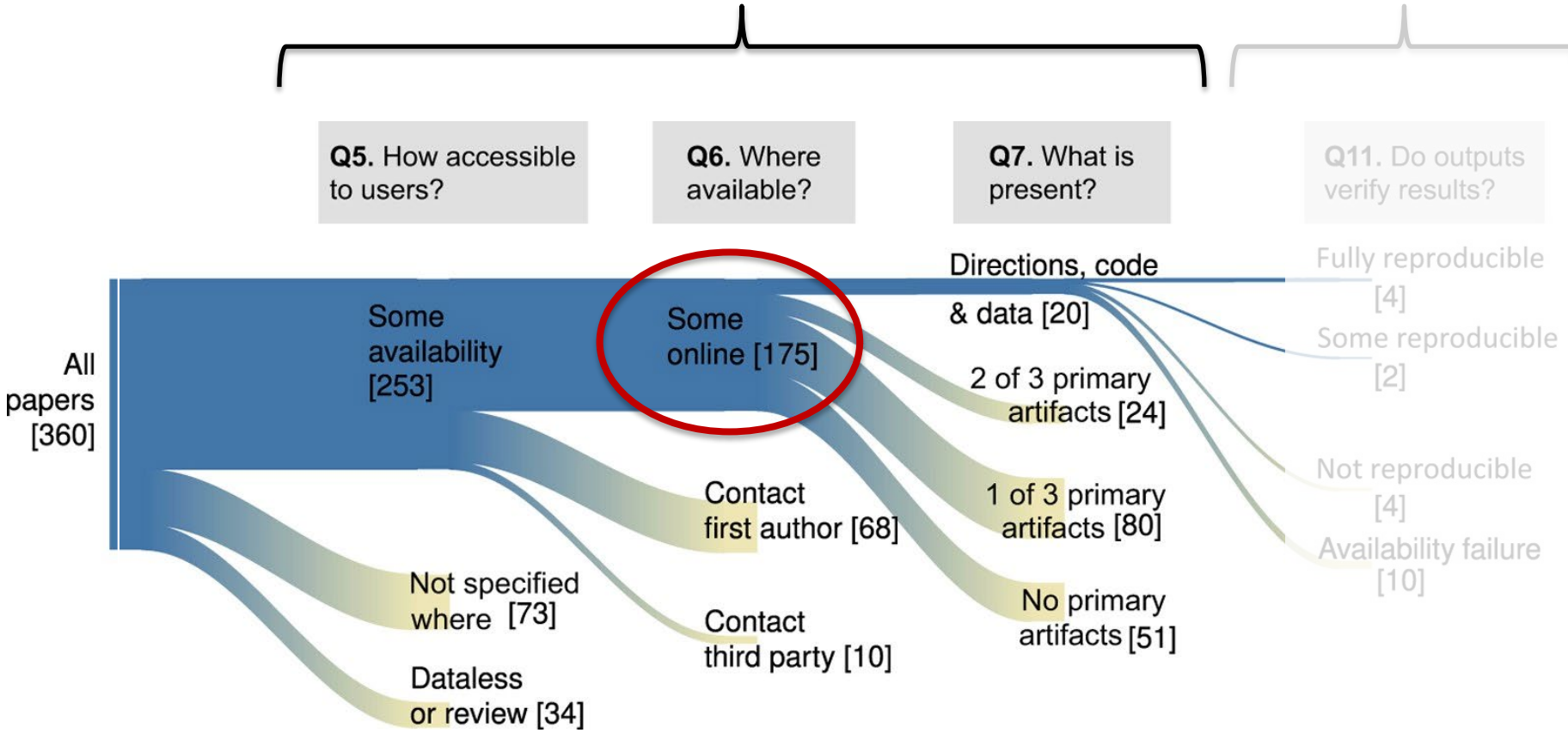


Stated Availability by Journal



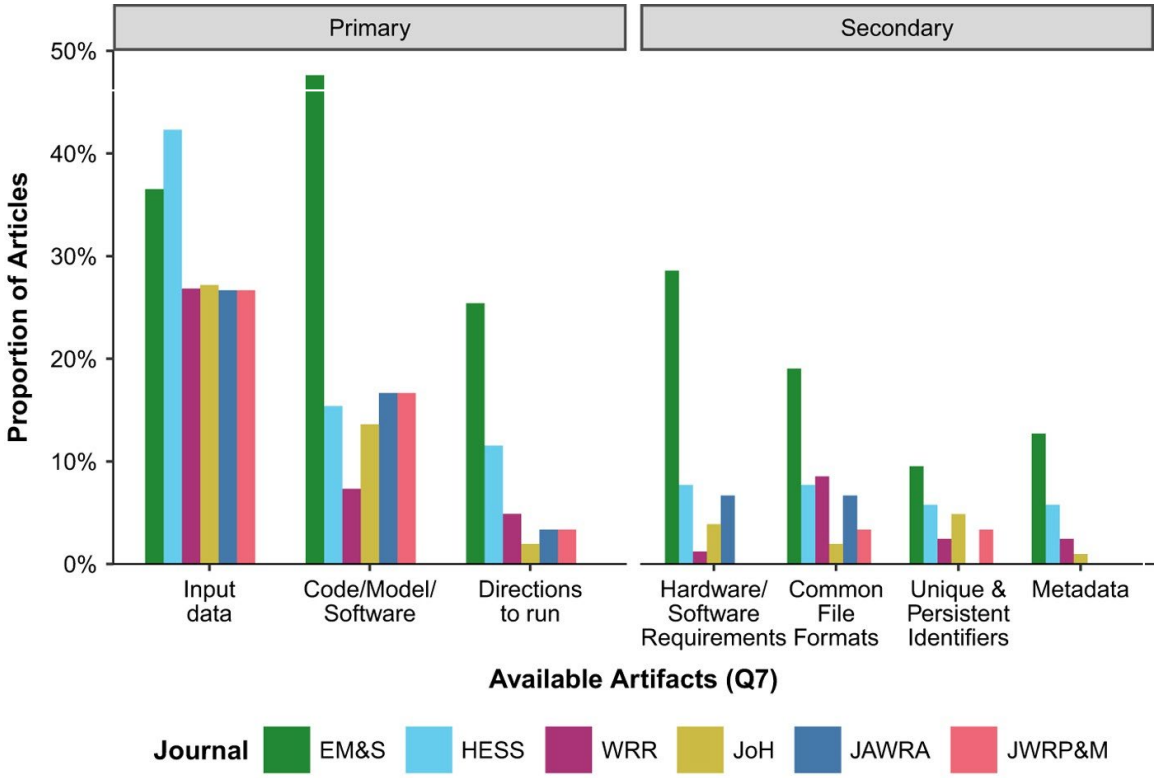
Availability

Reproducibility

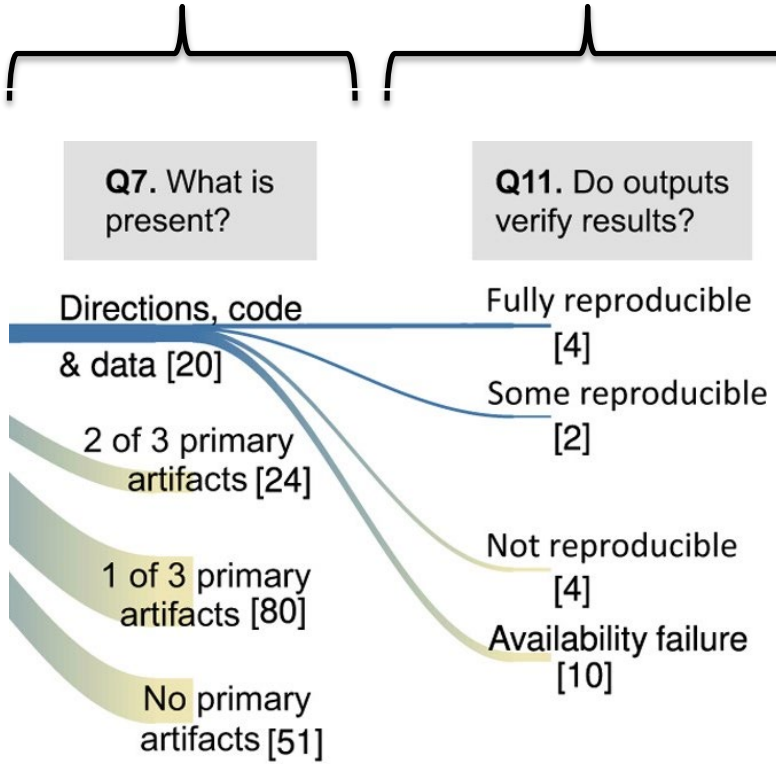


Available digital artifacts

- Primary (necessary) & Secondary (useful)
- Data most common artifact by far (25-40%)
- Followed by Code (15%) and Directions (5%)

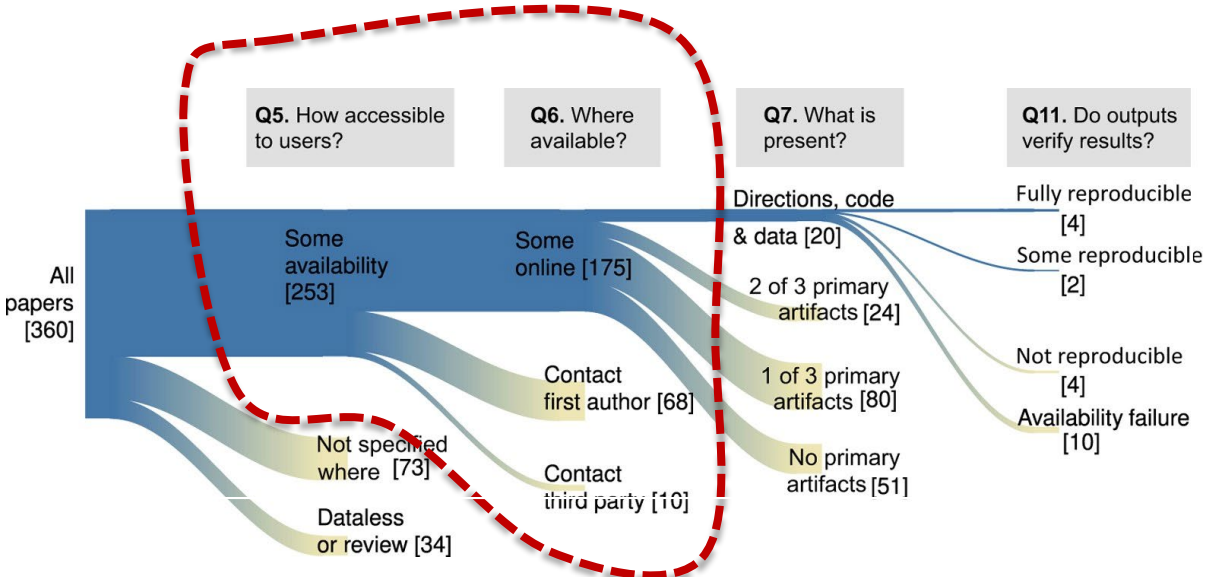


Availability Reproducibility



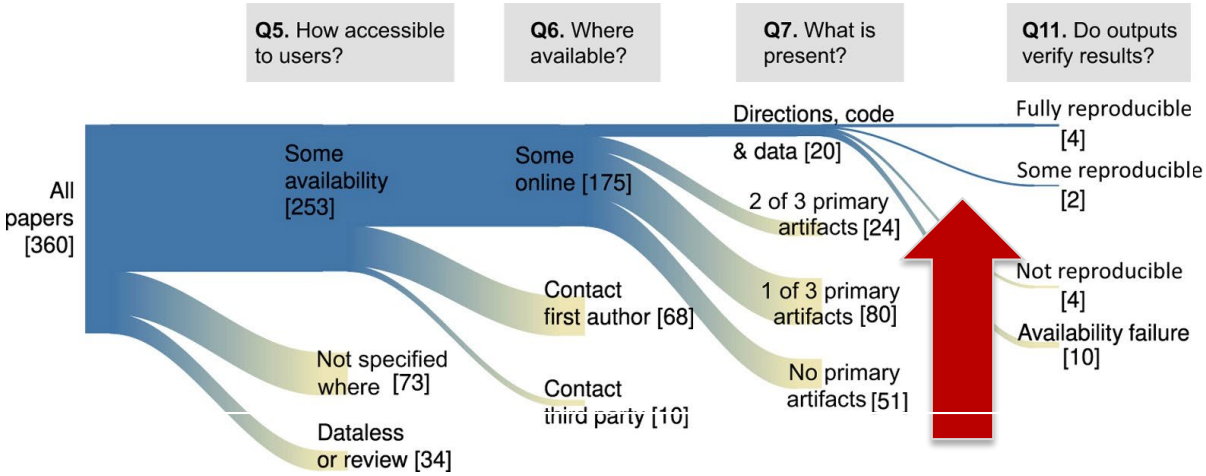
Conclusions: Bottlenecks or Opportunities

- **Data availability statement** appears to nudge towards availability



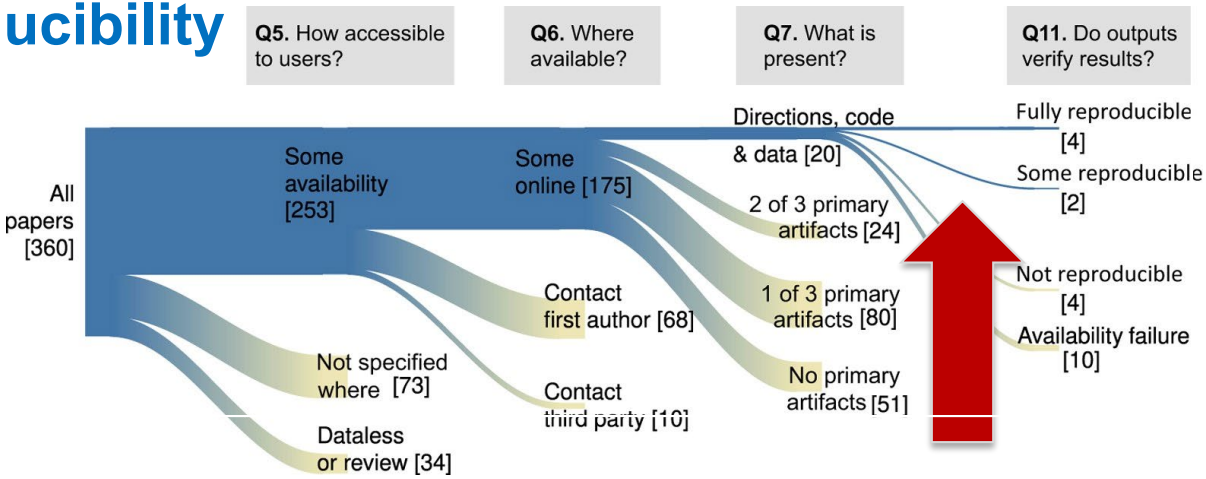
Conclusions: Bottlenecks or Opportunities

- Data availability statement appears to nudge towards availability
- **Experiment directions/documentation** vs. data dump
- Nudge towards documentation: repository tools, templates



Conclusions: Bottlenecks or Opportunities

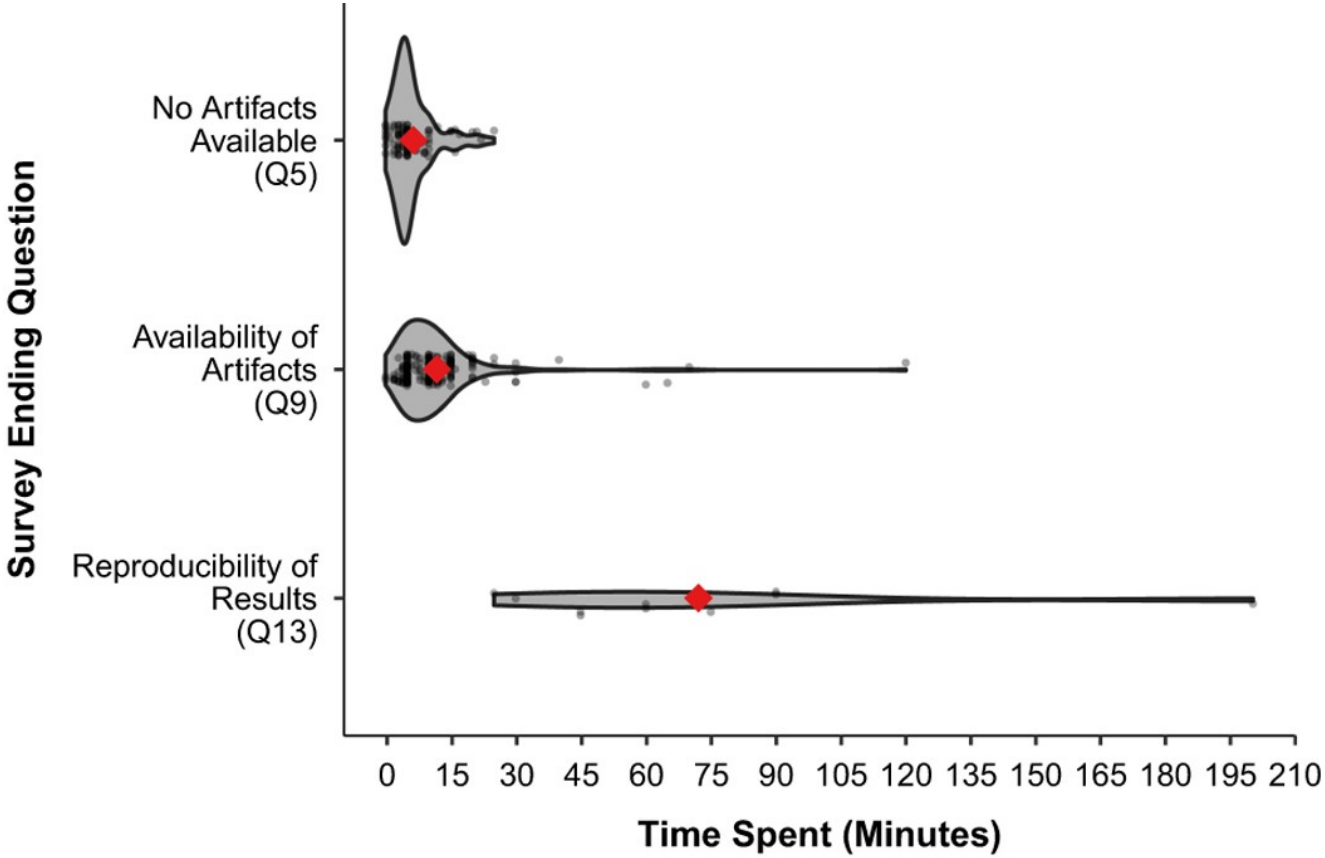
- Data availability statement appears to nudge towards availability
- Directions/documentation vs. data/code dump
- Nudge towards documentation: repository tools, templates
- When artifacts available, **high likelihood of reproducibility**



Acknowledgements

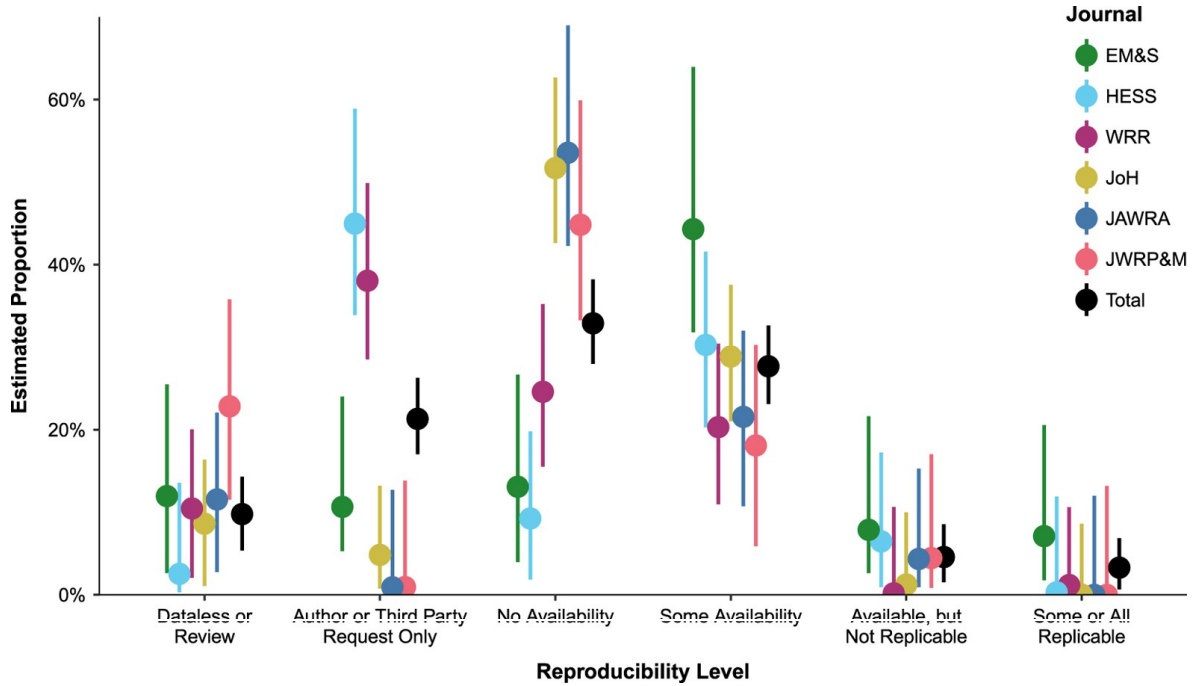
- Co-authors: David Rosenberg, Adel Abdallah, Hadia Akbar, Nour Attallah, and Ryan James
- Journals and Authors for making their work reproducible

Time Spent


Survey Ending Question
No Artifacts Available (Q5)
Availability of Artifacts (Q9)
Reproducibility of Results (Q13)
0 15 30 45 60 75 90 105 120 135 150 165 180 195 210
Time Spent (Minutes)

Population Estimates

- Oversampled for reproducible keywords
- 1.6% reproduced in sample
- 0.6 – 6.8% population estimate



Sampled articles

	EM&S		HESS		WRR		JoH		JAWRA		JWRP&M	
	2017	*Sample*	*2017*	*Sample*	*2017*	*Sample*	*2017*	*Sample*	*2017*	*Sample*	*2017*	*Sample*
Keyword	49	48	9	9	23	23	24	24	7	7	8	8
Non-keyword	181	15	319	43	511	59	645	79	102	23	111	22
Total	230	63	328	52	534	82	669	103	109	30	119	30