

A Survey Tool to Assess and Improve Data Availability and Research Reproducibility

J.H. Stagge^{1,2*}, D.E. Rosenberg¹, A.M. Abdallah¹, H. Akbar¹, N.A. Attallah¹, R. James¹

(1) **Utah State University**, Department of Civil and Environmental Engineering
(2) **The Ohio State University**, Department of Civil, Environmental and Geodetic Engineering
*Contact: stagge.11@osu.edu <http://www.jstagge.com>

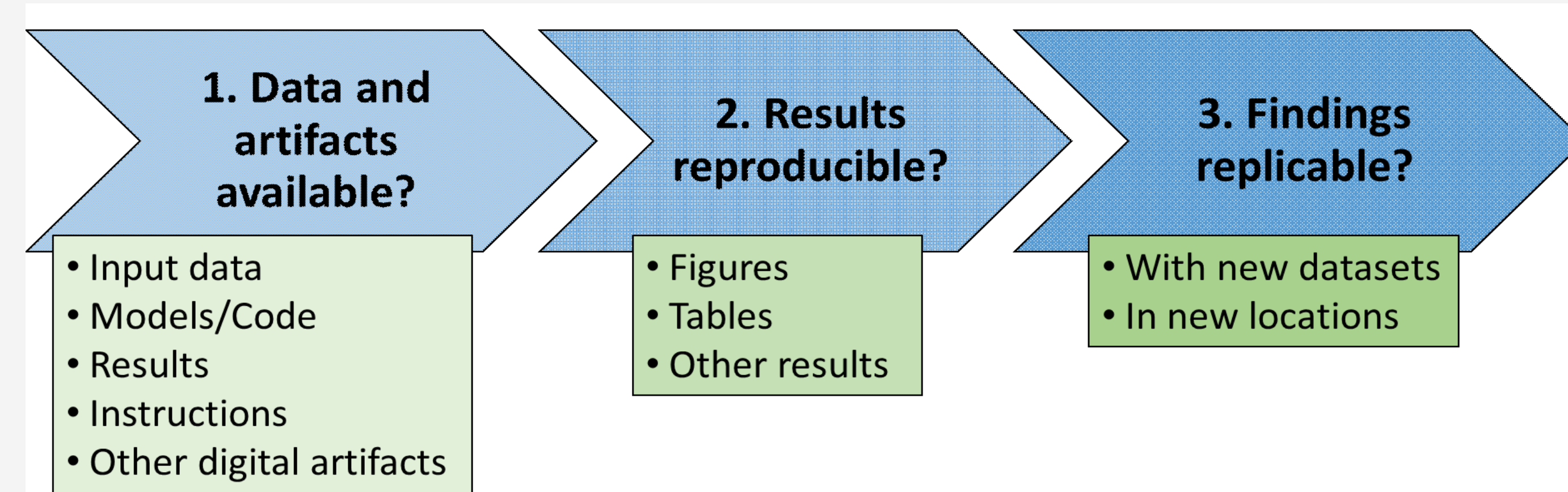
I. Introduction

The scientific community is broadly interested to improve the reproducibility of research, a cornerstone of the scientific process.

To further this effort, we:

- Developed a survey tool to assess the reproducibility of peer-reviewed publications
- Assessed 360 articles published in hydrology and water resources journals
- Identified common factors of reproducible papers and bottlenecks to reproducibility
- Recommended how authors, journals, funders, and institutions can encourage and reward reproducible research

Reproducibility is a continuum



Availability: all necessary research artifacts (data, model, code, directions, etc.) are made available for others to reuse

Reproducibility: ability to reproduce published results exactly using available data

Replicability: ability to replicate published conclusions using new data or techniques

III. Testing the Survey Tool

Sampling Approach

- 360 peer-reviewed articles were randomly sampled from the 1,989 articles published in 2017 by 6 hydrology and water resources journals.

- Sampling was approximately proportional to the number of articles published, with extra weight placed on articles that included a pre-determined set of reproducibility-keywords.

Table 1 Number of articles published in 2017 and number of articles sampled.

	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-key word	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

HESS: Hydrology and Earth Systems Sciences

EM&S: Environmental Modeling & Software

JAWRA: Journal of the American Water Resources Association

JoH: Journal of Hydrology

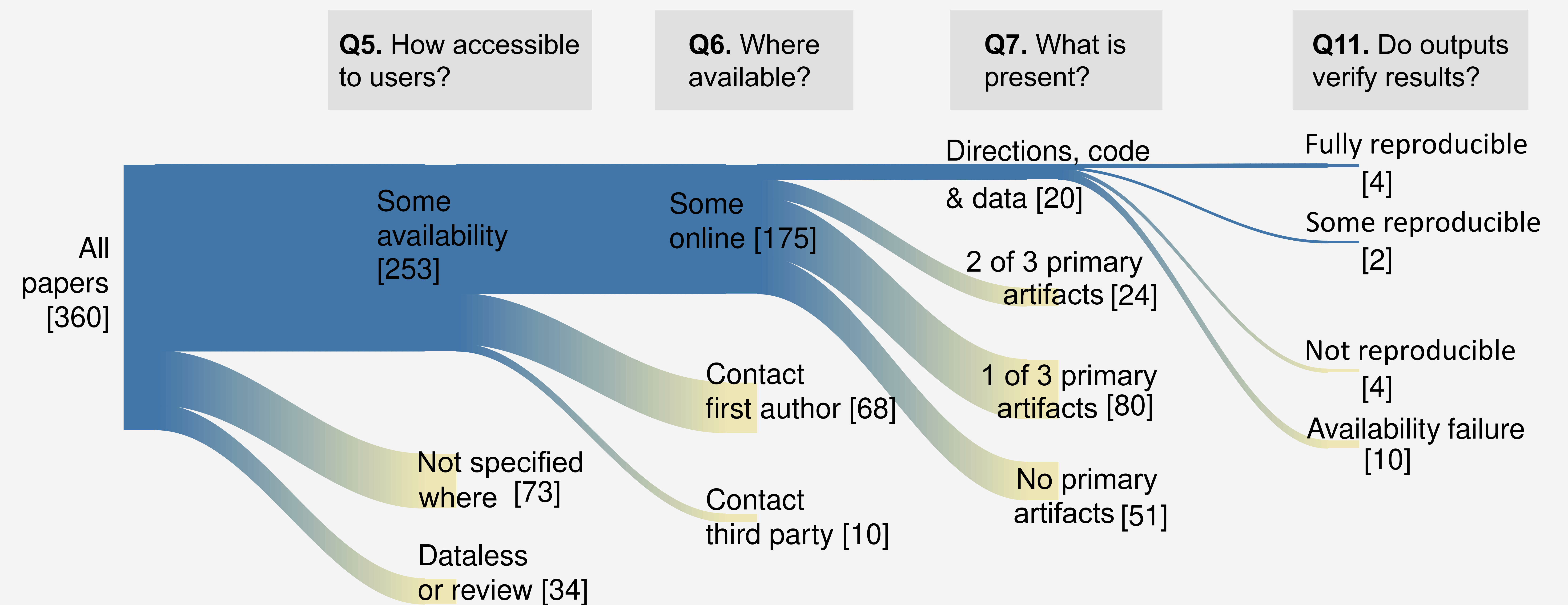
JWRP&M: Journal of Water Resources Planning and Management

WRR: Water Resources Research

Availability and Reproducibility Reviews

- Each author was randomly assigned 60 articles to test for availability.
- Articles were re-assigned for reproducibility testing based on software familiarity.
- All reviews followed the survey tool and were recorded online

IV. Results: Overview of papers progressing through the survey



II. 15-Question Survey Tool

<https://tinyurl.com/ReproduceSurvey>

Paper Metadata

Q1. Assessor's name
Q2. Journal name
Q3. Article DOI
Q4. Full paper citation

Availability

Q5. How accessible to users?
Some or all applicable | Not specified where | Not applicable

Q6. Where available?
All online | Third party | Author | In article

Q7. What is present?
Required: Input Data, Code / Directions, Software, Hardware / software requirements, File format
Optional: License, Metadata, Identifiers

Q8. Comments on availability [open response].

Q9. Do you estimate you and readers could use the available artifacts to generate results?
Yes | Not sure | Not familiar with resources | No

Q10. Continue to replicate results?
Yes | No

Reproducibility

Q11. Do the outputs verify published results (in text, figures, and tables)?
Yes (explain in Q12) | No (explain in Q13 and Q14)

Q12. If yes, explain what made the work reproducible and other comments [open response].

Q13. If no, why did reproducing the work fail?
Hardware / software errors | Did not generate results | Results differed | Unclear directions | Other

Q14. Other comments on why reproducing the work failed [open response].

Time to Complete

Q15: How many minutes did the survey take?

V. Results: Availability and Reproducibility Details

Stated Availability by Journal

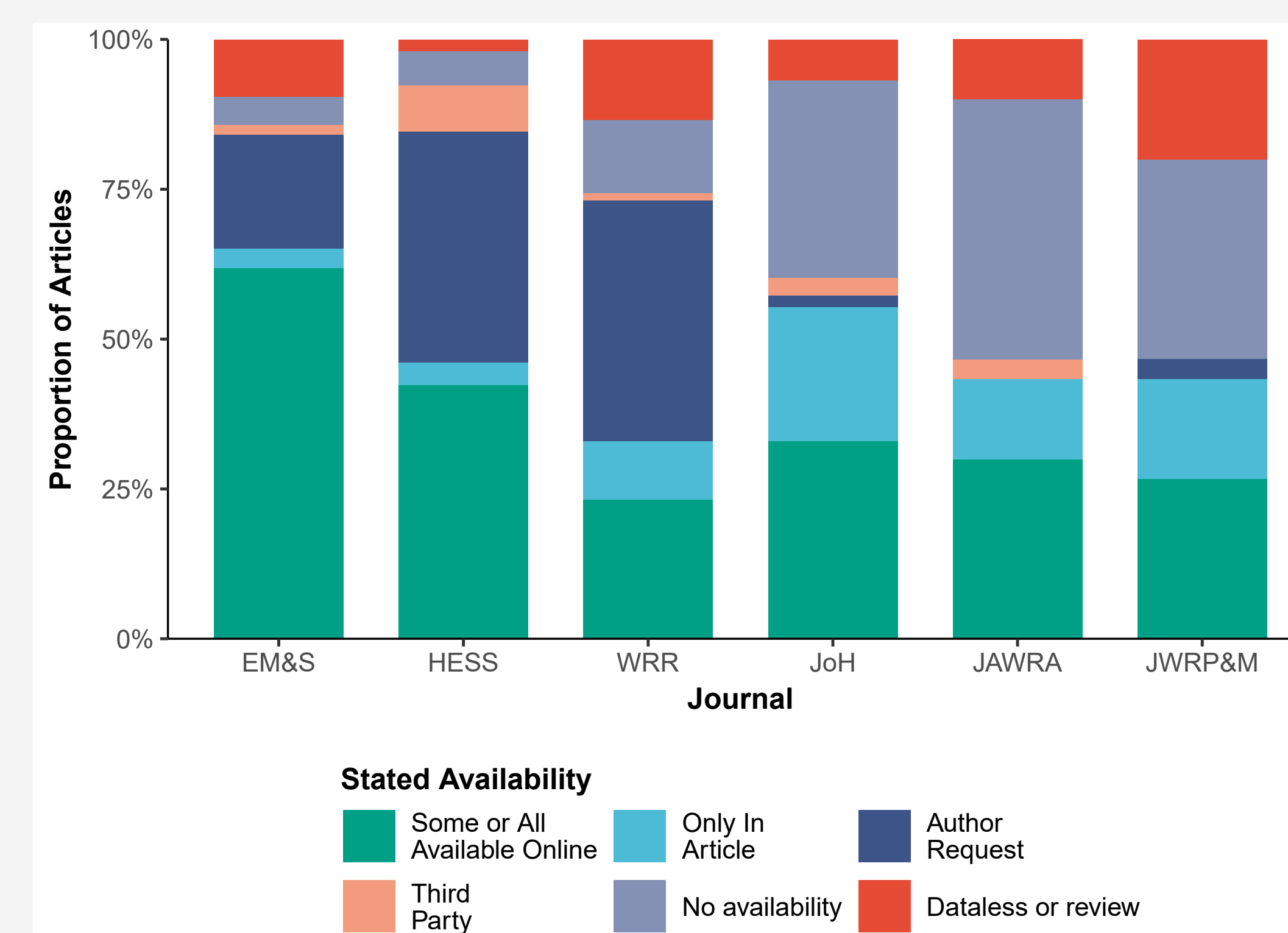


Figure 3 Data, model code availability by journal (summary of Q4 and Q5).

Available Artifacts

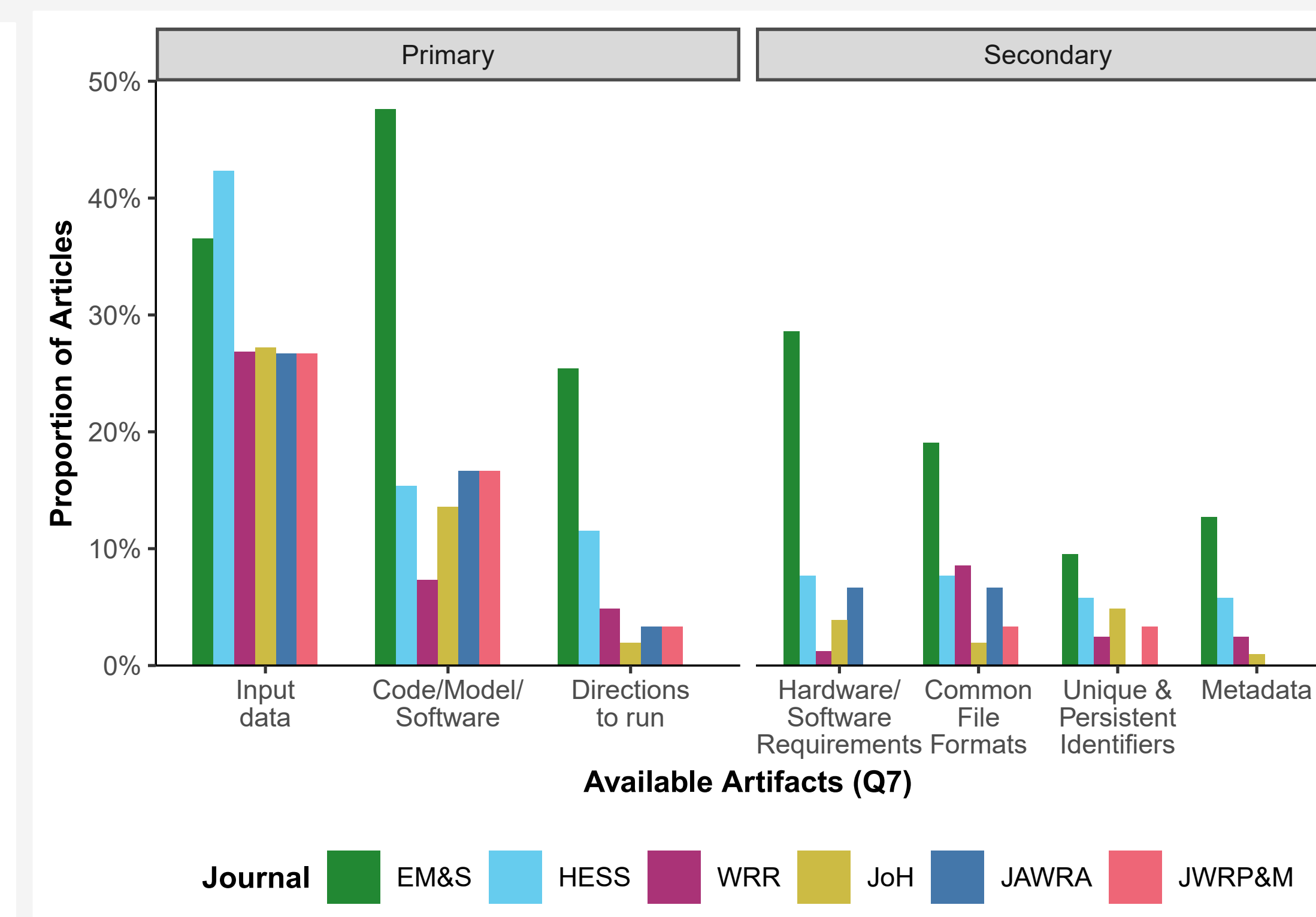


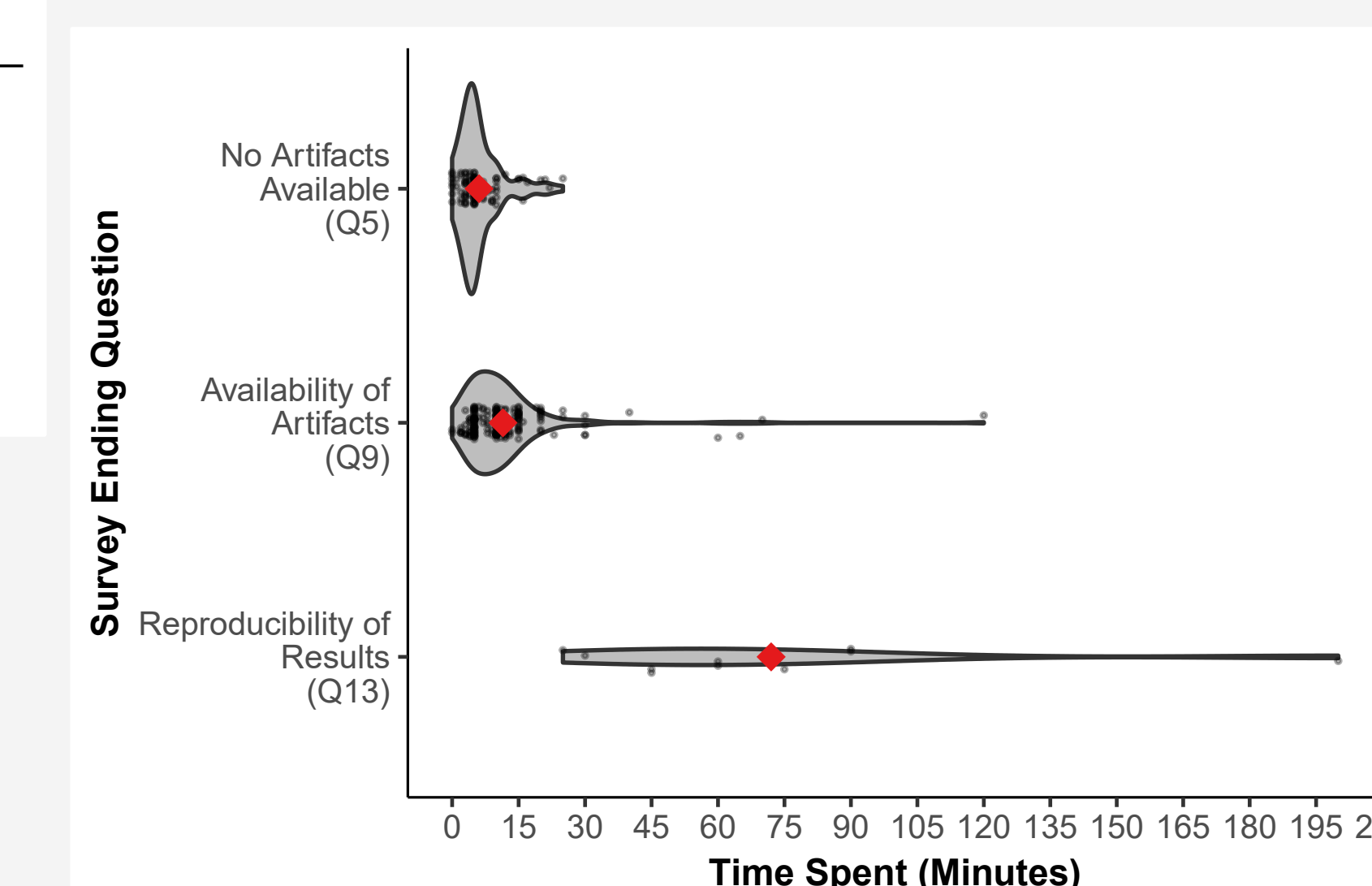
Figure 4 Availability artifacts organized by journal. All percentages are based on the total number of each journal.

Figure 5 Self-reported time to complete survey organized by the survey's ending question. Each reviewed paper is shown by a grey dot, while the mean is represented by a red dot.

Reproducibility

- Reasons for failure included:
- lack of all elements [10 articles]
 - unclear directions [4]
 - did not generate results [3]
 - hardware/software error [2]
 - results differed [1]

Time to Complete Survey



VI. Conclusions

Small changes could produce major improvement

Journal policies partially drive reproducibility of their publications.

Bottlenecks along the reproducibility continuum include:

- A significant fraction of artifacts were only available by request
- 2-3 times more publications included code/data than instructions to use them. Including instructions could potentially double "available" articles
- Once all artifacts were available, reproducibility was the most likely outcome (60%)

Recommendations

Authors: Self-assess before submission using survey. See reproducible papers for examples.

- Journals:**
- Reviewers or journals assess submissions and provide feedback to authors. Availability survey required only 5-15 minutes.
 - Acknowledge papers that meet reproducible standards (bronze, silver, gold).
 - Establish an Associate Editor for Reproducibility.

Funders and Institutions: Recognize and reward researchers that publish reproducible research.