

Enforcement of journal data archiving policies: Evidence from an interview-based study

August 8, 2018

Dan Sholler - University of California, Berkeley - rOpenSci project - @DanSholler

Carl Boettiger (ESPM); Karthik Ram (rOpenSci, Berkeley Institute for Data Science)



Have you published a dataset alongside a paper?

Journals are a key player in data archiving

> 75%



> 66%



Costello, M. J., Michener, W. K., Gahegan, M., Zhang, Z. Q., & Bourne, P. E. (2013). Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology & Evolution*, 28(8), 454-461.

Several different policy options



Several different policy options



Several different policy options



Several different policy options



Several different policy options



Who's responsible for enforcing data archiving policies?

Editorial staff?

*“We really **rely on the AE to make sure**, and the reviewers, but mainly the subject matter editor, to ensure that what’s archived is usable”
(editor-in-chief)*

Interview-based study

Role	Interviews
Editor/Editor-in-Chief	8
AE/Subject/Data Editor	6
Repository or publisher staff	6
Total	20

Reviewers?

*“So there are specific questions in the reviews [guide for reviewers] at [some journals] about whether all of the data should be available are available, **so we hope that that question gets looked at externally by reviewers or the academic editor.**” (AE)*

Authors?

INTERVIEWER: *Is there anyone at [data repository] that's checking the data for you, to see if it's complete, if it's appropriate for the paper, and that sort of thing?*

AE: ***It's completely on the author.*** ... and you're running on volunteers, it just blows up in your face.

Honor system

But we assume everybody's upstanding, and all we check is that data's been deposited. (data editor)

Honor system

*“Here’s the URL if you want the data.” So we know exactly if it has been deposited or not, **based on whether we got that URL or not.** (data editor)*

Future of journal data policies

“All of the quite junior people have become quite expert at fielding these types of inquiries” [about dataset quality and completeness]
(editor-in-chief)

Summary

Journals acknowledge the need for data archiving policies

Moving beyond “available upon request” policies

Summary

However, no real enforcement mechanisms, especially for quality and completeness

Trusting authors vs. enforcing standards

Rapid change vs. cultural change

Acknowledgments

Work funded by the Leona M. and Harry B. Helmsley Charitable Trust

Special thanks to the interview participants

Research conducted under UC Berkeley Office for the Protection of Human Subjects Protocol ID 2017-08-10194

Link to slides