

# #IAmAResearchParasite

In the midst of steady progress in policies for data sharing, a recent editorial expressed a contrarian view.\* The authors described the concern of some scientists about the rise of an underclass of “research parasites” who exploit data sets that are collected and curated by others. Even worse, these parasites might use such data to try to disprove the conclusions posited in the data’s original source studies. The editorial raised the points of how anyone not involved in the original study could use the data without misrepresenting it, and the danger of perhaps arriving at erroneous conclusions. The editorial advised instead that data sharing be implemented by involving the authors of the original study as coauthors in follow-up research. The research community immediately took to Twitter under the hashtag #IAmAResearchParasite to voice opposition to the editorial.

Much of what we know about the large-scale features of this planet is apparent thanks to widespread data-sharing practices and the early establishment of data banks in the geosciences. Aspects such as determining the shape of the ocean floor, ocean chemistry, the internal structure of Earth’s deep interior, the physics and chemistry of the atmosphere, and many other topics could not have been ascertained from a single investigator’s field program. One meta-analysis I published on the South Pacific benefited from observations of my own and those of others, including the 18th-century British explorer Captain James Cook. Involving Cook as a coauthor on my paper was clearly not an option, any more than it would have been feasible or desirable to include the dozens of others, living or dead, who had contributed to the data repository. Many fields, including the biomedical sciences, are now benefiting from meta-analyses of data to better understand the big picture.

Effective data sharing is not trivial or inexpensive

to implement, and it takes more than community acceptance of the practice. Agencies supporting research in oceanography have long funded data and sample repositories and have encouraged data and sample deposition by making new awards contingent on compliance. Repositories are *instrumental* in setting formats for data, so much so that standard programs and apps accept and output data in the standard format. The marine community supports data professionals who are responsible for the quality control of data collected on ships and from other major observing programs.

Often overlooked is the importance of community-established metadata, so that those not involved in the original research will know what the data mean. As an example, in an oceanographic temperature data base, the community had to agree on what T(0) meant. Was it temperature at atmospheric pressure? Temperature at the sea surface?

Communities must discourage low-quality data collection. A well-attended poster presentation at one prominent scientific meeting some years ago compared the crossover errors (misfits) of non-time-dependent measurements (such as depth soundings) from ships’ tracks where they intersected in the

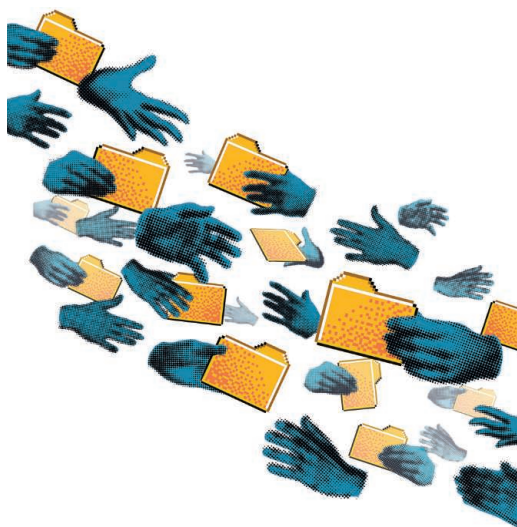
world’s oceans. Any discrepancy at a crossing could be attributed to poor data quality control on either ship, but with thousands of crossings, institutions with systematically more misfits than others stand out. The results did not escape the attention of the funding agencies that support ship time.

There are costs to implementing data reuse, but there are also costs for irreproducible research and for re-collecting data for new uses. And no amount of funding can reconstruct lost ephemeral or time-dependent phenomena for which the data were not well curated. No more excuses: Let’s step up to data sharing.

– Marcia McNutt



Marcia McNutt is  
Editor-in-Chief,  
Science Journals



*“There are costs...  
for re-collecting data  
for new uses.”*

\*D. L. Longo, J. M. Drazen, *N. Engl. J. Med.* **374**, 276 (2016).

## #IAmAResearchParasite

Marcia McNutt

*Science* **351** (6277), 1005.  
DOI: 10.1126/science.aaf4701

### ARTICLE TOOLS

<http://science.sciencemag.org/content/351/6277/1005>

### RELATED CONTENT

<http://science.sciencemag.org/content/sci/351/6277/1024.full>  
<http://science.sciencemag.org/content/sci/351/6277/1007.full>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.