The Availability of Research Data Declines Rapidly with Article Age

Description

Tweet

Summarised on SciDevNet, as â??*Most research data lost as scientists switch storage techâ?•* from this source:

Current Biology, 19 December 2013
Copyright Â© 2014 Elsevier Ltd All rights reserved.
10.1016/j.cub.2013.11.014

Authors

Timothy H. Vinessend email, Arianne Y.K. Albert, Rose L. Andrew, Florence DÃ©barre, Dan G. Bock, Michelle T. Franklin, Kimberly J. Gilbert, Jean-SÃ©bastien Moore, SÃ©bastien Renaut, Diana J. Rennison

Highlights

- We examined the availability of data from 516 studies between 2 and 22 years old
- The odds of a data set being reported as extant fell by 17% per year
- Broken e-mails and obsolete storage devices were the main obstacles to data sharing
- Policies mandating data archiving at publication are clearly needed

Summary

â??Policies ensuring that research data are available on public archives are increasingly being implemented at the government [1], funding agency [2,3,4], and journal [5,6] level. These policies are predicated on the idea that authors are poor stewards of their data, particularly over the long term [7], and indeed many studies have found that authors are often unable or unwilling to share their data [8,9, 10,11]. However, there are no systematic estimates of how the availability of research data changes with time since publication. We therefore requested data sets from a relatively homogenous set of 516 articles published between 2 and 22 years ago, and found that availability of the data was strongly affected by article age. For papers where the authors gave the status of their data, the odds of a data set being extant fell by 17% per year. In addition, the odds that we could find a working e-mail address for the first, last, or corresponding author fell by 7% per year. Our results reinforce the notion that, in the long term, research data cannot be reliably preserved by individual researchers, and further demonstrate the urgent need for policies mandating data sharing via public archives.â?•

Rick Davies comment: I suspect the situation with data generated by development aid projects (and their evaluations) is much, much worse. I have been unable to get access to data generated within the last 12 months by one DFID co-funded project in Africa . I am now trying to see if data used in a recent analysis of the (DFID funded) Chars Livelihoods Programme is available.

I am also making my own episodic attempts to make data sets publicly available that have been generated by my own work in the past. One is a large set of hosuehold survey data from Mogadishu in 1986, and another is household survey data from Vietnam generated in 1996 (baseline) and 2006 (follow up). One of the challenges is finding a place on the internet that specialises in making such data available (especially development project data). Any ideas?

PS 2014 01 07: Missing raw data is not the only problem. Lack of contact information about the evaluators/researchers who were associated with the data collection is another one. In their exemplary blog about their use of QCA Raab and Stuppert comment about their search for evaluation reports:

> â??Most of the 74 evaluation reports in our first coding round do not display the evaluatorâ??s or the commissionerâ??s contact details. In some cases, the evaluators remain anonymous; in other cases, the only e-mail address available in the report is a generic info@xyz.org. This has surprised us â?? in our own evaluation practice, we always include our e-mail addresses so that our counterparts can get in touch with us in case, say, they wish to work with us againâ?•

PS 2014 02 01 Here is another interesting article about missing data and missing policies about making data available: Troves of Personal Data, Forbidden to Researchers (NYT, May 21, 2012)

> â??At leading social science journals, there are few clear guidelines on data sharing. â??The American Journal of Sociology does not at present have a formal position on proprietary data,â?• its editor, Andrew Abbott, a sociologist at the University of Chicago, wrote in an e-mail. â??Nor does it at present have formal policies enforcing the sharing of data.â?•
>
> The problem is not limited to the social sciences. A recent review found that 44 of 50 leading scientific journals instructed their authors on sharing data but that fewer than 30 percent of the papers they published fully adhered to the instructions. A 2008 review of sharing requirements for genetics data found that 40 of 70 journals surveyed had policies, and that 17 of those were â??weak.â?•â?•

Category

1. Uncategorized

Date
16/11/2026
Date Created
22/12/2013
Author
admin