

Subject: Request for Information: Public Access to Federally Funded Research

Date: January 7, 2012 3:46:42 PM EST

To Whom It May Concern:

The quantity of research being produced and published has recently increased dramatically. The advent of the Internet allows scientists to quickly share new research and also reduces the burden of making results and publications publicly available. As the Internet has grown, we have the capacity not only to share written articles but raw data itself. This takes sharing and collaboration to new heights. Articles can be read but with the actual data, results can be double checked, other researchers in academia, industry, members of the public, anyone can look over the data and draw their own conclusions. Nevertheless, despite the capacity for sharing and making data and articles publicly available, the system is still holding on to the antiquated ways of the past---professional journals where no data is published and the articles themselves are only accessible to members of universities or those who can afford the high price of their own subscription.

This hampers research advancement. This restricts access to knowledge to the elite. The motto of the University of California system is fiat lux---let there be light. What good is light hidden away from the public?

The Office of Science and Technology Policy is in a unique position to address these issues and better science and the American public. First by requiring all publications resulting from federally funded research to be published in open access journals. Further, requiring the data collected with federal funds to be deposited in a public repository. One concern with publicly shared data is that the original collectors may not receive adequate credit. If there was a single, public data repository. Each researcher could be given their own, unique ID. Any data deposited to the repository would be linked to their ID. All federally funded articles would have to not only be open, but indicate what data set was used for the research. This would ensure that federally funded research had both the articles and data made publicly available. It would also mean that the original collectors of the data would have to be given credit, inhibiting data theft or plagiarism.

Why is it so important that data be publicly available? First, if the public pays taxes to support it, they should be able to access the

results. Second, with the data, it is easier to verify that analyses and conclusions drawn in articles accurately reflect the data. Third, individual studies are often small or conducted on specific populations. If many such studies are carried out, it is possible to use meta-analysis to combine results and drawn more reliable, accurate conclusions by pooling data from many studies. Theoretically many journals require authors to share data for purposes of verification or meta-analysis, but my experience is that reality falls woefully short of this ideal. Public access to data would greatly enhance the ability of researchers to accurately combine and synthesize the results of many studies. Obviously, there are many details that would need to be considered. Some data sets contain sensitive data and mechanisms would be required to maintain confidentiality in this case.

Some change can come organically from the ground up, other change requires a top down approach. There have been some efforts to develop data repositories and open access journals, but the most benefit would come from something similar to the www.pubmed.gov website. A centralized place for articles and data where researchers and public alike can read and learn.

Sincerely yours,

Joshua Wiley

--

Joshua Wiley
Ph.D. Student, Health Psychology
Programmer Analyst II, Statistical Consulting Group
University of California, Los Angeles
<https://joshuawiley.com/>