

**U.S. HOUSE OF REPRESENTATIVES
COMMITTEE ON SCIENCE, SPACE, AND TECHNOLOGY**

HEARING CHARTER

Scientific Integrity and Transparency

Tuesday, March 5, 2013

10:00 a.m. – 12:00 p.m.

2318 Rayburn House Office Building

Purpose

At 10 AM on Tuesday, March 5, 2013, the Subcommittee on Research will hold a hearing titled *Scientific Integrity and Transparency*. This hearing will provide Members an opportunity to understand the problem of access to underlying data from published research funded by the federal government, and why access to this underlying data is vital to scientific integrity and transparency for peer reviewed research. On March 29th, 2012 the Investigation and Oversight Subcommittee held a hearing entitled *Federally Funded Research: Examining Public Access and Scholarly Publication Interests*.¹ The focus of this past hearing was on open access to publications, whereas the focus of this hearing is on open access to data used in federal research.

Witnesses

- Prof. Bruce Alberts, Professor of Biochemistry, University of California San Francisco
- Prof. Victoria Stodden, Assistant Professor of Statistics, Columbia University
- Dr. Stanley Young, Assistant Director for Bioinformatics, National Institute of Statistical Sciences
- Mr. Sayeed Choudhury, Associate Dean for Research Data Management at Johns Hopkins University and Hodson Director of the Digital Research and Curation Center

Overview

The bedrock of the scientific process is the ability to replicate the experimental claims made by researchers. These claims include both the generation of data and the analysis of data by computer software and code. Scientists rarely reproduce the work of others since they neither have the time nor the resources to reliably replicate the work of their colleagues; instead, they often trust these claims and rely on the peer review process and their colleagues to share their data and analysis methods when needed. This exchange allows for scientists and companies to exploit the latest insights to develop new directions in their research, and allows them to maximize the impact of federal research investment. Thus, scientific progress cannot occur unless there is a strong culture of integrity and transparency.

¹ <http://science.house.gov/hearing/subcommittee-investigations-and-oversight-hearing-examining-public-access-and-scholarly>

Unfortunately, the current system has demonstrated several flaws. The current incentive system rewards researchers who publish in journals, but preparation of data for others' use is not an important part of this reward structure. The process of peer review, which the scientific community views as its primary means to check scientific integrity in journal publications, oftentimes does not try to replicate the results of submitted papers. Fellow researchers conducting the peer review for publication rarely ask for the original data of the submitted paper they are reviewing, and focus instead on whether the claims made in the paper are plausible. They simply assume the underlying data is valid. In a recent study by Young and Karr, upwards of 90% of clinical trial claims for new medicines cannot be replicated.² The inability to replicate published results is not unique to clinical trials and occurs across scientific disciplines.³

This hearing will attempt to understand the scope of the problem with scientific integrity, especially how thorough researchers deal with underlying data. This issue of scientific integrity should be differentiated from cases of scientists knowingly and intentionally committing scientific fraud, fabricating data, or plagiarism though these might be inter-related depending on individual circumstances. This hearing will focus primarily on how data is collected, shared, and analyzed by the scientific community and policies for what, how, and when federally funded research data should be shared, as well as the cost of making this data available to the scientific community and public.

Current federal laws governing the sharing of data include the Data Access Act (DAA) of 1999 and the Information Quality Act (IQA) of 2001.⁴ Introduced by Senator Richard Shelby, the DAA (sometimes known as “the Shelby Amendment” within the science community) requires that data from federally funded research be made available under the Freedom of Information Act procedures. The IQA requires the OMB to issue regulations for ensuring the quality and integrity of all information disseminated by federal agencies. However, the Government Accountability Office reported in September 2007 that federal agencies rarely monitor whether researchers make data available.⁵

In response to these aforementioned issues, the Office of Science and Technology Policy (OSTP) released guidance to federal agencies on February 22nd about increasing access to the results of federally funded scientific research which includes a discussion about access to non-classified digital data. In this memo, OSTP outlines the following principles for federal funding agencies to follow when issuing a data access plan⁶:

- Maximize access to scientific data created with federal funds;

² <http://science.house.gov/sites/repUBLICANS.science.house.gov/files/documents/hearings/HHRG-112-SY20-WState-SYoung-20120203.pdf>

³ “Again, and again, and again ...” p1225 Science Vol 334 2 December 2011

⁴ National Research Council, Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age (Washington, DC: National Academy Press), 2009.

⁵ <http://www.gao.gov/products/GAO-07-1172>

⁶ http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

- Ensure that researchers develop data management plans, and allow inclusion for costs in proposals along with proper evaluations of these proposals;
- Include mechanisms to ensure compliance with data management plans and policies;
- Promote the deposit of data in publicly accessible databases;
- Encourage cooperation with the private sector to improve data access and compatibility;
- Develop approaches for identifying/providing appropriate attribution to data sets;
- Support the training, education and workforce development related to data management; and
- Provide assessment of long-term needs for the preservation of scientific data.

This hearing will address how such principles might best be implemented by federal research agencies and members of the scientific community conducting such research.