

---

# Data Management Plans:

---

Bruce Herbert, Office of Scholarly Communications  
Texas A&M

# The Fourth Paradigm: Data-Intensive Scientific Discovery

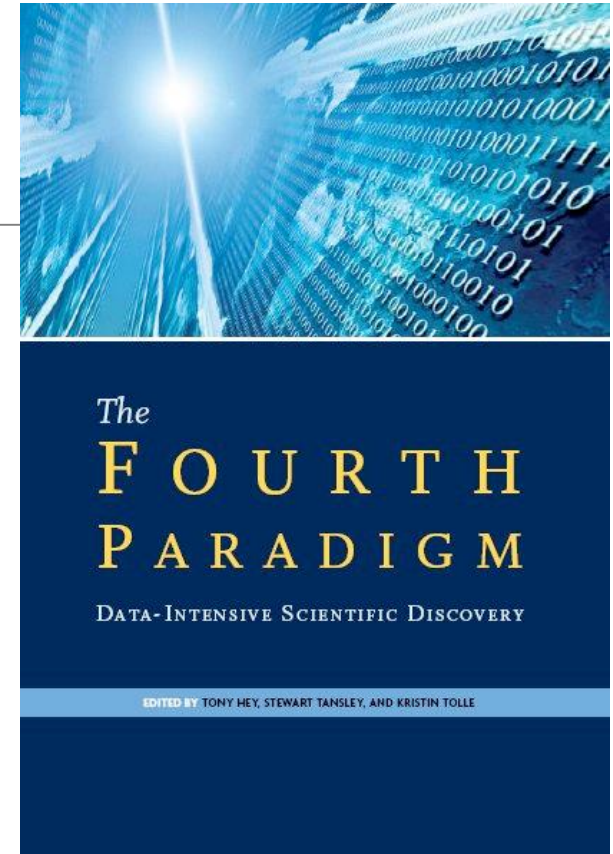
---

“Jim Gray described his vision of the fourth paradigm of scientific research.

He outlined a two-part plea for the funding of tools for **data capture, curation**, and analysis, and for a **communication and publication infrastructure**.

He argued for the establishment of modern stores for data and documents that are on par with traditional libraries.”

<http://research.microsoft.com/en-us/collaboration/fourthparadigm/>



# Federal Mandates For Public Access to Research



the WHITE HOUSE PRESIDENT BARACK OBAMA

Get Email Updates | Contact Us

BLOG PHOTOS & VIDEO BRIEFING ROOM ISSUES the ADMINISTRATION the WHITE HOUSE our GOVERNMENT

Home • The Administration • Office of Science and Technology Policy

Search WhiteHouse.gov

Office of Science and Technology Policy

About OSTP | Pressroom | OSTP Blog | Divisions | Initiatives | R&D Budgets | Resource Library | NSTC | PCAST | Contact Us

## Expanding Public Access to the Results of Federally Funded Research

Posted by Michael Stebbins on February 22, 2013 at 12:04 PM EST

E-Mail | Tweet | Share | +

The Obama Administration is committed to the proposition that citizens deserve easy access to the results of scientific research their tax dollars have paid for. That's why, in a policy memorandum released today, OSTP Director John Holdren has directed Federal agencies with more than \$100M in R&D expenditures to develop plans to make the published results of federally funded research freely available to the public within one year of publication and requiring researchers to better account for and manage the digital data resulting from federally funded scientific research. OSTP has been looking into this issue for some time, soliciting broad public input on multiple occasions and convening an interagency working group to develop a policy. The final policy reflects substantial inputs from scientists and scientific organizations, publishers, members of Congress, and other members of the public—over 65 thousand of whom recently signed a *We the People* petition asking for expanded public access to the results of taxpayer-funded research.

To see the new policy memorandum, please visit: [http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf)

To see Dr. Holdren's response to the *We the People* petition, please visit: <https://petitions.whitehouse.gov/response/increasing-public-access-results-scientific-research>

Michael Stebbins is Assistant Director for Biotechnology at OSTP

GIVE FEEDBACK ABOUT THIS PAGE

YOUR FEDERAL TAXPAYER RECEIPT

Launch the Receipt

WHITE HOUSE BLOGS

The White House Blog  
Middle Class Task Force  
Council of Economic Advisers  
Council on Environmental Quality  
Council on Women and Girls

## The Library Supports:

Publication repositories

Tools to create data management plans

TDL Data repository

Workflows, standards, & policies

<http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research>



# Use Case:

---

**Title:** Researcher needs to make their research data publicly available


**Primary Actors:**

PIs of federally funded research

Researchers working on unfunded research or funded research with no retention requirements

Graduate students working on theses, dissertations, or other data-generating projects.

# Sharing Detailed Research Data Is Associated with Increased Citation Rate

Heather A. Piwowar , Roger S. Day, Douglas B. Fridsma

Published: March 21, 2007 • DOI: 10.1371/journal.pone.0000308 • Featured in PLOS Collections

## Article

## Authors

## Metrics

## Comments

## Related Content



## Abstract

[Introduction](#)[Results](#)[Discussion](#)[Materials and Methods](#)[Supporting Information](#)[Author Contributions](#)[References](#)[Reader Comments \(6\)](#)[Media Coverage \(0\)](#)[Figures](#)

## Abstract

### Background

Sharing research data provides benefit to the general scientific community, but the benefit is less obvious for the investigator who makes his or her data available.

### Principal Findings

We examined the citation history of 85 cancer microarray clinical trial publications with respect to the availability of their data. The 48% of trials with publicly available microarray data received 85% of the aggregate citations. Publicly available data was significantly ( $p = 0.006$ ) associated with a 69% increase in citations, independently of journal impact factor, date of publication, and author country of origin using linear regression.

### Significance

This correlation between publicly available data and increased literature impact may further motivate investigators to share their detailed research data.

## Figures



545  
Saves

152  
Citations

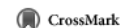
43,264  
Views

107  
Shares

Download PDF

Print

Share



CrossMark

Included in the  
Following Collection

[Open Access Collection](#)

## Subject Areas

Microarrays

Linear regression an...

Internet

Archives

Gene expression

Clinical trials

Confidence intervals

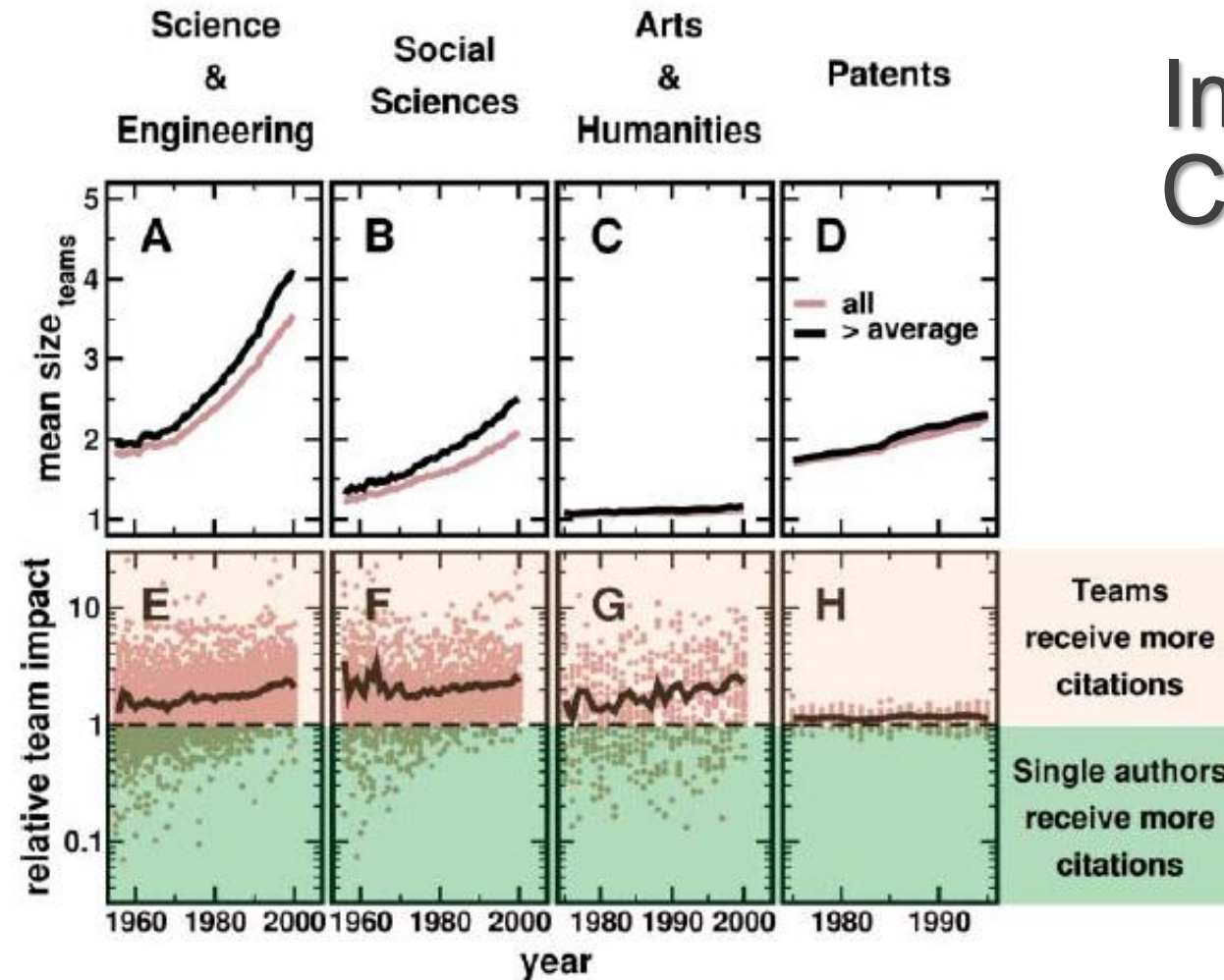
DNA sequence anal...

# Open Data Data Citation

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000308>



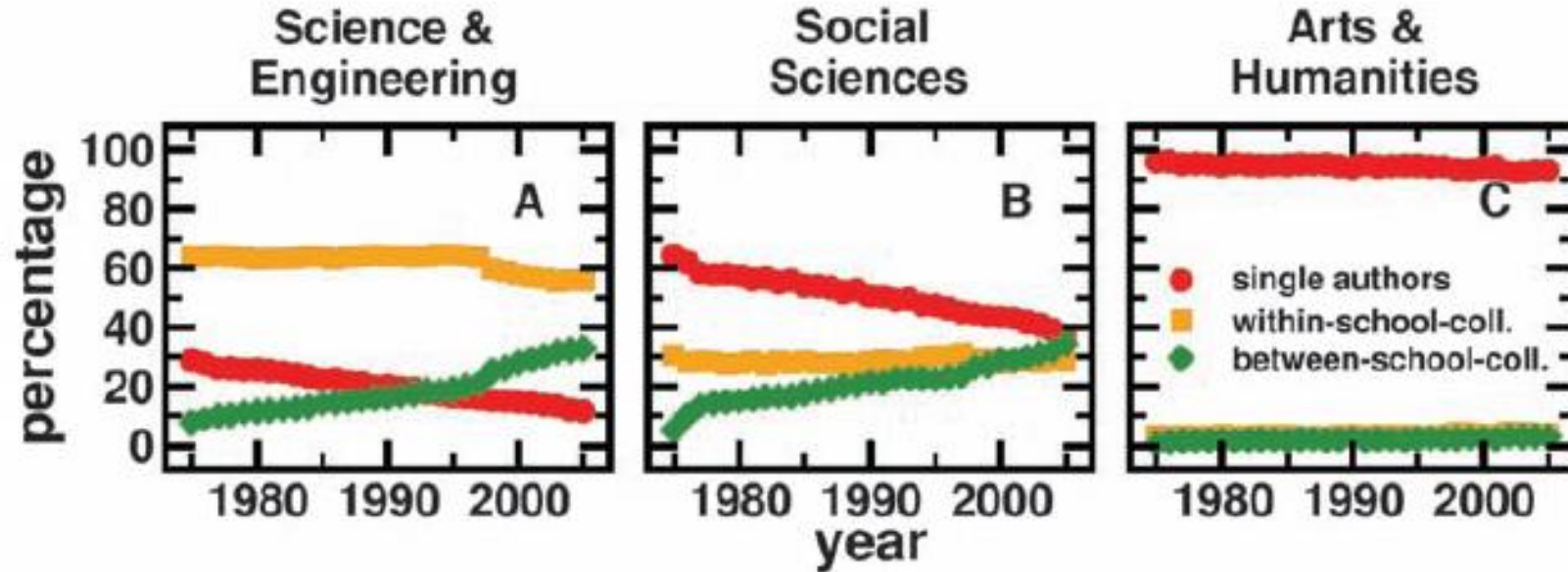
# Impact of Team Collaboration



**Fig. 2.** The relative impact of teams. (A to D) Mean team size comparing all papers and patents with those that received more citations than average in the relevant subfield. (E to H) The RTI, which is the mean number of citations received by team-authored work divided by the mean number of citations received by solo-authored work. A ratio of 1 indicates that team- and solo-authored work have equivalent impact on average. Each point represents the RTI for a given subfield and year, whereas the black lines present the arithmetic average in a given year.

Wuchty et al. (2007). *Science* 316(5827): 1036-1039.

# Collaboration Across Institutions



**Fig. 1.** The rise in multi-university collaboration. By comparing the incidence of papers produced by different authorship structures, we see that the share of multi-university collaborations strongly increases from 1975 to 2005. This rise is especially strong in SE (A) and SS (B), whereas it appears weakly in AH (C), in which collaboration of any kind is rare. The share of single-university collaborations remains roughly constant with time, whereas the share of solo-authored papers strongly declines in SE and SS.

# Use Case:

---

**Title:** Researcher shares active data within a trusted, collaborative network

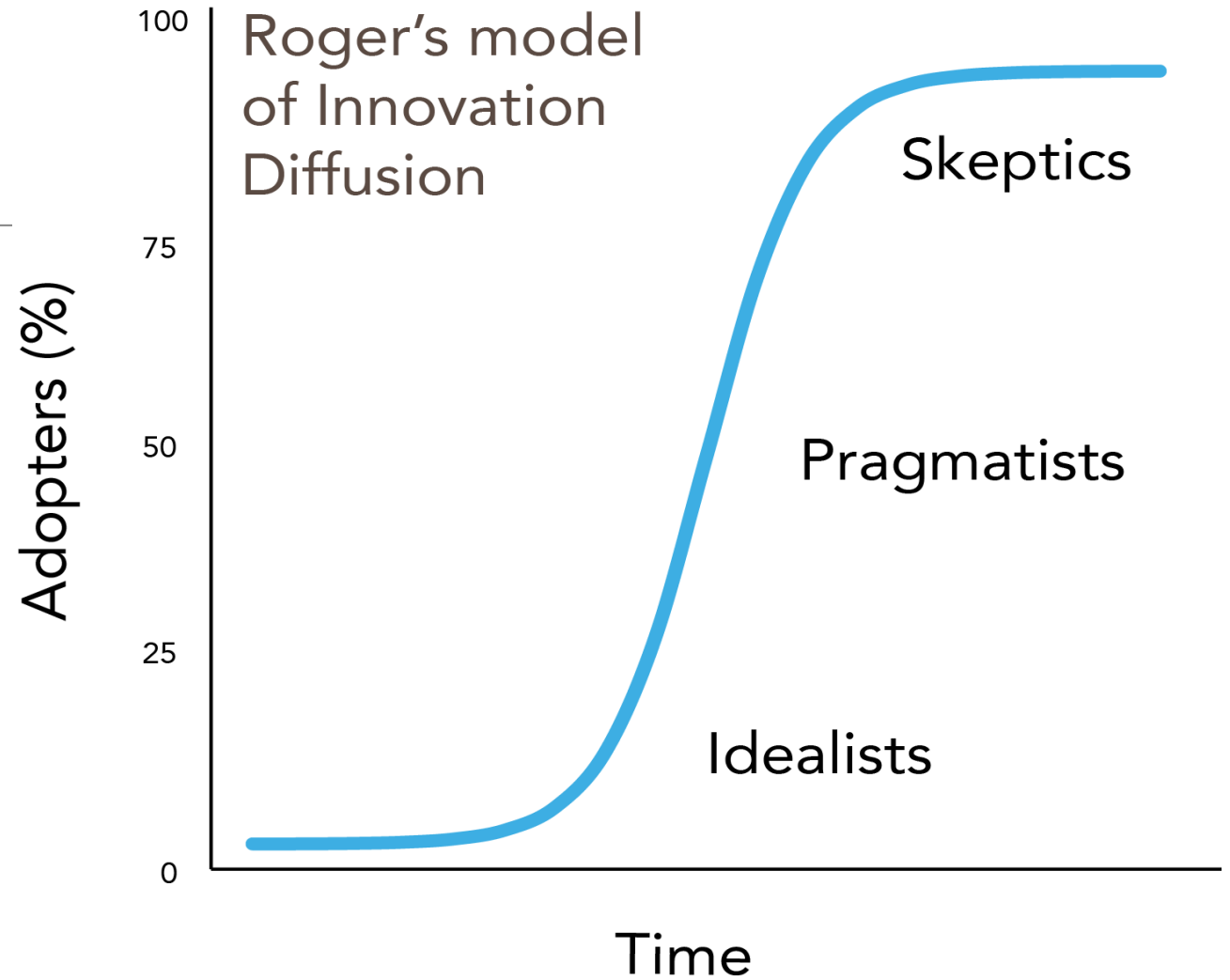
**Primary Actors:**

Researchers involved in collaborative teams or networks



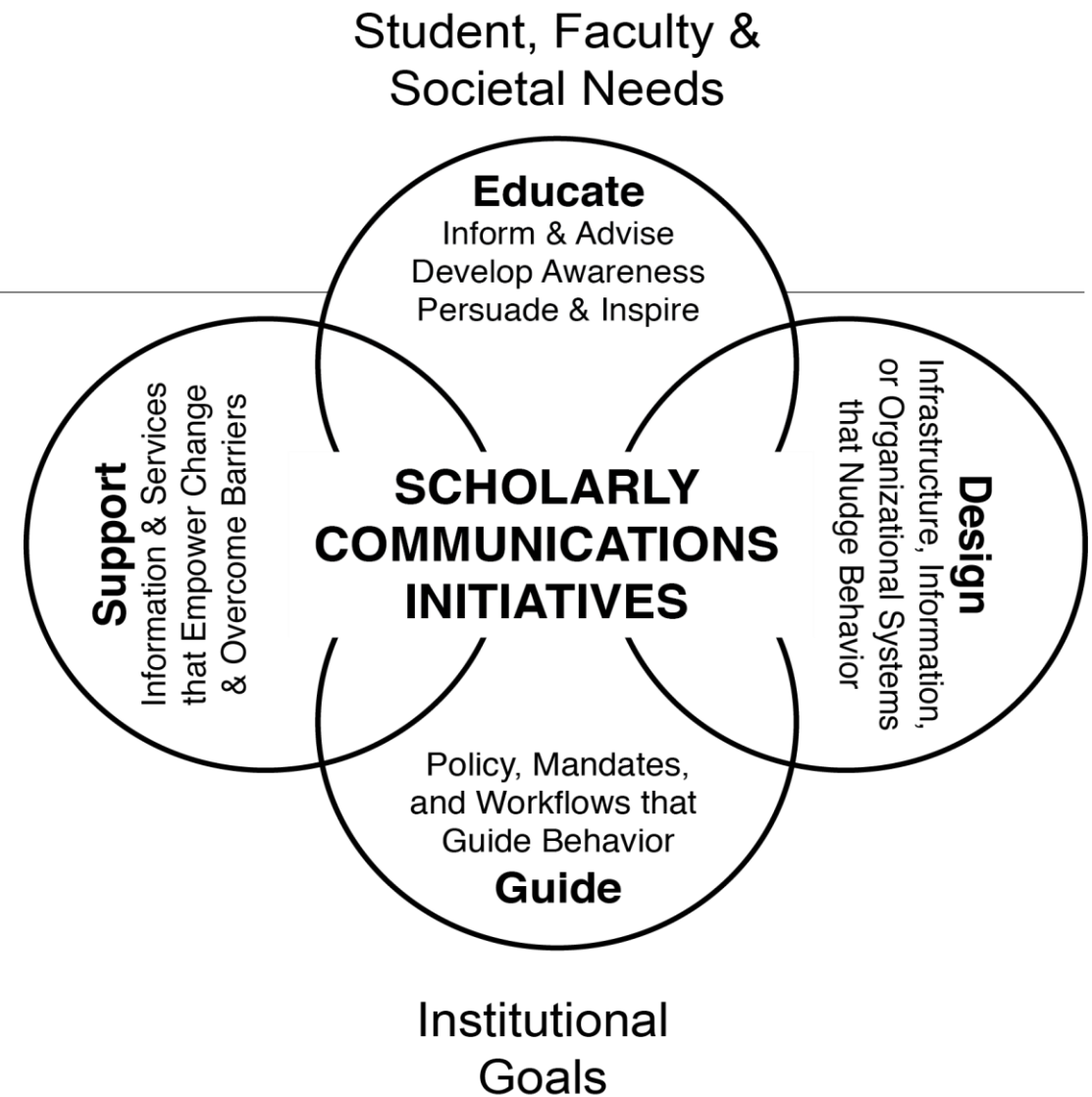
# Motivation

---



# Open Data Systemic Response

---



# Open Data TAMU Program Elements

Current Federal public access mandates for different agencies. I find the new SPARC database useful.

- <https://www.whitehouse.gov/blog/2016/10/28/federally-funded-research-results-are-becoming-more-open-and-accessible>
- <http://researchsharing.sparcopen.org/>

Reading requests for proposals for DMP instructions. Examples from NSF and NIH are below.

- NSF Grant Application Guide: [https://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/gpg\\_index.jsp](https://www.nsf.gov/pubs/policydocs/pappguide/nsf16001/gpg_index.jsp)
- NSF Data Management Plans FAQs: <https://www.nsf.gov/bfa/dias/policy/dmpfaqs.jsp>
- NIH Data Sharing Instructions: [https://grants.nih.gov/grants/policy/data\\_sharing/](https://grants.nih.gov/grants/policy/data_sharing/)
- NIH Grant Application Guide: <https://grants.nih.gov/grants/how-to-apply-application-guide.html#inst>

Introduction and use of the DMPTool: <https://dmptool.org/>

Examples of Data Management Plans: <https://www.lib.umn.edu/datamanagement/DMP/example>