

# Overview of Cornell University Policy 4.20 Research Data Retention

Mark Hurwitz, Ph.D., P.E.  
Chief Research Compliance Officer

## What is the policy?

Accurate and detailed research records are an essential component of any research project.

Cornell researchers and the University share responsibility in collecting, retaining, securing, accessing, and publishing or disseminating research data.

# Not What We Want!

The New York Times

## ***Duke University to Pay \$112.5 Million to Settle Claims of Research Misconduct***

March 25, 2019

THE UNITED STATES ATTORNEY'S OFFICE

MIDDLE DISTRICT *of* PENNSYLVANIA

**Former Penn State Professor Who Plead Guilty In \$3 Million Federal Research Grant Fraud Sentenced In Federal Court**

FOR IMMEDIATE RELEASE

November 30, 2012

## Cornell nutrition scientist resigns after retractions and research misconduct finding

By Kelly Servick | Sep. 21, 2018, 11:25 AM

(AAAS Science Website)

tctMD

News

Conferences

Slides

Specialties

Podcasts

Vid



NEWS • Daily News



### **Stem Cell Research—Shattered**



### **After Fabrication Scandal—**



### **Needs to Rebuild, Says EHI**



### **Editor**

SCIENTIFIC  
AMERICAN.



News Blog

## **Tennessee physicist sentenced to 4 years for sharing drone plans with foreign students**

By Brendan Borrell on July 2, 2009

# Who owns research data?

- **Cornell asserts ownership** of research data and related property rights arising from the activities of its researchers and others who use university resources, including those provided through an externally funded grant, contract, or other type of award or gift to the university.
- **PIs are the custodians** of their research data and are responsible to Cornell University for the proper use, access, and control of any research data under their management or supervision, including the use of data in scholarly publications and presentations.

# Reproducibility is essential to the advancement of science.

- **However:** The task of retrieving data by someone not familiar with the research can be difficult and expensive.
- Cornell recognizes that researchers have many ways of organizing their data and labeling their files and does not impose any specific required organization.
- The PI is only expected to maintain the research data in a form reasonably well organized according to the norms of the field of research.

## PI Responsibilities Include...

- Collecting, maintaining, retaining, and providing access to research data;
- Determining whether research data is public, confidential or otherwise restricted;
- Assuring compliance with any restrictions mandated by contract or federal regulations;
- Securing and controlling access as may be required;
- Determining (within any governing obligations) how data will be published or presented;
- Ensuring that data are available for review by the university, sponsors, journals, et cetera;
- Informing the Center for Technology Licensing, through the Technology Transfer Process, of any data supporting a new invention; and
- Providing a data retention statement to OVPRI and updating it as necessary.

## University Responsibilities Include...

- Protecting the rights of faculty, staff, and students;
- Complying with the terms of sponsored project agreements;
- Ensuring the appropriate use of animals, human subjects, recombinant DNA, biological agents, and radioactive materials;
- Securing the intellectual property rights of the university;
- Supporting, through OVPRI, the PI's research data management efforts;
- Resolving disputes among researchers over data control or access;
- Approving transfer of original research data off-campus;
- Facilitating the investigation of charges, such as research misconduct or conflict of interest; and
- Provide a suitable repository for PI data retention attestations.

## Attestation (CULearn Course)

- I have read and understood the Research Data Retention Policy 4.20 as presented in this course.
- I will abide by the policy in collecting, storing, and retaining data generated by or used in my Cornell research.

Agree

Return to  
Course



# Research Data Services at Cornell

Wendy Kozlowski, RDMSG Coordinator, wak57

[data.research.cornell.edu](http://data.research.cornell.edu) | [rdmsg-help@cornell.edu](mailto:rdmsg-help@cornell.edu)

2021-06-17

# Data Retention

## Store your data at Cornell

- Evaluate options for data storage: <https://finder.research.cornell.edu/>
- Learn best practices for storage and backup: <https://data.research.cornell.edu/content/data-storage-and-backup>
- Meet funder and publisher requirements for sharing and preservation

|   |  |  |
|---|--|--|
| <b>CAC Archival Storage</b><br>Single copy, non-mountable storage | <b>CAC Red Cloud Storage</b><br>Storage for Red Cloud compute instances                                | <b>CISER Research Servers</b><br>Storage for CISER computing environment   |
| <b>Cornell Box</b><br>Online file sharing and collaboration       | <b>Cornell Restricted Access Data Center</b><br>Storage for CRADC (confidential) computing environment | <b>EZ-Backup Static Storage</b><br>Archival storage and backup storage   |
| <b>LabArchives</b><br>Online electronic lab notebook              | <b>Shared File Services</b><br>File sharing between users and computers                                | <b>Shared File Services - Confidential</b><br>File sharing between users and computers for confidential data (non-WCM) |



[labarchives.cornell.edu](https://labarchives.cornell.edu)

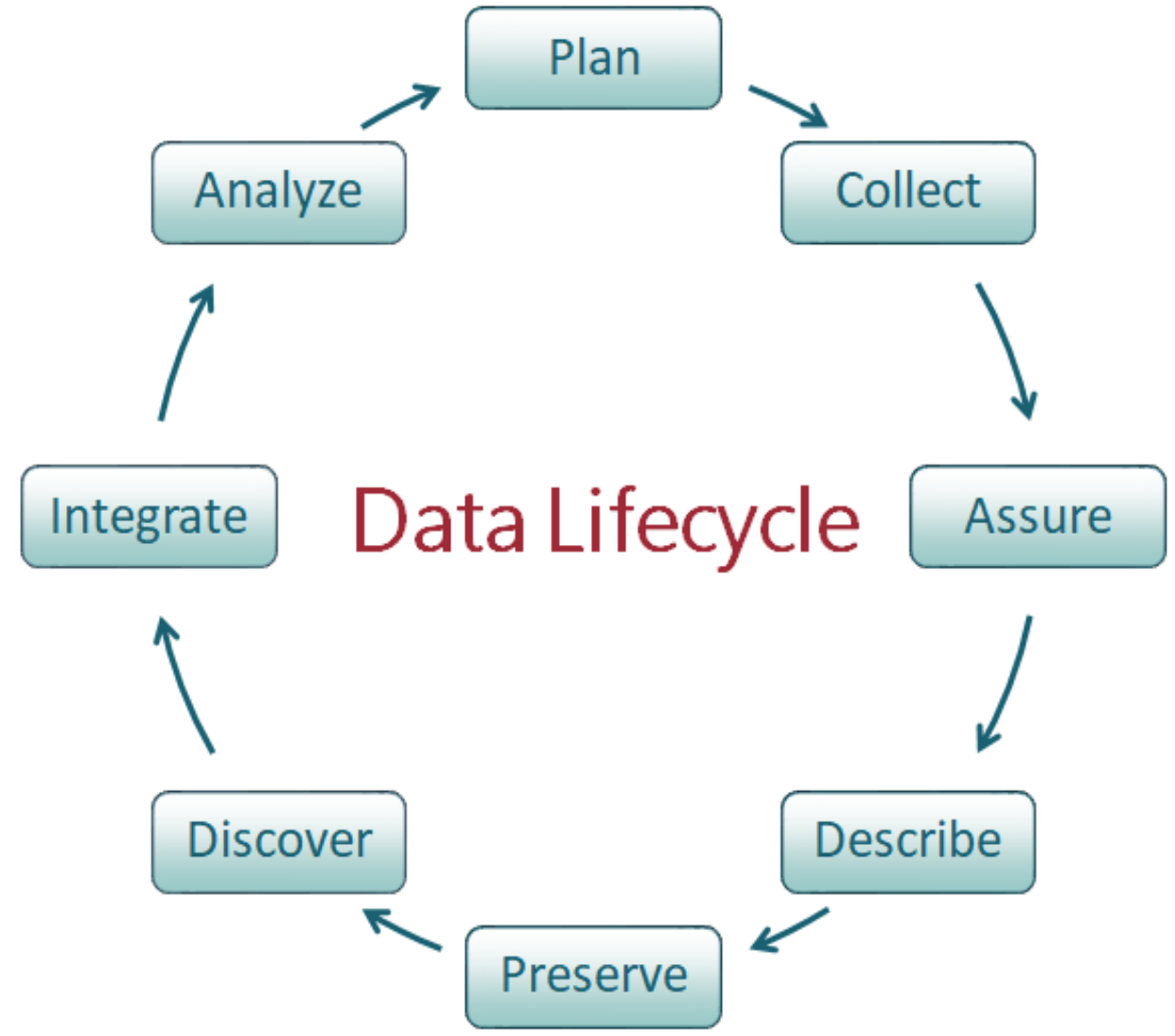


Open Science Framework

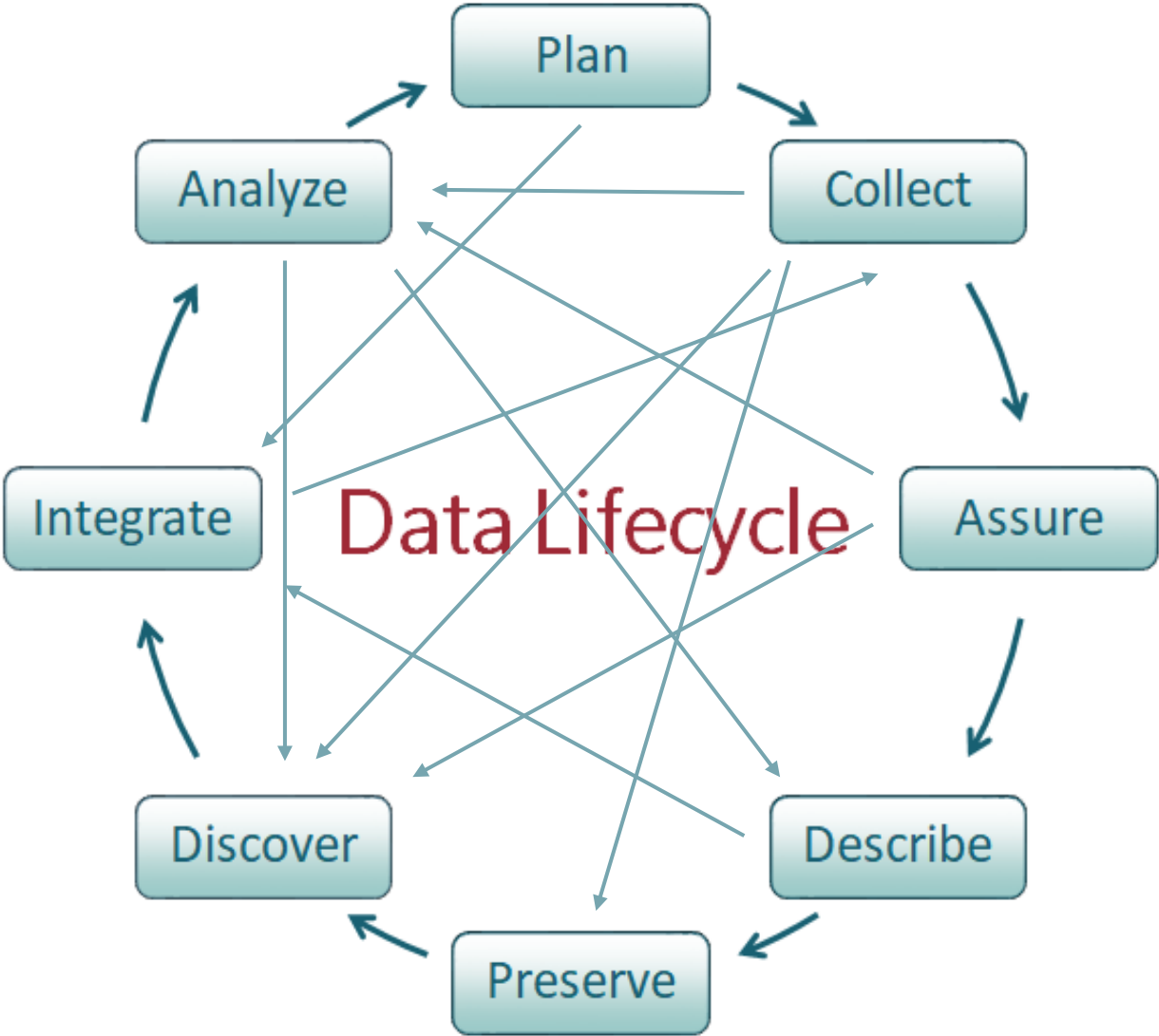
[osf.io](https://osf.io)



# Data Lifecycle



# Data Lifecycle in Reality



# Plan

## Data Management Planning Support

- DMP preparation consultation, editing and review by RDMSG consultants
- <https://data.research.cornell.edu/content/data-management-planning>
- Access to DMP Tool (via Cornell NetID) for online creation of and guidance on DMP writing
- Consultations for group, lab or project data management



**DMPTool**

Build your Data Management Plan

Collect

## Get the data you need

- Consult with experts for assistance finding and obtaining data



**The Qualitative Data  
Repository**



[data-services@ropercenter.org](mailto:data-services@ropercenter.org)

# Describe

## Get help creating quality metadata to document and describe your work.

- Find tools to help you create and maintain your documentation and records.
- <https://data.research.cornell.edu/content/writing-metadata>
- Readme template at:  
<https://data.research.cornell.edu/content/readme>



**DOCUMENT**



Informal ReadMe

Formal Schema

**Lower-Barrier**  
Fast  
Easy

**Lower-Potential**  
Irregular  
Incomplete

**Higher-Potential**  
Standardized  
Machine actionable

**Higher-Barrier**  
Slow  
Skilled

# Analyze



**CAC**  
*we enable your success*



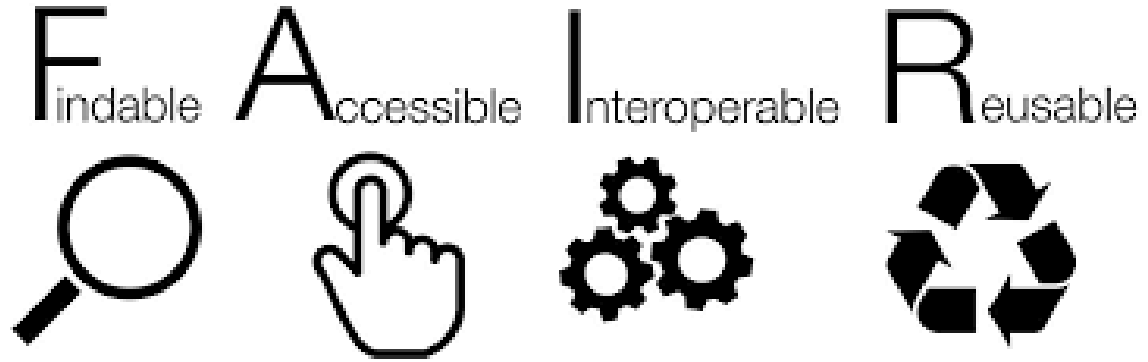
## Tools and infrastructure

- Cornell Restricted Access Data Center (CRADC): secure environment for working with confidential or restricted data
- Biotechnology Resource Center: HPC Computing with bioinformatics analysis software
- Cornell Advanced Computing (CAC): HPC and related services
- CUL Digital Co-Lab: Digital humanities, interpretive social sciences, computational text analysis, and privacy services
- Many service providers offer workshops, training and consultation on tools and best practices related to data analysis

<https://ciser.cornell.edu/data/secure-data-services/>  
<https://www.biotech.cornell.edu/>  
<http://cac.cornell.edu/>  
<https://digitalscholarship.library.cornell.edu/>



Assure



<https://data.research.cornell.edu/content/preparing-fair-data-reuse-and-reproducibility>



**Results Reproduction  
(R-Squared)**

CCSS's Results Reproduction service (R-Squared) computationally reproduces the results of your research to ensure Reproducibility and Transparency.

<https://ciser.cornell.edu/research/results-reproduction-r-squared-service/>

[ccssdata@cornell.edu](mailto:ccssdata@cornell.edu)

Preserve

<https://ecommons.cornell.edu>  
<https://cugir.library.cornell.edu/>  
<https://archive.ciser.cornell.edu/>



CORNELL UNIVERSITY LIBRARY

eCommons

Open scholarship at Cornell



CUGIR

Cornell University Geospatial  
Information Repository



CISER Data &  
Reproduction Archive

Preserve and share  
your data in Cornell-  
based repositories

- Curatorial review
- Open access
- Persistent identifiers
- Restricted access if needed



DATA  
CURATION  
NETWORK



# Cornell Research Data Management Service Group

Research data consultation and referral services for Cornell faculty, staff, and students.

<https://data.research.cornell.edu> | [rdmsg-help@cornell.edu](mailto:rdmsg-help@cornell.edu)