



Cornell University Policy Office
policy.cornell.edu

Interim Policy 4.20 Research Data Retention

Responsible Executive: Vice Provost for Research

Responsible Office(s): Office of the Vice Provost for Research

Interim Version Issued: August 7, 2020

Last Updated: August 7, 2020

Table of Contents

| | |
|---|-----------|
| 1. Policy Statement and Requirements | 1 |
| 1.1 Policy Statement | 1 |
| 1.2 Scope | 1 |
| 1.3 Policy Requirements | 2 |
| 2. To Whom This Policy Applies | 6 |
| 3. Definitions | 7 |
| 4. Policy Administration | 8 |
| 5. Appendix | 8 |
| 5.1 Related Resources | 8 |
| 6. Revision History | 10 |

1. Policy Statement and Requirements

1.1 Policy Statement

Accurate and detailed research records are an essential component of any research project. The long-term integrity of the record of research conducted at Cornell University requires that research data be preserved in sufficient detail and for an adequate period of time to enable appropriate responses to questions about accuracy, authenticity, authorship, adherence to funder/sponsor requirements, and compliance with laws and regulations governing the research. This policy defines the shared responsibilities of Cornell researchers and the university in collecting, retaining, securing, obtaining, and sharing research data.

1.2 Scope

Cornell University is committed to maintaining the highest standards of research and complying with all funder, federal, and state regulations related to research integrity. This policy is intended to ensure a common understanding of the responsibilities of researchers with respect to maintaining research data.

This policy addresses retention of research data only. Retention and protection of all other Cornell Institutional data are addressed by University Policy 4.7, Retention of University Records, and University Policy 5.10, Information Security.

1.3 Policy Requirements

The following key principles guide this document and the university's commitment to the responsible management of research data.

1.3.1 High-quality research, academic integrity and reproducibility: Sound stewardship of research data, which includes maintaining the highest standards in the generation, use, management, and retention of research data, is a fundamental requirement of ensuring high-quality research and academic integrity. In addition, public access to research data, within legal, ethical, and regulatory constraints, is an enabling factor in Cornell's mission to discover, preserve, and disseminate knowledge.

In collaborative research, whether with other Cornell researchers or with researchers at other institutions, lack of clarity concerning rights to data and means of sharing data can lead to acrimonious disputes within the collaboration and, in extreme cases, allegations of misconduct. This policy does not address specific means of data sharing, but it is noted that a written agreement, such as a formal data use agreement among all collaborators, can prevent misunderstandings and reduce the likelihood of later dispute.

1.3.2 Reproducibility of research results: Reproducibility, which is the ability to verify research findings by other members of the scientific community or by using other methods, is essential to the advancement of science. This ability requires access to relevant research data, materials, documents, protocols, methods, and procedures.

1.3.3 Stewardship of private, confidential, or proprietary data: Cornell is committed to safeguarding the privacy and security of confidential, restricted, identifiable, or otherwise sensitive data entrusted to its care. The loss or release of such data can lead to significant harms such as privacy violations, identity theft, and financial liability for the university and, in some cases, individual liability for the person who was the custodian of the data. Where the study involves human participants, the Institutional Review Board (IRB), Cornell Compliance Office, and the Vice President for Information Technology and Chief Information Officer have the authority to require safeguards appropriate for the protection of those participants. The university must also comply with external regulations that govern research data, such as Health Insurance Portability and Accountability Act (HIPAA), Family Education Rights and Privacy Act (FERPA), and General Data Protection Regulation (GDPR).

To ensure the integrity of the research process and to comply with sponsor and/or federal regulations, the university must retain research data in sufficient detail and for an adequate period of time to enable appropriate responses to questions about accuracy, authenticity, primacy, and compliance with laws and regulations governing the research.

Principal investigators (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.

Caution: University Policy 1.5, Inventions and Related Property Rights, states "Cornell asserts ownership of inventions and related property rights arising from the activities of its faculty, staff, students, 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." To maintain this ownership, any such invention must be disclosed promptly in writing to the Center for Technology Licensing, and a patent application filed, before any public presentation of the invention.

1.3.4 Responsibilities of Cornell University: 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. Cornell's responsibility for research data, and the scientific record arising from it, is based on federal regulation (such as 2 CFR §200315 Intangible property), other university policies, including University Policy 1.5, Inventions and Related Property Rights, sponsor requirements, and precedent. The responsibilities of the university in this regard include, but are not limited to:

1. Protecting the rights of faculty, staff, and students, including, but not limited to, their rights of access to data from research in which they participated;
2. Complying with the terms of sponsored project agreements;
3. Ensuring the appropriate use of animals, human subjects, recombinant DNA, biological agents, and radioactive materials;
4. Securing the intellectual property rights of the university;
5. Supporting, through the Office of the Vice Provost for Research, the PI's research data management efforts;
6. Resolving disputes among researchers over data control or access;
7. Approving any transfer of original research data off-campus for archival or other purposes;
8. Facilitating the investigation of charges, such as academic misconduct or conflict of interest; and
9. Provide a suitable repository within the Office of the Vice Provost for Research for PI data retention statements.

1.3.5 Responsibilities of the principal investigator (PI): The PI has the right and authority, within the limits set by the superseding authority of Cornell University, agreements with collaborators, and any applicable terms within sponsored agreements, to control the use of, and access to, any research data conducted under their management or supervision, including the use of data in scholarly publications and presentations.

Note: Many journals require co-authors to attest to the accuracy of, and their agreement with, the research results and outcomes.

The PI is responsible, with support from the Office of the Vice Provost for Research, for maintaining and retaining research data in accordance with this policy. Specifically, the PI is responsible for:

1. Collecting, maintaining, retaining, and providing access to research data for the periods required by this policy;
2. Determining, according to the sponsored research agreement, the data use agreement or other requirements, whether research data is public, confidential or otherwise restricted;
3. Assuring compliance with any restrictions mandated by federal International Traffic in Arms Regulations or Export Administration Regulations, including restrictions on publication or sharing with non-U.S. citizens;
4. Securing and controlling access to research data and ensuring that required protections can be provided;
5. Determining, consistent with obligations to sponsors, collaborators, and students, how data will be published or presented;
6. Ensuring that data are available for review by the university, sponsors, journals, and others as described in this policy, journal policies, and sponsor requirements;

7. Informing the Center for Technology Licensing, through the Technology Transfer Process, of any data supporting a new invention; and
8. Providing a data retention statement to the Office of the Vice Provost for Research and updating it as necessary.

1.3.6 Collection and retention:

- Collection, retention, and sharing of research data involving human subjects must comply with Cornell Institutional Review Board (IRB) policies as well as all applicable federal, state, and local laws, regulations, policies, and guidelines.
- The collection, retention, and sharing of research data that incorporates personal health information (PHI) must comply with all applicable Health Insurance Portability and Accountability Act (HIPAA) policies and processes, including security standards.
- Collection, retention, and sharing of research data involving vertebrate animals must comply with University Policy 1.4, Care and Use of Live Vertebrate Animals in Research and Teaching, as well as all applicable federal, state, and local laws, regulations, policies, and guidelines.
- Research data must be retained for a minimum of three years after the final project closeout. Complete data must be retained unless the Vice Provost for Research approves retaining summaries or other secondary data based on compelling justification in special cases. Longer periods of retention are required in some circumstances, such as:
 - The terms of a sponsored research agreement require a longer retention period (for example, New York State typically requires retention for six years after the close of a grant);
 - The Center for Technology Licensing deems retention is required to protect intellectual property;
 - Allegations regarding the research, such as academic misconduct or conflict of interest, arise and remain unresolved. Research data must be retained until such allegations are fully resolved;
 - Legal action, investigation, or official inquiry related to the research is ongoing. Research data must be retained until such issues are fully resolved.

Beyond any required period of retention, the destruction of research records is at the discretion of the PI within limitations imposed by college or department policy, the needs of collaborators or students, or the norms of their field. Records will normally be retained in the unit where they are produced. Research records must be retained in university facilities, or facilities mandated by the research sponsor, or other facilities commonly used for the purpose by peers in the same field of study so long as the data is maintained with appropriate oversight and reasonable means are available for access by Cornell faculty and administrative personnel when needed.

Research data owned by a government or other agency to which a researcher is granted limited access (for example, access through a data use or sponsored research agreement to government administrative data or health records) is not subject to this retention requirement but the steps by which the researcher gained access to such data must be made publicly available so others can apply for access to the same data.

Other private or personal storage of data is not acceptable unless the Vice Provost for Research grants specific permission based on a compelling justification and assurance that data will be maintained with appropriate oversight.

1.3.7 Data security:

- Research data that incorporates confidential information including, but not limited to, personally identifiable human participant data, trade secrets, or export-controlled information, must have the same security protections and be treated in the same manner as Institutional Information classified as “high risk” in University Policy 5.10, Information Security.
- Suspected or proven disclosure or exposure of confidential or otherwise restricted data must be immediately reported to the Vice President for Information Technology and Chief Information Officer.
 - If the data includes data from human participants in research, the Cornell Institutional Research Board must also be informed.
 - If the data was made available by an external sponsor as part of a sponsored research agreement or a limited-use data provider via a data use agreement, the Office of Sponsored Programs must be informed.
 - If the exposed data was export controlled, the Export Control Officer must be informed.
- *Note:* Costs associated with the preservation and security of research data during the term of a sponsored award are typically allowable direct costs of conducting research and Cornell University provides secure cloud storage at very affordable rates. Storage options freely provided to members of the Cornell community are likely to serve the needs of most researchers. Those faculty who work with extremely large data sets may want to take advantage of archives available to their discipline so long as such archives have privacy and security protocols sufficient to address any privacy and security requirements associated with the data.

1.3.8 Access:

- The university has the right to access all research data generated at the university or under its auspices, supported by university administered funds, or conducted using university facilities. Access to data provided to Cornell researchers under a data use agreement, nondisclosure agreement, or other contractual agreement is governed by the terms of such agreement.
- Access to export controlled data is limited to individuals approved by the Cornell Export Control Officer, in accordance with applicable laws and regulations.
- The university has the right to take custody of research data to ensure needed and appropriate access, for example, to facilitate a response to an allegation of research misconduct.
- A principal investigator (PI) may grant Cornell researchers and staff access to research data for research or administrative purposes, subject to all university rules, state and federal laws, and contractual obligations relevant to the data. Access to research data by researchers who are not Cornell University employees may be governed by additional agreements. In collaborative research, whether with other Cornell researchers or with researchers at other institutions, it is generally good practice to put data use agreements in writing, for example by executing a formal data use agreement with all researchers participating in the collaboration.
- Faculty and staff who have the authority to give researchers, whether at Cornell or at another institution, access to data must inform them, in writing, of any limitations or restrictions on the use or dissemination of the data. Faculty and staff may not allow access to any data

obtained under a sponsored research or data use agreement without working with the Office of Sponsored Programs to obtain approval from sponsor or data provider.

- A PI or co-investigator who leaves Cornell may take copies of research data for projects on which they have worked, unless this is prohibited by a data use agreement, sponsor agreement, federal or state law, or other applicable prohibition. Taking copies of data supporting an invention disclosure or an unpublished patent application must be approved by Cornell Technology Licensing. Taking copies of any data covered by a data use agreement or sponsorship agreement must be approved by the Office of Sponsored Programs. The purposes for which such data may be used are dependent on agreement with the PI, or as formally agreed-upon beforehand in a data use agreement. In all cases, the original research data must be retained at Cornell unless the Vice Provost for Research specifically authorizes moving it to another institution.

Caution: The PI is responsible, with support from the Cornell Export Control Officer and the Cornell Office of Sponsored Programs, for assuring compliance with any restrictions mandated by the Office of Foreign Assets Control, the International Traffic in Arms Regulations or Export Administration Regulations. This includes any agreed-upon terms from sponsors and data providers such as publication and sharing with non-U.S. citizen collaborators and/or students.

Note: The task of retrieving data in an intelligible form by someone not familiar with the research can be difficult and expensive. This policy requires the PI to retain research data but does not require the data to be organized in a manner such that a person with no familiarity with the research can immediately find and understand any specific data item. Although the PI is expected to maintain the research data in a form reasonably well organized according to the norms of the field of research, this policy recognizes that researchers have many ways of organizing their data and labeling their files and does not impose any specific required organization.

1.3.9 Publications: The principal investigator (PI) has the right and responsibility to ensure that research is accurately reported to the scientific and academic community, as well as to select the vehicle most appropriate for publication or presentation of research data and results. In the case of research conducted with a co-PI(s), the co-PI(s) jointly share this right and responsibility unless it is expressly written otherwise in the manuscripts. Many journals require co-authors to attest to the accuracy of, and their agreement with, the research results and outcomes.

Note: Cornell University handles disputes regarding authorship as an academic issue. Such disputes will be handled according to the procedures described in University Policy 1.2, Academic Misconduct. All persons listed as authors on publications or presentations should meet criteria accepted in their field for authorship credit, and

2. To Whom This Policy Applies

All units of the university

Who should read this policy:

All members of the Cornell community, including faculty, staff, and students, who are involved in the design, conduct, or reporting of research at, or under the auspices of, Cornell University.

3. Definitions

| Term | Definition |
|-----------------------------|---|
| Data retention statement | A brief statement of the principal investigator’s data retention practice describing how the PI complies with their obligations under this policy. |
| Final closeout | <p>For data generated under a sponsored project, <i>final closeout</i> means the award period has ended, and all deliverables have been submitted.</p> <p>For data acquired through a data use agreement, under a sponsored project, <i>final closeout</i> means the award period has ended, deliverables have been submitted and data has been destroyed or returned to the data provider according to the terms of the agreement.</p> <p>For non-sponsored projects, <i>final closeout</i> means all work has ceased, no further publications related to the project are anticipated, and any data acquired under a data use agreement has been returned to the data provider or destroyed according to the terms of the agreement.</p> |
| Principal investigator (PI) | <p>The individual who is responsible for the overall direction of the project. In the case of research projects led by students who are supported by the university or require significant university resources, the faculty advisor.</p> <p><i>Note:</i> Students who are not supported by the university and who are not using significant university resources in their research are the owners of their data and any intellectual property resulting from the research project.</p> |
| Research data | <p>The information needed to evaluate reported results of research.</p> <p>This policy applies the Federal definition broadly to data related to research, without regard to how the research is funded or how the data was acquired. As defined in 2 CFR 200.315 (e) (3): “Research data means the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This ‘recorded’ material excludes physical objects (e.g., laboratory samples). Research data also do not include: (i) Trade secrets, commercial information, materials necessary to be held confidential by a researcher until they are published, or similar information which is protected under law; and (ii) Personnel and medical information and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy, such as information that could be used to identify a particular person in a research study.”</p> |

| Term | Definition |
|-----------------------|---|
| | Nothing in this definition is intended to supersede an agreement with a human research subject to code or otherwise de-identify personal information or specimens they provide for research purposes. |
| Researcher | Any faculty or staff member, student, postdoctoral researcher, research associate or fellow, or other person involved in the design, conduct, or reporting of research. |
| University facilities | Facilities owned by the university or under contract by the university with a third party. |

4. Policy Administration

| Policy Clarification and Interpretation | Contact | Phone | Email/Web Address |
|---|---|----------------|--|
| Ithaca-based locations | Office of the Vice Provost for Research | (607) 255-7200 | vp_research@cornell.edu |

5. Appendix

5.1 Related Resources

| Topic | Office | Email/Web Address |
|--|--|--|
| Consultation about data storage and preservation systems, data classification, data management plans | Cornell Research Data Management Service Group (RDMSG) Cornell IT Security Office | finder.research.cornell.edu/storage University Policy 5.10, Information Security it.cornell.edu/regulated-data |
| Data management plans and data use agreements initiated within Cornell | RDMSG | data.research.cornell.edu/content/data-management-planning |
| Data use agreements initiated outside Cornell | Office of Sponsored Programs | osp_dua@cornell.edu |

| | | |
|---------------------------------------|--|---|
| General data management guidelines | HHS Office of Research Integrity NSF Institution and Award Support | ori.hhs.gov/data-management www.nsf.gov/bfa/dias/policy/dmp.jsp |
| Human participant data | Office of Research Integrity and Assurance | www.irb.cornell.edu/policy |
| Intellectual property rights | Center for Technology Licensing | University Policy 1.5, Inventions and Related Property Rights |
| Lab notebooks and lab archives | Cornell University Library | www.library.cornell.edu |
| Managing restricted access datasets | Cornell Restricted Access Data Center (CRADC) | ciser.cornell.edu/data/secure-data-services/cradc/? |
| Research Data Management Seminar | Cornell University Library | NTRES 6600 - guides.library.cornell.edu/c.php?g=32525&p=204729 |
| Rigor and reproducibility in research | Office of Research Integrity and Assurance | researchservices.cornell.edu/resources/rigor-and-reproducibility-planning |
| Standards on sharing data | The National Academies Press National Institutes of Health National Science Foundation | Sharing Publication-Related Data and Materials - www.nap.edu/catalog/10613/sharing-publication-related-data-and-materials-responsibilities-of-authorship-in Dissemination and Sharing of Research Results - www.nsf.gov/bfa/dias/policy/dmp.jsp NIH Data Sharing Policy and Implementation Guidance - grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm |

6. Revision History

| Date | Summary of Revisions |
|----------------|---|
| August 7, 2020 | <ul style="list-style-type: none"><li data-bbox="527 380 846 411">• Interim policy issued |