Responsible Conduct of Research, Scholarship and Creative Activities (RCRSCA)

Research Data, Data Management, and Peer Review

Greg M. Swain, Ph.D.
RCRSCA Education Coordinator – The Graduate School
Department of Chemistry
Topics of Discussion – Learning Objectives

- Understand core issues with types of data and data management practices that should be used in research and scholarship.

- Understand compliance issues relating to data ownership, data sharing and protection.

- Appreciate the importance of peer review and proper practices.

- Review of mentoring, misconduct and authorship.
Why is **Good** Mentorship (Coaching) Important?

- Good mentorship improves the quality and integrity of scientific research. *They show you the ropes!!*

- Good mentorship is essential for one’s professional development – preparation for a career.

- Different mentors for different areas.
What is a Mentor?

- The best mentors are advisors, coaches, counselors and supporters all at the same time.

- They are experienced faculty who guide your research/creative activity, but also challenge you to develop your independence.

- A good mentor will help you define your research goals, and then support you in your quest to achieve them. He or she will share knowledge, provide encouragement, and hopefully inspire you.
5 QUALITIES OF GOOD RESEARCH MENTORS

“A mentor is a person who has achieved career success and counsels and guides another for the purpose of helping him or her achieve like success.”

- **Respectful**: Demonstrates respect for all laboratory members, which reduces fear and unhealthy competitiveness.
- **Supportive**: Supports mentees by acknowledging accomplishments and challenging mentees to develop skills that advance their careers.
- **Available**: Establishes open and responsive communication with mentees, which promotes research integrity and discourages questionable research practices.
- **Prepared**: Anticipates the needs of mentees and is prepared to provide assistance and guidance.
- **Honest**: Sets high standards for honest reporting of data, regardless of whether the data supports the desired outcome.
The Mentor-Mentee Relationship

Mentee Responsibilities:

- Takes responsibility for establishing good lines of communication and actively participates in the relationship.
- Assesses academic/professional strengths, learning and developmental needs, values and short and long-term career goals (self evaluation).
- Develops a plan with the mentor for achieving these goals.
- Follows through on commitments and goals.
- Is receptive to feedback and coaching.
- Takes advantage of opportunities presented by the mentor.
Research Misconduct

- **Fabrication**: making up results and recording or reporting them.

- **Falsification**: manipulation of research materials, equipment, or processes, or changing or omitting results such that the research is not accurately represented in the record.

- **Plagiarism**: the appropriation of another’s ideas, processes, results, or words without giving proper credit.
Criteria for Research/Scholarship Misconduct

- Represents a significant departure from accepted practices;
- Has been committed intentionally, or knowingly or recklessly; and
- Can be proven by a preponderance of evidence.
Research/scholarship misconduct does not include honest error, misinterpretation of results or differences of opinion.
Ethical Standards to Guide Scholarship


- Honesty in reporting of scientific data and creative works.
- Careful recording and analysis of results to avoid error.
- Independent analysis and interpretation of results that is based on data – not external influences.
- Open sharing of methods, data, and interpretations through publication and presentation.
Ethical Standards to Guide Scholarship

- Sufficient validation of results through replication and collaboration with peers.
- Proper crediting of sources of information, data and ideas.
- Moral obligation to society in general and, in some disciplines, responsibility in weighing the rights of animals and human subjects.

MSU Authorship Guidelines

A person claiming authorship or being designated as an author of a creative endeavor should meet all of the following criteria:

- Participation in conception and design of the creative work, study, analysis or interpretation of any data.
- Participation in the drafting of the creative work or manuscript or in the editing of the creative work or manuscript.
- Final approval of the version of the creative work or manuscript to be published.
- Ability to explain and defend appropriate portions of the work or study in public or scholarly settings.

https://vprgs.msu.edu/announcement/msus-authorship-guidelines-updated
AUTHORSHIP PRACTICES TO AVOID CONFLICTS

Every field of study experiences conflicts with determining authorship on published papers

Implementing the following suggestions may help avoid potential authorship disputes:

- **BE PREPARED**
  Establish written authorship agreements with all members of the lab and other collaborators before preparing a manuscript or before starting a project.

- **DOCUMENT CONTRIBUTIONS**
  Authors should list their substantial contributions to the design of the study; the acquisition, analysis, or interpretation of data; and the contribution to the writing of the final paper.

- **BE CONSISTENT**
  Have clearly written expectations for authorship on publications and follow them.

- **COMMUNICATE OFTEN**
  As the project progresses, the authorship agreement may need to be revisited.

- **APPROVE THE MANUSCRIPT**
  All authors should review manuscripts and approve the final version.

**ACKNOWLEDGEMENTS**
Those who assisted\(^1\) with a manuscript but did not provide substantial contributions can be given acknowledgement.

\(^1\)This may include people who provide support such as: editorial assistance (e.g., proofreading), limited data collection, supervision of research tasks without contribution to the collection, analysis, or interpretation of data, or the writing of the publication, and technical support.
Types of Data

- **Institutional Data**
  Institutional Data are all data and records held by the University, in any form or medium, for the administration, operation, or governance of the University or any unit of the University. *(e.g., Research Data)*

- **Confidential Data**
  Institutional Data that could, by itself or in combination with other such data, be used for identity theft or related crimes. *(e.g., personal data)*.

- **Public Data**
  Public Data are Institutional Data that have become generally available to members of the public because a person with authority to do so has intentionally released or distributed them without restriction or limitation.
Responsible Use Requirements

A. Members of the University community may access and use Institutional Data only for University purposes.

1. Members of the University community may not use or disclose Institutional Data to obtain or provide others with a private benefit that is inconsistent with the University’s interests.

2. Members of the University community may alter, store, and distribute Institutional Data only for University purposes.

3. Each member of the University community may access Institutional Data only if, and then only to the extent that, he or she needs to do so for a University purpose.

https://tech.msu.edu/about/guidelines-policies/msu-institutional-data-policy/
Responsible Use Requirements

B. Institutional Data must be used, stored, transferred, disseminated, and disposed of in ways that minimize the potential for their improper disclosure or misuse.

1. Members of the University community must comply with all laws, University policies, and contracts that govern the use and release of Institutional Data, especially Confidential Data.

2. Records that contain Confidential Data and are no longer needed for University purposes must be disposed of promptly and properly. Best practices for record disposal are described in Appendix II.

https://tech.msu.edu/about/guidelines-policies/msu-institutional-data-policy/
Responsible Use Requirements

B. Institutional Data must be used, stored, transferred, disseminated, and disposed of in ways that minimize the potential for their improper disclosure or misuse.

3. Records that contain Confidential Data shall be properly secured to minimize the risk that the Confidential Data will be accessed, either intentionally or inadvertently, by individuals who do not need to see or use the Confidential Data for University purposes.

https://tech.msu.edu/about/guidelines-policies/msu-institutional-data-policy/
Responsible Use Requirements

C. Members of the University community are individually responsible for the security and integrity of Institutional Data in their possession or control, including their proper storage and disposal.

1. Members of the University community shall not knowingly create inaccurate or misleading Institutional Data, or deliberately alter or delete accurate Institutional Data to make those Institutional Data, or other Institutional Data, inaccurate or misleading.

2. Members of the University community may share Institutional Data only with individuals who need to access those Data for a University purpose.

https://tech.msu.edu/about/guidelines-policies/msu-institutional-data-policy/
Responsible Use Requirements

C. Members of the University community are individually responsible for the security and integrity of Institutional Data in their possession or control, including their proper storage and disposal.

3. Members of the University community are individually responsible for their own use, storage, dissemination, and disposal of the Institutional Data to which they have access.

4. Members of the University community who, for University purposes, make Institutional Data available to individuals who are not subject to this Policy should take appropriate action to provide for the proper use, storage, and disposal of those Institutional Data by those individuals, including, when necessary, contractual limitations on the further dissemination of the Institutional Data by those individuals.

https://tech.msu.edu/about/guidelines-policies/msu-institutional-data-policy/
"Research data, unlike other types of information, is collected, observed, or created, for purposes of analysis to produce original research results."

University of Edinburgh

"Research data is defined as recorded factual material commonly retained by and accepted in the scientific community as necessary to validate research findings; although the majority of such data is created in digital format, all research data is included irrespective of the format in which it is created."

Engineering and Physical Sciences Research Council (EPSRC)
The Data Cycle

Create

Process

Analyze

Data Life-Cycle Management

The diagram illustrates the life cycle of data management, including the stages of creating data, processing data, analyzing data, preserving data, and giving access to data. The processes are interconnected, emphasizing the cyclical nature of data management and the importance of data reuse throughout the cycle.

Research Data

Quiz: What Are Data?

True or False: In scientific research, only the information and observations that are made as part of scientific inquiry are considered data.

- True
- False

Answer: False. In fact, data also include the materials, products, procedures, and other data sources that are part of the research project. Essentially, data are considered to be anything and everything that informs the way in which individuals are able to understand and to process their world.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Collection

Data collection provides the information necessary to develop and justify research. A successful project collects **reliable and valid data!!**

- Enables those involved in the research to more accurately analyze and assess their work.
- Allows independent researchers to replicate the process and evaluate results.
- Provides justification to sponsors for expenditures and project decisions.
- Yields reliable and valid results, and hypothesis testing.

Data Management Plan

Obtaining quality research data starts with having a well conceived and organized data management plan!

1. **Research Study Design** – planning for all steps in the research cycle. Experimental plan is detailed, replicated and validated. Planning for all steps in the data cycle!
2. Core issues with data management that arise during the research process.
3. Methodological and legal regulatory considerations – HIPAA and FERPA.
4. Compliance issues relating to data ownership, sharing and protection.
Research Data Collection

Quiz: Data Collection

Data that are collected as part of a scientific research project ultimately prove or disprove the PI's hypotheses and justify a body of research to the public at large. Which statement is true about data collection in scientific research?

- Ensuring validity of the data is the key to successful research.
- Ensuring reliability of the data is the key to successful research.
- Ensuring reliability and validity are equally important.
- Data collection is actually not a key part of scientific research, since many researchers use previously collected data.

Answer: Ensuring reliability and validity are equally important. Ensuring reliability and validity of the data are equally important during data collection. When data collection is carried out according to these two rules, researchers will be able to accurately assess, replicate, and disseminate their results.

https://ori.hhs.gov/guidelines RESPONSIBLE DATA MANAGEMENT SCIENTIFIC RESEARCH
Design of the Research Study

- Student researcher together with PI and other collaborators **are responsible** for designing the study and executing the plan. **Proper control measurements are always a key consideration!!!**

- **Reliability** – replication/consistency of an outcome. Are the results reported reproducible??

- **Validity** - operational concepts and measurements. Refers to how well an instrument or technique measures what it is intended to measure. Refers to the accuracy of an instrument/measurement.
Research Data Collection Considerations

- Which data should be collected?
- By what means should data be collected to ensure reliability and validity?
- How many subjects or animals are required for adequate statistical power?
- Which collection methods will be used to reduce error or bias?
- Who will undertake each data related task?
- Who will supervise and how will quality and integrity of the research study be ensured?
Research Data Record Keeping

Diligent record keeping is essential to ensure the validity and reliability of data. Many research projects keep both written and electronic records to balance the benefits of each.

Written Record Keeping – Notebooks and Journals

- When data are actually collected, the records should attempt to accurately represent the progress of a project and answer such questions as what, how, and why data were collected or amended. Complete documentation!!!
- Records should be durable and accessible but safe from tampering or falsification. Keep in safe and secure place. Duplicate documentation is advisable.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Record Keeping

Electronic Records

Electronic records allow researchers to efficiently access and compare information from different sources and across similar projects. Security of electronic records is a significant concern, although there are methods for protecting electronic records.

One should be constantly aware of all the guidelines that might apply to the project's implementation and dissemination, including special regulations that involve human and animal subjects, hazardous materials, or other controlled biological agents.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Storage

Data storage is crucial to a research project for multiple reasons!

Type and Amount of Data to Retain

Generally speaking, enough data should be retained so that the findings of a project can be reconstructed with ease. **While this does not mean that a project needs to retain all the raw data that were collected, relevant statistics and analyses from this data should be saved, along with any notes or observations.** Furthermore, if research involves the use of biological specimens, care should be taken to retain them until their quality degrades.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Storage

- Storing data safeguards your research and your research investment.

- Storage allows future access to the data in order to re-create the findings, augment subsequent research, or establish a precedent.

- Enough data should be stored so that a project and its findings can be reconstructed with ease.
Case Vignette: Data Ownership

Dr. Smith works at The University and is the Principal Investigator on a large research project that is funded by the National Institutes of Health (NIH). However, while Dr. Smith wrote the original grant proposal, he does very little day-to-day work on the project. Instead, the Research Director, Betsy, oversees all aspects of the project, including staff supervision and all data management activities. In addition, Betsy has been lead author on several publications about the project's research findings.

Who owns the project and its data?

- The PI, Dr. Smith
- The Research Director, Betsy
- The University
- The National Institutes of Health
- No one person or organization

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Ownership

**Answer: The University.** Despite the PI's and the Research Director's work on the project, the sponsoring institution typically maintains ownership of a project's data as long as the PI submitted the grant through that institution and is employed by them.

However within the sponsoring institution, a PI is generally granted stewardship over the project data; he/she may control the course, publication, and copyright of any research, subject to institutional review.

Research Data Ownership

Ownership of research is a complex issue that involves the PI, the sponsoring institution, the funding agency, and any participating human subjects.

1. The Sponsoring Institution - Most often, the sponsoring institution/organization maintains ownership of a project's data as long as the PI is employed by that institution. The institution often controls all funding or the disbursement of government funding; consequently, it is also responsible for ensuring that funded research is conducted responsibly and ethically.

Within the sponsoring institution, a PI is granted stewardship over the project data; the PI may control the course, publication, and copyright of any research, subject to institutional review.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Ownership

2. The Funding Agency - Many research projects are funded by federal government agencies, philanthropic organizations, or private industries. These agencies often have specific stipulations for how data will be retained and disseminated: for example, they decide whether to publish the project's results or market a resulting product, rather than the PI.

3. The Principal Investigator – PI is a steward of a project's data, but may retain some ownership of the data. In businesses, it is assumed that rights and ownership of data remain with the business itself or with the funding agency, unless otherwise stipulated. In academic institutions, PIs are sometimes allowed to take their research and its data with them if they change institutions. Policies exist to ensure that such a transfer of data respects both the rights of the researcher and those of the institution(s).

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Protection

- In order to maintain the integrity of stored data, project data should be protected from physical damage as well as from tampering, loss, or theft. **This is best done by limiting access to it.**

- PIs should decide which project members are authorized to access and manage the stored data. Notebooks or questionnaires should be kept together in a safe, secure location away from public access, *e.g.*, a locked file cabinet.

- Regulatory considerations or proprietary information.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Protection

- **FERPA** – Family Educational Rights and Privacy Act - generally prohibits the improper disclosure of personally identifiable information derived from education records.

- **HIPAA** – Health Insurance Portability and Accountability Act - modernize the flow of healthcare information, stipulate how Personally Identifiable Information maintained by the healthcare and healthcare insurance industries should be protected from fraud and theft, and address limitations on healthcare insurance coverage.
Research Data Protection

- Data protection should be a part of every project's plan for data storage.

- The best way to protect data, whether in written or electronic form, is by limiting access to the data.

- Electronic data storage offers many benefits but requires additional consideration and safeguards.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Protection

Quiz: Data Protection

With the recent emergence of electronic databases, more scientific researchers are storing their data on their computer networks. However, data protection is an issue for both paper- and computer-based data. So what is the **best** way to protect data?

- Strip identifiers from human subjects data.
- Limit who has access to the data.
- Use an encrypted password system and assign new passwords quarterly.
- Destroy the written data after transferral to an electronic database.

**Answer: Limit who has access to the data.** This is the best way to protect data. Simple measures -- like keeping written data in a locked filing cabinet for which there is only one key -- will help minimize the chance that data could be corrupted or stolen. However, this is a complex issue and employing a multifaceted security approach is the best way to ensure that your data is protected.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Retention

How Long Should Data Be Kept?

- There is no set amount of time for which data should be stored. In some cases, the time period is at the discretion of the PIs; however, many sponsor institutions require that data be retained for a minimum number of years after the project ends.

- For instance, the USDHHS requires that project data be retained for **at least 3 years** after the funding period ends. Other sponsors or funders may require longer or shorter periods.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Retention

- Sponsor institutions and funding agencies often have their own requirements for how long data should be retained.

- Ultimately, the PI must decide when it is time to end data storage.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Destruction

- When the decision has been made to end data storage, data should be thoroughly and completely destroyed. Effective data destruction ensures that information cannot be extracted or reconstructed.

- Many document storage companies now offer onsite shredding and secure destruction of written and electronic records.

- For electronic data specifically, software products such as Eraser or CyberScrub are available.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Sharing and Reporting

• Data should be made as widely and freely available as possible while safeguarding the privacy of participants, and protecting confidential and proprietary data.

• Often data from a project are made available to the public for the first time at publication.

• Regardless of the mechanism used to share data, each dataset will require documentation. Proper documentation is needed to ensure that others can use the dataset and to prevent misuse, misinterpretation, and confusion.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Sharing and Reporting

As part of the scientific process, data are expected to be shared and reported. This serves several purposes, including the following:

- Acknowledging a study's implications
- Contributing to a field of study
- Stimulating new ideas

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Sharing and Reporting

- By sharing research results, a project may advance new techniques and theories and benefit other research. It encourages collaboration between researchers in the same field or across disciplines. Additionally, reporting of clinical research data can have a direct impact on the quality of health care provided to patients.

- **Data sharing usually occurs once a study has been completed.** Data reporting includes discussion of the data, the data analysis, and the authorship of a project. Data sharing and reporting are typically accomplished by publishing results in a scientific journal or establishing a patent on a product.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Research Data Sharing and Reporting

Sharing Data Prior to Publication

Before publication, there is often no obligation to share any preliminary data that have been collected. In fact, sharing at this stage is sometimes discouraged because of the following reasons:

- The implications for a set of data may not be understood while a project is still in progress. By waiting until a project is ready for publication, researchers ensure that what they share has been carefully reviewed and considered.

- There is fear that less scrupulous researchers will use shared research results for their own gain. This apprehension causes some researchers to refrain from disseminating their findings.

https://ori.hhs.gov/guidelines-responsible-data-management-scientific-research
Responsible Conduct of Research

- The Principal Investigator (PI) has overall responsibility for the conduct of the research or scholarly activity. The PI's responsibilities include, but are not limited to: (a) obtaining approval from regulatory committees such as Human Research Protection Program, Institutional Animal Care and Use Committee (IACUC), and Office of Environmental Health & Safety (EHS); (b) determining authorship and acknowledgement credit for presentations and publications; and (c) complying with MSU policies related to data ownership, control, and access.
In addition, the PI should ensure that all members of the research team have been trained to fulfill their responsibilities to the project, follow the approved plan for conducting the research, and comply with relevant laboratory and research policies. MSU requires that faculty members sign off as PIs for this purpose, but student and faculty investigators share the responsibility for complying with policies related to research ethics.
Data Ownership, Control and Access

Institution of Record

MSU shall be the institution of record for any publication, presentation, or other public dissemination of results for research conducted at MSU or using MSU resources. No other institutions, including future institutions where a student or faculty member works at the time of presentation/publication, shall be included in a presentation/publication unless the other institution contributed significantly to the project. *This policy does not preclude listing the student's or faculty member's current address for correspondence.*

http://www.com.msu.edu/Research/Research%20Integrity.htm
Data Ownership, Control and Access

Ownership of Research Data

- Research data include but are not limited to raw data recorded in any format, computer records and printouts, videotapes, audiotapes, photographs, laboratory notebooks, and field notes and journals.

- MSU is the owner of the scientific records for projects conducted at the University, under the auspices of the University, or with University resources. Therefore, the original copies of all such data must be maintained at MSU, including data from student-initiated research projects.

http://www.com.msu.edu/Research/Research%20Integrity.htm
Data Ownership, Control and Access

Management of Research Data

- The PI is the custodian of the data, with responsibility for ensuring the accuracy of the scientific record, the confidentiality of the data and the physical condition and security of the data. *In addition, the PI is responsible for retaining the research data for not less than three years after the submission of the final project report or publication (whichever occurs last).*

- Furthermore, the PI is responsible for protecting intellectual property resulting from the research and responding to allegations of misconduct in research or financial conflicts of interest (both situations may warrant retaining research data for long than three years).

http://www.com.msu.edu/Research/Research%20Integrity.htm
Data Ownership, Control and Access

Management of Research Data

- For student-initiated projects, research data must be retained at least until the degree is awarded or until it is clear that the student has abandoned the work.

http://www.com.msu.edu/Research/Research%20Integrity.htm
Data Ownership, Control and Access

Access to Research Data

- The PI is responsible for determining who may have access to research data and under which conditions. When students or faculty members involved in research projects at MSU leave the University, they may take copies of research data for projects on which they have worked.

http://www.com.msu.edu/Research/Research%20Integrity.htm
Peer Review Process

What is peer review?

Peer review is designed to assess the validity, quality and often the originality of articles for publication. Its ultimate purpose is to maintain the integrity of science by filtering out invalid or poor quality articles.

Did you know?

84% of researchers believe that without peer review there would be no control in scientific communication.

9/10 researchers feel that peer review improves the quality of their published paper.

(Source: University of Tennessee and CIBER Research Ltd, 2013.)
Peer Review Process

- Peer review is the evaluation of work by one or more people with similar competences as the producers of the work (peers).
- It functions as a form of self-regulation by qualified members of a profession within the relevant field.

The peer review process:

- Scientists study something.
- Scientists write about their results.
- Journal editor receives an article and sends it out for peer review.
- Peer reviewers read the article and provide feedback to the editor.
- Editor may send reviewer comments to the scientists who may then revise and resubmit the article for further review. If an article does not maintain sufficiently high scientific standards, it may be rejected at this point.
- If an article finally meets editorial and peer standards it is published in a journal.
Peer Review Process

Peer review is essential for the integrity of science and scholarship!!

https://authorservices.wiley.com/Reviewers/journal-reviewers/what-is-peer-review/index.html
Peer Review Process

As a reviewer, what would you be asked to comment on?

- Scientific importance of the work – *does it represent new science and advance the field?*
- Technical quality (*data, data analysis, data interpretation, methods, etc.*)
- Clarity - *is the paper well written with good language and grammar, and is the organization of the paper such that it can be understood by the reader?*
- Rigor of the statistical analysis and assessment of reproducibility.
Peer Review Process

Top tips for peer reviewers (Wiley):

Summary of Key Points

- Proper data/information acquisition (reproducibility and validity) are critical for quality science and scholarship.

- Starts with a well-organized research plan!

- Data ownership is with the university. PI is granted stewardship.

- Data management (record keeping, storage, ownership, protection, retention and sharing) are important to consider at the outset of a project. Must comply with all university and funding organization policies and requirements.

- Peer review is a responsibility and is essential for the integrity of science and scholarship.