Data Acquisition and Management

Stephen Kania, PhD
Professor of Biomedical and Diagnostic Sciences
College of Veterinary Medicine
2 Why is Data Management Important?
3 Failure to Properly Manage Data

- Lose your valuable data
- Become sanctioned by your funding agency
- Cost the university huge fines
- Forfeit your intellectual property
- Lose out on patents
4 Benefit of Properly Managing Data

- If you follow best practices for data management your life will be easy and you will be incredibly successful
- Easy access to data
- Impress funding agencies
- Enhance your reputation with colleagues
5 Data Management Topics

- Data Management plan
- Types of Data
- Ownership
- Collection/Recording
- Security/ Protection
- Retention
- Sharing
- Access
6 What is/are Data?

– OMB defines research data as “the recorded factual material commonly accepted in the scientific community as necessary to validate research findings...”
Many Granting Agencies Require Data Management and Sharing Plans

- Require data management and/or sharing plan
  - DOE
  - EPA
  - NASA
  - NIH
  - NOAA
  - NSF
  - USDA
8 Elements of a Typical Data Management Plan

- How data will be collected
- Who will collect the data
- Who will own the data
- How accuracy will be ensured
- How long data will be stored
- Where/how data will be stored
- What will happen if the PI leaves the institution
There are Two General Types of Data:

- Quantitative Data
- Qualitative Data
10 Quantitative Data

- Often stored on computers
  - Have three copies of all files
11 Quantitative Data

- Often stored on computers
- Have three copies of all files
12 Quantitative Data

- Often stored on computers
  - Have three copies of all files
  - Develop a file naming policy that facilitates retrieval
13 File Naming Best Practices

- Be descriptive
- Be consistent
- Avoid special characters
- Avoid upper case
- Use dates in file names
- Use version numbers
# File Naming

<table>
<thead>
<tr>
<th>Files without employing an naming convention:</th>
<th>Files with a naming convention:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Test_data_2013</td>
<td>- 20130503_DOEProject_DesignDocument_Smith_v2-01.docx</td>
</tr>
<tr>
<td>- Project_Data</td>
<td>- 20130709_DOEProject_MasterData_Jones_v1-00.xlsx</td>
</tr>
<tr>
<td>- Design for project.doc</td>
<td>- 20130825_DOEProject_Ex1Test1_Data_Gonzalez_v3-03.xlsx</td>
</tr>
<tr>
<td>- Lab_work_Eric</td>
<td>- 20130825_DOEProject_Ex1Test1_Documentation_Gonzalez_v3-03.xlsx</td>
</tr>
<tr>
<td>- Second_test</td>
<td>- 20131002_DOEProject_Ex1Test2_Data_Gonzalez_v1-01.xlsx</td>
</tr>
<tr>
<td>- Meeting Notes Oct 23</td>
<td></td>
</tr>
</tbody>
</table>

Scott Brandt, Purdue University
Qualitative data is anything that is not numerical. It can include words, pictures, sound recordings, video, etc.

- Physical Specimens
- Tissue Samples
Data Ownership
Who’s rules are we playing by?
17 Federal Regulations-Data Ownership

- NIH Grants Policy Statement -- “grantees own the data generated by or resulting from a grant supported project.”
- NSF -- “normally allows grantees to retain principal legal rights to intellectual property developed under NSF grants”
- Federal guidance typically supports institutional ownership of federal research
  - Federal agencies maintain rights to use of data
Ownership of non-federal research is subject to negotiation between the institution and the sponsor.

Industry often requires that the sponsor retain ownership as a condition of the award.

The ‘contract’ award mechanism is a common instrument for negotiating ownership.

Read your contract!
Data Accuracy

- Reliable data is dependent on reliable methods and appropriate controls

- Establish quality control procedures and document them!

http://www.vivianpartnership.co.uk
Data Collection/Recording

- Electronic data must be validated to prove it was recorded on the date stated

http://www.pronetic.co.uk
21 Data Collection/Recording

- Hard copy data should be collected into a numbered, bound notebook with all changes annotated.
22 Data Recording

- Data should be recorded contemporaneously to the extent possible
23 Why Secure and Retain Data?

- To confirm research findings
- To establish priority/ inventorship
  - Patents are now first to file
  - Must be the actual inventor
- To be reanalyzed by other researchers
- As a safeguard in case there is any question of misconduct
24 Data Security/ Integrity

- Researcher who collects or uses information is primarily responsible for data protection
- Back up computer files (Three copies)
  - Password protect
- Record data in indexed laboratory notebooks
  - Keep notebooks in a secure location
- Keep copies of original notebooks
- Store samples to prevent degradation


25 Data Needing Extra Protection

- Federal Classified Research
- Information regarding select biological agents
- Protected Health Information ex Health Insurance Portability and Accountability Act (HIPPA)
- Student information ex Family Educational Rights and Privacy Act (FERPA)
How Long is Long Enough?

- HIPAA records
  - 6 years
- Student assisted research
  - until degree is awarded or research is abandoned
- Patent applications
  - 20 year patent term plus extension
- FDA regulated research
  - 2 years after approval or after investigation is discontinued and FDA is notified
USDA Retention Requirement

(a) This provision sets forth requirements for record retention and access to records. As used in this provision, “records” includes books, documents, accounting procedures and practice, and other data, regardless of the type or format.

(b) Financial records, supporting documents, statistical records, and all other records pertinent to an award shall be retained for a period of 3 years from the date of submission of the final expenditure report or, for awards that are renewed quarterly or annually, from the date of the submission of the quarterly or annual financial report, as authorized by FAS. The only exceptions are the following:

1. If any litigation, claim, or audit is started before the expiration of the 3-year period, the records shall be retained until all litigation, claims or audit findings involving the records have been resolved and final action taken;

2. Records for tangible property acquired with Federal funds shall be retained for 3 years after final disposition;

3. When records are transferred to or maintained by FAS, the 3-year retention requirement is not applicable to the Recipient;

4. Indirect cost rate proposals, cost allocations plans, etc., as specified in paragraph (f) of this provision.

(c) Copies of original records may be substituted for the original records if authorized by FAS.

(d) FAS will request transfer of certain records to its custody from Recipients when it determines that the records possess long-term retention value. However, in order to avoid duplicate record keeping, FAS may make arrangements for Recipients to retain any additional records that are continuously needed for joint use.

(e) FAS, the Inspector General, Comptroller General of the United States, or any of their duly authorized representatives, have the right of timely and unrestricted access to any books, documents, papers, or other records of Recipients that are pertinent to the awards, in order to make audits, examinations, excerpts, transcripts and copies of such documents. This right also includes timely and reasonable access to a Recipient’s personnel for the purpose of interview and discussion related to such documents. The Recipient shall provide access to any program site(s) to FAS or any of its authorized representatives. The rights of access in this paragraph are not limited to the required retention period, but shall last as long as records are retained.

(f) No Recipient shall disclose its records that are pertinent to an award until the Recipient provides notice of the intended disclosure with copies of the relevant records to FAS.

(g) Indirect cost rate proposals, cost allocations plans, etc. Paragraphs (g)(1) and (g)(2) of this provision apply to the following types of documents, and their supporting records: indirect cost rate computations or proposals, cost allocation plans, and any similar accounting computations of the rate at which a particular group of costs is chargeable (such as computer usage charge back rates or composite fringe benefit rates).

1. If submitted for negotiation. If the Recipient submits to FAS or the sub Recipients submits to the Recipient the proposal, plan, or other computation to form the basis for negotiation of the rate, then the 3-year retention period for its supporting records starts on the date of such submission.
28 UT Retention Requirement

- UT Policy FI0120 – Records Management
- Intellectual Property Documentation, Including All Documents Relating to the Application for and Granting of All University Copyrights, Patents, and Trademarks must be retained permanently!
Data Sharing
Federal Agency

- NIH- Data should be made as widely and freely available as possible while safeguarding privacy of participants, and protecting confidential and proprietary data
- Investigators submitting NIH research application above $500k in direct costs annually are required to include plan for sharing final research data
- UT- Promotes prompt open exchange of research data
30 Data Sharing Considerations

- Preliminary data should not ordinarily be released
- Data can usually be kept confidential prior to publication
- Data is expected to be freely available after publication
31 Data Access
Freedom of Information

– Federally funded research data must be made available in response to Freedom of Information Act (FOIA) requests (OMB Circular A-110).
– “Sunshine Laws” in some states also necessitate sharing
– Institutions have some flexibility regarding the timing and content of what is released
There is a “tool” for Creating Data Management Plans

- [www.DMPTool.org](http://www.DMPTool.org)
- University of Tennessee has membership
- Log in with UT user name and password
- Creates agency specific plans, example USDA- NIFA
33 Summary

– Before you start a research project know:
  – Who will own the data
  – How it will be collected and stored
  – How long it needs to be retained
  – How the security and integrity of the data will be maintained
  – The requirements and process for sharing it
34 Resources

- UT Office of Research Integrity
  - Dr. Robert Nobles Associate Vice Chancellor for Research, Responsible Conduct of Research & Research Integrity Officer
- UTIA Office of Sponsored Programs
  - 865-974-7357 aggrant@utk.edu
- Chris Eaker, Data Curation Librarian at the University of Tennessee
  - www.DMPTool.org