IRB Issues and Archival Data: From Data Deposit to Data Preservation (Avoiding Loss)

Overview

I. Deposit Agreement: How it Functions

II. Ownership and Copyright

III. Practical Examples of Lost Data

IV. IRB and informed consent: Avoiding Loss
When to think about Data Archiving: Depositor Rights and Responsibilities
What is the Data Deposit Form?

• It is an ICPSR form given to data depositors for them to fill out to accompany a data submission

• We require a signature indicating assent with ICPSR’s terms of deposit.
What is it used for?

• Grants ICPSR permission to archive and disseminate the data (copyright is retained by the data creator)

• It is also used to create the study-level meta-data (i.e. the catalog information describing a study and its methodology)
ICPSR recently changed it…

We wanted to make more transparent to a data depositor the role that ICPSR takes in archiving and distributing the data.
Deposit Agreement Language

- I have implicit or explicit copyright to this work and have the right to make it publicly available through ICPSR.
Deposit Agreement Language

I give my permission for the Data Collection to be used by ICPSR for the following purposes, without limitation:

• To redisseminate copies of the Data Collection in a variety of media formats
• To promote and advertise the Data Collection in any publicity (in any form) for ICPSR
• To describe, catalog, validate and document the Data Collection
• To store, translate, copy or re-format the Data Collection in any way to ensure its future preservation and accessibility
• To incorporate metadata or documentation in the Data Collection into public access catalogues
Deposit Agreement Language

I give my permission to ICPSR to enhance, transform and/or rearrange to the Data Collection, including the data and metadata, for any of the following purposes:

- protect respondent confidentiality
- improve use
Deposit Agreement Language

To the extent allowable by law or permitted by the sponsor of the data collection, in preparing this data collection for public archiving and distribution, I have removed all information directly identifying the research subjects in these data, and I have used due diligence in preventing information in the collection from being used to disclose the identity of research subjects.
Deposit Agreement Language

I further agree to release and hold harmless ICPSR (including staff and the ICPSR Council) and the University of Michigan from any and all liability from claims arising out of any legal action concerning identification of research subjects, breaches of confidentiality, or invasions of privacy by or on behalf of said subjects.
Ownership of Research Data

PI vs. Institutional Ownership

Do PIs own the data?

Do Universities own the data?
Research Data & Copyright Issues

PI vs. Institutional Ownership

• "The products of research conducted by a faculty member or researcher in the course of employment and developed with institutional support...reside with the institutional employer." - Estelle Fishbein (General Counsel, JHU)

• "Research data generated under PHS funding generally is owned by the grantee institution, not the principal investigator or the researcher producing the data." - Office of Research Integrity
Example University Policies

University of Michigan copyright policy regarding research data indicates that ownership of research is governed by the federal sponsor (or private foundation).

Harvard has made the originating laboratory the custodian of research data. In that case, faculty who leave the university are allowed only to take a copy of their research data.
Data Accessibility

This issue is complicated by conflicts between an institution's policies on rights to copyrightable material and its responsibilities to assure access to data generated under federal awards.
NSF Data Sharing Policy

National Science Foundation Important Notice 106 (April 17, 1989) states: "[NSF] expects investigators to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections, and other supporting materials created or gathered in the course of the research. It also encourages awardees to share software and inventions or otherwise act to make such items or products derived from them widely useful and usable."
NIH Data Sharing Policy

The NIH expects and supports the timely release and sharing of final research data from NIH-supported studies for use by other researchers. Starting with the October 1, 2003 receipt date, investigators submitting an NIH application seeking $500,000 or more in direct costs in any single year are expected to include a plan for data sharing or state why data sharing is not possible.
What are expectations about retention of Research Data?

- NSF Circular [A-110] provides that financial records, supporting documents, statistical records, and all other records pertinent to an award shall be retained for a period of three years.

- American Psychological Association expects its members to retain data for a minimum of five years.
## LEADS Database at ICPSR

<table>
<thead>
<tr>
<th></th>
<th># Records Reviewed</th>
<th># Social Science Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recent NSF (1976+)</td>
<td>17,194</td>
<td>2,537</td>
</tr>
<tr>
<td>Historic NSF (Pre-1976)</td>
<td>96,403</td>
<td>4,019</td>
</tr>
<tr>
<td>NIH (1972+)</td>
<td>172,196</td>
<td>6,381</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>285,793</strong></td>
<td><strong>12,937</strong></td>
</tr>
</tbody>
</table>

### NSF & NIH Funded Data Collections: Where are they today?

- **Data Archived**: 6.3%
- **Has Copy of Data**: 65.8%
- **Data Lost**: 18.5%
Discarded Research Data

• One of the most common reasons that NSF and NIH PIs give for having discarded research data are interpretations of various policies…

Lost
Reasons Data are Intentionally Discarded

“The material…was considered sensitive data. Institutional review boards.. required us to promise to destroy the data after a certain period of time…”

“Data were erased and the floppy disks were destroyed 5 years after the papers were published (per the IRB approval agreement).
Reasons Data are "lost"

“… extensive precautions induced by IRB/Human Subjects procedures, led to careful lock-ups and eventual destruction to preserve anonymity of all those interviewed.”

“It was just too long ago, I generally keep data for something like 10 years beyond the last time I do something with them.”

“Data cards were destroyed long after several articles were published from them.”

“Destroyed, in accord with APA 5-year post-publication rule.”
IRB and Informed Consent

IRB approval (or not)

- Most, but not all, research data are collected with the approval of an Institutional Review Board
Informed Consent and Data Sharing

• Informed consent is the formal authorization by an individual of their agreement to participate in the proposed research.

• The human subjects involved in a project must participate willingly, having been adequately informed about the research.

• In preparing the informed consent document, investigators are asked to include a statement describing the extent to which confidentiality of records identifying the subject will be maintained. This has implications for one’s ability to share data with the research community.
Informed Consent and Data Sharing
Example: University of Michigan Language

“You will not be identified in any reports on this study. Records will be kept confidential to the extent provided by federal, state, and local law. However, the Institutional Review Board, the sponsor of the study (i.e. NIH, FDA, etc.), or university and government officials responsible for monitoring this study may inspect these records”

(www.irb.research.umich.edu/IRB_HSBS_Shared/consent.html).
Example: University of Florida Template

Information collected about you will be stored in locked filing cabinets or in computers with security passwords. Only certain people have the legal right to review these research records, and they will protect the secrecy (confidentiality) of these records as much as the law allows. These people include the researchers for this study, certain University of Florida officials, the hospital or clinic [delete references to hospital and/or clinic if not applicable – if applicable, specifically name the hospital and or clinic(s)] involved in this research, and the Institutional Review Board (IRB; an IRB is a group of people who are responsible for looking after the rights and welfare of people taking part in research). Otherwise your research records will not be released without your permission unless required by law or a court order.
Elements of Informed Consent Language: Data Sharing Issues

- Protect confidentiality of respondents and their information
- Avoid stating the data will be kept locked/not shared
- Data will be used by qualified researchers
- Avoid stating data will be used by PI only/research team
Every effort will be made to protect your statements and answers, so that no one will be able to connect them with you. Records will be kept confidential to the extent provided by federal, state, and local law. However, the Institutional Review Board, the sponsor of the study (i.e. NIH, FDA, etc.), or university and government officials responsible for monitoring this study may inspect these records. Any references to your identity that would compromise your privacy will be removed or disguised prior to preparation and production of publications and data files resulting from this study.
Informed Consent Language
Model II (DRAFT)

The information that we collect will be used only in ways that will not reveal your identity. You will not be identified in any publication or data released from this study. Records will be kept confidential to the extent provided by federal, state, and local law. However, the Institutional Review Board, the sponsor of the study (i.e. NIH, FDA, etc.), or university and government officials responsible for monitoring this study may inspect these records.
Thank you!

To contact me…
Amy Pienta
apienta@umich.edu
734-615-7957