Bridging the Divide

Photo by Mason Jones. https://unsplash.com/photos/INvs-DWLoEo
Hidden collections brought to light

“... seeking to improve discovery, access and long-term preservation of maternal, infant, child health research data through the creation of rich metadata.”

“... offering a descriptive approach that can help educate the researchers creating metadata for their studies early in the records lifecycle and inform the practices of archivists and data curators.”

Photo by Seabass Creatives
https://unsplash.com/photos/yIW287G3_U0
Project Partners

University of Alberta Libraries
Sharon Farnel, Kendall Roark (Purdue), Amanda Harrigan, Saurabh Vashishtha

Center for the History of Medicine (CHoM), Harvard
Emily R. Novak Gustainis, Amber LaFountain

WCHRI/MICYRN/CRIC Collaborators
Lawrence Richer, Rick Watts
Collections

Centre for the History of Medicine

- historical research records (HRR) in which datasets are embedded
- Harvard School of Public Health Longitudinal Studies of Child Health and Development records (1930-1979)
- Early intervention studies from the multisite Infant Health and Development Program in the papers of health services researcher Marie McCormick (1985-2014)
- Anne Becker’s social transition and risk for disordered eating in Fiji study (2004-2010)

University of Alberta Libraries

- contemporary, with entirely born digital research data (BDRD)
- selected 36 studies to be included in the pilot
  - 28 clinical trials, 2 cohort studies, 6 others
- variety of statuses:
  - some complete, ongoing, terminated, withdrawn trials.
Overview

- Special Collections Research Data Description
  - Harvard School of Public Health Growth Study Records
  - Special Collections Processing Workflow
  - Anticipated Metadata Workflows
- Contemporary Research Data Description
  - Created REDCap Project
  - Populated Dataverse REDCap Form
  - Contacted Researchers
    - Researcher Documents
    - Researcher Meetings
- Next Steps

Family Physical Characteristics Key, created during the Harvard School of Public Health Longitudinal Studies of Child Health and Development. From the Harvard Medical Library in the Francis A. Countway Library of Medicine.
Special Collections
Research Data Description
Special Collections Processing Workflow

- Rehousing
- Preservation
- Folder listing
- Identifying sensitive information
- Logging audio & audiovisual recordings
- Logging and imaging electronic records
- Rearrangement
- Finding aid and catalog record creation
Harvard School of Public Health Growth Study Records

- Founded by Harold Coe Stuart (center), with members of the HSPH Department of Maternal and Child Health.
- Multiple Studies:
  - Original (c. 1930-1957)
  - 30-year follow-up (c. 1960-1969)
  - 40-year follow-up (c. 1970-1979)
  - 50-year follow-up (c. 1970-1979)

- ~200 cubic feet of research data, administrative records, and publications
- Raw data (measurements, surveys, medical test results, physical examination records, interviews)
- Summarized data
- Assessment data (subject rankings and scores based on interpretation of raw data)
- Coded data
- Analyzed data (charts, graphs, calculations)
- Imaging records (photographs, negatives, x-rays)
- Regulatory records (protocols, methodologies, codebooks)
- Subject recruitment records
- Grant funding records
- Reports
Anticipated Metadata Workflows

- Anticipating multidisciplinary use
- Research Data Lifecycle
- Mapping existing archives and records management descriptive standards and metadata schemas to Dataverse
- Metadata gap analysis
- HMS Data Management Working Group

(UK Data Service’s Data Lifecycle Model)
https://www.ukdataservice.ac.uk/manage-data/lifecycle
Contemporary Research Data Description
Created REDCap Project
Populated Dataverse REDCap Form

Clinicaltrials.gov

MICYRN Birth Cohort Inventory

NCBI
Contacted Researchers

No Response or Declines

*Basic* study record in REDCap

Notify researcher via email, offer contacts if they want to follow up

Documentation

Use documents to complete record metadata in REDCap

Follow up on missing information

Rich study record with research documents in Dataverse, if approved

Meeting

Arrange meeting, go through form, update study record in REDCap

Discuss metadata

Follow up on missing information

Rich study record with research documents in Dataverse, if approved
Researcher Documents

Data Dictionary, Codebook and/or existing metadata

Protocol or Proposal

Research Instruments (surveys, questionnaires, scales, etc)

Researcher agreements, consent forms, ethics applications...

Research Data
Researcher Meetings

- Validate and add to the study record
- Researcher feedback on metadata
  - What do they need and want?
  - What information would help them find or understand data?
  - When they need data where do they go? Where would they look?
  - Are they using any standards already?
  - Will Dataverse work for these studies?
  - What’s missing? What’s irrelevant?
Next Steps

- Enhance DDI-based metadata
- Analyse and incorporate feedback
- Schema gap analysis
- Best practices
Any Questions?

You can find us at

Amanda Harrigan: aharriga@ualberta.ca
Sharon Farnel: sharon.farnel@ualberta.ca
Amber LaFountain: amber_lafountain@hms.harvard.edu
Emily Gustainis: emily_gustainis@hms.harvard.edu