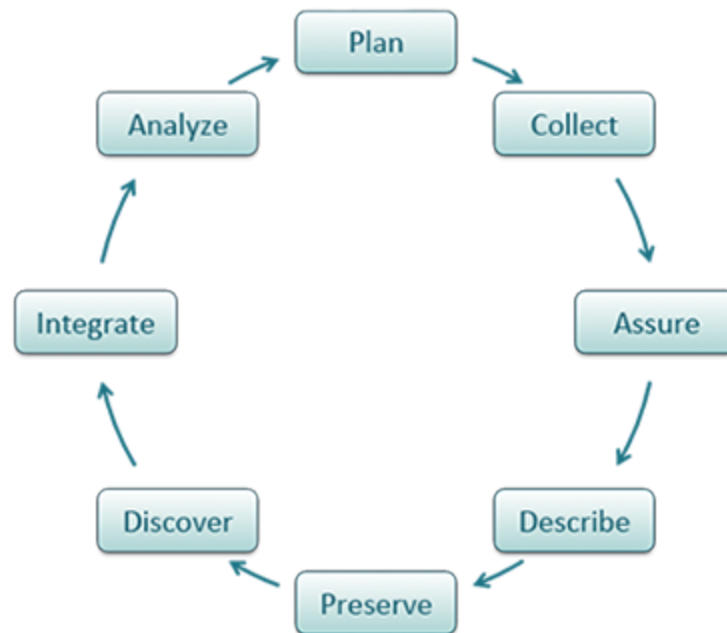


Collect



<https://old.dataone.org/data-life-cycle>

Ellie
Jason Ferrante
Mark Parsons
Sophie Hou

- 1) Why is this part of the Data Lifecycle Important to graduate students?
 - a) This is where the graduate students are making research happen.
 - b) This is the heart of the lifecycle from the graduate students' perspective.
 - c) This can be the most expensive or resource limited part of the project, especially if field work or sampling is involved, or if seasonality comes into play. May include permits, IACUC, legal access to a site, etc.
- 2) How can this be integrated into Graduate School? How to enhance/improve/optimize this experience?
 - a) Don't want to regret the formatting
 - b) How to find/organize it?
 - c) Don't have to do over
 - d) "Good handshakes and handoffs" from "Plan" and to "Assure" (and on to "Analyze")
 - i) "Assure" includes three phases: before collection (error prevention), during collection (error detection), and post collection (assessment)
 - ii) Assure at the moment you collect or you'll forget what you did.

- iii) Including “Assure” activities can help mitigating anxiety about missing anything from the “collect” stage (there might not be a chance to “do-over”).
- iv) Planning well can also help with making sure the “collect” stage goes as smoothly as possible.
- e) Students currently must take a methods class but might not have a data collection and management class. Specific disciplines may have more specific collection training.
 - i) The training process can be more formalized/consisten instead of “organic”.
- f) Planning is an essential part of collecting. Need a little stop sign between plan and collect to double check +1

3) What are the resources graduate students can use for this part of the Data Lifecycle?

- Tidy Data by Hadley Wickham 10.18637/jss.v059.i10
- “Tidy datasets are easy to manipulate, model and visualize, and have a specific structure: each variable is a column, each observation is a row, and each type of observational unit is a table.”
- Some Simple Guidelines for Effective Data Management by Elizabeth T. Borer et al. 10.1890/0012-9623-90.2.205
- Ten Simple Rules for the Care and Feeding of Scientific Data by Alyssa Goodman et al. 10.1371/journal.pcbi.1003542
- Support Your Data: A Research Data Management Guide for Researchers by John A Borghi et al. 10.3897/rio.4.e26439
- Data Management Training Clearinghouse <http://dmtclearinghouse.esipfed.org>
a favorite: “How to avoid a spreadsheet mess - Lessons learned from an ecologist”
- The Software Sustainability Institute <https://www.software.ac.uk>

Link to Assure:

https://docs.google.com/document/d/14lfr_UJukxhSINrtM1WQZAsOcXWFq2j_RtHfVu_yKa4/edit?usp=sharing