Agile data curation takes the principles of agile software development and maps them into data curation and management. The underlying principles to Agile Data Curation are a reapplication of agile software development principles to data management. A core principle of agile software development and data curation is the incremental creation of value through iterative development accompanied by frequent release. This is similar to the MPLP (More Product Less Process) idea [1] that advocates for minimal processing in order to reduce backlogs and increase access to collections.

After a brief introduction to what is meant by agile curation and MPLP, we will solicit exemplars of agile curation (though it might not have been conceptualized as such at the time) in practice from participants. This will include how users/agencies have handled finding and using existing data, in-project data management, strategies for developing data documentation, and transitioning data products and documentation into systems that enable preservation, discovery and reuse. In particular, we are seeking exemplars of how agencies/users have adopted a “get it done” attitude rather than “get it perfect” mentality.

Participants could be data curators, data managers, or data users, who are interested in sharing their experience.

Then we plan to map the exemplar practices to the foundational principles and through comparison learn lessons for future application and begin the process of translating principles into practices that are aligned with the best exemplars out there.


Agenda (11:00 - 12:30 Friday, July 22)
11:00 - 11:15 Conceptual Overview (Benedict, Lenhardt, Young)
This overview will provide the current state of the mapping of the concepts from agile software development into the data curation space with the goal of explicitly addressing a need to define the underlying conceptual foundation for data curation that embraces the values of flexibility, openness, responsiveness to user's (and reuser's) needs, and efficiency in the development of data curation workflows and strategies.

11:15 - 12:00 Case studies - examples and discussion of a strategy and technical approach for collecting case studies to inform the development of agile data curation design patterns (all attendees - moderated by Benedict & Hills)
This part of the session will focus on identifying some key examples of intentional or unintentional agile data curation case studies in an effort to have some concrete examples that we can work from in developing our strategy for collecting key elements of those case studies as input to identifying
common characteristics - ultimately feeding into a set of agile data curation design patterns that are reusable by new research data curation projects.

Presentations will include (will update as more are confirmed):

- Safeguarding our National Treasures - Nimbus Tape Recovery Challenges, by Ed Esfandiari (GEDS DISC)
- Unintentionally Agile - How a geological survey stumbled into agile data management, by Denise Hills (Geological Survey of Alabama)

12:00 - 12:30 Discussion and feedback on draft Values and Principles as derived from the agile software curation principles.


This will be a discussion that focuses on iteration on the draft language that has been developed for what the underlying values and principles of agile data curation should be. This will be the first time we have expanded to discussion of these values and principles outside of our research team and marks the beginning of an extended community dialog about them.

Attachments/Presentations: NotesAgileDataCurationintheWildWhatareyourstories.pdf [8]

Session Leads:

Name: Denise Hills [9]
Organization(s): Geological Survey of Alabama [10]
Email: dhills@gsa.state.al.us [11]

Name: Karl Benedict [12]
Organization(s): University Libraries, University of New Mexico [13]

Name: Nancy Ritchey [15]
Organization(s): NOAA/NCDC [16]
Email: nancy.ritchey@noaa.gov [17]

Name: Christopher Lenhardt [18]
Organization(s): RENCI [19]
Email: clenhardt@renci.org [20]

Presenters:

Name: Josh Young [21]
Organization(s): UCAR Unidata [22]
Email: jwyoung@ucar.edu [23]

Name: Karl Benedict [12]
Organization(s): University Libraries, University of New Mexico [13]

Name: Christopher Lenhardt [18]
Organization(s): RENCI [19]
Email: clenhardt@renci.org [20]

Name: Denise Hills [9]
Organization(s): Geological Survey of Alabama [10]
Email: dhills@gsa.state.al.us [11]