Recovering and Reusing Archival Data for Science: Investigating Curatorial Practices Across Disciplines

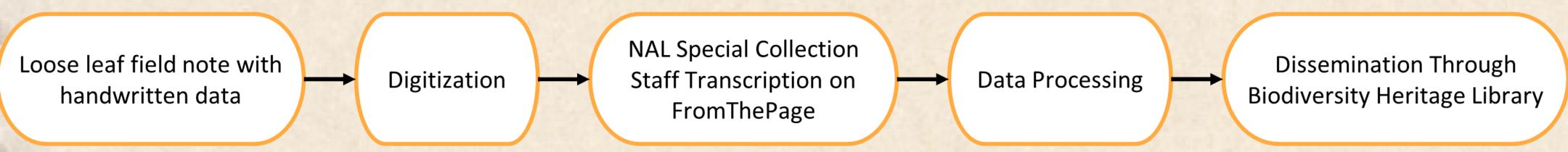
Hilary Szu Yin Shiue, Amanda Sorensen, Cooper Clarke, Katrina Fenlon (PI), University of Maryland, College Park Presented at the 18th Research Data Alliance Plenary Poster Exhibition (Virtual), 3-11 November, 2021

Research Questions:

- What opportunities and challenges confront the recovery and reuse of historical or defunct data for active and ongoing scientific research?
- What are the value and requirements of data reuse?
- What relationships are there among data recovery and reuse efforts across disciplines and organizational contexts?

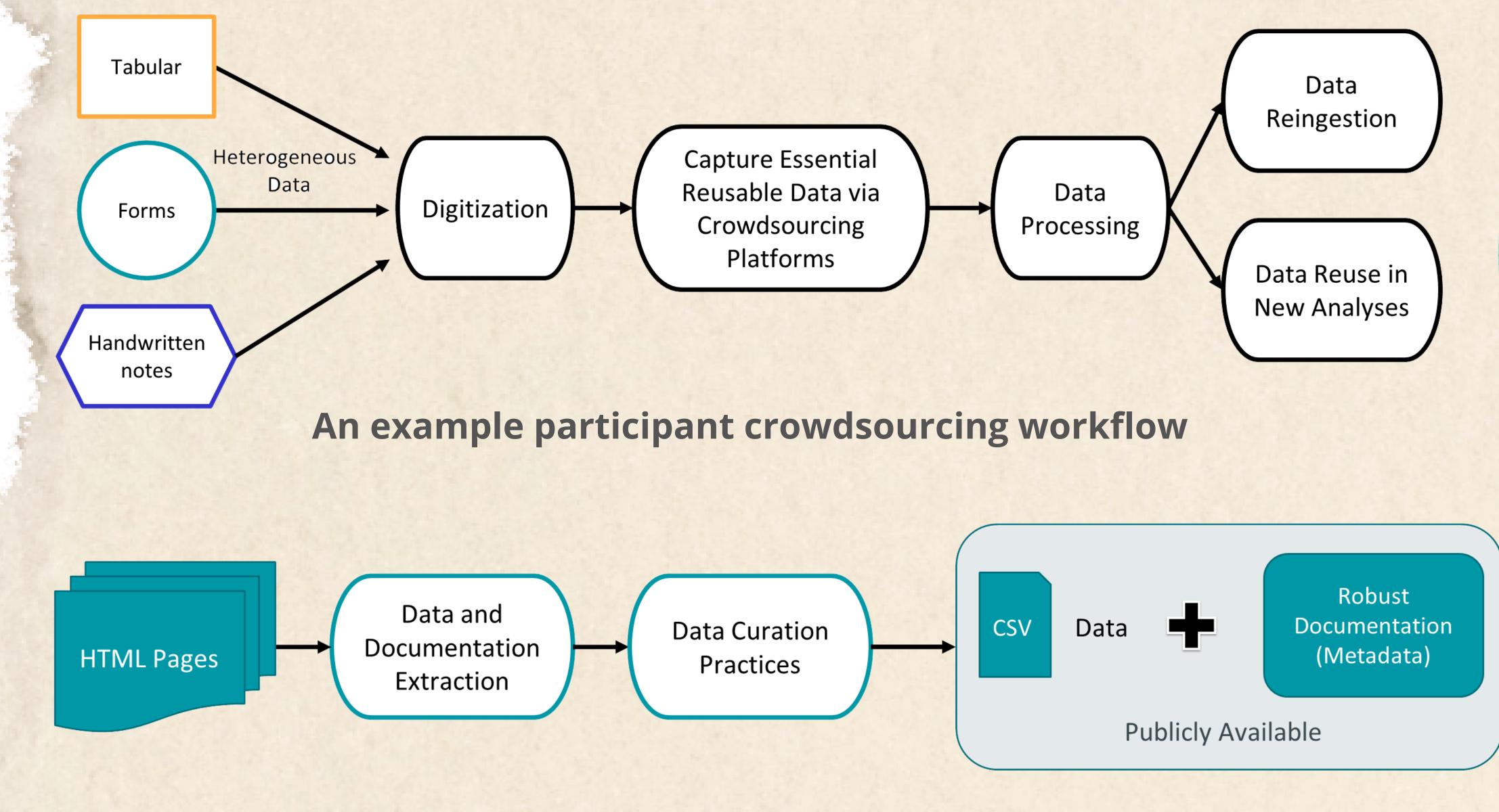
Methods:

- Semi-structured interviews with scientific reesearchers and data curation professionals across disciplines (ongoing, n=8 to date)
- Qualitative coding of interview transcripts to identify themes



National Agricultural Library case study workflow

Early Findings:

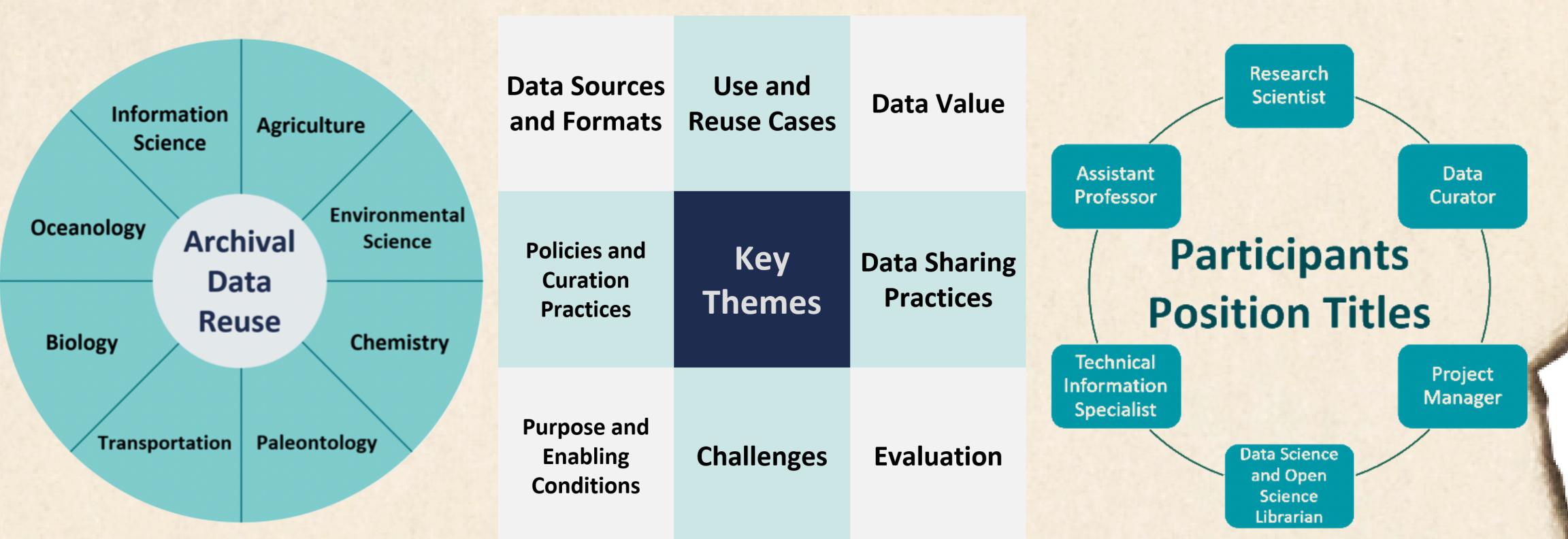


An example participant data recovery workflow

Research Process Observed the broad landscape Overview: of data rescue, recovery and reuse across domains Final Report and Recommendation Modern research data nterviews with Practitioners needs for data-intensive 18 Assessment 3 hybrid NAL Involved in Data Rescue, science & computer-**Factors** collections amenable collections **Recovery and Reuse Efforts** High-level tiered OAIS-compliant processing guide Iterative process

Context:

- Building on prior case studies of processing data-rich agricultural research collections
- Complex data types and formats: instrumental weather data, scientific photographs, log books, species survey records, grey literature, statistics data in PDFs, HTML files, and so on



Next Steps:

- Complete data collection and analysis (Summer 2021-Spring 2022)
- Produce pragmatic guidance for:
 - Data curators: What professional curators can learn from scientists' practices, toward promoting use and reuse of collections.
 - Domain Experts: Guidance for scientists doing data recovery and reuse, toward promoting the preservation of research data.





Supported by the University of Maryland College of Information Studies and by Non-Assistance Cooperative Agreement #58-8201-1-181 between the University of Maryland and the United States Department of Agriculture (USDA), Agricultural Research Service (ARS), National Agricultural Library with funding provided by the USDA, ARS, Office of National Programs.