

Repurposing Archival Theory in the Practice of Data Curation

Elizabeth Rolando | Wendy Hagenmaier | Susan Wells Parham



Introduction

- Expansion of data curation and digital archiving services at the Georgia Tech Library and Archives.
- How do data curation and archival science intersect?
- How can comparing data curation and archival science lead to improvements in local workflows and practices?

Methodology

- Process the same digital collection, once by data curator, once by digital archivist
- Data curation processing informed by OAIS Reference Model¹, ICPSR workflow², and UK Data Archive workflow³
- Archival processing informed by concepts, such as appraisal, *respect des fonds*, original order, and archival value⁴, as well documented practices at peer institutions
- Compare processing plans to discover areas of agreement and areas of conflict

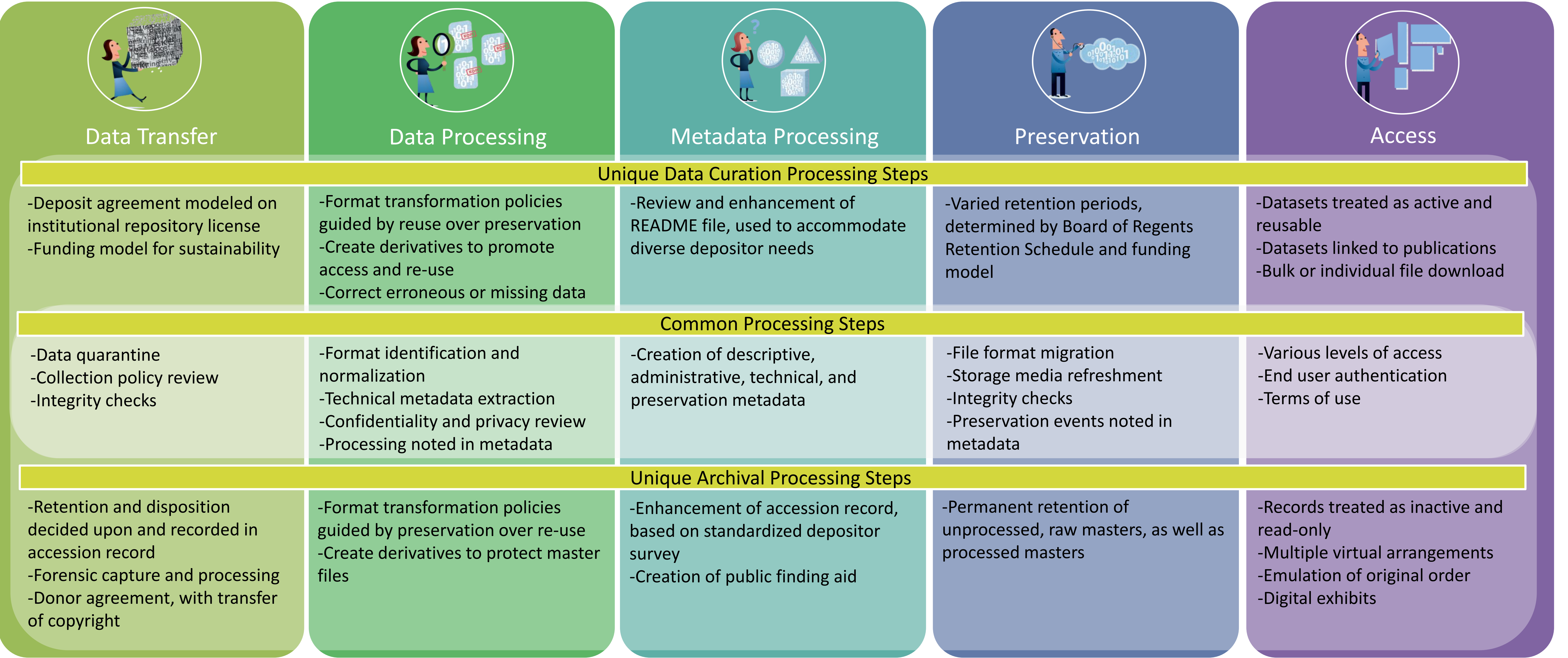


Figure 1: Highlights from comparison between archival and data curation processing plans. The first row lists those elements of the data curation processing plan that were unique, while the bottom row lists those steps in the archival processing plan that were distinct. The middle row identifies those elements of the processing plans that were common between the two. Images used in the diagram were created by Jørgen Stamp (www.digitalbevaring.dk) and are published under a Creative Commons Attribution 2.5 Denmark license.

Findings

What data curation might learn from archival science and processing:

- Forensic capture and processing may be valuable for certain data sets
- Existing repository license agreement models might not work for digital data sets
- Retention and disposition should be planned at the point of data transfer
- Creating virtual arrangements that emulate the data creator’s original environment could be valuable
- Data curators might question how much should be done to correct data in order to facilitate re-use--how much effort is enough?

What archival science and processing might learn from data curation:

- Establish a balance between supporting future access and use and maintaining the integrity of the record--do disk images support future access?
- Existing donor agreement and copyright transfer models might not work for digital archives acquisitions
- Funding model should be planned at the point of record transfer
- Processed records may not be “inactive”; the life of the record continues through re-use, which enriches the record and should be documented in the record itself

References

¹ Consultative Committee for Space Data Systems. (2012). *Reference model for an Open Archival Information System (OAIS)* (Magenta Book CCSDS 650.0-B-1). Retrieved from <http://public.ccsds.org/publications/archive/650x0m2.pdf>.
² Inter-university Consortium for Political and Social Research. (n.d.). *A Case Study in Repository Management*. Retrieved from <http://www.icpsr.umich.edu/icpsrweb/content/datamanagement/lifecycle/index.html>.
³ UK Data Archive. (2014). *How We Curate Data*. Retrieved from <http://www.data-archive.ac.uk/curate>.
⁴ Society of American Archivists, *Glossary of Archival and Records Terminology*: <http://www2.archivists.org/glossary>