Abstract

As part of the Mellon-funded AIMS project, the Universities of Virginia, Hull, Stanford and Yale have spent the last two years exploring the ramifications and distinct requirements of born digital archival materials in libraries. This presentation focuses on the partners' research and prototyping of tools, infrastructure and workflows necessary to provide an end-to-end environment for born digital archival materials.

In particular, it will detail the practical experiences gained in processing nine born digital collections from four different institutions, and on Hypatia, a Hydra-based application designed to give digital archivists a platform for arranging, describing, and providing access to born digital archives.

Hypatia: Research & Early Developments on a Platform for Managing Born Digital Archival Materials

As part of the Mellon-funded AIMS project (1), the Universities of Virginia and Hull, Stanford and Yale have spent the last two years focused on exploring the ramifications and distinct requirements of born digital archival materials in research libraries and archives. As manuscripts and records, drawings and data sets are increasingly produced, stored and offered to archives in digital form, memory institutions must develop the processes, agreements, infrastructure and workflows tooled to accommodating digital files.

This presentation will focus on the partners' research and prototyping of the tools, infrastructure and workflows necessary to provide an end-to-end environment for born digital archival materials. This includes forensically extracting files from their source media; enabling arrangement and description of the materials in a digital environment; and providing scholarly discovery and access to individual files.

The presentation will focus predominantly on the cooperative development done by the partners on Hypatia (Hydra Platform for Access To Information In Archives) (2), a Hydra-based Ruby-on-Rails application with a Fedora Repository back-end designed to provide digital archivists with a platform for managing, preserving and providing access to born digital archival materials. While Hypatia is still in the early stages of its development, through the process of specifying and developing it to accommodate nine born digital archives from four institutions, the partners have derived first hand, practical experience on some key issues surrounding born digital materials, including:

- the unique needs of digital archival materials relative to digital library objects
- the gaps and overlaps with EAD for managing these materials
- strategies to leverage forensic tools and processing, pre-ingest workflows and post-ingest / repository-based processing
- the necessity to support both detailed, item-level description and extremely lightweight processing, supporting the growing trend of MPLP, or "More Product, Less Process" in the archival world (3)
- data models for collection, series, and item objects that support both bulk and granular description and access levels, as well as preservation-ready objects
- requirements and initial designs for more robust arrangement and description tools, produced by the project's digital archivists as a byproduct of using existing tools on digital collections

In closing, the presentation will highlight both the successes and stumbling blocks encountered in the course of Hypatia's development, summarize the veins that seem most promising for further development, and survey the opportunities for continued collaboration on an open source solution in this space.

1 http://www2.lib.virginia.edu/aims/
2 https://wiki.duraspace.org/display/HYPAT/Home
3 http://ahc.uwyo.edu/documents/faculty/greene/papers/Greene-Meissner.pdf