

## Avoiding the Calf-Path: Digital Preservation Readiness for Growing Collections and Distributed Preservation Networks

» Martin Halbert, Woodruff Library, Emory University

» Gail McMillan, Digital Library and Archives, Virginia Tech

ETD repositories often start with very idiosyncratic and ad-hoc beginning data storage structures, driven by exigencies associated with creating an effective electronic workflow for accepting and securely storing digital copies of theses and dissertations as either a replacement or supplement to parallel workflows for print copies. ETD repositories also tend to grow in an effectively unbounded manner over time. Much like the story of the wobbling Calf-Path in the poem by Sam Walter Foss, these early idiosyncrasies and unbounded growth can subsequently cause enormous problems in systematic efforts to digitally preserve content of growing collections. The most effective preservation strategies incorporate pre-coordinated replication of content in distributed and secure locations; such replication strategies become increasingly difficult when the content is stored using irregular practices in directory structures, metadata, and file naming conventions.

This paper will address “Calf-Path” problems by providing practical guidelines, suggestions, and recommendations for ETD repositories. These recommendations are informed by five years of experience in operating the MetaArchive Cooperative, a distributed digital preservation cooperative of cultural memory organizations which has grappled with standardizing transfer mechanisms and developed cost/effective strategies for distributed preservation of ETDs based on the LOCKSS open source software. In the course of the past six years the members of the MetaArchive Cooperative have identified a series of best practices for digital preservation readiness.

These best practices can benefit start-up programs which have not yet established regular procedures and standards for directory structures, metadata, and file naming conventions. This paper will document relatively simple principles and guidelines for such programs that can greatly improve the subsequent likelihood of implementing successful distributed digital preservation programs.

## Framing Digital Assessts into Context: A Preservation Study in the Design of a Cryogenic Pressure Vessel Using “STEP Documentation”

» Samrat Sarovar, Graduate Student, Mechanical Engineering, West Virginia University

» Victor Mucino, Professor, Mechanical and Aerospace Engineering, West Virginia University

The process of development of a product involves innumerable activities and transactions, spread across a wide array of fields, departments, people etc. Right from the requirements of the customer, conceptualization of the design to manufacturing, there is an exchange of information punctuating all the activities and transactions. One of the main problems the industry faces is the documentation of this information so that it can be used as and when required.

In the long run, most of the information is either not retrievable, or is worthless due to proprietary issues, unavailability of programs used to create it, and other common situations. Even in the scenario that all the information is available and re-usable, there is always the missing link, which might be due to the absence of a particular designer, or in most cases, the statement of purpose, and the context of the design.

This study addresses the issue of Long Term Retention of data, the documentation procedures used during the design process,

and the development of context for an engineering problem. The context of an engineering design is the crux of the design process, as, it is the context that defines the problem, the different sets of solutions to that particular problem, and all the decisions taken in the life cycle of the product.

Also the context unifies each decision and each detail documented in the design. Thus, Loss of context renders all the information about the product useless. The design of a Cryogenic Pressure Vessel is used as a case study to understand the process of design, the flow of information, and the role of context in the design of a product.

A cryogenic pressure vessel that is designed to be used as a case study is designed to store and supply liquid nitrogen for a superconductor application. This is complex engineering problem, as the process of design and manufacture is a plethora of activities and transactions between people from different walks of life.

Using this design process as a reference to create Engineering Scenarios, we identify the flow of information through the various activities and transactions involved in the development of the product. These Engineering Scenarios also give us an idea on how the information is documented at each stage.

In the broader aspect, this project will help us identify a methodology for archiving information about engineering techniques and experiments, and be fundamental in providing guidelines to better preservation of complex engineering data. This will facilitate an environment conducive to easier and faster research on any given topic relevant to the end user.

This study is also part of wider digital preservation and archiving efforts lead by the U.S. National Archives and Records Administration (NARA). The common mission is to ensure the long-term retention and usefulness of digital data as well as the complex relationships and contextual metadata among digital assets. West Virginia University (WVU) is among a growing number of partner institutions working with the National Archives to provide comprehensive and integrated archival standards, methodologies, systems and solutions to guarantee the preservation of the digital information that becomes part of the historic national record.

Further, the WVU Libraries have offered the use of the institutional repository, “wvuScholar”, as a collaborative test bed environment for future usability studies in metadata subject and contextual analysis as well as for preservation and archiving using the application of the proposed methodologies of this study.

## BREAKOUT SESSION 2B

### COLLABORATIONS: INTER/INTRA DEPARTMENTAL COOPERATION

» Moderator: Suzie Allard, University of Tennessee

» Thursday, June 11 11:15 a.m.–12:45 p.m.

## Moving from a Decentralized to Centralized Model for Processing ETDs

» Angela McCutcheon, Director of Thesis and Dissertation Services, Graduate College, Ohio University

Ohio University in Athens, Ohio, used a decentralized graduate college model for 25 years. Moving towards a centralized ETD processing model required coordinating with the library, graduate council, several graduate studies offices, and the new Graduate College. Ohio University spent two years centralizing its operations.