Potential Strategies for Preserving Complex DH Projects

Corinne Guimont | Digital Scholarship Coordinator
Alex Kinnaman | Digital Preservation Coordinator
Best Practices Exchange 2020
Introduction

1. What is a DH project? What levels might we engage with? What project components can be identified?
2. What are early preservation-centric decisions and planning strategies?
3. What are potential pieces of documentation and planning measures?
4. VT Case study
Types of Projects
Small Projects

● Out of the box CMS (Omeka, WordPress, or Scalar)
  ○ Little development
  ○ Online exhibits

● Student support provided

● 1-2 semesters
Redlining Virginia

Congress created the Home Owners' Loan Corporation (HOLC) in 1933 to rescue the real estate industry and restructure the home mortgage market. In the process, HOLC also strengthened residential racial segregation. HOLC promoted racial covenants on deeds to prevent mixed-race neighborhoods and created color-coded maps of American cities that made race a key element in assessing financial risk. This exhibit explores the actions of HOLC in Virginia and its impact on Virginia cities.

It was originally an exhibit from 12/7/2016 to 2/17/2017 in Newman Library at Virginia Tech.

Featured Exhibit

Redlining in Virginia
Medium Projects

- Built on Content Management Systems or basic HTML
  - A little more customization / development
- More hands on work
- 1-2 years
The Southern Industrial Educational Association, Inc. 1905-1926

The Southern Industrial Educational Association, Inc. 1905-1926 was one of the first philanthropic organizations aimed at improving life in rural Appalachia. It is an organization that faded into obscurity, never receiving recognition for the work it accomplished during the first decade of the twentieth century. Kathleen Curtis Wilson, a Fellow at the Virginia Foundation for the Humanities, spent a decade uncovering a comprehensive history of the organization and its most influential members. An overall history of the organization is included in Wilson's publication, Uplifting the South, The Legacy of Mary Mildred Sullivan, Overmountain Press, 2008.

From 1909-1926, the association published a small booklet distributed to members and potential supporters. The Quarterly Magazine contained a wide variety of Appalachian-related information on subjects as diverse as health care, agriculture, speech patterns, book reviews, craft traditions, and night school education.

In 2014, under the guidance of the Institute for Advanced Technology in the Humanities at University of Virginia, Wilson digitized 51 of the 53 issues of the magazine. She provided the background information and formatted a framework for Virginia Tech to develop a comprehensive website that enables users to take full advantage of the fascinating new information found within the archive: www.kathleencurtiswilson.com.

Recurring themes:
1. Moral turpitude in cities;
2. Manual training can improve life for men, women, and children;
Large Projects

- Very customized
- Typically grant funded
- Multi-year projects
- Large amounts of data
LORD BYRON AND HIS TIMES

General Indexes

LBT is a growing digital archive of books, pamphlets, and periodical essays illustrating the causes and controversies that preoccupied Byron and his contemporaries. The documents, large and small, ephemeral and monumental, underscore the social dimensions of publishing in the romantic era; the archive uses notes, commentary, and links to highlight relationships among their readers and writers.

Byron was at once a social critic and an object of criticism, a brooding misanthrope and an icon of fashion. His life and writings were the subject of scores of books and pamphlets that sparked yet further controversies that often developed a life of their own. These quarrels over manners and morality, politics and religion were occasions for rhetorical exchanges that were sometimes brilliant, sometimes mean, and sometimes both.

If Byron was the focus of controversy, the focus of LBT is less the man than the controversies that enveloped him. The romantic era was, contemporary observers complained, an age of "personality." LBT pursues connections between people and publications using database technology to track the appearances of thousands of major and minor personages through a wealth of documents.

To access the archive, select CONTENTS in the menu to bring up a list of document-collections in the archive, or SEARCH PERSONS to find documents by or about a particular person. To read about the project at large, select ABOUT LBT.
Identifying Project Components
Common Components

- Grant Proposals
- Research Data
- Sustainability Plan
- Media
- Articles
- Documentation
Components: What are they?

- **Research Data:**
  - CSV files
  - XML
  - Text

- **Media:**
  - Images
  - Video
  - Audio

- **Documentation:**
  - Text
  - PDF

- **Articles:**
  - Text
  - PDF
  - HTML

- **Sustainability Plan:**
  - Text

- **Grant Proposals:**
  - Text
  - PDF

- **Overall Project Website:**
  - HTML/CSS
  - Other code
  - Look and feel of the site
Preservation
DH projects are their own ecosystem

Require their own mini preservation system
Preservation Challenges: Broad

- Technical support & funding
- Various types of content to manage
- File format migration
- Ongoing project backups and preservation
- Older projects need to be sunsettled
- Hosting, management, and general ownership
Preservation Challenges: Project Types

Small
- Lack of documentation, metadata
- Less ongoing support
- Possibly meant to be temporary
- Time limitations

Medium
- Limited documentation, metadata
- Less ongoing support in the long-term
- Time limitations

Large
- Larger amounts of content
- Higher content type diversity
- Multiple versions
- Older technologies
- Too much documentation to unravel
- Grant limitations
- Copyright limitations
- Ownership: Faculty members vs Library
Early Preservation Strategies

- Preservation planning & documentation
  - DMP
  - Preservation Profile
  - Succession Plan (in progress)

- Component-focused preservation
  - Evaluation of each component of a project for preservation challenges
  - Ongoing, regular preservation of projects segments as they are completed

- Web archiving
Planned Preservation Strategies

- Hosting projects at the Library
- Test exports from common DH tools like Omeka to develop standard workflows
- Develop DH-specific Levels of Preservation
  - Appraisal
  - Balance resources
  - Template workflow
  - Metadata and website-building requirements
Documentation & Planning
Documentation & Planning

How do you plan documentation for new projects?

What do you do when a project comes in with no documentation?

When do you start thinking about preservation planning?
Documentation & Planning

Project development
● Planning storage and backups
● Data management planning early
● Succession planning
● Decisions and general project documentation

Ongoing project
● Preservation planning and goals for each project segment
● Continue revising your DMP, succession plan
● Engage with any technical support you have
Why is this important?

- Understand your content
- Communication with stakeholders/project managers on roles, time commitments
- Grant requirements
- Review past decisions and maintain continuity
- Ability to accurately reconstruct or repair a project later
Case Study
Social Networks in Georgian Britain

- Built from previous project, Lord Byron and his Times (lordbyron.org)
- 30,000 individual records collected over the past couple of decades
- Ultimate Goal: Build an interactive social network that will allow users explore the data in both its raw XML format as well as through visualizations in a user interface
- Short term Goal: Publish the raw XML files in a public Git Repository
Social Networks in Georgian Britain

1. XML data
   a. Schemas
   b. ReadMe files

2. Grant proposal
   a. Data Management Plan

3. Documentation
   a. Preservation/Sustainability Plan
   b. Succession Plan

4. Visualizations

5. Final Website
   a. Archived Site
Social Networks in Georgian Britain

- Documentation and early planning has lead to:
  - Developing a standard schema that has revealed errors within data
  - Eased the grant writing / fundraising process
  - Ease decision making and conversations with faculty
  - Choosing a platform that allows for public access to data and documentation
Questions?

Corinnes Guimont gcorinne@vt.edu
Alex Kinnaman alexk93@vt.edu