



Mark Phillips

Digital Projects Department

University of North Texas

Annexation of Texas Project



Overview

"From Republic to State: Debates and Documents
Relating to the Annexation of Texas, 1836-1856"

- 6000 Digital Objects to be created
- PDF Files made from the scans
- Add digital representations to Catalog
- Preserve Master Files



Project Timeline

Date	Activity
Sept. 2002	Contract agreement and signing.
Oct. 2002	Hire and train student assistants. Begin scanning of titles and begin metadata entry
Nov. 2002 – March 2003	Continue Scanning
April – May 2003	Load files to Web server. Begin indexing by search engine.
June – Au 2003	Final overview and adjustments. Publicity and evaluation.
Sept – 2003	Write and submit the final report



Formats

- Fragile Books
- Large Bound Books
- Maps
- Certificates
- Newspaper clippings
- Pamphlets
- Handwritten Letters



Technology

- Minolta PS7000
- Microtek 6400XL
- Canon D30 digital camera
- Vidar Latitude large format scanner
- Photoshop 7.0
- OmniPage 12 Pro
- JEdit



Updated Project Outline

- Scan images to be saved as preservation masters
- Optical Character Recognition (OCR) on all text images
- XML markup of OCR'ed Text
 - TEI-Lite DTD
 - XSLT
 - Create HTML files using XSL Stylesheets



Funding

- Texas State Library and Archives Commission Grant
 - Tex-Treasures Grant – \$20,000
Covering student wages of 2 SLIS students from the Digital Imaging Management program.

- University Matching Funds
 - Technology
 - Training of Students
 - Supervision of project



Collaboration

- The Texana and Rare Books Collection
 - Rare books, pamphlets, maps, and official documents.
- UNT University Archives
 - Unique letters, and papers of Mary Jones, wife of Anson Jones.
- The Federal Documents Collection
 - U.S Congressional debates, correspondence between U.S., Mexico, and Texas officials



Training

- Technical Operation of Scanners
- Scanning standards for project
- File naming conventions
- OCR software training
- XML Mark-up training



Evaluation

The TexTreasure Grant requires statistical targets to be met as a part of the evaluation.

Formal evaluation is being performed in conjunction with the Portal to Texas History Project by a team from the Texas Center for Digital Knowledge. The project team is headed by Dr. Bill Moen.



Lessons Learned

- Workflow
- File Naming Conventions
- Quality Control
- XML Mark-up workflow
- OCR issues