Welcome to the Survey of Digitization class for Fall of 2012! Your instructors are Tanya Clement and Quinn Stewart with the gracious consultation of Zach Vowell, Digital Archivist for the Dolph Briscoe Center for American History and the assistance of Parker Fishel, our fearless TA.

We meet on Tuesdays, noon to 3pm, in UTA 1.210A.

This class is officially an introduction to the issues and trends in digitization initiatives and management, including project planning and management, asset delivery and management systems, interoperability and the importance of standards, copyright and other legal issues, metadata basics digital preservation, and specific digitization processes for documents, images, video, and sound. Three lecture hours a week for one semester. Unofficially, it's a fun class with lots of hands-on projects with real collections and archives and some theory (yes theory!) about the objects (text, images, sound, and moving images) we'll be digitizing.

Some questions we'll consider:

- Why digitize? What is the cultural significance of the collections or objects we'll be digitizing? And how will digitizing make them more accessible or preserve them?
- What are effective work flows that repurpose existing descriptive material, rather than creating new metadata?
- What are the most efficient work flows that keep costs reasonable but also follow prescribed standards for sustainability and preservation?
- How do we begin to create well-designed plans that evaluate the use of the digitized materials and the effectiveness of the methods employed in digitizing and displaying the materials?

Some collections with which we'll work are the Oral History of the Texas Oil Industry and the New York Journal America Newspapers collections from the Dolph Briscoe Center for American History, and poetry readings from the PennSound Poetry Archive.

Grades will be broken down as follows:

- Participation: 10%
- Digitization Portfolio: 55%
- Discussion leaders and synthesizers: 10%
- Final Grant Project: 20%
- Topic Presentations: 10%
- Final Presentations: 5%

We will use the following schedule in calculating final grades:

A = 95-100
A- = 90-94
B+ = 84-89
B = 79-83
B- = 74-78
C+ = 69-73
C = 60-68
F = <60
SOD Class 1- Text Digitization

Week 1 (9.4)- Text digitization You will find readings, tutorials, etc. in this module.

This class will introduce you to the instructors and the digitization classroom, and we'll go over the course syllabus and other "first day of class" matters.

Then, we get busy! There's a very good chance that in your journey to get to class today, for better or for worse, you owe a debt to a number of people who were gathered in Spindletop Texas in 1901. In 1951, these folks had a reunion, at which happened to be a historian from the University of Texas. This person realized that first person accounts of the cattle industry, and the railroad industry in Texas had been lost to time, but first person accounts of the oil industry in Texas could still be collected. Using new-fangled magnetic recording technology brought back from Nazi Germany after WWII, UT historians created the "Oral History of the Texas Oil Industry" 218 recorded interviews of people present at the birth of the modern oil industry. In 1980, ten students from our School created an index of 179 of these transcripts, and our current project to make this entire audio collection searchable and accessible online builds upon their work.

Now, it's your turn! We are down to the final 2000 transcript pages, and our hope is to scan these during this class period, and complete the OCR in subsequent classes and outside of class. Our goal for this class will be to complete the scanning during class, based upon this tutorial. Please review it before the first class, and also make sure you have signed up for an iSchool Account!

Discussion: SOD Class 1 - Introduction EXAMPLE

Sign up for Reading Presentations and Discussion Leaders and Synthesizers : Collaborations
SOD Class 2- Manuscript Digitization

Week 2 (9.11) Why Digitize? What’s digital preservation?

Discussion:

SOD Class 2

SOD Class 2 part B

Digitization Lab: Manuscript Pages Naming Convention

Tutorial: This is a continuation of the previous tutorial, but this link should take you to the new section: Scanning Manuscript Collation Tables and Pages


Hoccleve Archive: Scanning Phase

The Hoccleve Archive is a collection of resources meant to advance the study of the fifteenth-century London poet, Thomas Hoccleve. The largest part of the Archive is a collection of handwritten variation tables that contain transcriptions of the manuscripts of Hoccleve’s longest poem entitled "The Regiment of Princes," collated by lines of the text. These tables represent over a decade’s worth of labor by a half-dozen scholars. This group attempted to produce the first edition of Hoccleve’s poem to account for the variation among all 43 of its extant manuscripts, but they were unable to bring a full version of this edition to press. The Hoccleve Archive is attempting to resurrect this monumental editorial project in the digital realm. The first step of this project is to scan the hard copy collation tables into images suitable for both long-term preservation and web-display. The latter images will be part of crowd sourcing and pedagogical applications. For this class, you will be scanning the couple-hundred collation tables in the Archive that were handwritten on standard 8.5x11” paper. You will also be scanning and using OCR on a collection of dot-matrix printouts that served the original team of editors as an index to the earliest manuscript of the poem. If the OCR is successful, we hope to use this index as both a way to navigate the collection of scans and as the basis for our new hypertext critical edition of the poem.
SOD Class 3 - Copyright and Access: Carlos Ovalle

Week 3 (9.18) Copyright and Access : Carlos Ovalle, speaker

Guest Lecture: ALA Copyright Scholar Carlos Ovalle

Nyja Newspaper Clippings Instructions

Lab: NYJA newspaper clippings scanning  

Or, try the beta mobile version of the tutorial here.  

Discussion: SOD Class 3 - Copyright & Access and SOD Class 3 - Copyright & Access Part II
SOD Class 4- Digitization and Metadata

Week 4 (9.25) Digitization and Metadata :: Zach Vowell, speaker

Guest Lecture: Zach Vowell, Digital Archivist, Dolph Briscoe Center for American History

Zach will talk about the relationships between metadata and the digital objects created by digitization projects, including 1) the functions served by metadata, 2) professional metadata standards, 3) local implementations of metadata standards, 4) embedded metadata vs. external metadata, and 5) considerations for increasing the efficiency of digitization projects through minimal descriptive metadata. For specific examples, he will focus primarily on how metadata is applied to digital images, in preparation for the image digitization lab.

Lab: OCR continued: Ocr Project Expectations

Discussion: SOD Class 4- Digitization and Metadata
SOD Class 5- Image Digitization

**Week 5 (10.2) Image Digitization : Mary Alice Harper, speaker**

Mary Alice Harper- Harry Ransom Center

Lab: Photo (Negative) scanning of the Bob Bailey Collection

Tutorials: [Scanning a archival negative image](https://tutorials.ischool.utexas.edu/index.php/Scanning_a_negative_archival_master)

Metadata guide: [Bailey Metadata Standards, Fall 2012](https://tutorials.ischool.utexas.edu/index.php/Scanning_a_negative_archival_master)

Discussion: [SOD Class 5- Image Digitization](https://utexas.instructure.com/courses/709837/assignments/2047494)
SOD Class 6-Digitization in an Institutional Context

Week 6 (10.9) Digitization in an Institutional Context :: Amy Bowman and Hal Richardson, speakers

Guest Speakers: Amy Bowman and Hal Richardson, Briscoe Center for American History

Lab: Photo metadata and derivatives
SOD Class 7 - Digital Asset Management

Week 7 (10.16) Digital Asset Management :: Margie Foster and Zach Vowell, speakers

Zach Vowell is the Digital Archivist at the Dolph Briscoe Center for American History, and Margie Foster is the Digital Asset Management Librarian for Freescale Semiconductor.

Zach will discuss the critical role of digital asset management in an archive or library setting, as well as the various options that different digital asset management systems (or DAMS) provide for meeting this management need. He will also consider the differences, similarities, and possible interrelationships between a DAMS, an institutional repository, a digital library, a document management system, and a content management system.
SOD Class 8- Audio Digitization I

**Week 8 (10.23) Audio Digitization I**

Guest Lecture: Jeffrey Boruszak, PennSound

Lab: Cassette Digitization
SOD Class 9- Audio Digitization II- Audio digitization and Presidential archives outreach

Week 9 (10.30) Audio Digitization and US Presidential Archives Outreach

Guest Lecture: Presidential Library Archivists Heather Nice and Liza Talbot (Q books)

Lab- Cassette digitization and metadata continued, Q hearing test?
SOD Class 10- Video Digitization

Week 10 (11.6) Video Digitization

Guest Lecture: Megan Peck- Texas Archive of the Moving Image

Lab: VHS Digitization
SOD Class 11- Video and Film Digitization II

Week 11 (11.13) Video and Film Digitization

Guest Lecture: Steve Wilson, Film Curator, HRC, Justin Kovar, Briscoe Center (Q books)

Lab: Video metadata and derivatives
Course Modules

Syllabus

Week 1 (9.4) - Text digitization

Digitization Lab: Texas Oil Industry Transcripts

Tutorial

Introduction to ABBYY Finereader 11

ABBYY_FineReader_11_manual

Week 2 (9.11) Why Digitize? What's digital preservation?

Readings


Smith, Abby "Why Digitize?" CLIR Pub. 80, 1999


OPTIONAL: Association of Research Libraries Recognizing Digitization as a Preservation Reformatting Method. 2004


OPTIONAL- Price, Joseph LOC Optical Disk Project

Digitization Lab:

Tutorial

Week 3 (9.18) Copyright and Access - Carlos Ovalle, speaker

Readings


TED talk on fashion and copyright

Digitization Lab: New York Journal America Newspapers

Tutorial

Newspaper scanning tutorial

Mobile version of newspaper scanning tutorial

Week 4 (9.25) Digitization and Metadata - Zach Vowell, speaker

Readings


OPTIONAL: Dublin Core Metadata Element Set, Version 1.1

OPTIONAL: Briscoe Center Metadata Schema and Style Guide

OPTIONAL: Dublin Core Metadata Best Practices (CDP Metadata Working Group), Version 2.1.1 (Sections 2 and 3 (pgs. 4-17), and skim through Section 4)

**Digitization Lab: OCR scans**

OCR Project Expectations

Week 5 (10.2) Image Digitization: Mary Alice Harper, speaker

Readings


CAH Digitization_Process_08.doc

Moving Theory into Practice- Cornell Digital Imaging Tutorial

OPTIONAL: NMPW_imaging_specifications.pdf

OPTIONAL: Metadata_standards_ischool_Spring2012.doc

**Digitization Lab: Negative Scans**

Tutorial

Scanning a negative archival master

Week 6 (10.9) Digitization in an Institutional Context:: Amy Bowman and Hal Richardson, speakers

Readings


MPLP: It's Not Just for Processing Anymore

Organizing for Digitization at Oregon State University: a Case Study and Comparison with ARL Libraries

Success Factors and Strategic Planning: Rebuilding an Academic Library Digitization Program

Digitization and Researcher Demand

**Digitization Lab: Photo metadata and derivatives**

Tutorial

Creating derivative images from negative archival masters

Week 7 (10.16) Digital Asset Management:: Margie Foster and Zach Vowell, speakers

Readings

Choosing a Digital Asset Management System That's Right for You

National Library of Medicine Digital Repository: Functional Requirements (ch. 7 & 8)

Islandora: An Update on the Project & the Community (2010)

CONTENTdm Information Sheet

The Role of Digital Librarians in Digital Asset Management

Forrester Wave Digital Asset Management Q2 2012.pdf

The role of digital librarians in digital asset management

How to become litigation ready

Who Should Enter Metadata in Digital Asset Management

Asset Metadata cataloging

Week 8 (10.23) Audio Digitization I

2 of 5
Readings
Meeting the Challenge of Media Preservation: Strategies and Solutions; Indiana University Bloomington, Media Preservation Initiative Task Force, Public Version, September 2011
OPTIONAL: CLIR Capturing Analog Sound for Digital Preservation 2006

Digitization Lab: Audio

Week 9 (10.30) Audio Digitization and US Presidential Archives Outreach
Readings
Lee Ann Potter, "Education Programs in the Presidential Libraries," 2006 (Lee Ann is NARA's Head of Education)
Nancy Kegan Smith and Gary M. Stern, "Historical Review of Access to Records in Presidential Libraries," 2006 (Nancy is Director of the Presidential Materials Staff and Gary is NARA's General Council)
OPTIONAL: WebsiteSurvey.DOCX

Digitization Lab: Audio metadata

Week 10 (11.6) Video Digitization
Readings
CCAHA A Race Against Time: Preserving AV Media, Video Preservation 2009
Video Preservation website
Digital Tape Preservation Strategy
OPTIONAL: AV Artifact Atlas

Digitization Lab: Video

Week 11 (11.13) Video and Film Digitization
Readings
Kittler, “Film” Gramophone, Film, Typewriter 115-182, 1999. (Part 2)
CCAHA A Race Against Time: Preserving AV Media, Film Preservation Part 1, 2009
OPTIONAL: National Film Preservation Foundation Film Preservation Guide

Digitization Lab: Film

Week 12 (11.20) Digitization Funding :: Nancy J. Melley, NHPRC, speaker
Readings
Course Modules: SURVEY OF DIGITIZATION https://utexas.instructure.com/courses/709837/modules
Picturing Success Digitization Manual Amy Bowman.pdf

Outcome Based Evaluations (IMLS)

Tutorial
Lab: Grant workshop

Hide module contents    Add item to module

Week 13 (11.27) Grantwriting

Readings

Lab: Grant workshop

Hide module contents    Add item to module

Week 14 (12.4) Final Project Presentations

Hide module contents    Add item to module

Course Documents

Orphaned Readings

Hide module contents    Add item to module

Bush, Vannevar. As We May Think. 1945.
Licklider, J.C.R. Man-Computer Symbiosis. 1960
Microsoft Research: Context-Dependent Pre-trained Deep Neural Networks for Large Vocabulary Speech Recognition
Fortner, Robert Rest in Peas: The Unrecognized Death of Speech Recognition 2010.
?? North Carolina Exploring Cultural Heritage Online (ECHO) . Metadata 2007

Hide module contents    Add item to module

Assignments

Hide module contents    Add item to module

Communication

Hide module contents    Add item to module

Research Tools

Hide module contents    Add item to module

Alternative Tools

Adobe Connect
Participation

Participation means being an active member of the dialog. It consists of doing one or more of the following: being prepared for class and writing comments, by making observations about the readings and exercises, by asking questions, by staying on topic. In class, taking notes, actively working on in-class exercises instead of unrelated activities like email, chat and general web surfing, and actively listening lends to better discussion. It also means showing up and participating EVERY WEEK in the discussion online even when you are not the discussion leader or synthesizer.

Please plan to attend all classes and arrive on time. Please be courteous to the collegial community we are creating this semester by not conversing with others during class lectures; be mindful during the labs that this is time is designed to give you guided practice with some of the digital tools you will need in order to complete your digital assignments.

Unexpected problems happen (serious illness, etc.), it is important to let me know so we can make appropriate changes to your schedule. If you will miss class because of a religious holiday, or travel, let me know as soon as possible so we can make alternative arrangements for this as well.

In many of our classes we will have time in class to get some practical experience with some of the concepts we're learning during the lectures and through readings. These labs will not be collected but will be checked for completion.
Discussion Leaders

Discussion Leaders and Synthesizers (10%):
Sign up page on Collaborations will be opened up after the first day of class.

Your main job as leader and synthesizer is to keep the conversation going.

Leaders: Once in the semester you and a partner will be on the spot as leaders to the discussion for the readings for that week. As leaders, you will each post a thoughtful and productive set of questions as a POST in the discussion section. These questions should respond to the readings (instructions below) and should be posted by midnight on the WEDNESDAY before our class meeting. You should have the readings prepared before this meeting. You will also be responsible for raising these issues during the discussion in class (if we have one, time permitting). You may divide the readings up with your partner but all of the readings should be touched on in some way between you and your partner for that week.

Posting a threaded discussion:

1. Start a threaded discussion topic under "Discussions" and then under the heading already established for the week in which you are the discussion leader (ex: "SOD Class 2- Manuscript digitization")
2. Label the discussion according to the readings you are covering and/or the topic.
3. The second leader should start a new thread under the main thread.

Synthesizers: Once in the semester you and a partner will be on the spot as synthesizers to the readings discussion questions and the readings presentations that are posted on the discussion for a given week. As synthesizer, you will post your responses to the presentations as COMMENTS on the discussion page. It may be multiple comments on any post or on other comments, but your job is to synthesize the conversation. These should be posted by midnight on the THURSDAY before the class period. Consider: what can you add to the issues that you see coming up across the readings and the presentations? What connections do you see happening across them? You will also be responsible for raising these issues during the discussion.
Topic Presentations

In order for us to share our interests within the larger topic of digitization as well as keep one another up to date with cutting edge developments, each week a pair of students will give a 10-15 minute presentation on a current topic/event relevant to the focus of the week’s readings. Please plan to treat these presentations professionally. These are not book reports or summaries of the readings. It is expected that the class has done the reading. The pair will be expected to make an argument based on their research and the readings that corresponds to their interests or background in information science (loosely defined to mean a variety of concepts from computer programming and design to archival theory to library and technology services). Feel free to include your own research or work experience: how does your research and/or work speak to these readings? Does it dovetail or contrast? Why?

The presentation format will require a formal visual aid such as a powerpoint (OR a Prezi, OR an infographic you’ve created on the topic, OR a handout, etc.). You will also need to provide the class with at least 4 citations for resources you used in gathering your information. That way, if students are interested in the presentation topic, they know where to look for more information. Presenters should be prepared to field questions from the class. Presentation topics need to be submitted to the instructor at least two weeks before your presentation date for approval.

Presentation dates sign-up will be opened on Collaborations after the first day of class. Plan on a 5-10 minute discussion in class. Please post your presentation on the Presentations page.

SURVEY OF DIGITIZATION
https://utexas.instructure.com/courses/709837/assignments/1909900
Digitization Portfolio

Your Digitization Portfolio will consist of the items you have digitized during the semester. All of these items except for the video need to be put into your sod_fall_2012 folder on the NAS, so we can access them for grading.

It is our hope that many of these projects will lead to items you digitized having permanent URL's on the Web, and many items digitized during past iterations of this course have become publicly accessible exemplars of our shared cultural record.

Digitization Portfolio (50%)
Text/OCR (15%)
Images (10%)
Audio (10%)
Audio Sync (5%)
Video/Film (10%)
Final Grant Project

Final assignment: Write a grant proposal to obtain funds to digitize historical records.

Your proposal should take the form of a project narrative that conveys the ideas, objectives, and methods of your project based on what we have learned this semester. This proposal will be a response to the Digitizing Historical Records grant program sponsored by the National Historical Publications and Records Commission, a part of the National Archives and Records Administration (NARA). Your proposal will be based on the Digitizing Historical Records Grant Announcement.pdf.

Your proposal should

- include sections 1 – 6 described in the announcement
- include a rough budget based on the parameters described in the Budget Categories section
- be 15 to 20 double-spaced pages (12pt type with 1 inch margins) and
- demonstrate the following:
  1. the national significance of the collections or records series to be digitized;
  2. an effective work flow that repurposes existing descriptive material, rather than creating new metadata about the records;
  3. reasonable costs and standards for the project as well as sustainable preservation plans for the resulting digital records;
  4. well-designed plans that evaluate the use of the digitized materials and the effectiveness of the methods employed in digitizing and displaying the materials.

Because your peers on a review panel may not possess specialized knowledge of the proposed methods and historical field of study, the description should be free of jargon.