# Born Digital/Digital Collections
## UCLA California Rare Book School, August 2022

Instructor: Dr. Trevor Owens, @tjowens, trevorowens.org
Email: trevor.johnowens@gmail.com
Zoom link: https://ucla.zoom.us/my/digitalcollections

**Course Description:**
This course will take a holistic approach to digital content, technologies, standards and policies in special collections. Digital materials are an increasingly routine area of work in special collections, including, but not limited to collection development considerations and appraisal of born digital and digitized materials; acquisition and processing of born-digital materials; planning for management and access to digital surrogates for objects; online finding aids and reference activities; and ensuring long term access to content through digital preservation work. The first and last of these—born-digital materials and digital preservation—are areas of great activity and developing standards and best practices, and the course will include an emphasis on building basic core knowledge in these areas, together with the means to track developing standards and best practices. We will also consider the role of information technology services—either within the home institution or through external partnerships or consortia—in the management of born-digital materials and preservation of both born-digital and digitized content.

During the course we will also consider the role of digitization projects and services in building special collections of the future. Starting from the obvious premise that building collections of digital surrogates has value, we will consider the variety of options available for building services upon those digitized collections, including strategies for providing access and reference services, contextualizing digital content and enriching the collection through the participation of our users, and the increasingly important role of research in the digital humanities.

Since this course is being taught remotely, we will have a combination of synchronous and asynchronous activities. The synchronous portion of each day will be divided into three components: discussion (where we'll discuss readings and topics at a conceptual level), presentations and demonstrations (where we'll hear from experts working in the field on specific digital platforms or workflows), and open lab (where you'll have the opportunity to independently work on a project to apply the NDSA Levels of Digital Preservation to your organization or another organization's policies and ask questions of the instructor and each other.)

**Learning Objectives:**
- Develop an understanding of core functions and processes involved in the full digital content lifecycle for digital collections (acquisitions, ingest, processing, preservation, and access.)
- Become familiar with major software applications and platforms in use for work with digital collections
- Develop a high-level understanding of workflows, standards, and practices needed to support a cultural heritage digitization program
- Develop a high-level understanding of key aspects and features of born digital materials that libraries, archives, and museums are acquiring
- Gain hands on experience using the NDSA Levels of Digital Preservation as a practical tool for assessing the maturity of an organization's digital preservation/digital collections management program

## Schedule (Pacific Time)

**Before the first-class session read**

- o **Chapter one**, Beyond Digital Hype and Digital Anxiety **and chapter two** Artifact, Information, or Folklore: Preservation's Divergent Lineages in *Theory and Craft of Digital Preservation* (https://osf.io/preprints/lissa/5cpjt )
- o The NDSA Levels of Digital Preservation: An Explanation and Uses (https://osf.io/dpnqs/ ).

## Day 1: Fundamentals of Digital Collections, Mon Aug 1st

- 9:00 – 10:30: **Initial Discussion Session**
  - o Overview of course from instructor. Student and instructor introductions. Discussion of the first readings. Key concepts include defining digital preservation, artifactual, informational, and fixity.
- 11:00-12:30: **Workshop on Using the NDSA Levels of Digital Preservation**
  - o For this workshop, we will review the NDSA Levels of Digital Preservation Assessment Tool ( https://osf.io/m6j4q/ ) and begin working to attempt to use it for a model organizational self-assessment.
- 2:00-3:00: **Open Lab on NDSA Levels**
  - o Time to attempt applying the NDSA Levels tool to begin reviewing your organization, or if you do not have an organization, a time when you can connect with another student to collaborate on their attempts to use the tool for their organization. Instructor is present to answer questions and provide further explanation on any questions that arise.
- Your Time:
  - o In advance of tomorrow's discussion, read
    - ▪ **Chapter three** Understanding Digital Objects and **chapter four** Challenges and Opportunities of Digital Preservation in *Theory and Craft of Digital Preservation* (https://osf.io/preprints/lissa/5cpjt )
    - ▪ Rudersdorf, A. (2016). The proof is in the standards pudding [AVPreserve]. https://blog.weareavp.com/the-proof-is-in-the-standards-pudding
    - ▪ Arms, C., & Fleischhauer, C. (2005). Digital formats: Factors for sustainability, functionality, and quality. In *Archiving Conference* (Vol. 2005, pp. 222–227). Society for Imaging Science and Technology. http://memory.loc.gov/ammem/techdocs/digform/Formats_IST05_paper.pdf.
    - ▪ Skim -> Library of Congress Recommended Formats Statement https://www.loc.gov/preservation/resources/rfs/RFS%202021-2022.pdf
    - ▪ Skim -> Lavoie, B. F. (2014). *The Open Archival Information System (OAIS) Reference Model: Introductory Guide*. Retrieved from http://dx.doi.org/10.7207/TWR14-02

## Day 2: Understanding Digital Content and Formats, Tue Aug 2nd

- 9:00 – 10:30: **Discussion of Digital Content and Formats**
  - o We will discuss the readings. Key concepts include, bitstream, file, file format, compression, screen essentialism, embedded metadata, and disk images.
- 11:00-12:00: **Workshop on Glitching Files**
  - o For this workshop, we will work through examples of glitching files as outlined in this blog post (https://blogs.loc.gov/thesignal/2012/11/glitching-files-for-understanding-avoiding-screen-essentialism-in-three-easy-steps/ ). The instructor will demonstrate and then

- 2:00-3:30: **Demo and Presentation on Bitcurator**
  - ○ **Hannah Wang, Project Manager, BitCuratorEdu**
- Your Time:
  - ○ In advance of tomorrow's discussion, read
    - ▪ **Chapter five**, The Craft of Digital Preservation and **chapter six** Preservation Intent and Collection Development, in *Theory and Craft of Digital Preservation* (https://osf.io/preprints/lissa/5cpjt )
    - ▪ The Library of Congress Web Archiving Collection Policy Statement (https://www.loc.gov/acq/devpol/webarchive.pdf ) and Open Digital Content Collection Policy Statement (https://www.loc.gov/acq/devpol/opencontent.pdf ).
    - ▪ Dappert, A., & Farquhar, A. (2009). Significance Is in the Eye of the Stakeholder. In M. Agosti, J. Borbinha, S. Kapidakis, C. Papatheodorou, & G. Tsakonas (Eds.), *Research and Advanced Technology for Digital Libraries* (pp. 297–308). Springer Berlin Heidelberg. Retrieved from https://planets-project.eu/docs/papers/Dappert_Significant_Characteristics_ECDL2009.pdf
    - ▪ Skim DRAFT Technical Guidelines for Digitizing Cultural Heritage Materials - 3rd Edition

## Day 3: Preservation Intent, Digitization, & Collection Development, Wed Aug 3rd

- 9:00 – 10:30: **Discussion of Preservation Intent, Digitization, and Collection Development**
  - ○ We will discuss the readings. Key concepts include, digitization standards, preservation intent, significant properties, collection development policy
- 11:00-12:30: Digitization Workflows at UNT
  - ○ **Marcia McIntosh, Digital Production Librarian, University of North Texas**
- 2:00-3:30: **Demo of Archive-It and IA Vault**
  - ○ **Jefferson Bailey, Director of Archiving and Data Services, Internet Archive**
- Your Time:
  - ○ In advance of tomorrow's discussion, read
    - ▪ Chapter seven, **Managing Copies and Formats** in *Theory and Craft of Digital Preservation* (https://osf.io/preprints/lissa/5cpjt )
    - ▪ OSSArchflow resources https://educopia.org/ossarcflow/ read the Digital Dossiers and the *Guide to Documenting Born-Digital Archival Workflows* in full. Skim the other resources.
    - ▪ Rimkus, K., Padilla, T., Popp, T., & Martin, G. (2014). Digital Preservation File Format Policies of ARL Member Libraries: An Analysis. *D-Lib Magazine*, *20*(3/4). http://doi.org/10.1045/march2014-rimkus
    - ▪ Also, look at the digital preservation policy links for the afternoon session tomorrow.

## Day 4: Managing Copies and Formats for Preservation, Thur Aug 4th

- 10:00 – 11:00: **Discussion of Managing Copies and Formats for Digital Preservation**
  - ○ We will discuss the readings. Key concepts include, fixity checking, file format identification,
- 11:00-12:30: **Demo of Archivematica**
  - ○ Katie Punia, Systems Archivist, Artifactual Systems
- 2:00-3:30: **Digital Preservation Policy Lab/Review**
  - ○ Overview of the Library of Congress DCMC https://www.loc.gov/programs/digital-collections-management/about-this-program/

- o LC Recommended Formats Statement:
    https://www.loc.gov/preservation/resources/rfs/RFS%202022-2023.pdf
  - o LC Sustainability of Digital Formats Site:
    https://www.loc.gov/preservation/digital/formats/
  - o Duke University Libraries Digital Preservation Policy
    https://library.duke.edu/using/policies/digital-preservation
  - o NARA Digital Preservation Strategy https://www.archives.gov/preservation/digital-preservation/strategy
- Your Time:
  - o In advance of tomorrow's discussion, read
    - ▪ **Chapter eight**, Arranging and Describing Digital Objects and **chapter nine** Enabling Multi-modal Access and Use in *Theory and Craft of Digital Preservation* (https://osf.io/preprints/lissa/5cpjt )
    - ▪ DLF Levels of Born Digital Access https://osf.io/hqmy4/

## Day 5: Arranging, Describing, & Enabling Use of Digital Content, Fri Aug 5th

- 9:00 – 10:30: **Discussion of Arranging, Describing, and Enabling Use of Digital Content**
  - o We will discuss the readings. Key concepts include, MPLP, open access, rights restrictions, ethics of access,
- 11:00-12:30: **Demo of Library of Congress Digital Collections**
  - o Show access to Web Archive Collections https://www.loc.gov/web-archives/collections/
  - o Show access to Digitized Manuscript Collections
    https://www.loc.gov/manuscripts/collections/
  - o Show Access to Datasets https://www.loc.gov/collections/selected-datasets/about-this-collection/
  - o Show By the People Crowdsourcing Platform https://crowd.loc.gov/
- 2:00-3:00: **Closing Discussion**
  - o Come prepared to discuss one or two key take aways from the course
  - o Return to your initial work on the NDSA Levels of Digital Preservation Assessment tool. With the course completed, what do you see as the most critical next steps for advancing/developing your organization's digital collections program?