

Digital Archives Transfer and Preservation Plan

Background

Most LAUC Committees are now working collaboratively and sharing their documents on the cloud shared storage platform Box (<https://www.box.com/home>). Many of these documents are considered permanent records which should be carefully managed and transferred to the LAUC Archives.

One of the challenges in managing digital files in shared storage across a large organization with many users is ensuring consistency in routine housekeeping on documents stored in the shared location. A lack of consistency can result in the unnecessary and potentially costly storage of old drafts and other records without institutional value. The naming of files, folders, and sub-folders may be inconsistent or meaningless, and the folders and sub-folders may not be logically organized. Consistent, meaningful file names and file structures are especially hard to maintain with the addition of new employees, the loss and retirement of old employees, the regular turn-over of LAUC volunteer positions, and the close of projects or task forces.

Recommendations

The LAUC Archives Implementation Task Force therefore recommends that members submit their archival documents directly to their Division Chairs, who will organize material according to the Archives Guidelines for transfer to the LAUC Archives. The LAUC President or LAUC Secretary will manage file transfer for Board and statewide records. The LAUC Records Coordinator will work with said groups to transfer files, and responsible officers will complete the Records Transmittal Form, arrange for shipment of physical files and folders, and upload all digital documents to a specific Box folder for the LAUC Archivist to retrieve. The LAUC Archivist is currently the Associate University Archivist at UC Berkeley.

Sending digital files by physical media is not recommended due to cost, shipping, storage, and risk of damage to the media. Laptops are not collected by the LAUC Archives and should not be sent. Large files or sets of files can be transferred by zipping the files before upload.

This document outlines and explains the different steps and considerations of the transfer and preservation process for digital material in the LAUC Archives.

Tips for File Management

1. File and Folder Naming Conventions

Use meaningful file and folder names

- a) Accurate and meaningful file and folder naming requires that filenames (as well as folder structures) make sense to other users, present and future. To encourage continued accessibility of documents, structure the folders and sub-folders logically. Avoid creating hierarchies more than 4-5 folders deep to ensure that documents are not buried in many levels, and to prevent unnecessary duplication of documents.
- b) A file, folder, or sub-folder name should be:
 1. interpretable by future users
 2. distinguishable from files with similar subjects, and from different versions of the same file
 3. consistent across the department or workgroup
 4. In compliance with established naming conventions
- c) Avoid using special characters in a filename or folder:
 1. The computer may interpret special characters like: \ / : * ? " < > | [] & \$, . as specific to a task unrelated to the filename. The characters listed above are frequently used for specific tasks in an electronic environment. For example, a forward slash is used to identify folder levels in Microsoft products, while Mac operating systems use the colon.
 2. Periods are used in front of file-name extensions to denote file formats such as .jpg and .doc; using them in a filename could result in lost files or errors.
 3. To avoid "floating" folders to the top of alphabetical lists, don't use symbols or numbers as the first character. For example, use "Projects" instead of "_Projects". This will avoid losing folders and unnecessary duplication.
- d) Use underscores instead of periods or spaces:
 1. As mentioned above, periods already have a specific function in a filename, which is to tell the computer program where the file-name extension begins.
 2. Spaces are frequently translated in a Web environment to be read as "%20". For example, if the filename "Naming tutorial.doc" was published online, the filename would appear as "Naming%20tutorial.doc". This alteration can cause confusion in identifying the actual filename.
 3. Spaces in filenames can also cause broken links, because word processing tools like Microsoft Word, and e-mail clients like Microsoft Outlook, recognize spaces as an opportunity to move to another line.

- e) Use brief but descriptive file names:
1. Try to limit the number of characters in a filename to 25 or fewer. The filename should include all necessary descriptive information to identify the file without exceeding the maximum length set by your computer's operating system.
 2. Operating systems and software differ in acceptable length of file names; some allow up to 256 characters, while others allow far fewer. Keeping filenames short means you are less likely to experience errors when copying or sharing files with other people who may use different operating systems.
 3. Context is particularly important to include in the filename. For example, a file called "agenda.pdf" doesn't provide any clues as to the function or date of the agenda. A better filename would include a fuller description like:
"20161116_LAUC_Archives_agenda.pdf"
 - i. Notice the underscores used instead of spaces
 - ii. Notice the date at the beginning of the filename, in YYYYMMDD format (see section f, below)
- f) Include dates and format them consistently:
1. The best way to list the date is based on an international standard: ISO 8601 (<http://www.iso.org/iso/home/standards/iso8601.htm>). ISO 8601 specifies the numeric representations of date and time to be used in electronic format.
 2. The international standard date notation is: YYYY_MM_DD or YYYYMMDD. YYYY is the year, MM is the month of the year between 01 (January) and 12 (December), and DD is the day of the month between 01 and 31.
 3. This format allows ease of sorting and comparing files by date and prevents confusion with other date formats (especially in other formats that use just two digits for the year).
 4. Example: 20161116
- g) Include a version number on your documents:
1. The easiest way to do this is to use the letter "v" to represent "version number." Then, "v01, v02, v03" can be added as needed to a file and the main file name can stay the same.
 2. This is much more effective than other common additions like "update," "new," "old," etc.
 3. An exception to this rule is using "FINAL" to indicate the final version of the document. This can be helpful to quickly identify the most accurate version of the document. When using "FINAL", be sure to use it instead of the version number, rather than in addition to it.
 4. Examples:

20161116_LAUC_Archives_agenda_v01.pdf
 20161116_LAUC_Archives_agenda_FINAL.pdf

h) Be consistent:

1. The most important rule of file-naming is to be consistent. Some choices will need to be made about organization that affects the entire workgroup – where to include the date, what abbreviations to use, etc. Regardless of what the group decides, it is only effective if everyone follows the rules consistently.

2. File Formats

a) Identify and convert file formats:

- i. For long-term digital preservation, it is especially important to consider the suitability of the file formats. As versions of software change, the Archives may not be able to open the files in their original applications.
- ii. Acceptable files formats:
 1. Documents: PDF
 2. Images: JPEG, TIF
 3. Video: Quicktime .mov (uncompressed) H264

3. Transfer Description and File List

In order to check over time that your digital files are being preserved, it is necessary to record exactly *what* files are being transferred. This list should contain information such as file names, and if possible, locations, sizes, and format types.

a) Include the following information with every digital records transfer:

- i. “Readme.txt” text file document including:
 1. **Name:** Enter the name of the person making the transfer. If the Archives staff have questions or if the files did not transfer properly, they may need to contact you.
 2. **Department:** Name of the department that is transferring the records.
 3. **Description:** Briefly summarize the files.
 4. **Dates of records:** Provide either exact dates, or a range of dates.
 5. **Restrictions** (if any): Most collections in the Archives are open for research, but restrictions may be necessary if your documents contain private or sensitive data.
 6. **Copyright:** In most cases, LAUC owns the copyright for the files you will be transferring to the Archives. Some examples of when the copyright

owner is different could include works by photographers and other companies.

- ii. A copy of the directory structure and list of files (“inventory”)
 - 1. There are several free or open-source programs that allow users to quickly generate a detailed list of all files and sub-folders within a main folder.
 - 2. For those familiar with using the command line interface (CLI), directory lists can also be easily created with a few simple commands.
- iii. If you are transferring files on a physical object such as a CD, DVD, or USB flash drive, make sure to add a label on the outside, and if possible, add an internal label by “naming” the disc or drive in your operating system .

4. Transfer the Data

- a) Upload files to Box:
 - i. Log in, select the “File” tab from the menu bar, then click the little arrow next to the upload button and select "Bulk Upload".
 - ii. After accepting an applet, you'll be presented with the "Drag and Drop" box.
 - iii. Click "Add files" and then choose whichever folders and files you wish to upload.
 - iv. If file size is too big, try compressing files into a ZIP:
 - 1. On a Mac, right click your folder and select “Compress [name of folder]”
 - 2. On a PC, right click your folder and select “Send to – Compressed (zipped) folder”
- b) Contact the appropriate LAUC officer (LAUC Records Coordinator, Secretary, or President) and follow the guidelines for transferring records to the LAUC Archives using the Records Transmittal Form (<http://guides.lib.berkeley.edu/bancroft/transfer-university-records>)

5. Preserve the Data

(Or, what happens once files are uploaded to Box and the Records Transmittal Form is submitted?)

- a) The LAUC Archivist will be notified that an upload has been made and digital files are ready for retrieval.
- b) The digital files will be added to the Bancroft Library’s digital collections for long-term storage.
- c) Depending on the processing schedule, there will be a delay from when the records are received and when the collection record and finding aid are updated.

- d) Access copies of digital files will be added to the collection and available upon request through the reading room laptop. Copies of files in the LAUC Website Document Library are maintained separately by LAUC and will remain accessible to members.

LAUC Website Preservation

Captures of the LAUC website are found on Archive-It under the collection title Librarians Association of the University of California (LAUC). They are considered part of the UC Berkeley Libraries' collection. There are two collections, each with a seed URL:

1. <http://lauc.ucop.edu> includes captures from 6/4/2013 to date
2. <http://www.ucop.edu/lauc/> includes captures from 10/13/2007 to 7/4/2015

Crawls occur monthly. The crawl scope ranges between two and approximately four clicks, capturing html or pdf documents, in most cases. Some content falls outside the crawl scope, however:

LAUC blog: Information in the "Featured Content" section (e.g. <http://wayback.archive-it.org/all/20161110190705/https://lauc.ucop.edu/featured-content>) that provides highlights to members' activities is largely inaccessible. The main page content is captured, but generally no content beyond that point. The initial page of the 2007-2014 blog has been captured (e.g. <http://wayback.archive-it.org/all/20150904085634/http://laucassembly.blogspot.com/>), but it does not allow access beyond that page. The "Not in Archive" note provides URLs for searching the live web. The 2007-2014 content will remain available at <http://laucassembly.blogspot.com/>.

LAUC wiki: Most of the content beyond the home page is inaccessible, with the exception of the 2010 Campus Reports entry (http://wayback.archive-it.org/all/20150704071236/http://laucwiki.lib.ucdavis.edu/index.php/Campus_Reports). The "Not in Archive" notes provide alternative search options.