Preserve it or Lose it!

Preserving your personal pictures and documents

April 27, 2021
Millington Arbela District Library
Scott Witmer / Shannon Zachary
Today’s Agenda

- Books, photos & documents
- Personal digital archiving
- Q & A

Slides at “Caring for personal books and documents”

- Shannon Zachary, Department Head
- Scott Witmer, Digital Preservation Specialist

U-M Library, Department of Preservation & Conservation
Books, photos, & documents

- What is this stuff?
- What goes wrong?
- An ounce of prevention . . .
- Cures
What is this stuff?

- Plants
- Animals
- Minerals
- Plastics
- ... or a combination of the above

Black and white family photos and letters from the 1960s.
What goes wrong?

- Biological agents
  - Mold, insects, pests
- Chemical deterioration
  - Pollution, inherent chemicals
- Physical (mechanical) damage
  - Abrasion, wear, tear
- ... or a combination of the above
Biological damage

- Requires water to grow
- Pretty much everywhere
- Prefers dirty, undisturbed spaces
Chemical deterioration

- **External**
  - Pins, staples, clips, rubber bands, Post-Its, Scotch tape
  - Air pollution, light

- **Inherent**
  - Acid paper
  - Leather deterioration

Metal binding clip has left rust stains on the documents it held together.
Light-induced fading

- Cumulative, can’t be reversed
- *All* light can cause damage, not just UV light

1930s group photo showing light damage.
Acid paper

- Inherent acids
  - Acids from paper pulp
  - Acids added during papermaking
- Acid migration
  - Acidic boxes and folders
  - Acidic inserts
  - Air pollution

1935 Detroit newspaper showing yellowing and embrittlement from acid deterioration.
What to do about acids?

- Store in a cool, dry location
- Use archival, buffered boxes and enclosures
- Handle with care
- Make copies
- Deacidify?
Red rot on leather

- Inherent acid deterioration
- Can’t stop or reverse
- Conservator can apply a consolidant
- Leather dressing - DON’T!!
Physical/mechanical damage

- Conservator can repair
- IF the material is strong enough
- Or keep in a box

Book with torn paper from being dropped.
The mummy’s secret: how to last thousands of years

- Low/steady temperature or freezing
- Low moisture
- No light
- No air
- No pollution
- No people!
So you don’t have an Egyptian tomb handy . . .

- Store important stuff in living space; no attics, basements, or garages!
- 4-inch rule
  - 4 inches up off the floor
  - 4 inches away from any wall
- Put it in a box; select archival paper or archival plastic enclosures
- Handle with care
- Consult a conservator
Personal Digital Archiving

- **Digitization**
  The process of creating a digital version of a physical media source

- **Digital Preservation**
  The ongoing actions required to maintain access to digital materials over time
What are digital archives?

- **Digital**: Information stored and accessed on a computer system
  - Digitized: digital copies of physical (analog) sources
  - Born-Digital: files created on the computer
- **Archives**: Records of your life and activities, virtual memories
  - Variety of media formats: data, documents, photos, sound and video recordings
Why personal digital archiving?

- Risks to digital materials:
  - Computer crashes
  - Evolving hardware/software environments (obsolescence)
  - Data corruption
  - Overwhelming volume of digital content

- Many different reasons:
  - Long-term preservation and access
  - Sharing and reusing digital content
  - Organization and management
Everyone uses their digital materials in different ways. Do what’s best for you!
Creating digital files

- **File Formats**: Encoding standards and specifications for rendering digital content as types of media in a computing environment.
- Factors that archivists consider when selecting sustainable file formats for long-term preservation:
  - File Size vs. Storage Capacity
  - Compression: Lossy vs. Lossless
  - Proprietary vs. Open Standard
  - User Support
Digitization

- **Do-It-Yourself**
  - Cost of technology
  - Privacy/Security
- **Vendor**
  - Cost of expertise/service
- **Home Digitization Guide:**
  - [https://apps.lib.umich.edu/sites/default/files/departments/lit/DCU/quick_and_dirty_guide_for_home_digitization.pdf](https://apps.lib.umich.edu/sites/default/files/departments/lit/DCU/quick_and_dirty_guide_for_home_digitization.pdf)

Scanning Terms:

- **Bitonal**
  - two-tone black and white scans
- **Grayscale/Color**
  - palette of 256+ tones
- **Resolution**
  - finely spaced detail, a measure of a sharpness of an image
- **DPI or PPI**
  - a measure of dots of color or shading per inch, which affects the image resolution. The more dots per inch, the finer the image and the larger the file size.
Text Documents

Digitization

- **Good enough:**
  - Use default scanner settings
  - PDF
- **Recommended:**
  - TIFF
  - 150–300dpi (small text)
  - 8-bit greyscale, 24-bit color
- **OCR (optical character recognition)**
  - Converts images of text to editable text

Born Digital

- **Good enough:**
  - PDF
  - DOCX/XLSX
  - convert from pre-2007 DOC/XLS
- **Recommended:**
  - TXT/CSV
  - PDF/A (formatted)
Images

Digitization

- Good enough:
  - JPEG
  - Use default scanner settings, 72dpi
- Recommended:
  - TIFF
  - 150–300dpi (fine detail)
  - 8-bit greyscale, 24-bit color

 Born Digital

- Good enough:
  - JPEG
- Recommended:
  - TIFF
  - RAW (requires conversion to view)
Terms:

- **Sampling rate:**
  - Times-per-second sound wave signal is captured
  - 44.1 kHz = CD quality, good enough for human hearing

- **Bit depth:**
  - bits per sample; higher depth = wider range of sounds captured per sample
  - 16 bit = CD quality
  - 24 bit = DVD/Blu-ray audio

- **Bit rate:**
  - bits stored per second, kilobits per second (kbps)
  - 320 kbps

- **Channels:**
  - Audio streams/tracks
  - 2 channel = stereo

Graph showing analog audio wave in relation to digital representation measured by bit depth and sampling rate
Audio

- **Good enough:**
  - MP3
    - compressed (lossy)
    - common for streaming and mobile devices
  - AAC
    - compressed (lossy)
    - better quality than MP3

- **Recommended:**
  - WAV/BWF
    - uncompressed
  - FLAC
    - compressed (lossless)
    - not supported in iTunes, Apple proprietary = ALAC
Terms:

● Frame rate:
  ○ Frames/images displayed per second (fps)
  ○ 24 fps = film; 60 fps = HDTV

● Resolution:
  ○ Vertical and horizontal pixel dimensions
  ○ 640x480 = standard definition (SD)
  ○ 1280x720 = basic high def (HD)

● High quality uncompressed = very large file sizes! 40+ GB/hour for SD

Good enough:

○ MP4/M4V: compressed, common for streaming and mobile devices

○ Codecs: encoder/decoder compression standard
  ■ H.264
  ■ H.265: supports Ultra HD

Recommended:

○ AVI: uncompressed (lossless)
Preserving your digital archive

- Select
- Gather
- Organize
- Back Up
- Maintain
1. Select

- Start small
  - Collections of subjects
- What digital content is meaningful to you?
- What would you most regret losing if all of your digital content disappeared?

- Volume can be a problem!
- Don’t get burned out trying to preserve everything
- Weeding
  - Use
  - Quality
  - Duplication
  - Subject
2. Gather

- Copy all the files you want to preserve to a central computer
3. Organize

● **Metadata:**
  ○ Descriptive information will make your files easier to browse and search

● **Group by directory/folder**
  ○ Readme files: Short text files that describe the contents of a folder

● **File naming:**
  ○ Unique and descriptive
  ○ Short and simple: <25 characters, use acronyms
  ○ No spaces: use underscore _ or camelCasing
  ○ Avoid special characters: ^ ~ \ : * ? " < > | ! # % & $ , .
  ○ Start with date: YYYY-MM-DD
  ○ Whatever works best for you

● 2021-04-27_MALib_slides.pdf
4. Back Up

● 3-2-1 Rule:
  ○ 3 – Make 3 copies
  ○ 2 – At least 2 different types of storage media
  ○ 1 – Keep 1 in a different location

Types of digital storage

● External hard drives
  ○ Solid state vs. spinning disk

● Cloud services

● Not recommended:
  ○ Optical discs (cd-r, dvd-r)
  ○ USB thumb/flash drive

Digital storage is physical media: Keep it in conditions that you find comfortable. (~70 degrees, 40% humidity, away from dirt, moisture, and pests)
* Notes on the “Cloud”

- What are the ongoing costs?
- Where is your data stored?
- How secure is your data?

The OVHcloud data center in Strasbourg, France suffered a fire on March 10, 2021.
5. Maintain

- Check your backup storage annually
  - Test by copying and opening up a few files
- Refresh storage media every 5–7 years
More information


Scott Witmer switmer@umich.edu
Shannon Zachary szachary@umich.edu