

Donor guidelines for electronic or A/V materials coming into Woodson Research Center (WRC)

What to expect from Woodson:

- **A file inventory within approximately one month's time of delivery** including filename, size, format and indication of successful storage in our preservation environment. This environment follows our [Preservation Levels](#) and includes two levels:
 - Digital items in Rice's institutional repository, scholarship.rice.edu, freely accessible online access copies, dark master storage, stored redundantly at Rice with nightly file validity checks, plus more geographically diverse back-up via the MetaArchive (for online access copies) and DuraCloud (for dark masters).
 - Digital items not appropriate for freely accessible online access (often because of copyright or privacy issues), are stored on Fondren internal servers with regular tape back-up, accessible to researchers as appropriate per access guidelines.
- **A sampling of file checks** will be conducted by WRC staff to look for corrupted files or otherwise inaccessible materials. *Woodson cannot guarantee that every instance of corrupted materials would be found, but we will do our best to keep the files in working order for the long term.* Using our [Recommended File Formats](#) will help ensure the viability of your files.
- **Donations of materials over 500GB in size or donations which have formats other than our [Recommended File Formats](#)** may require special arrangements, to be discussed at time of delivery.
- Once the collection is fully processed, a finding aid describing it will go online, and if materials themselves are appropriate for online access via Rice's institutional repository, scholarship.rice.edu, they will be added as space allows.

What Woodson needs from the donor:

- Check your files before delivery, looking for any corrupted files which will not work.
- Follow our [Recommended File Formats](#) to help ensure the viability of your files.
- Please include descriptions of the materials (subjects, dates, software used)
- Use consistent file names that indicate the subject content, with no special characters or spaces (underscores or dashes are ok – see these file naming tutorials: [1](#), [2](#), [3](#), [4](#))
- Keep a copy for yourself when possible and do not discard it until Woodson can confirm that our copy is not corrupted and has been stored in our preservation environment.
- Deed of Gift which transfers physical ownership to Woodson and explains donor's preference on copyright / intellectual property status.

Method of transfer to Woodson:

- Delivered on external hard drive, by FTP or other media as coordinated with WRC.

VIDEO

Digital Video formats:

Preferred:

- Archival Masters – uncompressed 8 or 10 bit .avi or Quicktime (.mov) file format
- or original digital format such as .mts files (which can be exported into archival master format)

If also available, or if there is no uncompressed archival master or original digital file:

- DVD quality access copies – DVCPRO50, DV50, DV25 (prefer A264 video codec & aac audio codec)
- Web quality streaming files – MPEG-4 (prefer A264 video codec & aac audio codec)

Analog video formats such as original VHS, Beta, SuperBeta or Umatic tape can be accepted as well, with the understanding that preservation steps including conversion to digital format will be undertaken based on priorities on subject content, format obsolescence and age of the material.

For all video, see [Moving Images Preservation Levels](#) as described on Fondren's Digital Projects wiki.

AUDIO

Digital audio formats

Preferred:

- Archival masters in .wav format or broadcast wave @ 24 bit depth, 96 kHz sampling rate
- Access versions in .mp3 streaming format

Analog audio formats such as audiocassette and open reel audio can be accepted as well, with the understanding that these formats bring their own preservation challenges and will be handled according to our established [preservation levels](#).

Digital Images

Preferred:

- Archival format of TIFF (uncompressed) or JPEG2000 (lossless) (*.jp2), @ 24 RGB-color bit depth, 300 ppi scanned at actual size OR Visual arts materials capture at 3000 to 5000 pixels on the long sized

If also available, or if there is no archival master:

- Access quality files such as: JPEG/JFIF (*.jpg), GIF (*.gif), BMP (*.bmp), PNG (*.png)

Text

Preferred formats:

- Masters as plain text (.txt) or XML (.xml)
- Access copies as - Plain text (ISO 8859-1 encoding); PDF (*.pdf) (with embedded fonts) and OCR'd; Rich Text Format 1.x (*.rtf); HTML (include a DOCTYPE declaration); PDF/A; PDF/X; Open Office (*.sxw/*.odt)
- Woodson is able to accept Microsoft Word documents, but that proprietary format is not preferred because layouts, fonts and even information can be lost in the conversion to the long term preservation formats listed above.

Spreadsheets / databases

Preferred formats include:

- Comma Separated Values (*.csv); Delimited Text (*.txt); Excel (*.xls); DBF (*.dbf); OpenOffice (*.sxc/*.ods)

Presentations:

Preferred formats include:

- OpenOffice (*.sxi/*.odp); PowerPoint (*.ppt); PDF (*.pdf) (with embedded fonts) and OCR'd; PDF/A

Email

Overall:

- Donors should weed and organize their e-mail accounts before donating them if possible, as archivists will not have time to weed them and the donor knows what is worth preserving.

Donors can preserve e-mail of lasting value in several ways – see

http://www.digitalpreservation.gov/personalarchiving/documents/archive_email.pdf

- Donor can select and **print individual e-mails** and store them as paper records, including attachments
- Donor can select and **save individual e-mails as text files** and their attachments according to Recommended File Formats
- Donor can work with archivists **to preserve an entire e-mail account including its relationships between messages and folders in electronic format.**

- Scenario 1: Can use CERP's EMAP schema for e-mail preservation in XML format, preserving hierarchies of folders, messages and their relationships to attachments which are stored separately.
- Scenario 2: Can save e-mails as txt, pdf, or eml files. Use .txt if colored text does not factor in heavily, PDF if color / inline images matter. If saved as .eml files, these can be opened with Mozilla Thunderbird (free software). Attachments need to be saved and organized separately.

More technical e-mail guidelines and tools

- [PeDal project for Outlook mail](#) (.pst files)
- [Understanding where your e-mail is actually stored and purposefully organizing it](#) for preservation
- [XML-based E-Mail Account Preservation \(EMAP\) schema](#) developed by the Collaborative Electronic Records Project (CERP).

General Links:

Recommended file formats at Fondren Library:

<https://digitalriceprojects.pbworks.com/w/page/58487493/Recommended%20file%20formats>

Preservation levels at Fondren Library:

<https://digitalriceprojects.pbworks.com/w/page/59112512/Moving%20Images%20Preservation%20Levels>

Internal workflow of care for A/V materials at Fondren Library:

<https://digitalriceprojects.pbworks.com/w/page/57796817/Audio-Video%20processing%20workflow>

Example of [Woodson Research Center Deed of Gift](#)

Personal Archiving (Library of Congress)

<http://www.digitalpreservation.gov/personalarchiving/>