Digital Record AND Digitization Standards

THE CHURCH OF

JESUS CHRIST
OF LATTER-DAY SAINTS

Introduction

Digital records consist of two types—those that are born-digital (such as photographs taken with a digital camera) and those that are digital images of physical records that are digitized or "scanned."

Digital records of enduring, historical value are collected, housed, and stored in behalf of the Church at record preservation centers. Because they belong to the Church, these digital records should be sent for preservation to the Church History Library (CHL) in Salt Lake City, Utah, USA.

This document outlines the standards for digitizing Church history collections. The term "collection" refers to one or more objects (documents, books, artifacts, and so on), that are usually grouped and identified according to the name of their donor or creator. Before attempting to digitize a collection, you should catalog it using the cataloging worksheet provided in the Collecting Records Guide. Doing so will greatly facilitate the digitization process. It will help you keep the resulting files organized and correctly named.

When you send digital records to the Church History Library, they will be preserved in the Church's Digital Records Preservation System (DRPS). This sophisticated system was implemented to preserve the Church's records of enduring, historical value in perpetuity. If digital rights allow, the digital records will also be made available for viewing online via the Church History Catalog at history.lds.org.

To preserve digital records properly, follow certain standards when (a) digitizing physical records and (b) preparing to send either born-digital or digitized records to the Church History Library. The purpose of this document is to establish these standards and provide guidance for using them. If you do not follow these standards, the digital records you submit may not be accepted, and you will be asked to fix any errors and resubmit the records.

To select a particular scanner, follow the recommendations below. For more information, contact a representative of the Global Support and Acquisitions Division of the Church History Department.

A flatbed scanner is the preferred device for all digitalization efforts. However, if an item is too fragile or too large for a scanner, you may choose to use a digital camera and tripod. Using a camera poses many more technical challenges. Therefore you should contact the Global Support and Acquisitions Division for advice before beginning.

Environmental Considerations

1. Power: Make sure that your scanner meets the power requirements of your area.
2. Interface: Ensure that your scanner meets the proper interface requirements for your laptop or PC (USB, Firewire, and so on).
3. Drivers: Make sure that your scanner has supported drivers for the operating system on your laptop or PC.

Operational Considerations

1. Portability: Choose the type and size of scanner that meets your operational needs for portability. If you believe you will be transporting your scanner frequently, choose a model that has a locking feature for the scanning element.
2. Scanning Speeds: Determine what scanning speeds you require for your operational needs. These specifications are typically defined as "pages per minute (ppm)" or "images per minute (ipm)" and will change depending on the settings you use. Settings that can affect the speed are:
   a. Simplex—one side of a document.
   b. Duplex—both sides of a document.
   c. Color versus grayscale—color scanning typically requires more scanning time than grayscale.
   d. Resolution—the higher the resolution, the longer the time required to scan.
3. Support: Make sure that your scanner and scanning software come with technical support and available instruction.
4. Make sure that the scanner you choose can handle the various paper sizes you will be scanning (now and in the future).

Quality Considerations

1. Resolution: Pixels are the dots that make up an image on a display screen. Resolution is the number of pixels that the scanner can "copy" from the original record during a scan. The higher the resolution, the better the quality and detail of the image. Good resolution is important for both readability and future file format transformations. It is recommended that you choose a scanner with an OPTICAL resolution of at least 600 DPI (dots per inch).
2. Bit Depth: Bit depth is the amount of digital information (color, shading, sharpness) within each pixel that is scanned. The following specifications are recommended:
   a. Grayscale—minimum of 8-bit.
   b. Color—minimum of 24-bit.
3. The software that you use with your scanner should offer some level of color management capability.

The file formats shown in table 1 below are recommended preservation formats for digitized records. All materials sent to the Church History Library should conform to these standards. If you cannot conform to these standards, contact the Global Support and Acquisitions Division for advice.

Table 1—File Format Standards for Digitizing Records

Type of Record	File Formats	Filename Extension
Photographs and images of physical records	• Lossless JPEG 2000 (preferred format) • Uncompressed (lossless) TIFF (Tagged Image File Format) (alternate preferred format) • JPEG with the highest quality setting possible	.jp2 or .j2k .tif .jpg

For collected records that were created with a software application such as those shown in table 2 below, do not attempt to convert file formats to the standards given in table 1. Instead, contact the Global Support and Acquisitions Division for advice before you send the records.

Table 2—File Format Standards for Digitizing or Creating Digital Records

| Type of Record	Application	Filename Extension
Electronic Document	• Notepad or text editor • Microsoft Word • WordPerfect • OpenOffice • Adobe Acrobat	.txt .dot, .doc or .docx .wpd .odt .pdf
Electronic Spreadsheets	• Microsoft Excel • OpenOffice Calc	. xls or .xlsx .ods
Presentations	• Microsoft PowerPoint • OpenOffice Impress	.ppt or .pptx .odp

When you are digitizing a physical record, it is important to ensure that the record is not damaged during the digitization process. The following information gives the guidelines for properly digitizing a record according to the conditions of the document.
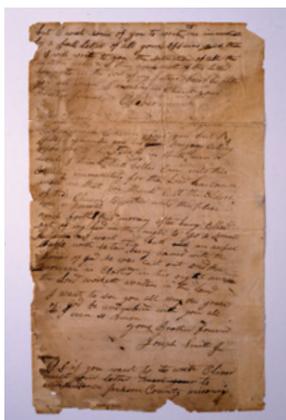
Record Types and Conditions
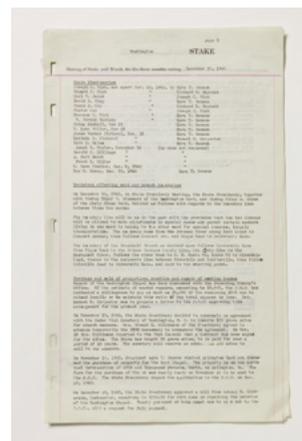
Example of a fragile, bound record

Example of a nonfragile, bound record
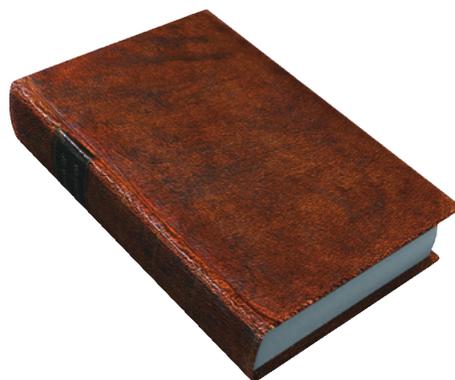
Example of a fragile, loose record

Example of a nonfragile, loose record

Determining Contrast Level

The word "contrast" describes the visible difference between characters and the background of a record. A record with high contrast has characters that are easily seen; a record with low contrast has characters that appear to blend with the background. You may need to adjust your scanner settings to help make medium or low contrast records easier to see in the resulting digital file.

High Contrast

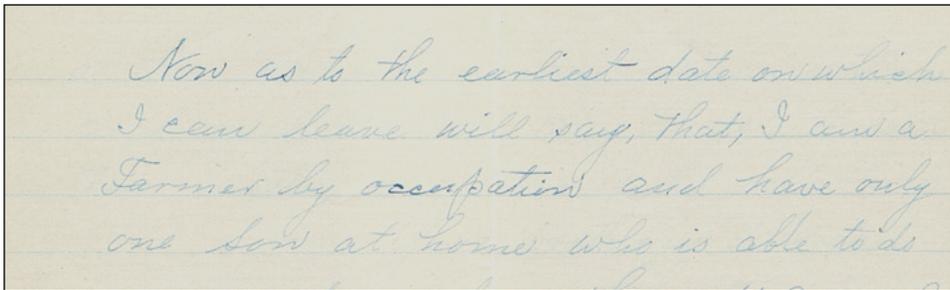*Medium Contrast*

Low Contrast

Scanner Settings

Digitizing Device Settings

Record Type	Contrast Level	Color Setting	Minimum DPI (Dots per Inch) Setting
Document	Medium or high	Grayscale or color	300
Document	Low	Grayscale or color	400
Black and white photograph	Not applicable	Grayscale or color	600
Color photograph	Not applicable	Color	600

Quality Standards

After capturing a record, view the resulting image to ensure that it meets quality standards as explained below. If quality is not sufficient, make the necessary changes, redigitize the record, and view the resulting image again to ensure its quality.

Focus

"Focus" is the quality of sharpness of an image. The top figure below is in focus (maximum sharpness), while the bottom figure is out of focus. If you attempt to scan a severely damaged document (folded, crunched, water damage, and so forth) with a flatbed scanner without closing the lid, the scanned image may be out of focus. Likewise, attempting to scan bound material with a flatbed scanner may also result in images that are out of focus (the area near the gutter will generally be out-of-focus).

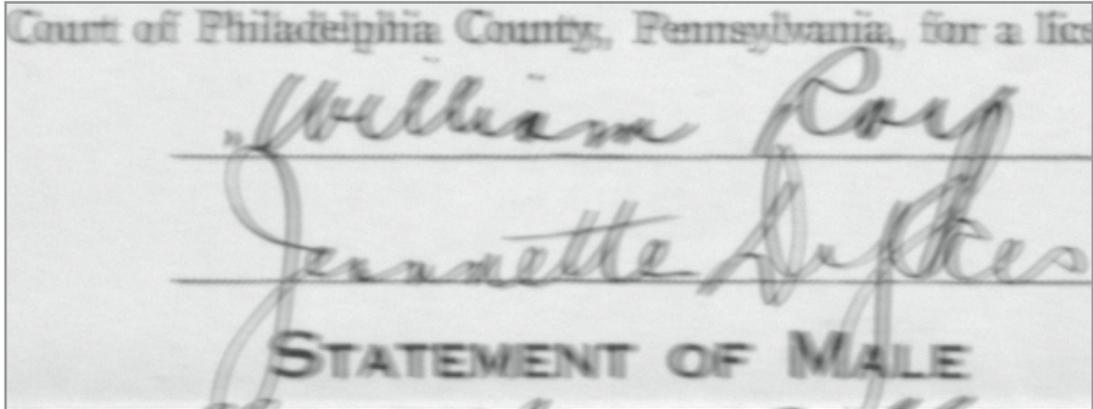*Example of a document scanned with proper focus*

Example of a document scanned with improper focus

Blur (applies mainly to digital camera capture)

Blurring results from a record or a camera being in motion during the capture process. Blurring is generally unidirectional in nature.

Example of an unacceptable blurred record that shows a double image

Exposure (applies only to digital camera capture)

"Exposure" is the total amount of light allowed to fall on the camera's image sensor during the process of capturing a record. An underexposed captured record is too black; an overexposed captured document is too white, as illustrated below.

A properly exposed record (left), and an overexposed record (right)
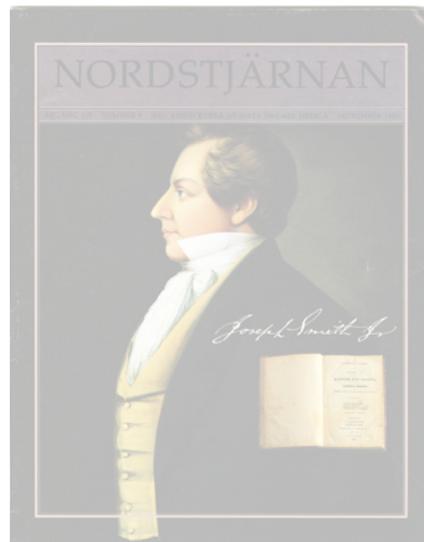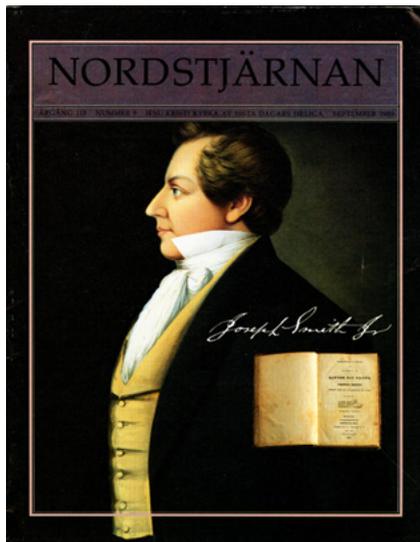
Cropping

Scan the record so that all edges are visible. Leave a sizable margin on all sides. When a record is completely scanned, all information from the original record will be present and readable or viewable. Margins that are too big should be cropped (or removed) from the resulting digital image, as illustrated below. If a record is too large to scan, contact the Global Support and Acquisitions Division for more information.
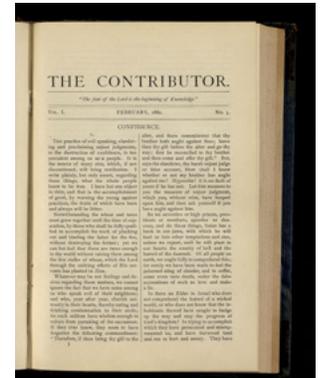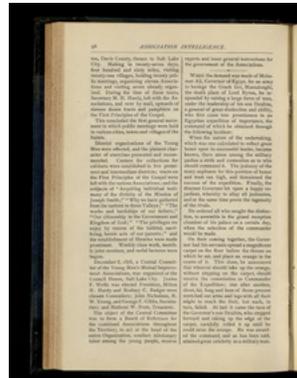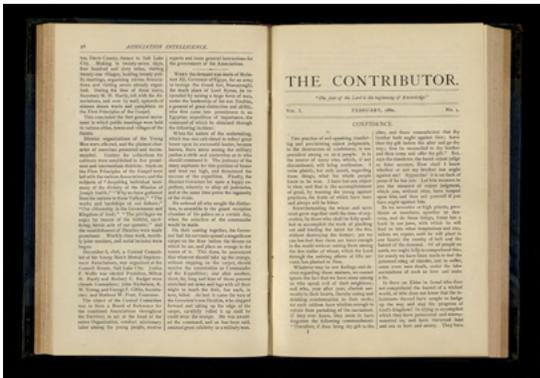
Proper cropping

Improper cropping

Examples of properly cropped bound records

Skew

Record image skew exists when a record is crooked. The image below on the left is correct and the image on the right needs to be rescanned.
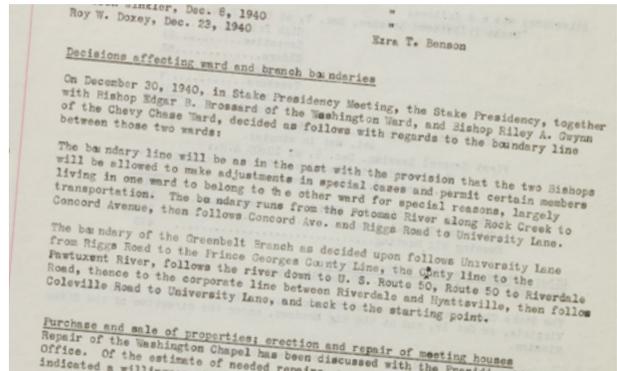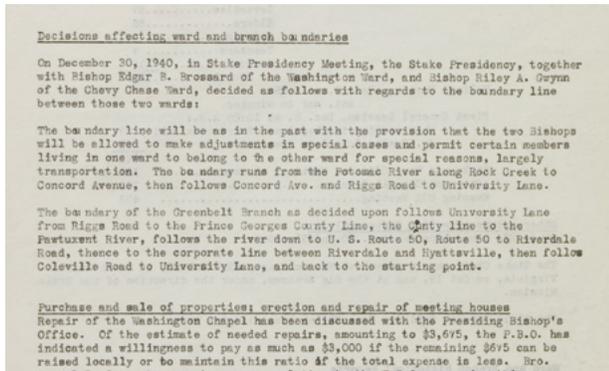
Complete Capture

There should be no covered or lost information in the final image. When a record is larger than the scanner, please contact the Global Support and Acquisitions Division for advice.

When you digitize a record that consists of more than one page or photograph, scan each page or photograph as a separate file. For example, a journal with 300 pages should be scanned as 300 individual lossless JPEG 2000 or uncompressed TIFF files. Likewise, an album of 76 photographs should be scanned as 152 individual lossless JPEG 2000 or uncompressed TIFF files (scan both sides of each photograph).

The examples of "incomplete capture" show where information is being lost because of a fold in the page, or covered because of a post-it note.

Complete capture   *Improper capture*
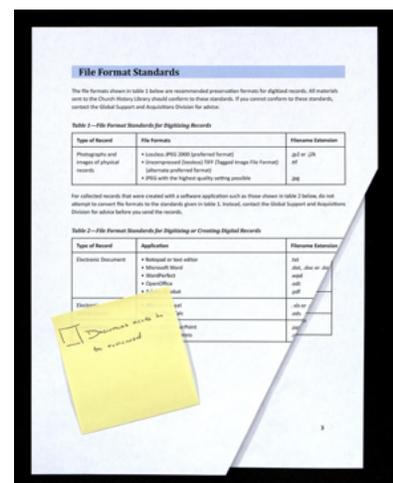
Quality Standards Checklist

The checklist below is to help you make sure that you consistently meet quality standards. After capturing a record, view the resulting image to ensure it meets all quality standards in the checklist. Even if only one checklist item does not meet the standard, redigitize the record, and view the resulting image again to ensure its quality.

Quality Standard	Meets Standard	Does Not Meet Standard
No covered or lost information		
Focus		
Blur		
Exposure		
Cropping		
Skew		

See the end of this document for a reproducible checklist.

Naming and Arranging Digital Files

Records That Are Stored in a Records Preservation Center

When you digitize a collection, it is important to organize the electronic folders in a way that mirrors the physical organization of the collection in the records preservation center. For example, if a journal is stored in a single folder, the images created from scanning the journal should be placed into a single electronic folder. The electronic folder represents the physical journal. Follow the file-naming standard described in the next few pages so files coming in from all parts of the world will not get mixed up at the Church History Library. Filling out the cataloging worksheet before you begin scanning will help you keep the collection organized and the files named correctly. See the following fictional examples from the country of Deseret.

Example 1

Juan wants to scan DS-02-00022, which is a collection that contains a 400-page loose-leaf journal and 27 photographs. This collection has been physically stored in two boxes. Box 1 holds the journal, which is divided into three folders. Box 2 holds the photographs in a single folder. To organize the scans, Juan creates a digital folder structure like the one below. Note that the filenames correspond with the digital folder structure.

Physical

FOLDER 1
FOLDER 2
FOLDER 3

FOLDER 1

Digital

DS-02-00022
- b0001
  - f0001
    - DS-02-00022_b0001_f0001_00001.tif
    - DS-02-00022_b0001_f0001_00002.tif
    - DS-02-00022_b0001_f0001_00140.tif
  - f0002
    - DS-02-00022_b0001_f0002_00001.tif
    - DS-02-00022_b0001_f0002_00002.tif
    - DS-02-00022_b0001_f0002_00185.tif
  - f0003
    - DS-02-00022_b0001_f0003_00001.tif
    - DS-02-00022_b0001_f0003_00002.tif
    - DS-02-00022_b0001_f0003_00075.tif
- b0002
  - f0001
    - DS-02-00022_b0002_f0001_00001.tif
    - DS-02-00022_b0002_f0001_00002.tif
    - DS-02-00022_b0002_f0001_00027.tif

| Call Number | Box Number | Folder Number | Scan Number |

Example 2

Hans has placed five small collections (DS-11-00229 through DS-11-00233) in a single box. When Hans scans DS-11-00231, since the collection does not span multiple boxes, he does not need to include the box number in the file name. He simply uses the call number and folder number.

The digital file name of the 55th scan of folder 1 is:

"DS-11-00231_f0001_00055.tif"

The digital folder structure looks like this:
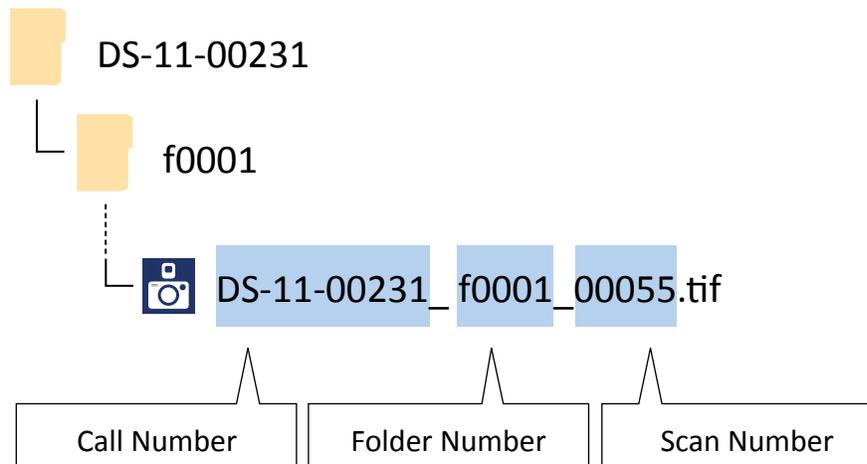
Example 3

Xiao Ming has cataloged a collection of 12 book volumes (DS-03-05271) and has placed them in 4 boxes. Since books are not placed in folders, he replaces "f0001" in the naming structure with "v0001." The letter "v" stands for "volume."

The digital file name of scan 1,154 from volume 10 is:

"DS-03-05271_b0003_v0010_01154.tif"

Volume 10 of DS-11-00231 is being digitized and is represented in the digital folder structure as follows:
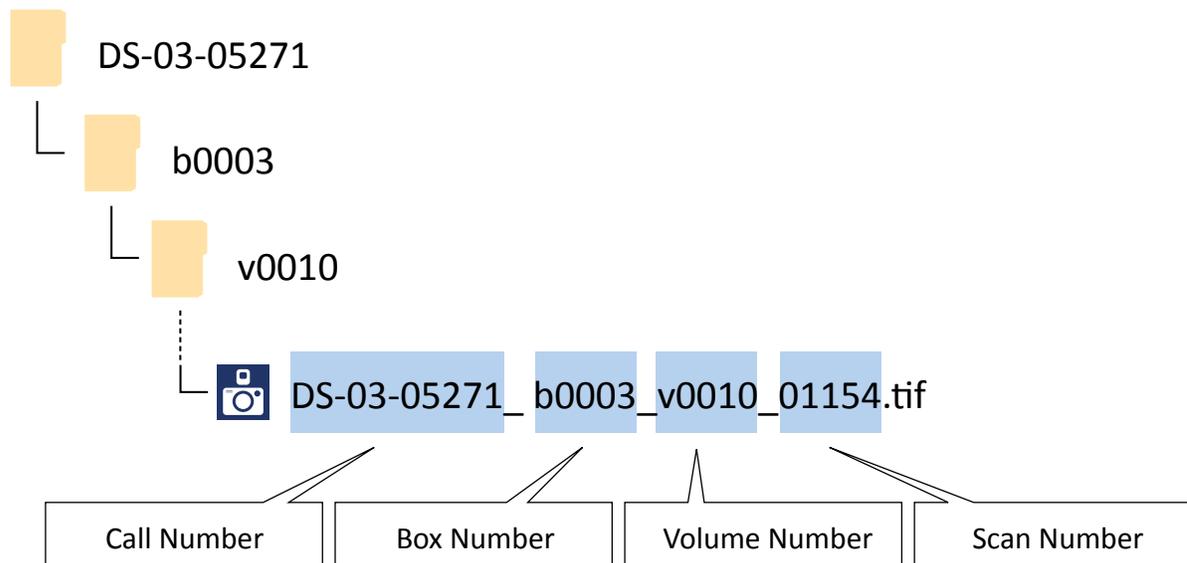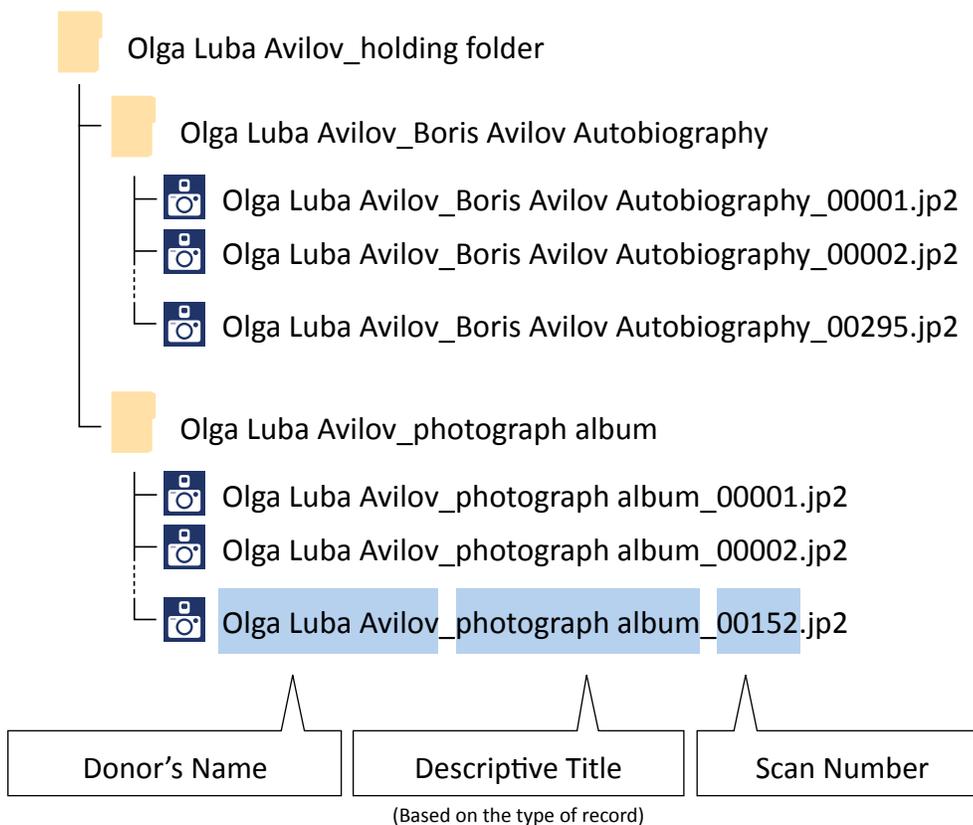
Records that Do Not Originate in a Records Preservation Center

When historical documents are scanned in a location outside a records preservation center, a temporary file naming structure must be used since the collection hasn't been assigned a call number yet. To name the images, instead of using the call number, as shown in the examples above, use the donor's name. Instead of using box and folder numbers, use descriptive titles that reflect the organization of the collection. After you send the files to the Church History Library in Salt Lake City, Utah, USA, the temporary file names you created will be replaced with new file names that include the assigned call number.

Example 1

Olga Luba Alvilov has donated the autobiography of her father, Boris Avilov. She has also donated a photograph album. Following the instructions contained in *Church History Guide: Collecting Records*, Fritz, the Church history adviser, creates a holding folder in which all the digital images will be placed. Fritz then creates a naming structure, using the donor's name and descriptive titles. He uses these same descriptive titles when filling out the cataloging worksheet (see example below) so it is easy to match the files with the cataloging worksheet.

Olga Luba Avilov_holding folder

    Olga Luba Avilov_Boris Avilov Autobiography

        Olga Luba Avilov_Boris Avilov Autobiography_00001.jp2

        Olga Luba Avilov_Boris Avilov Autobiography_00002.jp2

        Olga Luba Avilov_Boris Avilov Autobiography_00295.jp2

    Olga Luba Avilov_photograph album

        Olga Luba Avilov_photograph album_00001.jp2

        Olga Luba Avilov_photograph album_00002.jp2

        Olga Luba Avilov_photograph album_00152.jp2

Donor's Name	Descriptive Title	Scan Number

(Based on the type of record)

(Note: As explained in the "Digitizing Guidelines" section, when you scan a record that consists of more than one physical page, scan each physical page as a separate file. Also scan both sides of every photograph.)

Transferring Digital Records

As explained in *Church History Guides: Collecting Records*, all electronic (digital) records are to be transferred directly to the Church History Library in Salt Lake City, Utah, USA. Sending them electronically is the preferred method, but if you experience persistent connection problems over the Internet, you may mail them on an external hard drive (large collections), or a USB flash drive, CD, or DVD (smaller collections). Be sure the contents are securely packaged before mailing them. Use the courier service approved by your area office to mail the digital records. The Church History Library will return the hard drive in its transfer case but will not return USB flash drives, CD's or DVDs.

Also, keep at least one copy of these records until you receive confirmation that they have been successfully preserved in the Digital Records Preservation System. Once you receive confirmation, you may delete your copy (or copies).

Quality Standards Checklist

Feel free to print this page as needed.

Item or collection:

Date digitized:

Name of digitizer:

Quality Standard	Meets Standard	Does Not Meet Standard
No covered or lost information		
Focus		
Blur		
Exposure		
Cropping		
Skew		

Item or collection:

Date digitized:

Name of digitizer:

Quality Standard	Meets Standard	Does Not Meet Standard
No covered or lost information		
Focus		
Blur		
Exposure		
Cropping		
Skew		

Item or collection:

Date digitized:

Name of digitizer:

Quality Standard	Meets Standard	Does Not Meet Standard
No covered or lost information		
Focus		
Blur		
Exposure		
Cropping		
Skew		