

## Workflow overview for text files

This project requires the creation of at least three (usually four) versions of each file; these are saved in separate folders so they can be easily identified.

**Please see the Kansas State Historical Society standards for recommended resolution (ppi - pixels per inch) for scanning Archive files.**

The **Archive** version is the raw scanned image, saved in an uncompressed file format without any editing. Save as a .tif image file. This version is the one that is critical to include in a backup, as it is the closest to the original item.

The **OCR completed** folder contains working files from ABBYY FineReader, the program that uses Optical Character Recognition (OCR) to convert an image file into a full-text searchable file. You will use this version to double-check the OCR text, and make sure that it matches the text of your document. It is important to save each document in the FineReader Document format so that recognition will not need to be redone.

The **Upload** folder contains the PDF/A document that will be used for display online in Omeka. Once you have finished checking the OCR of your document in ABBYY FineReader, you will save it as a PDF/A document in your Upload folder.

The **Upload** folder will also contain a plain text file that is required to create keyword-searchable text in Omeka. The pdf document is internally searchable, but the derived text must be added to the Omeka item record for the content to be exposed to site searches.

The following table provides a quick comparison of the four versions.

<i>Folder</i>	<b>Archive</b>	<b>OCR completed</b>	<b>Upload</b>	<b>Upload</b>
<i>File type</i>	TIFF (.tif)	Fine Reader (folder)	PDF/A (.pdf)	Text (.txt)
<i>Resolution</i>	400-1500 dpi (see standards)	n/a	n/a	n/a
<i>File size</i>	Very large	Large (ex. 31 Files, 5 sub-folders)	Small	Very Small
<i>Use</i>	Original unedited scan	Working files, used for editing	Upload	Copy/paste text
<i>Example</i>	HPL_Letter_01.tif (41.8 MB)	HPL_Letter_01 folder (23.8 MB)	HPL_Letter_01.pdf (695 KB)	HPL_Letter_01.txt (5 KB)

