

Digitization Standards for BC Digital Commons*

In order to maintain consistency in the quality of digital work available balanced with accessibility in the BC Digital Commons, we have developed the following guidelines to aid in your content preparation.

These guidelines explain the minimum requirements for digitization standards, as well as recommended and acceptable file types. Whenever possible, all newly acquired materials should adhere to these minimum requirements, although higher standards may be desirable for some materials.

This document is not intended to address all technical issues related to digitization.

Our standards are based on best practices we have observed at other institutions. These standards are subject to change, given evolving digital formats and the future needs of the repository and our users.

If you need assistance converting your content from one file to another, please contact the IT.

*(*These standards are based on those developed at Linfield College.)*

Documents (typed or handwritten)

Any documents, whether typed or handwritten, that are not born-digital will need to be scanned. If OCR (optical character recognition) is desired to allow for full-text searching, then the document should be scanned and saved as a PDF. If OCR is not desired, the document can be scanned and saved as a JPEG. For handwritten documents, if OCR is used when creating the PDF, the text will be treated as an image (which means that full-text searching will not be possible). We recommend scanning handwritten documents in grayscale or color. If needed, backing paper may be used to enhance scan quality; refer to NARA's Technical Guidelines for Digitizing Archival Materials for Electronic Access (p. 42) for additional details, available at <http://www.archives.gov/preservation/technical/guidelines.pdf>.

Web Delivery	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Text (Scanned)	PDF/A JPEG	PDF PNG RTF docx	JPEG: 640x480 to 800x600 PNG: 640x480 to 800x600	PDF/A: 300 JPEG: 72 pixels per inch	PDF/A: minimal JPEG: medium	8-bit grayscale (for black and white) 24-bit color
Text (Born digital)	PDF/A	PDF RTF docx (Microsoft Word)				

Documents (typed or handwritten) continued

Archival Masters	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Text (Scanned)	TIFF	None	4000x2500 to 6000x4000	600 pixels per inch	lossless	8-bit grayscale (for black and white) 24-bit color
Text (Born digital)	XML+DTD (i.e., TEI) RTF ASCII/Unicode (TXT), using an editor that does not save to a proprietary format					

[DTD is document type definition which is a set of markup declarations that define a document type for an SGML-family markup language (SGML, XML, HTML). A Document Type Definition (DTD) defines the legal building blocks of an XML document. It defines the document structure with a list of legal elements and attributes. (https://en.wikipedia.org/wiki/Document_type_definition) TEI is Text Encoding Initiative which is a consortium that collectively develops and maintains a standard for the representation of texts in digital form. (<http://www.tei-c.org/index.xml>) RTF stands for Rich Text Format. RTF is a text file format used by Microsoft products, such as Word and Office. RTF, or Rich Text Format, files were developed by Microsoft in 1987 for use in their products and for cross-platform document interchange. RTF is readable by most word processors. (https://en.wikipedia.org/wiki/Rich_Text_Format)]

Photographs (Prints)

If the back side of the photograph contains markings or writing, note this in the metadata if significant to the provenance of the image. Make sure that the file name clearly indicates whether a file is the front side or the back side. See file naming in this document for additional information.

Web Delivery	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Photographs (prints)	JPEG	BMP GIF JPEG2000 PNG	640X480	72 pixels per inch	JPEG:medium	8-bit grayscale (for black and white) 24-bit color

Archival Masters	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Photographs (prints)	TIFF	none	4000x2500 to 6000x4000, excluding borders and mounts	600 pixels per inch	lossless	8-bit grayscale (for black and white) 24-bit color

Photographs (negatives or slides)

For more detailed information about specifications for digitizing transmissive originals (film, negatives, and slides), refer to the California Digital Library's Guidelines for Digital Images (pp. 6-8), available at http://www.cdlib.org/gateways/docs/cdl_dffr.pdf.

Web Delivery	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Photographs (prints)	JPEG	JPEG2000	800X600	150 pixels per inch	JPEG: medium	8-bit grayscale (for black and white) 24-bit color

Photographs (negatives or slides) continued

Archival Masters	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Photographs (prints)	TIFF	none	4000 pixels across the long dimension of the image area, excluding borders and mounts	3200 pixels per inch	lossless	8-bit grayscale (for black and white) 24-bit color

Large format posters or maps

Items larger than 17 by 24 inches should be digitized according to these specifications.

Web Delivery	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Photographs (prints)	JPEG	BMP GIF JPEG2000 PNG	1078 pixels per inch	150 pixels per inch	JPEG: medium	8-bit grayscale (for black and white) 24-bit color

Archival Masters	Recommended Formats	Other Acceptable Formats	Spatial Dimensions	Spatial (Capture) Resolution	Compression Level	Bit Depth
Photographs (prints)	TIFF	none	6000-8000 pixels across the longest side	600 pixels per inch (300-400 pixels per inch for larger items)	lossless	8-bit grayscale (for black and white) 24-bit color

Standards – Multimedia Audio

Audio settings will vary based on the file type. To reduce file size further, you may use mono rather than stereo tracks.

Web Delivery	Recommended Formats	Other Acceptable Formats	Bit Depth	Sample Rate
Audio (voice)	MP3	AIFF M4A WAV WMA	28-12Kbps	11 KHz
Audio (music)	MP3	AIFF M4A WAV WMA	128-224 Kbps	22-44.1KHz

Archival Masters	Recommended Formats	Other Acceptable Formats	Bit Depth	Sample Rate
Audio (voice)	AIFF WAV	NONE	8-16-bbit mono	22 KHz
Audio (music)	AIFF WAV	NONE	16-24-bit stereo	44.1-48 KHz

Video

Pixel size should generally be kept between 160x120 and 320x240, although higher sizes are acceptable (640x480 to 720x480).

Web Delivery	Recommended Formats	Other Acceptable Formats	Spatial Dimensions (pixels)
Video	H.264	MOV MPEG-1 MPEG-4	160x120 to 320x240

Archival Masters	Recommended Formats	Other Acceptable Formats	Spatial Dimensions (pixels)
Video	Raw video file MPEG-2 DVD-quality video file	none	720x480 or higher

Additional Considerations

Image Adjustment: Images should be rotated as needed post-scanning for ease of viewing. Depending on the project, cropping, color and contrast adjustments, sharpening, and other post-scanning adjustments may also be necessary. IT provides assistance with digital image editing.

Watermarks

Visible watermarks are not recommended for archival masters or web delivery access copies. We recommend using electronic or "invisible" watermarks to provide proprietary rights information as an alternative means to protect digital intellectual property.

File Naming

File names should be unique, consistently structured, and should use numbers and/or lowercase letters. Use underscores rather than spaces in file names, or use CamelCase. Do not use symbols in file names. If using a numerical scheme, use leading zeros for ease of sorting, and take into account the maximum number of items to be scanned so that you use an appropriate number of digits. Examples:
pumpkin_patch_back.jpg 0004_2010_Smith.jpg