All content types

Still Image

Resources predominantly visual in nature, such as two-dimensional visual arts (drawing, painting, graphics, photographic prints and negatives, plans), manuscript books and archival documents, early printed books, or books printed in alphabets or fonts for which Optical Character Recognition (OCR) is cannot be implemented.

Minimum Specifications

	Master	Processed	Access	OCR source	Text	Thumbnail	Print source??
File type	TIFF	TIFF	JPEG	n/a	n/a	JPEG	n/a
Resolution	400 PPI*						
Pixel array	4000 pixels on the long side						
Bit depth	24-bit RGB**						
Color space	eciRGBv2 or Adobe RGB 1998						
Retention	Permanent	Permanent	Create on ingest			Create on ingest	
Ratio	1:1						

* The resolution should be calculated from the dimensions of the object.

** Whenever possible, capture at 48-bit RGB (for color, monochrome, objects with stains or marks) and save at 24–bit RGB.

Textual

Resources that are expressed through a form of notation intended to be read, especially printed documents (books, serials, pamphlets, posters, broadsides, etc.) for which OCR (Optical Character Recognition)-extracted text should be provided as an output of the digitization workflow.

Minimum Specifications

	Master	Processed	Access	OCR source	Text	Thumbnail	Print source
File type	TIFF	TIFF	JPEG	TIFF	UTF-8 compliant .txt file	JPEG	PDF*
Resolution	300 PPI**				n/a	72 PPI	n/a
Bit depth	8-bit grayscale***			Bitonal, adjusted for			n/a

				contrast and brightness			
Retention	Permanent	Permanent	Create on ingest	Temporary	Permanent	Create on ingest	
Ratio	1:1						

* The maximum size for PDF files shlould not exceed 100 MB.

** The resolution should be calculated from the dimensions of the object and the size of the text. For large text, 10pt or higher, 300 PPI should suffice. A document with smaller text, 9pt or lower, may require 400-600 PPI.

*** Whenever possible, capture at 48-bit RGB (for color, monochrome, objects with stains or marks) and save at 24–bit RGB. Or capture at 16-bit grayscale (for most objects where color is not a concern) and save at 8-bit grayscale. Use 1-bit bitonal for clean, high-contrast documents with printed type only.

Audiovisual

Audiovisual materials at Yale University Library include analog audio and video formats, digital audio and video formats, and film. Below is a chart showing AV preservation format recommendations YUL is considering adopting in accordance with a consulting report submitted by Audiovisual Preservations Solutions in June 2013. Two competing standards are currently in use at YUL: (1) uncompressed 10-bit video, which results in large files, and (2) mxf-wrapper motion JPEG 2000, which is not 100% open source.

Minimum Specifications

Format	Preservation Master	Access Master	Access Copy
Audio-Analog	- Broadcast Wav File (BWF) wrapper - PCM uncompressed - 24-bit - 96kHz	n/a	- MPEG Audio Layer 3 (MP3) - Bitrate 256Kbps
Audio-Digital	- Broadcast Wav File (BWF) wrapper - Native uncompressed data at original sample rate and bit-depth	n/a	- MPEG Audio Layer 3 (MP3) - Bitrate 256Kbps
Video - Analog Standard Definition	- Quicktime wrapper (.mov extension) - Video encoded as 10-bit YUV 4:2:2 uncompressed (v210) - Audio encoded as uncompressed PCM, 48kHz - Maintain original aspect ratio, recording standard, interlacing, number of audio channels, and auxiliary information such	- Quicktime wrapper (.mov extension) - Video encoded as DV - Audio encoded as uncompressed PCM, 48kHz - Maintain original aspect ratio, recording standard, interlacing, number of audio channels, and auxiliary information such as original timecode and closed captioning	- MPEG4 wrapper (.mp4 extension) - Video encoded as H.264 - Audio encoded as uncompressed AAC, 44.1kHz, 256Kbps - Bitrate 5.0Mbps

	as original timecode and closed captioning		
Video-Digital	- Native encoding and data rate in Quicktime wrapper (.mov extension), e.g. DV for MiniDV and DVCam - Maintain original aspect ratio, recording standard, interlacing, number of audio channels, and auxiliary information such as original timecode and closed captioning	n/a	- MPEG4 wrapper (.mp4 extension) - Video encoded as H.264 - Audio encoded as AAC, 44.1kHz, 256kbps - Bitrate 5.0Mbps
Film	16 and 8 mm: - 2k 10-bit RGB 4:4:4 DPX log - Uncompressed 96kHz/ 24-bit Broadcast Wav File (BWF) for audio 35 mm: - 4k 10-bit RGB 4:4:4 DPX log - Uncompressed 96kHz/ 24-bit Broadcast Wav File (BWF) for audio	- MXF wrapper - AVC-Ultra	- MPEG4 wrapper (.mp4 extension) - Video encoded as H.264 - Audio encoded as uncompressed AAC, 44.1kHz, 256kbps - Bitrate of 5.0mbps

3D Objects

Objects, artifacts, and three-dimensional works of visual art encountered in archives, galleries, and museums (medals and badges, physical evidence from legal archives, some works of art).

Although the Yale Library collections include three-dimensional objects (such as artists' books in the Arts Library Special Collections, or globes in the Map Collections), guidelines for this format type are to be developed at a later time.

Born Digital

Born-digital materials arrive at Yale University Library from a variety of sources. Curators at different libraries acquire data tapes, floppy disks, CD-ROMs or hard drives from donors. Staff in different Yale offices transfer born-digital records to University Archives. Researchers work with staff at Marx Science and Social Science Library to preserve their datasets. The file formats and genre of these materials vary widely.

Born-digital objects should be preserved in their native file formats whenever possible. In addition, different preservation master copies and/or access copies may be created. The format of these copies will vary depending on the needs of staff and researchers, the library's digital preservation policy, and available resources. Guidelines about preferred file formats may be developed at a later date.

Minimum Specifications

Sample content	Sample formats
Text	.doc, .wpd, .txt, .rtf, .pdf, .xls
Notated music	
Visual	.pdf, .jpg, .tif, JPEG2000
Visual - Cartographic	.tif, JPEG2000, GeoTIFF, .drg
Visual - Plans and diagrams	.dxf, .dwg
Digital audio	.wav, .mp3
Digital video	.mov, mp4
Three-dimensional	
Database	.fp, .dat, .mdb, .mdf, .nsf
Software	
Datasets	
Email	.mbox, .pst, .eml, .msg
Websites	
Video games	