

Imaging Standards



Based on recommendations and standards established by the Federal Agencies Digitization Guidelines Initiative (2010) and Cornell University (2003).

File formats

- Items are always captured as .tiff files.
- Items that are fragile or present significant imaging challenges are preserved as .tiff files.

Bit depth

- Items that are text based (printed), have no color content, or are photocopies are captured in grayscale (8bit).
- Items that have color content or are photographic prints are captured in color (24 bit).
- Items that are fragile or present significant imaging challenges are captured in color.
- Photographic negatives are captured in color and convert to grayscale during QC.

Color space (ICC profiles)

The ICC profile for color images should be Adobe RGB 1998.

Resolution

Format	Size	Resolution
Rare materials- books and manuscripts	All sizes	300-400 ppi
General- books and manuscripts	All sizes	300 ppi
Oversize	19" x 22"	150-300 ppi
Prints and photographs	All sizes	600 ppi
Photographic transparencies and negatives	35mm to 4" x 5"	1200-3000 ppi
Photographic transparencies and negatives	Larger than 4" x 5"	800-1200 ppi
Paintings and 2D art (other than prints)	All sizes	Up to 8500 pixels on longest side
Newspapers	All sizes	300-400 ppi

Basic imaging workflow

1. Create images on capture device.
2. QC raw scans. Check for completeness, overall image quality, ability to meet image standards, and digital or physical artifacts. Refer to QC manual for more detailed instructions.
3. Complete re-scans if necessary.
4. Complete final organization and reconciling with finding aid when applicable.
5. Send copies to UK's backup system.
6. Files are prepared in the appropriate submission information package (SIP) structure for ingest into the digital preservation repository.
7. Automated process of creating archival information package (AIP) and dissemination information package (DIP) is launched, resulting in ingest of materials to ExploreUK.
8. Online content is QCd, spot checking for functionality and correct metadata mapping.

Digitization statistics

Manuscripts

- 30-40 hours to digitize and complete quality control (QC) of 1 cubic foot of content (approximately 2,000 pages)
- 1 technician can digitize and QC 40-50 cubic feet in 1 year (4 weeks of vacation/sick time included, 7.5 hour working days)

Photographic prints

- 0.66-1 minute to image, transport, and QC 1 photographic print
- 1 technician can scan and QC 60-90 photographic prints per hour (assuming cropping and straightening can be automated)

Digitization Cost Calculator (by the Digital Library Federation) <http://dashboard.diglib.org/>

Best practices

(Every rule has its exception.)

- Original material is scanned whenever possible.
- Reproductions are only scanned when a) it is the only copy or b) has content not included in the original (e.g. notations or if the original's quality has diminished).
- When a collection includes *exact* multiple copies of an item, the best version is selected, meaning the copy that is most complete, best condition, or greatest amount of content (including handwritten notes).
- Both items are captured if two versions differ.
- One item per image. Multiple items are not captured in one scan.
- All edges of an item are captured, with approximately ¼ inch borders.
- Materials containing folds are generally left as-is when scanning for access purposes.
- Bound handwritten pages are not split into separate images. However, bound pages of typed items are split for OCR purposes.
- Images are rotated so that the majority of content is readable.