

Scanning Standards

Joe Arthur, Eve Young and Debbie Emerson
Information Management, Information Services

Objective

To provide standards for scanning text based documents that are to be used for access by students and staff of the University of Melbourne. These standards do not include recommendations for cultural preservation of those documents.

Essential criteria for the resulting scanned document include:

- accessible format;
- good viewing and print quality;
- searchable document; and
- file size as small as possible to ensure easy downloading.

Recommendations

Document specifications

PDF format – PDF is widely used on campus and around the world and can be easily created with many scanning softwares. PDF files are viewable through a free reader (Adobe Acrobat Reader) and can be converted into other format files if required (e.g. TIFF or Word).

300-400 DPI – The resolution used will depend on the quality of the original document, with some poorer quality documents requiring the higher resolution. The higher the resolution used to scan, the larger the resulting file and the longer it will take to download by those accessing it.

Bitonal (black and white scanning) – Bitonal scanning will create the smallest possible file, while retaining high quality. The majority of documents will not require grey scale or colour reproduction, but where this is needed to ensure meaning is not lost, it could be used for individual pages such as colour charts.

Searchable document – Optical Character Recognition (OCR) tools can be applied to a document to make it searchable. This is a benefit to everyone, but especially people with disabilities as text to speech software is able to interpret the OCR enhanced file. Creating OCR enable documents is not difficult: Adobe Acrobat Professional 7 is able to create a searchable PDF with a single command.

Before scanning

Before scanning gently wipe the plate of the scanner with a soft cloth to remove any marks, as these will otherwise be transferred to every page of the scanned document. Check the original document and erase pencil and editing marks as required.

After scanning

Scanned documents should be checked to ensure they are complete with no missing or partial pages. Also check for correct page orientation, and that no pages are skewed as this reduces readability and affect the accuracy of the OCR.

Scanning Standards – draft 1

Use the cleanup features of scanning or editing software such as masking and cropping to remove grey gutters, plus despeckle to remove spots other marks.

Oversized originals

Most scanners can scan documents up to A4 in size. Documents that are larger than A4 may be able to be scanned by Special Collections. Please contact Pam Pryde, Curator Special Collections on ppryde@unimelb.edu.au to discuss individual requests.

Training options

Staff at the PBCLC are available to provide scanning assistance including a quick tutorial if you do your scanning at the PBCLC, 1st floor of the Baillieu Library. It may be necessary to book a computer in advance during peak periods.

Please note: Each scanner comes with slightly different software (drivers), so although staff can give you an introduction to scanning at computers in the PBCLC, the software installed on your local computer may be different.