The National Archives

Digital Records: Digitised Image Specification

Draft/Final	**Final**
Date	**4 April 2011**
Author	**Ross Spencer**
Reviewers	**Jess Ahmon, Natalie Bowbrick, Anna Buelow,**
	Andrew Fetherston, Tim Gollins, Alex Green,
	Stephen Jones, Matt Palmer, Gary Thorpe
Sign off	**Tim Gollins 14/12/2010**

Information Policy Team
The National Archives
Kew, Richmond
Surrey
TW9 4DU
Email: psi@nationalarchives.gsi.gov.uk

Any enquiries regarding the content of this document should be sent to pronom@nationalarchives.gsi.gov.uk

Contents

1. Introduction

These guidelines for the digitised image have been developed to provide a minimum specification to be used in digitisation projects undertaken or controlled by The National Archives. These guidelines shall be applied to all digitisation projects undertaken by the National Archives henceforth. It should be noted that this specification is only mandated for projects where there is potential for the disposal of the original paper version of a record. Guidance for surrogate images is available on request from The National Archives.

We have outlined an image specification for the most common types of paper record we may digitise but this list is not exhaustive and under other circumstances we may require a higher standard depending on the preservation and archival requirements of any particular project. If the particular record type being digitised is not explicitly covered or if there is any doubt as to which specification to follow, clarification should be sought to confirm the applicable standard.

This negotiation will form part of the standard contractual process for all digitisation that is carried out and will also be included in the standard start up activities of any in house digitisation projects so we will be completely transparent in how we approach this decision at the time.

Retrospective application of the specification: We are keen to highlight that the standards outlined in this document only apply to digitisation projects undertaken by The National Archives from the publication date of this paper. Projects currently being negotiated or undertaken will not be required to adopt these standards mid-process.

1.1. Who is this guidance is for?

Recognising the requirement for formal image specifications to ensure long term preservation of digitised records, this document is an internal and external standard for digitisation projects controlled by The National Archives. We recommend that it is also used by those government departments and organisations embarking on digitisation projects to provide a reference baseline from which to develop their own standards.

Where the ultimate destination of the digested form of record is to be accessioned as a Public Record by The National Archives any deviation from this standard should be discussed with The National Archives in advance of the project commencing digitisation work. Failure to comply with the specification outlined, (outside of any exemptions) will result The National Archives being unable to accession the resulting digitised record.

2. Guidance for digitisation projects undertaken or controlled by The National Archives

2.1. Scope

The full content of this specification applies to the digitisation of original paper records from where the digitised version will constitute the accessioned public record. When a digital copy serves only as a surrogate other less stringent standards may apply. Please contact The National Archives for further guidance on surrogate digital images.

Existing digital copies: If an organisation has digitised original paper records to use in a digitised form as part of normal business processes then providing the digital form meets digital accession rules this specification does not apply.

2.2. Image Quality Assurance

Images shall be tested for compliance with the documented specification. All files shall resolve to the expected format identifier through the use of DROID (Digital Record Object Identifier)[1]. All Files shall also validate successfully against the modules available in the JHOVE validation utility[2].

Spot checks shall be undertaken to ensure the following remain consistent across the full digitised output of any project:

- Correct metadata and file naming conventions
- Size and positioning of images
- Acceptable borders and margins on images
- Image completeness and cropping of images
- Files are created such that each file represents one single page
- File format including format extension, bit-depth, image type and resolution

[1] http://sourceforge.net/projects/droid/ - Accessed 17th November 2010

[2] http://hul.harvard.edu/jhove/ - Accessed 17th November 2010

2.3. Metadata

For digitisation carried out on contract, metadata and structuring of operating system folders created as part of the digitisation workflow shall be agreed with The National Archives as part of the contract negotiation process; for any in house digitisation project such agreement shall also be included in the standard start up activities of that project. In all cases a clear decision shall be made as to whether metadata shall be supplied as a separate file to the digitised image or whether it is to be encoded directly within the image file as in the case of using JPEG2000 as the format for digitisation (see below).

3. Image Specifications

The baseline file format and image requirements are outlined below. First we list acceptable file formats for accession before setting out image requirements for each type of record (e.g. documents, photographs, negatives). Each record type has its own specification section which details:

- Required image resolution (pixels per inch, ppi)[3]
- Colour management
- Compression

Because this specification applies to original paper records that will be destroyed the specification only supports colour scanning so that the maximum level of detail in the record is retained. This is re-iterated in the image specifications outlined below.

3.1. Acceptable file formats for digitisation

Our primary goal is to achieve high quality digitised images suitable for preservation, regardless of the size of the project undertaken. We have outlined a minimum image specification for multiple different object types to achieve this. We require all future projects to follow this specification unless specific negotiation has taken place to exempt the project from this standard. Any compression applied to an image shall be lossless and shall be encoded in one of the file formats described below:

3.1.1. Joint Photographic Expert Group 2000 (JPEG2000, JP2)

JPEG2000 is The National Archives' preferred format for digitisation. The qualities that the format presents by being a ratified ISO/IEC open source standard (ISO/IEC 15444) and providing high levels of compression for many image types suggest that it is suitable economically and structurally for preservation over long periods of time. Our primary requirement of any JPEG2000 file accessioned as part of this process is that it uses lossless compression, that is, the image should be mathematically lossless from that captured by the digital imaging device. Please see Appendix 1 for a more detailed description of our required JPEG2000 profile.

[3] PPI has been adopted to replace the previously inaccurate usage of DPI in more common parlance. Capture devices will be expected to capture *(n)* pixels per inch of the original paper record.

3.1.2. Tagged Image File Format (TIFF, TIF)

If it isn't possible to output JPEG2000 from a project then it is acceptable to output Tagged Image Format (TIFF) as an alternative. Any TIFF output as part of the digitisation process should remain uncompressed and not adopt any of the compression techniques outlined in the TIFF 6.0 standard.

4. Image Specification

The specifications outlined below represent a *minimum* standard for the type of image listed.

Note: Special cases require a more involved negotiation with The National Archives to determine the appropriate standard for digitisation and thus help ensure its longevity over time. These cases include:

- Architectural Plans
- Blueprints
- Diagrams
- Maps
- Textiles
- Artwork
- Artistic Designs

4.1. High contrast documents with text or graphics for which colour is not present or not essential and any images are line art

Resolution: 300 ppi
Type of image: Colour
Bit-depth: 24-bit
Colour management: Embedded ICC Colour Profile
Compression: Lossless

4.2. Documents where colour is present and important or documents with low contrast (e.g. faded text, coloured background)

Resolution: 300 ppi
Type of image: Colour
Bit-depth: 24-bit
Colour management: Embedded ICC Colour Profile
Compression: Lossless

4.3. Black and white photographs

Resolution: 600 ppi
Type of image: Colour
Bit-depth: 24-bit
Colour management: Embedded ICC Colour Profile
Compression: Lossless

4.4. Colour photographs

Resolution: 600 ppi
Type of image: Colour
Bit-depth: 24-bit
Colour management: Embedded ICC Colour Profile
Compression: Lossless

4.5. Black and white photographic transparencies (Slides / Negatives)

Resolution: 2400 ppi
Type of image: colour
Bit-depth: 24-bit
Colour management: Embedded ICC Colour Profile
Compression: Lossless

4.6. Colour photographic transparencies (Slides / Negatives)

Resolution: 2400 ppi
Type of image: Colour
Bit-depth: 24-bit
Colour management: Embedded ICC Colour Profile
Compression: Lossless

Bibliography

Document management — Portable document format — Part 1: PDF 1.7. (2008). *PDF 32000-1:2008* . Adobe Systems Incorporated.

Information Technology - JPEG 2000 image coding system: Core coding system. (2002). *ITU-T Recommendation T.800* . International Telecommunication Union.

Memory of the Netherlands Project Office. (2003, December). *Guidelines and procedures for the execution of projects for Memory of the Netherlands.* Retrieved November 17, 2010, from http://www.geheugenvannederland.nl/hgvn/webroot/files/File/PDF/richtlijnen/Microsoft%20Word%20-%20Richtlijnen%20en%20procedures%20%204.0%20engels.pdf

Public Record Office Victoria. (2010, January 1). *Digitisation: Image Requirements.* Retrieved March 10, 2010, from Public Record Office Victoria (PROV): http://www.prov.vic.gov.au/publications/publns/1002s2.pdf

TIFF Revision 6.0. (1992). *TIFF 6.0 Specification* . Adobe Systems Incorporated.

Van der Knijff, J. (2010, October 12). *TIFF to JPEG2000: report on migration and characterisation experiments.* Retrieved December 6, 2010, from http://www.udfr.org/jp2kwiki/images/4/4f/Jp2kMigrationCharacterisationKBExternal.pdf

Appendix 1: JPEG2000 Profile

The full JPEG2000 specification we have adopted for digital records is described below. Please refer to the JPEG2000 standard, ISO/IEC 15444-1 for more information on JPEG2000 profiles. Terminology used here should be consistent with any software implementing the JPEG2000 standard.

Conversion Software: Supplier Dependent
File Format: JP2 Part 1, compliant with ISO/IEC 15444-1
Lossy or Lossless: Lossless
Typical Compression: Image dependant
Tiling: Not Used
Progression Order: By Quality
No. of decomposition levels: N/A
Number of quality layers: 7
Code Block Size: N/A
Transformation: 5-3 reversible filter
Precinct Size: N/A
Regions of Interest: No
Tile Length Markers: N/A
Colour Management: Restricted ICC Profile

Note: Should the Restricted ICC Profile Method provided by the JP2 file format be considered inadequate for the purpose of some digitization projects, the use of the JPEG2000 JPX file format compliant with ISO/IEC 15444-2 and using the JP2 Part 1 compression algorithm can be discussed as part of the contract negotiation or project start up process.

Adobe Photoshop CS4: Images created in Photoshop CS4 cannot be accessioned by The National Archives. The tool is only capable of producing images conformant with the JPX standard however incorrect structuring of the format means the format will validate only as JP2. This presents a preservation risk. Use of Photoshop will be monitored an acceptance of the format created by this tool will be possible once JPEG2000 output has been addressed. More information can be found in (Van der Knijff, 2010).

For any other Adobe products being used we ask that you check with The National Archives before completing any work so that we can ensure that this issue is not present.