dri

Digital Repository of Ireland
Taisclann Dhigiteach na hÉireann

6

Producing high-quality images for digital preservation

fact sheet

For more information, please contact:

The Digital Repository of Ireland
Royal Irish Academy
19 Dawson St,
Dublin 2

web. www.dri.ie
email. dri@ria.ie

Issues to consider before digitising:

• What materials should you choose to digitise? Consider document attributes, preservation considerations, intended use, research potential.
• What legislation affects your materials? (e.g. copyright legislation or Data Protection).
• What are the financial capabilities or restrictions of the organisation/project.
• Do staff have the appropriate skills for the project and what training will be required?
• What is the remit of the digitising organisation and are there any gaps in institutional policy?
• What sort of file naming conventions and metadata schema will be used?

What equipment should you use to digitise?

• Possible equipment may include flatbed scanners, book scanners, digital cameras (including lenses, light source, tripod/camera stand, software, book cradle). Hardware such as a dedicated computers, monitors, and hard drives may also be necessary.
•Whichever hardware and software you choose to purchase, please ensure that you have fully researched the capabilities and restrictions of all systems. Please www.dri.ie/resources for more information.
• It is important to choose a digitisation system that reflects the materials you aim to digitise e.g. think about the preservation/conservation needs of the materials, how robust the materials are

and if the system can adapt to the needs of the collection. E.g. a flatbed scanner may be good for single sheets of paper but could damage a bound volume of letters; a digital camera system may be more suitable for a bound volume.

Issues to consider for digitisation

• **File Types and File Sizes**: It is necessary to choose the right file types when capturing, processing and storing your digitised images. It is also important to think of the size of files and the type of storage available to the digitising institution. Two primary types of files may be created: a master archival image file and a lower resolution surrogate image file. A master archival image may include a colour chart and ruler for reference purposes, but this may be cropped out for the surrogate file. Surrogate image files can be generated directly from the master archival file using a standard imaging software package. The Repository can also generate surrogates for dissemination on ingest please see DRI resources for more information.

• **Resolution:** Spatial resolution determines the amount of granularity in an image in terms of the number of picture elements or pixels per unit measurement. The specifications of the scanner or camera system, image processing and the characteristics of the item that is being digitised will inform the spatial resolution of the final image. Resolution is measured in dots per inch (dpi) or pixels per inch (ppi). A minimum of 300dpi is recommended for digitised images, although this can vary depending on the item, the

system used and the policies of the digitising organisation.

• **Bit-depth** defines the maximum number of shades and/or colours in a digital image file. The common standard for grey scale and colour images is to use 8-bits of data per channel and this provides a maximum of 256 shades or steps per channel ranging from pure black to pure white. High-bit or 16-bits per channel images can have a greater number of shades; a maximum of over 65,000. The greater the bit-depth, the larger the image files may be so this must be considered.

• **Colour Management** is a very important issue in digitisation; the main challenges are ensuring that the original colours of the digitised items are rendered successfully in the digital image and also ensuring that the digital image looks the same across a range of different output devices: the same image may look different when printed or when viewed on different screens. Colour management issues include:

- **Colour mode:** Colour images consist of three or more grey scale channels that represent colour and brightness information. Common colour modes include RGB (red, green, blue), CMYK (cyan, magenta, yellow, black), and LAB (lightness, red-green, blue-yellow).

- **ICC - Colour management** provides a context for objective interpretation of the numeric values that make up the colour and brightness of a digital image. It helps to compensate for differences between devices in their ability to render or display these values. International Color Consortium (ICC)-based colour management is the most widely implemented approach.

- **Colour profile** are sets of numbers that describe a colour space e.g. Adobe 1998. Files can be saved with any ICC-compliant profile that describes an input device, output device or colour space; although it is best practice to adjust the colour and tone of an image to achieve an accurate rendition of the original in a common, well-described, standard colour space.

- **Colour calibration**: Monitors and equipment should be calibrated on a regular basis to ensure correct colour management. There are specific hardware and software to assist in this process.

• **Reference targets** can be used to assist with adjusting scanners and image files to achieve objectively "good images" in terms of tone and colour reproduction, and to support the functions of device characterization for more advanced colour management. Examples include Kodak Q 13 target or the Greytag Macbeth Digital ColorChecker SG. A colour chart may be included in the master archival image for reference purposes

• **Backing**: Solid colour backing material may be placed behind imaged items to provide a uniform field, and to help define the boundaries of the physical item. The backing should be coloured neutral white, neutral black, or neutral 18% grey.

• **Cropping and Skew:** Digitised images may be cropped using standard image editing software. It is important not to crop out any of the original item, leaving approximately 0.1cm beyond the boundaries of the physical object. A small colour reference ruler may be included in master archival images, which may be cropped out for surrogate access images. Also items should be correctly aligned in digital images. Standard image editing software may correct any misalignment.

• **Quality control:** Images should be visually checked at capture and quality control audits carried out on processed master and surrogate images.

• **Watermarking** is the placement of a transparent or similarly subtle identifier in the image file; these marks are intended to support image ownership, authenticity, and copyright management activities. A watermark may be added to both master and surrogate images, just the surrogate image, or neither; depending on the policies and needs of the digitising institution.

• **Technical Metadata** refers to information that describes attributes of the digital image and is embedded in the image at time of creation. This should be considered before the digitisation process takes place.

• **Workflow:** It is important to create a workflow that is generic enough to provide a level of standardisation for all types of digitisation planned but also flexible enough to work with other types of capture that might be necessary in the future. The aims and outcomes of the digitisation project should be considered to inform the creation of the workflow.