DIGITISATION FOR PRESERVATION AND ACCESS
A technical perspective
Definition of a digitisation project

The conversion of analog material (hardcopy materials) e.g. written manuscripts, printed material, photos, video clips, film, sound recordings, tapes, microfiche, glass negatives, maps, paintings to an electronic and easily accessible format
**Investment – is it worth it?**

- Investing in digital conversion only makes sense if institutions are prepared to provide long-term access to digital collections and have a sound project plan to start with.

- Scanning process 1/3 of actual digitisation cost.
Criteria for a digitisation project

- Selection criteria
- Preservation
- Standards
- Equipment
- Sustainability
- SWOT analysis
Preservation aspects

- Fragility of digital representations
- Preservation strategy vital
- The key to digital preservation is the establishment of a managed environment
- Searchable metadata and well documented technical aspects
Preservation methods

• Basic archival preservation
• Digitisation best practices and standards
• Future migration
• Technical standards are maintained and developed by international organizations, i.e. (ISO); (IEC); (IEEE); (ITU); (W3C)
LET'S JUST SEE HOW BROTHER DOMINIC IS GETTING ON WITH TRANSFERRING OUR ARCHIVES ON TO CD-ROM!
Digitisation the project

1. Project identification
2. Selection criteria
3. Dismantling of material and copyright clearance
4. Basic preservation
5. Scanning
6. Archival server
7. Conversion process
8. Web-ready (eBooks, images, journals, sound, video clips)
9. Metadata
10. UPSpace
11. Technical data
12. Final documentation
Project identification

- Is there a known potential audience for the materials that are planned to be digitised?
- Will digitisation increase access, functionality or intellectual control?
- Will digitising these materials fill a need that is currently unmet?
- Are the materials in the public domain or can proper rights be secured?
Why collaborate

- Scarce funding
- Richer digitised collections can be created in a collaborative project
- Increase in the perceived values of collections
- Fulfilling research and educational goals and objectives at the institutional level
- Collaboration allows each partner to contribute its strengths
- Background linking to each others collections without user knowing
Selection criteria

- Break down to administrative and operational issues
- Type of material in collection i.e. photos, docs
- Will the complete collection be digitised?
- Copyright clearance
- OCR necessary?
- What is the current condition of the collection?
- Could the project be done in phases
- W3, restricted or in-house availability
- Outsource or in-house
- More than one copy of the same source?
Preservation basic

- Take off any damaging material i.e. elastic bands, staples, clips
- Mend tears in hardcopy source
- Tuck in loose leaves of books
- Dust and clean the source i.e. book, slide, photo
- Get rid of mould with a clean brush
- Document what you did and what still need to be done to preserve original source
Copyright clearance

- Never -- ASSUME --
- If clearance is necessary start immediately
- Write courtesy letters if possible to stakeholders even if out of Copyright
- Keep track – be persistent – follow-up
- Document everything on database, dates, names and addresses
- Life + 50 years standard copyright length in South Africa
- Publishers Rights run from the 1st edition + 50 years
Scanning

- If done correctly, the digital image will represent the visual information of the original source.
- Resolution is not the only factor that will affect the quality of an image file:
  -- scanning conditions
  -- scanner type
  -- quality of source
  -- settings
  -- skill of the scanning operator
  -- quality of the final display image (derivative)
Scanning technical

- Archival scanning – resolution size
- Calibrate work station
- Scanning of books, slides, maps – different scenarios
- Saving in Tiff, Jpeg, Gif, WAV, MP3
- Sound – Audacity programme
- QUALITY CONTROL
- Document it
Digital Services and Development Unit

Image Quality Calculator

Image Dimensions

<table>
<thead>
<tr>
<th>Width</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.5</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Options: inches or centimeters

Notes:
You must supply the dimensions of the original document or object in order to calculate the size, resolution, and quality. If you plan to use a flatbed scanner, it is also important to measure the document to be certain that it will fit on the scanning bed. Most commercial flatbed scanners range in dimensions from 8.5" x 14" to 12" x 17".

Limitations: The formulas for the Image Quality Calculators are intended for text and manuscript scanning. The IQC can be used for other types of materials, but we do not recommend this practice.

Enter width and height information in inches using decimal values and/or fractions of inches, tenths, eighteenths and sixteenths only. For centimeters, use integers with up to four digits to the right of the decimal point.

Next Step

Digital Services and Development Unit
Granger Engineering Library
University of Illinois at Urbana-Champaign
1301 W. Springfield, Urbana, IL 61801 (MC-274)

Copyright © 2000 University of Illinois at Urbana-Champaign
Digital Services and Development Unit

Image Quality Calculator

Smallest Significant Character or Visual Element
(measured in millimeters)

Height: 3 mm

Notes:
The goal in identifying the smallest significant character or visual element is to find the smallest feature of the picture that must be resolved clearly in the digital image. For example, in a text document the legibility of a character that is smaller than 6-point type may be critical. In a handwritten manuscript, it may be important to identify the width of a pen stroke. Use a 10x scale loops to obtain this measurement in millimeters. There is an inverse relationship between the smallest significant feature and the resulting file size and resolution. If you choose a very small significant feature, the file size may be very large.

For millimeters, use integers with up to four digits to the right of the decimal point.

Digital Services and Development Unit
Grainger Engineering Library
University of Illinois at Urbana-Champaign
1301 W. Springfield, Urbana, IL, 61801 (MC-274)

Copyright © 2000 University of Illinois at Urbana-Champaign
Digital Services and Development Unit

Image Quality Calculator

Notes:

Quality Index is a standard developed by the micrographics industry that relates system resolution and text legibility (ANSI/AIIM MS23-1991). The values used here are rough predictors of image quality in the following manner: Low (QI = 3.6), Medium (QI = 5), and High (QI = 8). It is important to note that by choosing a high quality index value, the smallest significant character or visual feature will be easily discernible. Ultimately, there is a positive correlation between the quality index and both the resolution and file size. The higher the quality index, the higher the file size and resolution. However, high resolution and large file size do not necessarily warrant a high quality image. (See About the IQM)

Digital Services and Development Unit
Georges Engineering Library
University of Illinois at Urbana-Champaign
1301 W. Springfield, Urbana, IL 61801 (MC-274)

Copyright 2000 University of Illinois at Urbana-Champaign
Digital Services and Development Unit

Image Quality Calculator

Bit Depth

- Black & White (1-bit)
- Standard Grayscale (8-bit)
- Enhanced Grayscale (12-bit)
- Standard True Color (24-bit)
- Enhanced True Color (36-bit)

Notes:

Bit depth is the number of bits needed to represent one pixel in a digital image. In the case of a black and white image, a bit depth of one (1 bit) is necessary to represent black and white. Grayscale (8-bit) and color (24-bit) images possess higher bit depths. A high bit depth allows more tones to be represented in an image. File size increases with bit depth in a linear fashion.

Next Step  Start Over

Digital Services and Development Unit
Geraghty Engineering Library
University of Illinois at Urbana-Champaign
1301 W. Springfield, Urbana, IL, 61801 (MC-274)

Copyright © 2000 University of Illinois at Urbana-Champaign
Digital Services and Development Unit

Image Quality Calculator

**Input:**

- Image Width = 12.5 cm
- Image Height = 9.5 cm
- Quality Index = 5
- Smallest Significant Character = 3 mm
- Bit Depth = 24-bit

**Results:**

- Recommended Resolution = 64 dpi
- Approximate File Size = 0.211 Mb
- Horizontal Pixel Dimension = 237 Pixels
- Vertical Pixel Dimension = 312 Pixels

**Notes:**

The recommend resolution (measured in dots per inch or dpi) is the minimum resolution necessary to capture all of the detail you desire to capture in the digital image. If, after scanning the image at the recommended resolution, you still desire to see greater detail, we recommend that you re-calculate for using a smaller measurement for the smallest significant feature.

File size stated here is for an uncompressed image file.

The image size that has been calculated is an estimate that assumes you will be saving in an uncompressed TIFF format. Slight variations (+/-5%) in file should be anticipated due to the unlikelihood of scanning at precise document dimensions.

Digital Services and Development Unit
Grainger Engineering Library
University of Illinois at Urbana-Champaign
1301 W. Springfield, Urbana, IL, 61801 (MC-274)

Copyright © 2000 University of Illinois at Urbana-Champaign.
Conversion: why and how

- Keep archival copy untouched – separate
- Crop, de-scew, cleaning, colour management
- Size of archival image ±50 – 60 MB
- Size of normal photo on web 50 KB
- 24 sec average download time per image
Storage archival and other

• One “sacred” archival copy – not to be touched
• Preferable: Server with search engine (Lucerne)
• 1 x derivated copy in PDF or Jpeg to collection
• 1 x derivated copy PDF and Jpeg in digitisation office
• Gold dye CD-R best
• Test CDs/DVDs once a year for data loss
• Should be able to migrate if necessary
Web-ready

- Lower resolution images derived from master file for web-ready copies
- Resolution of Jpeg derivatives between 72 ppi and 150 ppi
- PDF documents for web display - truly portable across different platforms
- PDF retains all of the formatting and detail present in the original
- PDF can be compressed into smaller file sizes for the Web
Metadata the finding aid

• METADATA is the information needed for researchers to evaluate whether or not a visit to the repository will be necessary

• Librarians are now challenged to create eye appealing informational products, preferably with a Google-like box to provide searching capabilities across many formats of data

• Along with the plethora of digital assets comes the need to attach meaningful metadata to provide access to items in their collections
Technical metadata

• A most necessary and well thought through technical database has to be compiled and uploaded onto the web platform

• Use descriptive field in DSpace for the technical metadata
<table>
<thead>
<tr>
<th>Image</th>
<th>Date Scan</th>
<th>Size of Original Canvas (cmxcm)</th>
<th>dpi</th>
<th>Scanned Size (w x h)</th>
<th>Colour</th>
<th>%</th>
<th>Date Scan</th>
<th>Software</th>
<th>Format</th>
<th>Image Size (w x h)</th>
<th>Quality</th>
<th>Size in Pixels</th>
<th>Size in Megabytes</th>
<th>General Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>cl_001</td>
<td>16-Mar-07</td>
<td>355x255</td>
<td>600</td>
<td>650 x 480 pixels</td>
<td>100%</td>
<td>100%</td>
<td>24-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 528</td>
<td>128</td>
<td>560 x 334</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>cl_002</td>
<td>16-Mar-07</td>
<td>355x255</td>
<td>600</td>
<td>1040 x 700 pixels</td>
<td>100%</td>
<td>100%</td>
<td>24-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 540</td>
<td>128</td>
<td>560 x 384</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>cl_003</td>
<td>16-Mar-07</td>
<td>355x255</td>
<td>600</td>
<td>1040 x 700 pixels</td>
<td>100%</td>
<td>100%</td>
<td>24-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 540</td>
<td>128</td>
<td>560 x 384</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>cl_004</td>
<td>16-Mar-07</td>
<td>410x280</td>
<td>600</td>
<td>6224 x 4728 pixels</td>
<td>100%</td>
<td>100%</td>
<td>06-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 578</td>
<td>128</td>
<td>560 x 382</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>cl_005</td>
<td>16-Mar-07</td>
<td>355x255</td>
<td>600</td>
<td>6224 x 4728 pixels</td>
<td>100%</td>
<td>100%</td>
<td>06-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 578</td>
<td>128</td>
<td>560 x 382</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>cl_006</td>
<td>16-Mar-07</td>
<td>550x280</td>
<td>600</td>
<td>1800 x 1100 pixels</td>
<td>100%</td>
<td>100%</td>
<td>24-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 573</td>
<td>128</td>
<td>560 x 331</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>cl_007</td>
<td>16-Mar-07</td>
<td>410x410</td>
<td>600</td>
<td>9272 x 6048 pixels</td>
<td>100%</td>
<td>100%</td>
<td>06-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 573</td>
<td>128</td>
<td>560 x 331</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>cl_008</td>
<td>16-Mar-07</td>
<td>410x280</td>
<td>600</td>
<td>7782 x 10014 pixels</td>
<td>100%</td>
<td>100%</td>
<td>06-Mar-07</td>
<td>Adobe Photoshop</td>
<td>.jpg</td>
<td>840 x 573</td>
<td>128</td>
<td>560 x 331</td>
<td>9.7</td>
<td>90%</td>
</tr>
<tr>
<td>Image Code</td>
<td>Image Name</td>
<td>Format</td>
<td>Size in Pixels (w x h)</td>
<td>General Info (Dimensions w x h)</td>
<td>Ppi</td>
<td>Safe for web</td>
<td>Download Information</td>
<td>Adobe Acrobat Distiller 6</td>
<td>Watermark (Arial 12 Shadow, Italic)</td>
<td>Resolution</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>--------</td>
<td>------------------------</td>
<td>---------------------------------</td>
<td>-----</td>
<td>------------</td>
<td>--------------------</td>
<td>------------------------</td>
<td>--------------------------------</td>
<td>------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_001</td>
<td>Elephant</td>
<td>jpg</td>
<td>590 x 344</td>
<td>537 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>22.7 K b</td>
<td>22.7 K b</td>
<td>Watermark created in MSWord converted to pdf</td>
<td>6.0 Standard</td>
<td>Medical view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_002</td>
<td>Elephant</td>
<td>gif</td>
<td>550 x 359</td>
<td>497 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>23.7 K b</td>
<td>23.7 K b</td>
<td>From Arial 12 pts</td>
<td>6.0 Standard</td>
<td>Lateral view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_003</td>
<td>Elephant</td>
<td>gif</td>
<td>550 x 421</td>
<td>537 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>31.8 K b</td>
<td>31.8 K b</td>
<td>Arial 12 pts</td>
<td>6.0 Standard</td>
<td>Casual view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_004</td>
<td>Elephant</td>
<td>gif</td>
<td>550 x 421</td>
<td>537 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>30.7 K b</td>
<td>30.7 K b</td>
<td>University of Pretoria Veterinary Science Anatomy and Physiology</td>
<td>6.0 Standard</td>
<td>Ventral view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_005</td>
<td>Elephant</td>
<td>gif</td>
<td>550 x 421</td>
<td>537 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>30.7 K b</td>
<td>30.7 K b</td>
<td>University of Pretoria Veterinary Science Anatomy and Physiology</td>
<td>6.0 Standard</td>
<td>Dorsal view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_006</td>
<td>Elephant</td>
<td>gif</td>
<td>550 x 421</td>
<td>537 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>31.8 K b</td>
<td>31.8 K b</td>
<td>University of Pretoria Veterinary Science Anatomy and Physiology</td>
<td>6.0 Standard</td>
<td>Casual view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_007</td>
<td>Elephant</td>
<td>gif</td>
<td>550 x 421</td>
<td>537 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>32.7 K b</td>
<td>32.7 K b</td>
<td>University of Pretoria Veterinary Science Anatomy and Physiology</td>
<td>6.0 Standard</td>
<td>Ventral view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>el_008</td>
<td>Elephant</td>
<td>gif</td>
<td>550 x 421</td>
<td>537 x 328mm</td>
<td>90</td>
<td>✓</td>
<td>30.7 K b</td>
<td>30.7 K b</td>
<td>University of Pretoria Veterinary Science Anatomy and Physiology</td>
<td>6.0 Standard</td>
<td>Dorsal view of elephant skull, juvenile</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
UP Space at the University of Pretoria: Item 2263/2213 - Microsoft Internet Explo...

Please use this identifier to cite or link to this item: http://hdl.handle.net/2263/2213

Title: Adult elephant skull, caudal view
Author(s): Saegers, Christine D.
LC Subjects: Veterinary anatomy
Keywords: Elephant skull, Caudal view
Issue Date: 11-Apr-2007
Creation Date: 1994
Abstract: Anatomical representation of a caudal view of the adult elephant skull.
Description: Black and white pencil sketch; Original canvas size: (w)51.0 x (h)36.0 cm. Original scanned size in pixels: 6960 x 1640 (600 dpi). Final size in pixels: 550 x 461 (150 dpi). Estimate download time: 51 sec @ 28.8 kbps.
URI: http://hdl.handle.net/2263/2213
Rights: ©University of Pretoria. Faculty of Veterinary Science, Dept of Anatomy and Physiology (original). ©University of Pretoria, Dept of Library Services (digital). Provided for educational purposes only. It may not be downloaded, reproduced, or distributed in any format without written permission of the original copyright holder. Any attempt to circumvent the access controls placed on this file is a violation of copyright laws and is subject to criminal prosecution. Please contact the collection administrator for copyright issues.
Type: Image
Language:
Appears in Collections: The Elephant

Files in This Item:

- **oll_006.pdf**
  - Description: 135Kb
  - Format: Adobe PDF
  - View/Open

Note: No thumbnail available if PDF document
In Conclusion ...

High quality/ proper metadata can …
- Facilitate & improve retrieval of info
- Increase precision
- Increase search recall
- Increase trustedness of digital collection & items
- Also important for:
  - Administrative control
  - Security
  - Personal information
  - Management information
  - Content rating
  - Rights management
  - Preservation
Bibliography

• Making of America. http://www.hti.umich.edu/m/moagrp/moa_faq.html
• Oklahoma State University Library. http://digital.library.okstate.edu/manual/toc%20page.html
• Digitization 101 Blog of Jill Hurst-Wahl, MLS
• Cornell University Library; http://www.library.cornell.edu/iris/dpo/publications.html
• DigitFutures, Kings College, UK. Notes taken at workshop. October, 2006
Project selection criteria

The most common types of materials reported to be digitized:

- Photographs (50.4%)
- Bound materials such as books and journals (37.6%)
- Unbound sheets of paper smaller than 8.5” by 11” (33.9%)
- Oversized unbound sheets of paper larger than 8.5” by 11” (32.1%)
- Photograph slides or negatives (32.1%)