

ERPANET
Workflow in Digital
Preservation Budapest 13-15
Oct 2004

Managing Mass Digitisation of Manuscript
Documents

Rob Mildren – National Archives of Scotland

ERPANET WORKSHOP ON WORKFLOW

Project Background

- Originally planned as in index only
- Testaments – what they are
- Testaments – scope
- Project plan – Digitise, link to index and market via e-commerce

Managing Mass Digitisation of Manuscript Documents

Type of Material

- Bound Volumes
- Loose leaf material

Managing Mass Digitisation of Manuscript Documents

Selection and Assessment of Material

- Preservation Concerns
- Conservation assessment
- Conservation input before and after digitisation
- Storage of Manuscript Volumes
- Conservation Report
<http://www.scan.org.uk/aboutus/Reports/conservationreport.pdf>

Managing Mass Digitisation of Manuscript Documents

Preparation and pagination

- Pagination
 - Completeness
 - Duplication
 - Missing pages
- Quality Control Check

Managing Mass Digitisation of Manuscript Documents

Proper handling by trained camera operators

- Handling guidelines
- Part of training program
- Part of contract

Managing Mass Digitisation of Manuscript Documents

Image capture software that minimised operator intervention

- Staff with little ICT experience
- Fast image capture
- Optimised capture to PC
- Minimise capture time overheads

Managing Mass Digitisation of Manuscript Documents

Image Quality – Fit for Purpose

- CCD array cameras
- Qualitative assessment
- Capture sufficient pixels to reproduce penstrokes

Managing Mass Digitisation of Manuscript Documents

Formal quality control procedures

- Separate procedure
- Results recorded
- 30% check

Managing Mass Digitisation of Manuscript Documents

Software for data back ups

- Retain Colour TIFF
- Simple procedures for Camera Operators
- Written to DLT

Managing Mass Digitisation of Manuscript Documents

Resilience

- Online Storage
- RAID 5 protection and Hot Spare
- Tape copies
- Off Continent Storage

Managing Mass Digitisation of Manuscript Documents

Creating links between the finding aid and the images

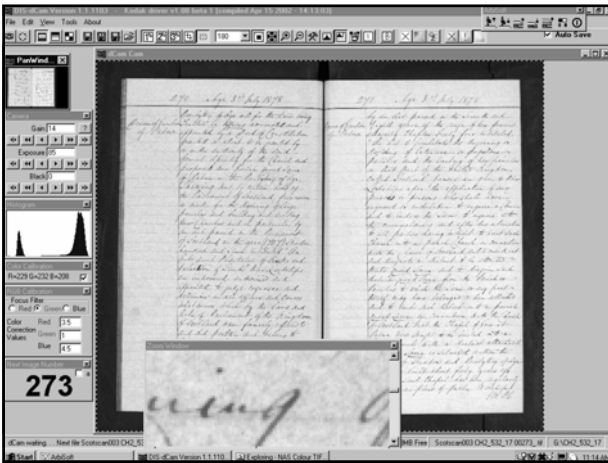
- Need more granular access to volumes of images
- Create a Comprehensive Index
- Image to Index linking program
- Special Provisions
 - Multiple testaments on a page
 - Testaments starting on the same page

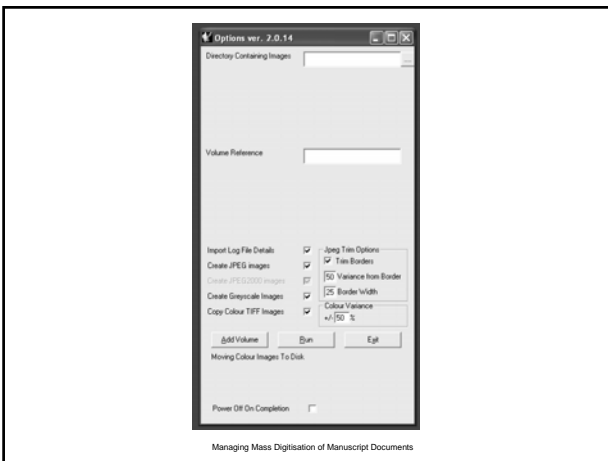
Managing Mass Digitisation of Manuscript Documents

E-Commerce Access

- Remote access via e-Commerce
- Index free of charge
- Customers retain access to images they have purchased

Managing Mass Digitisation of Manuscript Documents





What Do We Know About Our Images?

- Information from Capture
 - Operator and Camera
 - Camera Settings
 - Speed of capture
- Information from Image File Conversion
 - File creation times

Managing Mass Digitisation of Manuscript Documents

What Do We Know About Our Images?

- Information from Quality Control
 - QC operator
 - QC Status
 - Number of Retakes required
- Information from Image Linking
 - Linking operator
 - Page numbers
 - Sequence of index entries
 - Statistics

Managing Mass Digitisation of Manuscript Documents

What Do We Know About Our Images?

- Information from Tape Backup
 - Location of tape for Tiffs
 - Location of tape sent to SLC
 - Location of Jpeg tape
- Information from Virtual Volumes
 - Number of Orders made

Managing Mass Digitisation of Manuscript Documents

Conclusions

- Successful approach to large Scale digitisation
- High throughput
- Preservation friendly
- Value for Money

Managing Mass Digitisation of Manuscript Documents
