Creating an End-to-End Process for Implementing a Digital Archiving Workflow: How we are putting theory into practice

Leo Konstantelos & Emma Yan
Thu 21st Sep 2023, 13:30-15:00
Pilot project ran from Dec 2022 – Sep 2023 to produce an end-to-end process for implementing our digital archiving workflow.

Develop, test and evaluate the necessary framework for delivering robust digital preservation and digital archiving services.
Proposal submission gives more detail on what we intended to achieve

http://bit.ly/3LmEbYZ
Key areas

- Develop policies, workflows, procedures and tools for digital archiving
- Review existing ASC policies and donor agreements
- Update digital asset register, create risk register
Outcomes – Archival Forensics Lab access policy

- Policy provides conditions of access for individuals seeking access to the Archival Forensics Lab (AFL).
- Follows and/or complies with the specifications defined in:
  - Interpol Global Guidelines For Digital Forensics Laboratories
  - UK Gov Forensic science providers: codes of practice and conduct, 2021
  - AFL access agreement
  - AFL access application form
Outcomes – Digital holdings

• Updated a previous digital scoping survey to build a digital asset register

• Need to think about integration with our collection management system, particularly around location logging
Outcomes – Quarantine and conservation

• Guidelines for quarantine checks and conservation of computer storage media

• Covers magnetic disks, floppy disks, optical media, cassettes and tapes, solid state media

• Based on British Standards, community resources and own expertise

• Will be released under CC Licence Q1 2024
Outcomes – archival forensics workflow

Archival Forensics Workflow (Storage media deposit)

Preparation: Consult conservation report, image back-up, and verify media authenticity.

Imaging: Write blocking, disk image generation, image verification, and analysis.

Processing: Collect evidence, process records, forensic examination, and verify check digital evidence.

Analysis: Check data agreement, data cleaning, and analysis.

Exporting: Apply search filters, apply labels, and export reports.

Collections Management System: Preservation logs for media characterization and management configuration.

Digital Archiving workflow: Nulled imaging export, digital archiving steps, and export to logical image.

http://bit.ly/3EGzlSw
Outcomes – Digital archiving logs
Outcomes – Risk register

47 risks identified across 9 root causes:

- Management
- Legal liabilities
- Operations
- Staffing
- Technical infrastructure
- Security
- Acquisition and ingest
- Preservation and archival storage
- Metadata

Based on:

- Digital Repository Audit Method (DRAMBORA)
- Risk assessment handbook and self-assessment tool (TNA)
- Managing digital continuity (TNA)
- Digital Preservation Handbook, Risk and change management (DPC)
Outcomes – End-to-end digital archiving case study

Evidence-based approach to delivering digital preservation and digital archiving services.

Implement an end-to-end digital archiving case study using one of our collections to test workflows, explore benefits of archival forensics.

Highlight areas where improvements and further investment are required.
Outcomes – End-to-end digital archiving case study

| Description | Dance House Glasgow was a creative arts organisation involved in supporting the city’s professional dance sector and offering community development programmes for over 20 years. In 2018, it lost its Creative Scotland funding and ceased operating. |
| Variability | The whole collection consists of records relating to Dance House Glasgow dating from c.1990 to 2018. It includes governance, financial, staff and project records, along with photographs, audio and video, press cuttings, and promotional material. |
| Volume (digital) | 6.3TB |
| Representative ness | Hybrid – both born digital and paper records. Digital records across three HDDs and 90 optical media discs. At least one of the hard drives and some of the CD-ROMs not functioning, as per a survey of the material in 2021. |
| Legal issues | The records came to UofG ASC via the Business Archives Surveying Officer and were gifted to us in 2019. As a creative arts collection, we expected issues around IPR and rights for the music used in the collection. As a business collection, we expected issues around personal information, confidential records etc. |
Outcomes – End-to-end digital archiving case study

Dance House Glasgow collection: computer storage media
Outcomes –
End-to-end digital archiving case study

F.R.E.D workstation imaging a corrupt hard drive
Outcomes – End-to-end digital archiving case study
Outcomes – End-to-end digital archiving case study

Examples of born-digital items in the collection

boogie babies and mini movers

Fridays 16 Sep - 2 Dec 2011
Boogie Babies (age 0 - 2) 11.30 - 12.15
Mini Movers (age 2 - 5) 12.30 - 13.15

Partick Burgh Halls
9 Burgh Hall St Glasgow G11
For booking & info please call:
0141 552 2442

£50 for 12 weeks or £5 per session

These fun, energetic parent & child classes are devised to promote coordination, movement skills and imaginative play using movement, storytelling and games. Children of all abilities are welcome and adults attending the children’s classes are expected to join in!
Forensic analysis – FTK and BitCurator:
• Data carving
• Decryption and password recovery
• Personal data analysis (Bulk extractor)
• Document content analysis
• Format identification and validation (FTK, Droid, JHOVE)
• Duplicate flagging
• Index search in FTK

FTK – Weeding:
• Duplicates
• Deleted files
• Free space / file slack
• Recycle bin items
• OS / system files

Outcomes – End-to-end digital archiving case study

<table>
<thead>
<tr>
<th>Selected file categories</th>
<th>Pre-appraisal</th>
<th>Post appraisal</th>
<th>Difference</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archives</td>
<td>1572</td>
<td>48</td>
<td>-1524</td>
<td>Difference includes deleted and temporary files</td>
</tr>
<tr>
<td>Databases</td>
<td>2</td>
<td>1</td>
<td>-1</td>
<td>Difference includes deleted and temporary files</td>
</tr>
<tr>
<td>Documents</td>
<td>1802</td>
<td>219</td>
<td>-1583</td>
<td>Difference includes deleted and temporary files</td>
</tr>
<tr>
<td>Email</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Executables</td>
<td>510</td>
<td>0</td>
<td>-510</td>
<td></td>
</tr>
<tr>
<td>Graphics</td>
<td>24225</td>
<td>15278</td>
<td>-8947</td>
<td>Difference includes deleted and temporary files</td>
</tr>
<tr>
<td>Internet</td>
<td>35</td>
<td>35</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Multimedia</td>
<td>1456</td>
<td>979</td>
<td>-477</td>
<td>Difference includes deleted and temporary files</td>
</tr>
<tr>
<td>OS/File system</td>
<td>1081</td>
<td>0</td>
<td>-1081</td>
<td></td>
</tr>
<tr>
<td>Presentations</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Slack/free space</td>
<td>20512</td>
<td>0</td>
<td>-20512</td>
<td></td>
</tr>
<tr>
<td>Spreadsheets</td>
<td>12</td>
<td>11</td>
<td>-1</td>
<td>Difference includes deleted and temporary files</td>
</tr>
</tbody>
</table>

Selected file status:
| Deleted files            | 557           | 9              | -557       |       |
| Duplicate items          | 14995         | 3607           | -11388     |       |
| KFF Alert files          | 37            | 0              | -37        |       |
| KFF Ignore files         | 1288          | 0              | -1288      |       |
| From recycle bin         | 3060          | 0              | -3060      |       |

- Total volume (GB): Pre-appraisal 6363, Post appraisal 423, Difference -5941, Total volume reduced by 5.94TB
- Total file size: Pre-appraisal 55936, Post appraisal 17954, Difference -37982, Total files reduced by 37,982
Outcomes - final thoughts

What we didn’t complete:
• Donor agreement updates
• Collection Development policy update
• Research data end of life

Moved to planning for 2023/2024 academic year.

By the end of the pilot, we achieved what we needed to, plus more.
• Digital archiving guides

Read about our work on the new DPC Digital Preservation Documentation Guide
Come and talk to us!
Leo.Konstantelos@glasgow.ac.uk
Emma.Yan@glasgow.ac.uk