LIFE3: Predicting Long Term Preservation Costs

Paul Wheatley
Digital Preservation Manager
The British Library
LIFE projects overview

- Collaboration between University College London (UCL), the British Library (BL) and HATII at the University of Glasgow

- Co-funded by Joint Information Systems Committee (JISC) and the Research Information Network (RIN)

- The LIFE Project:
  - 1 year project
  - Completed in April 2006

- The LIFE² Project:
  - 1.5 year project
  - Completed August 2008

- The LIFE³ Project:
  - 1 year project
  - Began August 2009
Overview of the LIFE Projects so far

■ The LIFE Project:
  ▪ Aim: to explore a lifecycle approach to costing the preservation of digital materials
  ▪ The Project developed:
    ▪ A generic model of the digital preservation lifecycle
    ▪ A methodology for assessing lifecycle costs against this model
    ▪ 3 case studies, examining and costing a range of digital lifecycles

■ The LIFE² Project:
  ▪ Aim: to evaluate, refine and further develop the techniques developed in phase one of LIFE
  ▪ Key elements:
    ▪ Review by external economics expert
    ▪ Revision of the lifecycle model and costing methodology
    ▪ 3 new lifecycle case studies
The LIFE³ Project:

- Aim: To develop the ability to estimate preservation costs across the digital lifecycle
- The Project is developing:
  - A series of costing models for each stage and element of the digital lifecycle
  - An easy to use costing tool
  - Support to enable easy input of data
  - Integration to facilitate use of the results
LIFE3 costing tool inputs

- **Content Profile**
  - Type of Content
  - File format
  - Complexity
  - Volume

- **Organisational Profile**
  - Existing infrastructure
  - Preservation policy
  - Legal constraints

- **Context**
  - Inflation
  - Hardware costs and trends
  - Staff costs
### LIFE³ costing tool outputs – estimated costs

<table>
<thead>
<tr>
<th>Lifecycle Elements</th>
<th>Lifecycle Stage</th>
<th>Acquisition</th>
<th>Ingest</th>
<th>Bit-stream Preservation</th>
<th>Content Preservation</th>
<th>Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Check-in</td>
<td></td>
<td>Selection</td>
<td>Quality Assurance</td>
<td>Repository Admin</td>
<td>Preservation Watch</td>
<td>Access Provision</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Submission Agreement</td>
<td>Metadata</td>
<td>Storage Provision</td>
<td>Preservation Planning</td>
<td>Access Control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IPR &amp; Licensing</td>
<td>Deposit</td>
<td>Refreshment</td>
<td>Preservation Action</td>
<td>User Support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ordering &amp; Invoicing</td>
<td>Holdings Update</td>
<td>Backup</td>
<td>Re-ingest</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Obtaining</td>
<td>Reference Linking</td>
<td>Inspection</td>
<td>Disposal</td>
<td></td>
</tr>
</tbody>
</table>
Integration

- DROID
- Planets Content Profile
- FITS
- DRAMBORA
- Planets Preservation Policy
- Data Audit Framework
- Context

Cost Estimation Tool

- DRAMBORA
- Plato
- JISC Framework
Template approach

- Templates for typical content and organisational profiles
- Lower barrier of access for none DP experts
- Auto completion of specific inputs
- Custom profiles
Strategic Issues

- Challenging context
  - Hybrid world
  - Non-digital not dying, in fact usage is increasing
  - Greater variety of digital content
  - Scale

- Allocation of resource: ratio of digital to non-digital spending
  - Non-digital preservation : Digital preservation

- Digital as a preservation medium for non-digital content
  - Preservation of non-digital content
  - Risk, cost

- Supporting the lifecycle approach
  - Evidence of efficiencies over the medium to long term
Preservation Planning

- Collection management decision making
  - Whether to purchase/acquire/digitise?

- Selecting an appropriate preservation solution
  - Plato
  - Cost – Risk – Value
  - Preservation requirements

- Budgeting for expected preservation costs
LIFE-SHARE

- LIFE-SHARE Project
  - Focus on digitisation
  - Activity costing and analysis
  - Skills audit
  - Supporting a preservation and lifecycle approach to digitisation

- www.leeds.ac.uk/library/projects/lifeshare (available soon)
Key challenges, and request for help

- Content complexity
  - Categorisation of content type / complexity and impact on effort required to preserve

- Data, activity costing
  - Capturing / contributing costing data

- Trialling the models, feedback
A finite amount of funding is available for digitisation, ingest and preservation of a collection. How many items should be digitised without overspending?

A digital collection is due to be ingested into an organisation’s digital repository. Migration to a new file format offering superior compression and savings in storage cost is a possibility, but the operation itself will also have a cost. Should the organisation migrate the collection?

An organisation is considering outsourcing the storage, preservation and access of a digital collection. The service provider gives a quote. Will outsourcing save the organisation money?

A digitisation project within an organisation is not following best practice. What will be the cost of picking up the pieces in 5 years time?