

PASIG Dublin

17-19 October 2012

Cost of Archiving - Current Practices

Costs Versus Benefits

Neil Grindley

JISC

Programme Manager – Digital Preservation & Curation



[Home](#) » [Blogs](#) » [paul's blog](#) » Digital Preservation Cost Modelling: Where did it all go wrong?

Digital Preservation Cost Modelling: Where did it all go wrong?

By [paul](#) on 29 June 2012 - 2:08pm

I recently spoke at a [workshop on digital preservation costing](#) organised by the lovely people at Knowledge Exchange and Nordbib. After briefly covering some of the work I was previously involved in as part of the [LIFE Projects](#), I talked about why I think that

Digital Preservation Costing Initiatives

- LIFE 1, 2 and 3. Projects to explore digital preservation costing, and develop costing models.
 - <http://www.life.ac.uk>
- Cost Model for Digital Preservation (CMDP): Project at the Royal Danish Library and the Danish National Archives to develop
 - <http://www.costmodelfordigitalpreservation.dk>
- Keeping Research Data Safe 1 and 2 (KRDS): Cost model and benefits analysis for preserving research data
 - <http://www.beagrie.com/krds.php>
- Presto Prime cost model for digital storage
 - <http://prestoprime.it-innovation.soton.ac.uk>
- Cost Estimation Toolkit (CET): Data centre costing model and toolkit, from NASA Goddard
 - http://www.pv2007.dlr.de/Papers/Fontaine_CostModelObservations.pdf
- Cost Model for Small Scale Automated Digital Preservation Archives (Strodl and Rauber)
 - http://www.ifs.tuwien.ac.at/~strodl/paper/strodl_ipres2011_costmodel.pdf
- APARSEN Project activity focused on digital preservation costing
 - <http://www.alliancepermanentaccess.org/index.php/knowledge-base/digital-preservation-business-models/costbr>
- EPSRC and JISC study on Cost analysis of cloud computing for research
 - http://www.jisc.ac.uk/media/documents/programmes/research_infrastructure/costcloudresearch.pdf
- Cost forecasting model for new digitization projects (Excel and web tool under development) (Karim Boughida, Martha WI)
 - http://www.cni.org/wp-content/uploads/2011/12/cni_cost_boughida.pdf
- DP4lib is developing a business and cost model for a digital preservation service
 - <http://dp4lib.langzeitarchivierung.de/downloads/DP4lib-One-Pager-08-eng.pdf>
- DANS Costs of Digital Archiving Volume 2 Project, focusing on preservation and dissemination of research datasets
 - <http://www.dans.knaw.nl/en/content/categorieen/projecten/costs-digital-archiving-vol-2>
- Blue Ribbon Task Force on Sustainable Digital Preservation and Access
 - <http://brtf.sdsc.edu>
- Economic Sustainability Reference Model
 - <http://unsustainableideas.wordpress.com/economic-sustainability-ref-model-page/>
- ENSURE Project - Enabling kNowledge Sustainability Usability and Recovery for Economic value
 - <http://ensure-fp7-plone.fe.up.pt/site>
- 4C. EU funded project on costing. Due to commence in 2012. Led by JISC.
 - TBC
- Cost Model for Electronic Health Records (Bote, Fernandez-Feijoo, and Ruizb)
 - <http://www.sciencedirect.com/science/article/pii/S2212017312004434>

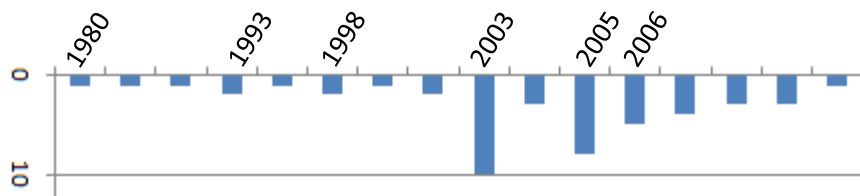
A Digital Asset Sustainability and Preservation Cost Bibliography

June 26th, 2012 by [Butch Lazorchak](#)



Back in the [early days](#) of the NDIIPP program we had a series of cross-cutting initiatives that we called "affinity groups." These groups addressed areas of interest such as "collection and selection," "technical architecture" and "rights and restrictions" that cut across all the different projects.

I was interested in the work of the "economic sustainability" working group, and for no particular reason I



Contingent Valuation
 Ecosystem Valuation
 Intangible Assets
 Investments
 Information Markets
 Sustainability
 Incentives
 Lifecycle Modelling
 Loss Costing
 Business Modelling
 Cost-Driven Design

Cost/Benefit Analysis
 Net Financial Benefit
 Net Economic Benefit
 Net Social Benefit
 Shadow Pricing

Relevant JISC-funded work

- 3 Phases of the LIFE Project (2005 – 2010)
- Piloting the LIFE Costs Tool in UK HEI's (2011)
- 4 Phases of Keeping Research Data Safe (2008 – 2011)
- The Costs Observatory Study (2011)
- The Blue Ribbon Task Force on Sustainable Digital Preservation and Access (2008-2010) (contributed towards funding)
- A Draft Economic Sustainability Reference Model (2011) (with OCLC)

IPRES 2011 Presentation

<http://www.slideshare.net/neilgrindley/ipres-2011-the-costs-and-economics-of-preservation>

Support and Enabling Project - £250k (2 years)

SPRUCE Project (Sustainable Preservation Using Community Engagement)

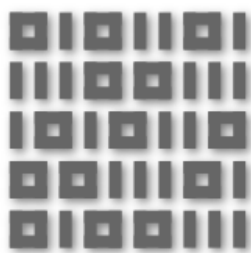
(University of Leeds/British Library/Open Planets Foundation/Digital Preservation Coalition/London School of Economics)

Engage, support and enable the community to contribute to:

The Business Case for Digital Preservation

SPRUCE making a very practical contribution ...

Datasets



These are the Datasets or collections that relate to specific preservation Issues which in turn (may) have Solutions developed for them. The Datasets are categorised by their media type.

[Click this link to create a new Dataset, then edit the italicised text](#).

Audio datasets

- Label: [audio](#)
- [User-generated audio field recordings](#)
- [Audio Collection \(York\)](#)
- [LAVC audio](#)
- [Endangered Archives Programme \(EAP\)](#)

Disk image datasets

- Label: [disk_image](#)
- [Disk Images](#)

11 Categories
60 Items

Issues



These are the preservation or other business driven Issues that relate to a specific Dataset and may have one or more specific Solutions developed to solve them. The aim of an Issue page is to provide a detailed description of the preservation challenge and the requirements of the Issue Owner that will help to inform development of a Solution that solves the Issue.

[Click this link to create a new Issue, then edit the italicised text](#).

Unsolved issues

- Issues that do not have linked solutions. Why not suggest or contribute a solution?
- Label: [unsolved_issue](#)
- [Classification of files within a disk image](#)
 - [Extraction of keywords \(and images\) from large collections of text based files](#)
 - [Checking significant properties of images have been retained after migration](#)
 - [Correct File Formats](#)
 - [Checking significant properties of documents have been](#)

19 Categories
159 Items

Solutions



These are Solutions that address specific Issues encountered in particular Datasets. Solutions are typically quite specific to a particular Issue and Dataset but many will have a wider application. For details of tools utilised in a Solution, either follow links from individual Solution pages or see the [Tool Registry](#).

[Click this link to create a new Solution, then edit the italicised text](#).

Appraisal and assessment solutions

- Solutions for assessing or appraising datasets.
- Label: [appraisal_assessment](#)
- [Analysis of Lucene Index Word Frequency](#)
 - [Characterising Externally Generated Content](#)

Bit rot detection and repair solutions

- Solutions for detecting and possibly repairing bit rot Issues in Datasets.
- Label: [bit_rot_detection](#)
-

14 Categories
80 Items

Support and Enabling Project - £250k (2 years)

SPRUCE Project (Sustainable Preservation Using Community Engagement)

(University of Leeds/British Library/Open Planets Foundation/Digital Preservation Coalition/London School of Economics)

Engage, support and enable the community to contribute to:

The Business Case for Digital Preservation

Support
Actions

Active Case Studies - £75k (4 months)

Carcenet Press

Manchester University

Digital Directorate

Institute of Education

Future Proofing

Uni of London Computer Centre

POPE

Institute of Education

POPS

Uni of Central Lancashire

Mashup /
Hackathon
Events

Case Studies

Enabling
Grants (£5k)

Mashup /
Hackathon
Events

Enabling
Grants (£5k)

Case Studies

Wiki
Evidence
Base



Evidence of Value

Proof of Impact

Benefits of Investment

Sample Business Cases

Enhancing Capability - £100k (9 months)

DataSafe Bristol University

DICE London School of Economics

SHARD Uni of London Computer Centre

PrePARE Cambridge University

Project Blogs
Learning Materials
Case Studies

Carcenet Project

(Archiving email)

Increased credibility
 Decreased risk of loss
 Established capability
 Increased confidence
 Progressed strategy
 Increased capacity
 Enhanced reputation
 Promoted collaboration
 Furthered research

Digital Directorate

(Records Management)

Established influence
 Improved Process
 Opened communication
 Increased knowledge
 Developed procedures
 Enabled future action
 Provided blueprint
 Enhanced collaboration

Future Proofing

(Rationalising ad hoc stored files)

Created infrastructure
 Started strategy development
 Established template
 Saved money
 Demonstrated options
 Enhanced sustainability
 Improved processes

POPE

(Fixing link rot)

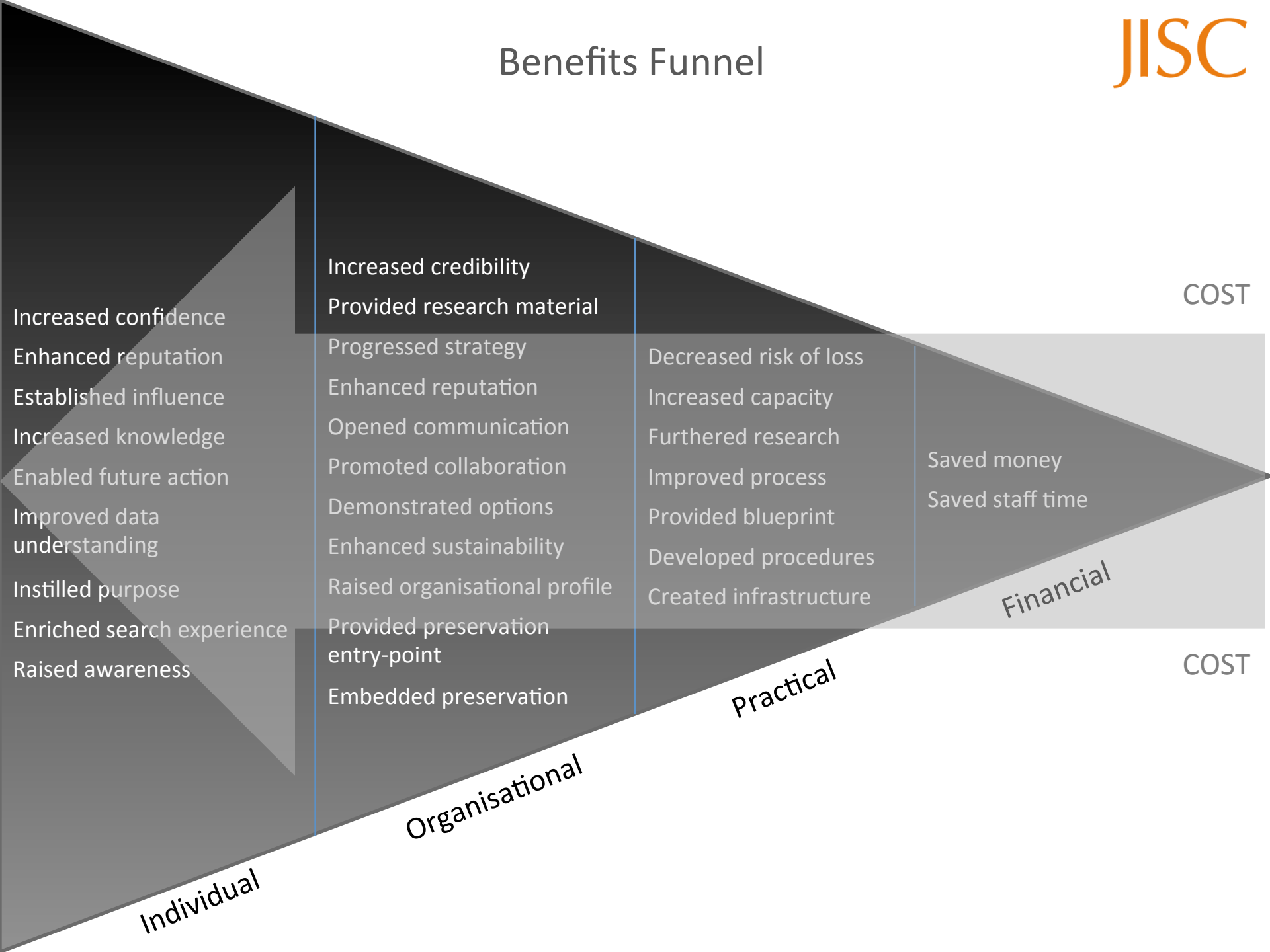
Defined future work
 Improved data understanding
 Enhanced planning
 Instilled purpose
 Saved staff time
 Enriched search experience
 Raised organisational profile
 Established sustainability
 Provided preservation entry-point

POPS

(Publishing from repository)

Embedded preservation
 Raised awareness

Benefits Funnel



A Collaboration to Clarify the Costs of Curation



Project Summary

The Collaboration to Clarify the Costs of Curation (4C) project will help organisations across Europe to more effectively invest in digital curation and preservation. Making an investment inevitably involves a cost and existing research on cost modelling provides the starting point for the 4C work. But the point of an investment is to realise a benefit, so work on cost must also focus on benefit, which must then encompass related concepts such as ‘risk’, ‘value’, ‘quality’ and ‘sustainability’. Organisations that understand this will be more able to effectively control and manage their digital assets over time, but they may also be able to create new cost-effective solutions and services for others.

Existing research into cost modelling is far from complete and there has been little uptake of the tools and methods that have been developed and very little integration into other digital curation processes. The main objective of the 4C project is, therefore, to ensure that where existing work is relevant, that stakeholders realise and understand how to employ those resources. But the additional aim of the work is to closely examine how they might be made more fit-for-purpose, relevant and useable by a wide range of organisations operating at different scales in both the public and the private sector.

Subject to final agreement



A Collaboration to Clarify the Costs of Curation



Project Summary

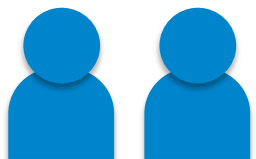
The Collaboration to Clarify the Costs of Curation (4C) project will help organisations across Europe to more effectively invest in digital curation and preservation. Making an investment inevitably involves a cost and existing research on cost modelling provides the starting point for the 4C work. But the point of an investment is to realise a benefit, so work on cost must also focus on benefit, which must then encompass related concepts such as 'risk', 'value', 'quality' and 'sustainability'. Organisations that understand this will be more able to effectively control and manage their digital assets over time, but they may also be able to create new cost-effective solutions and services for others.

Existing research into cost modelling is far from complete and there has been little uptake of the tools and methods that have been developed and very little integration into other digital curation processes. The main objective of the 4C project is, therefore, to ensure that where existing work is relevant, that stakeholders realise and understand how to employ those resources. But the additional aim of the work is to closely examine how they might be made more fit-for-purpose, relevant and useable by a wide range of organisations operating at different scales in both the public and the private sector.

2 year Co-ordination Action scheduled to start February 2013

Subject to final agreement





Assessment

Tasks

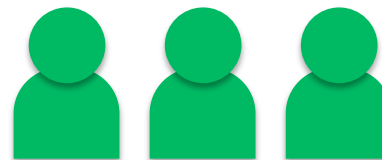
- Assess cost models & strategies
- Examine good practice
- Analyse requirements
- Integrate components
- Produce guidance & briefing materials
- Setup costs exchange



Engagement

Tasks

- Engage stakeholders
- Raise awareness
- Organise meetings
- Promote Research & Innovation
- Build community network



Collaboration to Clarify the Costs of Curation



Networking & Coordination



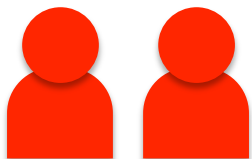
Affiliate Partners & Stakeholders



Enhancement

Tasks

- Examine and refine related concepts
- Value
 - Risk
 - Benefits
 - Sustainability
 - Economic Reference Model



JISC

Project Coordination

Tasks

- Project meetings
- Project reporting
- EC liaison
- Budget oversight
- Outputs QA



Outputs



Reports for General Dissemination



Curation Costs Exchange



Reports for European Commission



Submission of Roadmap to the EC

Conclusion

Organisations need to routinely use (embed) cost/benefits analysis when specifying, implementing and evaluating preservation processes.

... And do so in a granular way

... And be prepared to share that analysis!

Thankyou