

The LIFE² Project

0110110101010101100111010011011010101011001
1101001101101010101011001110100110110101010
110011101001101101010101011001110101100110110
101010101100111010011011010101010110011101001
101101010101011001110100110110101010101100111



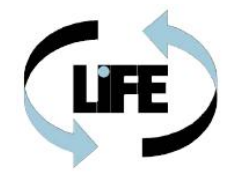
Rory McLeod
LIFE Project SRO,
The British Library



Overview

- ▶ LIFE¹ and LIFE² Case Studies
- ▶ BL Newspapers Case Study
- ▶ Analogue and Digital Lifecycles
- ▶ The LIFE model



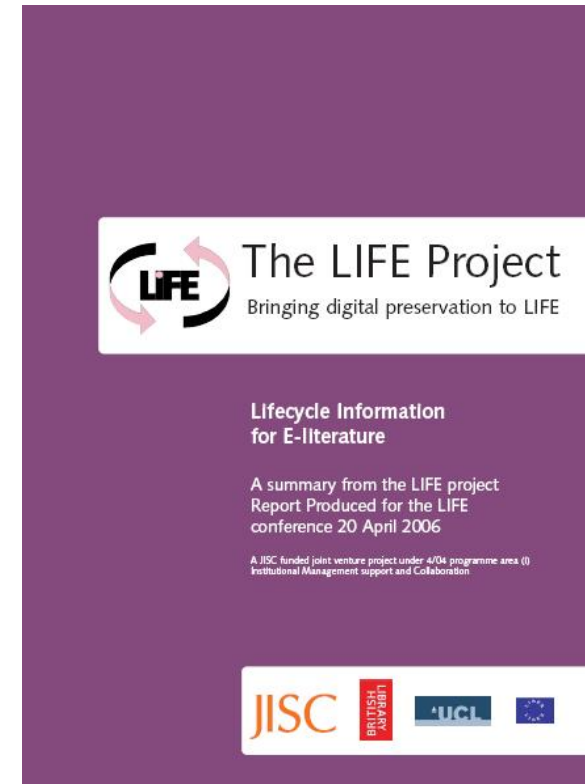


LIFE¹ Case Studies

e-Journals

Web Archiving

Voluntary Deposit of Electronic Publications



→ www.life.ac.uk/1

LIFE² Case Studies

Digitisation as Surrogacy:

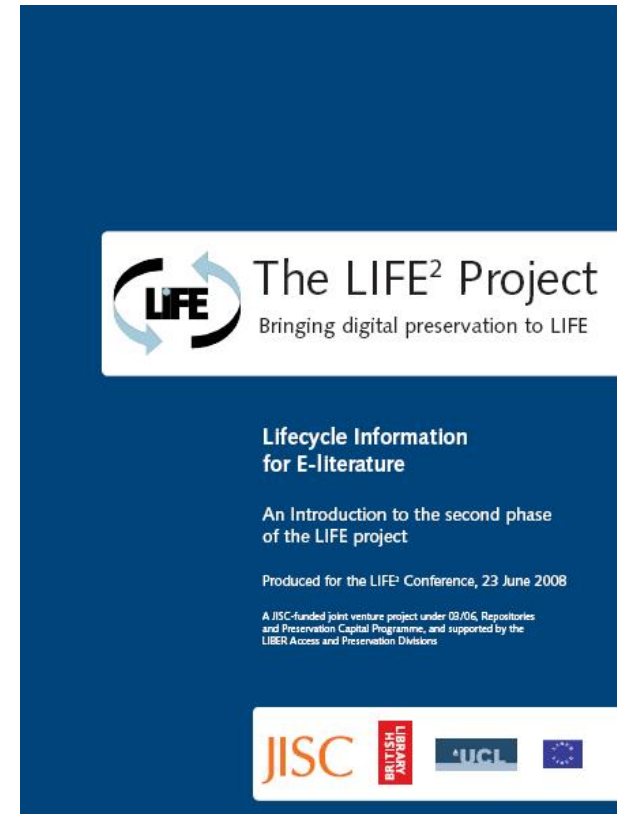
- ▶ BL Newspapers


Institutional Repositories:

- ▶ SHERPA-LEAP
- ▶ SHERPA-DP

Primary Data:

- ▶ Medical Research Council







 The LIFE² Project
Bringing digital preservation to LIFE

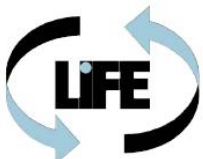
**Lifecycle Information
for E-literature**

An Introduction to the second phase
of the LIFE project

Produced for the LIFE² Conference, 23 June 2008

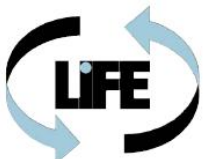
A JISC-funded joint venture project under D3.0.6, Repositories
and Preservation Capital Programme, and supported by the
LIBER Access and Preservation Divisions



LIFE Model v1.1: Stages and Elements

Lifecycle Stage	Creation or Purchase	Acquisition	Ingest	Metadata Creation	Bit-stream Preservation	Content Preservation	Access
	Selection	Quality Assurance	Re-use Existing Metadata	Repository Admin	Preservation Watch	Access Provision
....	Submission Agreement	Deposit	Metadata Creation	Storage Provision	Preservation Planning	Access Control	
....	IPR & Licensing	Holdings Update	Metadata Extraction	Refreshment	Preservation Action	User Support	
Lifecycle Elements	Ordering & Invoicing	Reference Linking		Backup	Re-ingest	
		Obtaining			Inspection		
		Check-in					



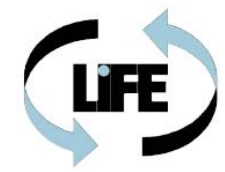
Case Study Outputs

- ▶ Full Case Study Write-up
- ▶ Workflow diagram
- ▶ Costing Spreadsheet

Lifecycle Stages	Practical Explanation	Cost notes	% of Staff time	Year 1 (Individual costs)	Year 1	Year 2
Creation or Purchase						
	Cost to digitise the total archive plus create associated project information	creation of digital archive		0.00		0.00
					250,445.00	272,713.00
	Sub Total				250,445.00	272,713.00
Acquisition						
Selection						
Selection Policy (policy/procedure)		10% Curatorial Grade B	10.0%	5,556.32	5,556.32	0.00
Selection (action)	Sort Microfilm into correct order	50% Curatorial Grade C, 8 months	25.0%	10,713.15	10,713.15	0.00
Selection Metadata (metadata)						
Submission Agreement						
Submission Agreement (policy/procedure)	Setting up of contract with third party	10% project Grade A	10.0%	6,647.62	6,647.62	6,647.62
Negotiation of Submission (action)						
Submission Metadata (documentation)	Recording of metadata relating to submission requirements.	5% Legal support Grade A	5.0%	3,323.81	3,323.81	0.00
IPR & Licensing						
IPR & Licensing (policy/procedure)	contract	5% Product development Grade B	5.0%	2,778.16	2,778.16	0.00
Negotiation of Rights						

- ▶ Website
- ▶ LIFE Blog

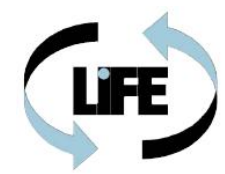
www.life.ac.uk
www.life.ac.uk/blog



Digitisation as surrogacy in detail

To develop a method of comparing
analogue and digital lifecycles

and then test the approach
by applying it to analogue and digital lifecycles
from British Library collections

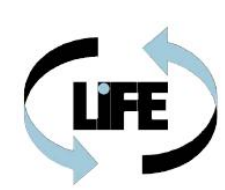


The Burney Collection

- ▶ Purchased by the British Library in 1818 for £13,500
- ▶ 1,100 volumes of the earliest known newspapers
- ▶ 1,000,000 pages from 17th, 18th and 19th Centuries.
- ▶ Re-scanning or re-microfilming is not possible.
- ▶ Microfilmed in the 1970s
- ▶ Digitisation started in 1995-96 and ran until 2004.

**19TH CENTURY
BRITISH LIBRARY NEWSPAPERS**

Morning Chronicle
LONDON, THURSDAY, SEPTEMBER 27, 1832.
SIR WALTER SCOTCHPOPE



Issues that arise from Newspapers Study

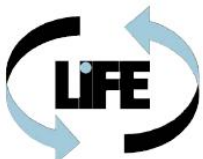
- ▶ Comparing digital and analogue lifecycles
- ▶ What is the lifecycle cost to an institution of producing digitised surrogates?
- ▶ What are the key preservation issues common across digitisation projects of differing scales?





Case Study Process

- ▶ Develop analogue and digital workflows
- ▶ Terminology issues
- ▶ How can we get a meaningful comparison?
- ▶ Future digitisation costs (Creation Stage)



What's compared

Analogue

Digital

Legal Deposit Newspapers

Burney Collection



Per-entity cost

	A	B	C	D	E	F
1						

Burney Digital Collection- L

5	Lifecycle Stages	Practical Explanation	Cost notes	% of Staff time	Year 1 (individual costs)	Year 1
---	------------------	-----------------------	------------	-----------------	---------------------------	--------

6						
---	--	--	--	--	--	--

7 Creation or Purchase

8					0.00	
---	--	--	--	--	------	--

9		Cost to digitise the total archive plus create associated project information	creation of digital archive			250,445.00
---	--	---	-----------------------------	--	--	------------

10						
11	Sub Total					250,445.00

12						
----	--	--	--	--	--	--

13 Acquisition

14 Selection

15	Selection Policy (policy/procedure)		10% Curatorial Grade B	10.0%	5,556.32	5,556.32
----	-------------------------------------	--	------------------------	-------	----------	----------

16	Selection (action)	Sort Microfilm into correct order	50% Curatorial Grade C, 6 months	25.0%	10,713.15	10,713.15
----	--------------------	-----------------------------------	----------------------------------	-------	-----------	-----------

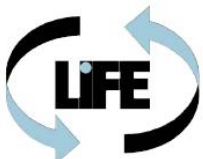
17	Selection Metadata (metadata)					
----	-------------------------------	--	--	--	--	--

18 Submission Agreement

19	Submission Agreement (policy/procedure)	Setting up of contract with third party	10% project Grade A	10.0%	6,647.62	6,647.62
----	---	---	---------------------	-------	----------	----------

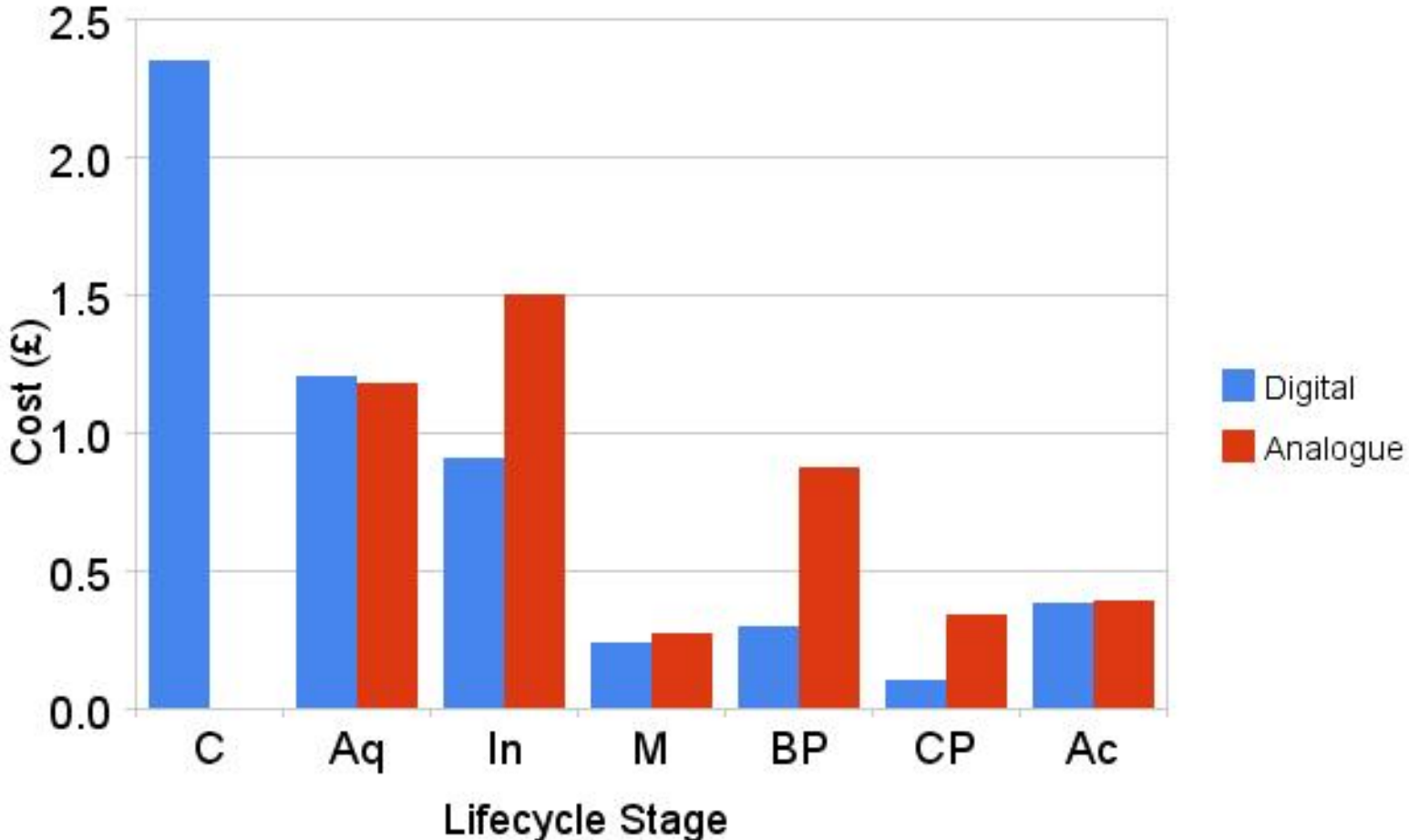
20	Negotiation of Submission (action)		5% Legal support Grade A	5.0%	3,323.81	3,323.81
----	------------------------------------	--	--------------------------	------	----------	----------

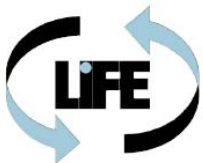
	Submission Metadata	Recording of metadata relating to submission				
--	---------------------	--	--	--	--	--



Overview of Costs

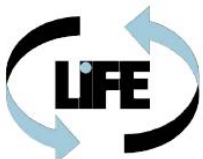
Entity Cost by LIFE Stage



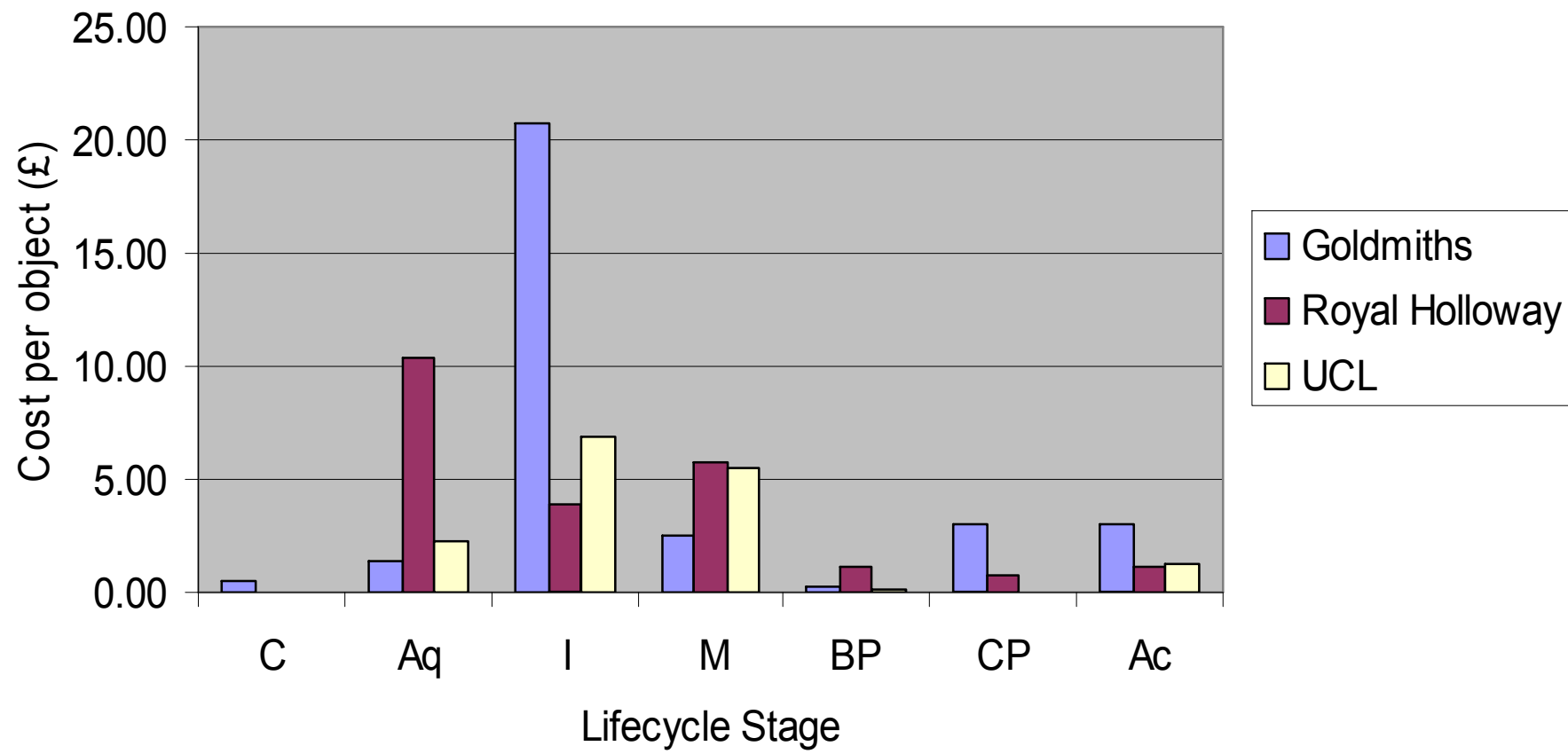


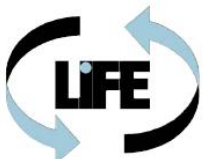
Overview of Costs

LIFE Stage	Aq	In	M	BP	CP	Ac	Total
Digital	£1.20	£0.91	£0.24	£0.30	£0.10	£0.38	£3.13
Analogue	£1.18	£1.50	£0.27	£0.87	£0.34	£0.39	£4.56



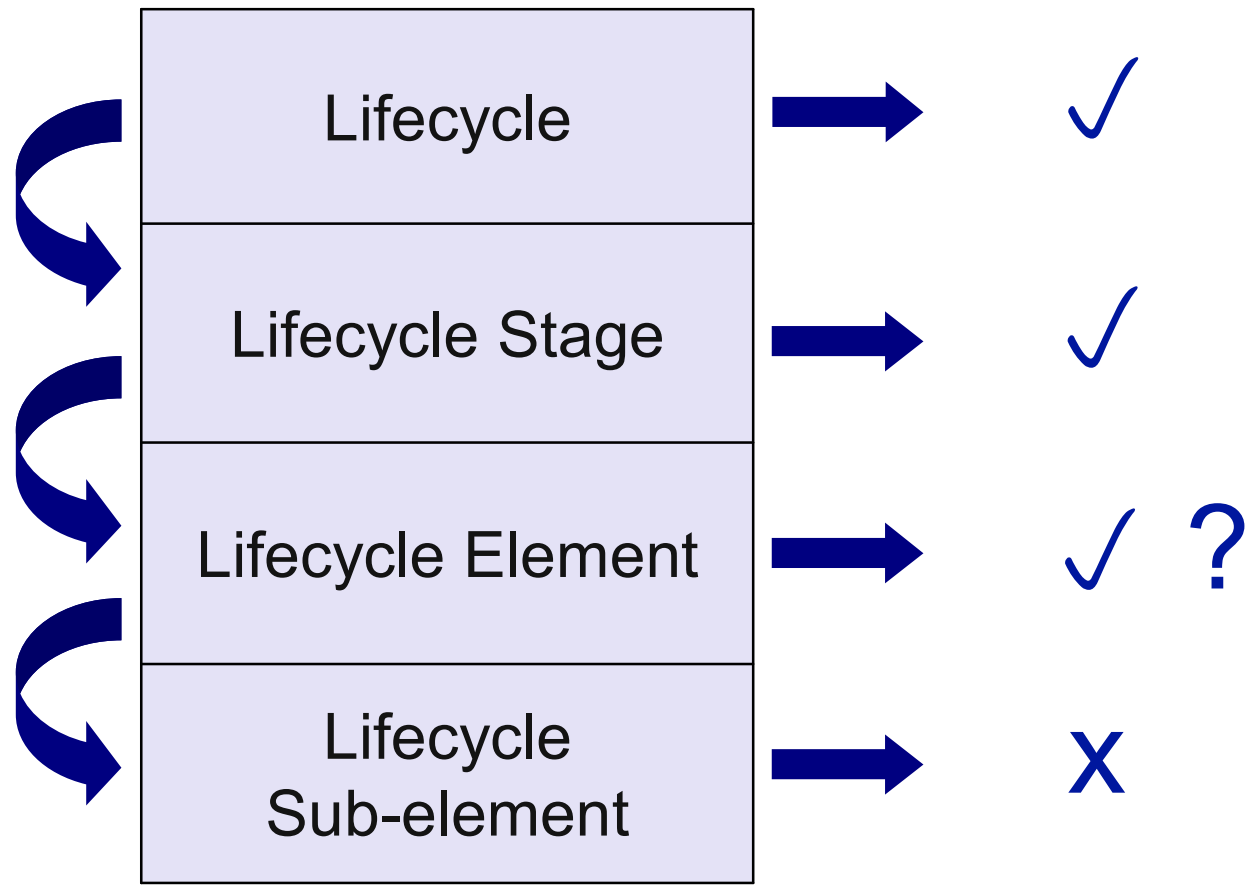
Repository lifecycle costs

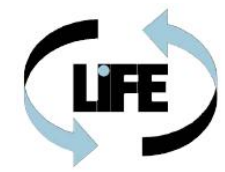




Applicability of the model

$$L_T = C + Aq_T + I_T + M_T + BP_T + CP_T + Ac_T$$

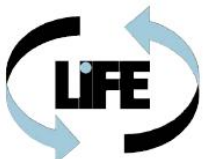




the LIFE Model in more detail

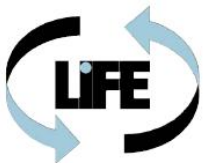
What it does

1. Standard way of representing the key functions in a lifecycle
2. Provides sufficient detail to enable useful analysis of lifecycles
3. Remains broadly high level to ensure relevance across different lifecycles and content types
4. Enables like with like comparison between different lifecycles



LIFE Model v1.1: Stages and Elements

Lifecycle Stage	Creation or Purchase	Acquisition	Ingest	Metadata Creation	Bit-stream Preservation	Content Preservation	Access
	...		Selection	Quality Assurance	Re-use Existing Metadata	Repository Admin	Preservation Watch
...		Submission Agreement	Deposit	Metadata Creation	Storage Provision	Preservation Planning	Access Control
...		IPR & Licensing	Holdings Update	Metadata Extraction	Refreshment	Preservation Action	User Support
Lifecycle Elements	...	Ordering & Invoicing	Reference Linking		Backup	Re-ingest	
		Obtaining			Inspection		
		Check-in					



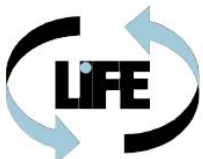
LIFE Model v1.1: Sub-elements

Ensuring the Model is clear and unambiguous to apply:

- Detailed definitions
- Sub-element descriptions

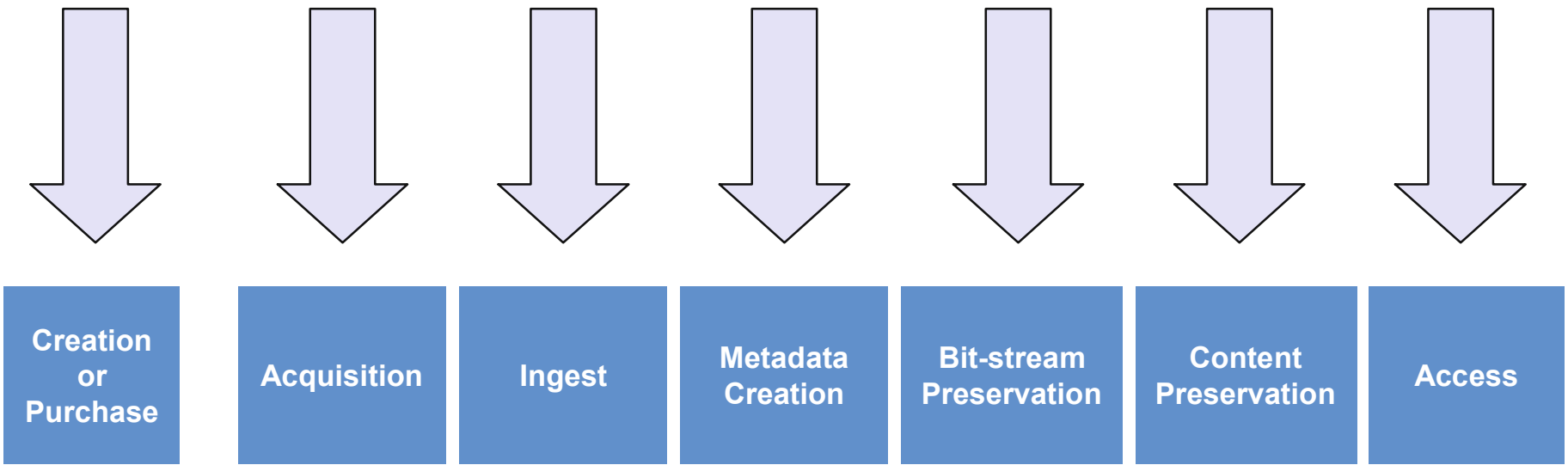
Sub-elements are suggested functions or activities only

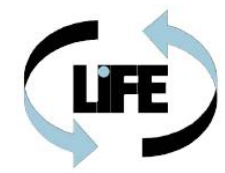
Sub-elements of Quality Assurance	Explanation / notes
QA Policy (policy/procedure)	Description of quality requirements and required mitigation actions should quality requirements not be met. Policy for sampling of objects for QA (if appropriate)
QA Characterisation (action)	Characterisation of the digital object. Identification of file format, and assessment of whether the object is valid, well formed, and/or renders correctly with current access software
Content Examination (action)	Assessment of whether the content of the digital object is of an expected, agreed or sufficient level of quality. Typically, a manual process on a sample of the ingested objects
Mitigation (action)	Action to mitigate quality issues (might include virus cleaning or re-ordering or obtaining the digital object)
QA Metadata (metadata)	Record QA metadata



Looking ahead: "LIFE3"?

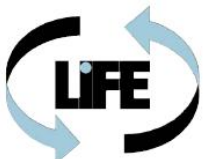
Estimative models for each stage of the lifecycle





Some Conclusions at this point

- ▶ Comparison is complex but workable. LIFE model provides a good mind-map across complex business and archival functions.
- ▶ Retrospective costing adds complications- knowledge is lost.
- ▶ Similar costs across a number of LIFE Stages- much more data required.
- ▶ Analogue lifecycles are well established compared to digital.
- ▶ Digitisation costs coming down - should we scan later?
- ▶ Preservation costs still largely estimated and unknown
- ▶ Sustainability therefore uncertain - If you don't know how much something costs, its hard to predict its future.



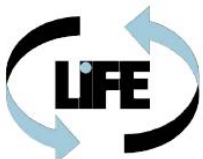
Thank you.

LIFE Team Acknowledgements

- ▶ Richard Davies
- ▶ Paul Ayris
- ▶ Rui Miao
- ▶ Helen Shenton
- ▶ Paul Wheatley

01101101
 01010101
 10011101
 00110110
 10101010
 11001110
 10011011
 01010101
 01100111
 01001101
 10101010
 10110011
 10100110
 11010101
 01011001
 11010110

e rory.mcleod@bl.uk
 t +44 (0) 20 7412 7182



LIFE's broader strategic role

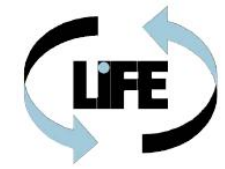
Blue Ribbon Task Force
on Sustainable Digital Preservation and Access



*This is the only group I know of
that is chartered to help the community understand
the economic issues surrounding sustainable
repositories and identify candidate solutions*

- Lucy Nowell,
Program Director
Office of Cyberinfrastructure, NSF

See <http://blueribbontaskforce.sdsc.edu/>

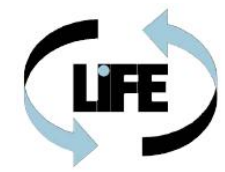


Blue Ribbon Task Force on Sustainable Digital Preservation and Access

BRTF-SDPA

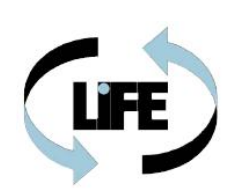
Funded by the National Science Foundation (NSF) and the Andrew W. Mellon Foundation, in partnership with the Library of Congress, the JISC (UK), the Council on Library and Information Resources (CLIR), and the National Archives and Records Administration (NARA – US)

Created in late 2007 to run for 2 years



4 Key Questions being addressed

1. How will we ensure the long-term preservation and access to our digital information, growing exponentially with each passing day?
2. How will we successfully migrate data as technology moves from one preservation medium to the next?
3. Who should determine which digital data should be saved, and what criteria will be used to make those decisions?
4. How do we ensure economic (and digital) sustainability



Economic Sustainability in a Digital Preservation Context- The LIFE aim.

The set of business, social, technological, and policy mechanisms:

- ▶ Encourage the gathering of important information assets into digital preservation systems
- ▶ Support the indefinite persistence of digital preservation systems, thus securing access to and use of information assets into the long-term future

Economically-sustainable digital preservation requires:

- ▶ Recognition of the benefits of preservation by key decision makers, as part of a process of selecting digital materials for long-term retention
- ▶ Appropriate incentives to induce decision makers to act in public interest
- ▶ Mechanisms to secure an ongoing allocation of resources, both within and across organizations, to digital preservation activities
- ▶ Efficient use of limited preservation resources
- ▶ Appropriate organization and governance of digital **preservation activities**