Co-funded by the European Union under FP7-ICT-2009-6

APA RSEN

Alliance Permanent Access to the Records of Science in Europe Network

Results of analysis and testing of cost models

Kirnn Kaur, kirnn.kaur@bl.uk
The British Library

iPres 2013 Workshop, Lisbon, 6th September 2013

Co-funded by the European Union under FP7-ICT-2009-6

Introduction

- APARSEN project overview
- Survey
- Cost model analysis
- Cost model testing
- Future of cost modelling
- Further information

Co-funded by the European Union under FP7-ICT-2009-6

APARSEN PROJECT

- A **Network of Excellence** in digital preservation
- Objective is to set up a **Virtual Centre of Excellence**
- Funded by European Commission
  - 7th Framework Programme - Digital Libraries and Digital Preservation
- January 2011 to December 2014 (4 years)
- Coordinated by Science and Technology Facilities Council (UK)
- **31 partners from 13 countries**

APARSEN Network of Excellence

Alliance Permanent Access to the Records of Science in Europe Network

Science & Technology Facilities Council

stm

UNIVERSITÀ DEGLI STUDI DI TRENTO

UK • DATA ARCHIVE

In|Con|Tec

Data Archiving and Networked Services
DANS

FTK Forschungsinstitut für Telekommunikation und Kooperation

AIRBUS

BRITISH LIBRARY

INES Centre Informatique National de l'Enseignement Supérieur

KB Koninklijke Bibliotheek National Library of the Netherlands

inmark estudiosyestrategias

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ UNIVERSITY OF PATRAS

HELMHOLTZ ASSOCIATION

esa

LIBER LIBER

cini

Fondazione RINASCIMENTO digitale nuove tecnologie per i beni culturali

Tessella Technology & Consulting

DEUTSCHE NATIONAL BIBLIOTHEK

LULEÅ TEKNISKA UNIVERSITET

DigitalPreservationCoalition

ИКИ

secure sba-research.org

APA ALLIANCE FOR PERMANENT ACCESS

GLOBIT

FORTH Institute of Computer Science

IBM

CSC

Österreichische Nationalbibliothek

Co-funded by the European Union under FP7-ICT-2009-6

APARSEN PROJECT

APARSEN defines four topics in which it undertakes research in digital preservation:

- **Trust**
- **Sustainability**
- **Usability**
- **Access**

Co-funded by the European Union under FP7-ICT-2009-6

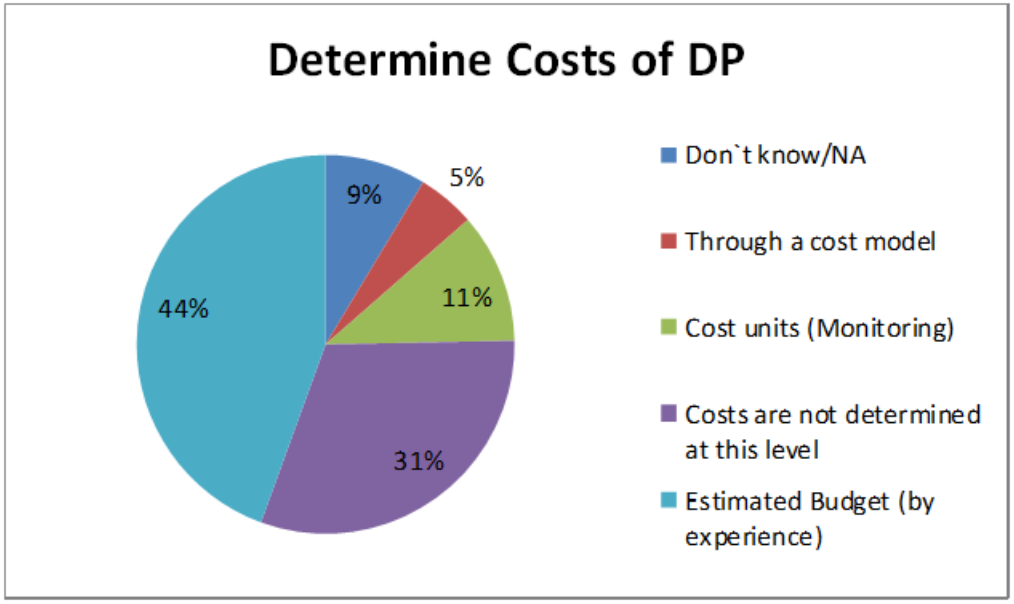
Cost model survey: methodology

- European-wide online survey
- Research libraries (members of LIBER)
- Conducted in collaboration with Business Cases WP
- Questions specifically related to cost models included
- Sample size approx. 100

Co-funded by the European Union under FP7-ICT-2009-6

Cost model survey: highlights (1 of 3)

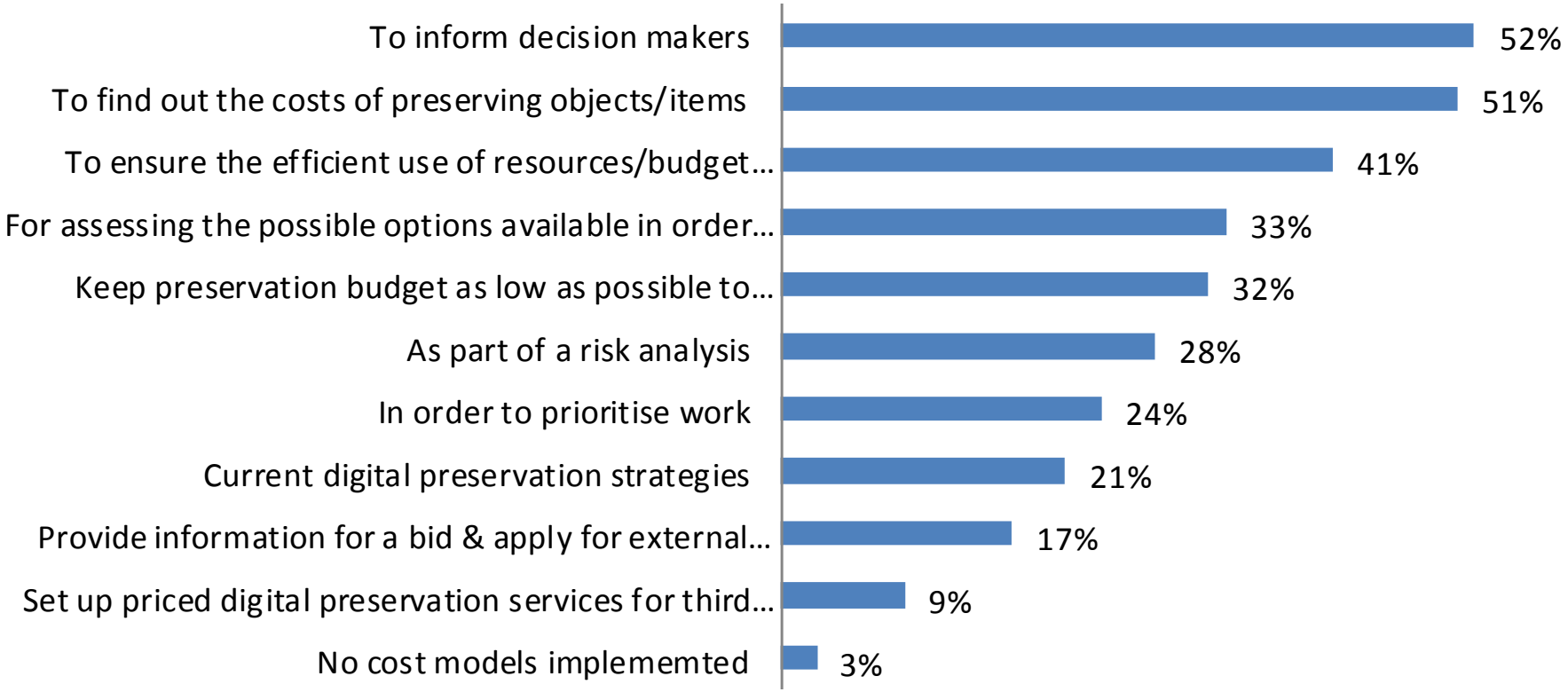
How do you determine the costs of Digital Preservation within your organisation?



Co-funded by the European Union under FP7-ICT-2009-6

Cost model survey: highlights (2 of 3)

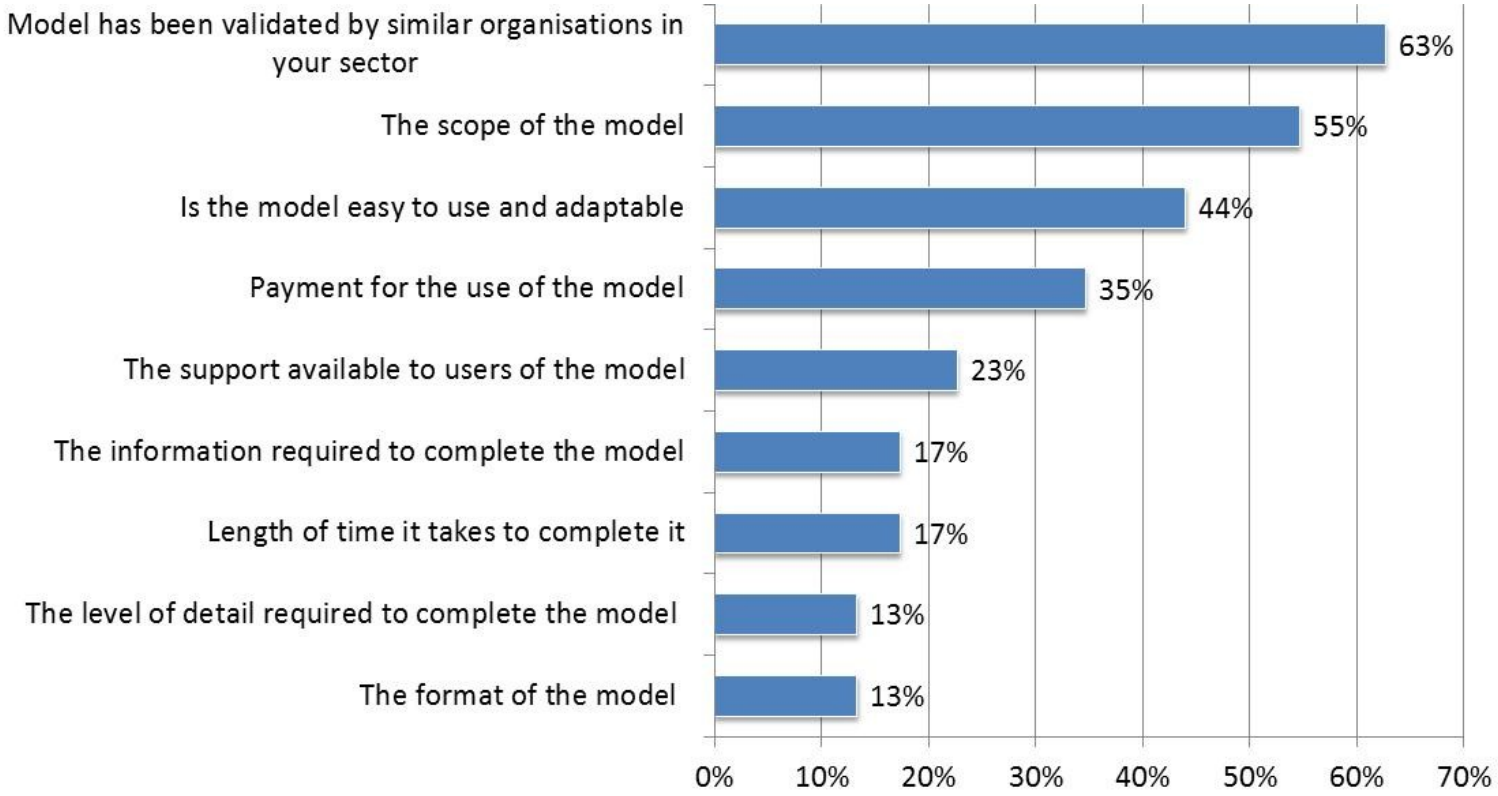
Reasons for using a cost model?



Co-funded by the European Union under FP7-ICT-2009-6

Cost model survey: highlights (3 of 3)

On what basis would you select a cost model?



Co-funded by the European Union under FP7-ICT-2009-6

Cost model survey: results

Overview: Limited use of cost models although advantages widely accepted

- **Funding** – cost models provide **cost predictions**
- **Management** – involve **key stakeholders** in future development of cost models
- **Benefits and added value** – cost models can provide **tools**
- **Controlling costs** – very **limited use** although **good understanding** of the reasons for use of cost
- **Cost model requirements** – **validated**, cover the digital preservation **lifecycle**, **easy to use** and **freely available**

Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis – methodology

Analysis of cost models

- To determine **scope**, **access**, **strengths** and **weaknesses** of published cost models

Mapping of cost parameters

- ISO16363 to ascertain **concentration** of parameters and identify **gaps** and areas for **further investigation** and **development**
- Assessing parameters against the activities audited would show whether cost models included specific **functional elements** of the ISO

Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis: scope

CMDP - Cost Model for Digital Preservation
- Developed by the Royal Library of Denmark and the Danish National Archives

DANS cost model
- Developed by DANS, Data Archiving and Network Services, Netherlands

DP4lib - Digital Preservation for libraries
- Developed by the DNB

ENSURE project
- Being developed by EC FP7 project, ENSURE (Feb 11 – Jan 14) http://ensure-fp7-plone.fe.up.pt/site

ISIS facility model
- Developed as part of Cranfield University MSc project in collaboration with STFC

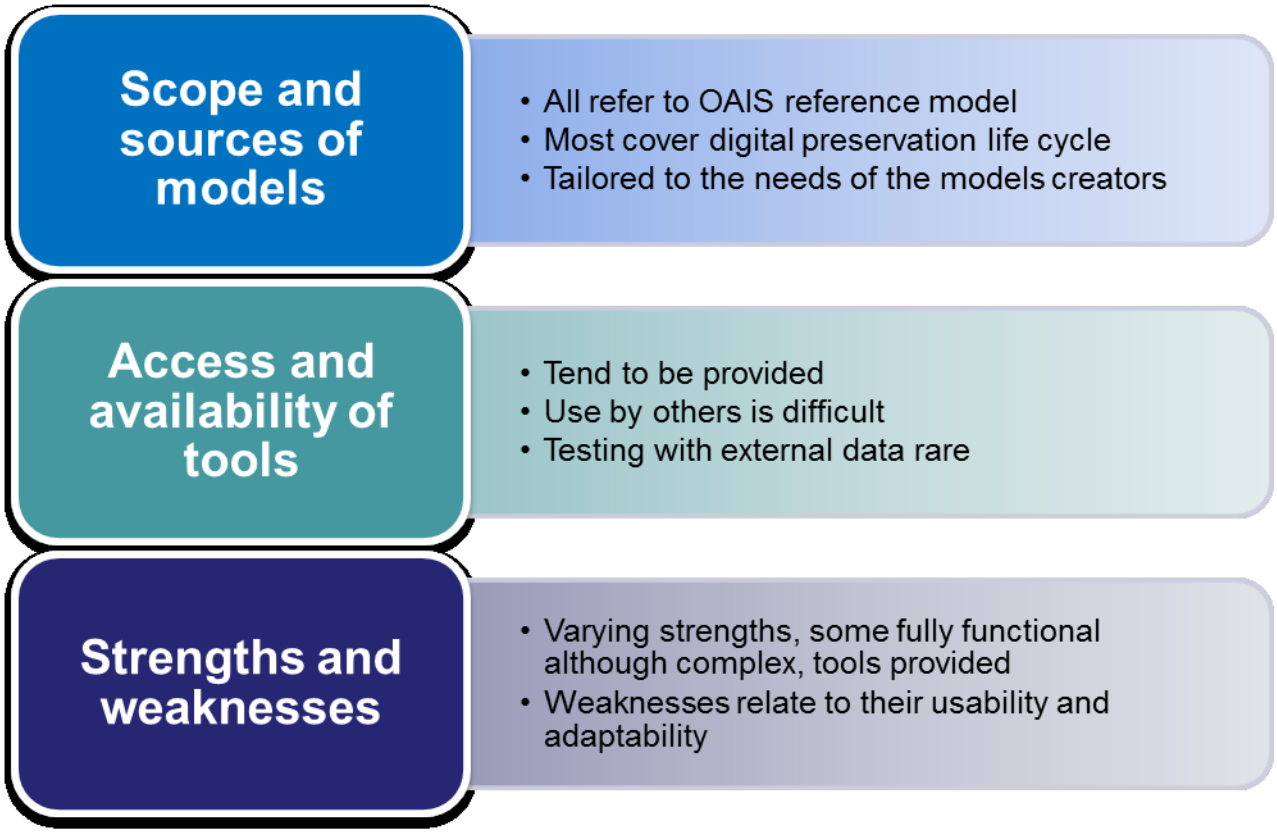
KRDS – Keeping research data safe (KRDS + KRDS 2)
- Development of toolkits funded by JISC partners in project include Charles Beagrie Ltd, UKOLN, DCC, UCL, UKDA, ADS, OCLC

LIFE^3 – Life Cycle Information for E-literature
- Developed by UCL and BL, project funded by JISC and RIN

Presto PRIME – cost model for digital storage
- Developed within EC FP7 project http://www.prestoprime.eu/

Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis: high level



Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis: cost paramaters

Benchmarked against ISO16363:

- cost parameters mapped under specific headings
- comparisons could be made across the models
- similarities and differences can be analysed

Gaps show areas:

- where parameters may not be needed as out of scope or not measurable
- for further investigation
- for development

Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis: gap analysis (1 of 3)

- Inconsistent coverage of parameters across all areas
- KRDS provides best coverage

Organisational Infrastructure						
Cost Model	Sub-heading	Governance, organisational viability	Organisational structure and staffing	Procedural accountability, preservation policy	Financial sustainability	Contracts, licenses and liabilities
CMDP			X	X		
DANS			X			X
DP4lib			X	X	X	X
ENSURE						
ISIS						
KRDS		x	X	X	X	X
LIFE3			X	X	X	X
PrestoPRIME		x		X	X	

Key: x = cost parameter partially mapped
X = cost parameter fully mapped

Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis: gap analysis (2 of 3)

- Consistent coverage across most areas, as would be expected for cost models related to digital preservation
- LIFE3 and ENSURE models provide parameters across all areas

Digital Object Management							
Cost Model	Sub-heading	Ingest: acquisition of content	Ingest: creation of AIP	Preservation planning	AIP preservation	Information management	Access management
CMDP		X	X	X	X		x
DANS		X	X	X		X	
DP4lib		X	X		X	X	X
ENSURE		X	X	X	X	X	X
ISIS		X	X	X	X	X	
KRDS		X	x	X		X	X
LIFE3		X	X	X	X	X	X
PrestoPRIME		x	x			x	X

Key: x = cost parameter partially mapped
X = cost parameter fully mapped

Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis: gap analysis (3 of 3)

- Good coverage across the sections, although not consistent
- Every cost model has cost parameters which match at least one of the two subheadings

Infrastructure and security risk management			
Cost Model	Sub-heading	Technical infrastructure risk management	Security risk management
CMDP		X	X
DANS			X
DP4lib		X	
ENSURE		X	
ISIS		X	
KRDS		X	
LIFE3		X	X
PrestoPRIME		X	X

Key: x = cost parameter partially mapped
X = cost parameter fully mapped

Co-funded by the European Union under FP7-ICT-2009-6

Cost model analysis – summary

Mapping of cost parameters

- Parameter definitions not easily aligned
- Areas of standard may not be realistically measurable through a cost parameter

Areas for further investigation and development

- In some cases the **gaps** identified can be **justified** due to the scope of the cost model
- Areas requiring **additional focus** identified
- For cost models to become more **useful to a wider audience** – cost parameter definitions should be provided in a clear, concise and understandable form

Co-funded by the European Union under FP7-ICT-2009-6

Cost model testing: methodology

Cost model selection criteria

- In-depth **knowledge** and **understanding** of the cost model
- Ideal situation where **cost model owner/creator** involved e.g. **DANS**, **DP4lib, LIFE3** selected as data and case studies used to set up the model were available

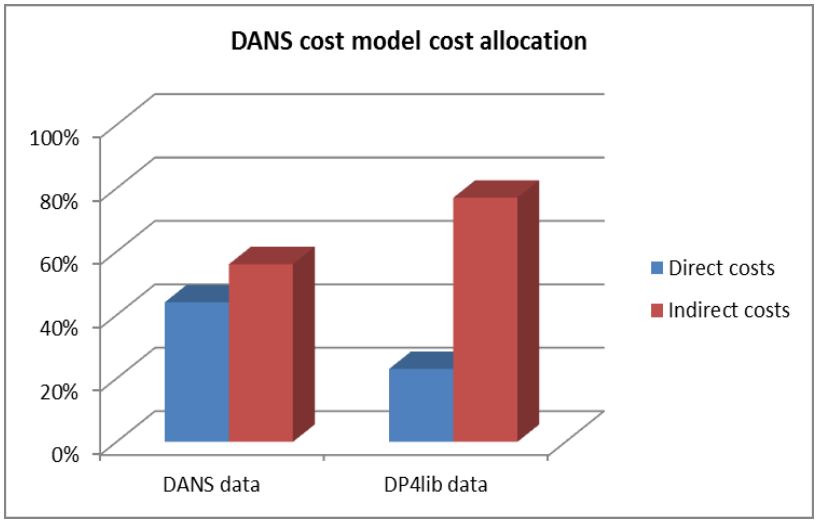
Test data

- Test data was used to carry out **cost allocations** across the models using various assumptions for **apportionment** of costs
- To test the **workability** of a model or its **flexibility** for adaptation to other costs of services and workflows
- **Transferability** of costs to another model would provide insights into how **'usable'** the cost model would be
- Where data entry not possible this informs **development** of cost models for owners and those wishing to set up their own cost models

Co-funded by the European Union under FP7-ICT-2009-6

Cost model testing: DANS with DP4lib data

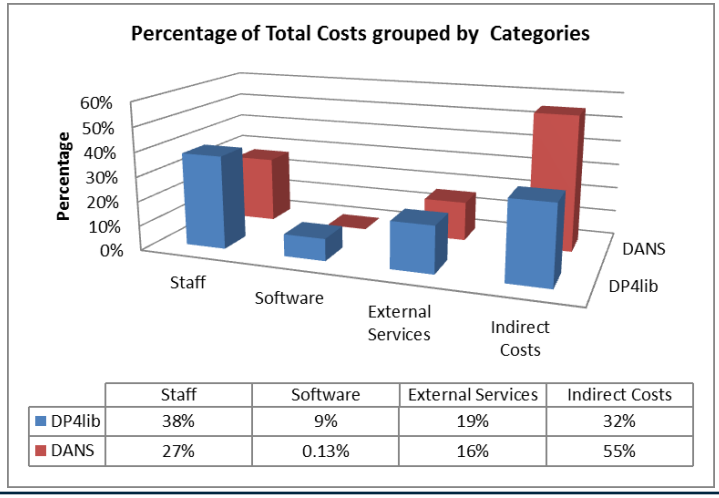
- Overall division of **direct** and **indirect costs** is very different (see below)
- The differences may be due to services, activities and related costs , or because **cost categories** and **allocations** are **not fully compatible**
- **Comparison** between the two models was **possible** at a high level
- Generally, when using another cost model, it is necessary to make **explicit** all the **underlying assumptions** of the model and test data
- **Testing** cost data in another model **provides valuable insights** into both models as well as to the contexts in which these cost models were created



Co-funded by the European Union under FP7-ICT-2009-6

Cost model testing: DP4lib with DANS data

- Only **45%** of all costs were able to be **allocated** as the DANS data **indirect cost** category was so high
- DP4lib allows for **other cost categories** (indirect costs), providing 100% apportionment of all DANS at a high level
- **Comparison** at high level **feasible** as both models are based on **activity based costing** allowing allocation of costs between the two models
- For any **future testing** or validation it is recommended that data from within the organisation, where the **same processes** have been implemented, be used for this exercise
- Where **external data** is used a **clear understanding** of the test data is required including any **assumptions** made in relation to cost allocations. It may be a useful exercise to look at the data being used by another organisation within the context of these models as lessons can be learned about how costs are measured, distributed and allocated

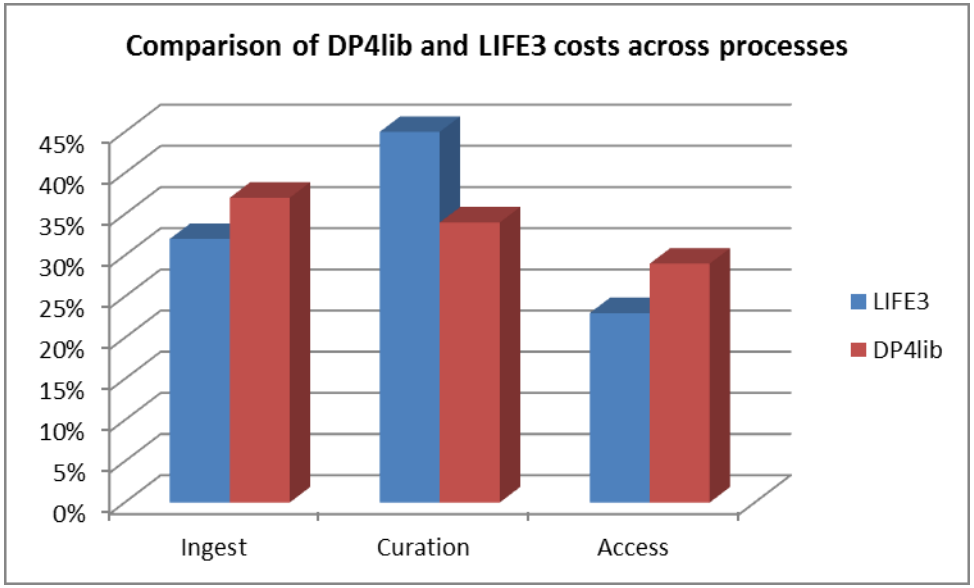


	Staff	Software	External Services	Indirect Costs
DP4lib	38%	9%	19%	32%
DANS	27%	0.13%	16%	55%

Co-funded by the European Union under FP7-ICT-2009-6

Cost model testing: LIFE3 with DP4lib data

- 67% cost allocation which left 34% of costs unallocated from the DP4lib model as no clear match between the processes or activities
- When comparing the allocation of costs across the three processes of ingest, curation and access – they differ across the two models
- One third of costs could not be allocated due to processes not being the same across the organisations for which the models were created
- The LIFE3 model could be simplified and additional case study data would improve the model



Co-funded by the European Union under FP7-ICT-2009-6

Cost model testing: summary

- **Overview**: A common basis for comparison and apportionment of costs although possible proved to be a difficult exercise
- Good **understanding of costs** as well as their breakdown is needed
- Cost models are quite **specific** to the organisations where they were created in terms of **activities**, **services** and **workflows**
- **DANS** and **DP4lib** models still under development – useful tools for developing 3rd party preservation services as additional revenue streams for repositories
- **LIFE 3** provides detailed lifecycle costing of DP workflows

Co-funded by the European Union under FP7-ICT-2009-6

Future of cost modelling

- Further development
- Advice and guidance needed
- Benchmarking to compare specific activities
- Usable, adaptable models

Co-funded by the European Union under FP7-ICT-2009-6

Further information

- Further work within APARSEN on sustainability, common vision and VCoE
- Project website www.aparsen.eu
- Reports
  - D32.1 Report on cost parameters for digital repositories (Feb 2013)
  - D32.2 Report on testing of cost models and further analysis of cost parameters (June 2013, with PO)
  - Sustainability booklet (approx. Sept 2013)