Collaboration to Clarify the Cost of Curation

D2.1—Baseline Study of Stakeholder & Stakeholder Initiatives

Deliverable Lead:	KEEP SOLUTIONS
Related Work package:	WP2 — Engagement
Author(s):	Miguel Ferreira (KEEPS), Luís Faria (KEEPS), Hélder Silva (KEEPS)
Dissemination level:	Public
Submission date:	31st of July 2013
Project Acronym:	4C
Website:	http://4cproject.eu
Call:	FP7-ICT-2011-9
Project Number	600471
Instrument:	Coordination action (CA)—ERA-NET
Start date of Project:	01 Feb 2013
Duration:	24 months

Project funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Version History

Version	Date	Changed pages / reason	Modified by
0.1	2013-03-18	Document structure defined	Luís Faria
0.2	2013-07-04	Content added to the document. First draft for review.	Miguel Ferreira
0.3	2013-07-09	Version shared with partners in WP2	Miguel Ferreira
0.4	2013-07-10	Registry of relevant initiatives added to the text	Miguel Ferreira, Hélder Silva
0.5	2013-07-11	Internal review	Hélder Silva
0.6	2013-07-17	Internal review	Joy Davidson
0.7	2013-07-22	Internal review	Magdalena Getler
0.8	2013-07-23	Internal review + Stakeholder characterisation	Katarina Haage
0.9	2013-07-24	Internal Review	Patrick McCann
1.0	2013-07-31	Release version	Paul Stokes

Acknowledgements

This report has been developed within the project "Collaboration to Clarify the Cost of Curation" (4cproject.eu). The project is an ERA-NET co-funded by the 7th Framework Programme of the European Commission.

The 4C participants are:

Participant organisation name	Short Name	Country
Jisc	JISC	UK
Det Kongelige Bibliotek, Nationalbibliotek Og Kobenhavns Universitetsbibliotek	KBDK	DK
Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa	INESC-ID	PT
Statens Arkiver	DNA	DK
Deutsche Nationalbibliothek	DNB	DE
University of Glasgow	HATII-DCC	UK
University of Essex	UESSEX	UK
KEEP SOLUTIONS LDA	KEEPS	PT
Digital Preservation Coalition Limited by Guarantee	DPC	UK
Verein Zur Forderung Der It-Sicherheit In Osterreich	SBA	AT
The University of Edinburgh	UEDIN-DCC	UK
Koninklijke Nederlandse Akademie van Wetenschappen -Knaw	KNAW-DANS	NL
Eesti Rahvusraamatukogu	NLE	EE

Author(s):	Miguel Ferreira (KEEPS), Luís Faria (KEEPS), Hélder Silva (KEEPS)

Table of Contents

Figures

Executive Summary

To initiate the Engagement work package in 4C, a baseline group of stakeholders was identified and an analysis of significant cost modelling and economics-related work in the field of digital curation has been carried out. Also, a small questionnaire has been sent to stakeholders in order to engage them in the project and to better understand their current state of practice in assessing digital curation costs.

As such, this document reports on task 2.1 of the 4C project, i.e. Baseline study of stakeholders and initiatives on the domain of digital curation costs; and includes the results of the following subtasks:

1. A collection of relevant work on cost modelling activities in the context of digital curation;
2. An initial registry of stakeholder groups and contacts;
3. The results of the application of a questionnaire sent to stakeholders to grasp the state of practice and current needs in the field of digital curation costs.

1 Introduction

The Engagement work package of the 4C project aims at identifying, involving and building partnerships with individuals, groups and institutions that have a particular interest in economic issues relating to digital curation. Based on the outcomes of this initial consultation with stakeholders' groups, the Engagement work package will identify and provide channels (virtual and physical) for interaction between the various project members and representatives from these groups. The impact and success of these interactions will be gauged initially by the number of those indicating their willingness to participate in the survey and from the evaluation and feedback received from those participants.

The Engagement work package will also review relevant work on cost modelling in the area of digital curation and maintain an up-to-date registry describing these initiatives. Furthermore, this work package will refine and disseminate the outputs of the project by tailoring them towards specific stakeholder audiences.

This deliverable will inform the creation of an effective on-going register of relevant work and stakeholders, which will be maintained and updated throughout the duration of the project (mainly by task 2.2) and also on the results of an early consultation made to stakeholders.

2 Relevant initiatives on cost modelling for digital curation

This section provides an initial registry of recent cost modelling and economics-related work in, or relevant to, the field of digital curation. The report includes information about on-going and past projects, publications, presentations, posters and events.

For each type of information, different metadata has been extracted and documented. For instance, projects are characterised by distinct metadata fields, while publications are depicted as citation references in the APA formatting style. Whenever possible, links to the original works have also been included in the index.

An online version of this registry will also be made available on the project's Website as soon as an adequate platform to support it is set up. The registry will be kept up-to-date throughout the lifetime of the project (mainly by task 2.2).

This section is organised in 5 subsections each of which reporting on a particular type of information:

1. **Projects** – Includes research or other kind of projects limited in time that focus partially or entirely on the topic of digital curation costs;
2. **Publications** – Includes published works in the form of technical reports, journal articles and conference papers;
3. **Presentations** – Includes public talks focused on digital curation costs;
4. **Posters** – Includes posters presented on events;
5. **Events and event reports** – Includes a list of events such as conferences, workshops or any type of public gathering of people and also relevant information about those events such as event reports, participants lists, webcasts, etc.;

The information included in the following sections has been obtained from the projects' websites or their funding agencies. Parts of the descriptions included bellow are verbatim copies of the information published on those sites.

2.1 Projects

This section reports on research or other kind of projects limited in time that focus partially or entirely on the topic of digital curation costs.

4C—Collaboration to Clarify the Cost of Curation

Project acronym	4C
Project title	Collaboration to Clarify the Cost of Curation
Description[1]	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. These objectives will be achieved by a coordinated programme of outreach and engagement that will identify existing and emerging research and analyse user requirements. This will inform an assessment of where there are gaps in the current provision of tools, frameworks and models. The project will support stakeholders to better understand and articulate their requirements and will clarify some of the complexity of the relationships between cost and other factors. The outputs of this project will include various stakeholder engagement and dissemination events (focus groups, workshops, a conference), a series of reports, the creation of models and specifications, and the establishment of an international Curation Costs Exchange framework. All of this activity will enable the definition of a research and development agenda and a business engagement strategy which will be delivered to the European Commission in the form of a roadmap. The consortium undertaking this project includes organisations with extensive domain expertise and experience with curation cost modelling issues. It includes national libraries and archives, specialist preservation and curation membership organisations, service providers, research departments and SME's. It will be coordinated by a national funding organisation that specialises in supporting the innovative use of ICT methods and technologies.
Start & end date	From 2013-02-01 to 2015-01-31 (24 months)
Subjects	Cost modelling of curation processes
Coordinator	Jisc (UK)
Participants	Jisc (UK) (Project Co-ordinator), Danish National Archives, DANS – Data Archiving and Networked Services (KNAW/NWO) (Netherlands), Deutsche Nationalbibliothek (Germany), Digital Curation Centre – University of Edinburgh (UK), Digital Preservation Coalition (UK), Humanities Advanced Technology and Information Institute - University of Glasgow (UK), INESC-ID – Institute for System and Computer Engineering (Portugal), KEEP SOLUTIONS (Portugal), National Library of Estonia (Estonia), The Royal Library – National Library of Denmark (Denmark), Secure Business Austria (Austria), University of Essex (UK)
Website	http://4cproject.eu
Funding	FP7-ICT, project 600471 (Coordination and support actions)

[1] http://4cproject.net/overview/

APARSEN—Alliance Permanent Access to the Records of Science in Europe Network

Project acronym	APARSEN
Project title	Alliance Permanent Access to the Records of Science in Europe Network
Description[2]	Digital preservation offers the economic and social benefits associated with the long-term preservation of information, knowledge and know-how for re-use by later generations. However, digital preservation has a great problem, namely that preservation support structures are built on projects which are short lived and fragmented. The unique feature of APARSEN is that it is building on the already established Alliance for Permanent Access (APA), a membership organisation of major European stakeholders in digital data and digital preservation. These stakeholders have come together to create a shared vision and framework for a sustainable digital information infrastructure providing permanent access to digitally encoded information. To this self-sustaining grouping APARSEN will bring a wide range of other experts in digital preservation including academic and commercial researchers, as well as researchers in other cross-European organisations. The members of the APA and other members of the consortium already undertake research in digital preservation individually but even here the effort is fragmented despite smaller groupings of these organisations working together in specific EU and national projects. APARSEN will help to combine and integrate these programmes into a shared programme of work, thereby creating the pre-eminent virtual research centre in digital preservation in Europe, if not the World. The APA provides a natural basis for a longer term consolidation of digital preservation research and expertise. The Joint Programme of Activity will cover: - technical methods for preservation, access and most importantly re-use of data holdings over the whole lifecycle; - legal and economic issues including costs and governance issues as well as digital rights; - outreach within and outside the consortium to help to create a discipline of data curators with appropriate qualifications;
Start & end date	From 2011-01-01 to 2014-12-31 (48 months)
Subjects	Information, Media, Innovation, Technology Transfer
Coordinator	Science and Technology Facilities Council (UK)
Participants	University of Essex (UK), STFC (UK), Alliance Permanent Access (NL), CERN (CH), International Association of Scientific, Technical and Medical Publishers (NL), Forschungsinstitut für Telekommunikation (DE), CSC - Tieteen tietotekniikan keskus Oy (FI), German National Library (DE), Digital Preservation Coallition (UK), AFPUM (DE), British Library (UK), European Space Agency (FR), KNAW-DANS (NL), Netherlands National Library (NL), The Stichting LIBER Foundation (NL), CINI - Consorzio Interuniversitario Nazionale per l'Informatica (IT), ICT - InConTec GmbH (DE), FORTH - Foundation for Research and Technology-Hellas (GR), GLOBIT - Globale Informationstechnik GmbH (DE), Microsoft Research Limited (UK), Philips Consumer Lifestyle (NL), Airbus Operations SAS (FR), INMARK Estudios y Estrategias (ES), Fondazione Rinascimento Digitale (IT), Luleå University of Technology (SE), University of Trento (IT), Tessella (UK), IBM Israel (IL), Secure Business Austria (SBA), Space Research Institute of the Russian Academy of Sciences (RU), Austrian National Library (AT), University of Patras, Library & Information Center (GR), University of Essex (UK), CINES (FR)
Website	http://www.alliancepermanentaccess.org/index.php/aparsen/
Funding	FP7-ICT, project 269977 (Networks of Excellence)

[2] http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=11743835

ENSURE—Enabling kNowledge Sustainability Usability and Recovery for Economic value

Project acronym	ENSURE
Project title	Enabling kNowledge Sustainability Usability and Recovery for Economic value
Description[3]	Ensuring long-term usability for the spiralling amounts of data produced or controlled by organizations with commercial interests is quickly becoming a major problem. Drawing on motivation from use cases in aerospace, health care, finance and clinical trials, ENSURE will significantly extend the state of the art in digital preservation which to-date has focused on relatively homogeneous cultural heritage data. Our use cases bring up a large number of issues which have yet to be fully addressed: 1) safely leveraging scalable pay-as-you-go infrastructure such as clouds 2) having businesses understand the economic implications of preservation, 3) conforming to regulatory, contractual and legal requirements as part of a whole workflow 4) managing long term integrity and authenticity significant intellectual property or highly personal data and 5) using off-the-shelf IT technologies for preservation to support different types of digital resources. Building on prior work, ENSURE will address these issues with innovative approaches and tools: Cost and Value Evaluate the cost and benefit of different quality solutions. Preservation Lifecycle Management Build on industry standard lifecycle management approaches to manage the preservation lifecycle, ensuring regulatory compliance, allowing changes in the preservation approach to reflect environmental changes, addressing evolution of ontologies and managing the quality of the digital objects over time. Content-Aware Long Term Data Protection Ensure long-term, content-aware data protection, addressing changes in personally identifiable information, new and evolving regulations, managing user identities over decades, etc. Leveraging Wider ITC Evaluate the costs/risks/benefits and demonstrate how to use emerging, commonly available information technology, to enable scalable solutions for digital preservation, considering in particular cloud storage and virtual application image capture.
Start & end date	From 2011-02-01 to 2014-01-31 (48 months)
Subjects	Evaluate costs and technologies to provide preservation and protection of data in use cases outside the cultural heritage domain.
Coordinator	BM ISRAEL - SCIENCE AND TECHNOLOGY LTD
Participants	JRC Capital Management Consultancy & Research GMBH (DE), Philips (NL), Fraunhofer-Gesellschaft zur Foerderung der Angewandten Forschung E.V (DE), Atos Spain (ES), Lulea Tekniska Universitet (SE), Maccabi Healthcare Services (IS), Centro Superior de Investigacion en Salud Publica (ES), Universidade do Porto (PT), Tessella (UK), Cranfield University (UK), Custodix NB (BE), STFC (UK)
Website	http://ensure-fp7-plone.fe.up.pt/site/
Funding	FP7-ICT, ICT-2009.4.1, project ref. 270000 (Collaborative project)

[3] http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=11841469

TCP—Total Cost of Preservation

Project acronym	TCP
Project title	Total Cost of Preservation
Description[4]	Information technology and resources are thoroughly integrated with, and indispensable to, today's web-based culture, commerce, science, education, and entertainment. The digital assets underpinning those activities, however, are inherently fragile with respect to ever increasing disruptive technological change. Without effective and affordable curation management, today's digital assets will not remain viable and useful in the future. To address this concern, UC3 has developed an analytical framework for modelling the full economic costs of preservation, the "total cost of preservation" (TCP).
Start & end date	From 2011 to today (on-going)
Subjects	Cost model for digital preservation
Coordinator	University of California Curation Center (UC3)
Participants	University of California Curation Center (UC3)
Website	https://wiki.ucop.edu/display/Curation/Cost+Modeling
Funding	Self-funded by the University of California Curation Center (UC3)

CMDP—Cost Model for Digital Preservation

Project acronym	CMDP
Project title	Cost Model for Digital Preservation
Description[5]	The concrete objective is to develop a tool that calculates present and future costs of cultural heritage institutions' digital collections based on various user inputs, such as the amount and type of data. The project has finished 3 phases, which are: 1) Preservation Planning and digital migration, 2) Ingest and 3) Archival Storage.
Start & end date	CMDP 1: In 2009 CMDP 2: In 2010 CMDP 3: In 2012
Subjects	Cost model for digital preservation
Coordinator	N/A
Participants	Royal Danish Library and the Danish National Archives
Website	http://www.costmodelfordigitalpreservation.dk/
Funding	Funded by the Royal Danish Library and the Danish National Archives

[4] https://wiki.ucop.edu/display/Curation/Cost+Modeling

[5] http://www.costmodelfordigitalpreservation.dk

PrestoPRIME—Keeping Audiovisual contents Alive

Project acronym	PrestoPRIME
Project title	PrestoPRIME - Keeping Audiovisual contents Alive
Description[6]	Audiovisual content collections are undergoing a transformation from archives of analogue materials to very large stores of digital data. As time-based digital media and their related metadata are edited, re-used and re-formatted in a continuously evolving environment, the concept of the unique original loses its meaning and we require dynamic processes that can preserve indefinitely not only the audiovisual signal but also its evolving associations, context and rights. PrestoPRIME will research and develop practical solutions for the long-term preservation of digital media objects, programmes and collections, and find ways to increase access by integrating the media archives with European on-line digital libraries in a digital preservation framework. This will result in a range of tools and services, delivered through a networked Competence Centre. The project will deliver a preservation framework, complete with risk management and content quality and corruption control measures, capable of supporting audiovisual signal migration and multivalent preservation methods using federated services for distributing and storing content. It will create a metadata conversion and deployment toolkit, with a novel and efficient process for metadata vocabulary alignment, annotation and services for user-generated content metadata. A rights management system and audiovisual fingerprint registry will make it possible to track and manage content at all stages of its lifecycle, in all contexts of use The project will demonstrate and evaluate an integrated prototype of the preservation Framework and software in the networked Competence Centre. The Competence Centre and the European Association for Audiovisual Archives will be established to provide business models, registry and best practice services and training.
Start & end date	From 2009-01-01 to 2012-11-30
Subjects	Education, Training, Information, Media, Information Processing, Information Systems, Telecommunications
Coordinator	Institut National de L'Audiovisuel (FR)
Participants	British Broadcasting Corporation (UK), Stichting Nederlands Institut voor Beeld en Geluid (NL), University of Liverpool (UK), EURIX (IT), Joanneum Research Forschungsgesellschaft (AT), Highlands Technologies (FR), Osterreichischer Rundfunk (AT), University of Southampton (UK), Ex Libris (IS), Stichting Europeana (NL), RAI-Radiotelevisione Italiana (IT), Vereniging voor Christelijk Hoger Onderwijs Wetenschappelijk Onderzoek en Patientenzorg (NL), Universitaet Innsbruck (AT)
Website	http://www.prestoprime.org http://prestoprime.it-innovation.soton.ac.uk/
Funding	Funded by FP7-ICT-231161 (Collaborative project)

[6] http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=10389705

Costs of Digital Archiving vol. 2

Project acronym	Costs of Digital Archiving vol. 2
Project title	Costs of Digital Archiving vol. 2
Description[7]	This project aims at generating a cost model for archiving and disseminating digital scholarly datasets relevant to the circumstances of DANS. It is a follow-up of the work done by Kevin Heerema and Anna Palaiologk, which resulted in a first draft cost model based mainly on the OAIS reference model. It is anticipated that the new and improved model will assist the management of DANS in achieving economic sustainability. Material costs need more attention, but the most challenging task in the new project is implementation of the model described within the organisation, in particular in the current time reporting systems and other planning tools. For public sector organisations the term sustainability has to be examined through the prism of fixed budget. In the case of data archiving entities, like DANS, the challenge lies in the fact that there are continuous, disproportionate increases in both the quantity of data and the complexity of datasets relative to the funding. Further obstacles to accurate forecasting arise from frequently changing preservation strategies. A cost model is an essential tool for effective cost management which, along with the other elements of a proper business model, ensures economic sustainability. The model that is being developed in DANS is an ABC Activity Based Cost model as salary is the main cost driver. All DANS activities which surround the core archiving process are taken into consideration and combined using formulae. The variables used determine the estimated costs per dataset. For DANS a dataset is a collection of digital objects coming from a single research project and includes documentation describing the data and their relation. In this project a special matrix is used to rank the complexity of a dataset (formats, number of files, size, metadata, etc.). Finally, it is used to examine the influence of the dataset complexity to the costs. The ABC Model is applied in combination with the BSC Balanced Scorecard Method. This method translates the mission of an organisation and the existing business strategy into a limited number of specific strategic objectives that can be linked and measured operationally. This enables us to balance the costs of the various strategic objectives of DANS. The main areas for attention of the project are the relevance of the generated cost model to the reality of archiving practice, its usability (employees understand it and managers can effectively use it) and its accuracy (no discrepancies with reality). The richness of the problem offers many interesting directions in which to expand the scope of the model.
Start & end date	From 2008 to 2011
Subjects	Cost model for archiving and disseminating scholarly datasets
Coordinator	Data Archiving and Networked Services (NL)
Participants	N/A
Website	http://www.dans.knaw.nl/en/content/categorieen/projecten/costs-digital-archiving-vol-2
Funding	Funded by DANS

[7] http://www.dans.knaw.nl/en/content/categorieen/projecten/costs-digital-archiving-vol-2

DP4lib—Digital Preservation for Libraries

Project acronym	DP4lib
Project title	Digital Preservation for Libraries
Description[8]	The project Digital Preservation for libraries (DP4lib) project is funded by the Deutsche Forschungsgemeinschaft (DFG). The goal of this project is to evaluate the feasibility of all options for establishing and running a ready-to-operate service for long-term preservation (LTP). In addition, the preceding conceptual work will be implemented in a piece of prototype software.
Start & end date	From 2009 to 2012
Subjects	Service for long-term preservation
Coordinator	N/A
Participants	Deutsche Nationalbibliothek (DE), SUB Göttingen (DE)
Website	http://dp4lib.langzeitarchivierung.de
Funding	Funded by Deutsche Forschungsgemeinschaft (DFG)

KRDS—Keeping Research Data Safe

Project acronym	KRDS
Project title	Keeping Research Data Safe
Description[9]	Keeping Research Data Safe has been developed in three major phases funded by the Joint Information Systems Committee. The first Keeping Research Data Safe study (KRDS1) completed in 2008 made a major contribution to the study of preservation costs by developing a cost model and identifying cost variables for preserving research data in UK universities. That work has had considerable impact and received international interest. The second Keeping Research Data Safe project (KRDS2) completed in December 2009, built on this previous work and identified and analysed longitudinal data on preservation costs and benefits associated with long-lived data. The final phase has focussed on transferring knowledge from the research into practice through development of a Factsheet, User Guide, and Benefits Analysis Toolkit.
Start & end date	from 2007 to 2011
Subjects	Cost model and benefits analysis for preserving research data
Coordinator	Charles Beagrie Ltd
Participants	Charles Beagrie Ltd (UK), OCLC Research (USA), the UK Data Archive (UK), the Archaeology Data Service, the University of London Computer Centre (UK), and the universities of Cambridge (UK), King's College London (UK), Oxford (UK) and Southampton (UK).
Website	http://www.beagrie.com/krds.php
Funding	Funded by JISC

[8] http://dp4lib.langzeitarchivierung.de

[9] http://www.beagrie.com/krds.php

CFM—Cost Forecasting Model for New Digitization Projects

Project acronym	CFM
Project title	Cost Forecasting Model for New Digitization Projects
Description[10]	Current cost model studies in the field (both in the United States and in Europe) are helpful case studies in providing libraries and cultural institutions with an understanding of the cost implications for digitizing book collections. Because these projects are far-reaching and comprehensive, however, they offer up only a broad generalization of what cost variables to consider. To contribute to the dialogue of digitizing library book collections, the George Washington University Libraries will share their cost model with the community, which is based on the current production workflow setup at the Gelman Library using robotic arm technology, and is funded by the Institute of Museum and Library Services and donor contributions.
Start & end date	2011
Subjects	Cost model for digitization projects
Coordinator	N/A
Participants	George Washington University (USA)
Website	http://www.cni.org/topics/digital-libraries/cost-forecasting-model/
Funding	Funded by the Institute of Museum and Library Services (Grant US-IMLS NLG 2008)

[10] http://www.cni.org/topics/digital-libraries/cost-forecasting-model/)

LIFE—Life Cycle Information for E-Literature

Project acronym	LIFE
Project title	Life Cycle Information for E-Literature
Description[11]	LIFE (Life Cycle Information for E-Literature) is a collaboration between University College London (UCL) and the British Library. The LIFE Project has developed a methodology to model the digital lifecycle and calculate the costs of preserving digital information for the next 5, 10 or 20 years. For the first time, organisations can apply this process and plan effectively for the preservation of their digital collections. The third phase of LIFE commenced in August 2009, and will run for one year with funding from JISC and RIN. By producing a predictive costing tool, LIFE3 will significantly improve the ability of organisations to plan and manage the preservation of digital content. The project will expand its existing Generic Preservation Model to create a comprehensive suite of models covering all life cycle stages, providing greater accuracy and assurance in estimation. The predictive costing tool will be made available towards the end of 2010, as both a web application and an Excel-based model. The project team would be delighted to hear from organisations interested in assisting with trials of the tool.
Start & end date	LIFE 1: From 2006-04 to 2007-04 LIFE 2: From 2007-03 to 2008-08 LIFE 3: From 2009-08 to 2010-08
Subjects	Analysis and costing of the lifecycle and preservation of digital assets
Coordinator	University College London (UK)
Participants	University College London (UK), British Library (UK)
Website	http://www.life.ac.uk
Funding	Funded by Joint Information Systems Committee (JISC) and Research Information Network (RIN)

[11] http://www.life.ac.uk

Piloting the LIFE costs tool in UK HEIs

Project acronym	N/A
Project title	Piloting the LIFE costs tool in UK HEIs
Description[12]	The LIFE tool was developed by HATII as part of the LIFE3 project , which ran from August 2009 to September 2010. While the project did include user testing, potential end-users of the LIFE tool in UK HEIs would benefit from a greater wealth of practical user experiences and a broader range of cost data to draw upon. Over this 3 month project, the DCC led a phase of more detailed user testing of the LIFE tool to capture and disseminate a wider range of user experiences and cost data. The current LIFE model costs were derived from a number of case studies that were carried out primarily in national libraries and similar institutions. The DCC aimed to capture additional cost data from institutional repositories - who may not necessarily view digital preservation as a core activity - to provide a more balanced picture of the potential costs associated with preserving content at the institutional level. By characterising costs according to the contexts within which they are evident the tool will better serve more diverse user communities and offer more representative default cost values.
Start & end date	May 2011 – July 2011
Subjects	Analysis of the effectiveness of the LIFE tool and recommendations on future developments
Coordinator	DCC at HATII, University of Glasgow
Participants	Selection of UK HEI repositories
Website	http://www.dcc.ac.uk/projects/life
Funding	Funded by Joint Information Systems Committee (JISC)

[12] http://www.dcc.ac.uk/projects/life

BRTF-SDPA—Blue Ribbon Task Force on Sustainable Digital Preservation and Access

Project acronym	BRTF-SDPA
Project title	Blue Ribbon Task Force on Sustainable Digital Preservation and Access
Description[13]	To address issues like the digital information long-term preservation, access and its economic sustainability, the Blue Ribbon Task Force on Sustainable Digital Preservation and Access was created in late 2007, and in early 2010 published its Final Report, called "Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information". The report provides: general principles and actions to support long-term economic sustainability; context-specific recommendations tailored to specific scenarios analysed in the report; and an agenda for priority actions and next steps, organized according to the type of decision maker best suited to carry that action forward. Following publication of the report, the Task Force SDPA proposed a Grand Challenge recommendation for the U.S. Office of Science and Technology Policy's submission website to ensure that the knowledge of today is available for use tomorrow, while fostering innovation for sustainable growth and creating high-quality jobs. That report was submitted to the OSTP in mid-April.
Start & end date	From 2007 to 2010
Subjects	Long-term economic sustainability of digital preservation
Coordinator	N/A
Participants	Several individual people: http://brtf.sdsc.edu/members.html
Website	http://brtf.sdsc.edu
Funding	Funded by the National Science Foundation and the Andrew W. Mellon Foundation, in partnership with the Library of Congress, the Joint Information Systems Committee of the United Kingdom, the Council on Library and Information Resources, and the National Archives and Records Administration.

[13] http://brtf.sdsc.edu/about.html

UK Data Service data management costing tool and checklist

Project acronym	DMP-ESRC
Project title	Data management planning for ESRC research data-rich investments
Description[14]	The UK Data Archive will work closely with selected ESRC research centres and programmes to help develop and implement effective data management planning in the research life cycle and increase individual and institutional data managing and sharing capacity by providing best practice guidance, support and training. After assessing existing data management practices and auditing data assets, data management plans will be developed alongside protocols to implement and monitor them and overall data management strategies. This will help develop better data management practices and procedures in research and increase the potential for long-term use and validity of research data. The UK Data Service has prepared this costing tool and checklist to help formulate research data management costs in advance of research starting, for example for inclusion in a data management plan or in preparation for a funding application. This tool considers the additional costs - above standard planned research procedures and practice - that are needed to preserve research data and make them shareable beyond the primary research team. The checklist indicates the activities to consider and cost to enable good data management. Such additional activities may require extra researcher or administrative staff time input, equipment, software, infrastructure or tools. There are no hard and fast rules for costing data sharing requirements, as some research projects will pay more attention to detailed data documentation, organisation and formatting than others as part of routine fieldwork or preparation before analysis. Much also depends on the long-term storage, preservation and publication plans beyond the duration of the research itself. When data are deposited with a professional data centre or repository, such as the UK Data Archive, data preservation and dissemination activities are covered by the data centre/repository.
Start & end date	2011-2011
Subjects	Costs associated with research data management planning
Coordinator	UK Data Archive
Participants	ESRC research centres and programmes
Website	http://data-archive.ac.uk/create-manage/projects/jisc-dmp
Funding	Funded by Joint Information Systems Committee (JISC)

[14] http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmp/esrc.aspx

2.2 Publications

This section includes project reports, journal articles and conference papers that relate to cost modelling in digital curation. References are grouped by year of publication and sorted by author's last name.

2013

Kaur, K. et al. (2013). Report on Cost Parameters for Digital Repositories. Deliverable D32.1 of the APARSEN project. Retrieved from http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2013/03/APARSEN-REP-D32_1-01-1_0.pdf

Rosenthal, D. S. H., & Vargas, D. L. (2013). Distributed Digital Preservation in the Cloud. *International Journal of Digital Curation*, *8*(1), 107–119. Retrieved from http://www.ijdc.net/index.php/ijdc/article/view/8.1.107/300

2012

Addis, M., Jacyno, M., Hall-May, M., & Wright, R. (2012). Storage strategy tools. *International Association of Sound and Audiovisual Archives Journal*, (38). Retrieved from http://eprints.soton.ac.uk/273039/

Addis, M., Jacyno, M., Hall-May, M., McArdle, M. and Phillips, S. (2012) Planning and Managing the 'Cost of Compromise' for AV Retention and Access. Society of Motion Picture & Television Engineers Journal, 121(1) pp. 32-38. Retrieved from http://eprints.soton.ac.uk/273158/

Badawy, M., Shehab, E., Baguley, P., Wilson, M. (2012). Towards a cost model for long-term digital preservation. In Proc. 2012 ISPA/SCEA Joint International Conference: Assuring cost efficiency: global solution, Brussels, Belgium. Related to the ENSURE project, Retrieved from: http://epubs.stfc.ac.uk/bitstream/7711/Towards%20a%20Cost%20Model%20for%20Long%20Term%20Digital%20Preservation.pdf

Beagrie, N., Duke, M., Hardman, C., Kalra, D., Lavoie, B., Patel, M., Lyon, L., et al. (2012). The KRDS Benefit Analysis Toolkit: Development and Application. International Journal of Digital Curation, 7(2), 64–67. doi:10.2218/ijdc.v7i2.230

Bøgvad Kejser, U. (2012). Cost of preserving research data. PhD Thesis. Retrieved February 22, 2012, from http://www.diku.dk/research/phd-studiet/phd/thesis_20111215.pdf/

Bote, J., Fernandez-Feijoo, B., & Ruiz, S. (2012). The Cost of Digital Preservation: A Methodological Analysis. *Procedia Technology*, *5*(null), 103–111. Retrieved from http://dx.doi.org/10.1016/j.protcy.2012.09.012

Bote, J., Fernandez-Feijoo, B., Ruiz, S. (2012). The Cost of Digital Preservation: A Methodological Analysis. In CENTERIS 2012 proceedings, available at: http://www.sciencedirect.com/science/article/pii/S2212017312004434

Charles Beagrie. (2012). Economic Impact Evaluation of the Economic and Social Data Service. Retrieved from http://www.esrc.ac.uk/_images/ESDS_Economic_Impact_Evaluation_tcm8-22229.pdf

Hawtin, R., Hammond, M., Gillam, L., Curtis, G. (2012). Cost analysis if cloud computing for research. Report for EPSRC and JISC, available at: http://www.jisc.ac.uk/media/documents/programmes/research_infrastructure/costcloudresearch.pdf

Kiefer, S., & Wilson, M. (2012). Ensuring Profitability of Commercial Long Term Digital Preservation. *ERCIM News*, *91*, 19–21. Retrieved from http://ercim-news.ercim.eu/en91/ri/ensuring-profitability-of-commercial-long-term-digital-preservation

Palaiologk, A. S., Economides, A. A., Tjalsma, H. D., & Sesink, L. B. (2012). An activity-based costing model for long-term preservation and dissemination of digital research data: the case of DANS. *International Journal on Digital Libraries*, *12*(4), 195–214. doi:10.1007/s00799-012-0092-1

Schimtt, K., Klaproth, F. (2012). Cost model for a long-term preservation service (German). Report of the DP4lib project, available at: http://dp4lib.langzeitarchivierung.de/downloads/DP4lib-Kostenmodell_eines_LZA-Dienstes_v1.0.pdf

UC Curation Center (2012). Total Cost of Preservation (TCP) - Cost Modelling for Sustainable Services - Rev. 2.0 (draft). Report of the TCP project, available at: https://wiki.ucop.edu/download/attachments/163610649/TCP-cost-modeling-for-sustainable-services-v2.pdf

2011

Addis, M., Allasia, W., Bailer, W., Boch, L., Gallo, F., Phillips, S., & Schallauer, P. (2011). Digital preservation of audiovisual files within PrestoPRIME. Retrieved from http://eprints.soton.ac.uk/273042/1/virtualgoods2011_submission_4.pdf

Addis, M., Jacyno, M., Hall-May, M., McArdle, M. and Phillips, S. (2011) Planning and Managing the 'Cost of Compromise' for AV Retention and Access. In: 2011 Conference of the International Broadcast Convention, 6-11 September, Amsterdam. Available at: http://eprints.ecs.soton.ac.uk/23158/1/23158.pdf

Addis, M., Wright, R. and Weerakkody, R. (2011) Digital Preservation Strategies: The Cost of Risk of Loss. Society of Motion Picture & Television Engineers Journal, 120(1) pp. 16-23. Retrieved from http://journal.smpte.org/content/120/1/16.abstract

Beagrie, C. (2011). User Guide for Keeping Research Data Safe: Assessing Costs/Benefits of Research Date Management, Preservation and Re-use. Charles Beagrie Limited (p. 44). Retrieved from http://www.beagrie.com/KeepingResearchDataSafe_UserGuide_v2.pdf

Beagrie, N. (2011). A GUIDE TO THE KRDS BENEFITS FRAMEWORK V3 (JULY 2011). Keeping Research Date Safe. Retrieved December 15, 2011, from http://www.beagrie.com/KRDS_BenefitsFramework_Guidev3_July 2011.pdf

Eddis, M., & Wright, R. (2011). Creating Preservation Scenarios, Understanding Risks and Estimating Costs. Retrieved from https://www.prestocentre.org/stfcon-2011#2

Gantz, J., & Reinsel, D. (2011). Extracting value from the Chaos. Retrieved from http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf

Kejser, U. B., Nielsen, A. B., & Thirifays, A. (2011). Cost Model for Digital Preservation: Cost of Digital Migration. The International Journal of Digital Curation, 13. doi:http://dx.doi.org/10.2218/ijdc.v6i1.186

McCann, P. (2011) Piloting the LIFE costs tool in UK HEIs - report on the effectiveness of the LIFE tool and provide recommendations on future developments. http://www.dcc.ac.uk/sites/default/files/documents/life_pilot_final_2012.pdf

Rusbridge, C., & Lavoie, B. (2011). Draft economic sustainability reference mode. Wordpress. Retrieved December 15, 2011, from http://unsustainableideas.files.wordpress.com/2011/09/reference-model-0-6-clean.pdf

Snyders, M., Westerhof, H., Ubois, J. (2011). Financial Models and Calculation Mechanism. Deliverable D6.3.1 of the PrestoPRIME project, available at: https://prestoprimews.ina.fr/public/deliverables/PP_WP6_D6.3.1_FM_calculation_R0_v1.01.pdf

Strodl, S., Rauber, A. (2011). A cost model for small scale automated digital preservation archives. In iPRES 2011 proceedings, available at: http://www.ifs.tuwien.ac.at/~strodl/paper/strodl_ipres2011_costmodel.pdf

Xue, P., Badawy, M., Shehab, E., Baguley, P. (2011). Cost modelling for long-term digital preservation: Challenges and issues. Proceedings of the 9th International Conference on Manufacturing Research ICMR 2011, Glasgow Caledonian University, Glasgow Sep 2011, pp 187-192. Available at: http://cranfield.academia.edu/EssamShehab/Papers/1285084/COST_MODELLING_FOR_LONG-TERM_DIGITAL_PRESERVATION_CHALLENGES_AND_ISSUES

2010

Addis, M., Allasia, W., Bailer, W., Boch, L., Gallo, F., & Wright, R. (2010). 100 million hours of audiovisual content: digital preservation and access in the PrestoPRIME project. Proceedings of the 1st International Digital Preservation Interoperability Framework Symposium (pp. 3:1–3:8). New York, NY, USA: ACM. doi:10.1145/2039263.2039266

Beagrie, N., Lavoie, B., & Woollard, M. (2010). *Keeping Research Data Safe 2* (p. 89). Retrieved from http://www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf

Berman, F., et al. (2010). Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information. Final report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access project, available at: http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

Brebner, P., & Liu, A. (2010). Modeling Cloud Cost and Performance. Proceedings of the International Conference on Cloud Computing Virtualization 2010 CCV 2010, (Ccv), 79–86. doi:10.5176/978-981-08-5837-7_154

Hole, B., Wheatley, P., Lin, L., & McCann, P. (2010). LIFE 3 : A PREDICTIVE COSTING TOOL FOR DIGITAL COLLECTIONS. LIFE. Retrieved December 19, 2011, from http://www.life.ac.uk/3/docs/Ipres2010_life3_submitted.pdf

Keskitalo, E. (2010). Costs and Benefits of a Shared Digital Long-Term Preservation System. National Library of Finland. Retrieved December 15, 2011, from http://www.doria.fi/bitstream/handle/10024/69928/liber paper keskitalo.pdf

Meeting, E. (2010). Expert Meeting: Price Tags of Digital Preservation Policy Choices. Digital Heritage Netherlands (DEN), the Netherlands Coalition for Digital Preservation (NCDD) ,the Koninklijke Bibliotheek, National Library of the Netherlands (KB). Retrieved December 15, 2011, from http://www.ncdd.nl/en/documents/20100916PriceTagsConferenceReportfinal.pdf

Poole, N. (2010). The Cost of Digitising Europe's Cultural Heritage : A Report for the Comité des Sages of the European Commission. the Collections Trust. Retrieved December 15, 2011, from http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/annexes/digiti_report.pdf

Walters, T., & Skinner, K. (2010). Economics, sustainability, and the cooperative model in digital preservation. Library Hi Tech, 28, 259–272. doi:http://dx.doi.org/10.1108/07378831011047668

Wilkin, J. (2010). Hathi trust Cost Rationale 2013. University of Michigan Library. Retrieved December 15, 2011, from http://www.hathitrust.org/documents/hathitrust-cost-rationale-2013.pdf

Zeller, J. (2010). Cost of digital archiving: Is there a universal model? 8th European Conference on Digital Archiving. Geneva.

2009

Beagrie, N., Beagrie, R., & Rowlands, I. (2009). Research Data Preservation and Access: The Views of Researchers. *Ariadne*, *60*(60). Retrieved from http://www.ariadne.ac.uk/issue60/beagrie-et-al/

Mageto, D. (2009). *Cost Factors in Digital Preservation*. International Master in Digital Library Learning. Retrieved from http://research.dnv.com/LongRec/Intranet/ResearchResults/RecommendedPractices/CostFactorsForDigitalPreservation.pdf

Neuroth, H., Oßwald, A., Strathmann, S., & Jehn, M. (2009). Nestor Handbuch Kapitel 14.2 Kosten. *nestor*. Retrieved December 15, 2011, from http://nestor.sub.uni-goettingen.de/handbuch/artikel/nestor_handbuch_artikel_310.pdf

2008

Addis, M., Wright, R., & Miller, A. (2008). The Significance of Storage in the "Cost of Risk" of Digital Preservation. The International Journal of Digital Curation, 4(3), 104–122. doi:http://dx.doi.org/10.2218/ijdc.v4i3.125

Ayris, P., Davies, R., McLeod, R., Miao, R., Shenton, H., & Wheatley, P. (2008). The LIFE2 final project report. LIFE (p. 122). London. Retrieved from http://eprints.ucl.ac.uk/11758/

Beagrie, N. (2008). KRDS Benefits Taxonomy Version 2.0. Keeping Research Date Safe. Retrieved December 15, 2011, from https://docs.google.com/viewer?a=v&q=cache:4goSfdAMGC8J:www.beagrie.com/KRDSBenefitsTaxonomyv2.doc+krds+benefits+taxonomy&hl=de&gl=de&pid=bl&srcid=ADGEESgsLULa_ynSndft3-OTRpUKkPjXdHCsF2-UMF_0uS64dpr0z0KY_wbAT5NI3TE_0jTwXkIsLG0BODXrfxHu-6Srcx4QFv-VC_aN_TYezmDzipYhFRBmjaWuMa47B3JD3GGkqRby&sig=AHIEtbSknAf3GXXK2yKc8nhNH7lHABctpg

Dickmann, F. (2008). AP 5 - Kosten der elektronischen Langzeitarchivierung. Universitätsmedizin Göttingen Abteilung Medizinische Informatik. Retrieved December 15, 2011, from http://kolawiss.uni-goettingen.de/projektergebnisse/AP5_Report.pdf

Lavoie, B. (2008). The Fifth Blackbird: Some Thoughts on Economically Sustainable Digital Preservation. *DLib Magazine*, *14*(3/4). Retrieved from http://www.dlib.org/dlib/march08/lavoie/03lavoie.html

Lavoie, B., Chruszcz, J., & Beagrie, N. (2008). *Keeping Research Data safe - A Cost Model and Guidance for UK Universities*. *Charles Beagrie Limited* (p. 169). Retrieved from http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf

Schoger, A., Hagel, H., Minkus, M., Rödig, P., Lang, S., Beinert, T., & Borghoff, U. M. (2008). Development of Organisational and Business Models for the Long-Term Preservation of Digital Objects. iPres presentation. Retrieved December 15, 2011, from http://www.bl.uk/ipres2008/presentations_day1/04_Lang.pdf

The Blue Ribbon Task Force. (2008). *Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable Digital Preservation*. *The Blue Ribbon Task Force* (p. 72). Retrieved from http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf

2007

Björk, B.-C. (2007). Evaluation of the costing activities and economic models for digital curation using the life methodology. *Finland*. Retrieved December 16, 2011, from http://eprints.ucl.ac.uk/7684/1/7684.pdf

Fontaine, K., Hunolt, G., Booth, A., Banks, M. (2007). Observations on cost modelling and performance measurement of long-term archives. NASA research paper in PV2007 Conference Proceedings (2007), available at: http://www.pv2007.dlr.de/Papers/Fontaine_CostModelObservations.pdf

Piwowar, H., Day, R., & Fridsma, D. (2007). Sharing Detailed Research Data is Associated with Increased Citation Rate. Retrieved from http://precedings.nature.com/documents/361/version/1/files/npre2007361-1.pdf

Plasmon. (2007). Total-Cost-of-OwnershipAnalyse von Archivierungslösungen. *Plasmon*. Retrieved December 14, 2011, from http://www.plasmon.com/downloads/pdf/tcode.pdf

Wright, R. (2007). Annual Report on Preservation Issues for European Audiovisual Collection. Retrieved February 24, 2012, from http://www.prestospace.org/project/deliverables/D22-8.pdf

2006 and older

Adler, P. S. (1989). When knowledge is the critical resource, knowledge management is the critical task. *Engineering Management, IEEE Transactions on*, *36*(2), 87–94. doi:10.1109/17.18822

Baskerville, R., & Pries-Heje, J. (1999). Knowledge capability and maturity in software management. *SIGMIS Database*, *30*(2), 26–43. doi:10.1145/383371.383374

Booth, B., Banks, M., & Hunolt, G. (2006). Cost estimation tool enhanced operational comparables database. Retrieved from http://opensource.gsfc.nasa.gov/projects/CET/Doc.zip

Bradway B., R. S. (2000). Bradway B., Ross S.: Measuring corporate customer profitability: the role of activity-based cost analysis. *Corp. Cust. Manag*, *4*(6), 1–10.

Brown, A. (2005). Cost Modelling: The TNA Experience. Retrieved from http://www.dpconline.org/docs/events/050726brown.pdf

Chapman, S. (2004). Counting the Costs of Digital Preservation: Is Repository Storage Affordable? *Journal of Digital Information*, *4*(2). Retrieved from http://journals.tdl.org/jodi/article/view/100

Cokins, G. (2000). Overcoming the obstacles to implementing activity-based costing. *Bank Accounting and Finance*, (1), 47–52. Retrieved from http://scholar.google.pt/scholar?hl=pt-PT&q=Overcoming+the+obstacles+to+implementing+activity-based+costing&btnG=&lr=#1

Cokins, G. (1996). *Activity-Based Cost Management Making It Work: A Manager's Guide to Implementing and Sustaining an Effective ABC System* (p. 192). McGraw-Hill. Retrieved from http://www.amazon.com/Activity-Based-Cost-Management-Making-Work/dp/0786307404

ERPANET. (2003). ERPANET: Cost orientation tool. *Erpanet*. Retrieved from http://www.erpanet.org/guidance/docs/ERPANETCostingTool.pdf

Granger, S., Russell, K., & Weinberger, E. (2000). Cost Elements of Digital Preservation. Retrieved from http://www.webarchive.org.uk/wayback/archive/20050409230000/http://www.leeds.ac.uk/cedars/colman/costElementsOfDP.doc

Hendley, T. (1998). *Comparison of Methods and Costs of Digital Preservation*. *British Library Research and Innovation Report* (pp. 10–6). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.136.1118

Plasmon. (2004). Archival Storage Total Cost of Ownership Analysis. *Plasmon*. Retrieved December 14, 2011, from http://www.dataarchivecorp.com/PDF/Total Cost of Ownserhip v2.0.pdf

Roztocki, N., Valenzuela, J. F., Porter, J. D., Monk, R. M., & Needy, K. L. (1999). A Procedure for Smooth Implementation of Activity Based Costing in Small Companies. *ASEM National Conference Proceedings* (pp. 279–288). Virginia Beach. Retrieved from http://www2.newpaltz.edu/~roztockn/virginia99.pdf

Sanett, S. (2002). Toward Developing a Framework of Cost Elements for Preserving Authentic Electronic Records into Perpetuity. *College & Research Libraries*, *63*(5), 388–404. Retrieved from http://crl.acrl.org/content/63/5/388.abstract

Sanett, S. (2003). The Cost to Preserve Authentic Electronic Records in Perpetuity: Comparing Costs across Cost Models and Cost Frameworks. *RLG DigiNews*, *7*(4). Retrieved from http://worldcat.org/webharvest/h1250009815353/viewonline

Shenton, H. (2003). Life Cycle Collection Management. *LIBER Quarterly*, *13*(3/4), 254–272. Retrieved from http://liber.library.uu.nl/index.php/lq/article/view/URN%3ANBN%3ANL%3AUI%3A10-1-113343/7808

Slat, J., & Verdegem, R. (2005). Cost Model for Digital Preservation. *Nationaal Archief of the Netherlands*. Retrieved December 13, 2011, from http://dlmforum.typepad.com/Paper_RemcoVerdegem_and_JS_CostModelfordigitalpreservation.pdf

Stephens, A. (1994). The application of life cycle costing in libraries: A case study based on acquisition and retention of library materials in the British Library. *IFLA journal*. Available at http://cat.inist.fr/?aModele=afficheN&cpsidt=4180961

2.3 Presentations

This section includes presentations given at scientific events about cost modelling initiatives in the context of digital curation. References are grouped by year of publication and sorted by author's last name.

2012

Abrams, S., Cruse, P., Kunze, J. (2012). Total cost of preservation: Cost modelling for sustainable services. Screening the Future: Pause, Play, and Press Forward, Los Angeles, May 21-23, 2012. Available at: https://wiki.ucop.edu/download/attachments/163610649/Screening-the-future-UC3-cost-model.pptx

Abrams, S., Cruse, P., Kunze, J. (2012). Cost modelling for sustainable curation services. Preservation and Archiving Special Interest Group (PASIG), Dublin, October 16-19, 2012. Available at: https://wiki.ucop.edu/download/attachments/163610649/PASIG-cost-modeling-for-sustainable-services.pptx

Abrams, S., Cruse, P., Kunze, J. (2012). Pay once, preservation forever: a 'paid-up' cost model for long-term preservation," CNI Spring 2012 Membership Meeting, Baltimore, April 2-3, 2012. Available at: https://wiki.ucop.edu/download/attachments/163610649/CNI-UC3-cost-model.pptx

Abrams, S., Cruse, P., Kunze, J. (2012). Total cost of preservation (TCP): Cost modeling for sustainable services," Joint UC Council of University Librarians (CoUL) / Systemwide Operations and Planning Advisory Group (SOPAG) meeting, February 16, 2012. Available at: https://wiki.ucop.edu/download/attachments/163610649/TCP-cost-model-v04.pptx

Friese, Y. (2012). Cost Factors of Digital Preservation - Their importance and how to detect them. *Nordbib* (p. 12). Copenhague. Retrieved from http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=/Files/Filer/downloads/Primary+Research+Data/Cost+models+workshop/Yvonne+Friese+Presentation+Nordbib.pdf

Friese, Y. (2012). Cost Factors of Digital Preservation - Their importance and how to detect them. Nordbib (p. 12). Copenhague. Available at: http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fdownloads%2fPrimary+Research+Data%2fCost+models+workshop%2fYvonne+Friese+Presentation+Nordbib.pdf

Wheatley, P. (2012). Some thoughts on the challenges of costing digital preservation. Copenhagen. Retrieved July 9, 2013, from http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=/Files/Filer/downloads/Primary+Research+Data/Cost+models+workshop/Paul+Wheatley+Costing+Digital+Preservation2.pdf

Wright, R. (2011). Preservation: Scenarios, Risks, Costs. Retrieved August 28, 2012, from http://www.prestocentre.org/library/resources/preservation-scenarios-risks-costs

2011

Addis M. (2011) Storage and Services: Planning and managing cost, quality and risk. Presentation at the FIAT/IFTA world congress 2011 as part of the PrestoPRIME pre-conference workshop. 28 Sep 2011. Turin, Italy. Available at: http://eprints.ecs.soton.ac.uk/23043/1/PrestoPRIME_FIAT_2011v3.pdf

Addis, M. (2011). Cost, Risk, Loss and Other Fun Things. Retrieved August 28, 2012, from http://www.prestocentre.org/library/resources/cost-risk-loss-and-other-fun-things

Boughida, K., Whittaker, M., Colet, L., Chudnov, D. (2011). Cost forecasting model for new digitization projects (Excel and web tool under development). Available at: http://www.cni.org/wp-content/uploads/2011/12/cni_cost_boughida.pdf

Hall-May, M. (2011) Storage and Services: Planning and managing cost, quality and risk. Association of Moving Image Archivists Conference. Nov 2011. Austin, Texas. Available at: http://eprints.ecs.soton.ac.uk/23118/1/PrestoPRIME_AMIA_2011.pdf

Wright, R. (2011) Preservation: Scenarios, Risks, Costs. Screening the Future. Available at: http://www.prestocentre.org/library/resources/preservation-scenarios-risks-costs

2009

Bellinger, M. (2009). Cost and business models for digital preservation: Developing digital life cycle management services at OCLC. *4th Forum of Digital Preservation Coalition "Preservation of e-Learning Materials and Cost Models for Digital Preservation"*. London. Retrieved from http://www.dpconline.org/graphics/events/presentations/pdf/BellingerDPCForum_CostsBusinessModels.pdf

Wheatley, P. (2009). Presentation: LIFE3: Predicting Long Term Preservation Cost. iPres presentation. Retrieved December 15, 2011, from http://www.cdlib.org/services/uc3/iPres/presentations/Wheatley.pdf

2.4 Posters

This section includes posters presented at conferences. References are grouped by year of publication and sorted by author's last name.

2012

Badawy, M. (2012). Cost Modelling for Long Term Digital Preservation, DTC 2012 Conference, 17 Jan 2012, Cranfield University. Available at: http://ensure-fp7-plone.fe.up.pt/site/DTCposterPortrait.pdf

2011

Lefort, A., Conway,E., Shehab, E., Bagley, P., Xue, P., Badawy, M., Wilson,M. (2011). Modelling Digital Preservation costs for ISIS Instrument Data Proc.7th International Digital Curation Conference, Bristol, UK, 05-08 Dec 2011. Related to the ENSURE project. Available at: http://epubs.stfc.ac.uk/bitstream/7170/Poster%20DCC.pdf

Alhawas, A. (2011). Towards Cost Modelling for Long Term Digital Preservation, MSc project poster, Cranfield University, Sep 2011. Related to the ENSURE project. Available at: http://ensure-fp7-plone.fe.up.pt/site/ALHAWAS.pdf

Lefort, A. (2011). Modelling Digital Preservation Costs for ISIS Instrument Data. MSc project poster, Cranfield University, Sep 2011. Related to the ENSURE project. Available at: http://ensure-fp7-plone.fe.up.pt/site/Poster.pdf

Shehab, E. (2011). Cost Modeling for Long Term Digital Preservation. UK National Manufacturing Debate 2011, Cranfield University. Related to the ENSURE project. Available at: http://ensure-fp7-plone.fe.up.pt/site/copy2_of_ENSURE_ManufacturingDebate_Poster.pdf

2.5 Events and event reports

This section includes names of events, event reports, and other relevant information such as webcasts, participants' lists, etc. from events dedicated to cost modelling and economic issues in the context of digital curation. References are grouped by year of publication and sorted by author's last name.

2013

Exchange Knowledge. (2013). Knowledge Exchange workshop - Price of keeping knowledge: financial streams for digital preservation. Amsterdam. Conference report at: http://www.knowledge-exchange.info/Default.aspx?ID=570

Screening the Future 2013 - Crossing the Boundaries for AV Preservation. By PRESTOCentre, on May 7-8, 2013, London, UK. Site available at: https://prestocentre.org/calendar/screening-future-2013-conference

2012

Screening the Future 2012 - Play, Pause and Press Forward. By PRESTOCentre, on 2012. Site available at: https://www.prestocentre.org/calendar/screening-future-conference-2012-play-pause-and-press-forward, blog post available at: http://www.ncdd.nl/blog/?p=1849

Workshop: 'The Costs and Benefits of Keeping Knowledge: economic models for digital preservation'. By Knowledge Exchange, on June 11, 2012, Copenhagen, Denmark. Site available at: http://www.knowledge-exchange.info/Default.aspx?ID=512

2011

UK Data Archive Seminar Data Management Planning and Practices for Research Centres and Programmes17, on 4 May 2011, at Royal Statistical Society, London. Seminar website: http://www.data-archive.ac.uk/news-events/events.aspx?id=2807

2010

Blue Ribbon Task Force. (2010). A National Conversation on the Economic Sustainability of Digital Information. Washington D.C., USA. Conference report at: http://www.oclc.org/research/news/2010/04-08.html

Expert Meeting: Price Tags of Digital Preservation Policy Choices. By the National Library of the Netherlands, on September 16, 2010, The Hague, Netherlands. Conference report at: http://www.ncdd.nl/en/documents/20100916PriceTagsConferenceReportfinal.pdf

JISC MRD Programme workshop: Costs, Benefits and Sustainability18, on2-3 November 2010 in Bristol, Workshop website: http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmevents/mrdworkshop.aspx

3 Stakeholders registry

The Engagement work package acts as an interface between external stakeholders and the 4C project. To establish that interface, an initial set of stakeholder categories has been created by means of internal brainstorming. These categories are particularly important, as they define professional audiences to which all project outputs should be tailored for. The initial list of categories will further be refined as the project progresses.

For each stakeholder category, we have set a degree of difficulty in reaching that particular type of stakeholder and the impact that the 4C project will have towards that particular community. The combination of these aspects results in a final score that allowed us to prioritise the effort spent towards reaching each of these communities.

The difficulty and the impact are evaluated according to the following numeric scale:

- **Low –** assigned a value of value of 1
- **Medium** - assigned a value of value of 2
- **High** - assigned a value of value of 3

The final prioritisation score is given by the formula: $\text{Priority} = \text{Impact} \times (4 - \text{Difficulty})$.

Table 1 lists all the stakeholder categories, the perceived difficulty in engaging them, the impact that reaching that stakeholder would have in terms of uptake of project and the priority score according to the previous formula.

Stakeholder category	Description	Difficulty	Impact	Priority
Research funders	Institutions that provide funding for scientific research.	Low	High	9
Big data science	Institutions for scientific research that deal with large amounts of data, e.g. space and high-energy physics research.	Medium	High	6
Digital preservation vendors	Companies that deliver products or services in the area of digital preservation, e.g. storage vendors, software providers, digital preservation consulting.	Low	Medium	6
Universities	Higher education institutions with responsibility for maintaining digital collections.	Low	Medium	6
Government agencies	Public administration institutions that must maintain data for long periods of time, e.g. central banks, medical records, police, health care, cartography, local authorities.	Medium	Medium	4
Publishers & content producers	Publishers of books, scholarly materials, and other types of media (e.g. audio, video) as well as related services. This category also includes the actual producers of the materials.	Medium	Medium	4
Industry	Companies that keep data to support their business, e.g. aviation, space, bioinformatics, cartography, automotive, banks & finance, pharmaceutical, defence industry.	High	High	3
Memory institutions and content holders	Institutions whose main mission is to preserve cultural heritage, e.g. libraries and archives.	Low	Low	3
Small medium enterprises	Enterprises that are legally or operationally compelled to maintain data for long periods of time, e.g. escrow services.	High	High	3
Other	Contacts that should also be addressed but do not fit easily under any of the previous stakeholder group contacts.	-	-	-

Table 1—Prioritized list of stakeholder categories.

For each stakeholder category, a list of personal contacts has been collected amongst all partners in the project. Additional general contacts (e.g. digital preservation mailing lists) have been added to the list under the "Other" category. All of the collected contacts have been invited to participate in an initial stakeholder consultation composed by a three-parted questionnaire (more details on Section 4).

The contacts compiled so far, and all additional contacts collected during the stakeholders consultation will be imported into an online CRM system that will help with managing all future engagement activities and monitoring our impact with each community. This system will be available to specific project members only to protect the privacy of the stakeholders.

Table 2 depicts the metadata that has been collected for each stakeholder contact.

# | Name	e.g. Mary Smith
Motivation	Why this person is a good representative of the community
Stakeholder category	From the list of categories
Role/position	The role or position this contact has on the organization defined below
Organization	**Name:** Name of the organization this contact belongs to **Description:** Small description of the organization **Particular interests:** List of organization main interests that relate to 4C **Country:** (Main) country the organization belongs to
Email, phone or other	e.g. mary.smith@stakeholder.org, +55 555 555 555
Contact owner	4C partner that owns this contacts and has better chance of getting a reply e.g. Miguel Ferreira (KEEPS), mferreira@keep.pt
Reason for contact	Explanation on what kind of information is expected from the contact person. Also indicate task id for greater reference and date. e.g. T2.1 - early consultation of stakeholder interests and initiatives - 2013-05-01
Log and schedule	Registry of past and future contacts by 4C members, e.g.: **DONE** 2013-03-20 - List of 6 questions sent by email. No reply yet. **TODO** 2013-04-20 - Ask for a reply if not yet received.

Table 2— Metadata collected for each stakeholder contact.

The characterisation of stakeholder types is depicted on the following sections. For each category we have registered the perceived community size, possible output channels and input channels, awareness about digital preservation issues, motivation to be involved in 4C, main barriers in reaching that group and mitigation strategy, existing sub-communities, number of contacts collected and their countries.

For privacy reasons, personal contacts collected will not be included in this report. Instead, a few statistics will be provided on the number of contacts that have been collected per stakeholder category.

3.1 Research funders

Research funders	
Institutions that provide funding for scientific research.	
Community size	Medium (a targeted number of funders will be contacted in the course of 4C)
Channels out	Links with UK funders via DCC, links with DFG, SURF, DEFF and CFC via involvement in Knowledge Exchange; links with IMLS in the US via DCC and Jisc;
Channels in	Invitations to participate in focus group meetings; review of key findings
DP issues knowledge	Knowledge of requirements for data management planning and data sharing
4C involvement motivation	The relative interests of research funders in this area will vary greatly. At one end of the scale, some funders run their own data archives and hence are directly exposed to the costs and benefits of preservation; others expect data to be preserved by others in a way that is not a cost to the research project. Their interest in costs is therefore close to zero. Funding bodies have a direct interest in the long-term sustainability of project outputs to increase impact and justify initial investment. RCUK funders have, in their Common Principles, stated that the use of public funds to support research data management and sharing are acceptable in-project costs. Accordingly, funders will have an indirect interest in making assessment on the validity of costs requested in new grant proposals. In the longer term, there will need to be better agreement on how in-project costs and longer term costs can be met by funders and institutions.
Barriers & mitigation	There may be difficulty in engaging funders directly as they may see costs as more of an institutional concern.
Priority sub-communities	Collaborative bodies between research funders and direct funders of research.
Number of contacts	13
Contact countries	United Kingdom, Netherlands, Germany, Portugal, Denmark, Finland, USA,

3.2 Big data science

Big data science	
Institutions for scientific research that deal with large amounts of data, e.g. space and high-energy physics research.	
Community size	Small
Channels out	Homepage, technical reports in relevant literature
Channels in	Invitations to participate in focus group meetings; review of key findings
DP issues knowledge	High
4C involvement motivation	Medium
Barriers & mitigation	n/a
Priority sub-communities	The Big Data Science organisations with a membership in APA, the Alliance for Permanent Access to the Records of Science
Number of contacts	15
Contact countries	Italy, Switzerland, United Kingdom, Germany, Portugal, Netherlands, USA,

3.3 Digital preservation vendors

Digital preservation vendors	
Companies that deliver products or services in the area of digital preservation, e.g. storage vendors, software providers, digital preservation consulting.	
Community size	Medium
Channels out	Talks, brochures, published materials, training sessions
Channels in	Invitations to participate in focus group meetings; review of key findings
DP issues knowledge	A digital preservation will have a high level of knowledge on the issues and the state of the art in terms of digital preservation solutions.
4C involvement motivation	Vendors will be interested in the results of the 4C project to learn more about the wholeset of activities that certain groups of customers carry on regarding digital curation. New business opportunities may arise from that knowledge. Learning about cost quantification enhances vendors' ability to benchmark their prices with its competitors as well as advise their customers on the best options available for them. Fruitful partnerships may also result from this knowledge as complementary services may be coupled together by the single fact that companies recognising their existence.
Barriers & mitigation	It may be difficult to obtain feedback from this group of stakeholders.
Priority sub-communities	Hardware vendors, software vendors, consultancy firms.
Number of contacts	14
Contact countries	United Kingdom, Israel, USA, Sweden, France, Germany, Ireland,

3.4 Universities

Universities	
Higher education institutions with responsibility for maintaining digital collections.	
Community size	Large
Channels out	Specific mailing list for HEI senior managers in the UK (UCISA, RUGIT)
Channels in	Links with 4C partner institutions (University of Edinburgh, University of Glasgow, University of Essex); Links with a number of UK HEIs via DCC Institutional Engagement activity (22+); Links with Russell Group in the UK via DCC partners;
DP issues knowledge	Little awareness generally amongst senior management; better amongst heads of IT services
4C involvement motivation	In the UK, recent EPSRC requirements have generated more interest amongst senior management in developing and sustaining research data management and curation infrastructure. Several UK HEIs have developed business cases in the last few months in advance of EPSRC's deadline for compliance to its research data framework requirements in May 2015.
Barriers & mitigation	Most senior managers in HEIs are extremely busy and would need to be convinced of the value of spending time shaping/contributing and/or commenting on the work of 4C.
Priority sub-communities	Jisc MRD Programme projects; UK HEIs
Number of contacts	8
Contact countries	United Kingdom

3.5 Government agencies

Government agencies	
Public administration institutions that must maintain data for long periods of time, e.g. central banks, medical records, police, health care, cartography, local authorities.	
Community size	Large
Channels out	Homepage, link from other websites, Surveys, Mailing Lists, Facebook, Twitter, Verbal Contact, "Word of Mouth"
Channels in	User opinions, Facebook, Twitter, Verbal Contact, "Word of Mouth"
DP issues knowledge	Knowledge of importance of DP exists
4C involvement motivation	low / medium
Barriers & mitigation	Lack of knowledge but understanding of importance to do digital curation
Priority sub-communities	n/a
Number of contacts	11
Contact countries	Denmark, United Kingdom, Scotland, Ireland

3.6 Publishers & content producers

Publishers & content producers	
Publishers of books, scholarly materials, and other types of media (e.g. audio, video) as well as related services. This category also includes the actual producers of the materials.	
Community size	Small
Channels out	Homepage, link from other websites, Surveys, Mailing Lists, Facebook, Twitter, Verbal Contact, "Word of Mouth".
Channels in	STM publishers, LOCKSS, Portico
DP issues knowledge	Medium/varying
4C involvement motivation	Medium
Barriers & mitigation	As industry enterprises, publishers might even be more hesitant than public institutions to share cost information.
Priority sub-communities	STM publishers (via Eefke Smit, member of the 4C Advisory Board), publishers involved in LOCKSS, Portico.
Number of contacts	10
Contact countries	Netherlands, Germany, United Kingdom,

3.7 Industry

Industry	
Companies that keep data to support their business, e.g. aviation, space, bioinformatics, cartography, automotive, banks & finance, pharmaceutical, defense industry.	
Community size	Large
Channels out	Homepage, social media (Facebook, Twitter) and other media (papers, magazines etc.) , surveys
Channels in	Homepage, surveys, social media (Facebook, Twitter)
DP issues knowledge	Medium/low
4C involvement motivation	Our assumption is that industry should be motivated to get involved with 4C because 4C tools and results can help companies to avoid needless costs and return a better profit on company's data assets.
Barriers & mitigation	Little awareness of the data curation issue.
Priority sub-communities	n/a
Number of contacts	12
Contact countries	United Kingdom, Denmark, Germany, France, Portugal, International

3.8 Memory institutions and content holders

Memory institutions and content holders	
Institutions whose main mission is to preserve cultural heritage, e.g. libraries and archives.	
Community size	Large
Channels out	Homepage, link from other websites, Surveys, Mailing Lists, Facebook, Twitter, Verbal Contact, “Word of Mouth”
Channels in	User opinions, Facebook, Twitter, Verbal Contact, “Word of Mouth”
DP issues knowledge	High
4C involvement motivation	High
Barriers & mitigation	Only self-interest, not interested in co-operating but receiving results
Priority sub-communities	University Libraries, Municipal Libraries, Students, Scientists
Number of contacts	27
Contact countries	USA, Portugal, Denmark, Germany, United Kingdom, Ireland

3.9 Small and medium enterprises

Small medium enterprises	
Enterprises that are legally or operationally compelled to maintain data for long periods of time, e.g. escrow services.	
Community size	Very big
Channels out	Homepage, link from other websites, Surveys, Mailing Lists, Facebook, Twitter, Verbal Contact, “Word of Mouth”
Channels in	User opinions, Facebook, Twitter, Verbal Contact, “Word of Mouth”
DP issues knowledge	Low
4C involvement motivation	Low
Barriers & mitigation	Only self-interest, not interested in co-operating but receiving results
Priority sub-communities	Service providers, hardware vendors, consulting companies
Number of contacts	1
Contact countries	Germany

3.10 Other

Other	
Contacts that seem relevant and should be addressed in the initial consultation but do not fit easily in any of the previous stakeholder categories. This includes mailing lists.	
Community size	Large
Channels out	Varying
Channels in	Varying
DP issues knowledge	Unknown
4C involvement motivation	Unknown
Barriers & mitigation	Unknown
Priority sub-communities	N/A
Number of contacts	185
Contact countries	Varying

3.11 Summary and statistics

The total number of personal contacts collected amongst the consortium was 111 direct email contacts. These are organised by stakeholder category, and for each contact an owner has been assigned, i.e. the person with the highest probability of engaging successfully with that external contact.

Additionally, 185 contacts that do not fit under any of the existing categories have been added to the list under the "Other" category totalling 296 contacts. These also include mailing lists, which can potentially broaden the list of contacts significantly (Figure 1).

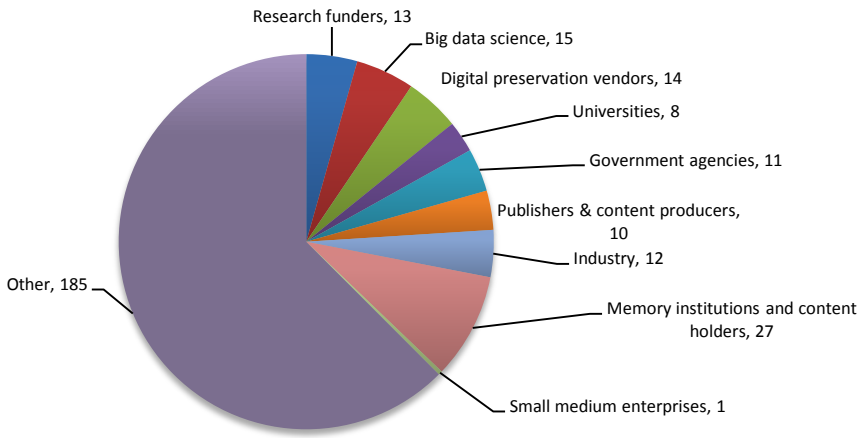

Figure 1 – Number of contacts by stakeholder category.

Figure 2 depicts the number of contacts by top-level domain. These figures include all the contacts including the ones categorised as “Other”.

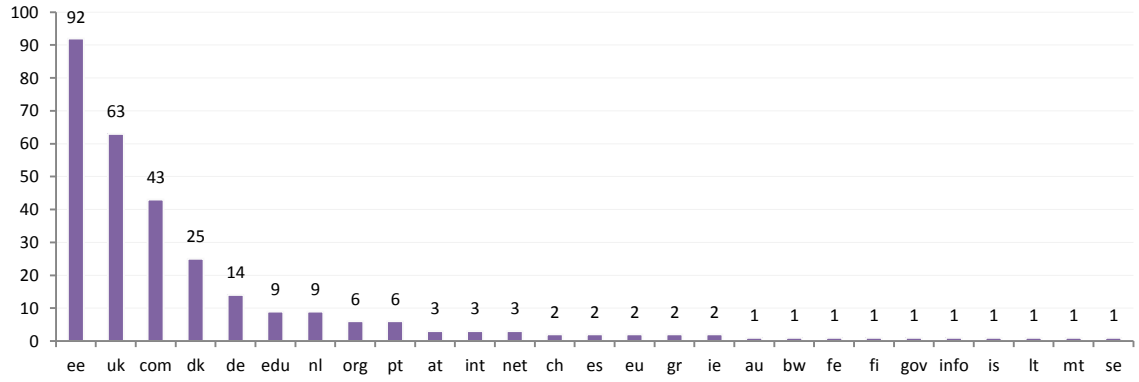

Figure 2 – Number of contacts per top-level domain.

4 Consultation of stakeholders

A series of questions was asked to assess the state of practice in gauging digital curation costs, as well as to obtain additional information regarding the most prominent challenges and needs in this area.

The questions were provided by the Engagement Group (WP2) and the Assessment Group (WP3) as the two main internal beneficiaries of the outcomes of this task.

The following sections briefly describe the methodology used in the consultation and present the questionnaire results.

4.1 Methodology

The methodology used in the consultation was as follows:

1. A call for questions was sent to all partners in the project;
2. Questions were analysed, combined and reshaped to make sure they worked well together in a single questionnaire;
3. To reduce the number of questions, an internal poll was conducted among partners in the project so that individuals could vote on questions they considered most relevant;
4. An online survey tool was deployed to support the questionnaire and to collect answers. The tool was supported by the open-source software LimeSurvey[15];
5. The online questionnaire was set up to include the most voted questions;
6. A call for participation, in the form of an email, was written and reviewed by partners of Work Package 2 (Appendix 5.1);
7. The questionnaire was tested internally by project partners to fine-tune the language, the interaction, the clarity of the questions and the closed-set answers. Additional assistance texts were also added to the questionnaire;
8. The call for participation with a total of 296 invitations was sent by each contact owner on 20 May 2013 . A reminder was sent on 13 June and the consultation closed on 21 June 2013;
9. Additionally, a blog post[16] encouraging participation was published and a link to the consultation was added to the project website[17].

[15] http://www.limesurvey.org

[16] http://4cproject.net/2013/05/13/be-part-of-the-action-collaborate-with-4c-and-help-to-clarify-the-costs-of-curation/

[17] http://4cproject.net/initial_consultation/

4.2 Questionnaire

The questionnaire was composed of three main sections:

1. **Organisation—**In this section we asked for general information to help identify and categorise the consultees.
2. **Be a part of the 4C network!—**In this section consultees were given the opportunity to leave their contact details (full name, email, organisation, position/role, main interests) so they could be contacted by the 4C project in the future.
3. **Additional questions—**This section was integrated with the first part of the consultation and allowed for a more detailed insight into the consultees' digital curation and cost modelling practices.

A complete questionnaire is available in Appendix 5.2 on page 67.

4.3 Results

As previously stated, 296 calls for participation were sent by contact owners, trusted and known to the consultees.

Of 296 calls for participation, 164 contacts opened the questionnaire (55%), with only 76 submitting their answers (25%). For the purpose of this report incomplete questionnaires were not analysed.

The following sections report on the responses provided. Some answers were removed or anonymised, and will be pointed out whenever that is the case.

Q1. What is the description that best fits your organisation?

Answer	Count	Percentage
Research funder (A1)	4	5,26%
Big data science (A2)	5	6,58%
Digital preservation vendor (A3)	7	9,21%
Government agency (A4)	10	13,16%
Publisher or content producer (A5)	3	3,95%
Data intensive industry (A6)	5	6,58%
Memory institution or content holder (A7)	18	23,68%
Small or medium enterprise (A8)	2	2,63%
University (A9)	11	14,47%
Other	11	14,47%
No answer	0	0,00%

Consultees who chose “Other” provided the following additional information:

1. University and archive
2. National health service
3. Combination of museums, libraries, archives, and scientific research organizations
4. Municipal Archive
5. Consultancy
6. Organisation for small and medium enterprise
7. Private sector
8. Research center within an industrial company
9. Infrastructure organization for the social sciences
10. Consultant researcher
11. International NGO

Q2. In which country does your organisation reside?

Answer	Count	Percentage
Austria (14)	1	1,32%
Canada (38)	1	1,32%
Denmark (57)	17	22,37%
Estonia (66)	5	6,58%
France (72)	1	1,32%
Germany (78)	7	9,21%
Ireland (100)	1	1,32%
Israel (101)	1	1,32%
Italy (102)	1	1,32%
Netherlands (145)	1	1,32%
New Zealand (148)	1	1,32%
Portugal (166)	5	6,58%
Sweden (199)	1	1,32%
Switzerland (200)	2	2,63%
United Kingdom (219)	26	34,21%
United States (220)	5	6,58%

Q3. What is your organisation's core business activity?

Answer	Count	Percentage
Public administration (A1)	11	14,47%
Research and/or education (A3)	25	32,89%
Law (A5)	1	1,32%
Social and health services (A2)	2	2,63%
Commerce and services (A4)	6	7,89%
Information technologies (A8)	6	7,89%
Consulting services (A9)	6	7,89%
Training services (10)	0	0,00%
Industry and manufacturing (A7)	3	3,95%
Other	16	21,05%
No answer	0	0,00%

Consultees who chose “Other” provided the following additional information:

1. Cultural heritage / Memory institution
2. Public archiving
3. Library services
4. Archive
5. Preservation services
6. National Library
7. Publishing
8. Information sector
9. Financial sector
10. Scientific data infrastructure
11. Archiving
12. Arts & Culture
13. Archive Storage & Services
14. National Library
15. Intellectual Property Rights
16. Digital preservation

Q4. Who are the users or customers of your organisation?

Answer	Count	Percentage
Research institutions (SQ002)	36	47,37%
Researchers and students (i.e. people) (SQ008)	51	67,11%
Archives, libraries or museums (SQ003)	23	30,26%
Citizens (SQ005)	39	51,32%
Other	22	28,95%

Consultees who chose “Other” provided the following additional information:

1. Commercial archaeologists, commercial architects, secondary and tertiary education (taught)
2. Business End Users
3. Government agencies
4. Content owners
5. Public and private sector archaeologists
6. Companies
7. Anyone interested in heritage be they academic, commercial or simply a member of the public
8. Administration
9. Government agencies
10. Enterprises
11. Publishers
12. Libraries, etc.
13. Currently internal, but may extend to customers and supply chain
14. Large variety of the above and more
15. Research projects
16. Space Agencies
17. Engineers
18. Government & Private organizations
19. Public organizations
20. Research funders
21. Research Funders
22. International Scientific Unions, data centres

Q5. What are the main funding sources for your organisation?

Answer	Count	Percentage
Public funding (SQ001)	54	71,05%
Private donations (SQ002)	12	15,79%
User fees (SQ003)	19	25,00%
Business profit (SQ004)	18	23,68%
Other	6	7,89%

Consultees who chose “Other” provided the following additional information:

1. Private sector business
2. Consultancy
3. Sales
4. Consultancy fees
5. Membership fees (these may originate in many cases from public funds)
6. ESF

Q6. What is the global annual budget for your organisation?

Answer	Count	Percentage
< 2.000 € (A1)	1	1,32%
2.001 - 10.000 € (A8)	0	0,00%
10.001 - 50.000 € (A7)	2	2,63%
50.001 - 100.000 € (A2)	1	1,32%
100.001 - 500.000 € (A3)	7	9,21%
500.001 - 1.000.000 € (A4)	9	11,84%
1.000.001 - 50.000.000 € (A5)	20	26,32%
> 50.000.000 € (A9)	20	26,32%
Don't know (A6)	16	21,05%
No answer	0	0,00%

Q8. Would you be willing to share curation cost information under confidential conditions with the 4C project?

Answer	Count	Percentage
Yes (Y)	47	61,84%
No (N)	29	38,16%
No answer	0	0,00%

Q9a. Under what conditions would you be willing to share cost information with a wider community?

Answer	Count	Percentage
Would share with no special conditions (A1)	10	21,28%
Data should be anonymized (A2)	25	53,19%
Data should not be anonymized and attribution is required (A3)	3	6,38%
Would not share with wider community (A4)	4	8,51%
Other	5	10,64%
No answer	0	0,00%

Consultees who chose “Other” provided the following additional information:

1. I have nothing to share at the moment - we do not know what our costs are
2. Not sure would need to seek permission from senior management, but in principle willing to share
3. This would be dependent on the kind and form of information required and should be discussed further
4. Filled in text here that then disappeared. Contact me.
5. Would need to discuss further

Q10. Are you interested in being contacted by the 4C project in the next couple of months for further questions and engagement activities?

Answer	Count	Percentage
Yes (Y)	61	80,26%
No (N)	15	19,74%
No answer	0	0,00%

Q11. Would you like to be informed about future activities of the 4C project?

Answer	Count	Percentage
Yes (Y)	63	82,89%
No (N)	13	17,11%
No answer	0	0,00%

Q12. Please leave your name and email for further contact.

61 people left email addresses (80% of consultees). The answers to this question are not included in this report to protect consultees' privacy.

Q13 Would you be willing to answer some additional questions about your digital curation activities, accounting and cost modelling?

Answer	Count	Percentage
Yes (Y)	46	60,53%
No (N)	14	18,42%
No answer	16	21,05%

Q14. What are the main funding sources for your digital curation activities?

Answer	Count	Percentage
Public funding (SQ14_01)	32	69,57%
Private donations (SQ14_02)	6	13,04%
User fees (SQ14_03)	9	19,57%
Business profit (SQ14_04)	9	19,57%
Other	5	10,87%

Consultees who chose "Other" provided the following additional information:

1. Don't know
2. Research and development grants
3. Should be covered by contract
4. N/A
5. Client development and support fees

Q15. What is the annual budget for your digital curation activities (including operational costs)?

Answer	Count	Percentage
< 2.000 € (A1)	3	6,52%
2.001 - 10.000 € (A2)	0	0,00%
10.001 - 50.000 € (A3)	4	8,70%
50.001 - 100.000 € (A4)	1	2,17%
100.001 - 500.000 € (A5)	5	10,87%
500.001 - 1.000.000 € (A6)	6	13,04%
> 1.000.000 € (A7)	8	17,39%
Don't know (A8)	13	28,26%
No answer	6	13,04%

Q16. How important are the digital curation activities when compared with your other business activities?

Answer	Count	Percentage
A core activity (A1)	28	60,87%
Medium importance activity (A2)	9	19,57%
Minor activity (A3)	3	6,52%
Other	3	6,52%
No answer	3	6,52%

Consultees who chose “Other” provided the following additional information:

1. Core for me; minor, if visible, for other staff members.
2. They are not an end in themselves but part of good data management
3. Organisation as a whole hasn't engaged with these issues in an integrated way to my knowledge, so it's hard to give a definitive response

Q17. Are the digital curation activities performed in-house or outsourced? Please specify the pricing model in the comment.

Answer	Count	Percentage
In-house (A1)	28	60,87%
Outsourced (A2)	3	6,52%
Partially outsourced, partially in-house (A3)	11	23,91%
Comments	16	34,78%
No answer	4	8,70%

Other comments:

1. By me in regard to collections management & digital resources, with occasional technical support from another member of staff; institutional, internally, generated data is not subject to any institutional-wide procedures or policies.
2. Still setting things up at the moment
3. Nothing is at scale enough yet to merit outsourcing except in one case - submitting theses for digitisation by the Library, which is paid for at a fixed cost per thesis.
4. In assigning international standard identifiers to sound recordings (for use in commerce, reporting and archiving) we coordinate a global network of national agencies and users themselves.
5. Our repository is hosted. Our scanning is outsourced. The data mining & metadata creation is done in-house.
6. The legally required activities pertaining to public administration materials are outsourced. Activities regarding private materials and digitised materials are performed in-house.
7. We buy commodity hardware but work with open-source software. We fund internal developers and digital curators, and are attempting to embed digital workflows across the organisation to transform areas of practice rather than create new, large teams.
8. We are managing our own long-term preservation archive and related applications around, storage and servers are provided via IaaS
9. Outsourcing: Fixed rate for agreed-upon work.
10. For the maintenance of the storage, an external data center is commissioned. All other activities are done in-house.
11. We do both, in-house and outsourced. Outsourced is paid per produced digital unit.
12. As a fractal organization, we are unlikely to have a single solution
13. It varies. We out-source digitisation of print sources (pay as we go, then sell to institutions), but manage feeds and other development of our primary research articles in-house. We are working with Dryad (and other repositories) to support curation and linking of research datasets (Dryad cost model is just emerging). The peer review process involves data curation by peer review managers, editors and reviewers (mixture of internal and external) (paid for by subscriptions and/or APCs). This last isn't managed with an overall policy or strategy in place, rather different teams, locations, workflows have gradually evolved.
14. Cost based pricing, cost pool distribution to various business areas.
15. I represent a consulting firm which is not curating itself but works with clients who do.
16. Partnership agreement with labour and overhead sharing.

Q18. What infrastructure does your organisation use for digital curation? Please give details in comment.

Answer	Count	Percentage
Digital repository system (e.g. DSpace or Eprints) (A1)	13	28,26%
In-house developed system (A4)	16	34,78%
Folder or file based system (A2)	8	17,39%
Other (A3)	6	13,04%
Comments	20	43,48%
No answer	3	6,52%

Other comments:

1. Still investigating infrastructure at the moment, but we have fedora for access to some content - no preservation back end at the moment.
2. We have a Digitool Software to a Digital Library.
3. We use the Fedora digital repository system, employing the Hydra framework over this for workflow and interface management.
4. Most nodes in the network operate a simple registry of allocated organization codes. Users themselves maintain registries of recording codes. This is being changed to a central registry which will be developed by a technical partner.
5. Fedora + Hydra, plus various tools in lightweight workflows for ingest and other forms of processing.
6. Long-term preservation system Rosetta (by Ex Libris, Israel)
7. “Own development”" dating back to 1990s.
8. As defined by the state archives. We also have what you could call a folder based system.
9. IT-Infrastructure is based on in-house data storage, which serves access, and on external data storage run by a third party provider for archive purposes. Both, the in-house and the external data storage constitute the digital curation infrastructure.
10. There is not one system for the whole organisation, but they tend to be custom build to the project/domain of the data.
11. As can be seen from my response above, we use several infrastructures, internal, external and ad-hoc.
12. File-based system in combination with a software for data management developed in-house and available under a public license.
13. Combination of DSpace repository and file based system.
14. "Centralized electronic data storage. Q19 below: Access to recently created data (max. 5 % of data pool): 10 - 20 times. Access to older (>2 weeks) data: at least 95 % of data pool: never or almost never)".
15. Digital Asset Management System based on Fedora and part of the BL system for managing Legal Deposit Material.
16. Combined with an in-house developed data management and monitoring system.
17. I represent a consulting firm which is not curating itself but works with clients who do.
18. Inhouse developed system based on a series of open source systems.
19. A variety of bespoke or community developed data management systems.
20. The LOCKSS software.

Q19. How often are assets accessed by consumers?

Answer	Count	Percentage
Never (dark archive) (A1)	5	10,87%
1 to 10 times a year (A2)	2	4,35%
11 to 100 times a year (A3)	1	2,17%
101 to 1.000 times a year (A4)	5	10,87%
1.000 to 10.000 times a year (A5)	2	4,35%
10.001 to 100.000 times a year (A6)	4	8,70%
100.000 to 1.000.000 times a year (A7)	8	17,39%
more than 1.000.000 times a year (A8)	8	17,39%
No answer	11	23,91%

Q20. How does your organisation currently breakdown the costs of digital curation activities?

Answer	Count	Percentage
The costs of digital curation is not separated from other business activities (A1)	22	47,83%
The costs of digital curation is separated from other business activities but not broken further down (A2)	7	15,22%
The costs of digital curation is broken down in several activities (A3)	6	13,04%
The costs of digital curation is broken down is several entities (e.g. by department) (A4)	6	13,04%
No answer	5	10,87%

Q21. What types of information assets do you need to curate?

Answer	Count	Percentage
Records of business activity (SQ001)	22	47,83%
Cultural heritage data (SQ002)	26	56,52%
Research data (SQ003)	25	54,35%
Digitised materials (SQ004)	34	73,91%
Scholarly publications (SQ005)	22	47,83%
Other	12	26,09%

Consultees who chose “Other” provided the following additional information:

1. Multimedia
2. Identity management data for sound recordings
3. Research data to come
4. Municipality data
5. Web-Sites, net publications
6. Design data and safety cases
7. Please not that GESIS curates other digital material as well, however, with regard to digital curation we are choosing to focus on the GESIS Data Archive here which curates mostly research data and accompanying materials (documentation etc.)
8. Engineering Data
9. Legal Deposit material
10. Personal archival material
11. None
12. Software and emulation environments

Q22. What is the motivation for keeping these assets?

Answer	Count	Percentage
Business requirement (SQ001)	20	43,48%
Legal requirement (SQ002)	28	60,87%
Ensure availability of public good (SQ003)	30	65,22%
Other	9	19,57%

Consultees who chose “Other” provided the following additional information:

1. For research data - because funders require it
2. Scientific
3. Scientific Need
4. Funder requirement
5. If we don't, we cease to be a research library in the medium-long term
6. Could be for publishing partner, such as a learned society
7. Personal interest
8. None
9. Rendering of preserved assets

Q23. Who are the producers of the assets for curation?

Answer	Count	Percentage
Ourselves (SQ001)	30	65,22%
Publishers (SQ002)	14	30,43%
Researchers (SQ003)	28	60,87%
Public administration (SQ004)	17	36,96%
Companies (SQ005)	13	28,26%
Other	12	26,09%

Consultees who chose “Other” provided the following additional information:

1. Anyone who deposits digital material in the archive (can include external companies and organisations)
2. Individuals and organisations (such as campaign groups) who deposit archives. Potentially, anyone who publishes on the Web.
3. Private organisations
4. General public
5. Repositories
6. Donators
7. Customers, claimants etc.
8. Public
9. Citizens
10. N/A
11. Private archives
12. Artists

Q24. Who are the consumers of the assets curated?

Answer	Count	Percentage
No consumers (SQ001)	0	0,00%
Ourselves (SQ002)	27	58,70%
Scholars and researchers (SQ003)	36	78,26%
Students (SQ004)	31	67,39%
Public administration (SQ005)	21	45,65%
Country citizens (SQ006)	30	65,22%
Other	9	19,57%

Consultees who chose “Other” provided the following additional information:

1. Anyone who wants to
2. Commercial archaeologists and architects
3. Mostly commercial users of the underlying assets (which we do not curate themselves)
4. Companies
5. Libraries
6. May be required by the courts
7. R&D companies, indexers, repositories, funders
8. Citizens from other countries
9. N/A

Q25. What benefits do the assets represent to your organisation?

Answer	Count	Percentage
Fulfill the institutional mission (SQ001)	37	80,43%
Fulfill institutional secondary objectives (SQ002)	6	13,04%
Fulfill legal requirements (SQ003)	28	60,87%
Provide direct monetary profit (i.e. in short-term) (SQ004)	10	21,74%
Can provide monetary profit or reduce costs in the long-term (SQ005)	10	21,74%
Document the history of the organisation (SQ006)	16	34,78%
There are no outstanding benefits (SQ007)	0	0,00%
Other	4	8,70%

Consultees who chose “Other” provided the following additional information:

1. Increase stature and reputation of the institution
2. Showcase the institutions results
3. Personal fulfilment
4. Document cultural heritage

Q26. Over what timescales does your organisation need to maintain access to the assets?

Answer	Count	Percentage
Short term storage (1-5 years) (A1)	1	2,17%
Medium term storage (5-20 years) (A2)	6	13,04%
Long term storage (infinite) (A3)	32	69,57%
Other	5	10,87%
No answer	2	4,35%

Consultees who chose "Other" provided the following additional information:

1. To end of product life - 40 to 70 years
2. Varies according to asset
3. Varies depending on Line of Business
4. The timescale depends on the nature of the assets, some assets needs long term storage
5. Highly variable, though in practise we would be planning in terms of a 5-20 yr time scale

Q27. What is the current volume and the projected yearly increase for the next 5 years of the assets kept by your organisation?

Number of files

Answer	Count	Percentage
0 to 10 (A1)	0	0,00%
11 to 100 (A7)	0	0,00%
101 to 1.000 (A8)	1	2,17%
1.001 to 10.000 (A2)	3	6,52%
10.001 to 100.000 (A9)	3	6,52%
100.001 to 1.000.000 (10)	5	10,87%
1.000.001 to 10.000.000 (A3)	7	15,22%
10.000.001 to 100.000.000 (A4)	2	4,35%
100.000.000 to 1.000.000.000 (A5)	2	4,35%
more than 1.000.000.000 (A6)	4	8,70%
No answer	19	41,30%

File increase for the next 5 years

Answer	Count	Percentage
0 to 10 (A1)	0	0,00%
11 to 100 (A7)	0	0,00%
101 to 1.000 (A8)	0	0,00%
1.001 to 10.000 (A2)	1	2,17%
10.001 to 100.000 (A9)	4	8,70%
100.001 to 1.000.000 (10)	6	13,04%
1.000.001 to 10.000.000 (A3)	5	10,87%
10.000.001 to 100.000.000 (A4)	5	10,87%
100.000.000 to 1.000.000.000 (A5)	2	4,35%
more than 1.000.000.000 (A6)	4	8,70%
No answer	19	41,30%

Volume in GB

Answer	Count	Percentage
0 to 10 (A1)	0	0,00%
11 to 100 (A7)	3	6,52%
101 to 1.000 (A8)	3	6,52%
1.001 to 10.000 (A2)	8	17,39%
10.001 to 100.000 (A9)	5	10,87%
100.001 to 1.000.000 (10)	3	6,52%
1.000.001 to 10.000.000 (A3)	1	2,17%
10.000.001 to 100.000.000 (A4)	5	10,87%
100.000.000 to 1.000.000.000 (A5)	0	0,00%
more than 1.000.000.000 (A6)	2	4,35%
No answer	16	34,78%

Volume increase for the next 5 years

Answer	Count	Percentage
0 to 10 (A1)	0	0,00%
11 to 100 (A7)	1	2,17%
101 to 1.000 (A8)	3	6,52%
1.001 to 10.000 (A2)	6	13,04%
10.001 to 100.000 (A9)	6	13,04%
100.001 to 1.000.000 (10)	3	6,52%
1.000.001 to 10.000.000 (A3)	4	8,70%
10.000.001 to 100.000.000 (A4)	3	6,52%
100.000.000 to 1.000.000.000 (A5)	1	2,17%
more than 1.000.000.000 (A6)	2	4,35%
No answer	17	36,96%

Q28. For what purposes does your organisation need financial information related to digital curation?

Answer	Count	Percentage
Accounting (A1)	16	34,78%
Documentation to meet external legal requirement (A2)	9	19,57%
Documentation for internal financial management (A3)	20	43,48%
Budgeting; i.e. balancing expenditures and financing whether this implies increasing, reducing or maintain status quo (A4)	34	73,91%
Calculation of past costs (ex post) (A5)	17	36,96%
Projection of future costs (ex ante) (A6)	37	80,43%
Charging (A7)	16	34,78%
Increase efficiency; i.e. enhance activities without compromising quality, e.g. by exploiting economies of scale or economies of scope (A8)	23	50,00%
Comparison of costs (and benefits) of alternative scenarios to support decision making (A9)	29	63,04%
Other	0	0,00%

Q29. Who is responsible for accounting and budgeting for digital curation in your organisation?

Answer	Count	Percentage
General financial / accounts manager (SQ001)	16	34,78%
Department director (SQ004)	28	60,87%
Repository manager (SQ002)	16	34,78%
Asset owner (SQ003)	3	6,52%
Other	6	13,04%

Consultees who chose “Other” provided the following additional information:

1. President
2. CEO
3. Senior managers within the university
4. Fragmented
5. Consultants
6. Executive Director

Q30. How do you determine the costs of curation in your organisation?

Answer	Count	Percentage
Never tried	16	34,78%
Tried before but failed, please specify why	2	4,35%
Experience based	18	39,13%
Checklist of costs	7	15,22%
Cost model, please specify which	7	15,22%
Other	5	10,87%

Consultees provided the following additional information:

1. Never tried
    a. Tentative investigation made, but without real depth
    b. Not part of my responsibility
2. Tried before but failed, please specify why
    a. Used LIFE in their pilot via HATII. Cost categories didn't map to our day-to-day activities for vast majority of people involved in digital curation (too much research data focus?)
    b. Too much components to calculate exact numbers
3. Experience based
    a. Number of staff involved, workload on their time, making the case for new roles and skills training
    b. Worked out staff costs of developing in house repository
    c. Internal costs of personnel (effort based)
4. Checklist of costs
    a. External costs (federal data center)
5. Cost model, please specify which
    a. CMDP for archival storage
    b. KRDS
    c. Internally developed
    d. We use the DP4lib cost model, developed in the DP4lib project
    e. Backfiles: gather costs, project sales, compile business case
    f. Adaptation of Life
    g. KRDS
6. Other
    a. Essentially 100% of costs relate to this activity
    b. Budget
    c. Don't know. Not my remit.
    d. Peer review, not calculated to my knowledge
    e. Management Accounting (Kosten-Leistungs-Rechnung, KLR)

Q31. How often does your organisation need to prepare accounts and budgets for digital curation?

Answer	Count	Percentage
Never (A1)	8	17,39%
Annually (A2)	24	52,17%
Every three to five years (A3)	2	4,35%
Other	4	8,70%
No answer	8	17,39%

Consultees who chose "Other" provided the following additional information:

1. Ad hoc, when making cases, e.g. for storage replacement cycle or new staff posts
2. Monthly/quarterly
3. Don't know. Not my remit.
4. When clients ask for it

Q32. What type of costs does your organisation need to account for?

Answer	Count	Percentage
Full economic costs (FEC) / Total costs of ownership (TCO) / lifecycle costs (SQ001)	21	45,65%
Investment costs (SQ002)	16	34,78%
Operation and maintenance costs (SQ003)	30	65,22%
Overhead costs (indirect costs) (SQ004)	23	50,00%
Other	3	6,52%

Consultees who chose "Other" provided the following additional information:

1. Don't know. Not my remit.
2. Not sure
3. Usually depends on client requirements

Q33. How do you think your organization is likely to benefit from digital curation cost modelling?

Answer	Count	Percentage
1 (1)	0	0,00%
2 (2)	2	2,82%
3 (3)	10	14,08%
4 (4)	19	26,76%
5 (5)	10	14,08%

Sum (Answers)	41
Number of cases	46
No answer	5
Arithmetic mean	3,9
Standard deviation	0,83

Q34. Select the 3 main reasons for your organisation to use a cost model.

Answer	Count	Percentage
To inform decision makers (SQ001)	36	78,26%
To find out the costs of preserving assets (SQ002)	33	71,74%
For assessing the possible options available in order to carry out digital curation activities (SQ003)	25	54,35%
Keep digital curation budget as low as possible to enable collection development while performing digital curation (SQ004)	10	21,74%
To provide information for a bid to apply to external funding (SQ005)	11	23,91%
As part of risk analysis (SQ006)	17	36,96%
In order to prioritise work (SQ007)	15	32,61%
To ensure the efficient use of resources (SQ008)	28	60,87%
To set up priced digital curation services for third parties (SQ009)	12	26,09%
Other	1	2,17%

Consultees who chose “Other” provided the following additional information:

1. Budget planning

Q35. On what basis would you select a cost model?

Answer	Count	Percentage
Model has been validated by similar organisation in your sector (SQ001)	29	63,04%
The scope of the model; e.g. covering the digital curation lifecycle (SQ002)	26	56,52%
Length of time it takes to complete it (SQ003)	11	23,91%
The information required to complete the model (SQ004)	17	36,96%
The format of the model; e.g. online tool or paper based (SQ005)	12	26,09%
Payment for the use of the model (SQ006)	13	28,26%
The support available to users of the model (SQ007)	12	26,09%
The level of detail required to complete the model (high level with limited information requiring a breakdown in costs as specified by the model) (SQ008)	14	30,43%
Is the model easy to use and adaptable (SQ009)	30	65,22%
Other	2	4,35%

Consultees who chose “Other” provided the following additional information:

1. Best practice
2. Likely applicability of the model

Q36. Have you ever tried a cost model for digital curation?

Answer	Count	Percentage
Yes (please specify which in the comment) (A1)	9	19,57%
No (A2)	30	65,22%
Comments	12	26,09%
No answer	7	15,22%

Comments provided by consultees:

1. LIFE, CMDP, NASA CET
2. http://archaeologydataservice.ac.uk/advice/chargingPolicy
3. LIFE - during the pilot by HATII
4. LIFE lifecycle cost equation
5. We analysed several cost models (LIFE3, DANS). Currently we used the DP4lib cost model.
6. I have looked at LIFE and ENSURE, which we are involved in, has also looked at cost modelling
7. We are obliged to use the KLR/Management Accounting system. This, however is not a model tailored to cost modelling in the field of digital curation/preservation; is strongly focused on products, but not curation activities.
8. Had some involvement in KRDS and have evaluated LIFE for some digitised collections
9. Adapted Life Model
10. LIFE models
11. DANS ABC model
12. KRDS

Consultees who answered “No” to this question were asked to respond to a follow up question:

What features could be drivers for using a cost model for digital curation in the future?

Answers	Count	Percentage
Checklist of asset management activities that incur cost (to know which costs are included and which are not) (SQ001)	16	53,33%
Assessment of past costs (ex post) (SQ002)	11	36,67%
Projections of future costs (ex ante) (SQ003)	22	73,33%
Assessment of benefits/value (SQ004)	19	63,33%
Comparison of costs and benefits of alternative scenarios to support decision making (SQ005)	20	66,67%
Other	0	0,00%

Q37. Was the cost model effective?

Answer	Count	Percentage
Yes (Y)	6	66,67%
No (N)	3	33,33%
No answer	0	0,00%

Consultees who answered “No” to this question provided the following additional information:

1. Imprecise, difficult to use, unadapted
2. The activities in the model didn't map to the day-to-day work of the majority of people involved in digital curation (much was logged under "other") - perhaps due to the KRDS/research data origins of the model, and that the categories seemed to be abstracted randomly - in some cases grouping several activities under one heading when more would be useful, in others giving too much irrelevant granularity.
3. Costs unrealistically volume-sensitive, and failed to account adequately for complexity of object (both ways).

Q38. How could the cost model be improved?

Consultees offered the following suggestions:

1. Better definition of tasks/activities; usability; simpler formulas
2. Need to further split out costs of activity directly related to OAIS model
3. By being relevant to the practical context we operate in, not a seemingly theoretical imagination of what is actually done.
4. More detail
5. With an effective monitoring system for all cost elements
6. LIFE was designed for library material and whilst there is some cross-over for data, it is not the same. However I was not using the results for any real life purposes, just investigating the applicability of the model to my organisation & research
7. Ours only covers cost of storage and should be extended
8. More clearly distinguish fixed and variable costs; more accurately represent costs for more complex objects...
9. Refine cost drivers, allocate the other-than-staff costs to activities, experiment with other cost objects, develop the "matrix of dataset complexity", apply economic adjustments, test reliability and accuracy, develop/customise software to make ABC easy to use

Q39. To your knowledge, is this model used by other organisations?

Answer	Count	Percentage
Yes (please specify which in the comment) (A1)	4	44,44%
No (A2)	2	22,22%
Comments	3	33,33%
No answer	3	33,33%

Comments provided by consultees:

1. Digital Antiquity in USA, and DANS in Netherlands have both adopted it for their archaeological data archiving activities
2. The cost model is perfectly matched to our own workflows. The probability this can be transferred to other organisations depends on the similarities of the workflows, preservation activities and the underlying IT-Infrastructure.
3. It is used by DANS and has been used as an input for other organisations

Q40. What is the origin of the cost model?

Answer	Count	Percentage
Existing general model (A1)	4	44,44%
Existing general model adapted for your organisation (A2)	2	22,22%
Custom made model (A3)	2	22,22%
Other	1	11,11%
No answer	0	0,00%

Consultees who chose "Other" provided the following additional information:

1. DANS ABC model was made by compiling many models in one plus custom made sections and functions

Q41. Does the model cover the activities required by your organisation in the right grouping and the right level?

Answer	Count	Percentage
Yes (A1)	1	11,11%
No (please comment why) (A2)	7	77,78%
Comments	7	77,78%
No answer	1	11,11%

Consultees who answered "No" to this question provided the following additional information:

1. Not all groups are covered, i.e. Access is missing in CMDP. Level OK.
2. As above - not refined enough in terms of groupings
3. No - see comment above
4. Further detail would be good
5. Right grouping is not by 100% (OAI). The model reflects the practice in the organisation.
6. Not sufficiently comprehensive
7. See above; also it is a corporate model rather than personal

Q42. Is this digital curation cost model integrated with other cost models used by your organisation?

Answer	Count	Percentage
Yes (please specify with which in the comment) (A1)	1	11,11%
No (A2)	8	88,89%
Comments	3	33,33%
No answer	0	0,00%

The consultees provided the following additional comments:

1. No
    a. We sometimes bid for money to catalogue collections (but not preserve them digitally - although we may like to in some circumstances) - we also have some inclusion for library support in large research funding proposals but digital curation is not part of this (at the moment)
    b. It is unclear whether this should be done
2. Yes
    a. Accounting system of DANS and its Balanced Scorecard

Q43. What features does the model include?

Answer	Count	Percentage
Checklist of asset management activities that incur cost (to know which costs are included and which are not) (SQ001)	5	55,56%
Assessment of past costs (ex post) (SQ002)	4	44,44%
Projections of future costs (ex ante) (SQ003)	7	77,78%
Assessment of benefits/value (SQ004)	0	0,00%
Comparison of costs and benefits of alternative scenarios to support decision making (SQ005)	0	0,00%
Other	3	33,33%

Consultees who chose "Other" provided the following additional information:

1. CMDP has a simplified checklist following OAIS
2. Record of staff time
3. Case by case basis

Q44. Do you have any request for additional features to the cost model?

Answer	Count	Percentage
Checklist of asset management activities that incur cost (to know which costs are included and which are not) (SQ001)	3	33,33%
Assessment of past costs (ex post) (SQ002)	2	22,22%
Projections of future costs (ex ante) (SQ003)	2	22,22%
Assessment of benefits/value (SQ004)	6	66,67%
Comparison of costs and benefits of alternative scenarios to support decision making (SQ005)	4	44,44%
Other	1	11,11%

Consultees who chose “Other” provided the following additional information:

1. The aim of the cost (budgeting, accounting, charging)

5 Conclusions and next steps

This deliverable reports on the work carried out in the 4C project to meet the objectives of Task 2.1 in the Engagement Work Package.

The goals of this task were three-fold:

1. To collect and analyse recent cost modelling and economics-related work in the field of digital curation;
2. To select an initial diverse group of stakeholders;
3. To consult stakeholders and determine their willingness to engage in future activities of the 4C project, to assess their current state of practice in curation cost modelling, and to discover gaps in existing cost models in order to improve them in the future.

To meet the first goal, a desktop research and literature review were carried out to build a baseline registry of cost modelling initiatives. The registry currently includes detailed information on 12 projects dating from 2006, 54 publications from 2007, 18 publications from 2006 or older, 15 presentations, five posters and six event reports.

The registry of all relevant initiatives will be published on the project's website and maintained throughout the projects' lifetime to keep the community well informed of any new initiatives in the field. This work will be carried out in task 2.2 in the Engagement Work Package.

In order to meet the second goal, ten different categories of stakeholders were identified and characterised; these were: research funders, big data science, digital preservation vendors, universities, government agencies, publishers & content producers, industry, memory institutions and content holders, small medium enterprises, and others.

A total of 111 personal and professional contacts were selected for each of these categories and several individual contacts from existing relations within the consortium. Additionally, 185 general contacts that included mailing lists and general-purpose contacts such as institutional emails.

The contacts were used to send a call for participation in the first 4C consultation, a questionnaire that aimed primarily at assessing the current state of practice in curation cost modelling, discovering gaps in existing cost models and inviting consultees to engage in future 4C activities (e.g. workshops, focus groups).

74 people responded to the consultation and 61 expressed their willingness to participate in future 4C activities. Responses arrived from a variety of countries, i.e. Austria (1), Canada (1), Denmark (17), Estonia (5), France (1), Germany (7), Ireland (1), Israel (1), Italy (1), Netherlands (1), New Zealand (1), Portugal (5), Sweden (1), Switzerland (2), United Kingdom (26), and United States (5).

After a brief analysis of the responses, a proposal was developed to merge the categories "Small Medium Enterprises" and "Industry" into a new one called "Small medium enterprises and industry" as the motivations and issues behind these organisations were essentially the same.

It was also suggested that a new category called "Cost model experts" should be included in the registry to represent institutions or people that developed or implemented a digital preservation cost model and were willing to use this knowledge to sell consultancy or training services to others.

This indicates that the stakeholder categories are by no means finalised. They will continue to evolve alongside engagement activities in 4C, and our growing understanding of the real issues and needs of these groups.

A recent reasoning about the stakeholder groups let us to believe that these could be merged into a smaller set of categories as it seems to group suitably the needs and channels of engagement among those participants. This assumption will be validated throughout the lifetime of the project, especially in workshops and focus groups. The new set of stakeholders is composed by:

1. **Commerce** - digital preservation vendors, publishers and content producers, small and medium enterprises;
2. **Culture** - memory institutions and content holders;
3. **Education** - universities, cost model experts;
4. **Science** - research funders, big data science;
5. **Government** - government agencies

Stakeholder categories will influence the way results and overall project communication will be delivered to these communities. Distinct “marketing” approaches are currently being discussed to address each of the stakeholder groups individually.

Appendixes

5.1 Call for participation in the consultation

The 4C consultation remains open!

The 4C project is issuing an open consultation to find people interested in clarifying the costs of curation for their organisation. It consists of an online questionnaire that will start a communication channel for further engagement. Although the first phase of the consultation closed on **June 21st**, you can still contribute so join in now!

Join the consultation

Who are we?

4C is a European research project that will help organisations across Europe to invest more effectively in digital curation and preservation. We are a consortium led by Jisc and composed of 12 other institutions including national libraries, archives, universities and small & medium enterprises (SMEs). For more information see http://4cproject.eu.

With this consultation we intend to identify and categorize organisations and people that want to participate in this project and be a part of the 4C stakeholder network.

Why participate?

We will establish and maintain an on-going dialogue with a wide-range of users, producers, funders and practitioners; in other words we will engage directly with the very people who will benefit from the resources our project will produce. Our goal is to make the results of collecting and identifying relevant initiatives and projects working in the field of long-term preservation cost models accessible to interested groups.

Who should participate?

This consultation is especially aimed at: research funding institutions, cost model experts, big data science institutions, digital preservation vendors, government agencies, publishers, content producers & holders, data intensive companies, memory institutions or in fact any person or institution that has an interest in identifying and clarifying the(ir) costs of curation.

So, if you are interested in knowing more about digital curation costs this consultation is for you!

How long does it take?

There are only 13 basic questions that will take you on average **5 minutes**. The optional set of up to 33 additional questions (to identify more precisely your needs) will take you on average 25 minutes, but only if you choose to answer them.

Remember, you can still contribute so, so join in now!

Join the consultation

The 4C project is co-funded by the European Union in the 7th Framework Programme, contract 600471.

Contact us at:

sarah@dpconline.org

contact us | check our website

5.2 Consultation questionnaire

Q1. What is the description that best fits your organisation?

Choose one of the following answers

- Research funder
- Big data science
- Digital preservation vendor
- Government agency
- Publisher or content producer
- Data intensive industry
- Memory institution or content holder
- Small or medium enterprise
- University
- Other: _____

Q2. In which country does your organisation reside?

Choose one of the following answers:

- Afghanistan
- Albania
- Algeria
- …

Q3. What is your organisation's core business activity?

Choose one of the following answers

- Public administration
- Research and/or education
- Law
- Social and health services
- Commerce and services
- Information technologies
- Consulting services
- Training services
- Industry and manufacturing
- Other: _____

Q4. Who are the users or customers of your organisation?

Check any that apply

- Research institutions
- Researchers and students (i.e. people)
- Archives, libraries or museums
- Citizens
- Other: _____

Q5. What are the main funding sources for your organisation?

Check any that apply

- Public funding
- Private donations
- User fees
- Business profit
- Other: _____

Q6. What is the global annual budget for your organisation?

Choose one of the following answers

- < 2.000 €
- 2.001 - 10.000 €
- 10.001 - 50.000 €
- 50.001 - 100.000 €
- 100.001 - 500.000 €
- 500.001 - 1.000.000 €
- 1.000.001 - 50.000.000 €
- > 50.000.000 €
- Don't know

Q8. Would you be willing to share curation cost information under confidential conditions with the 4C project?

- Yes
- No

Q9a. Under what conditions would you be willing to share cost information with a wider community?

Choose one of the following answers

- Would share with no special conditions
- Data should be anonymized
- Data should not be anonymized and attribution is required
- Would not share with wider community
- Other: _____

Q10. Are you interested in being contacted by the 4C project in the next couple of months for further questions and engagement activities?

- Yes
- No

Q11. Would you like to be informed about future activities of the 4C project?

- Yes
- No

Q12. Please leave your name and email for further contact.

- Full name: _____
- Email: _____
- Organisation: _____
- Position/Role: _____
- Main interests: _____

Q13. Would you be willing to answer some additional questions about your digital curation activities, accounting and cost modelling? There are up to 33 extra questions and will take you on average 25 minutes.

- Yes
- No
- No answer

Q14. What are the main funding sources for your digital curation activities?

Check any that apply

- Public funding
- Private donations
- User fees
- Business profit
- Other: _____

NOTE: Digital curation involves pre-ingest (appraisal, selection, preparation), ingest, data management, archival storage, preservation planning, access, repository administration and general management. Digital curation activities include but are not restricted to:

- Selection and appraisal of content to be ingested by creators and archivists;
- Content creator negotiation and legal agreement;
- Content transfer;
- Digitisation of analog content;
- Development and maintenance of descriptive classification plan;
- Repository ingest;
- Metadata extraction or production;
- Evolving provision of intellectual access;
- Storage including redundancy and backup;
- Data transformations;
- and, for some materials, a commitment to long-term preservation such as in file format migration, emulation, preservation planning and preservation watch.

Q15. What is the annual budget for your digital curation activities (including operational costs)?

Choose one of the following answers

- < 2.000 €
- 2.001 - 10.000 €
- 10.001 - 50.000 €
- 50.001 - 100.000 €
- 100.001 - 500.000 €
- 500.001 - 1.000.000 €
- > 1.000.000 €
- Don't know
- No answer

Q16. How important are the digital curation activities when compared with your other business activities?

Choose one of the following answers

- A core activity
- Medium importance activity
- Minor activity
- Other: _____
- No answer

Q17. Are the digital curation activities performed in-house or outsourced? Please specify the pricing model in the comment.

Choose one of the following answers

- In-house
- Outsourced
- Partially outsourced, partially in-house
- No answer

Q18. What infrastructure does your organisation use for digital curation? Please give details in comment.

Choose one of the following answers

- Digital repository system (e.g. DSpace or Eprints)
- In-house developed system
- Folder or file based system
- Other
- No answer

Q19. How often are assets accessed by consumers?

Choose one of the following answers

- Never (dark archive)
- 1 to 10 times a year
- 11 to 100 times a year
- 101 to 1.000 times a year
- to 10.000 times a year
- 10.001 to 100.000 times a year
- 100.000 to 1.000.000 times a year
- more than 1.000.000 times a year
- No answer

Q20. How does your organisation currently breakdown the costs of digital curation activities?

Choose one of the following answers

- The costs of digital curation is not separated from other business activities
- The costs of digital curation is separated from other business activities but not broken further down
- The costs of digital curation is broken down in several activities
- The costs of digital curation is broken down is several entities (e.g. by department)
- No answer

Q21. What types of information assets do you need to curate?

Check any that apply

- Records of business activity
- Cultural heritage data
- Research data
- Digitised materials
- Scholarly publications
- Other: _____

Q22. What is the motivation for keeping these assets?

Check any that apply

- Business requirement
- Legal requirement
- Ensure availability of public good
- Other:

Q23. Who are the producers of the assets for curation?

Check any that apply

- Ourselves
- Publishers
- Researchers
- Public administration
- Companies
- Other: _____

Q24. Who are the consumers of the assets curated?

Check any that apply

- No consumers
- Ourselves
- Scholars and researchers
- Students
- Public administration
- Country citizens
- Other: _____

Q25. What benefits do the assets represent to your organisation?

Check any that apply

- Fulfill the institutional mission
- Fulfill institutional secondary objectives
- Fulfill legal requirements
- Provide direct monetary profit (i.e. in short-term)
- Can provide monetary profit or reduce costs in the long-term
- Document the history of the organisation
- There are no outstanding benefits
- Other: _____

Q26. Over what timescales does your organisation need to maintain access to the assets?

Choose one of the following answers

- Short term storage (1-5 years)
- Medium term storage (5-20 years)
- Long term storage (infinite)
- Other: _____
- No answer

Q27. What is the current volume and the projected yearly increase for the next 5 years of the assets kept by your organisation?

	Number of files	File increase for the next 5 years	Volume in GB	Volume increase for the next 5 years
0 to 10	○	○	○	○
11 to 100	○	○	○	○
101 to 1.000	○	○	○	○
1.001 to 10.000	○	○	○	○
10.001 to 100.000	○	○	○	○
100.001 to 1.000.000	○	○	○	○
1.000.001 to 10.000.000	○	○	○	○
10.000.001 to 100.000.000	○	○	○	○
100.000.000 to 1.000.000.000	○	○	○	○
more than 1.000.000.000	○	○	○	○
No answer	⊙	⊙	⊙	⊙

Q28. For what purposes does your organisation need financial information related to digital curation?

Check any that apply

- Accounting
- Documentation to meet external legal requirement
- Documentation for internal financial management
- Budgeting; i.e. balancing expenditures and financing whether this implies increasing, reducing or maintain status quo
- Calculation of past costs (ex post)
- Projection of future costs (ex ante)
- Charging
- Increase efficiency; i.e. enhance activities without compromising quality, e.g. by exploiting economies of scale or economies of scope
- Comparison of costs (and benefits) of alternative scenarios to support decision making
- Other: _____

Q29. Who is responsible for accounting and budgeting for digital curation in your organisation?

Check any that apply

- General financial / accounts manager
- Department director
- Repository manager
- Asset owner
- Other: _____

Q30. How do you determine the costs of curation in your organisation?

Check any that apply

- Never tried
- Tried before but failed, please specify why:
- Experience based
- Checklist of costs
- Cost model, please specify which:
- Other: _____

Q31. How often does your organisation need to prepare accounts and budgets for digital curation?

Choose one of the following answers

- Never
- Annually
- Every three to five years
- Other: _____
- No answer

Q32. What type of costs does your organisation need to account for?

Check any that apply

- Full economic costs (FEC) / Total costs of ownership (TCO) / lifecycle costs
- Investment costs
- Operation and maintenance costs
- Overhead costs (indirect costs)
- Other: _____

Q33. How do you think your organization is likely to benefit from digital curation cost modelling?

Choose one of the following answers (1 = Not at all likely; 5 = Completely likely)

- 1
- 2
- 3
- 4
- 5
- No answer

Q34. Select the 3 main reasons for your organisation to use a cost model.

Check any that apply

- To inform decision makers
- To find out the costs of preserving assets
- For assessing the possible options available in order to carry out digital curation activities
- Keep digital curation budget as low as possible to enable collection development while performing digital curation
- To provide information for a bid to apply to external funding
- As part of risk analysis
- In order to prioritise work
- To ensure the efficient use of resources
- To set up priced digital curation services for third parties
- Other: _____

Q35. On what basis would you select a cost model?

Check any that apply

- Model has been validated by similar organisation in your sector
- The scope of the model; e.g. covering the digital curation lifecycle
- Length of time it takes to complete it
- The information required to complete the model
- The format of the model; e.g. online tool or paper based
- Payment for the use of the model
- The support available to users of the model
- The level of detail required to complete the model (high level with limited information requiring a breakdown in costs as specified by the model)
- Is the model easy to use and adaptable
- Other: _____

Q36. Have you ever tried a cost model for digital curation?

Choose one of the following answers

- Yes (please specify which in the comment)
- No
- No answer

Q37. Was the cost model effective?

- Yes
- No
- No answer

Q38. How could the cost model be improved?

Free text answer.

Q39. To your knowledge, is this model used by other organisations?

Choose one of the following answers

- Yes (please specify which in the comment)
- No
- No answer

Q40. What is the origin of the cost model?

Choose one of the following answers

- Existing general model
- Existing general model adapted for your organisation
- Custom made model
- Other: _____
- No answer

Q41. Does the model cover the activities required by your organisation in the right grouping and the right level?

Choose one of the following answers

- Yes
- No (please comment why)
- No answer

NOTE: "Right grouping" means here whereas the model puts costs in the correct the functional entity and role, e.g. ingest, archival storage, data management, administration or preservation planning. "Right level" means here whereas the correct level of abstraction is used on the model, i.e. how far are the activities broken down, are they only broken down at the "grouping" level, if they going into the functional level, or are they broken down even further. The right groupings and levels are put into you own perspective and consideration.

Q42. Is this digital curation cost model integrated with other cost models used by your organisation?

Choose one of the following answers

- Yes (please specify with which in the comment)
- No
- No answer

Q43. What features does the model include?

Check any that apply

- Checklist of asset management activities that incur cost (to know which costs are included and which are not)
- Assessment of past costs (ex post)
- Projections of future costs (ex ante)
- Assessment of benefits/value
- Comparison of costs and benefits of alternative scenarios to support decision making
- Other: _____

Q44. Do you have any request for additional features to the cost model?

Check any that apply

- Checklist of asset management activities that incur cost (to know which costs are included and which are not)
- Assessment of past costs (ex post)
- Projections of future costs (ex ante)
- Assessment of benefits/value
- Comparison of costs and benefits of alternative scenarios to support decision making
- Other: _____