

Collaboration to Clarify the Cost of Curation



D2.6—Report on Communications Activities

<i>Deliverable Lead:</i>	DPC
<i>Related Work package:</i>	WP2
<i>Author(s):</i>	Sarah Norris (DPC)
<i>Dissemination level:</i>	Public
<i>Submission date:</i>	31 January 2014
<i>Project Acronym:</i>	4C
<i>Website:</i>	http://4cproject.eu
<i>Call:</i>	FP7-ICT-2011-9
<i>Project Number</i>	600471
<i>Instrument:</i>	Coordination action (CA)—ERA-NET
<i>Start date of Project:</i>	01 Feb 2013
<i>Duration:</i>	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
0.01	17 Jan 2014	First draft	SN
0.02	24 Jan 2014	Second draft	SN
0.03	28 Jan 2014	Third draft	SN
0.04	30 Jan 2014	Edit	PLSS
1.00	31 Jan 2014	Released version	PLSS
1.01	07 Mar 2014	Updated logo	PLSS

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, Investigacao 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

Disclaimer: The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

{Third party acknowledgements}



D2.6—Report on Communications Activities by 4cproject.eu is licensed under a [Creative Commons Attribution-ShareAlike 3.0 Unported License](https://creativecommons.org/licenses/by-sa/3.0/).

This document reflects only the authors’ view. The European Community is not liable for any use that may be made of the information contained herein.

<i>Author(s):</i>	Sarah Norris, DPC
-------------------	-------------------

Table of Contents

Figures	6
Tables.....	8
Executive Summary	9
1 Introduction	10
2 Communication and information exchange with EC-funded and other projects and organisations	11
2.1 Agreements	11
2.2 Project Communications	12
2.3 Conferences and Webinars.....	14
2.4 Meetings and knowledge sharing.....	15
3 Stakeholder focus groups and workshops	17
3.1 Workshop and focus group (1)	18
3.2 Focus group (2)	20
4 Advisory Board Meetings	21
5 Project Website	23
5.1 Blog posts.....	25
5.2 Website analytics.....	26
5.2.1 <i>INTERIM WEBSITE (MONTHS 1 TO 6)</i>	26
5.2.2 <i>DEVELOPED WEBSITE (MONTHS 6 – 12)</i>	27
5.2.3 <i>WEBSITE STATISTICAL ANALYSIS</i>	30
6 Social Media and Analytics.....	31
6.1 Twitter statistical analysis.....	34
7 Conferences and Events	35
8 Publications	37
9 Reports	39
10 Summary of activities.....	40
11 A Tactical Look Ahead (Months 12-24)	46
11.1 Communication and information exchange with EC-funded and other projects and organisations	46
11.2 Stakeholder focus groups and workshops.....	47
11.3 Advisory Board Meetings.....	47
11.4 Project Website	47
11.5 Social Media.....	49
11.6 Conferences and Events	49
11.7 Publications	50
11.8 Reports.....	50
A Minutes of cost data meeting on 07-08 October 2013 at KB in The Hague	51
B Minutes of Collaboration Meeting with ENSURE Project, Cranfield University, 4 th November..	60
C Minutes of CERN Meeting, 7 th —8 th November 2013, CERN, Geneva, Switzerland	72
D Report on 4C Focus Group at iPRES 2013 in Lisbon, Portugal	77

E Notes from Focus Group #2—Industry Group, 12th December 2013—Jisc, Brettenham House, London..... 80

F Public report of the 1st 4C Advisory Board Meeting 11th June 2013, Jisc Office in London, UK.. 86

G Annex to D2.5: 4C Project Communications Plan—Draft ‘CCEx Communications Plan’ 89

Figures

4C entry in the SCAPE newsletter.....	12
4C entry in the TIMBUS newsletter	13
Part of the 4C presentation to APARSEN.....	14
Ensure and 4C Project Team at Cranfield University.....	16
Focus group results.....	17
iPRES invitation.....	18
The Focus Group game screenshot	24
Focus Group game results screenshot	25
Blog comments screenshot	26
Wordpress stats from the initial web site	26
Monthly activity from initial website	27
Monthly activity for the current website	28
Bounce rate of the current website	28
Visitor geographic distribution of the current website	29
4C Twitter account	31
Re-tweet statistics	32
Twitter mentions	33
Twitter “favourited” tweets	33
KB Organisation	54
ENSURE partners	60
Cost model challenges.....	61
Long term Digital preservation framework	62
Long term Digital preservation lifecycle.....	62
ENSURE work breakdown structure	62
Generic cost breakdown structure	63
Long term Digital preservation lifecycle.....	64
4C overview	65
4C timeline.....	65
Motivations for keeping data	66
Financial information need.....	66
Model evaluation—Type	67
Model evaluation—Cost variables	67

IED game results 68

Draft ESRM model 69

CCEX user types 69

Possible steps for CCEX input routine..... 70

Tables

Table 1—Summary of activities	9
Table 2—iPRES Workshop agenda	19
Table 3—iPRES Focus group agenda	19
Table 4—London Focus group agenda	20
Table 5—The Hague Advisory Board agenda (day1)	21
Table 6—The Hague Advisory Board agenda (day2)	22
Table 7—Website visits	27
Table 8—website visitor acquisition and behaviour	27
Table 9—Top website page views	29
Table 10—Geographical visitor distribution.....	30
Table 11—Geographic breakdown of Twitter followers	31
Table 12—Twitter statistics.....	32
Table 13—The most ‘retweeted tweets’	32
Table 14—Tweets most 'favourited'	34
Table 15—Conferences and events.....	35
Table 16—Publications	38
Table 17—Summary of activities.....	45
Table 18—Future focus groups and workshops.....	47
Table 19—Blog schedule	48
Table 20—Future conferences and events.....	49
Table 21—KB agenda.....	51
Table 22—KB storage tiers	56
Table 23—The results of the IED exercise at the KB	58
Table 24—KB meeting action summary	59
Table 25—ENSURE Agenda	60
Table 26—CERN agenda	72
Table 27—iPRES Focus group agenda	77
Table 28—Proposed success measures for the CCEX communications plan	90
Table 29—Key CCEX messages	91
Table 30—CCEX communications timings.....	91
Table 31—CCEX communications push strategy.....	92
Table 32—CCEX communications pull strategy.....	95

Executive Summary

This document is a review and report of communications activities over the first 12 month period with a tactical look forward across months 12 to 24.

For the purposes of this report, communications activities may be defined as those communications made with parties external to the project team and the EC.

The report follows the structure of intended activities laid out in the 4C Project Communications Plan and outlines the communications activities undertaken by partners of the 4C Project, between month 1 and Month 11¹—1 February 2013 to 31 December 2013.

A summary of activities is provided as follows, with a full detailed register of activities in Section 10—Summary of activities:

Communication and information exchange with EC-funded and other projects and organisations	Engagement with fourteen EU Projects and/or other organisations
Stakeholder focus groups and workshops	Face to face engagement with stakeholders on three occasions
Advisory Board Meetings	One Advisory Board Meeting held, with two more planned for 2014
Project Website	Thirty one substantive posts to the 4C Project Website, including deliverables, news items and blogs
Social Media	Two hundred and seventy six tweets
Conferences and Events	4C representation at twelve conferences and events
Publications	Twenty two articles published
Reports	Six reports published

Table 1—Summary of activities

The project CRM system, hosted on the project website contains 324 contacts with whom 4C has engaged. Furthermore, web statistics suggest these followers of the project are most interested in new resources to assist in their own digital curation projects. This is particularly evidenced by the response to the upload of D3.1—Evaluation of Cost Models and Needs & Gaps Analysis (MS12 Draft) which demonstrates an enormous appetite for information in this area.

The demographic reports for twitter and the web also show that there is a wide audience and appetite for the 4C Project and the information it disseminates, as well as a potential to expand stakeholder engagement by targeting contacts within identified regions.

Results of the communications activities show a good level of interest and engagement in the project and its outcomes, and generally the project is achieving its goals for communications targets set in the 4C Project Communications Plan.

¹ The report was compiled in month 12 for submission on 31 January 2014. Activities for month 12 will be added as an addendum by mid February 2014

1 Introduction

The 4C Project objectives are being achieved by a coordinated programme of outreach and engagement that are identifying existing and emerging research and analysing user requirements. This has informed an assessment of where there are gaps in the current provision of tools, frameworks and models. The project is supporting stakeholders to better understand and articulate their requirements and is clarifying some of the complexity of the relationships between cost and other factors. The outputs of this project include various stakeholder engagement and dissemination events (focus groups, workshops, and 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 enables the definition of a research and development agenda and a business engagement strategy which will be delivered in the form of a roadmap.

4C is classified by the European Commission as a ‘Coordination Action’. As such it is different from many of the large collaborative initiatives which have come before. ‘Coordination Actions’ are not funded to undertake primary research, but to assist the coordination and networking of existing projects, programmes and policies. This has two implications for 4C. Firstly it is a relatively small and relatively short-lived project that makes the most of existing research and adds to it, allowing partners to share and compare know-how rather than inventing new ways to consider the problem. Secondly, 4C is by its nature an outward-looking project that seeks to engage a large and diverse number of stakeholders. Consultation and participation—and thus communication—is a repeated theme.

As a ‘Coordination Action,’ with an emphasis on this critical theme of communication, the fulfilment of the project objectives are, therefore, the responsibility of all project partners.

This document reports activities against all the channels, audiences, information and content disseminated by the project, as well as the engagement and input it has received from its stakeholder groups. It reports the number and type of activities taken and aims to provide an analysis of their success.

2 Communication and information exchange with EC-funded and other projects and organisations

2.1 Agreements

APARSEN

The 4C Project has implemented an informal ‘communication cooperation’ agreement with the APARSEN project, which covers:

- Co-organised common webinars on specific pertinent topics, an example being the webinar delivered on 13th June 2013
- Co-organised common workshops over the course of events for knowledge and experience exchange on specific topics in digital curation, for example at the iPRES 2013 Conference.
- Distribution of newsletters and news releases using respective mailing lists (APARSEN newsletter 3 times a year)
- Promotion of events on the respective websites
- Descriptions of project/initiatives on respective websites
- Entry in the APARSEN Interactive Map of stakeholders in digital preservation

2.2 Project Communications

SCAPE

- Issue 5 of the SCAPE newsletter on 26th June 2013:
<http://us4.campaign-archive1.com/?u=20cef0f757e3840df2769745b&id=114edecf55>



Seven European countries joined together in February this year to launch the EC funded 4C Project. Designed to help public and private organisations invest more effectively in digital curation and preservation, the project aims to sustain the long-term value of all types of digital information.

4C is an 'open and social' project, and rather than waiting for polished results, partners are blogging and sharing findings as they go. The team hopes that this will encourage engagement and debate, increasing the likelihood that findings and guidance are useful. Already in six months in, 4C can tick off some major milestones. The project has undertaken a baseline study of its potential stakeholder groups and existing initiatives, and has just concluded an initial consultation designed to investigate the wide ranging needs of the practitioner community.

Still to come, 4C is running a [workshop at iPRES 2013](#), as well as organising more focus groups and outreach events over the next two years. Activities will culminate in a conference and Roadmap document to share results at the end of the process.

For more information about 4C, visit the project website and follow the blog: <http://4cproject.net>

Interview

Who are you?

My name is Krešimir Đuretec. I work as a project assistant at the Department of Software Technology and Interactive Systems, Vienna University of Technology with focus on [SCAPE](#) and [BenchmarkDP](#) projects. I am also pursuing a PhD at the same department.

What is your role in SCAPE?

My primal focus in the SCAPE project is in the Planning and Watch sub-project. There, I have been involved in the development of Scout – preservation watch component. Furthermore, I am responsible for the development of the simulation environment.

Recently, I took over the subproject lead so now my responsibilities are shifting from the development to coordination. My task will be to make sure the Planning and Watch sub-project products (Scout, Plato, C3PO, policies) integrate nicely with each other and with other products from the rest of the SCAPE project (repositories, component catalogue, web archives).

Why is your organisation involved in SCAPE?

With the knowledge in the digital preservation field and SCAPE being the follow up of the Planets project (where we were also participating) it was a logical choice for Vienna University of Technology to be part of the SCAPE project. The biggest benefit for us, as a research institution, from participating in this project is the contact with potential users of our research products. This enables us to get immediate feedback on our results and also allows us to drive our research towards real users needs.

[Read more](#)

4C entry in the SCAPE newsletter

TIMBUS

- Volume 2 Issue 1 of the TIMBUS Times Newsletter on 28th June 2013:
<http://timbusproject.net/about/publicity-material>

TIMBUS TIMES



Front Page

European Partnerships

TIMBUS is working closely with other European projects. For example, the training for advanced practitioners is organised by TIMBUS, APARSEN, SCAPE, EUDAT and the IMPACT Centre, and the training days on research process preservation are conducted by TIMBUS and Wf4Ever. We have even agreed on a common use case that lets us directly compare our tools, approaches and demos. Here are some news from some partners.

In the first week of July the 4C Project Team (Collaboration to Clarify the Costs of Curation) convenes in Frankfurt for their second quarterly meeting. Items on the agenda include a review of the information collected through the recent consultation and the cost models evaluation, as well as planning the next big steps – the workshop at iPRES2013, the Curation Costs Exchange and the Roadmap. Keep up to date with 4C's progress by following the team's blog: <http://4cproject.net/news-and-comment/>



12

4C entry in the TIMBUS newsletter

2.3 Conferences and Webinars

APARSEN Webinar

- Project Co-ordinator Neil Grindley presented an introduction to the 4C Project as part of the APARSEN Webinar entitled ‘Sustainability and Cost Models for Digital Preservation’ on the 13th June 2013:
<http://www.alliancepermanentaccess.org/index.php/aparsen/webinars/>



Collaboration to Clarify the Costs of Curation



The 4C Project – A Collaboration to Clarify the Costs of Curation

APARSEN Webinar: 13 June 2013

Neil Grindley – 4C Project Coordinator
 Jisc Programme Manager: Digital Preservation and Curation



Part of the 4C presentation to APARSEN

iPRES 2013 Conference

- 4C presented a ‘Minute Madness’ and Poster session at IPRES, and was awarded ‘Best Poster’ for the conference. Both sessions and the poster may be viewed on the 4C Project website:
Poster—<http://www.4cproject.eu/community-resources/focus-groups/ipres-workshop/4c-poster?highlight=WyJwb3NOZXliXQ==>
Poster Session—<http://vimeo.com/74101447>
Minute Madness—<http://vimeo.com/74097837>

ASIS&T PASIG

- PASIG Webinar—Presentation of 4C Project by Neil Grindley, ‘Implementing Sustainable Digital Preservation,’ 22nd October 22 2013, 11:30am-12:30pm (EDT):
<http://www.asis.org/Conferences/webinars/Webinar-PASIG-10-22-2013-register.html>

EUDAT

- Presentation of 4C Project by Kevin Ashley at EUDAT 2nd Conference, 28-30th October, Rome. Parallel Track III—Policy & Sustainability Issues:
<http://www.eudat.eu/system/files/ASHLEY%20EUDAT%2030OCT2013.pdf>

ANADP II

- Action Session 3: ‘4C Case Studies and Quantitative Data’. Implementing digital preservation solutions costs money and needs to be underpinned by sound economic arguments and robust business cases. The action session engaged participants by asking them to share stories about making the business case for digital preservation and to consider what data could usefully be shared, 19th November 2013, Barcelona
<http://www.educopia.org/events/ANADP/program>
- A summary of the Action session outcomes is available at:
<http://www.slideshare.net/neilgrindley/anadp-4-c-action-session-summary-annotated>

2.4 Meetings and knowledge sharing

APARSEN

- Meeting between 4C and APARSEN to discuss overlap between the projects and to compare approaches to cost models analysis. The outcome of this meeting helped to determine the shape and approach to T3.2—Needs & Gap Analysis report, 10th April 2013
- Contribution of APARSEN input into 4C iPRES workshop, 6th September 2013, Lisbon:
<http://ipres2013.ist.utl.pt/ws4-4C-iPRES%20Workshop%20Agenda.pdf>

Archivematica

- Meeting between 4C and Archivematica to understand more about their economic perspective on offering digital preservation services, 30th September 2013. (Notes are available to project team in the 4C Project [WP2/Meetings/Archivematica] Dropbox folder)

Koninklijke Bibliotheek (KB – National Library of the Netherlands)

- Meeting between 4C and KB, 7-8th October 2013, KB, The Hague. Minutes in Appendix A.

ENSURE Project

- Meeting between 4C and ENSURE Project, 4th November 2013, Cranfield University. Minutes in Appendix B.



Ensure and 4C Project Team at Cranfield University

CERN

- Meeting between 4C and CERN, 7th—8th November 2013, CERN Switzerland. Minutes in Appendix C.

MiLOS Project

- Meeting between 4C and the MiLOS Project, 29th November 2013, Bad Marienberg, Germany. (MiLOS have asked that slides from this meeting are kept confidential).

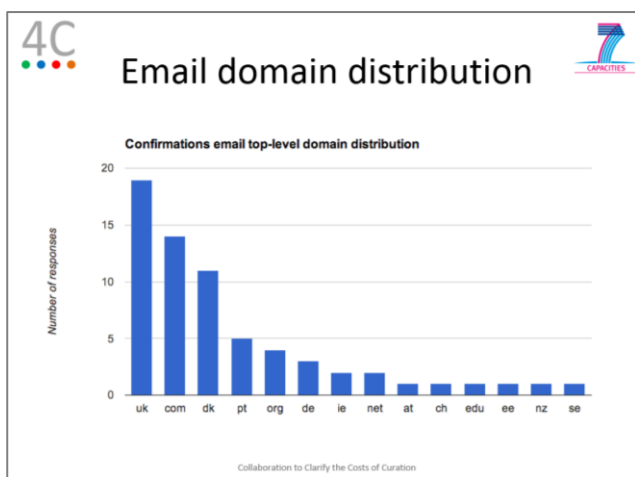
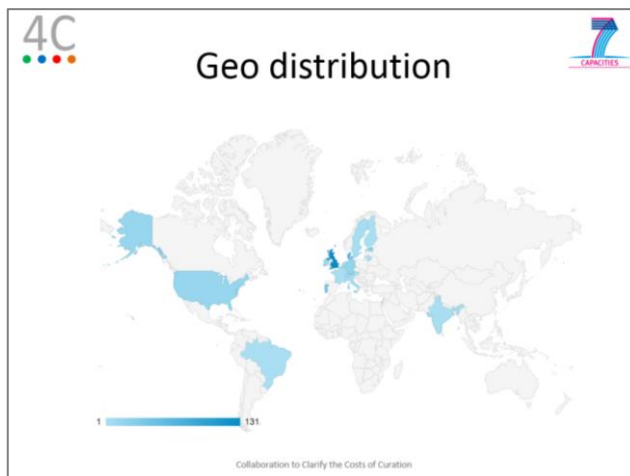
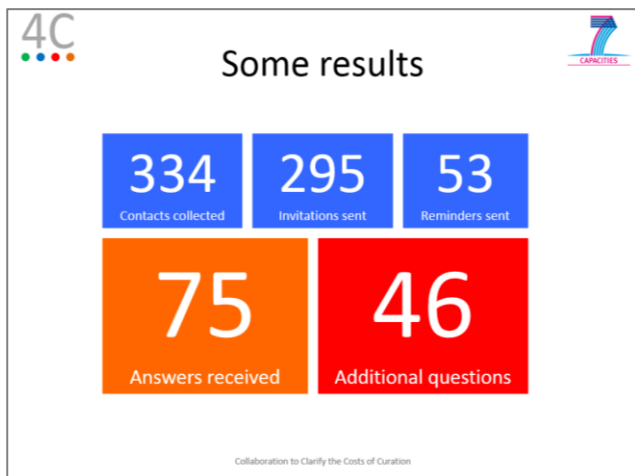
UNESCO, IFLA and ICA

- The 4C Project was featured at high-profile event organised by UNESCO, IFLA and ICA to discuss and launch a Roadmap for Digital Preservation, 5th—6th December 2013, The Hague, Netherlands
<http://www.slideshare.net/neilgrindley/unesco-30493950>
http://www.unesco.nl/sites/default/files/uploads/Comm_Info/digital_roadmap_-_report.pdf

3 Stakeholder focus groups and workshops

Early engagement was made with stakeholder groups through an Initial Consultation between 17th May and 21st June 2013. The consultation sought an indication of willingness to become involved in further knowledge exchange throughout the project duration: http://4cproject.eu/initial_consultation/

This initial consultation returned the following results, as reported by KEEPS at the 4C Project Meeting on 2nd—3rd July 2013:



Responses per stakeholder category

Type	Responses
Research funders	5
Big data science	6
Cost model experts	0
Digital preservation vendors	9
Universities	14
Government agencies	15
Publishers & content producers	5
Industry	10
Memory institutions and content holders	25
Small and medium enterprises	4
Miscellaneous contacts / Other interested parties	18

Collaboration to Clarify the Costs of Curation



Focus group results

Using the information generated by the initial consultation, invitations were sent to those stakeholders who indicated a willingness to be involved in the project, inviting them to join the 4C Project Workshop and Focus Group at iPRES in September 2013:

Dear colleague,

With regard to this year's iPRES Conference in Lisbon, the 4C project has arranged a public workshop on September 6th, 9:00 am – 1:00 pm which we would be delighted if you could attend. Following this session, we would also like to invite you to join us for a dedicated Focus Group session from 2:00-4:30 pm.

The 4C project, "Collaboration to Clarify the Cost of Curation," is a consortium of 13 partners in 7 different countries that will help organisations across Europe to gain a better understanding of digital curation costs through collaboration, and invest more effectively in digital preservation.

4C has worked on an evaluation of the relative importance of all economic determinants of digital curation. Our preliminary list includes: value, risk, benefits and sustainability as well as efficiency, reputation, interoperability, flexibility, transparency, trustworthiness, confidentiality, sensitivity etc.

With you and a handful of other experts in digital preservation and data curation, we would like to share some of our early key results and assess the priority of each of these determinants.

All of these concepts feed into the draft "Economic Sustainability Reference Model" that is being developed to support the design of strategy, to influence and standardise terminology, and to assist with the declaration of roles and responsibilities in relation to curation and preservation.

We would appreciate and value your opinion on these early results and gather your ideas on how they should be perpetuated to most fit the needs and requirements of users/customers. Additionally we would like to invite you to make suggestions on what else is to be discussed during this experts round.

Please note that this is a special and exclusive invitation to the 4C focus group in September 2013! There will be no charge for attending.

We hope you will be able to join us and look forward to hearing your thoughts.

With best regards

iPRES invitation

Invitation communications resulted in 34 attendees at the 4C Workshop and 7 attendees at the 4C Project Focus Group, covering representatives from all stakeholder groups.

3.1 Workshop and focus group (1)

The first 4C workshop and focus group were held at the 10th International Conference on Preservation of Digital Objects (iPRES) in Lisbon, Portugal on 2nd–6th September 2013.

This focus group targeted stakeholders from the following groups who were expected to attend iPRES:

- Research funders
- Cost model experts
- Universities
- Government agencies
- Data intensive industry
- Memory institutions

The agenda ran as follows:

Time	What?	Who?
09:00-09:05	Welcome	Katarina Haage, DNB
09:05-09:20	Introduction of 4C approaches, goal, priorities etc.	Neil Grindley, Jisc
09:20-09:50	Presentation of APARSEN results on analysis and testing of cost models	Kirnn Kaur, British Library
09:50-10:20	Presentation of initial results from 4C's work on assessing current cost models and tools	Ulla Bogvad Kejser, KBDK / Alex Thirifays, DNA
10:20-10:45	Q & A	Neil Grindley, Jisc
10:45-11:15	Coffee Break	
11:15-11:30	Overview of Knowledge Exchange activities with regard to digital (data) curation costs	Angela Holzer, DFG
11:30-11:45	Outcomes from the recent DCC Research Data Management Forum	Kevin Ashley, UEDIN-DCC
11:45-12:00	Summary of the Economic Sustainability Reference Model work	Neil Grindley, Jisc
12:00-12:15	Short presentation of CERN cost data	Jamie Shiers, CERN
12:15-12:50	Open discussion	William Kilbride, DPC
12:50-13:00	Wrap up	

Table 2—iPRES Workshop agenda

The agenda for the focus group ran as follows:

Time	What?	Who?
14:00-14:15	Welcome	Katarina Haage, DNB
	Short introduction of 4C approaches, goal, priorities etc.	Neil Grindley, Jisc
	Short self-introductions	
14:15-15:15	Presentation of Focus Group "Game" results	Raivo Ruusalepp, NLE
	Brief introduction of the concept	
	Group discussion in smaller groups about specific questions according to the determinants	
	Discussion of the group results in plenum	
15:15-16:15	Presentation of ESRM	Neil Grindley, Jisc
	Brief introduction to the model	
	Introduction of the 4C ESRM self-assessment questionnaire based on the model	
	Discussion on relevance and potential of the model	
	Capture recommendations from the discussion	
16:15-16:30	Wrap up	Sabine Schrimpf, DNB

Table 3—iPRES Focus group agenda

A report from focus group #1 may be found at Appendix D. A summary of the feedback contained therein is as follows:

- 'Non-memory institutions' had problems with the concept of Indirect Economic Determinants (IEDs), and how this fits with cost modelling
- 'Memory institutions' reported that they (as practitioners) would probably have a different view/prioritised assessment of the IEDs than their senior management teams, for example.
- The ESRM was not generally applicable to the group
 - As Public Sector institutions, they do not have control over many of the issues presented by the ESRM
- The ESRM was too long and too complex, and unlikely to be read by any senior management
 - This group would be better served by a ten-page checklist
- The purpose and value to be derived from the ESRM needs to be clearer.

3.2 Focus group (2)

The second focus group was held with stakeholders representing industry and commercial enterprises, in London on 12th December 2013.

An introductory webinar was held with all participants on 26th November 2013, providing an overview of the project to date, also providing more information about the focus group held a couple of weeks later.

Invitations to participate in the focus groups were sent to 56 contacts from various commercial entities, and resulted in 7 positive responses.

The agenda for the focus groups ran as follows:

Time	What?	Who?
13:00-13:10	Welcome Short introduction of 4C approaches, goal, priorities etc.	Sarah Norris, DPC Katarina Haage, DNB
13:10-13:30	Self-introduction	all participants
13:30-14:15	3 open questions as a start for the discussion Reminder Economic Sustainability Reference Model (ESRM) Introduction of the 4C ESRM self-assessment questionnaire based on the model	William Kilbride, DPC Sabine Schrimpf, DNB
14:15-14:30	Coffee break	
15:15-16:15	ESRM Exercise Discussion on relevance and potential of the model Capture recommendations from the discussion	all participants
16:15-16:30	Wrap up	Sarah Norris, DPC Sabine Schrimpf, DNB

Table 4—London Focus group agenda

A report from focus group #2 may be found at Appendix E. A summary of the feedback contained therein is as follows:

- The term ‘curation’ is not widely used in ‘industry’
- ESRM lifecycle not necessarily applicable to ‘industry’
 - Often begins upon data creation and may be no end point
- Main motivators for curation: risk avoidance, re-use and preserving corporate memory for marketing/branding etc
- The concept of value and ROI does not necessarily relate to cost
 - Avoiding cost of loss
 - Avoiding reputational loss
- Comments on the ESRM sections included:
 - Why is there no section on risk?
 - Non-homogenous file types are not necessarily chaotic
 - Resources apply not just to people, but to skills, expertise and knowledge
 - ‘Selection’ was the most troublesome issue for all parties
 - ‘Incentives’ and conflict of interest not such a great issue for ‘industry’—they build their own archives to do what they want them to do
 - Good potential for the ESRM to add value—good if it could emulate OAIS

4 Advisory Board Meetings

The first Advisory Board Meeting was held on 11th June 2013 at the Jisc Office in London, UK. In attendance were:

- Neil Grindley, Jisc
- Paul Stokes, Jisc
- Rachel Bruce, Jisc
- David Rosenthal, LOCKSS
- Matthew Addis, Arkivum
- Ron Dekker, Netherlands Organisation for Scientific Research
- Sabine Schrimpf, Deutsche National Bibliothek
- Raivo Ruusalepp, National Library of Estonia
- Sean Barker, BAE Systems
- Hildelies Balk, Koninklijke Bibliotheek
- Alex Thirifays, Danish National Archive
- Brian Lavoie, OCLC

A ‘Highlights’ report from the meeting has been published on the 4C Project website (http://www.4cproject.eu/component/docman/doc_download/11-highlights-of-advisory-board-meeting-june-2013?Itemid=) and is available in Appendix F.

The next Advisory Board Meeting is due to be held in The Hague on Wednesday 22nd January—Thursday 23rd January, 2014. The meeting will follow the third face to face project meeting on Tuesday 21st January—Wednesday 22nd January 2014.

The planned Agenda for the Advisory Board Meeting is as follows:

Wednesday 22nd January			
#	Time	Activity	Led by
	12:30 – 13:30	LUNCH	
16	13:30 – 14:00	Do you two know each other? Project Team & Advisory Board get to know each other	Neil
17	14:00 – 15:30	Project presentation Project presents what it has done and what we expect to achieve (good practice for the Review meeting in March!)	WP2
	15:30 – 15:45	BREAK	
18	15:45 – 16:30	Breakouts with the advisors Advisory Board members invited to sit down with Task Leaders and small groups to have more detailed discussions about anything they feel needs clarifying or challenging	All
19	16:30 – 17:00	Reporting and wrapping up, planning for next meeting, calendar	
	17:00 – 18:00	RECEPTION	
	19:30	ADVISORY BOARD DINNER	

Table 5—The Hague Advisory Board agenda (day1)

Thursday 23rd January			
#	Time	Activity	Led by
20	09:30 – 10:15	So what did you think? Each AB member invited to share thoughts and reactions to previous day's information	Neil
21	10:15 – 11:00	Hands on session Using some of the resources (ESRM, IED's, Gap Analysis, CCEX template, terminology list, website, forum, etc.) Advisors to work one-to-one/in small groups with 4C Team members	ALL
	11:00 – 11:30	BREAK	
22	11:30 – 13:00	If I was in charge ... Everyone gets a chance to re-write the work plan, change the budget and devise new deliverables	NG/All
	13:00 – 14:00	LUNCH	
23	14:00 – 15:00	A Roadmap to Where Exactly? Defining what strategic elements and directions should feature in the Roadmap document with reference to H2020 and other relevant initiatives	Paul
24	15:00 – 15:30	That's more than enough about 4C! Board members (and 4C partners) to report about work in the broader curation domain: new projects, new research, funding and other emerging opportunities.	
	15:30	Close the meeting	

Table 6—The Hague Advisory Board agenda (day2)

5 Project Website

The project established an early web presence by creating a temporary site using WordPress in month 1. A developed 4C Project website was completed in line with the deliverable deadline and went live on 31st July 2013 (<http://www.4cproject.eu/>)

The site structure has been developed to include features and suggestions made by the Advisory Board and Project Partners, and now comprises the following structure:

- Home
- About 4C
 - Partners
 - Advisory Board
 - Contacts
 - Credits
- Work Packages
 - Engagement
 - The 4C consultation remains open!
 - 4C workshop: “What does it cost? – EU to Assess the Cost of Digital Curation”
 - Assessment
 - Enhancement
 - Roadmap
- Community Resources
 - Outputs and Deliverables
 - Focus Groups
 - Focus Group #1 - iPRES 2013
 - 4C Poster
 - 4C Flyer
 - Gallery
 - Videos
 - Focus Group #2 - London/Frankfurt
 - Stakeholder Participation
 - ESRM Self-Assessment
 - Related Projects
 - Glossary
- News and Comment
 - Current News
 - 4C at the Ninth International Digital Curation Conference (IDCC)
 - 4C Project submits Deliverable 4.1 to the European Commission
 - 4C Project Wins Best Poster Award at iPRES 2013
 - 4C Blog
 - Press
 - Press Release: New EU collaboration to clarify the costs (and benefits) of curation
- News and Comment
- Login
- Register
- Site Map

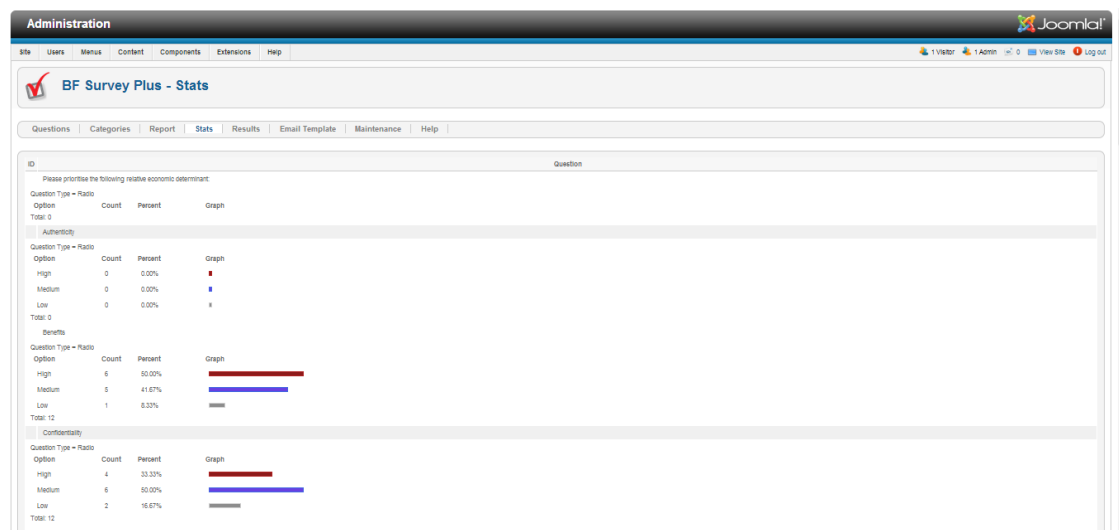
The web site will continue to grow to include more content as the project progresses and the team fulfils its objectives. Following its general release, subsequent months saw the addition of blog posts within the ‘News and Comment’ section, as well as completed deliverables uploaded to the ‘Community Resources’ section:

- D2.1—Baseline Study of Stakeholder & Stakeholder Initiatives
- D2.5—Project Communication Plan
- D3.1—Evaluation of Cost Models and Needs & Gaps Analysis (MS12 Draft)
- D4.1—A prioritised assessment of the indirect economic determinants of digital curation
- T4.2—Draft Economic Sustainability Model (Summary)
- MS7—Functioning Information Dependency Profile
- MS9—Draft Economic Sustainability Reference Model

The ‘Community Resources’ section also contains information from focus groups and workshops, as well as the ‘4C Focus Group Game’ for registered users. The 4C Focus Group Game was created as a means for participants in Focus Groups to indicate their prioritised Indirect Economic Determinants (IED).

The Focus Group game screenshot

Results for each IED are provided on the backend of the website as percentage ratios of ‘high,’ medium’ and ‘low,’ and have fed into the deliverable D4.1—A Prioritised Assessment of the Indirect Economic Determinants of Digital Curation.



Focus Group game results screenshot

5.1 Blog posts

Blogs by project team members and guests have continued to be posted on a regular basis throughout the year. The following blogs were posted in months 1 to 11:

- Let's Collaborate! by Neil Grindley, 13th March 2013
- 'There's room for everyone @4C' by William Kilbride, 18th March 2013
- 'The Age of Exploration and the Curation Costs Exchange' by Alex Thirifays, 5th April 2013
- 'Digital curation cost models for everybody' by Sabine Schrimpf, 17th April 2013
- 'Cache in the Attic' by William Kilbride, 29th April 2013
- 'Be part of the action – Collaborate with 4C and help to Clarify the Costs of Curation' by Luis Faria, 1st May 2013
- 'A very pragmatic European enterprise - reflections on cross border project involvement' by Paul Stokes, 07th May 2013
- 'Guest Blog: Digital Lifecycles and the Costs of Curation' by Paul Wheatley, 2nd June 2013
- 'Collaborating our way to success' by Kathrine Hougaard, 9th June 2013
- 'Call for Curation Cost Models' by Ulla Bøgvad Kejser, 14th June 2013
- 'Communication is key...' by Sarah Norris, 20th June 2013
- 'Nothing is Static' by Katarina Haage, 10th July 2013
- 'How do I get to where I want to be (starting from Lisbon and going via Frankfurt)?' by Paul Stokes, 18th July 2013
- '4C's Cost Model Evaluation' by Joy Davidson, 19th August 2013
- 'What the 4C Project Learnt in Lisbon' by Neil Grindley, 11th September 2013
- 'The Case of the Curious Machine' by Sarah Norris, 19th September 2013
- 'How to cut costs and keep the quality of service?' by Raivo Ruusalepp, 9th October 2013
- 'The Future of Curation Costs' by Heiko Tjalsma, 21st October 2013
- 'The Carrot and the Stick' by Matthew Addis, 28th October 2013
- 'No such thing as free digital preservation' by Jan Dalsten Sørensen, 8th November 2013
- Friends and family (or "Is there anybody out there"), 4th December 2013
- 'Please help us draw a map!' by Alex Thirifays, 17th December 2013

The blogs have elicited a number of comments as shown below:

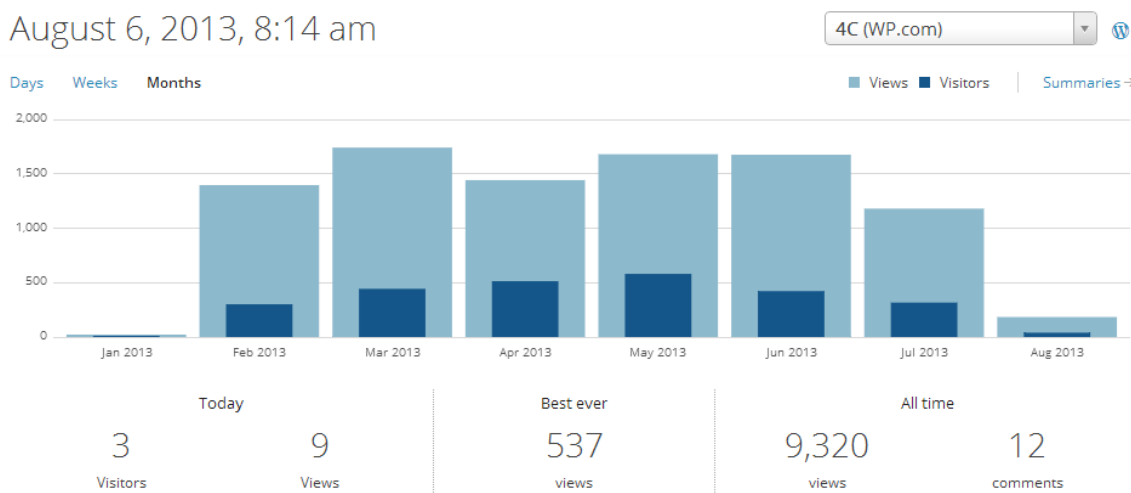
Great post Matthew! I agree that we need to emphasise the 'carrot' of digital curation as well as the 'stick'. I've come from the archive sector, where the focus is often on compliance and the legal r...	✓ Anna Henry 83.244.224.98	'The Carrot and the Stick' by Matthew Addis	[com_content]	2013-10-30 11:38:57
I agree that there are more activities involved in digital information management/curation than covered by the OAIS Reference Model; and as you note particularly prior to ingest. This is also why, in...	✉ Ulla Bøgvad Kejser 188.65.97.60	'Call for Curation Cost Models' by Ulla Bøgvad Kejser	[com_content]	2013-08-01 09:54:27
Although you define curation as equivalent to digital information asset management, you then go on to define the activities of curation in terms of OAIS. This is a very serious mistake, as the OAIS is...	✉ Chris Rusbridge 188.65.97.60	'Call for Curation Cost Models' by Ulla Bøgvad Kejser	[com_content]	2013-08-01 09:53:21
There are a few more here... http://wiki.o-pf-labs.org/display/CDP/Home... but you know about that list already *:-)	✉ Paul Wheatley 188.65.97.60	'Call for Curation Cost Models' by Ulla Bøgvad Kejser	[com_content]	2013-08-01 09:50:29
1) The title is a play on the BBC daytime TV show called 'Cash in the Attic' where contestants find unwanted objects in their homes and sell them at auction, and the idea of a computer 'cache'. The qu...	✉ 4C Project 188.65.97.60	'Cache in the Attic' by William Kilbride	[com_content]	2013-08-01 09:48:42
1) Is there a hidden meaning for the title? 2) 'Our goal to have an impact on the real world, now and in the future' is a must for a world that judges performance in limited time frames. You mention t...	✉ Lert 188.65.97.60	'Cache in the Attic' by William Kilbride	[com_content]	2013-08-01 09:20:45
Hi Eliot, thanks very much for getting in touch and making sure we have got ENSURE on our radar. yours is definitely one of the projects that 4C needs to take account of and build on over the next cou...	✉ 4C Project 188.65.97.60	There's room for everyone @4C	[com_content]	2013-08-01 08:54:44
ENSURE (www.ensure-fp7.eu) is an FP7 project in long term digital preservation, and is now in its third year. We have been developing a research prototype to create a system for LTDP, which uses cost...	✉ Eliot Salant 188.65.97.60	There's room for everyone @4C	[com_content]	2013-08-01 08:46:58

Blog comments screenshot

5.2 Website analytics

5.2.1 Interim Website (Months 1 to 6)

During the first 6 months of the project, WordPress Stats shows the following overview of activity on the temporary WordPress site:



Wordpress stats from the initial web site

Month by month activity is shown in the chart below, with March as the month with the most traffic through the site.

The March peak correlates with the co-ordinated issue of news releases² by all thirteen project partners around this time, which generated initial interest in the project.

² News releases issued by Project Partners are detailed within Section 7 – Publications.

August 6, 2013, 8:18 am

4C (WP.com)

[« Return to Stats](#)

Months and Years

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2013	22	1,396	1,741	1,442	1,680	1,675	1,180	184					9,320

Average per Day

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Overall
2013	1	50	56	48	54	56	38	35					45

Monthly activity from initial website

5.2.2 Developed Website (Months 6 – 12)

Once the developed site went live on 31st July 2013, Google Analytics was used to provide an overview of the new website activity. During the period 31 July to 31 December 2013, the website has seen 2,315 hits, with a breakdown as follows:

	Visits
Returning Visitors	1,219
New Visitors	1,096

Table 7—Website visits

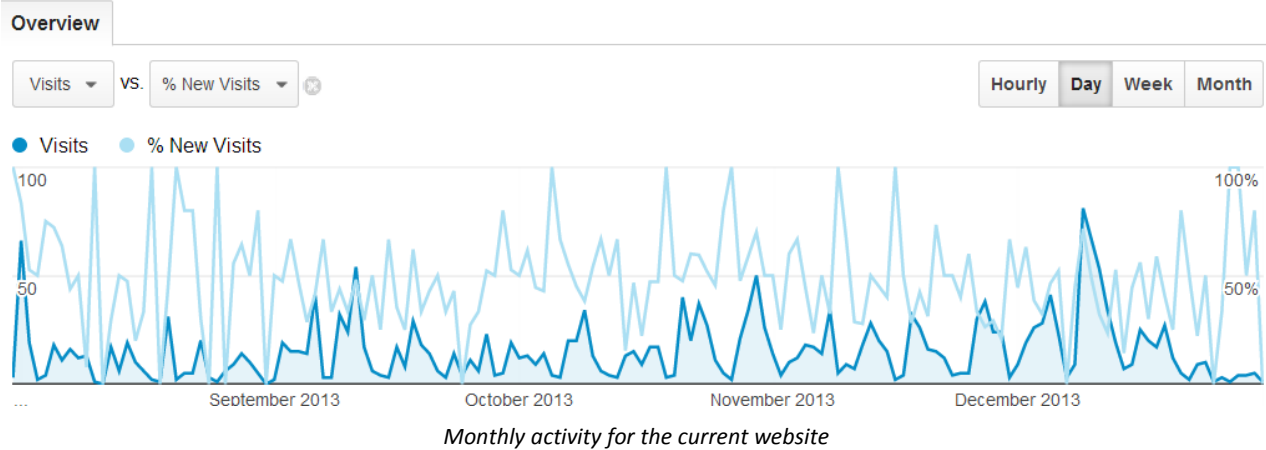
Based on 558 hits per month, as an indicator for a small enterprise³, the 4C Project website is generating traffic commensurate with the average for an organisation of its type.

A breakdown of visitor acquisition and behaviour is shown below:

	Acquisition			Behaviour		
	Visits	% New visits	New visits	Bounce rate	Pages/visit	Average visit duration
	2,315	47.34%	1,096	49.29%	3.62	00:04:04
Direct	939			50.59%		
Referral	636			37.26%		
Social	445			55.28%		
Organic Search	290			61.38%		
Email	5			100%		

Table 8—website visitor acquisition and behaviour

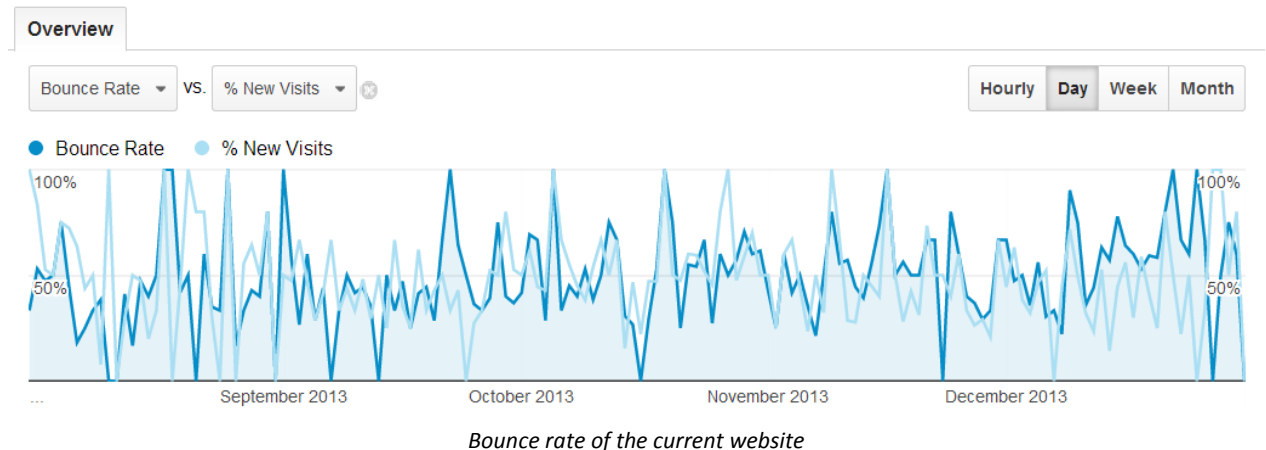
³ Jones, M, 'How many visitors should your site get?' <http://blog.hubspot.com/blog/tabid/6307/bid/5092/How-Many-Visitors-Should-Your-Site-Get.aspx>, 2009



Month by month activity is demonstrated in the chart above, showing the following spikes in activity:

- 1st August 2013 (July 31st)—launch of the new site
- 11th September 2013—upload of 4C's iPRES information
- 30th October 2013—Matthew Addis's blog post 'The Carrot and the Stick'
- 9th December 2013—upload of D3.1—Evaluation of Cost Models and Needs & Gaps Analysis

'Bounce rate' is expressed as a percentage and represents the proportion of visits that end on the first page of the website that the visitor sees.



High bounce rates typically indicate that the website is not doing a good job of attracting the continued interest of visitors. Between 40 to 60%⁴ bounce rate is average; anything in excess of 80 per cent represents a major problem. The site's average bounce rate is 49.9%, which is normal for this type of site.

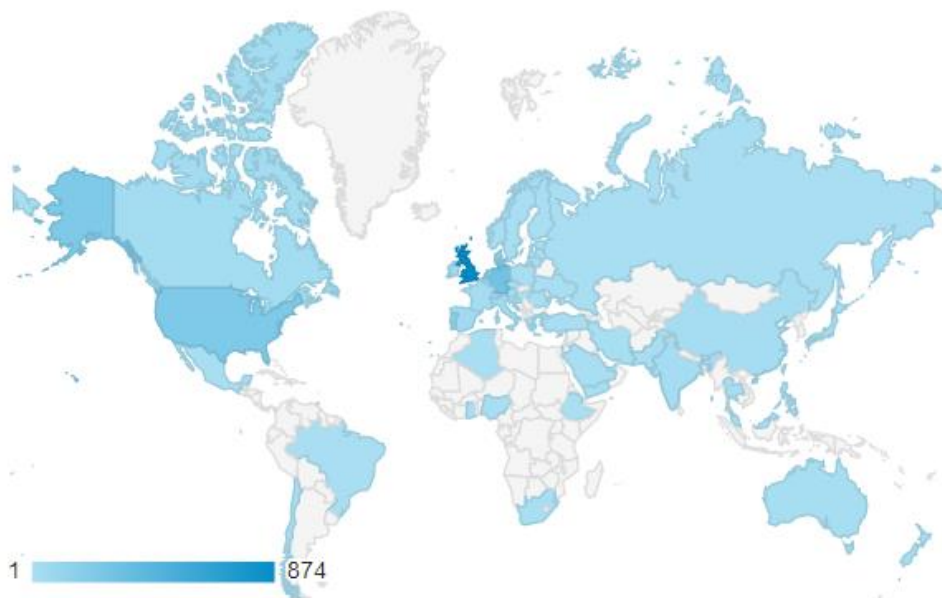
⁴ Anders Analytics, 'What is an average bounce rate,' <http://www.andersanalytics.com/blog/153-what-is-an-average-bounce-rate>, 2010

Pages on the site were viewed a total of 8,384 times. The top ten highest page-views are as follows:

Page Title	Page views	% Page views
1. Home	1854	22.11
2. Community resources	583	6.95
3. About 4C	449	5.36
4. 4C Blog	359	4.64
5. Work packages	310	3.70
6. News and Comment	309	3.69
7. D3.1—Evaluation of Cost models and Needs & Gap Analysis	262	3.12
8. The 4C Focus Group Game	176	2.10
9. Login	160	1.91
10. 'The Carrot and the Stick' by Matthew Addis	150	1.79

Table 9—Top website page views

The majority of visitors within the period 31st July to 31st December 2013 come from the UK, with the top ten visitor origins shown below.



Visitor geographic distribution of the current website

The majority of top ten countries are those represented by the 4C Project team. United States, Belgium, Spain, Canada and Ireland (highlighted in green on the table below) represent visitors independent of the 4C Project.

	Acquisition		
	Visits	% New Visits	New Visits
	2,315 % of Total: 100.00% (2,315)	47.34% Site Avg: 47.34% (0.00%)	1,096 % of Total: 100.00% (1,096)
United Kingdom	874	34.90%	305
Germany	290	33.45%	97
United States	229	77.73%	178
Netherlands	162	54.32%	88
Denmark	149	33.56%	50
Portugal	132	39.39%	52
Belgium	67	68.66%	46
Spain	41	39.02%	16
Canada	34	70.59%	24
Ireland	34	73.53%	25

Table 10—Geographical visitor distribution

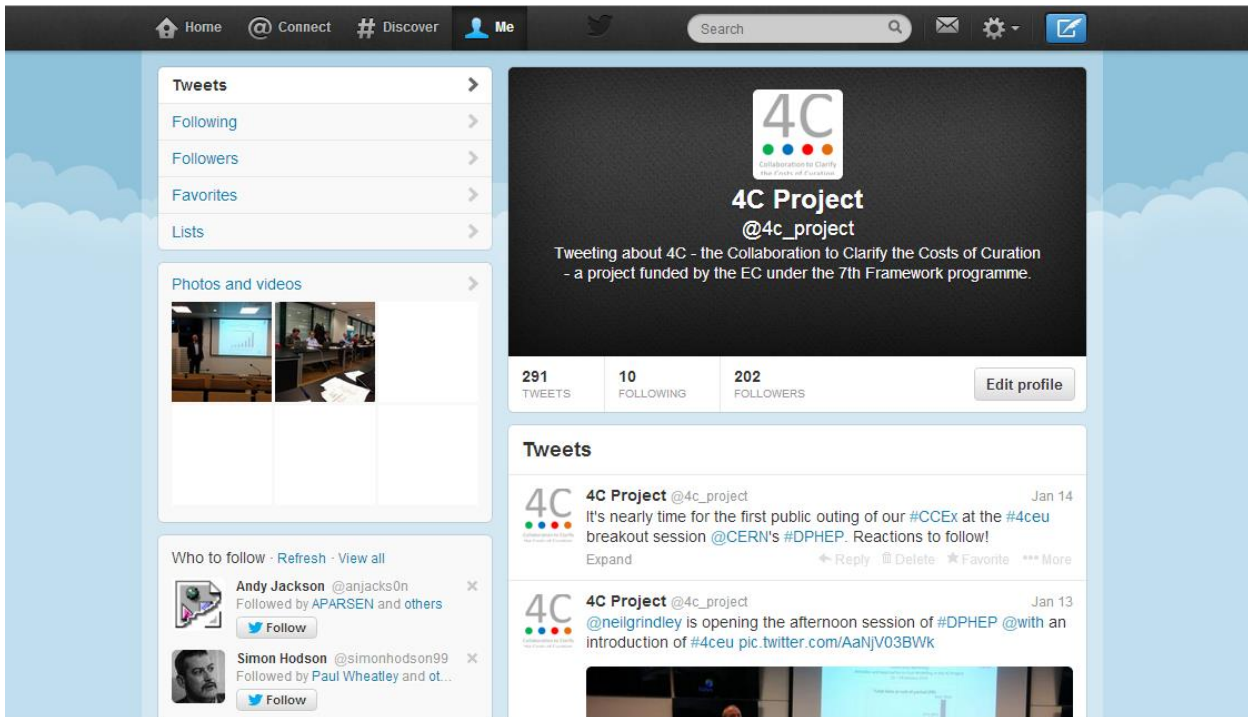
5.2.3 Website statistical analysis

The statistics show suggest the visitors are predominantly driven by the addition of new content to the website, as publicised through twitter and digital preservation email discussion lists. This is particularly evidenced by the response to the upload of D3.1—Evaluation of Cost Models and Needs & Gaps Analysis (MS12 Draft) which demonstrates an enormous appetite for information in this area.

The demographic report also highlights the potential to expand stakeholder engagement by targeting contacts within the United States, Belgium, Spain, Canada and Ireland.

6 Social Media and Analytics

The project has established a twitter account; '@4c_project' and a hashtag; '#4ceu.'



4C Twitter account

In the first 12 months of the project, the 4C Project has made 276 tweets (on average, just under 1 tweet per day) and gained 199 followers. A geographical breakdown of followers is:

Country of Origin	Number	Country of Origin	Number
UK	64	Spain	3
USA	34	Czech Republic	2
Germany	15	Finland	2
Netherlands	15	France	2
Europe	8	International	2
Greece	6	Italy	2
Ireland	6	New Zealand	2
Portugal	6	Unknown	2
Denmark	5	Greenland	1
Belgium	4	Mexico	1
Canada	4	Norway	1
Australia	3	Peru	1
Austria	3	Sweden	1
Japan	3	Tonga	1

Table 11—Geographic breakdown of Twitter followers

This social media presence has been used to promote website content and project activity, and is represented in the following breakdown of hashtag uses, retweets, mentions and favourites:

Total tweets	276
Tweets per day	0.83
User mentions	166 (0.60% per tweet)
Links	116 (0.42% per tweet)
Tweets retweeted	115 tweets retweeted (41.67% of tweets), a total of 198 times (1.72 per retweeted tweet)
Tweets favourited	32 tweets favourited (11.59% of tweets), a total of 38 times (1.19 per favourite tweet)
Hashtag uses	368 (1.33 per tweet)

Table 12—Twitter statistics

The most ‘retweeted tweets’ were those shown below, with a spike in activity corresponding to a number of tweets about iPRES being retweeted once or twice:
















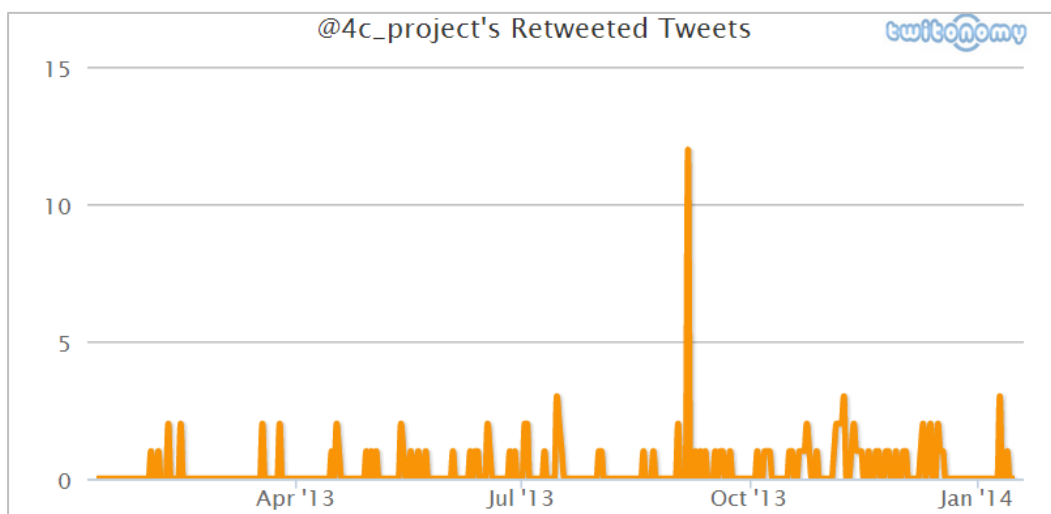
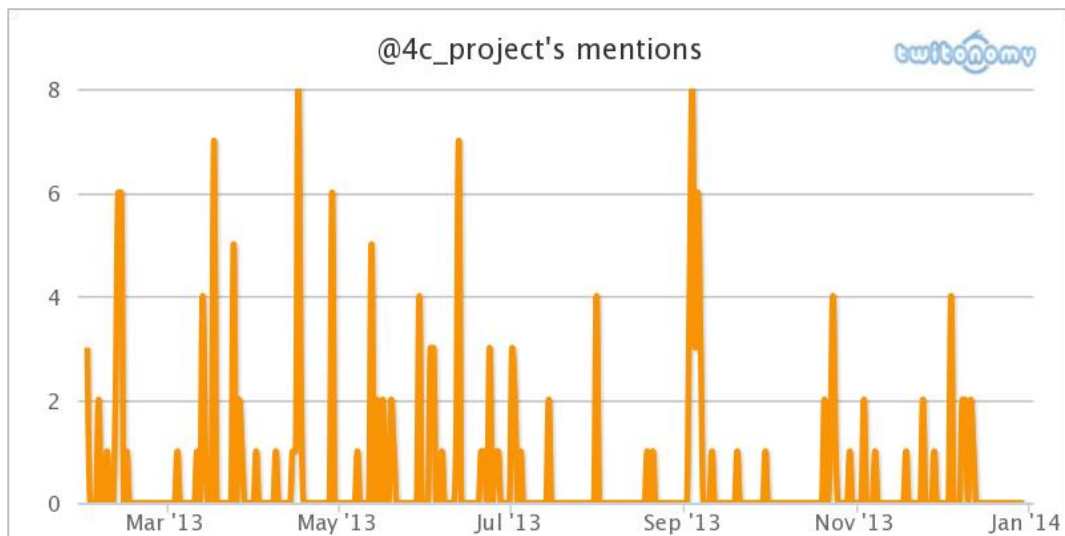
- 
4C Project @4c_project March 25, 2013, 1:05 am via WordPress.com  9  0
 Press Release: New EU collaboration to clarify the costs (and benefits) of curation wp.me/p336nA-26
- 
4C Project @4c_project October 23, 2013, 12:48 pm via TweetDeck  9  2
 We've added a registry of work on digital curation costs to our website. Find out more and leave your suggestions: 4cproject.eu/community-reso...
- 
4C Project @4c_project February 8, 2013, 3:07 pm via TweetDeck  8  4
 There's a lot written about costs in digital preservation. the list by [@prwheatley](#) and [@anjacks0nbit.ly/11uuE8d](#) is very handy!
- 
4C Project @4c_project December 9, 2013, 9:37 am via TweetDeck  8  1
 A new [#4ceu](#) deliverable completed! Read 'Evaluation of Cost Models & Needs & Gaps Analysis' & tell us your thoughts: 4cproject.eu/community-reso...
- 
4C Project @4c_project May 20, 2013, 9:35 am via TweetDeck  5  1
 The 4C Consultation is now open! Get involved and have your say by visiting the website now: 4cproject.net/initial_consul... [#4ceu](#)

Table 13—The most ‘retweeted tweets’



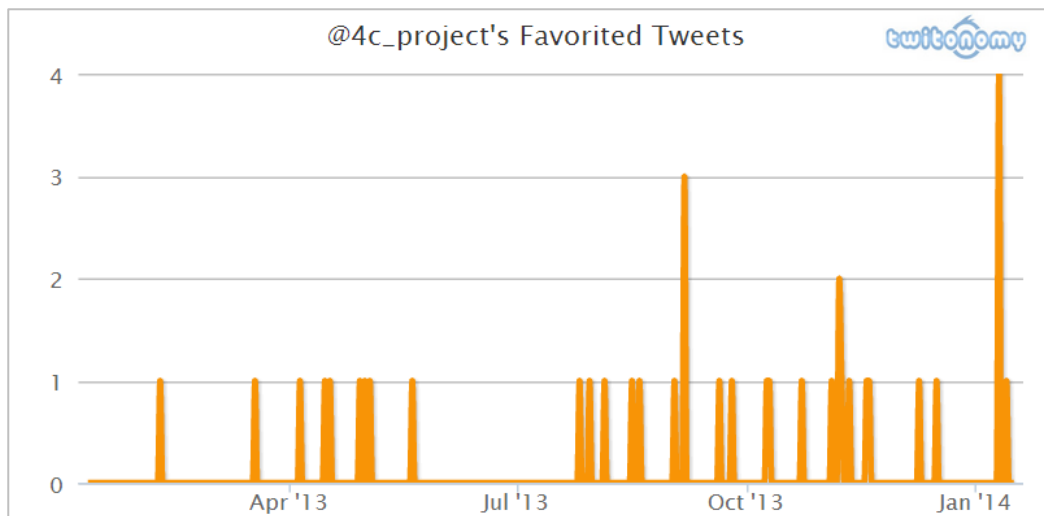
Re-tweet statistics

Month by month 'mention' activity is demonstrated in the chart below, with the spike in activity on 17th April 2013 coinciding with the post of 'Digital Curation Cost Models for Everybody' blog by Sabine Schimpf and mentions by @TIMBUSProject which were retweeted, as well as on 4th September 2013 corresponding with 4C Project's award of 'Best Poster' at iPRES 2013.



Twitter mentions

Month by month 'favourite' activity is shown in the chart below, with the spikes in activity corresponding to a number of tweets about iPRES being each being favourite on 6th September 2013, and tweets directed at the Business Archives Council during their Conference as well as a tweets relating to a visit to CERN on 7th November 2013:



Twitter "favourited" tweets

Tweets most favoured:



4C Project @4c_project February 8, 2013, 3:07 pm via TweetDeck 8 4
There's a lot written about costs in digital preservation. the list by [@prwheatley](#) and [@anjacks0nbit.ly/11uuE8d](#) is very handy!



4C Project @4c_project November 19, 2013, 9:39 am via TweetDeck 0 2
[#4ceu](#) is at [#anadp13](#) today. We're inviting you to share stories about making business cases for [#digitalpreservation](#): [educopia.org/events/ANADPII](#)



4C Project @4c_project October 23, 2013, 12:48 pm via TweetDeck 9 2
We've added a registry of work on digital curation costs to our website. Find out more and leave your suggestions: [4cproject.eu/community-reso...](#)



4C Project @4c_project August 19, 2013, 10:48 am via TweetDeck 5 2
A new blog hot off the press for your Monday morning! Tell us what you think about our Cost Model Evaluation: [4cproject.eu/news-and-commen... #4ceu](#)



4C Project @4c_project August 5, 2013, 9:30 am via TweetDeck 0 1
Another set of deliverables completed! Check out our new [#4ceu](#) website & 'Baseline Study of Stakeholder Initiatives': [4cproject.eu](#)

Table 14—Tweets most 'favoured'

6.1 Twitter statistical analysis

The twitter statistics corroborate those from the website which suggests the greatest interest is in content driven tweets, directing followers to useable resources and information.

The demographic report for twitter also shows that there is a wide audience and appetite for the 4C Project and the information it disseminates, as well as a potential to expand stakeholder engagement by targeting contacts within the identified regions.

7 Conferences and Events

The 4C project has been represented at the following conferences and events:

Conference/ Event	Location	Date	4C Partner & Affiliation	Activity
IS&T Archiving Conference 2013	Washington DC, USA	2 nd April 2013	Diogo Proença, INESC-ID	Paper – ‘The role of risk analysis to support cost models for digital preservation’
DigCurV Conference	Florence, Italy	6 th May 2013	Neil Grindley, Jisc	Keynote Presentation - ‘Economics of Digital Curation Training and Education’
APARSEN Training Webinar	Web	13 th June 2013	Neil Grindley, Jisc	Paper - ‘Sustainability & Cost Models for Digital Preservation
IFLA	Singapore	17 th -23 rd August 2013	Krista Kiisa, NLE	Flyers distributed
iPRES 2013	Lisbon, Portugal	2 nd – 6 th September 2013	Katarina Haage, DNB	Poster Session, Minute Madness, Workshop and Focus Group
TPDL2013	Malta	23 rd – 26 th September 2013	Raivo Ruusalepp, NLE	Flyers distributed
Frankfurt Book Fair 2013	Frankfurt, Germany	9 th – 13 October 2013	Katarina Haage, DNB	Flyers distributed
ASIS&T PASIG Webinar	Web	22 nd October 2013	Neil Grindley, Jisc	Paper - ‘Implementing Sustainable Digital Preservation
EUDAT 2nd Conference	Rome, Italy	29 th October 2013	Kevin Ashley, UEDIN-DCC	Paper - ‘Parallel Track III - Policy & Sustainability Issues
Business Archives Conference Council 2013	Nottingham, UK	7 th November 2013	Sarah Norris, DPC	Flyers distributed
ANADP II Conference	Barcelona, Spain	18 th – 20 th November 2013	Neil Grindley, Jisc Raivo Ruusalepp, NLE	Panel Session: “How can we employ the resources we have available to us most economically to achieve our digital preservation objectives?” Action Session: ‘4C Case Studies and Quantitative Data.’
Institute of Historical Research	London, UK	28 th November 2013	Neil Grindley, Jisc	Panel session on ‘Sustainability’
UNESCO Roadmap for Long Term Access to Digital Heritage Meeting	The Hague, Netherlands	5 th – 6 th December 2013	Neil Grindley, Jisc	Invited Expert Presentation – to provide an economic perspective on digital preservation

Table 15—Conferences and events

Other informal dissemination on the 4C Project's activities was made at the following events:

- APARSEN, EUDAT, Impact, SCAPE and TIMBUS APC Training in Glasgow, UK on 15th to 19th July 2013
- DPC's 'Getting Started in Digital Preservation' Workshops in London, Glasgow, Belfast, Dublin and Aberystwyth, UK between March and December 2013
- Archiving Tomorrow 2013 in Manchester, UK on 22nd to 23rd November 2013

8 Publications

The aims, objectives and descriptions of the 4C Project have been published in various articles, news releases and newsletters on partner and other websites.

Title	Publication/ link	Date	Author/ Affiliation
New EU project examines the costs of digital preservation	http://digitalbevaring.dk/nyt-eu-projekt-undersoger-omkostningerne-ved-digital-bevaring/	7 th February 2013	Kathrine Hougaard, DNA
Press release on the EU project 4C	http://digitalbevaring.dk/pressemeddelelse-4c/	23 rd March 2013	Kathrine Hougaard, DNA
4C Project—the Collaboration to Clarify the Costs of Curation—a project within the Digital Preservation area	http://www.sba-research.org/research/data-security-and-privacy/digital-preservation/4c-project-the-collaboration-to-clarify-the-costs-of-curation/	25 th March 2013	Stephan Strodl, SBA
Project 4C—Collaboration to Clarify the Costs of Curation	http://www.keep.pt/projeto-4c-collaboration-to-clarify-the-costs-of-curation	25 th March 2013	Miguel Ferreira, KEEPS
DPC joins new EU collaboration to clarify the costs (and benefits) of curation	http://www.dpconline.org/newsroom/latest-news/978-dpc-joins-major-new-eu-initiative-to-understand-the-costs-and-benefits-of-digital-curation	25 th March 2013	William Kilbride, DPC
Collaboration to Clarify the Costs of Curation (4C)	http://www.dcc.ac.uk/projects/4c	25 th March 2013	Joy Davidson, DCC
13 partners from across Europe join together to improve digital curation	http://www.jisc.ac.uk/news/stories/2013/03/4C.aspx	25 th March 2013	Paul Stokes, JISC
Project 4C: the Collaboration to Clarify the Costs of Curation	http://www.dans.knaw.nl/content/categorieen/projecten/project-4c-collaboration-clarify-costs-curation	25 th March 2013	Heiko Tjalsma, DANS
Major EU initiative to clarify the costs and benefits of digital curation	http://www.data-archive.ac.uk/news-events/news.aspx?id=3466	27 th March 2013	Herve L'Hours, UKDA
Digital preservation—what it will cost and what the benefits are?	http://www.nlib.ee/en/mis-kasusaab-digitaalsest-sailitamisest/	01 st April 2013	Raivo Ruusalepp, NLE
Major new EU initiative to understand the costs and benefits of digital curation	http://www.nlib.ee/major-new-eu-initiative-to-understand-the-costs-and-benefits-of-digital-curation/	01 st April 2013	Raivo Ruusalepp, NLE
PROJECT STARTED—4C	http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/Kurz-meldungen/4C.html	18 th April 2013	Katarina Haage, DNB

Title	Publication/ link	Date	Author/ Affiliation
4C - EU project started on the costs and benefits of digital preservation	http://files.d-nb.de/nestor/newsletter/nestor-Newsletter_28.pdf	01 st May 2013	Sabine Schrimpf, DNB
The 4C consultation is now open!	http://www.dpconline.org/newsroom/latest-news/1014-4c-consultation-now-live	17 th May 2013	Sarah Norris, DPC
The 4C consultation is now open!	http://www.keep.pt/estudo-para-a-quantificacao-de-custos-associados-a-preservacao-digital-convite-a-participacao	20 th May 2013	Miguel Ferreira, KEEPS
Danish national Archives and the Royal Library appreciate the cultural heritage:	http://www.sa.dk/content/dk/om_statens_arkiver/nyhedsoversigt/statens_arkiver_og_det_kongelige_bibliotek_satter_pris_pa_den_digital_e_kulturarv	26 th June 2013	Kathrine Hougaard, DNA
4C—Collaboration to Clarify the Costs of Curation	Issue 5—SCAPE Newsletter: http://us4.campaign-archive1.com/?u=20cef0f757e3840df2769745b&id=114edecf55	26 th June 2013	Sarah Norris, DPC
European Partnerships	TIMBUS Times: http://timbusproject.net/about/publicity-material	28 th June 2013	Sarah Norris, DPC
New EU project examines the costs of digital preservation	http://digitalbevaring.dk/nyt-eu-projekt-undersoger-omkostningerne-ved-digital-bevaring/	02 nd July 2013	Kathrine Hougaard, DNA
4C—Collaboration to Clarify the Costs of Curation	http://www.dlib.org/dlib/july13/07inbrief.html	13 th July 2013	Sarah Norris, DPC

Table 16—Publications

9 Reports

Reports for project meetings with other projects/organisations as well as the Advisory Board, are provided at Appendices A to F:

- Appendix A—Minutes of cost data meeting on 07-08 October 2013 at KB in The Hague
- Appendix B—Minutes of Collaboration Meeting with ENSURE Project, Cranfield University, 4th November
- Appendix C—Minutes of CERN Meeting, 7th—8th November 2013, CERN, Geneva, Switzerland
- Appendix D—Report on 4C Focus Group at iPRES 2013 in Lisbon, Portugal
- Appendix E—Notes from Focus Group #2—Industry Group, 12th December 2013—Jisc, Brettenham House, London
- Appendix F—Public report of the 1st 4C Advisory Board Meeting

10 Summary of activities

Date	Location/ Country	Title	Audience	Partner
1 Communication and information exchange with EC-funded and other projects and organisations				
10 th April 2013	UK	APARSEN Meeting	APARSEN Project	Jisc
26 th June 2013	UK	Issue 5—SCAPE Newsletter	SCAPE Project	DPC
28 th June 2013	Europe	Volume 2 Issue 1- TIMBUS Times	TIMBUS Project	DPC
13 th June 2013	Europe	Sustainability & Cost Models for Digital Preservation	APARSEN Project	Jisc
6 th September 2013	Lisbon, Portugal	APARSEN Presentation, 4C Workshop, iPRES 2013: http://ipres2013.ist.utl.pt/ws4-4C-iPRES%20Workshop%20Agenda.pdf	Stakeholder Groups	ALL
2 nd —6 th September 2013	Lisbon, Portugal	Poster, Poster Session and Minute Madness at IPRES 2013: Poster: http://www.4cproject.eu/community-resources/focus-groups/ipres-workshop/4c-poster?highlight=WyJwb3NOZXliXQ== Poster Session: http://vimeo.com/74101447 Minute Madness: http://vimeo.com/74097837	Stakeholder Groups	Jisc
30 th September 2013	UK	Archivemata Meeting (Notes stored in 4C Project Dropbox folder)	Archivemata	Jisc
22 nd October 2013	Online	ASIS&T PASIG, Webinar - 'Implementing Sustainable Digital Preservation:' http://www.asis.org/Conferences/webinars/Webinar-PASIG-10-22-2013-register.html	Public	Jisc
28-30 th October 2013	Rome, Italy	EUDAT 2nd Conference, Parallel Track III - Policy & Sustainability Issues: http://www.eudat.eu/system/files/ASHLEY%20EUDAT%2030OCT2013.pdf	Stakeholder Groups	DCC
4 th November 2013	Cranfield, UK	ENSURE Meeting (See Appendix B)	Project Team/ Ensure	Jisc, DPC
7-8 th October 2013	The Hague, Netherlands	Koninklijke Bibliotheek meeting (KB—National Library of the Netherlands) (See Appendix A)	Project Team/ KB	Jisc, DNA
7-8 th November 2013	Geneva, Switzerland	CERN Meeting (See Appendix C)	Project Team/ CERN	DNB
29 th November 2013	Bad Marienberg, Germany	MiLOS Project Meeting - Presentation of the 4C project and discussion about the MiLoS project and further engagement/cooperation opportunities	Project Team/ MiLOS	DNB
5 th —6 th December 2013	Netherlands	UNESCO, IFLA and ICA Meeting	Project Team/ UNESCO, IFLA and ICA	Jisc

Date	Location/ Country	Title	Audience	Partner
2 Stakeholder focus groups and workshops				
17 th May— 21 st June 2013	Europe	4C Initial Consultation	Stakeholder Groups	All
6 th September 2013	Lisbon, Portugal	4C Workshop and Focus Group, iPRES 2013 Conference	Stakeholder Groups	ALL
12 th December 2013	London	4C Focus Group #2 (Industry)	Industry stakeholder group	DPC
3 Advisory Board Meetings				
11 th June 2013	London	4C Advisory Board Meeting (1) –(See Appendix F)	Advisory Board	All
4 Project Website and Blog posts				
13 th March 2013	International	Blog post - Lets Collaborate!	Public	All
18 th March 2013	International	Blog post - There's room for everyone @4C	Public	All
5 th April 2013	International	Blog post - "The Age of Exploration and the Curation Costs Exchange" by Alex Thirifays	Public	DNA
14 th April 2013	International	Blog post - 'Digital curation cost models for everybody' by Sabine Schrimpf	Public	DNB
29 th April 2013	International	Blog post - 'Cache in the Attic' by William Kilbride	Public	DPC
1 st May 2013	International	Blog post - Be part of the action—Collaborate with 4C and help to Clarify the Costs of Curation	Public	KEEPS
7 th May 2013	International	Blog post - 'A very pragmatic European enterprise - reflections on cross border project involvement' by Paul Stokes	Public	Jisc
2 nd June 2013	International	Guest Blog Post: Digital Lifecycles and the Costs of Curation by Paul Wheatley	Public	Guest
9 th June 2013	International	Blog post - 'Collaborating our way to success' by Kathrine Hougaard	Public	DNA
14 th June 2013	International	Blog post - 'Call for Curation Cost Models' by Ulla Bøgvad Kejser	Public	KBDK
20 th June 2013	International	Blog post - 'Communication is key...' by Sarah Norris	Public	DPC
10 th July 2013	International	Blog post - "Nothing is Static" by Katarina Haage	Public	DNB
18 th July 2013	International	Blog post - 'How do I get to where I want to be (starting from Lisbon and going via Frankfurt)?' by Paul Stokes	Public	Jisc
31 st July 2013	International	4C Project Website	Public	DPC
19 th August 2013	International	Blog post - '4C's Cost Model Evaluation' by Joy Davidson	Public	DCC
2 nd August 2013	International	Community Resources - D2.1—Baseline Study of Stakeholder & Stakeholder Initiatives	Public	KEEPS
19 th August 2013	International	4C Focus Group Game	Stakeholder Groups	DPC
19-Aug-13	International	Blog post - '4C's Cost Model Evaluation' by Joy Davidson	Public	DCC

Date	Location/ Country	Title	Audience	Partner
9 th September 2013	International	Blog post - 'How to cut costs and keep the quality of service?' by Raivo Ruusalepp	Public	NLE
19 th September 2013	International	Blog post - 'The Case of the Curious Machine' by Sarah Norris	Public	DPC
11 th September 2013	International	Blog post - 'What the 4C Project Learnt in Lisbon' by Neil Grindley	Public	Jisc
5 th September 2013	International	Community Resources - T4.2—Draft Economic Sustainability Reference Model (Summary)	Public	Jisc
6 th September 2013	International	Community Resources - MS9—Draft Economic Sustainability Reference Model	Public	Jisc
6 th September 2013	International	Community Resources—IPRES Resources: http://www.4cproject.eu/community-resources/ipres-workshop	Public	ALL
21 st September 2013	International	Community Resources—Related Projects: http://www.4cproject.eu/community-resources/related-projects	Public	KEEPS
28 th October 2013	International	Guest Blog post - 'The Carrot and the Stick' by Matthew Addis	Public	Advisory Board
21 st October 2013	International	Blog post - 'The Future of Curation Costs' by Heiko Tjalsma	Public	DANS
7 th October 2013	International	Community Resources - D4.1—A prioritised assessment of the indirect economic determinants of digital curation	Public	NLE
8 th November 2013	International	Blog post - 'No such thing as free digital preservation' by Jan Dalsten Sørensen	Public	DNA
4 th December 2013	International	Blog post - Friends and family (or "Is there anybody out there")	Public	Jisc
17 th December 2013	International	Blog post - 'Please help us draw a map!' by Alex Thirifays	Public	DNA
5 Social Media				
1 February – 31-Dec-13	International	Posts made using the @4c_Project may be found at: https://twitter.com/4c_project	Public	DPC
6 Conferences and Events				
2 nd April 2013	Washington DC	'The role of risk analysis to support cost models for digital preservation' at the IS&T Archiving Conference 2013	Stakeholder Groups	INESC-ID
6 th May 2013	Florence, Italy	'Economics of Digital Curation Training and Education' at the DigCurV Conference	Stakeholder Groups	Jisc

Date	Location/ Country	Title	Audience	Partner
13 th June 2013	Web	'Sustainability & Cost Models for Digital Preservation' at the APARSEN Training Webinar	Stakeholder Groups	Jisc
17 th —23 rd August 2013	Singapore	Flyers distributed at IFLA	Stakeholder Groups	NLE
2 nd —6 th September 2013	Lisbon, Portugal	Poster Session, Minute Madness, Workshop and Focus Group	Stakeholder Groups	DNB
23 rd —26 th September 2013	Valletta, Malta	Flyers distributed at TPDL 2013	Stakeholder Groups	NLE
9 th —13 th October 2013	Frankfurt, Germany	Flyers distributed at Frankfurt Book Fair 2013	Stakeholder Groups	DNB
22 nd October 2013	Web	'Implementing Sustainable Digital Preservation' at the PASIG ASIS&T Webinar	Stakeholder Groups	Jisc
29 th October 2013	Rome, Italy	'Parallel Track III - Policy & Sustainability Issues' at the EUDAT 2nd Conference	Stakeholder Groups	UEDIN-DCC
7 th November 2013	Nottingham, UK	Flyers distributed at the Business Archives Conference Council 2013	Stakeholder Groups	DPC
18 th —20 th November 2013	Barcelona, Spain	Panel Session: Chaired by Neil Grindley (JISC), the Resource Alignment panel "How can we employ the resources we have available to us most economically to achieve our digital preservation objectives?" Action Session: 4C Case Studies and Quantitative Data	Stakeholder Groups	Jisc
28 th November 2013	London, UK	Panel session on 'Sustainability' at the Institute of Historical Research	Stakeholder Groups	Jisc
7 Publications				
7 th February 2013	Denmark	New EU project examines the costs of digital preservation http://digitalbevaring.dk/nyt-eu-projekt-undersoger-omkostningerne-ved-digital-bevaring/	Public	KBDK/ DNA
23 rd March 2013	Denmark	Press release on the EU project 4C http://digitalbevaring.dk/pressemeddelelse-4c/	Public	KBDK/ DNA
25 th March 2013	Austria	4C Project—the Collaboration to Clarify the Costs of Curation - a project within the Digital Preservation area http://www.sba-research.org/research/data-security-and-privacy/digital-preservation/4c-project-the-collaboration-to-clarify-the-costs-of-curation/	Public	SBA
25 th March 2013	Portugal	Project 4C—Collaboration to Clarify the Costs of Curation http://www.keep.pt/projeto-4c-collaboration-to-clarify-the-costs-of-curation	Public	KEEPS

Date	Location/ Country	Title	Audience	Partner
25 th March 2013	UK	DPC joins new EU collaboration to clarify the costs (and benefits) of curation http://www.dpconline.org/newsroom/latest-news/978-dpc-joins-major-new-eu-initiative-to-understand-the-costs-and-benefits-of-digital-curation	Public	DPC
25 th March 2013	UK	Collaboration to Clarify the Costs of Curation (4C) http://www.dcc.ac.uk/projects/4c	Public	DCC
25 th March 2013	UK	13 partners from across Europe join together to improve digital curation http://www.jisc.ac.uk/news/stories/2013/03/4C.aspx	Public	JISC
25 th March 2013	Netherlands	Project 4C: the Collaboration to Clarify the Costs of Curation http://www.dans.knaw.nl/content/categorieen/projecten/project-4c-collaboration-clarify-costs-curation	Public	KNAW DANS
27 th March 2013	UK	Major EU initiative to clarify the costs and benefits of digital curation http://www.data-archive.ac.uk/news-events/news.aspx?id=3466	Public	UK Data Archive/ Essex
1 st April 2013	Estonia	Digital preservation - what it will cost and what the benefits are? http://www.nlib.ee/en/mis-kasu-saab-digitaalsest-sailitamisest/	Public	NLE
1 st April 2013	Estonia	Major new EU initiative to understand the costs and benefits of digital curation http://www.nlib.ee/major-new-eu-initiative-to-understand-the-costs-and-benefits-of-digital-curation/	Public	NLE
18 th April 2013	Germany	PROJECT STARTED - 4C http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/Kurzmeldungen/4C.html	Public	DNB
18 th April 2013	Germany	PROJECT STARTED - 4C http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/Kurzmeldungen/4C.html	Public	DNB/ Nestor
1 st May 2013	Germany	4C - EU project started on the costs and benefits of digital preservation http://files.d-nb.de/nestor/newsletter/nestor-Newsletter_28.pdf	Nestor Readership	DNB
17 th May 2013	UK	The 4C consultation is now open! http://www.dpconline.org/newsroom/latest-news/1014-4c-consultation-now-live	Public	DPC
20 th May 2013	Portugal	The 4C consultation is now open! http://www.keep.pt/estudo-para-a-quantificacao-de-custos-associados-a-preservacao-digital-convite-a-participacao	Public	KEEPS
26 th June 2013	Denmark	State Archives and the Royal Library appreciate the cultural heritage: http://www.sa.dk/content/dk/om_statens_arkiver/nyhedsoversigt/statens_arkiver_og_det_kongelige_bibliotek_satter_pris_pa_den_digitale_kulturarv	Public	DNA/ KBDK
26 th June 2013	Europe	4C - Collaboration to Clarify the Costs of Curation in Issue 5—SCAPE Newsletter: http://us4.campaign-archive1.com/?u=20cef0f757e3840df2769745b&id=114edecf55	SCAPE Readership	DPC
28 th June 2013	Europe	European Partnerships in TIMBUS Times: http://timbusproject.net/about/publicity-material	TIMBUS Readership	DPC

Date	Location/ Country	Title	Audience	Partner
2 nd July 2013	Denmark	New EU project examines the costs of digital preservation: http://digitalbevaring.dk/nyt-eu-projekt-undersoger-omkostningerne-ved-digital-bevaring/	Public	DNA
13 th July 2013	US	4C - Collaboration to Clarify the Costs of Curation http://www.dlib.org/dlib/july13/07inbrief.html	D-Lib Readership	DPC
10 th August 2013	Germany	"Nothing is static" in "Dialog mit Bibliotheken": http://www.dnb.de/DE/Service/Publikationen/dialog201302.html	Stakeholder Groups	DNB
8 Reports				
7—8 th October 2013	The Hague, Netherlands	KB Meeting Minutes—Appendix A	Project team	DNA, Jisc
4 th November 2013	Cranfield, UK	Ensure Meeting Minutes—Appendix B	Project teams	Jisc, DPC
7—8 th November 2013	Geneva, Switzerland	CERN Meeting Minutes—Appendix C	Project teams	DNB
6 th September 2013	iPRES, Lisbon	Report from Focus Group 1—Appendix D	Public sector stakeholders	DNB
12 th December 2013	Jisc Offices, London	Report from Focus Group 2—Appendix E	Industry stakeholders	DPC
11 th June 2013	London	Highlights of Advisory Board Meeting (1)—Appendix F	Project team and stakeholders	All

Table 17—Summary of activities

11 A Tactical Look Ahead (Months 12-24)

Within the next 12 months, the 4C Project will deliver:

- Two more focus groups (months 14 and 17)
- One workshop (month 17)
- A 4C Conference (month 21)
- The Curation Costs Exchange (CCEX) (month 24)
- The register of Stakeholders & Stakeholder Initiatives (month 24)
- A Final Stakeholder Report (month 24)
- A Final report on Outreach Events (month 24)

These and other planned activities for the next 12 months are outlined below.

In addition to general 4C communications activities, a Communications Plan relating particularly to optimising the uptake of the Curation Costs Exchange (CCEX) has also been developed, and forms an Annex to the Project Communications Plan (included in this document as Appendix G). The statement of principle behind the CCEX Communications Plan is to facilitate access to collaborative tools that anyone can use, in the knowledge that information is shared without commercial exploitation, with a view to sustaining the currency and relevancy of the CCEX output.

Derived directly from the 4C project objectives, the plan and the activities outlined within the plan, aim to meet this statement of principle, specifically to:

- Communicate the benefits of the CCEX to 4C stakeholder groups in order to create demand - Measurable in the uptake of the CCEX, i.e. in the number of people or institutions that use the CCEX and the amount of data cost sets shared through the CCEX
- Achieve an initial sharing of 12 -25 cost data sets within the remainder of the project lifetime.
- Sustain the life of the CCEX through ongoing sharing of 2-5 cost data sets per month, post project completion.

The fulfilment of these objectives and the application of an ‘open and social’ communications model will ultimately facilitate the achievement of the main 4C project objectives by engaging users in sustainable dialogue throughout the lifetime of the project.

The full CCEX Communications Plan (Annex to the 4C Project Communications Plan) may be found at the end of this document as Appendix G: Annex to D2.5: 4C Project Communications Plan—Draft ‘CCEX Communications Plan’.

11.1 Communication and information exchange with EC-funded and other projects and organisations

The 4C Project plans to engage with the following projects as a means to seek stakeholder input on the deliverables to be developed throughout the year:

- Digital Preservation Coalition members
- Nestor
- APARSEN
- ENUMERATE
- The British Postal Museum & Archive

This is not a static list and additional projects are being constantly sought out.

11.2 Stakeholder focus groups and workshops

The following opportunities for focus groups have been identified and proposals submitted. Further opportunities for holding additional focus groups and workshops are listed under Section 11.6 Conferences and Events, and will continue to be developed throughout the year:

Conference/ Event	Location	Date	4C Partner & Affiliation	Activity
DPHEP Topical Workshop on "Full Costs of Curation"	Geneva, Switzerland	13 th —14 th January 2014	Katarina Haage, DNB	Costs and cost model workshop at CERN with "Big data science" as target group
EuroCRIS	Rome, Italy	13 th —15 th May 2014	Joy Davidson, HATII-DCC	Paper/ Focus Group
IS&T Archiving 2014	Berlin, Germany	13 th —16 th May 2014	Katarina Haage, DNB	Paper/ Focus Group

Table 18—Future focus groups and workshops

11.3 Advisory Board Meetings

The next Advisory Board Meeting is due to be held on 22nd—23rd January 2014, with a highlights report to follow which will be published on the 4C Website.

One further meeting is planned for July 2014.

11.4 Project Website

The 'Community Resources' section of the website will continue to be populated with deliverables and other resources as these are completed by the project team through the year.

The CCEx will also be accessible through the 4C Project website, along with supporting resources and information for stakeholders to use.

Blog posts will continue to be posted on a weekly basis, where possible. A proposed schedule for the year is as follows:

Date (w/c)	Partner	Author	WP	Topic
Jan 6		~		
13		~		
20	Jisc	NG/PS	1	Feedback from CERN? /Draft Sustainability & Benefits Plan?
27	DANS	HT	1	Advisory Board/F2F -outcomes (as the host)/Guest blog from Advisory Board?
Feb 3	INESC-ID/ NLE	DP/RR	4	IED
10	DNA	ABN	3	CCEx - call for data
17	DNB	KH	2	Plan for outreach/a look ahead for the year
24	HATII-DCC	JD	2	IDCC
Mar 3	UESSEX	HLH	3	Cost concept model

Date (w/c)	Partner	Author	WP	Topic
10	SBA	Sstrodl	4	costs to business models
17	DPC	SN	2	CCEX - the pilot
24	Jisc	NG	1	Review Meeting
31	DNB	KH	2	RDA Plenary
April 7	DANS	HT	3	CCEX
14	NLE	RR	4	
21	Jisc	PS	1	
28	KBDK	UBK	3	cost concept/gateway requirement spec
May 5	INESC-ID	DP	4	benefit, impact, value
12	DNA	AT	3	CCEX
19	DNB	SS	2	Archiving 2014
26	HATII-DCC	JD	2	DCC Webinar/Eurocris outcomes
June 2	DNB	KH	2	Library Day Bremen
9	SBA	Sstrodl	4	
16	DPC	SN	2	CCEX
23	Jisc	PS	5	Draft Roadmap
30	DNA	ABN	3	
July 7	DANS	HT	3	CCEX
14	Jisc	PS	5	Public consultation on Roadmap (MS22)
21	INESC-ID	DP	3	Cost Concept Model & Gateway Requirement Spec
18	DNA	KHJ	4	Report on Risk, Benefit, Impact and Value (D4.4)
Aug 4	Guest		1	Advisory Board 3?
11	DPC	SN	2	CCEX (at the 4C Conference?)
18	DNB	SS	2	IFLA Lyon
25	Host	TBC	1	Project Meeting 4
Sep 1	UESSEX	HLH	4	ESRM
8	KEEPS	LF	3	CCEX
15	DPC	SN	2	Prep for 4C Conference
22	SBA	DW	4	
29	Jisc	PS	5	
Oct 6	DANS	HT	3	CCEX Framework (D3.3)
13	NLE	RR	4	
20	Jisc	PS	5 and 2	4C Conference/ Roadmap workshop
27	KBDK	UBK	2	
Nov 3	INESC-ID	DP	3	
10	DNA	AT	3	CCEX
17	DNB	KH	2	
24	HATII-DCC	JD	2	
Dec 1	KBDK	UBK	2	Final Stakeholder Report
8	Jisc	NG	1	Final Sustainability & Ben Plan
15	DPC	SN	2	CCEX/Report on Outreach Events (DNB)
22		~		Christmas
29		~		New Year

Table 19—Blog schedule

11.5 Social Media

Tweets will continue to be made on a daily basis, where possible, as a means to publicise web content (deliverables, blogs and so on) as well as activities (events, engagement).

11.6 Conferences and Events

The end of project conference will be a good opportunity to demonstrate resources produced, including the CCEx, as well as finalising input into the Roadmap. The conference will be used to invite stakeholders to reflect upon the project's implications for the future and a workshop will be designed as part of the programme.

Other potential events and conferences are as follows:

Conference/ Event	Location	Date	4C Partner & Affiliation	Activity
DPHEP Topical Workshop on "Full Costs of Curation"	Geneva, Switzerland	13 th —14 th January 2014	Katarina Haage, DNB	Costs and cost model workshop at CERN with "Big data science"
9th International Digital Curation Conference	San Francisco, USA	24 th —27 th February 2014	Joy Davidson, HATII-DCC	Paper
Nordic Conference National Archives, "NorDig"	Copenhagen, Denmark	May 2014	Alex Thirifays, DNA	Paper/ Workshop
EuroCRIS	Rome, Italy	13 th —15 th May 2014	Joy Davidson, HATII-DCC	Paper/ Focus Group
IS&T Archiving 2014	Berlin, Germany	13 th —16 th May 2014	Katarina Haage, DNB	Paper/ Focus Group
International Internet Preservation Consortium (IIPC) Conference	Paris, France	19 th May 2014		
CENL	Moscow, Russia	May 2014		
Librarian Day 2014	Bremen, Germany	3 rd —6 th June 2014	Katarina Haage, DNB	Working group
IASSIST Conference	Toronto, Canada	3 rd —6 th June 2014		
LIBER Conference	Riga, Latvia	2 nd —5 th July 2014		
JCDL/ TPDL	London, UK	8 th —12 th September 2014		
RDA Plenary 4	Amsterdam, Netherlands	22 nd —24 th September 2014		
iPRES 2014	Melbourne, Australia	6 th —10 th October 2014		
Alliance for Permanent Access (APA) Conference	Brussels, Belgium	21 st —24 th October 2014		
4C Conference	Brussels, Belgium	Pre/post 21 st —24 th October 2014		
Data Lifecycle Management Conference	Lisbon, Portugal	November 2014		

Table 20—Future conferences and events

11.7 Publications

Opportunities for articles will be sought throughout the year, as a means to publicise activities and engage particular stakeholder groups.

11.8 Reports

Reports highlighting the salient points from collaborative and Advisory meetings will be produced throughout the year for the benefit of the project team, and published on the 4C website as a resource for stakeholders.

A Minutes of cost data meeting on 07-08 October 2013 at KB in The Hague

A.1 Attendees

KB:

Tanja de Boer—Head Collection Care Department

Hildelies Balk—Head Research Department

Trudie Stoutjesdijk—Senior Advisor Digital Collection Care

Barbara Sierman—Digital Preservation Manager

Jeffrey van der Hoeven—Business Analyst

Marcel Ras—Program Manager

4C:

- Neil Grindley (JISC)
- Alex Thirifays (Danish National Archives)

A.2 Agenda

Monday 7 October			
13.00 – 14:30	<ul style="list-style-type: none"> • Introductions • What the 4C project hopes to achieve • What the KB is currently doing • 4C project (stakeholder consultation, economic modelling, assessment work, costs gathering) 	Tanja, Hildelies, Barbara, Trudie, Marcel	B4 (lunch)
14:30 – 16:00	KB – Storage and digitization	Trudie, Marcel,	B4
16:00 - 17:00	KB – business plan international e-Depot	Marcel, Barbara	B4
Tuesday 8 October			
9:00 – 10:00	KB storage costs model	Jeffrey, Rogier, Marcel	Kamer Marcel
10:00 – 11:00	4C to re-present KB data to align with 4C modelling approaches (such as they are!) <ol style="list-style-type: none"> Economic Sustainability Reference Model Indirect Economic Determinants 4C Cost concept Model 4C Curation Costs Exchange (CCEX) 	Tanja, Hildelies, Barbara, Trudie, Jeffrey, Marcel	Kamer Tanja of Hildelies
11:00 – 13:00	<ul style="list-style-type: none"> • Discussion to refine approaches and objectives of 4C modelling activities • Discussion around next steps and additional engagement opportunities 	Tanja, Hildelies, Barbara, Marcel, Jeffrey, Trudie	Kamer Tanja of Hildelies (lunch)

Table 21—KB agenda

A.3 Session 1 day 1—KB storage and digitization

Participants: 4C: Alex Thirifays; KB: Trudie Stoutjesdijk, Marcel Ras.

Topic: Costs of digitization and storage of the digitized material.

Context

Digitization

KB performs a value assessment of the physical material to be digitized by attributing primary and secondary values (for example information value, vulnerability, frequency of use). This value assessment differentiates the material in 5 levels, where level 1 material won't be digitized, level 2-3 will exist in an access copy only, while 4-5 represent vulnerable material that only supports one digitization process and the output of it is 2 copies: a master copy and an access copy.

To calculate the costs of digitization, KB uses TCO (Total Cost of Ownership) that permits to look at return on investment and total profitability over time.

Archival Storage

Once digitized, the material enters into a storage hierarchy with 4 levels (steel, bronze, silver and gold). This division is both to satisfy user and financial requirements. The copies are stored in two geographic locations (70 km apart), the KB has a disaster plan and the material is stored on different types of media according to the 4 levels (Silver Cube, tape, server, but no optical disks). The management of copies has not been consistent, so sometimes 8 copies exist of the same master file, even though their strategy only dictates 2. As volume is a cost driver in Archival storage, this is of essence. Note also that some of the costs of Access are included in KB Archival Storage because of their 4 level storage hierarchy that privileges their users.

For the costs of digitization and storage, they have developed their own cost models and tried to get financial information from their financial department. This exercise has proven to be difficult, because the financial department does not break down the costs of digital curation in any way (neither by activity nor by department).

They develop in-house solutions but they also buy finished products. This is a pragmatic approach that privileges the short-term solving of problems, but that also can result in lack of interoperability and difficulties getting a clear overview of the systems (for example, they have bought the Tivoli Storage Management System and have developed an in-house system in favour of an IBM system (this was because they encountered scalability issues (volume; formats) with the IBM system).

On E-journals and Pre-Ingest

KB receives e-journals from publishers. They have specified a reception format (SIP) that conforms to standards (METS, NLM, PDF), but the publishers are not specialists and the KB spends a lot of time managing faulty SIPs and transforms them into AIPs (40 days pr. Batch – before it was 80 and their objective is to bring it down to 20 → this number covers the whole process).

On education

Marcel Ras claims that there is an educational problem: There aren't enough 'in-betweens', that is people with combined it and librarian knowledge.

On awareness

This is linked to an awareness issue:

- Schools and universities don't offer relevant courses;
- the strategic level is slow to transform the physical library to a digital one and it is not possible to re-allocate resources from the former to the latter;
- and the political level hasn't fully understood the digital challenges yet, which makes it impossible for the strategic level to accomplish the needed changes (the National Archives, for example, are only just preparing the reception of digital material from governmental agencies, because the 'digitization' of this has been slow; also in the past, the agencies delivered the digital material they had to data centres⁵ owned and controlled by the ministries).

On cost models, visibility and efficiency

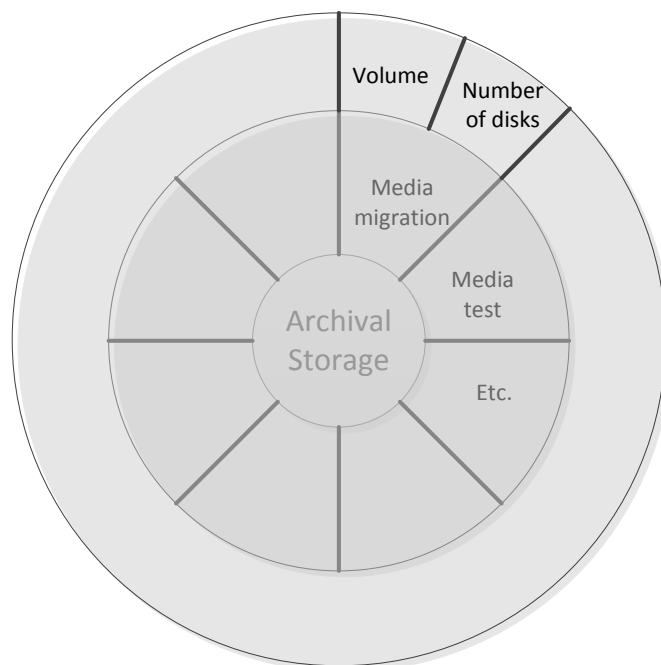
KB needs cost models to make their costs visible (their building costs 50% of the total KB budget = physical archiving is extremely expensive because of floor space).

KB needs business/analysis tools to measure performance, analyse and optimize the management of their collection, but also to receive user feedback and to base decisions on these analyses and statistics. This is something the OAIS mentions.

KB needs a tool to predict the increase in volume of their collection (they have now built one).

On cost concept models

4C presented a concept cost model (the Moon, the Flower with petals, the Onion, The Wheel (this could roll representing the 'over time' aspect...) or the Cog wheel model):



The inner circle represents a functional entity from OAIS, the middle one a number of activities attached to this function and the outer circle holds the cost drivers of each activity. This figure only shows one of

⁵ There are 60 in the Netherlands, but a fusion is taking place bringing the number to 4.

the functional identities—ideally there should be 6 and visually—if it’s a cog wheel—the cost drivers would be the teeth of the cog wheels that touch each other, making each other spin).

4C explained that it could also be feasible to display other parameters necessary for costing, for example financial adjustments (depreciation; inflation) and the type of costs (indirect, direct, maintenance, one-off, etc.).

KB was positive about this approach, but expressed a need for a cost model they could use in their context, not necessarily a *concept* model.

4C explained that the advantage of a generic model would be that it would increase awareness on the topic across borders and organizations and help highlighting challenges of this nature.

A.4 Session 2 day 2—KB cost models

KB to present their cost models for storage and business plan for international e-Depot.

KB set out their ‘primary processes’

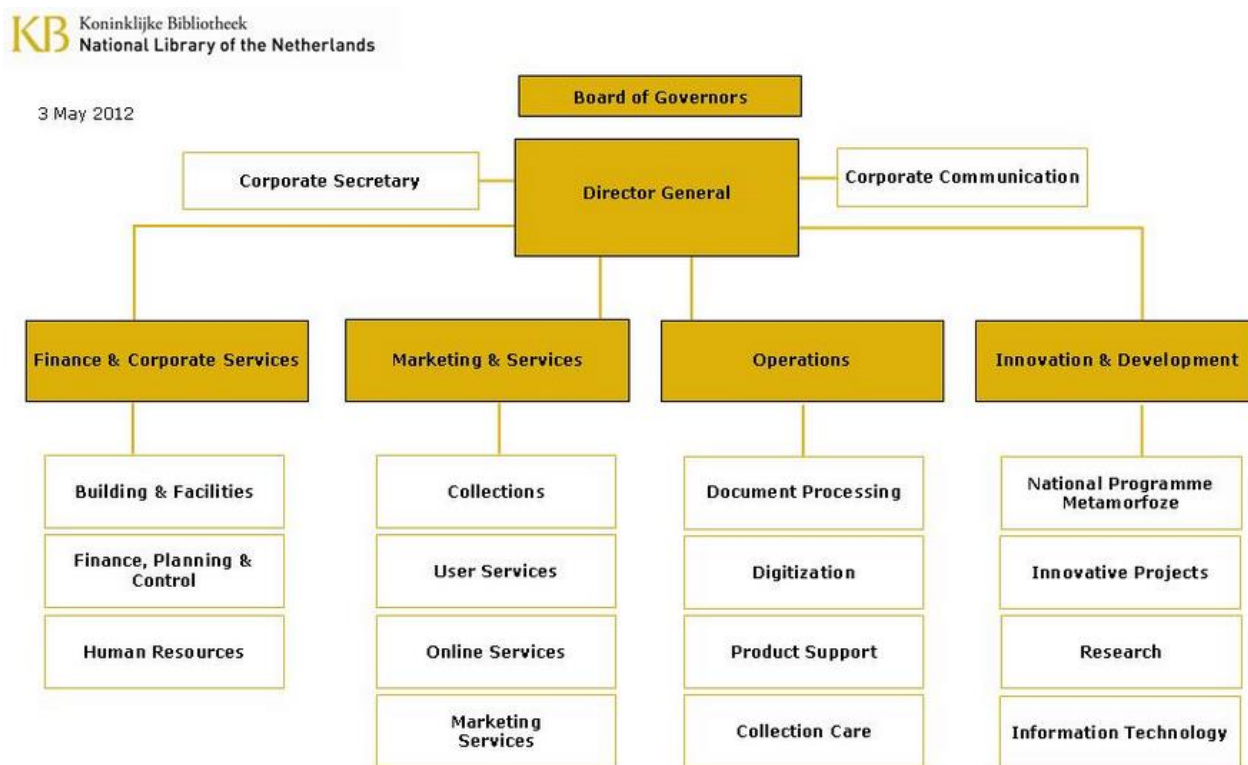
- Selection
- Processing (includes digitisation and adding metadata)
- Storage / Preservation
- Access

Marketing is a function that is covered by the Operations group within KB.

Note: Strategic plan (2010-2013) available at:

http://www.kb.nl/sites/default/files/docs/KBstratPI_print.pdf

KB Organogram (from: http://www.kb.nl/sites/default/files/organogram-en_0.jpg)



KB Organisation

The majority of KB is focused on analogue materials but Digital Preservation takes place as part of 'Operations'. But the Collections department (marketing & services) need to have knowledge of preservation when they deal with publishers. And digital preservation research takes place in the context of 'Innovation & Development'. And the budgets are all assigned and understood by 'Finance & Corporate Services' – so DP is linked to all parts of the organisation. KB is interested in providing digital services to external organisations (e.g. the Metamorfoze project offers external organisations digitisation services).

One of the big drivers for looking into the cost of preservation is to be able to offer realistic and cost effective preservation services to others. This reinforces an important foundational concept of the 4C project.

KB storage cost model

Context: DANS made a cost model—its raison d'être was the economic crisis and the inability for archaeological companies to pay for the datasets stored by DANS.

Jeffrey van der Hoeven presents the KB storage cost model which includes a presentation of current storage needs (newspapers 260TB; Books 29TB; Pictures 121TB ... , or: newspapers 60%, pictures 27%, books 7%, etc), predicted storage needs and current and predicted costs.

KB aren't interested in the insurance cost as the collections aren't insured.

They need to know predictive figures of potential storage requirements because either a) they WILL need to make the case for procurement, or b) because they might NOT wish to make the case. If it looks impossible or prohibitively expensive, they may want to decline to select a collection for long-term preservation. **Understanding costs can support strategic planning.**

They calculate the costs for storage per unit (a unit being 1 TB) per year. Using TCO and quantifiable components taken from Hitachi model ("Principles for Reducing TCO"). The Hitachi model comprises 30 components – KB has used 15, for example: hardware, software, depreciation, maintenance, support, power and cooling, floor space, monitoring (does this include error handling? -AT), waste and duplication, off-site locations, networks.

A cost driver is volume, but not different formats, for example. Another driver is LTP (Long-Term Preservation) requirements:

One conclusion is that the price hugely depends on supply and demand (price on a storage system went from 250.000€ to 40.000€ in 2 years).

The KB produces new figures each year.

Storage Tiers

Name	Current Size of Assets Stored (TB)	Nature of storage	Objective of Tier	Technology
Gold	88	Very fast, very expensive	Indexing, databases	Oracle cluster, Lucene
Silver	459	Fast, expensive	Web hosting, data processing	RAID arrays
Steel	64	Slow, moderately priced	Long term preservation	RAID-like systems
Bronze	721*	Very slow >45 secs to access, very cheap	Backup, archiving	Tape

Table 22—KB storage tiers

* Not including offsite backup

KB wish to move away from having backup and archiving in the same tier.

If you can quantify the costs between the tiers, you can make the right choices about exactly the level of service you can afford to offer to your users. The sweet spot is to meet (or just surpass) the expectations of users with the cheapest acceptable tier of storage. **Understanding costs can support tactical decision-making.**

Storage Costs

KB looked back over 41 invoices over a period of time to try and establish the costs of storage. This involved digital assets amounting to 1.3 TB and involved 2,747 parts.

Jeffrey's presentation <TCO storage KB 2013 EN.pdf> contains all the detailed figures, some of which are embargoed until KB finance department approves the numbers.

What was remarkable...

- Strong price drop of gold & silver storage: Old: HP EVA 8400 chassis cost EUR 276.228 ... New: HP EVA P6500 chassis cost EUR 38.993;
- Strong price drop of disk drives (HDD): Old (1 TB): EUR 1.554 ... New (1 TB): EUR 532;
- Archival storage solution is not so cheap as people often think: due to extra LTP requirements;
- Tape extremely cheap (by long depreciation period and large volume);
- Yearly cost of power consumption of KB: EUR 575.000,-

Not included in the presentation is the cost of the dedicated fibre link to the backup location in Hilversum, 70km away = 50k euros per year

One of the big questions is 'how many bits do we want to preserve?' This is bigger picture strategic decisions that are yet to be resolved. The issue of duplication is important. How many copies of a master file are needed?

One of the most important factors in driving down cost is procurement and the deals that can be struck. Scale and timing are important. Also the accounting principles. If you decide that hardware depreciates over 10 years rather than 5, your costs really go down. A dramatic decrease in costs per TB can be shown between 2008 – 2013. But this has to be understood in different ways.

Idea for CCEx: Could include information on cost model methodologies.

Cloud cost comparison

KB did a TCO archival storage comparison with cloud storage providers. KB's figures were in the middle of a number of providers. (see presentation)

It was based on archival storage but there is a much larger issue to do with (the cost of) access where one would expect the KB to really pull away from the cloud providers. KB infrastructure is designed to provide access as an implicit part of the cost of hosting the data. The cloud providers will have different access arrangements and separate pricing. **Understanding costs can provide evidence of cost-effectiveness and value**

Motivations for sharing costs data

KB wish to be as open with their data as possible as it aligns with their public role and their publicly funded remit. If the figures are accurate and plausible (and based on efficiently procured, effectively configured and well managed systems) then being open and candid about them should increase the credibility of the organisation.

Clarifying and publishing the cost of digital curation can be used to enhance an organisation's credibility. But this must be done along with the context of how the costs were calculated.

Seeing the costs are interesting but it is important to look behind the figures.

E-Depot

E-journal archiving on an international basis: By 2018, KB will receive e-journals from 80 publishers (today the number is 12).

Marcel Ras spent 1 year making a cost model, and based some of his calculations on Jeffrey van der Hoeven's and Trudie Stoutjesdijk's work.

He identified 5 cost areas: Staffing, storage, infrastructure, research and 'other'.

By predicting how many publishers will be added each year, predictive storage and staffing is possible. To raise the profile of the E-Depot, KB are going to have to make it clear that they wish to deliver services to publishers and libraries.

Marcel used a pragmatic framework to try and arrive at the costs. Broke down the costs into:

- Staff
- Storage
- Infrastructure
- Research
- Other

Most costs are in staffing. Used an activity based model, e.g. someone to negotiate with publishers; someone to analyse material; someone to ingest and monitor data; someone to monitor if a trigger event occurs.

Used Jeffrey's model for storage. Did some predictive staffing based on the current ways of working and likely future publisher engagement. (Better to work with large publishers rather than small ones because the overhead of working out how to ingest one journal from one publisher is as onerous as working out how to ingest 50 journals from a larger publisher).

18m articles stored in E-Depot taking up 17TB. This cost 120k euros in 2013 and is likely to cost 127k euros in 2014.

The efficiency of ingesting data streams has improved over the years, now publishers are mostly using the NLM DTD. They aren't using it in a standard way but they are at least converging on that schema.

If charges were passed onto publishers (the activity is currently being supported by the Dutch tax payer) then deviation from a standard schema could incur a financial cost. There would be a charge for long-term archival storage and a startup fee for getting the processes set up.

Indirect Economic Determinants (IEDs)

Tanja stated that consideration of the Indirect Economic Determinants might have made a useful management exercise when they recently sat down to update the KB strategic plan.

We explained that going through the IED's and having different members of staff from the same organisation trying to prioritise them may shed interesting light on what level of shared understanding there was across the organisation about motivations and priorities. The exercise took about 10 minutes.

IED	Marcel	Trude	Tanja	Barbara*	Alex**	Notes
<i>Authenticity</i>	High	High	Low		High	
<i>Benefit</i>	Medium	Medium	Low		Low	
<i>Efficiency</i>	Medium	Low	Medium		Medium	
<i>Impact</i>	High	Medium	High		High	
<i>Innovation</i>	Medium	Medium	High		Medium	Tanja very focused on this
<i>Interoperability</i>	Low	Medium	Low		Low	
<i>Quality</i>	High	High	Medium		Low	
<i>Reputation</i>	High	High	High	High	High	Barbara very focused on this
<i>Risk</i>	High	Medium	Medium		High	
<i>Sensitivity</i>	Medium	Medium	Low		High	Found this term difficult
<i>Skills</i>	High	High	High	High	Medium	
<i>Sustainability</i>	High	High	High	High	High	
<i>Transparency</i>	High	Medium	Low		Low	
<i>Trustworthiness</i>	High	High	High	High	High	
<i>Value</i>	Medium	Medium	High	High	Low	

Table 23—The results of the IED exercise at the KB

* Barbara arrived late and only had 2 minutes to look at this

** Alex filled it in (again) to demonstrate potential variation across organisations

Interestingly with the KB respondents the only IED that provoked answers that were non-adjacent (i.e. that included responses that were ‘high’ and ‘low’—rather than just ‘high’ and ‘medium’ or ‘medium’ and ‘low’) was **Transparency**. This seems to demonstrate a good level of consensus and understanding between the staff about the KB’s motivations and priorities for investment.

Understanding economic drivers can help to strategically align an organisation

Economic Sustainability Reference Model

Introduced the ESRM and its component parts and KB responded that it looked interesting. They requested a copy of the questionnaire to be sent after the meeting. They would try to take a look at it towards the middle or end of October and give feedback.

CCEx

The Curation Costs Exchange was discussed.

KB’s input:

- Contextual information needed to interpret cost data.
- It should be a platform for information exchange
- It should encourage people to make separate budgets for digital curation and to introduce these for the financial departments that should be considered as a new stakeholder group by the 4C-project.
- It should make it clear that digital curation affects the rest of the organization and its costs.
- It should provide information on value (benefits)
- Different types of information should be accessible
- The idea of a more exclusive space (club) where people might share more sensitive information could be interesting (Barbara)

Action Summary

Item	Description	Who	Owner	Due by
1	Show that KB is an affiliate partner to the 4C Project and show this prominently on the website	Sarah	Neil	1/11/2013
2	Do a blog post about the meeting	Neil/Alex	Neil	1/11/2013
3	Send KB updated version of ESRM	Neil/Paul	Neil	1/11/2013
4	Place material into costs dropbox folder and alert project partners	Neil	Neil	1/11/2013
5	Talk to WP2 about potentially adding Finance Officers as a 4C stakeholder group	Neil	Neil	1/11/2013

Table 24—KB meeting action summary

B Minutes of Collaboration Meeting with ENSURE Project, Cranfield University, 4th November

B.1 Attendees

ENSURE:

- Essam Shehab - Supervisors
- Paul Baguley - Supervisors
- Isaac Sanya - Project Manager
- Zaira Fernandez Ortiz - Research Associate
- Mohamed Badawy - PhD Student

4C:

- Neil Grindley – Project Co-ordinator (Jisc)
- Paul Stokes – Project Manager (Jisc)
- Sarah Norris – WP2 (DPC)

B.2 Agenda

Monday 4 November		
13:30 - 14:00	Arrival at Cranfield University	
14:00 - 14:30	In Room Buffet Work Lunch	
14:30 – 15:00	Coffee and Open Collaboration Possibilities Discussion	All
15:00 - 15:30	Introducing ENSURE Project	Mohamed Badawy
15:30 - 15:45	Tea/Coffee Break	
15:45 - 16:15	Introduction to 4C Project	Neil Grindley
16:15 - 16:30	Introduction to Digital Preservation Coalition (DPC)	Sarah Norris
16:30 - 17:00	Collaboration Possibilities Discussion continue.	All

Table 25—ENSURE Agenda

B.3 Introducing ENSURE Project

The ENSURE project (Enabling kNowledge Sustainability Usability and Recovery for Economic value) involves 13 partners, and has €11.73 million of funding from EU-FP7.



ENSURE partners

The aim of the project has been to:

“Develop a framework to predict the whole life cycle cost of long-term digital preservation in the cloud. The developed framework and its output cost model will serve three business sectors; healthcare, financial and clinical trials”

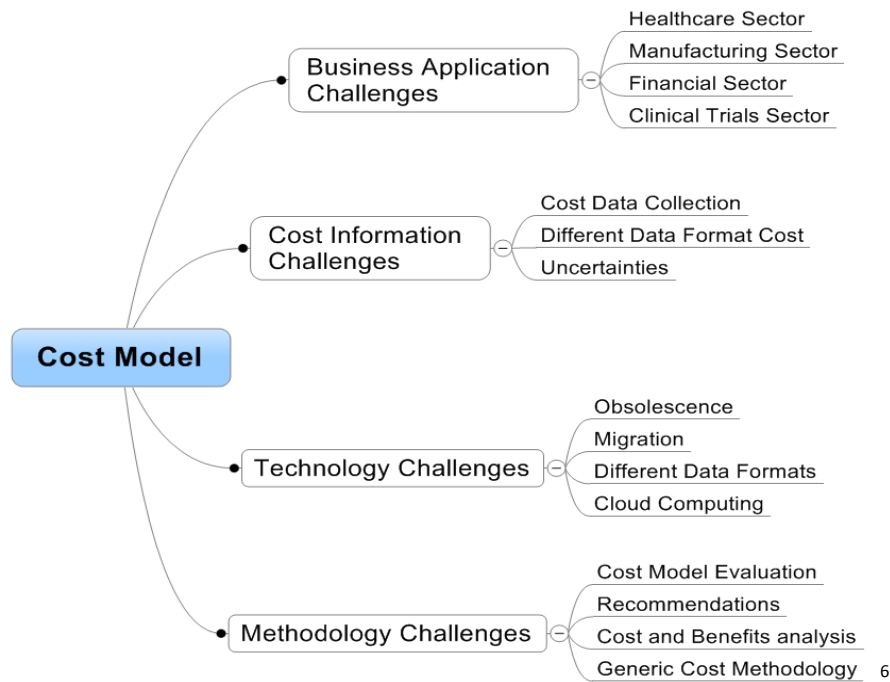
Partners handle data for Healthcare, Clinical Trials and Financial business sectors.

The project has produced three Optimisation Engines provide critical cost, economic performance and quality information for decision makers utilising Cloud Computing technology (public and private) for storage as well as computing for Long term Digital preservation (LTDP).

The benefits to be derived from the outcomes of the project are:

- Reduces the impact of hardware obsolescence
- Reduces the infrastructure start-up cost
- No expense for new features and facilities that are added regularly
- Eco-Friendly
- Pay-As-You-Go
- Computation power are hired
- Easy access over the internet

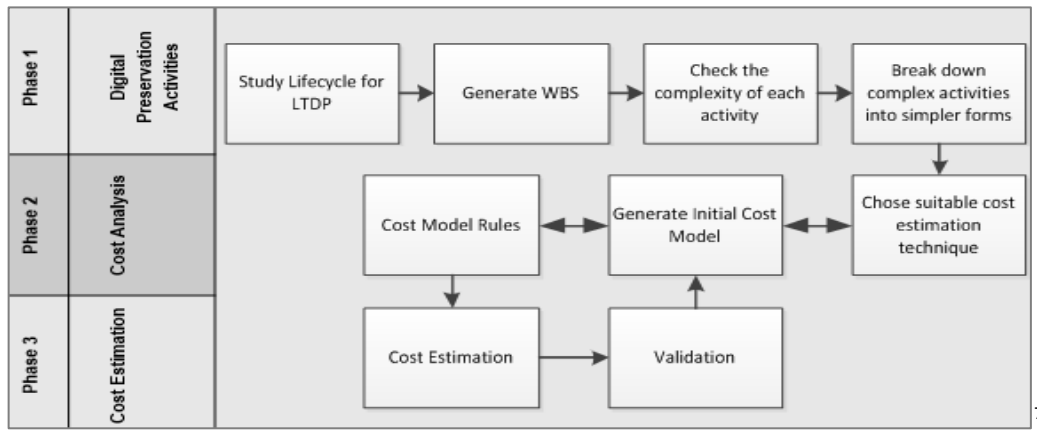
Challenges of LDTP identified by the project are as shown below:



Cost model challenges

The Project’s ‘Cost Engine’ considers uncertainties in Long term Digital preservation (LTDP) and obsolescence issues, the framework for which is shown below:

⁶ Pengfei Xue, Essam Shehab, Paul Baguley, Mohamed Badawy (Cranfield University), Cost modeling 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.



Long term Digital preservation framework

The Cost Engine considers the LTDP Lifecycle in the following identified stages:

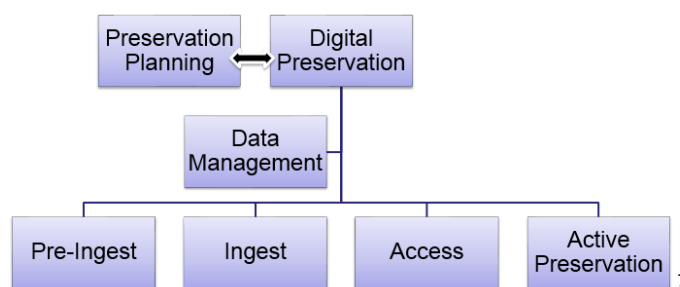


Long term Digital preservation lifecycle

As well as differences between the three identified sectors (healthcare, finance and clinical trials):

- Legal requirements
- Preservation duration (retention period)
- File type
- File format
- Access rate
- Copy rights

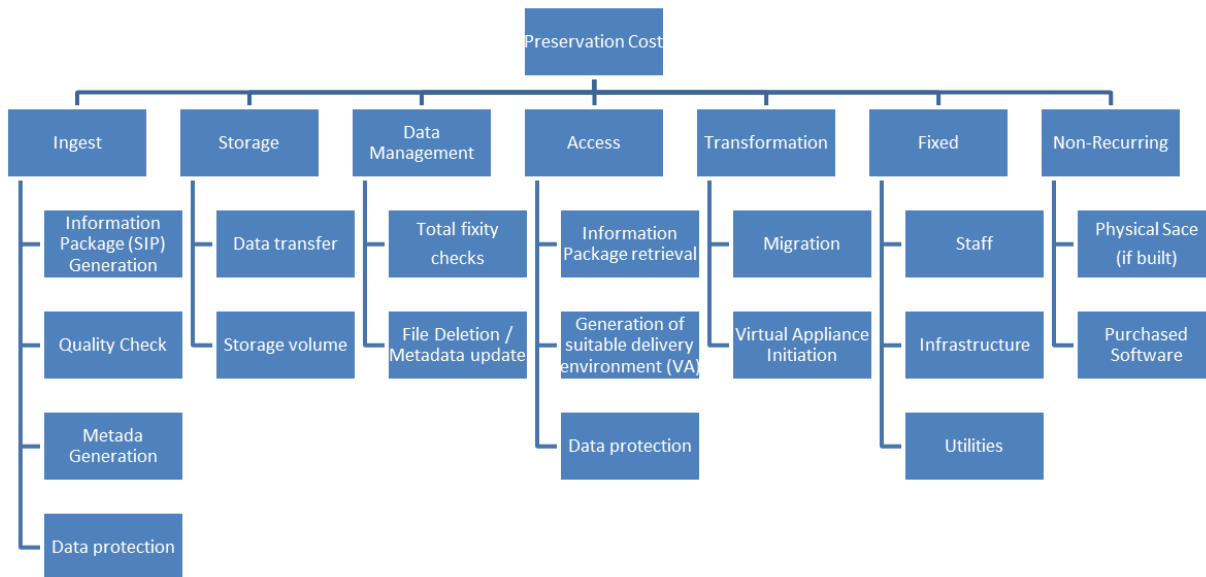
Despite sector differences, the project has identified a common work breakdown structure (WBS) for the preservation of digital data, as shown below:



ENSURE work breakdown structure

From the WBS, the project has developed a generic cost breakdown structure which overlays these activities:

⁷ Badawy M., Shehab E M, Sanya I. and Baguley P. (2013) " TOWARDS A FRAMEWORK FOR PREDICTING WHOLE LIFE-CYCLE COST FOR LONG-TERM DIGITAL PRESERVATION " Proceedings of The 11th International Conference on Manufacturing Research (ICMR 2013), Cranfield University, Cranfield UK, 19-20 September 2013, pp 151-156, ISBN 978-1-907413-23-0, ISSN 2053-3373.



8

Generic cost breakdown structure

Uncertainties considered are defined as:

- A state marked by the inability to specify an entity (outcome, event, or occurrence) with precision.
- It is the lack of certainty, a state of having little or no knowledge about the existing state, or future outcome.
- Uncertainty does not always imply loss or damage, they sometimes create opportunity for value creation.

And grouped into:

- Software
- Hardware
- Human Skills
- Preservation Plan

⁸ Braud, M., Edelstein, O., Rauch, J., Rabinovici-Cohen, S., Voets, D., Sanya, I., Randers, F., Droppert, A., Klecha, M. ENSURE: Long term digital preservation of Health Care, Clinical Trial and Financial data. iPRES 2013.

And the following categories:



Long term Digital preservation lifecycle

Considering all of these factors, the Cost Optimisation Engine will process parameters entered by users from within the healthcare, finance and clinical trials sectors and generate a recommended cloud based preservation plan for their digital data, with an estimated investment cost.

The ENSURE Project is drawing to a close and is due for completion in February 2014.

B.4 Introduction to 4C Project

- Understand the overlaps between ENSURE and 4C
- Test out our thinking
- Check our understanding
- Refine our methods and messages
- Can we share data?

Project Summary

The Collaboration to Clarify the Costs of Curation (4C) project will help organisations across Europe to more effectively invest in digital curation and preservation.

Vision

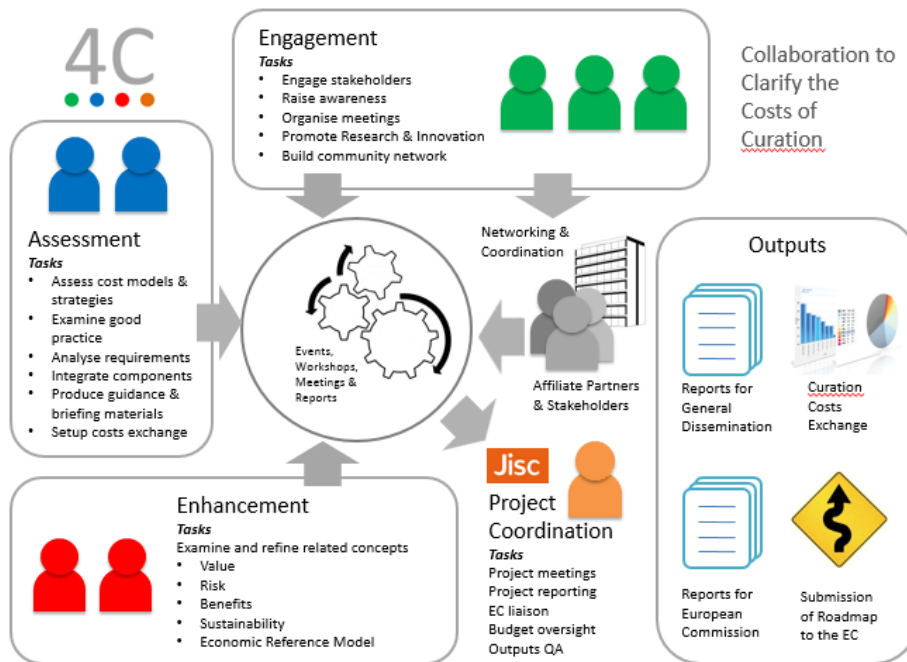
The 4C vision is to create a better understanding of digital curation costs through collaboration.

Mission

Our mission is to provide useful, useable resources which support the process of cost management in digital curation.

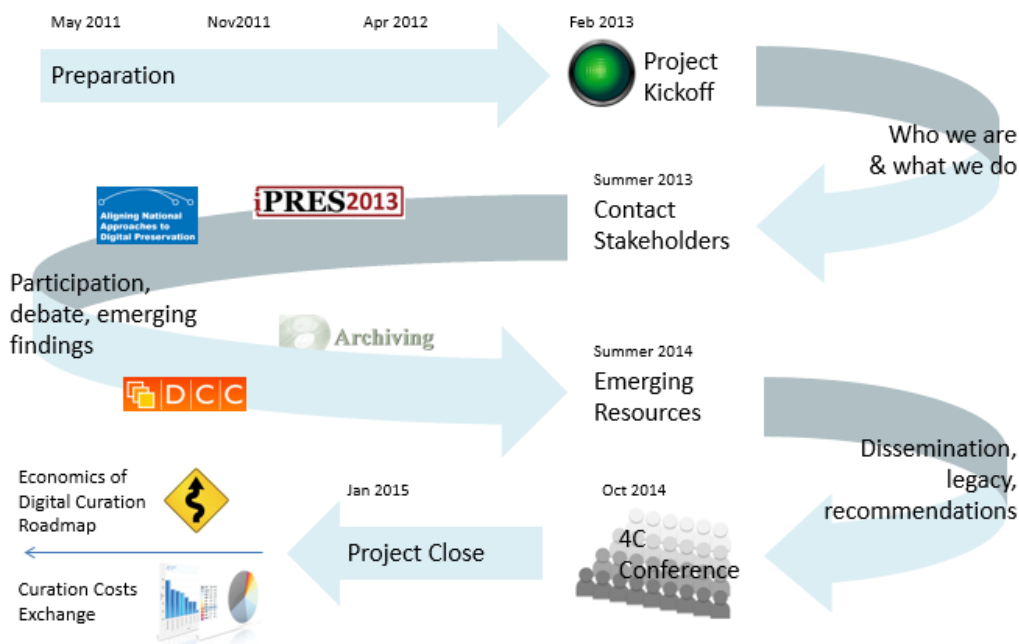
The Project is broken down into five work packages, including Project Management:

⁹ Shehab E M, Chuku G and Badawy M. (2013) "A Framework for Identifying Uncertainties in Long-Term Digital Preservation" Proceedings of The 11th International Conference on Manufacturing Research (ICMR 2013), Cranfield University, Cranfield UK, 19-20 September 2013, pp 151-156, ISBN 978-1-907413-23-0, ISSN 2053-3373



4C workpackage overview

The project started in February 2013 and will conclude in January 2015:



4C timeline

The 4C Project has undertaken an initial consultation to understand the motivation for needing/wanting to understand the costs of digital curation within the following identified stakeholder groups:

- Research funders,
- Cost model experts,
- Big data science,
- Digital preservation vendors,
- Universities,
- Government agencies,

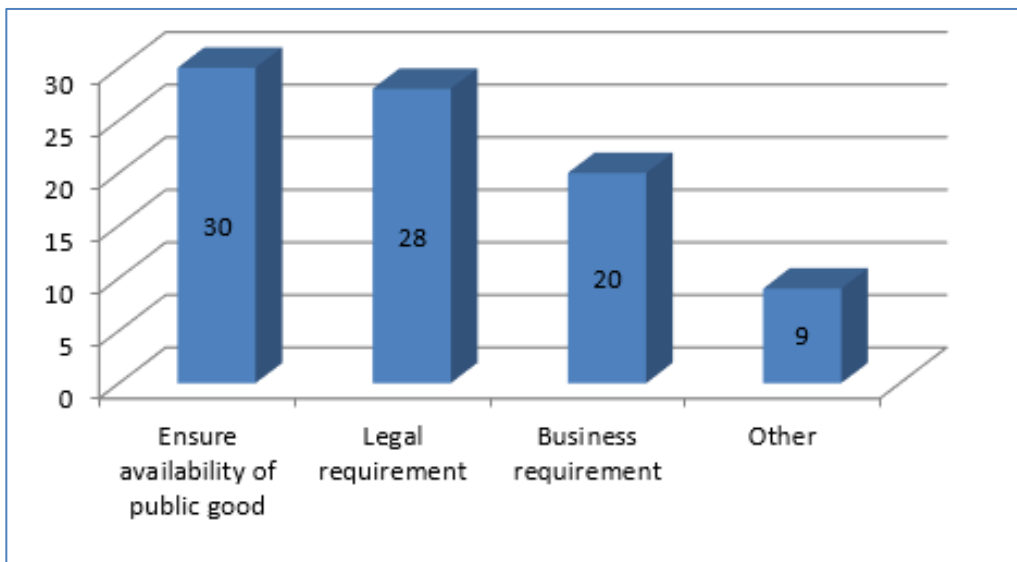
- Publishers & content producers,
- Data intensive industry,
- Memory institutions,
- Small medium enterprises,
- Media

Feedback includes:

- Understanding the cost of preservation may mean we can offer realistic and cost effective curation services to others.
- Understanding costs can support strategic planning.
- Understanding costs can support tactical decision-making.
- Understanding costs can provide evidence of cost-effectiveness and value.
- Clarifying and publishing the cost of digital curation can be used to enhance our organisation’s credibility. But this must be done along with the context of how the costs were calculated
- Understanding economic drivers can help to strategically align an organisation

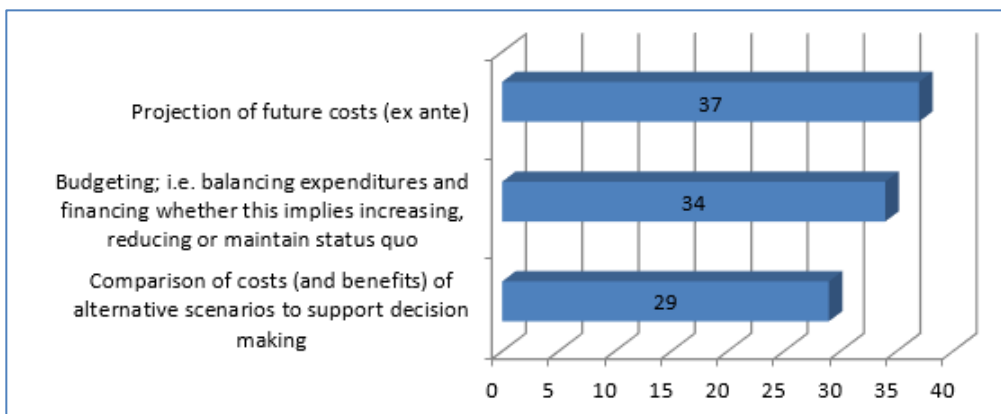
Other results include:

Motivation for keeping data



Motivations for keeping data

Reason for requiring financial information related to digital curation



Financial information need

4C aims to help by carefully analysing all of the information we assemble and making sure that it is passed onto our stakeholders through ...

- A series of state of the art reports
 - Cost model evaluation and a needs & gap analysis
 - Trust and quality (in relation to the cost of curation)
 - Risk, benefit, impact and value (in relation to the cost of curation)
 - From cost models to business models
 - A roadmap for future economic considerations in relation to digital curation
- New frameworks and models to assist with designing new approaches and building future tools
 - Indirect economic determinants
 - An economic sustainability reference model
 - A gateway specification for future cost models
 - A cost concept model for digital curation
- A Curation Costs Exchange

The project has already completed a Cost Model Survey and Needs & Gap Analysis, which evaluates 10 models against 79 criteria:

Results of the model evaluation for the characteristic “Model type”

Model type	T-CMDP	NASA-CET	LIFE3	KRDS	CMDA	CMDP	DP4lib	PP-CMDS	CDL-TPC	EMLTS
Economic model	No	No	No	No	No	No	No	No	No	No
Benefit model	No	No	No	Yes	Yes	No	No	No	No	No
Cost model	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Activity checklist	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Past/current costs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Future costs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Future costs - short term	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Future costs-medium term	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes
Future costs-long term	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes

Model evaluation—Type

Results of the model evaluation for the characteristic “cost variables “

Cost variables	T-CMDP	NASA-CET	LIFE3	KRDS	CMDA	CMDP	DP4lib	PP-CMDS	CDL-TPC	EMLTS
Quantification of assets	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Consideration of access frequency	No	Yes	Yes	Yes	No	No	No	Yes	No	No
Support of different asset types	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Limits of asset numbers	No	Yes	Yes	No	No	Yes	No	No	Yes	No
Quality of activities	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No
Quality of repository	No	Yes	No	Yes	Yes	No	No	No	No	No

Model evaluation—Cost variables

The results of the stakeholder consultation also showed that users need models for budgeting more than for accounting, and that they are generally more concerned with being able to account for running costs than for investment cost. These findings suggest that many organisations are able to get money for one-time investments, but have more trouble justifying the on-going costs, which also explains why they focus

on how to plan for future costs rather than past costs. To this end, the stakeholder consultation pointed out that users have a strong need for associating costs of digital curation with benefits. It is not enough to be able to make accounts and budgets; users also need to express the relation between the investments in digital curation and potential benefits that can be derived, and as mentioned this is about financial as well as more intangible, non-financial benefits. Thus, cost and benefit models for digital curation are important for decision making on alternative solutions and strategic planning, including risk management.

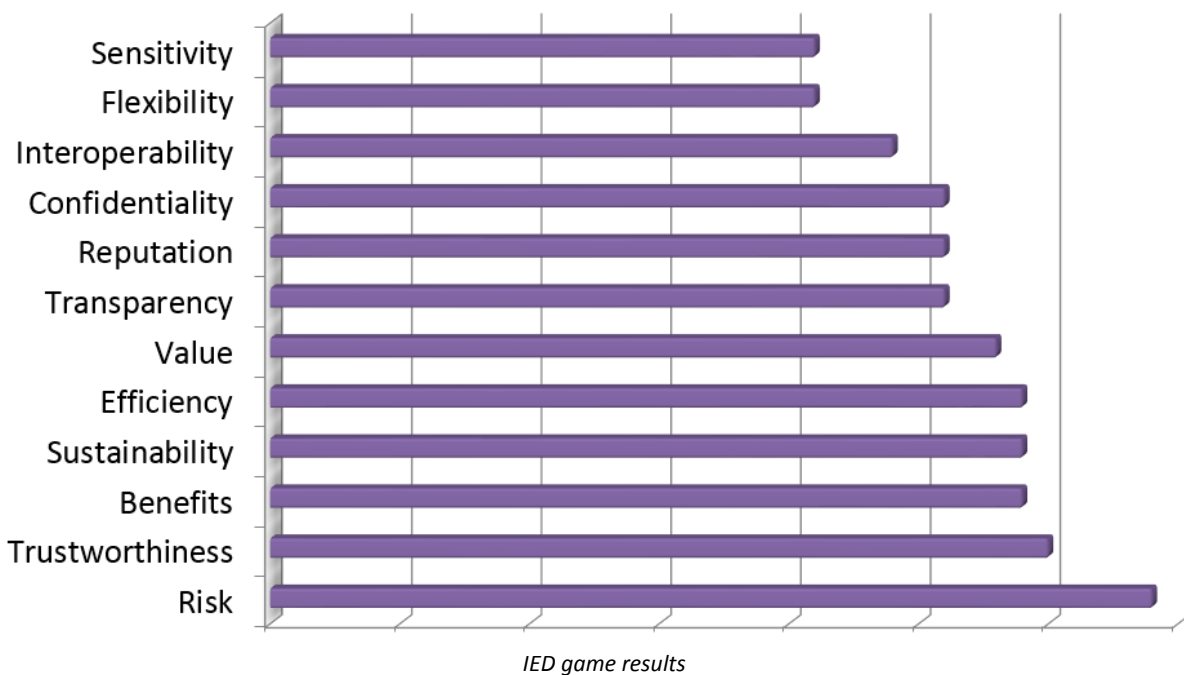
Recommendation 7:

Upgrade existing cost models with benefit modules, or develop models that integrate the assessment of costs and benefits.

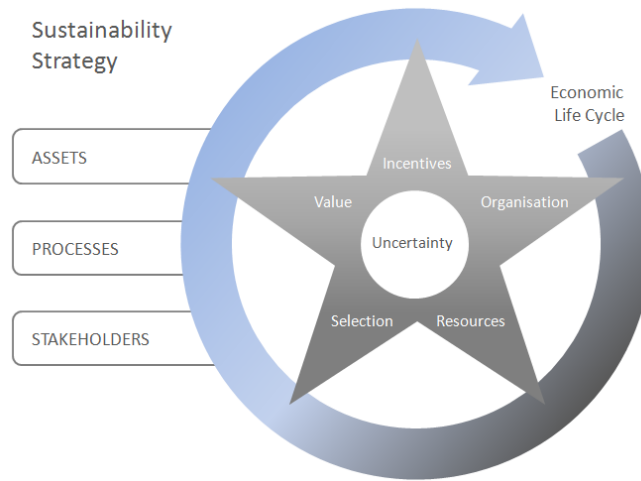
The project has identified a set of 'Indirect Economic Determinants' as a high level benefits framework, and has asked stakeholders to prioritise the following 15 values using an online 'game' on the project website:

- authenticity
- benefit
- efficiency
- impact
- innovation
- interoperability
- quality
- reputation
- risk
- sensitivity
- skills
- sustainability
- transparency
- trustworthiness
- value

The results show that risk is a constantly high priority in digital curation considerations:



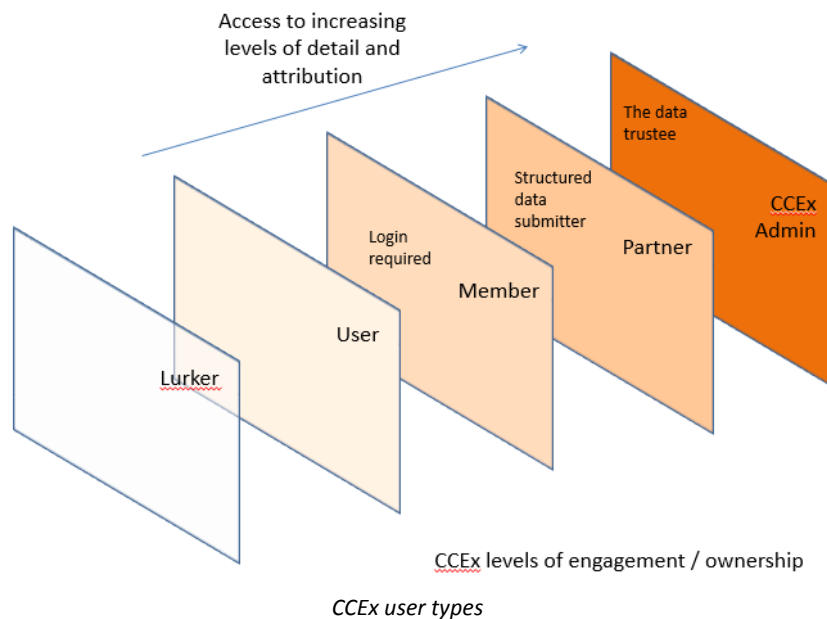
Building on these outcomes, the other project deliverables and the work already completed by the Blue Ribbon Task Force, the project will develop an Economic Sustainability Reference Model (ESRM) which has at its core the following model/sustainability strategy:



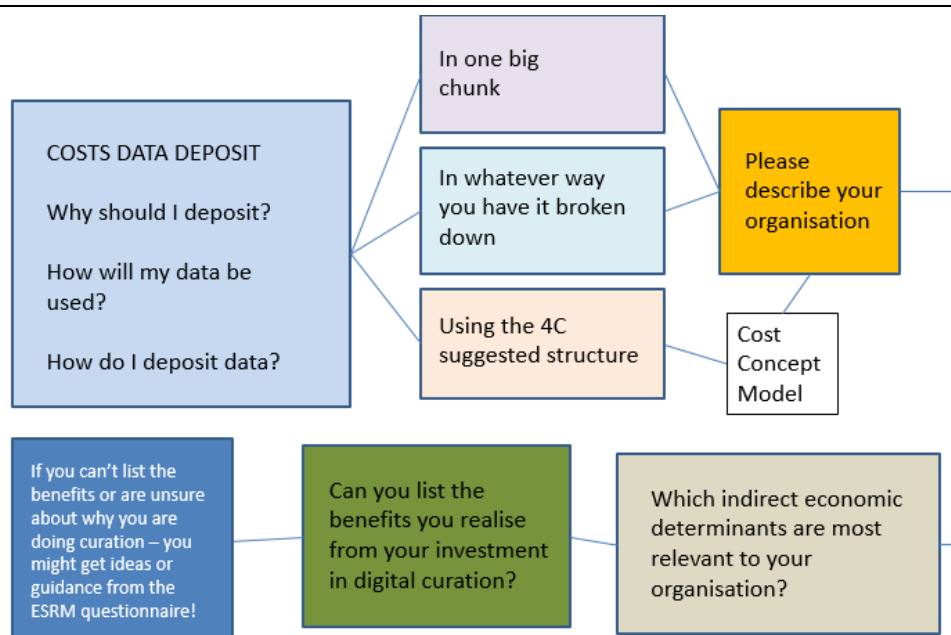
Draft ESRM model

The CCEx is intended to be an online, virtual community platform for the exchange of curation cost information. The CCEx will be used to gather cost information from partner organisations and stakeholders, submitted to the exchange using a Submission Form/Template. The form will aim to capture calculation processes, metrics, effort statistics, value calculations, from stakeholders in order to underpin future activity with empirical knowledge.

At the end of the project we could setup the access levels as below ... In the meantime we want barriers to deposit as low as possible.



Possible steps for CCEx input routine...?



Possible steps for CCEX input routine

B.5 Introduction to Digital Preservation Coalition (DPC)

DPC is a member organisation with 14 full and 26 associate members across the UK and Ireland.

Activities are arranged into 5 strands:

- **Workforce Development**
 - Training, scholarships and curriculum development based on research and best practice
 - Student placements, staff exchange and analysis of the digital preservation job market
- **Knowledge Exchange**
 - Practitioner network, briefing days and events to disseminate information – a platform to enable the dissemination of ENSURE Project data and findings for the benefit of members
 - Research projects – like 4C, TIMBUS, APARSEN
 - Publications
- **Developing Assurance & Practice**
 - Development of standards – using research outcomes and establish of working parties to develop best practice
 - Liaison with developers to communicate user needs using network of contacts developed through partnerships like this
 - Accreditation of digital preservation training & practitioners
- **Advocacy**
 - Demonstration of digital preservation benefits identified through stakeholder engagement
 - Public policy development, and embedding policy and regulation
 - Recognition of best practice & innovation
- **Partnership & Sustainability**
 - Relationships with existing and emerging partners – like ENSURE, other EU and other projects/organisations

- Extending relevant organisational membership – aim to develop market representation to include private sector, commercial and business entities through network of contacts

C Minutes of CERN Meeting, 7th—8th November 2013, CERN, Geneva, Switzerland

C.1 Attendees

CERN:

- Jamie Shiers (JS),
- Germán Cancio Melia (GCM), section leader IT
- Andrea Valassi (AV), IT-DB

4C:

- Anders Bo Nielsen (ABN), Danish National Archives
- Katarina Haage (KH), German National Library

C.2 Agenda

Thursday 7 November		
14:00 – 14:30	<ul style="list-style-type: none"> • Introductions • Presentation of the 4C project (stakeholder consultation, economic modelling, assessment work, costs gathering) 	Jamie, German, Andrea, Anders Bo, Katarina
15:00 – 18:00	Focus on CERN: <ul style="list-style-type: none"> • What CERN is currently doing • Bit preservation costs and storage issues • Software porting, data format migration • Response from HEPiX to the questionnaire; report from use cases and associated business cases for APARSEN and RDA 	German Andrea Jamie
Friday 8 November		
9:00 – 11:00	Focus on 4C - Assessment (WP3): <ul style="list-style-type: none"> • Cost concept model (practical test with CERN information) • CCEx (get feedback from CERN on this) 	Anders Bo Katarina
11:00 – 12:00	Focus on 4C - Enhancement (WP4): <ul style="list-style-type: none"> • Indirect economic determinants exercise “Focus Group Game”) • Discussion on audit and risk assessment costs 	Anders Bo and Katarina
13:00 – 13:30	<ul style="list-style-type: none"> • Discussion around next steps and additional engagement opportunities, e.g. → DPHEP costs and cost model workshop in January 2014 at CERN → Idea: CERN as participant in CCEx	Jamie, Anders Bo, Katarina
13:30 – evening	<ul style="list-style-type: none"> • Time for more discussion 	Jamie, German, Andrea, Anders Bo, Katarina

Table 26—CERN agenda

C.3 Session 1 day 1—CERN storage and digitization

Participants: CERN: Jamie Shiers, Andrea Valassi, German Cancio Melia; 4C: Anders Bo Nielsen (DNA), Katarina Haage (DNB).

Topic: Storage of the digitized material, bit stream preservation and costs for a digital preservation at CERN (example)

Context

This meeting was an in-depth discussion on costs in digital preservation. A first contact between CERN and the 4C project was in the scope of this year's iPRES conference in Lisbon at the 4C workshop. There Jamie Shiers presented the 4C members and the other workshop participants CERN cost data of DPHEP (data preservation in High Energy Physics) and suggested strategies how digital archiving of huge amounts of data could be handled/ managed at CERN.

Terminology

The terminology in the field of digital preservation differs from subject or kind of organisation. It is therefore important to clarify before the beginning of a discussion if both parties mean the same when saying a particular term. The 4C glossary can help to clarify eventual confusion beforehand.

Data preservation

JS thinks it would be more reasonable to say: "Data preservation has a cost" than "Data preservation is costly" when you have to justify the dire necessity of doing it towards the management, for example.

At CERN so far there is no structured organisation of data preservation; there are computer specialists but no archivists or specialists for digital preservation; the aim is to build up a department, to develop a policy and a structured approach.

JS also refers to the last report¹⁰ (May 2012) of the DPHEP study group on Sustainable Data Preservation in High Energy Physics. The former, less comprehensive report¹¹ (2009) addressed general issues related to data analysis and management beyond the lifetimes of collaborations. The following recommendations can be drafted:

- urgent action is needed for data preservation in HEP;
- the preservation of the full capacity to do analysis is recommended such that new scientific output is made possible using the archived data;
- the stewardship of the preserved data should be clearly defined and taken in charge by data archivists, a new position to be defined in host laboratories;
- a synergic action of all stakeholders appears as necessary;
- the activity is best steered by a lightweight organisation at international level.¹²

The cost of data preservation in order to facilitate reuse of data should be proposed as a small percentage of all project budgets based on payment up front.

¹⁰ <http://arxiv.org/pdf/1205.4667>

¹¹ <http://arxiv.org/pdf/0912.0255>

¹² Status report of the DPHEP working group, p. 6; <http://arxiv.org/pdf/1205.4667>

Data storage at CERN

There seems to be no central data storage policy at CERN. CERN is very much project based, and even though the projects last decades it appears that each project decides how it will storage and preserve its data. There is a general agreement that tape due to the data amounts is the only rational choice as storage media.

Data preservation at CERN is not only about data storage or archival storage in OAIS term.

Migration of data formats and even programs is a central preservation task. Unfortunately, programs do not appear to be designed to run in the long term with the need for format migration. Apparently independent data specification (i.e. apart from the programming language) is not a central part of data preservation. C, C++ and Java seem to be the most dominant languages.

Cost Model/ Cost Concept Model

CERN does not have/work with a cost model or a cost concept model.

Germán Cancio Melia presented three different possible cases for the cost of a 20 year digital archive (not specifically for CERN).

His presentation refers to a generic archive and considers the costs of hardware, media, maintenance and electrical power. The three scenarios are:

- a) 100PB growing @ 50PB / year (“CERN-like”)
- b) 100PB + 50PB initially, then +15% / year
- c) 100PB, no further data

Furthermore, the examples are based on limitations, such as:

The archive is tape based with a disk cache front-end

- the disk cache capacity is for 10% of the archive
- the archived data is not compressible / de-duplicable and the tapes are working at 100% capacity
- the tape media is kept for 2 cycles, then they are obsolete
- the model is based on 7x3 year cycles (21 years) as a result of a combination of the lifetime of tapes libraries and drives.
- duty cycle of 30% for both disk and tape servers
- enterprise tape is used, not LTO.

Parameters which are not included in these reflections:

- Manpower costs
- Software development / licensing costs
- General DC operations / floor space cost
- No assumptions on HW/media resale
- No inflation / interest rates; payments done upfront

The analyses of all three cases are shown in a spreadsheet¹³; one tab is for global parameters, one for each scenario including graphs when scrolling down and the green fields show the input data.

Additionally, they are visualised in Germán’s presentation¹⁴ in graphics from page 9 onwards.

¹³ <https://www.dropbox.com/s/ai2ew3682sbg5k5/20-year-archive-cost-v2.xlsx>

¹⁴ <https://www.dropbox.com/s/sw856vwtfrzo19y/20-year-archive-cost.pptx>

In short, the cost of storage on tape is expected to decrease drastically in the next three generations due to technological advances.

HEPIX bit preservation working group

Germán also shared information about the HEPPIX bit preservation working group between DESY and CERN. Background of this working group is to share ideas and experiences on bit preservation activities across sites providing long-term and large-scale archive services. For this purpose a questionnaire¹⁵ was developed.

The results can be summed up as follows¹⁶:

- Archiving has become a reality by fact rather than by design
- Often no clear understanding, SLA or agreement of how long archived data should live
- Bit rot implying data loss is a reality. Missing QoS or detailed SLA's defining acceptable data loss rates. Users seem to expect "no data loss"
- Redundancy of data is a way for reducing data loss
- Regular archive audits help improving reliability & reducing migration troubles
- No inter-site replica recovery; sites rely on users/VO's for this

The survey gives a proper overview of HEP bit-level archiving. For the future of the working group it means that more best practices are to be collected, documented and shared at the forthcoming HEPPIX meetings and to tackle automation of data recovery for distributed archives.

The whole presentation of the results can be found here¹⁷.

Awareness

For CERN it is important to call attention to the topic within the community and towards the decision makers of the organisation and get things started quickly. The cooperation with a project like 4C is a welcomed opportunity to raise awareness from another point of view.

C.4 Session 2 day 2—4C and CERN cooperation

CCEx

The Curation Costs Exchange was discussed after first ideas from the CCEx Design Group were presented on how the CCEx could look like.

CERN's input:

- Has to be available and running over the lifetime of the 4C project
- CCEx as platform for the exchange within the community; not only data but experience, approaches etc.
- CERN would contribute to this as a member or even a partner (as to be seen in NG graphics¹⁸ on Google Drive)

¹⁵ <http://w3.hepix.org/bit-preservation/lib/exe/fetch.php?id=bit-preservation%3Aintroduction&cache=cache&media=bit-preservation:questionnaire-hepix-bit-preservation-v1.pdf>

¹⁶ See p. 9: <https://indico.cern.ch/getFile.py/access?contribId=45&sessionId=3&resId=1&materialId=slides&confId=247864>

¹⁷ <https://indico.cern.ch/contributionDisplay.py?contribId=45&confId=247864>

¹⁸ <https://drive.google.com/?tab=wo#folders/0B-2-xr2TAgEKa21waUotekwzNlk>

Economic Sustainability Reference Model (ESRM)

The ESRM was not discussed at this meeting but could be introduced/ presented to the community at the workshop on DPHEP in January 2014 at CERN?

Indirect Economic Determinants

The indirect economic determinants are being valued online through the focus group game on the 4C website so the results can already feed into the statistics and graphics. Logins have been provided to JS, AV and GCM.

Audit and risk assessment costs

CERN so far never did any audit or a risk assessment, yet. There is also no disaster plan or preparedness.

Further Engagement

CERN and 4C aim for further engagement in the sector of cost of digital curation to raise awareness for the necessity of acting, especially in the big data science community. JS is planning a DPHEP workshop¹⁹ on full costs of curation on January 13/14 2014 at CERN where 4C is going to contribute with a talk and a discussion. The 4C project also plans to contribute with a focus group as a satellite event attached to the workshop with focus on the CCEX.

As mentioned above, CERN would be interested in actively taking part in the development of the CCEX under certain conditions.

¹⁹ <https://indico.cern.ch/conferenceDisplay.py?ovw=True&confId=276820>

D Report on 4C Focus Group at iPRES 2013 in Lisbon, Portugal

D.1 Attendees

4C:

- Alex Thirifays, DNA
- Sabine Schrimpf, DNB
- Katarina Haage, DNB
- Diogo Proença, INESC-ID
- Sarah Norris, DPC
- Paul Stokes, Jisc
- Neil Grindley, Jisc
- Raivo Ruusalepp, NLE

Participants:

- Anna Henry, TATE
- Yvonne Fries, ZBW Kiel
- Artur Caetano, INESC-ID
- Kirnn Kaur, British Library (APARSEN)
- Paul Wheatley, University of Leeds
- Catherine Jones, Science + Technology Facilities Council
- Christina Bankhardt, AbbVie
- Sheila Morrissey, Portico

Agenda

Time	What?	Who?
14:00-14:15	Welcome Short introduction of 4C approaches, goal, priorities etc. Short self-introductions	Katarina Haage, DNB Neil Grindley, Jisc
14:15-15:15	Presentation of Focus Group “Game” results Brief introduction of the concept Group discussion in smaller groups about specific questions according the determinants Discussion of the group results in plenum	Raivo Ruusalepp, NLE
15:15-16:15	Presentation of ESRM Brief introduction to the model Introduction of the 4C ESRM self-assessment questionnaire based on the model Discussion on relevance and potential of the model Capture recommendations from the discussion	Neil Grindley, Jisc
16:15-16:30	Wrap up	Sabine Schrimpf, DNB

Table 27—iPRES Focus group agenda

Participants motivations to take part and expectations

The participants represented most stakeholder groups that are of interest to the 4C project. From the motivations and expectations expressed in the introductory round, it became clear that motivations ranged from “experience exchange” to “hope to get more clarity on economically relevant concepts” and “hope to find out if [my institution] does preservation in an efficient way”.

Indirect Economic Determinants (“IED”)

The 4C concept of “Indirect Economic Determinants” was presented by Raivo Ruusalepp (NLE). He explained that the seemingly rather unfamiliar term “IED” was chosen in the attempt to reach out to the higher level managers and decision makers, who are assumedly more familiar with business terms than with DP terms. The intention of 4C is to help them seeing the business case in digital preservation.

Most DP managers will be familiar with the terms “direct costs” and “indirect costs”. There are costs, however, that do not easily fit into these 2 categories (e.g., the costs of an audit to become certified as a Trustworthy Digital Archive). These kinds of costs have not been well described yet. The 4C projects sets out to describe them more clearly and calls them IED for that purpose. By working through the list of IEDs, an organization that is tasked with DP can assumedly capture its unique organizational context and thereby get a better understanding of its mission. Ultimately, they are hoped to help determine the benefits of digital curation.

In preparation of the focus group meeting, the participants had been asked to rank the list of IEDs. The combined ranking led to the following order: 1. Risk, 2. Trustworthiness, 3. Benefits, 4. Sustainability, 5. Efficiency, 6. Value.

Break out group discussion on IEDs:

In a break-out session, the participants were roughly sorted into a “memory institutions group” and a “non-memory institutions group”. Both groups discussed separately about the IEDs.

The “non-memory institutions group” reported back that they had some difficulties with the IEDs. Some of the IEDs were considered a mere precondition for some organizations and an organization doesn’t have much choice in prioritizing or even selecting from the list of IEDs. The participants raised the question if it can’t even be assumed that the most general, most well understood, terms, will likely always be on top in any organization. The participants also raised the question how the IEDs fit into cost models. Neil Grindley explained that they are supposed to help making business cases as they help to create a narrative around cost modeling.

The “memory institutions group” reported back that they thought it was likely that managers and practitioners will have different views on the importance of the IEDs and that it would help to have scenarios to highlight the conceptual terms. Although this is not in the scope of the 4C project, it was registered as input into the roadmap.

The Economic Sustainability Reference Model (ESRM)

The ESRM was presented by Neil Grindley (Jisc). He explained that different people have, due to different roles, different perspectives on digital assets, their values and the benefits from preserving them. Like the IEDs, the ESRM is mostly targeted to high level managers and decision makers that decide about time and effort spent in digital preservation. Intentionally, the ESRM keeps clear of costs. It is designed more as a managerial tool and shall serve as a tool to make first steps into more detailed cost/benefit considerations.

“ESRM exercise”—interviews along the ESRM appendix between 4C members participants and non-member participants

All ESRM appendices with notes were collected by Neil Grindley for further analysis. In a feedback round, participants reported that they had some difficulties with the questions, and that they found them only “more or less useful” altogether. Specifically, it was mentioned that public organizations have little or no control over some of the issues that the questionnaire touches upon. There should be an answer option: “Not applicable“, or “not in our control“. Furthermore, the term “issue” was not quite clear to all participants. The whole ESRM document was found rather too long and detailed to be read by the high level managers and decision makers that it is targeted at. It was suggested that this target group would better be served with a 10 pages checklist.

The final recommendation of the participants was that the purpose and the value of completing the ESRM exercise needs to be made clearer and needs to be pitched by the 4C project.

E Notes from Focus Group #2—Industry Group, 12th December 2013—Jisc, Brettenham House, London

Attendees

Neil Grindley, Jisc

William Kilbride, DPC

Sarah Norris, DPC

Carol Jackson, DPC

Margaret Katny, BBC

Anne Archer, Lloyds Banking Group

Chris Fryer, Northumberland Estates

Sean Barker, BAE Systems

Richard Wright, Consultant to BBC

Matthew Addis, Arkivum

Agenda

1. Welcome, Short introduction of 4C approaches, goal, priorities etc.
2. Self-introduction
3. (a) Reminder Economic Sustainability Reference Model (ESRM)
(b) Introduction of the 4C ESRM self-assessment based on the model
(c) 4 open questions as a start for the discussion
4. ESRM Exercise

Discussion on relevance and potential of the model

Capture recommendations from the discussion

For the purposes of feedback, these notes capture the outputs of points (3) and (4)

3. (a) Reminder Economic Sustainability Reference Model (ESRM)

Benefits

In response to the ESRM ‘reminder’ slides the group observed that the phrase ‘This resource allocation decision must be based on a thorough understanding of the long-term costs of digital curation—i.e., the required investment—as well as the anticipated benefits from curation—i.e., the expected return on investment’ did not fully reflect the anticipated benefits.

Had it been written ‘e.g. the expected return on investment,’ this may have been more accurate.

The return on investment is only one kind of benefit, others include the compliance to legal obligations for example.

Avoiding the negative implications for not complying with legal obligations – ‘the cost of loss’ was seen to be another benefit, loss could be defined as financial loss through the imposition of fines, reputational loss, or loss of license to operate.

‘Curation’

The group was asked to reflect on the term curation and what this meant in their own contexts. The group returned the following observations:

- It is not ‘file and forget’
- Active management of digital assets with some value
- ‘Continuity’
- Digital sustainment
- Lifecycle
- Preservation for usability

The group noted that the term was not used widely within their own contexts, and saw it as a term used mostly within specialist communities.

ESRM Lifecycle

Discussing the ESRM’s note that ‘the general pattern of economic decision-making includes two endpoints’ the group observed that the model was very ‘repository-centric’ and the points identified would be too late, particularly for the aerospace industry.

Ideally, within aerospace what users need/want should be considered at top level, and this cascaded down to the executors of a task. Therefore tools used are chosen on the basis of their sustainability. Investment decisions are made when sustainability criteria are known – it would be very costly to change tools half way through the project, and have to verify the design.

Within pharmaceuticals, it is expected that data is ‘inspection ready’ from the moment of creation, and that the authenticity of drugs on trial can be demonstrated immediately, and throughout its lifecycle.

For banking the ESRM model works better, as the top down approach described by aerospace and pharmaceuticals is not taken. Banking archives start to look after data once they have been designated of archival interest... but they may not have the same endpoint however.

For broadcasting, the aim is that data (a programme) should get to the archive before it gets transmitted, and that the archive is involved from the point of commissioning. But there is no endpoint. Programmes are never removed from the archive once archived.

The group also noted that archives have their own economic lifecycles, and that often institutions or organisations holding data collapse, and data is lost.

3. (c) 4 open questions as a start for the discussion

1. What is the main motivation for your organisation to “afford” digital curation?

When making a recent case for funding for an archive within banking, the focus was on risks, namely – litigation, regulatory, reputational/operational. While there is also a recognition that the archives contribute towards CSR and corporate memory, there is no channel to exploit this.

Estates also noted the main motivation for digital curation was risk avoidance, particularly in preserving business critical information and enhancing core functions, improving day to day operations through records management.

Broadcasting noted that the archive supports the core function of the BBC, for reuse and for heritage purposes. They are also obliged under the conditions of the BBC charter to preserve data, which provides a clear route to make the case for investment. It was observed that this might be different for ‘harder nosed’ commercial archives, who won’t keep data if they don’t see in commercial value in it.

Aerospace observed that electronic records are much cheaper to keep, and reuse is not that high, especially for military customers. They are always looking at ways to reuse the data, if that is permitted, and also need to retain information on spares for repairs.

2. Can you categorise the digital objects that you are in charge of being either ASSETS or LIABILITIES?

Most of the group saw their data as both ‘asset’ and ‘liability.’

- Liability – if you don’t keep it you get fined, there are negative consequences if it is lost.
- Asset – if you can make more money from it that it costs to preserve.

Banking noted that for ‘assets,’ this would be that the benefits or risks mitigated by preserving the asset would need to outweigh the costs of preservation as financial gain could not be made directly from the assets.

3. What benefits or outcomes does your organisation expect from digital curation?

The group generally felt that this was covered by the first question... product liability, certification and reuse.

It was also felt this could be approached from the outcomes which were expected to be avoided, e.g. you won’t get fined, and you won’t have your license revoked.

However, it was noted that it is easier to build a strong business case with positive outcomes.

4. Do the costs of curation actually matter in your organisation? And/or the potential Return of Investment?

All agreed that yes, the costs of curation matter, although cost is not the only important factor—see earlier discussion on return on investment.

Sometimes it’s what the organisation exists to do – therefore all costs are associated with this activity. Do you therefore attribute all of your staff costs into the ‘cost of curation?’ It makes it more expensive that way.

Archiving is seen as a way to save costs. It does cost money, but it saves money long term.

What is the cost of moving archives? Those who used external archiving service providers don’t expect their service to survive as long as they need it, and is therefore not so repository focused as service focused. It is about sustaining the data not the service. These two approaches represent different costs.

The archival service provider builds in succession planning, business planning and an exit strategy for customers – with a view to achieving portability of data at the lowest cost.

The group was interested to hear about the CCEX. They agreed that they would be keen to use such a resource, if they could take something away from it that was endorsed by expertise, and would strengthen a business case.

They also agreed that there is a huge value in cost data (shared by others), and a ‘Trip Advisor’/recommender option would have little value without real cost data. There may be other ways to do this—e.g. headcounts? Tangible quantities? Budgets and costs are too hard for many people.

4. ESRM Exercise

General observations

The group asked why the ESRM Assessment did not contain a section on ‘risk.’ NG explained that ‘risk’ had become ‘uncertainties’ and included those not just economic.

It was suggested that the checklist should contain a ‘have you identified uncertainties...?’ section, including:

- Risk of whether you may/ may not get funding
- Inspection/ audit—ability to pull back content
- Some rare yet huge consequences if/when they do happen and looking at the life cycle of the content

The modern definition of risk includes ‘opportunities’—the opportunity to monetise assets vs the risk of loss.

The assessment required some members of the group to manipulate certain concepts to fit their contexts. The concept of value, for example, could not always be seen as positive, but often in terms of the value of ‘not having’ the asset, or in terms of sociological value rather than economic.

Value also depends on levels of quality—there is a premium to be paid for higher levels of quality.

The group thought that practitioners would be willing to do the work (the assessment) as long as the results for their customers/users were clear and some questions may need to be asked differently depending on who and where you are in ‘the process.’

NG explained that the assessment had been designed for those managing data in a managed environment.

It was suggested that the introduction should make this clear, and either provide greater context or make the questions more generic...if possible? The value section also, needs to allow for different kinds of value. Cost and value should be delineated more, the group got the feeling that the terms ‘cost’ and ‘value’ were used interchangeably—that is not the case.

Assets

With reference to the term ‘manageable formats’ – this is contextually dependent. Industry puts a lot of money into making assets manageable. This is not a passive process. Another question might be ‘who/how do you make the assets available in the formats you want?’

The rights issue only arises at the point of access for broadcasting.

This is not necessarily the case for other archives—you can often end up with orphan works, orphan software, orphan users...

With reference to the term ‘homogenous’—this is not necessarily chaotic. In broadcasting, there are many different file formats, but the volume of each means that batch processing is still practical.

Stakeholders

‘Curation role’ is not a necessarily a term used in industry, definitely not in pharmaceuticals.

With reference to the question ‘are the curation roles in the ecosystem clear to everyone?’ the group felt that there were too many/mixed metaphors in use.

Could who you are influence the slant of the self-assessment?

Processes

Emerging technology influences the way the group curates and vice versa, it is a two way process.

Emerging technology is both an opportunity and a risk, any change has to be validated so it is much harder to adapt. When adapting, it needs to be done with confidence.

If processes can't adapt—are they really curation processes?

This all depends on the users of your process.

Value

See earlier discussion on Value.

Resources

The group felt that they often had to work hard to justify investment in resources, so this section was good to see, and highlighted the issue well.

It was suggested that this section should also include intellectual resources – e.g. skills, expertise, knowledge; not just human resources. Perhaps this could be split into a separate question.

It was noted that the cost of getting the correctly skilled people is also a factor, and this is significantly more when you are a small contractor.

Selection

The group agreed that this is one of the most difficult aspects of digital curation. You can never have enough parameters, you can only make assumptions, and you never know whether you have enough information to make a decision.

Deduplication is a part of an agreed framework, although this is not necessarily the selection of the best copy.

Sometimes the cost of not making a selection is much higher than making a selection. It was suggested that there should be a question asking whether users knew what the cost of selection is?

This section provoked the group to think about many other questions. NG explained that the ESRM aimed to be simple, and that the assessment should be boiled down to the simplest core questions, which prompt you to think further in other areas.

Organisation

Policy is what, procedure is how. The questions are more focused on the how rather than the what.

This should also reflect reliability.

Incentives

From an archive point of view is there a conflict of interest between what the archive does/curates, why this is curated and what users need? This is less of an issue for commercial companies – they set up archives to do what they want them to. The issues of value and organisation also feed into this.

More general observations

The British Library and BBC would have to split this assessment into cases: e-print, sound archives, web archives etc.

It was felt by the group, that in some ways the scope of the ESRM was too great, and in others too small:

- It would be too complex for SMEs, big businesses like pharmaceuticals could manage it, but it's in the wrong language for them.
- Could we provide broad definitions? Tailored versions?
- Small/ large?
- Country specific approaches to reflect different accounting practices?

If the ESRM could emulate the OAIS that would be a good start and would provide a coherent story.

F Public report of the 1st 4C Advisory Board Meeting 11th June 2013, Jisc Office in London, UK

The first Advisory Board Meeting was held on 11 June 2013 at the Jisc Office in London, UK. In attendance were:

1. Neil Grindley, Jisc
2. Paul Stokes, Jisc
3. Rachel Bruce, Jisc
4. David Rosenthal, LOCKSS
5. Matthew Addis, Arkivum
6. Ron Dekker, Netherlands Organisation for Scientific Research
7. Sabine Schrimpf, Deutsche Nationalbibliothek
8. Raivo Ruusalepp, National Library of Estonia
9. Sean Barker, BAE Systems
10. Hildelies Balk, Koninklijke Bibliotheek
11. Alex Thirifays, Danish National Archive
12. Brian Lavoie, OCLC

Some highlights of the meeting:

Context, background, introduction to project

Provided by Neil Grindley:

- The project is a co-ordination action, intended to clarify and synthesise current information to make it more usable
- In the curation domain complex models are rife
- The chosen term for the process is “curation”
- Curation is an investment
- The EU is keen to use projects such as 4C to drive economic recovery through the engagement of SMEs

Reaction from Board members about high-level aims

Particular discussion points included:

- The Curation Costs Exchange (CCEX)
- The role of SME in relation to larger companies
- The sustainability of the project
- The governance of the project
- The Advisory board agreed to maintain an informal flow of information in the intervening period and to allow some information regarding the board and its remit to be published on the project web site.

Summary of Work Packages and their Objectives

WP1—The question of the openness and published status of documents relating to the project was discussed. At present most documents/outputs in development are not open, but when they are complete they will be. It was suggested that the 4C glossary is a document that should be open even during development.

WP2—Board members were invited to engage with the on-line consultation and provide feedback on the stakeholder analysis. The board noted that commercial stakeholder should include more than just SMEs. Other stake holders were identified including: Media companies; Pharma; Oil and gas; Scientific archivist groups (who meet on a regular basis); “stake holders we don’t know yet”— potentially the biggest group. Other related projects mentioned include BBC, Presto, APARSEN, TAPE (a Finish project).

WP3—The Work Package covers three areas: 1) Finding out what (cost model data) is out there; 2) Establishing what stake holders want; 3) Establishing the gaps and discrepancies between the models and the wants. The appropriate number of models (or frameworks) needed to cover “all” use cases was discussed, and the number and types of parameters for those models and the expectation regarding the use of model results. There was also discussion regarding the increasing adoption of “models (complex digital objects and the proprietary software required to access them) as the unit of curation (as opposed to “documents”) and how that might affect any curation models. Once again it was noted that the different use and ascribed meaning of the same word—“model”—in alternative domains could lead to confusion.

WP4—The Work Package sets out to fill some gaps in the existing cost models. It was noted that, although the CCEx could be a rich source of information for WP4, it would not be available in time. The board discussed the issues surrounding the reasons for curation, specifically the differences between keeping materials because of their intrinsic value, and keeping materials for compliance reasons. Comparisons were made between research data kept for its future research value and pharmaceutical data kept for compliance (although it was acknowledged that in most cases data will be kept for both reasons even if the relative weighting is different). Further comparisons were made between the standards in the US and in Europe. The Board recommended developing a map of the reasons for data curation.

WP5—Discussion of Work Package 5 was passed over in favour of increasing the time devoted to discussion of the project challenges.

Discussion of the main challenges the project faces

- Terminology—how do we most clearly explain to stakeholders what we want from them? - There is a potential tension between the use of the glossary internally (and in externally facing communications) to ensure that we put across a consistent, internally understood message and the life of the glossary beyond the lifetime of the project. What may be useful and pragmatic within the 4C arena may not be suitable for or accepted by a wider audience.
- Models and specifications—how can we ensure that they are useful and understandable? – Neil Grindley introduced the Economic Sustainability Reference Model (ESRM), currently in its 0.6 iteration. The glossary describes a model as “a simplified version of reality”. Neil Grindley’s hypothesis is that a reference model has a process that enables the user to change things for the better. The board discussed possible uses for the ESRM (and compared it to the OAIS model). It was postulated that a reference model should contain the building bricks to allow the creation of new models, or as a tool for comparing differences in models, but the ESRM as described is a tool to help the user to think about their needs and requirements in a different way (and is particularly

focused on the preservation of digital objects). A public draft of the model is to be made available on the web site for comment.

- Gathering Costs Data—who will give it to us and how should we analyse it? - It was proposed that the 4C consortium members should lead by example with all partners pledging to come up with 2 sets of costs each (giving a starting point of 26 sets). The board agreed that 4C should be smart in the data gathering process, using secondary indicators such as effort, using statistical methods, making the metrics gathered as easy for the supplier to source as possible and embracing “ball park” figures. Relative costs were deemed to be useful, especially in situations where costs and prices are being obfuscated by other supplier/customer arrangements.
- The Curation Costs Exchange—what is it and how do we make it useful? - A paper prototype with a group of “friendly” participants was deemed to be a good idea in order to get some unbiased opinions of needs (what they want from the CCEX) and concerns (what they fear about the CCEX). The people who had indicated their willingness to be contacted from the initial consultation exercise might form just such a group. Questions were raised about quality control, specifically would there be checks on who put in data and checks on data validity.
- Another potential use case could be those who wish to offer services using the CCEX to see where such services could be best targeted

G Annex to D2.5: 4C Project Communications Plan—Draft 'CCEx Communications Plan'

G.1 Introduction

The CCEx is intended to be an online, virtual community platform for the exchange of curation cost information. The CCEx will be used to gather cost information from partner organisations and stakeholders, submitted to the exchange using a Submission Form/Template.

The form will aim to capture calculation processes, metrics, effort statistics, value calculations, from stakeholders in order to underpin future activity with empirical knowledge.

This plan pertains particularly to optimising the uptake of the Curation Costs Exchange (CCEx) through the application of communications activities, and identifies the channels, audiences, information and content to be disseminated in order to achieve this objective.

It aligns key project messages with those particular to the CCEx, the frequency with which communications will take place, milestones for communications, quality controls and performance indicators, as well as responsibilities for undertaking these activities.

As a 'Coordination Action,' with an emphasis on this critical theme of communication, the fulfilment of the 4C project objectives are the responsibility of all project partners. Where specific deliverable tasks have been allocated, however, these are highlighted throughout the Project Communications Plan.

G.2 Objectives and Purpose

This Communications Plan sets out the activities designed to achieve the communications objectives relating to the uptake of the Curation Costs Exchange (CCEx). The statement of principle behind the Communications Plan is to facilitate access to collaborative tools that anyone can use, in the knowledge that information is shared without commercial exploitation, with a view to sustaining the currency and relevancy of the CCEx output.

Derived directly from the 4C project objectives, the plan and the activities outlined herein, therefore aim to meet this statement of principle, specifically to:

- Communicating the benefits of the CCEx to 4C stakeholder groups in order to create demand - Measurable in the uptake of the CCEx, i.e. in the number of people or institutions that use the CCEx and the amount of data cost sets shared through the CCEx
- Achieving an initial sharing of XX cost data sets within YY months/ weeks
- Sustaining the life of the CCEx through ongoing sharing of XX cost data sets within YY months/ years

The fulfilment of these objectives and the application of an 'open and social' communications model will ultimately facilitate the achievement of the main 4C project objectives by engaging users in sustainable dialogue throughout the lifetime of the project.

Measures of Success

The success of the CCEx related communications activities may be measured by the levels of:

- Initial stakeholder participation engaged.
- Sustained stakeholder participation

A good indicator of what we can expect is approximately the number of individuals who expressed their willingness to share cost information during the initial consultation (around 50). If just half of these individuals use the CCEx, this could be deemed successful.

Proposed success levels for the activities described are:

	Initial (Remainder of Project Lifetime)	Sustained (Post Project)
Good	25	5/ month
Acceptable	12	2/ month
Poor	5	<1/ month

Table 28—Proposed success measures for the CCEx communications plan

These measures are to be discussed and agreed by the CCEx Design Group.

G.3 Key messages

Messages pertaining to optimising the uptake of the CCEx may otherwise be identified as those categorised within the 4C Project Communications Plan, taking place in the ‘middle stages.’

Participation, debate, emerging findings

- resources to support the process of calculating the cost data preservation will be made available, through collaborating and information sharing between stakeholder groups;
- shared information relating to the costs and benefits of digital preservation, and used to facilitate the availability of cost modelling tools, will be used without commercial exploitation;

Specifically, communications relating to optimising the uptake of the CCEx will elaborate on these messages and should be supported by the following proposition and key messages which highlight the benefits and value which may be derived in participating with the CCEx:

Message	Applicable stakeholder group
Share background knowledge, best practice experiences and have access to ‘open data’ in order to:	
<ul style="list-style-type: none"> • support strategic planning and tactical decision making thus optimising budgets 	All stakeholder groups
<ul style="list-style-type: none"> • compare how similar organizations proceed to support decision making with regard to investment in digital preservation, on a political level 	All stakeholder groups
<ul style="list-style-type: none"> • increase the ability to manage the digital assets over time and enable the creation of new cost-effective solutions 	All stakeholder groups
<ul style="list-style-type: none"> • support the development of (internal/external) business cases for a digital curation strategy 	All stakeholders (except funders?)
<ul style="list-style-type: none"> • evaluating proposals for funding 	Funders

<ul style="list-style-type: none"> • keep track of your costs once a strategy is implemented 	All stakeholders
<ul style="list-style-type: none"> • benchmark your costs against others 	All stakeholders
<ul style="list-style-type: none"> • to maintain competitiveness 	Corporate enterprise

Table 29—Key CCEX messages

G.4 External Communications – Strategy and Timing

There are two key ways of using communications to encourage stakeholder participation in the CCEX:

- **Push strategy:** 4C uses its communications channels to ‘push’ the concept of the CCEX directly towards targeted stakeholder groups, supported by incentives for use and messages outlining long term benefits
- **Pull strategy:** 4C communications channels are used much more widely to convey the benefits of the CCEX, creating a demand for participation and causing stakeholders to ‘pull’ more information and opportunities to become involved.

Typically the most effective/appropriate forms of communication for each strategy are as follows:

Push

Personal communication through:

- Personal invitation and direct contact through email/ phone
- Exhibitions and demonstrations
- Offer of Incentives

Pull

Mass communication through:

- Social media – twitter, linked in
- PR – website news releases, publications

The CCEX communications plan will employ a combination of both strategies.

Timing

In order to optimise the participation in the CCEX development, it is anticipated that communications efforts should be focused within the following timeframes:

Months	Focus of effort	Strategy
12 to 24	<ul style="list-style-type: none"> • Creation of demand 	Pull
12 to 24	<ul style="list-style-type: none"> • Encouraging sharing of data – population of the CCEX • Shaping/testing prototype 	Push
24 to beyond project completion	<ul style="list-style-type: none"> • Sustaining demand 	Pull

Table 30—CCEX communications timings

G.5 External Communications—Means and Method

All communications should include the key messages outlining the benefits of participating in the CCEx:

- supports strategic planning and tactical decision making thus optimising budgets
- helps compare how similar organizations proceed to support decision making with regard to investment in digital preservation, on a political level
- increases the ability to manage the digital assets over time and enable the creation of new cost-effective solutions
- supports the development of (internal/external) business cases for a digital curation strategy
- helps evaluating proposals for funding
- keeps track of your costs once a strategy is implemented
- benchmarks costs against others
- maintains competitiveness

Push strategy

Personal communication through:

Means	Method
1. Personal invitation and direct contact through email/ phone	<ul style="list-style-type: none"> • Contact individuals who have previously expressed interest in testing and shaping the CCEx • Contact all individuals who had previously expressed willingness to share cost information in the initial consultation • Contact all individuals who had previously expressed staying in contact with the 4C Project following the initial consultation • Contact attendees at the following events to encourage participation
2. Exhibitions and demonstrations	<ul style="list-style-type: none"> • DPHEP, CERN Switzerland, 13-14 January 2014 • IS&T Archiving 2014, Berlin Germany, 13-16 May 2014 • Librarian Day 2014, Bremen Germany, 3-6 June 2014 • EuroCRIS, Rome Italy, 13-15 May 2014 • Third RDA plenary meeting, Dublin UK/Ireland, 26-28 March 2014 • IFLA Conference 2014, Lyon France, 16-22 August 2014 • Others...
3. Offer of Incentives	<ul style="list-style-type: none"> • See below

Table 31—CCEx communications push strategy

Personal invitation and direct contact through email/ phone

Email draft:

“As you may know [PARTNER NAME] is working with an EC-funded project called 4C the ‘Collaboration to Clarify the Costs of Curation’. Our thinking is that everyone’s job would be a lot easier if the costs of data curation were clarified, so the sorts of tools and services which 4C will produce aim to do this.

The project team is currently developing a specification for a Curation Costs Exchange (CCEx) which is intended to be an online, virtual community platform for the exchange of curation cost information. The CCEx will be used to gather cost information from anyone with some experience and an interest in digital curation / the cost of digital curation. We believe that through sharing cost data, the CCEx will be able to:

- *support strategic planning and tactical decision making thus optimising budgets*
- *compare how similar organizations proceed to support decision making with regard to investment in digital preservation, on a political level*
- *increase the ability to manage the digital assets over time and enable the creation of new cost-effective solutions*
- *support the development of (internal/external) business cases for a digital curation strategy*
- *help evaluating proposals for funding*
- *keep track of your costs once a strategy is implemented*
- *benchmark costs against others*
- *maintains competitiveness.*

So that this is immediately useful, I wonder if we could ask two favours:

Would your institution be willing to share a set of ‘test’ cost data for digital curation?

Retrieving as much cost data as possible to populate the CCEx will also make it more useful for those who are interested in the costs of curation. One of the objectives of the CCEx is to enable comparison of cost data across organisations and borders, because it can lead to effectiveness, and one of the most important drivers for using cost models that we identified in the stakeholder analysis is to reduce costs by enhancing efficiency. If the CCEx offers this display of cost data, users can start comparing costs.

All cost data shared with the 4C project will be treated confidentially and will be used solely for the purposes of building up data sets for the CCEx. Even as outputs for the CCEx, your data will remain anonymous.

Would you be prepared to help us by taking part in a focus group for testing the CCEx and its workflows this year?

4C is an ‘open and social’ project and we rely on input from stakeholders and practitioners to make sure we put together the right resources and that these are useful. Sharing your thoughts on what we’re developing would be provide a great deal of useful guidance.”

Exhibitions and demonstrations

The concept of the CCEx is still new, and the idea of sharing cost data might still be daunting to stakeholder groups. If the interface and benefits may be explained and demonstrated in person, with the opportunity to ask questions, it is likely that understanding of the concept and a keenness to participate will be improved.

The following framework for testing may be used within smaller focus groups:

- Having been introduced to the concept of the CCEx and its benefits, do you think this is a tool you would be interested in using?
- If so, which particular element of the CCEx would you be interested in? What would drive you to use it?
- The premise of the exchange is that you share your data to unlock greater access to more information, namely, other cost data.
- Would you share your data?
- If not, what would stop you?

- Too complicated
- Commercial confidentiality
- Ambiguity of data use
- We have seen that cost data is generally collected/managed by organisations in three ways:
 1. By labour/ capital costs
 2. By activity (for digital curation this might be: ingest, archival storage, access, data management, admin, preservation planning)
 3. Other
- Do you know how your organisation groups its data?
- We are exploring the ways in which data could be uploaded and shared with the CCEx. One suggestion is to provide the three options above.
 - Which would you choose?
- Selecting the ‘other’ option would require 4C project personnel to work through the data with you to categorise it (by activity) to make this comparable or useable – out of context this will be very difficult, and not sustainable once the project has finished.
 - We may need to contact you again to provide that contextual information if this was your selection, how would you feel about that?
 - If this is not acceptable, what incentives could the project offer you in order for the data to be organised according to a standard submission template?
- What would you like to see in return for sharing your data?
- Are there any ‘showstoppers’ or issues you can see in the process?
- What would the most resource intensive (time/user effort) part of the process for you?

Offer of Incentives

Despite greater understanding, some stakeholders may still be unwilling to share data (e.g. for reasons of complexity or confidentiality). The following incentives to do so may be offered to encourage cost data exchange:

- Assure all contributors that data may be shared anonymously
- Offer to visit stakeholder to work through data available if too complex to breakdown independently
- Offer early access to shared cost data
- Offer immediate access to sector specific information on cost modelling
- Opportunity to build networks and collaborations between similar organisations/institutions on basis of the CCEx

Pull

Mass communication through:

Means	Method
1. Social media	<ul style="list-style-type: none"> • Twitter • Linked in
2. PR	<ul style="list-style-type: none"> • Mailing lists • Website news releases: • Publications

Table 32—CCEX communications pull strategy

Social Media

Aligned with the 4C Project Communications Plan, and in order to reach a wide audience, the project will use the 4C website and blogs, plus Twitter feed and a #4ceu hashtag to publicise and create demand for the CCEx, whilst encouraging debate and participation.

Twitter

In order to raise awareness of the CCEx, maintain a continuous conversation about the tool, and create interest for more information, the following tactics may be used:

- Updates—content driven weekly updates, which link to blog posts on the 4C and other websites on the topic of the CCEx (outlining benefits and encouraging participation) and/or a CCEx section on the 4C website which maps progress, provides (controlled) access to the prototype for testing, screen shots/wire frames and commentary.
- Discussions—weekly/monthly twitter conversations either at planned and publicised times, or ‘impromptu’ conversations started by friends of the project asking pertinent and leading questions about the benefits of the CCEx.
- Hashtag—establishment of the hashtag #CCEx to enable those interested to follow updates and discussions.

To enable the best coverage of events where the CCEx may be showcased, all 4C colleagues re-/presenting the project at conferences or workshops etc. may have access to the 4C Twitter account. This may be beneficial not only for the CCEx but the whole project.

LinkedIn

- Discussions - The CCEx may also be advertised through the LinkedIn Groups of related projects, e.g. APARSEN, although a specific 4C LinkedIn Group will not be established in order to maintain the ethos of an ‘open and social’ project.

PR

Mailing lists – specific and particular audiences may be reached through subscribed email discussion lists. This channel would be best suited for inviting attendance at events where the CCEx will be exhibited and demonstrated, or for inviting volunteers for participation in cost data sharing and prototype testing.

Potential lists include:

- DPC-Discussion
- Digital-preservation
- Nestor

- APARSEN
- ENUMERATE
- The British Postal Museum & Archive
- Others...

Website news releases—specific and particular audiences may also be reached through news releases online. This channel would also be best for inviting attendance at events where the CCEx will be exhibited and demonstrated, or for inviting volunteers for participation in cost data sharing and prototype testing.

Websites include:

- 4C News pages
- Project partner news pages
- Related project new pages

Blogs and website—aligned with the 4C Project Communications Plan, project partners will continue to contribute blog posts in order to maintain an enduring web presence and awareness of the project. Blog posts on the progress of the CCEx development, with thanks to/posts by contributors (as desired) may be made on a monthly basis. The blog will not only provide a commentary on the progress of the CCEx, but will support the provision of content for twitter through which this activity will also be publicised.

Publications—periodic partner and related project newsletters may also be used to report progress on the CCEx development, describe the benefits of participation and encourage cost data sharing.

G.6 Roles and responsibilities

Aligned with the 4C Project Communications Plan, this section defines the roles and responsibilities related to CCEx communications activities and includes and expands upon the roles and responsibilities defined with the Description of Work (DoW).

All 4C Partners

All partners will:

- be responsible for communication activities;
- assist in the implementation of the CCEx Communications Plan as defined in this document by carrying-out communications tasks as directed by the Project Co-ordinator, Work Package (WP) or Task Lead;
- include relevant communication on organisational websites;
- ensure communications reflect the CCEx messages and benefits;
- include the 4C Project web address and contact details in external communications related to the project;
- use 4C appropriate templates for relevant CCEx-related communications;
- use their own organisations' contacts and established communication channels in support of the CCEx as appropriate;
- include the #CCEx hashtag when mentioning the project on Twitter;
- acknowledge EU funding through communications, as specified in the grant agreement.
- WP2 'Engagement' Lead

The Lead Partner for WP2 Engagement will:

- manage the undertaking of all external CCEx-related communication activities defined in this document;
- act as the central point of contact for all external communication activities defined in this document;
- delegate particular communications tasks to WP2 participants as required;
- monitor, update and add to the CCEx Communication Plan as required; and
- Delegate a designated alternate who will provide additional support and cover for communications management duties - To be decided in consultation with Sabine and the WP2 members.

G.7 Work Package Leads

To aid in the communication of information about the CCEx, Leads will:

- provide regular updates on the progress of relevant work packages at the scheduled project team meetings in order to enable external communications;
- ensure that all relevant deliverables include an accessible summary section that can be repurposed for communication purposes and similar 4C activities; and
- provide information and content on the work carried out within relevant work packages as required by WP2 Participants producing communications outputs.