Collaboration to Clarify the Cost of Curation

D2.3—Final Stakeholder Report

Deliverable Lead:	German National Library
Related Work package:	WP2—Engagement
Author(s):	Katarina Haage (DNB), Sarah Middleton (DPC), Neil Grindley (Jisc), Hervé L'Hours (UEssex) Paul Stokes (Jisc)
Dissemination level:	Public
Submission date:	12th February 2015
Project Acronym:	4C
Website:	http://4cproject.eu
Call:	FP7-ICT-2011-9
Project Number	600471
Instrument:	Coordination action (CA)—ERA-NET
Start date of Project:	01 Feb 2013
Duration:	24 months

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

Version History

Version	Date	Changed pages / reason	Modified by
0.01	14 July 2014	First draft	KH
0.02	9 Jan 2015	Completed draft	KH
0.03	16 Jan 2015	Commented and revised draft	SLM/NG
0.04	22 Jan 2015	Edited draft	KH
0.05	23 Jan 2015	Updated numbers	KH
0.06	30 Jan 2015	Format updated	KH
1.00	12 Feb 2015	Released version	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.

D2.3—Final Stakeholder Report by 4cproject.eu is licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

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.

Author(s):	Katarina Haage (DNB), Sarah Middleton (DPC), Neil Grindley (Jisc), Hervé L'Hours (UEssex), Paul Stokes (Jisc)

Table of Contents

Acknowledgements 3
Table of Contents 4
Figures 6
Tables 8
Executive Summary 9
1 Introduction 10
1.1 Definition of the Final Stakeholder Report 11
2 Brief summary of Baseline Study of Stakeholders and stakeholder initiatives (D2.1) 12
3 Summary of all stakeholder engagement activities 14
3.1 Communication and information exchange with EC-funded and other projects 15
3.1.1 Meetings, discussions and other types of engagement activities with external parties 16
3.1.2 Memoranda of understanding with other projects, organisations or initiatives 18
3.2 Stakeholder Focus Groups 19
3.2.1 Focus Group #1 19
3.2.2 Focus Group #2 20
3.2.3 Focus Group #3 20
3.2.4 Focus Group #4 20
3.2.5 Focus Group #5 20
3.2.6 Focus Group #6 21
3.3 Webinars 21
3.4 Advisory Board Meetings 22
3.5 First Review Meeting 23
3.6 Project Website 24
3.6.1 Interim Website (Months 1 – 6) 24
3.6.2 Developed Website 24
3.6.3 Blog posts 28
3.7 Social Media 29
3.8 Conferences and Events 31
3.9 Publications 31
3.10 Information Material 31
4 Conclusion 33
References 35
Appendices 36
Appendix A adidas meeting Report 37
Appendix B AV preserve meeting Report 41
Appendix C Bedern Group meeting Report 45
Appendix D Nestor working group on costs meeting Report 50
Appendix E Focus Group 1 Report 54
Appendix F Focus Group 2 Report 59
Appendix G Focus Group 3 Report 65
Appendix H Focus Group 4 Report 70

Appendix I Focus Group 5 Report 78
Appendix J Focus Group 6 Report 88
Appendix K Summary of Engagement Activities 93
Appendix L Printed Information and Materials 104
4C FLYER 104
CCEX FLYER 106
ROADMAP 107

Figures

Figure 1—Interim website stats 24
Figure 2—Interim website stats 24
Figure 3—4C final website 25
Figure 4—Google analytics results 26
Figure 5—Website bounce rate 26
Figure 6—Top performing web pages 27
Figure 7—Geographic spread of website visitors 27
Figure 8—Geographic spread of website visitors 28
Figure 9—4C twitter account 30
Figure 10—4C twitter account retweets 30
Figure 11—4C twitter account 'favorites' 31
Figure 12—4C project work packages 46
Figure 13—4C Flyer—English 104
Figure 14—4C Flyer—German 105
Figure 15—CCEx Flyer 106
Figure 16—Roadmap booklet 119
Figure 17—Actions for Curation Practitioners—English 120
Figure 18—Actions for Curation Researchers—English 121
Figure 19—Actions for Data Users—English 122
Figure 20—Actions for Managers—English 123
Figure 21—Actions for Member Organisations—English 124
Figure 22—Actions for Policy Makers—English 125
Figure 23—Actions for Solution Providers—English 126
Figure 24—Actions for Curation Practitioners—German 127
Figure 25—Actions for Curation Researchers—German 128
Figure 26—Actions for Data Users—German 129
Figure 27—Actions for Managers—German 130
Figure 28—Actions for Member Organisations—German 131
Figure 29—Actions for Policy Makers—German 132
Figure 30—Actions for Solution Providers—German 133
Figure 31—Actions for Curation Practitioners—French 134
Figure 32—Actions for Curation Researchers—French 135

Figure 33—Actions for Data Users—French 136
Figure 34—Actions for Managers—French 137
Figure 35—Actions for Member Organisations—French 138
Figure 36—Actions for Policy Makers—French 139
Figure 37—Actions for Solution Providers—English 140
Figure 38—Actions for Curation Practitioners—Portuguese 141
Figure 39—Actions for Curation Researchers—Portuguese 142
Figure 40—Actions for Data Users—Portuguese 143
Figure 41—Actions for Managers—Portuguese 144
Figure 42—Actions for Member Organisations—Portuguese 145
Figure 43—Actions for Policy Makers—Portuguese 146
Figure 44—Actions for Solution Providers—Portuguese 147
Figure 45—Actions for Curation Practitioners—Dutch 148
Figure 46—Actions for Curation Researchers—Dutch 149
Figure 47—Actions for Data Users—Dutch 150
Figure 48—Actions for Managers—Dutch 151
Figure 49—Actions for Member Organisations—Dutch 152
Figure 50—Actions for Policy Makers—Dutch 153
Figure 51—Actions for Solution Providers—Dutch 154

Tables

Table 1—Brief summary of activities 14
Table 2—Communication and information exchange with projects 16
Table 3—Engagement activities with external parties 17
Table 4—Webinars 22
Table 5—Social media activity 31
Table 6—Summary of Engagement Activities 103

Executive Summary

The main objective of the work package 2 “Engagement” has been to connect with a wide range of stakeholders from memory institutions through to data intensive research and industry, understand their needs and requirements and actively involve them in the process of working on the issue of curation costs. This also fits in the context of one of the main objectives of the 4C Project (4C) to ensure that where existing work is relevant, that stakeholders realise and understand how to employ those resources.

Throughout the whole project lifecycle 4C talked with a wide range of stakeholders to receive their opinions, statements and views. These we listened to, gathered together and implemented within the project work and its outcomes. It is this input from external parties that has, perhaps above all, made the project a ‘lively place to work’ and most importantly made its achievements relevant and valuable for the whole digital curation community.

This report begins with a summary of the foundation of all stakeholder engagement work, namely the Baseline Study of Stakeholders and Stakeholder Initiatives (D2.1)[1]. A chronological list and description of the engagement activities undertaken by the team within the last 24 months follows and illustrates the impact on the further work of 4C as well as the findings that were discussed in depth with the different stakeholder groups and interested parties. The work included focus groups, webinars, telephone interviews, face-to-face meetings, test runs, consultations and surveys to collect feedback.

We also took the opportunity to capitalise on further engagement opportunities—opportunities highlighted in this report as well. These opportunities included collaboration with other projects, initiatives and organisations that are active or interested in the issue of curation costs, and have fostered a better understanding of the issue amongst the digital curation preservation community.

A summary and variety of engagement activities is provided in Appendix K; the number of dissemination opportunities taken by the consortium during the 2 years is shows coverage of the project work. The interest generated in this period is an indicator of the potential future impact. All of this work has fed into and shaped one of the important outputs of the project, the 4C Roadmap.

Activities/channels included:

- Communication and information exchange with EC-funded and other projects and organisations
- Stakeholder focus groups and workshops
- Webinars
- Advisory Board Meetings
- Project Website
- Social Media
- Conferences and Events
- Publications

[1] D2.1 - Baseline Study of Stakeholders & Stakeholder Initiatives: http://4cproject.eu/d2-1-stakeholders

1 Introduction

"Collaboration is key" was an early assertion that captured the ethos of the 4C project; namely to enable two-way interactions between the project and its wide range of stakeholders.

This report ties together all stakeholder engagement and communication activities which were organised and performed during the lifetime of the 4C project. The engagement work goes back to the very early work of the Baseline Study of Stakeholders and Stakeholder Initiatives (D2.1) which was published in month 6 of the project and builds upon this foundation for all further engagement activities. The different approaches and formats of gathering input from external stakeholders as well as from the Advisory Board members and the whole 4C consortium resulted in a rich and varied potpourri of input relating to the tasks and issues that arose.

As part of the stakeholder engagement organised and implemented in Work Package 2, a series of focus groups and webinars were organised over the duration of the 4C project in order to share, disseminate and discuss 4C outcomes and recommendations. Participants at these events came from many stakeholder groups including: related or relevant projects; initiatives and coalitions in the area of digital curation; consumers of digital curation services; and service providers. In the presentations and discussions these participants were able to bring in their views, introduce their own approach and make 4C aware of what happens in their arenas.

The web statistics suggest 'followers' of the project are most interested in new resources to assist in their own digital curation projects. The demographic reports for twitter and the web also show that there was a wide audience and appetite for the 4C Project and the information it disseminated.

The results of the communications activities show a good level of interest and engagement in the project and its outcomes, and generally the project achieved its goals for communications targets set in the 4C Project Communication Plan (D2.5)[2].

Key DOW quotes

"Based on the stakeholder analysis (M1-6) 4-6 audiences will be targeted. These might include: data intensive industry; big data science; digital preservation solution/storage vendors; small and medium enterprises (SMEs); publishers; memory institutions; government agencies and research funders. Stakeholder engagement with different groups might require a flexible methodology, i.e. we may target our audiences individually or in groups, we can do structured telephone interviews, email surveys, discussion groups, feedback rounds… An extensible framework interview template will be crated to facilitate in-depth interviews / mini-consultations with selected representative stakeholders (starts in M4). Focus group meetings will be organized for each stakeholder group to understand their needs and requirements and to gain a better understanding of their views on the nature of cost, benefit, value, sustainability, etc. The focus groups will ideally be attached to a key event that is of relevance to the respective stakeholder group, e.g. iPRES 2013 in Lisbon, CeBit, International Conference on Electronic Publishing, etc. As far as possible input will be gathered from stakeholders according to the requirements set out in the Information Dependency Profile (T3.1). This will act as a checklist of useful categories of metrics that will facilitate effective quantitative information gathering. Useful intelligence (including qualitative data) relating to digital curation cost determinants will be passed to the Assessment group for analysis and synthesis. A synthesis, summary and evaluation of engagement activity will be reported at the close of the project in the form of a Stakeholder Report (D2.3)."

[2] D2.5 - Project Communication Plan, http://4cproject.eu/component/docman/doc_download/42-4c-project-communications-plan-1

1.1 Definition of the Final Stakeholder Report

This deliverable, the Final Stakeholder Report is defined in the description of work (DOW) as

D2.3 Final Stakeholder Report: A synthesis, summary and evaluation of engagement activities throughout the lifetime of the project and is principally focusing on the gathering of input into the project.[3]

Over and above the synthesis/summary in the core of the document we have also provided some of the raw data and outputs in the appendices.

[3] 4C - Description of Work, page 9

2 Brief summary of Baseline Study of Stakeholders and stakeholder initiatives (D2.1)

As a starting point for the Engagement work package in 4C, a baseline group of stakeholders was identified and an analysis of significant cost modelling and economics-related work in the field of digital curation was carried out. In addition, a questionnaire was sent to stakeholders in order to engage them in the project at an early stage and to better understand their current state of practice in assessing digital curation costs.

The deliverable included the baseline study of stakeholders and initiatives on the domain of digital curation costs; and includes the results of the following subtasks:

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

After reconsideration within the second quarter of the project the initially identified stakeholder groups were merged into a smaller set of categories as it seemed to group suitably the needs and channels of engagement among those participants. This assumption has been validated throughout the lifetime of the project, especially in workshops and focus groups. The revised and enduring set of stakeholder groups comprises:

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

Identifying these different stakeholder groups that were relevant to approach for the project work provided the important foundation for all engagement activities that were to follow.

As outlined in this report, we have been engaging with the stakeholder groups that were defined to target with the 4C project and its outputs. Since the variety of the stakeholder groups covered all kinds of backgrounds in digital curation, differences between the several groups could be (and were) observed. Although the overall impression of or reaction from stakeholder groups was multi-facetted it was also uniformly encouraging and giving.

The culture (memory institutions and content holders) and education (universities, cost model experts) groups were very proactive in approaching the project themselves which made their interaction between the project, its outputs and the stakeholder group fruitful and diverse. Science (research funders, big data science) showed sustainable interest in the topic and provided great input into the work with their knowledge and their views on the topic. The lessons learnt in interaction with the commercial side of digital curation (vendors, publishers, solution providers) helped to define and bring forward the sustainability ideas that we have verified with these stakeholders.

In conclusion one can say that, although different approaches were required, all stakeholders had sufficient common ground to form part of a community. We are optimistic that this community can be

fostered using mechanisms such as the Curation Costs Exchange (CCEx)[4] and that future initiatives will be able to maintain and grow and the network established by 4C.

[4] http://www.curationexchange.org

3 Summary of all stakeholder engagement activities

This part of the Final Stakeholder Report summarises all engagement and communication activities undertaken by partners of the 4C project, from month 1 to 24, February 2013 to January 2015.

It follows the structure of intended activities laid out in the Project Communications Plan. For the purposes of this report, the engagement activities may be defined as those made with external parties and also communication activities with the Advisory Board and the European Commission.

Brief Summary of Activities:

Communication and information exchange with EC-funded and other projects and organisations	Engagement with seventeen EU Projects and/or several other organisations.
Stakeholder focus groups and workshops	Engagement with stakeholders on ten occasions, including six Focus Groups, three Workshops, one Conference.
Webinars	Two CCEx specific and two Roadmap specific Webinars held; 4C project presentation in several SERI-COSA webinars by Neil Grindley.
Advisory Board Meetings	Three Advisory Board Meetings held.
Project Website	Eighty six posts to the 4C Project Website, including 13 deliverables, 15 community resources, 19 news items, 36 blogs and 3 others (Press, Engagement, 4C Focus Group Game)[5].
Social Media	Six hundred and forty five tweets[6].
Conferences and Events	4C representation at forty five conferences and events plus ten own events.
Publications	Twenty four news articles published.

Table 1—Brief summary of activities

[5] Latest figures from21st January 2015. Web site updates and interactions have continued.

[6] Latest figures from21st January 2015. Project related tweets have continued

3.1 Communication and information exchange with EC-funded and other projects

The 4C project engaged in communication and information exchange with the following projects and organisations:

Name	Occasion	Date
APARSEN http://www.alliancepermanentaccess.org/index.php/aparsen/	APARSEN Webinar "Sustainability and Cost Models for Digital Preservation"	13th June 2013
	Contribution of APARSEN Input into 4C iPres Workshop	6th September 2013
SCAPE http://www.scape-project.eu/	Issue 5 of SCAPE Newsletter	26th June 2013
	Presentation of 4C Project in SCAPE and OPF Seminar	2nd April 2014
TIMBUS http://timbusproject.net/	Volume 2 Issue 1 of TIMBUS Times Newsletter	28th June 2013
Koninklijke Bibliotheek http://www.kb.nl/en	Meeting between 4C and KB	7th – 8th October 2013
ASIS&ST PASIG http://www.preservationandarchivingsig.org/	PASIG Webinar "Implementing Sustainable Digital Preservation"	22nd October 2013
EUDAT http://www.eudat.eu/	Presentation of 4C Project at EUDAT 2nd Conference	28th – 30th October 2013
ENSURE http://ensure-fp7-plone.fe.up.pt/site/	Meeting between 4C and ENSURE Project	4th November 2013
CERN http://home.web.cern.ch/	Meeting between 4C and CERN; CERN contribution in several 4C event (1st Focus Group and Workshop at iPres, 4C Conference)	7th - 8th November 2013
MiLoS http://www.eurekanetwork.org/project/-/id/7360#	Presentation of 4C to the MiLoS Project Consortium	28th November 2013
	Presentation of 4C and CCEx Mock-ups to MiLoS Project members and Cinevation staff	24th February 2014
nestor http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/home_node.html	Presentation of 4C Project and CCEx to the nestor working group of costs	25th March 2014

Name	Occasion	Date
NCDD http://www.ncdd.nl/en/	Contribution to 4C CCEx Focus Group at ARCHIVING 2014; and part of CCEx sustainability team	13th May 2014
Inspire http://www.inspirefp7.eu/	4C Conference was included in the e-Newsletter	July 2014
Presto4U https://www.prestocentre.org/4u	4C Conference was included in the 'Presto4U' Friday Rewind	11th July 2014
UNESCO http://www.unesco.nl/sites/default/files/uploads/Comm_Info/digital_roadmap_-_report.pdf	Roadmap Meeting	5th – 6th December 2014

Table 2—Communication and information exchange with projects

With all these projects, initiatives and organisations a strong relationship and cooperation base was built and will most likely remain post-project as also stated further down in chapter 3.1.2 Memoranda of understanding with other projects organisations or initiatives.

3.1.1 Meetings, discussions and other types of engagement activities with external parties

Throughout the whole lifetime of the project partners continually engaged with external parties and colleagues through different kinds of types, channels and opportunities such as email, phone calls, conference calls, face to face discussions, conversations at conferences, workshops, webinars and other community events as well as participating the overall community discourse on digital curation in general and the costs of curation in particular. The following chart shows examples of other occasions of engagement and gives an impression on how multi-facetted the project outreach has been and how committed and dedicated the project team was feeding into the work of the engagement work package:

Occasion/external party	Topic	Date
ADA Summer School	Full day Workshop/presentation of the 4C project in general and the CCEx and ESRM in particular	4th July 2014
adidas	Follow up engagement on CCEx and how digital preservation is managed in this area of industry[7]	7th February 2014

[7] For details see minutes of the meeting in Appendix A

Occasion/external party	Topic	Date
ANADP2 Barcelona	How can we employ the resources we have available to us most economically to achieve our digital preservation objectives?[8]	18th – 20th November 2013
Archivematica	One to One Engagement Meeting for information exchange on community building and approaching costs of curation	30th September 2013
AV preserve	Further engagement after introduction both to CCEx (4C) and the 'Cost of Inaction' Tool (AV preserve)[9]	18th March 2014
Bedern Group	Introduction to 4C project in general and CCEx in particular[10]	22nd May 2014
Deutsche Bahn Archives	Follow up engagement on CCEx and how digital preservation is managed in this area of industry	March 2014
nestor working group on costs	The main purpose of this meeting was an exchange of experiences. The nestor working group has also tried to collect cost data and they have done some general cost model research.[11]	25th March 2014
Research Information Magazine	Interview with Neil Grindley and Dr Rebecca Pool about digital preservation, activities and industry developments	December 2014

Table 3—Engagement activities with external parties

Future activities that were already planned before the end of the project, but which will take place after 31st January 2015 include:

- IDCC 2015, 9th - 12th February 2015: Jisc and DCC
- Sussex Research Hive Seminar, 26th February 2014: Neil Grindley (Jisc)
- Presentation of the CCEx at Jisc Digifest 2015, 9th – 10th March 2015, [12]
- NCDD plan to take the 4C results and CCEx as starting point for their new project on Cost Management starting in April 2015

[8] See Blog post: ANADP II Action Session - 4C Case Studies and Quantitative Data Session' by Neil Grindley and Raivo Ruusalepp— http://4cproject.eu/news-and-comment/4c-blog/93-anadp-ii-action-session-4c-case-studies-and-quantitative-data-session-by-neil-grindley-and-raivo-ruusalepp

[9] For details see minutes of the meeting in Appendix B

[10] For details see minutes of the meeting in Appendix C

[11] For details see minutes of the meeting in Appendix D

[12] http://www.jisc.ac.uk/events/jisc-digital-festival-2015-09-mar-2015

- Cooperation and further exchange with vendors (Preservica, Arkivum, ExLibris) relating to the CCEx and Roadmap
- Approaching the European Data Science Academy (EDSA) to make CCEx part of their training offer, http://edsa-project.eu/
- Approaching the Digital Democracy Commission (http://www.parliament.uk/business/commons/the-speaker/speakers-commission-on-digital-democracy/) to promote the 4C work

3.1.2 Memoranda of understanding with other projects, organisations or initiatives

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

SCAPE

4C established an informal agreement with the SCAPE project to support and cooperate in terms of distributing and sharing newsletters and to support one another in outreach and engagement activities.

Presto4U

A similar arrangement was established with Presto4U. 4C set up an informal agreement to support and cooperate in terms of distributing and sharing newsletters and to support one another in outreach and engagement activities.

NCDD

NCDD has agreed to use the Curation Costs Exchange (CCEx) in their current project on digital preservation about managing costs, which starts in 2015. Part of the project would be to submit and compare Dutch costs by means of the CCEx.

nestor

Nestor, as network for digital preservation initiatives in Germany (and 4C partner on through the DNB), has agreed to maintain the CCEx together with the DPC as hosting organisation of the website. This maintenance includes taking care of the content on the website.

DPC

DPC, as network for digital preservation initiatives in the UK and 4C partner, has agreed to host and assist with the maintenance of the Curation Costs Exchange website. This maintenance includes taking care of the content, day to day operation and management of any technical issues relating to the content management system.

KEEP Solutions

KEEP Solutions, as solution provider and developer in digital preservation and 4C partner, has agreed to provide technical support the custom development elements of the CCEx that might arise in the future.

3.2 Stakeholder Focus Groups

The DoW stipulated four Focus Groups were to be held within month 8-17. To satisfy demand we were able to hold two more focus groups before the project end in February 2015. These additional focus groups were used to gather more input and valuable opinions on results and outputs of the project work, in particular the CCEx and the Roadmap.

Early engagement was undertaken with stakeholder groups through an Initial Consultation between the 17th May and the 21st June 2013. The consultation sought an indication of willingness to become involved in further knowledge exchange throughout the project duration. As described in the deliverable D2.1[13] (M6) the initial consultation returned good results. Using the information generated by this initial consultation, invitations were sent to those stakeholders who indicated a desire to be involved in the project, inviting them to join the focus groups. The reports of the focus groups are provided in the Appendices.

Some of the focus groups were attached to satellite events or combined with national or international events to reach out to a multi-facetted audience.

3.2.1 Focus Group #1

The first focus group was held at iPres 2013 on the 6th September 2013 as a half-day event. Invitation resulted in 7 attendees, covering representation from all stakeholder groups. From the 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”.

In preparation of the focus group meeting, the participants had been asked to rank the list of Indirect Economic Determinants (IEDs, now Indirect Cost Drivers). The 4C concept of “Indirect Economic Determinants” was presented in the meeting by Raivo Ruusalepp (NLE) and as an outcome from the Focus Group the combined ranking led to the following top 5:

1. Risk
2. Trustworthiness
3. Benefits
4. Sustainability
5. Efficiency, 6. Value.

In a break-out session, the participants were roughly sorted into a “memory institutions group” and a “non-memory institutions group”. Both groups discussed the IEDs separately. Their input was used in the development of the deliverable D4.1—A prioritised assessment of the indirect economic determinants of digital curation.[14]

[13] 4C Project, deliverable D2.1 Baseline Study of Stakeholders & Stakeholder Initiatives: http://4cproject.eu/d2-1-stakeholders

[14] Appendix E—Focus Group 1 Report

3.2.2 Focus Group #2

This focus group was tailored for industry stakeholders and was organised in Frankfurt for the 11th December 2013 and in London for 12th December 2013. Initially, it was planned to have the same scope for the Frankfurt meeting and the London meeting. Due to the fact that most of the participants for the Frankfurt meeting called in sick the day before the meeting was scheduled for, it was decided in consultation with WP2 to cancel the Frankfurt Focus Group. Sabine Schrimpf and Katarina Haage from DNB contacted some of the planned external attendees afterwards by phone to gather input relating to the planned outcomes for the focus group meeting and reported it to the consortium. The London focus group took place as planned with ten participants on the 12th December 2013. After an overview to the 4C project, its approaches, goals and priorities the participants were introduced to the Economic Sustainability Reference Model (ESRM) and a self-assessment questionnaire based on the model. This was followed by an ESRM Exercise and the participants went into discussion on relevance and potential of the model and recommendations from the discussion were captured[15].

3.2.3 Focus Group #3

The third focus group was a satellite event attached to the DPHEP workshop at CERN in Geneva on the 13th -14th January 2014. It was held on the second day of the workshop as a breakout session where the CCEx was presented to the audience in order to gather feedback on the idea and the tool. The feedback for the 4C project and its cost exchange was very positive. It was remarked in the discussion that cost information is available in different organisations from different areas, but so far too little effort was made to collect and consolidate this information. The need and the positive effects of having cost comparable information was agree by all[16].

3.2.4 Focus Group #4

This event was attached to the ARCHIVING 2014 in Berlin and took place on the 13th May 2014. Four 4C members and by nine external participants took part. Next to presentations about the CCEx and the Cost Concept Model there were two guest speakers from the NCDD to give an introduction to their project on digital preservation and underline the relation of this work to the 4C project. At the end of this meeting it was agreed that further engagement between 4C and the NCDD and Presto4U would take place[17].

3.2.5 Focus Group #5

This event took place on November 3rd 2014 in London at The Wesley Hotel and focused on getting input on the CCEx. Its main objectives were to find out what users think of the CCEx; to identify what users want/expect from the tool; to find out what the common problems are when using it; to collect recommendations for improvement; and to find out if they would have concerns about sharing their cost data (and if so, why); what could be done to mitigate these concerns? In attendance were four 4C members and seven external participants with Research Data Management background and an interest in the topic[18].

[15] Appendix F—Focus Group 2 Report

[16] Appendix G—Focus Group 3 Report

[17] Appendix H—Focus Group 4 Report

[18] Appendix I—Focus Group 5 Report

3.2.6 Focus Group #6

This event on the 15th January 2015 in London focused on getting input on the CCEx and also the 4C Roadmap from the digital preservation vendor's and solution provider's point of view. It was set up close to the end of the project to present the CCEx as close to its final version as possible, whilst at the same time yet leaving enough time to make minor changes if desired. It was attended by five 4C members and five external participants. After a welcome by project coordinator Neil Grindley (NG) and a brief introduction round, NG gave an overview to the 4C project and the project results to date were presented. Following this Luis Faria (LF) gave a live introduction to the Curation Costs Exchange (CCEx)[19] website, focusing on the Cost Comparison Tool (CCT). The discussion that followed gave us a rich view of the perspectives as a vendor; about take-up, comparison, value, benchmarks, maturity. Also future possibilities regarding the CCEx and Roadmap were discussed and agreements have been made on future information exchange re the CCEx, CCT and the Roadmap[20].

3.3 Webinars

Online webinars proved to be a very useful and popular alternative to face to face meetings. They were a fertile ground for discussions and information exchange with a wide range of stakeholders groups.

The 4 webinars attracted audiences not only from Europe but also from the US. Topics varied for each webinar. Each was tailored to gain input on a specific topic as well as external views and opinions on project results and outputs (mainly relating to the CCEx and the 4C Roadmap, both important outputs for the project). If necessary, the participants were provided with information—such as a draft documents or access to a draft website or tool—beforehand to allow them to prepare.

The structure of the webinar itself was kept simple and followed the same format for each event. The start and welcome would be made by one of the 4C colleagues, followed by a short introduction round by all attendees. Then a brief introduction to the 4C project in general would be given to inform the participants and to ensure common starting conditions. After this the specific topic or product would be presented by 4C colleagues. This presentation would then be followed by the discussion of either prepared questions or an open Q&A sessions (or both) in order to collect impressions and opinions on the presented topic.

These webinars also proved to be an excellent way of generating a better insight into and understanding of existing gaps in the area of costing digital preservation, identifying the areas where clarification is needed. The webinars also helped identify serious gaps in current digital preservation environments and practices. The ideas communicated by the webinar participants fed back into the further iterations of project outputs and into the reflections of the project as a whole.

Project coordinator Neil Grindley additionally gave several webinars and presented the 4C project in the course of a SERI-COSA training series in the US - http://www.statearchivists.org/seri/.

[19] http://www.curationexchange.org

[20] Appendix J—Focus Group 6 Report

Webinar	Audience	When	Topic
1.	AVpreserve	18th March 2014	CCEx and Cost of Inaction Tool
2.	Bedern Group	22nd May 2014	Introduction to 4C and CCEx
3.	All stakeholders	10th June 2014	Roadmap
4.	All stakeholders	17th June 2014	Roadmap
5.	SERI-COSA	2013-2014	4C Introduction

Table 4—Webinars

3.4 Advisory Board Meetings

Within the lifetime of the project three Advisory Board (AB) meetings were planned and held.

The first AB meeting was on the 11th June 2013 at the Jisc Office in London, UK. In attendance were five 4C representatives and seven AB members. Highlights from this meeting were: the context, background, introduction to the project and the purpose of the AB; the reaction from AB members about high-level aims; the summary of work packages and their objectives; and the discussion of the main challenges the project faces, such as: terminology, models and specification, gathering costs data, and the Curation Costs Exchange.

The second AB meeting was held on 22nd January 2014 in The Hague, following the third face to face project meeting. The objectives of this meeting were to review progress made on the project; to address any problems or challenges that had arisen; to reach a shared understanding about the work that needs doing; to discuss and agree changes that we may need to introduce into the work plan; to prepare and review our readiness for the project review meeting in March 2014; and to effect introductions between the Project Team and the Advisory Board members. In attendance were eight AB members and six 4C representatives.

This meeting was the first where a joint meeting between the AB and the larger Project Board (PB) took place. The AB had at the previous board meeting suggested that a joint working session would be useful. This proved to be the case and the joint session was repeated at the following AB meeting (which was shifted to coincide with a planned PB).

The third and last AB meeting took place on the 30th June 2014, before the fourth face to face project meeting. The purposes for the AB meeting were as follows: to review progress made on the project; to address any problems or challenges that had arisen; to reach a shared understanding about the work that needs doing; to discuss and agree changes that we may need to introduce into the work plan; to review the draft deliverables due in the latter part of the project, in particular the Roadmap; and to gain data from the Advisory Board for the current data gathering exercises and in the process to review/validate the current tools, in particular the CCEx.

In attendance were six 4C representatives and seven AB members. Some highlights resulting from the meeting were: the announcement of “excellent progress” rating awarded to the 4C project by the European Commission reviewers in the first project review in March 2014; the Curation Costs Exchange, the introduction of the Core Concept Model;the introduction of the Roadmap; and the announcement of the 4C Final Conference on 17th – 18th November 2014. Particular discussion points between the AB and

the project team included the CCEx, the trust and the cost of auditing digital archives, and the 4C Roadmap.

3.5 First Review Meeting

The first Review Meeting took place on 27th March 2014 at the German National Library in Frankfurt. I n attendance were:

- Alex Thirifays, DNA
- Frédéric Blin
- Katarina Haage, DNB
- Manuela Speiser
- Maurice Van den Dobbelsteen
- Neil Grindley, Jisc
- Paul Stokes, Jisc
- Raivo Ruusalepp, NLE
- Sabine Schrimpf, DNB
- Sarah Norris, DPC York
- Ulla Bøgvad Kejser, KB-DK
- William Kilbride, DPC Glasgow

The purpose of the meeting was to review progress made on the project; and to address any issues raised by the project reviewers.

After the private pre meeting between the Project Officer and the Reviewers William Kilbride moderated the introductions and outline of the day's proceedings. A brief overview of the deliverables and milestones achieved to date were represented by the project coordinators Neil Grindley and Paul Stokes followed by detailed Work Package reports from the work packages leads. After a private meeting, the Project Officer and Reviewers had the opportunity to provide verbal feedback and announce the result of the review meeting; the project was marked with "excellent progress".

Additional stakeholder work has been arranged as a result of the review , for example incorporation of vendors, translation of certain communication material, such as the 4C flyer and the Roadmap Postcards[21]. The suggestion to approach the vendor and solution provider stakeholder group to gather views and input on the economic and monetary issues in costing curation has been taken up and successfully fulfilled as outlined in this report. Communication material in English, Dutch, French, German and Portuguese was produced.

[21] See Appendix L—Printed Information and Materials

3.6 Project Website

3.6.1 Interim Website (Months 1 – 6)

In the first few weeks an interim website was produced using WordPress to provide an early web presence. This was used for the first six months of the project. WordPress Stats provides the following overview of activity on the temporary WordPress site:

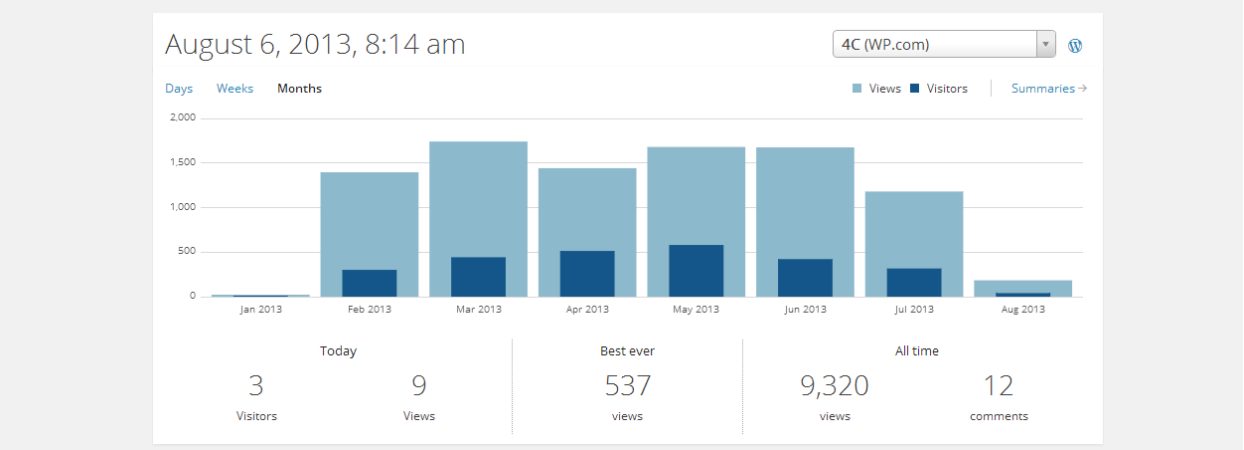
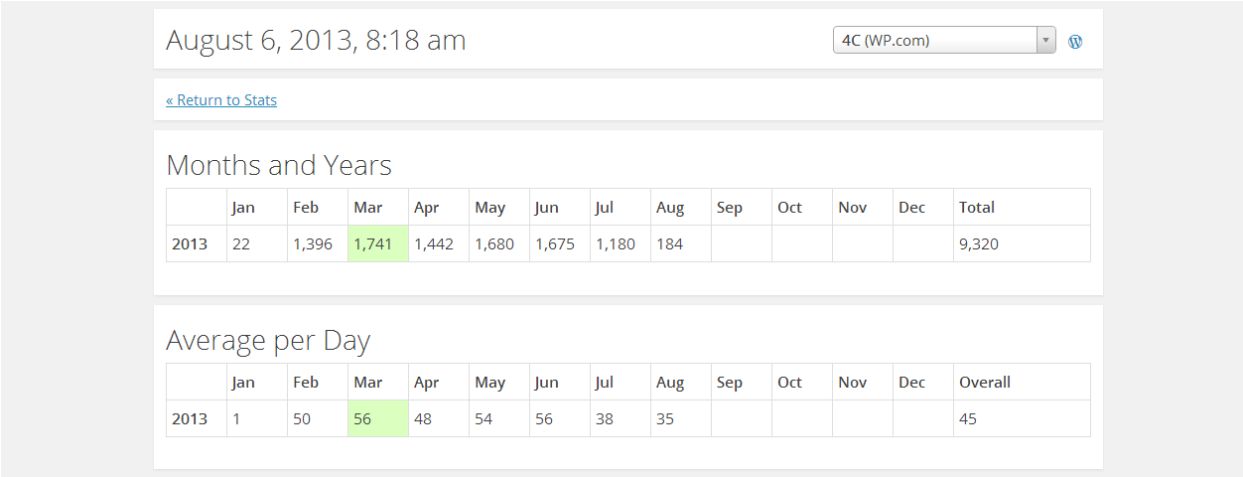


Figure 1—Interim website stats

Month by month activity is demonstrated in the charts above and below, showing March as the month with the most traffic through the site.

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



	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

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

Figure 2—Interim website stats

3.6.2 Developed Website

The developed 4C project website was completed in line with the deliverable D2.7 deadline in month 6 and published on the 31st July 2013. The sitemap of the site was developed to include suggestions and features made by the AB and project partners.

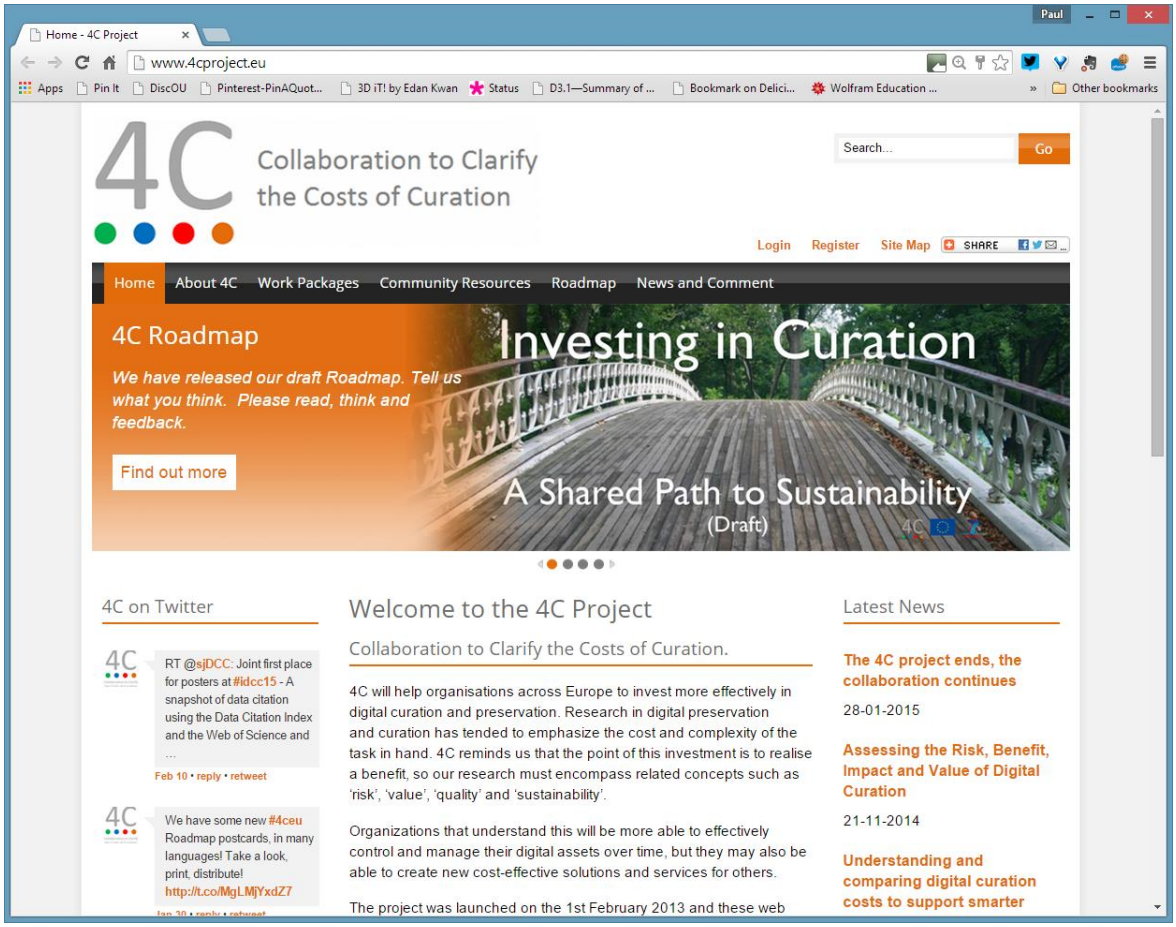


Figure 3—4C final website

The web site grew throughout the project's lifetime and came to include a great deal more content as each team fulfilled their objectives and tasks and published their results. All deliverables and many milestone reports were published in the public "Community Resources" area on the 4C Website. The number of blog posts from project members and guest bloggers also rose steadily. By the end of the project we had in excess of 38 blog posts. A summary of these activities can be found in Appendix K.

Google analytics were used to provide an overview of the new website activity between months 7 and 24 of the project. In this period, the website saw 9,859 visitors (16,239 sessions, 47,488 page views), of which 60.7 % were new visitors[22]. A breakdown of visitor acquisition and behaviour is shown below.

[22] Latest numbers from 22 January 2015

	Acquisition			Behaviour		
	Sessions ↓	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration
	16,239	60.68%	9,854	50.45%	2.92	00:02:52
1 Direct	7,869			43.83%		
2 Organic Search	4,066			57.92%		
3 Referral	3,003			53.91%		
4 Social	1,288			58.93%		
5 Email	13			84.62%		

Figure 4—Google analytics results

Bounce rate is a measure of the effectiveness of a website in encouraging visitors to continue with their visit. It 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. 50 per cent bounce rate is average; anything in excess of 80 per cent represents a major problem. The site's average bounce rate for the period from month 7-24 is 50, 45 %[23], which is within the average range.

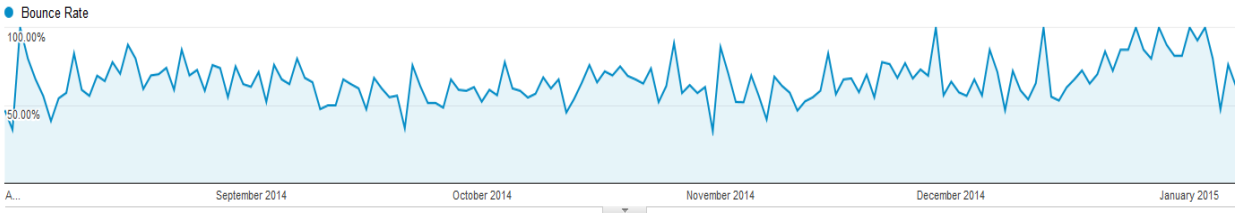


Figure 5—Website bounce rate

[23] Latest numbers from 22 January 2015

The top ten best performing pages for the period from month 7-24 were[24]:

	Page	Pageviews	% Pageviews
1.	/	8,084	17.02%
2.	/community-resources	2,394	5.04%
3.	/about-us	2,129	4.48%
4.	/community-resources/outputs-and-deliverables	1,664	3.50%
5.	/work-packages	1,374	2.89%
6.	/d5-1-draft-roadmap	1,155	2.43%
7.	/news-and-comment	1,057	2.23%
8.	/news-and-comment/4c-blog	929	1.96%
9.	/community-resources/investing-in-opportunity-conference	705	1.48%
10.	/community-resources/outputs-and-deliverables/d3-1-evaluation-of-cost-models-and-needs-gaps-analysis-ms12-draft	636	1.34%

Figure 6—Top performing web pages

The average duration of visits is 02:52minutes, with the longest time spent on pages containing on Community Resources.

The majority of visitors come from Germany and UK, with the top ten visitor origins shown below.

Half of the top ten countries are those represented by the 4C Project team (Germany, UK, Denmark, Portugal, and Netherlands).

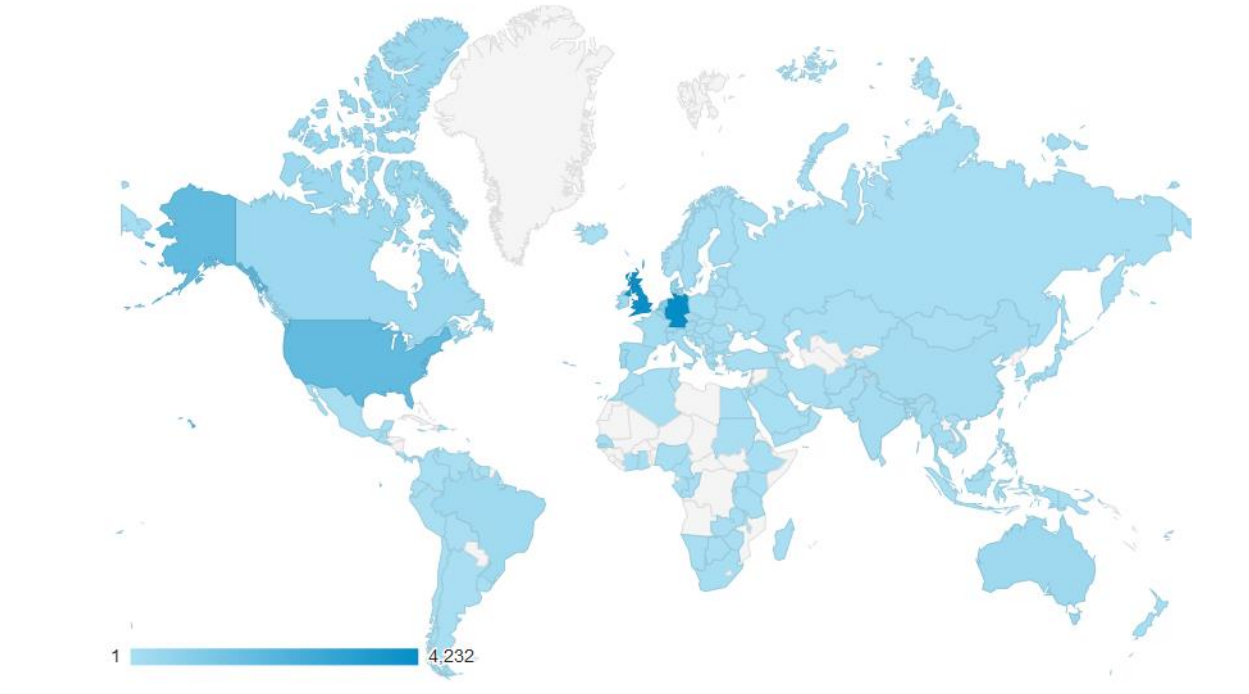


Figure 7—Geographic spread of website visitors

[24] Latest numbers from 22 January 2015

Visitors independent of the 4C team were from the US, Canada, Australia, Brazil, and Belgium; order and behaviour as shown below:

Country	Acquisition			Behaviour		
	Sessions ↓	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration
	16,239 % of Total: 100.00% (16,239)	60.69% Avg for View: 60.68% (0.02%)	9,856 % of Total: 100.02% (9,854)	50.45% Avg for View: 50.45% (0.00%)	2.92 Avg for View: 2.92 (0.00%)	00:02:52 Avg for View: 00:02:52 (0.00%)
1. Germany	4,232 (26.06%)	71.95%	3,045 (30.89%)	25.59%	2.82	00:02:15
2. United Kingdom	4,058 (24.99%)	42.19%	1,712 (17.37%)	52.34%	3.76	00:04:38
3. United States	1,809 (11.14%)	70.81%	1,281 (13.00%)	67.77%	2.40	00:02:04
4. Netherlands	802 (4.94%)	48.63%	390 (3.96%)	47.01%	3.41	00:03:30
5. Denmark	623 (3.84%)	32.91%	205 (2.08%)	40.45%	3.45	00:03:36
6. Portugal	368 (2.27%)	44.29%	163 (1.65%)	54.89%	2.68	00:02:22
7. Canada	349 (2.15%)	74.79%	261 (2.65%)	65.33%	2.61	00:02:38
8. Australia	302 (1.86%)	69.87%	211 (2.14%)	71.19%	2.01	00:01:39
9. Brazil	264 (1.63%)	94.70%	250 (2.54%)	93.56%	1.30	00:00:32
10. Belgium	241 (1.48%)	63.90%	154 (1.56%)	65.98%	2.59	00:02:09

Figure 8—Geographic spread of website visitors

3.6.3 Blog posts

Blogs by project team members and guests have continued to be posted on a regular basis over the two years. The following 36 blogs were posted in months 1 – 24, including 4 guest posts:

- 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, 7th 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, 2th October 2013

- ‘No such thing as free digital preservation’ by Jan Dalsten Sørensen, 08th November 2013
- 'Please help us draw a map!' by Alex Thirifays, 17th December 2013
- 'Lessons from the Half Way Point' by Neil Grindley, 3rd February 2014
- 'Whistle for the start of the second half' by Katarina Haage, 14th February 2014
- 'Uncertainty: the final frontier' by José Borbinha, 27th February 2014
- 'Crunching Numbers and Comparing Costs' by Matthew Addis, 17th March 2014
- 'Models and Methods and Tools, Oh My!' by Hervé L'Hours, 7th April 2014
- 'Valuable feedback from Berlin' by Alex Thirifays, Katarina Haage and Hervé L'Hours, 29th May 2014
- 'Digital curation buys us options—invest in opportunity' by Sarah Middleton, 26th June 2014
- 'Ready for take-off' by Alex Thirifays and Sarah Middleton, 18th July 2014
- '3 (more) reasons to head to iPRES 2014' by Sarah Middleton, 1st October 2014
- 'Collaborating on sustainable services for curation' by Matthew Addis, 27th October 2014
- 'Trust, certification, sustainability and framework agreements' by Matthew Addis, 11th November 2014
- 'The Curation Costs Exchange unveiled and challenged' by Alex Thirifays, 5th December 2014
- ‘Zettabyting off more than we can chew’ by Paul Stokes, 17th January 2015
- 'How Time Flies?!' by project coordinator Neil Grindley, 23rd January 2015
- ‘Shaping the Curation Costs Exchange: sharing your feedback' by Magdalena Getler, 28th January 2015
- 'Why Cost Models are Risky' by Sean Barker, 4C Project Advisory Board, 30th January 2015

Links to all blog posts are provided in the Appendix K—Summary of Engagement Activities.

3.7 Social Media

The 4C project established a twitter account; ‘4c_project’ and a hashtag; ‘#4ceu’. Within the 24 months of the project, the 4C Project has made 647 tweets (on average, just under tweet a day) and gained 368 followers in total[25].

The project also used the hashtag #IIO2014 for the 4C conference. This tag had over 1,300 mentions in the period immediately before, during and after the conference.

[25] Latest numbers from 22 January 2015

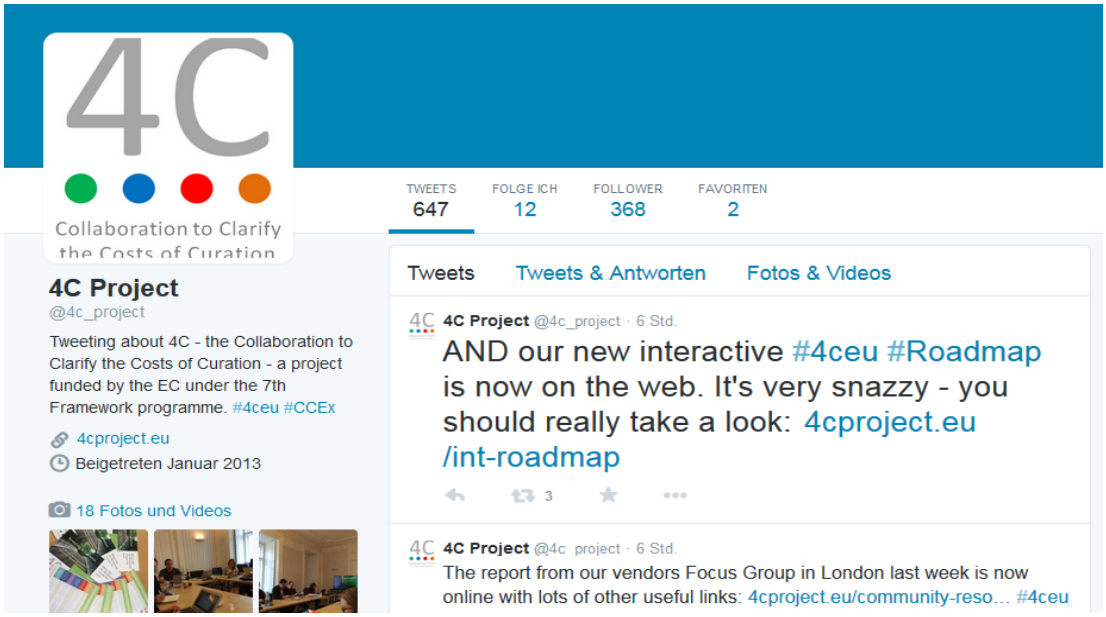


Figure 9—4C twitter account

Month by month retweet activity is shown in the chart below, with the spikes in activity corresponding to a number of tweets being retweeted from conferences and other events:

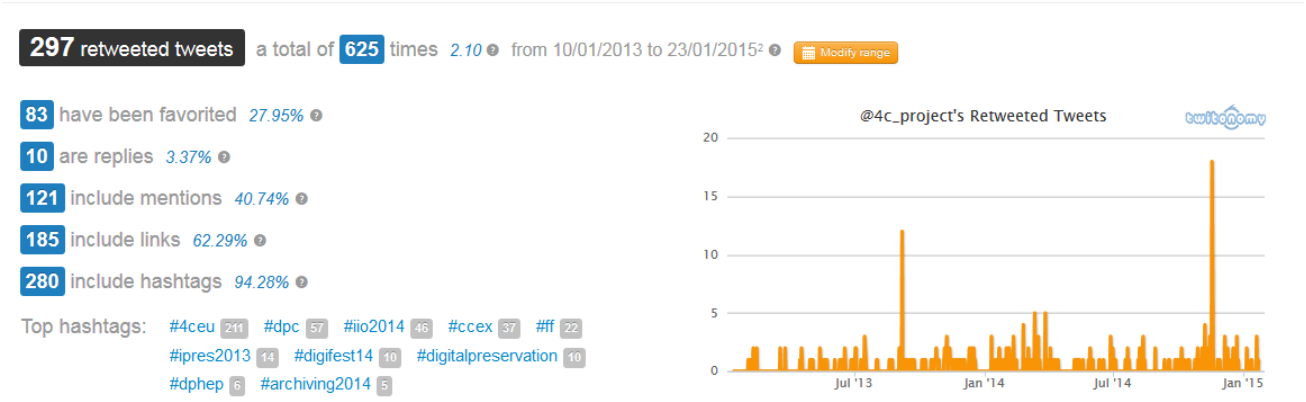


Figure 10—4C twitter account retweets

Month by month 'favorite' activity is shown in the chart below, with the spikes in activity corresponding to a number of tweets about conferences and events being 'favorited':

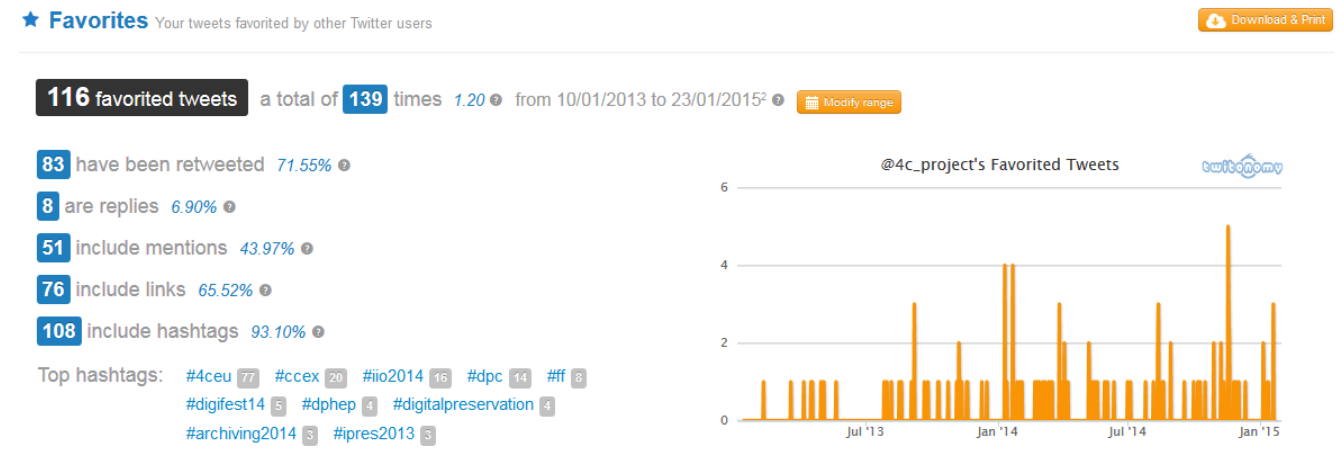


Figure 11—4C twitter account 'favorites'

This social media presence has led to dissemination of project activity through the main project twitter account. The breakdown of hashtag uses, retweets, mentions and favourites is a s follows[26]:

Activity	Hashtag use	Mention	Retweet	Favourite
Number	388	368	625	139

Table 5—Social media activity

3.8 Conferences and Events

The 4C project has been represented at 53 events, including 45 conferences and 8 other events over the 2 years of the project. These events have provided opportunties to disseminate the project in general and the different results and outputs in particular. A list of all events, including the project's own events can be found in Appendix K.

3.9 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. A summary of all 24 publications with links to the sites is provided in Appendix K.

3.10 Information Material

Throughout the lifetime of the project some information material has been printed and disseminated. The materials include:

- 4C Flyer (English and German)
- CCEx Flyer (English)
- Roadmap Brochure (English)
- Roadmap Actiomns Postcards (English, German, Dutch, French)

[26] Latest numbers on 23 January 2015

A collection of these can be found at the end of this report in Appendix L—Printed Information and Materials.

4 Conclusion

The 4C project, classified by the European Commission as a 'Coordination Action', was 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 work, programmes and policies. The implications of this for 4C were twofold. Firstly, it was a relatively small and relatively short-lived project that made the most of existing research and added to it, allowing partners to share and compare know-how rather than inventing new ways to consider the problem. Secondly, 4C was by its nature an outward-looking project that sought to engage a large and diverse group of stakeholders.

The topics covered and presented through the engagement activities, in the focus groups and other events carried out within the realm of WP2 provide a comprehensive overview of the views, opinions and state of the art when it comes to costing digital curation. Thanks to the active participation of the many external stakeholders, these engagement activities provided an excellent source of knowledge and inspiration enabling a better understanding of the current concerns, risks, needs and requirements and also the gaps that need to be tackled within the next years.

One of the key engagement activities was to collect feedback for the Roadmap document that was to be produced between M6-24 through the organisation of a period of public consultation on the draft version of the Roadmap in M19-21. This was to *"encourage the broadest possible endorsement and community buy-in for the roadmap, a period of public consultation will occur using an early draft version of the text and making it available for comment as an online annotatable document. All project partners, particularly where they are representative of particular sectors (e.g. SMEs, national libraries, funders, research organisations etc.) were to feed their perspectives into the report. Active consultation started on a draft version of the report in M17; the task lead Jisc and engagement group partners started to raise awareness of its availability and elicit input. Affiliate partners, lead stakeholders and the Advisory Board members were the first invited to comment on and contribute to the report*."[27] The wider community was given the chance to contribute via an online feedback feature, via email and though face to face feedback at the Roadmap Workshop, the Roadmap Webinars and any other dissemination opportunity. The results from these interactions between stakeholders and the project team also fed into part of the 4C Roadmap that is reported upoin elsewhere.

The fulfilment of the overall objectives and the application of an 'open and social' communications model within the engagement work ultimately facilitated the achievement of the main 4C project objectives by engaging users in sustainable dialogue throughout the lifetime of the project.

Looking at the numbers of engagement activities shown in this report one can say we have done a lot to raise awareness, get involved and communicate with our stakeholders. We feel—now we have reached the end of the funded portion of the project—that there is still more that we could do. We believe that we could go on for another twelve months doing what we were doing—engagement and dissemination. We have learnt a lot from our stakeholders and received great feedback. Although the project has come to an end, the creation of the Curation Costs Exchange and the sustainability plans that have been put into action should allow us to continue; building the community and providing the partners with the

[27] See Task description 5.1 in DoW, page 18

opportunity to keep on with the good work, keep in touch and help people understand better the costs in digital curation.

References

4C Project, Description of Work (DoW)

D2.1—Baseline Study of Stakeholders & Stakeholder Initiatives: http://4cproject.eu/d2-1-stakeholders

D2.4—Final Report on Outreach Events

D2.5—Project Communication Plan, http://4cproject.eu/component/docman/doc_download/42-4c-project-communications-plan-1

D2.6—Report on Communications Activities

D2.8—Curation Costs Exchange

D5.2—Final Roadmap Report

Appendices

Appendix A adidas meeting Report 37
Appendix B AV preserve meeting Report 41
Appendix C Bedern Group meeting Report 45
Appendix D Nestor working group on costs meeting Report 50
Appendix E Focus Group 1 Report 54
Appendix F Focus Group 2 Report 59
Appendix G Focus Group 3 Report 65
Appendix H Focus Group 4 Report 70
Appendix I Focus Group 5 Report 78
Appendix J Focus Group 6 Report 88
Appendix K Summary of Engagement Activities 93
Appendix L Printed Information and Materials 104

Appendix A adidas meeting Report

Collaboration to Clarify the Cost of Curation

Minutes of meeting with adidas on February 7th 2014 via Skype

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	10 Feb 2014	First draft	KH
1.00	28 Feb 2014	Ratified version	KH

Attendees

- Christina Block, adidas
- Maria Scherrers, adidas
- Katarina Haage, DNB

Agenda

This meeting took place as a follow up to the second 4C focus group on industry bodies in early December 2013. In the run up to this focus group a webinar was held on November 26th 2013 via WebEx with regard to background and context of the 4C project and a special focus on the *Economic Sustainability Reference Model* (ESRM), developed by project coordinator Neil Grindley .

Following questions were presented to the webinar participants to prepare for the subsequent focus group, held in London and Frankfurt:

1. What is the main motivation for your organisation to "afford" digital curation?
2. If you categorise the digital objects that you are in charge of being either ASSETS or LIABILITIES… is that possible?
3. What benefits or outcomes does your organisation expect from digital curation?
4. Do the costs of curation actually matter in your organisation? And/or the potential Return of Investment?

Unfortunately, the event in Frankfurt had to be cancelled at short notice, due to the fact that most participants called in sick that day. This meeting was held instead.

Minutes

Adidas's archive of the digital object contains pictures and videos etc. of the fashion collections and also catalogues and advertising posters that are being digitized and uploaded so they can be seen by visitors of the online archive https://www.adidas-archive.org/#/home/. The archive consists of 1.675 objects so far and shows 147 exhibitions. Adidas started digitizing their archive objects 5-6 years ago; it is an active and steadily collecting. The storage happens with "TMS" software which is similar to "MuseumPlus" in close cooperation with the in-house IT department. They have three servers and do double storage; the photos are archived as jpegs; the original photos are stored as tif separately to be able to have access to them at any time.

All in all, the adidas archive has three criteria when it comes to what they include in their collection:

a) important object like the sports outfit of a sports legend
b) technical and design innovations
c) typical outfits from a certain decade

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

- company's history management
- internal usage; i.e. for exhibitions or studies regarding the fabric and patterns etc.
- legal department; for patents etc.

2. If you categorise the digital objects that you are in charge of being either ASSETS or LIABILITIES… is that possible?

It is both; for the history department it is more an asset as well as for the internal usage. The physical objects are definitely an asset of high value for the company. Of course, for the legal department the archived objects have an important position when it comes to patent discussions.

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

See above.

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

The archive or rather curation department has more or less “power of decision” when it comes to the budget; they have a steady budget “framework” for continuously digital archiving work.

Appendix B AV preserve meeting Report

Collaboration to Clarify the Cost of Curation

Minutes of WebEx meeting with AVPreserve representatives, 18th March 2014

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	20 Mar 2014	First draft	SS

Attendees

AVPreserve:

- Chris Lacinak
- Bertram Lyons
- Rebecca Chandler

Cinevation:

- Riccardo Lery

4C:

- Sabine Schrimpf, German National Library
- Katarina Haage, German National Library
- Alex Thirifays, Danish National Archives
- Luis Faria, KEEPS

Agenda

- Introductions
- Brief overview of 4C initiative and projects
- (Demonstration of 4C CCEx plans and drafts)
- Brief overview of AVPS
- Demonstration of Cost of Inaction Calculator

Minutes

4C introduction

Katarina Haage started the meeting with a short introduction of the 4C project: The purpose of the project is to create a better understanding of digital curation costs through collaboration. The project goal is to provide useful, useable resources which support the process of cost management in digital curation.

Curation Cost Exchange

Next, Luis Faria and Alex Thirifais (AT) presented some key resources that are to be created within the course of the project: the Curation Cost Exchange platform (CCEx) with mock-ups and the envisaged submission template. The CCEx is generally designed to cover costs for all kinds of digital materials.

Chris Lacinak (CL) asked if the people with whom the 4C project is in touch are familiar with cost modelling and have their cost information ready. AT explained that on the contrary, it is very tedious for them to collect this information. The 4C project notes that there is quite a distance between the people who are responsible for curation and the financial departments of their institutions. CL confirmed that AVPreserve has also made this observation.

CL marked that the submission templates looked quite complex and the results that the CCEx produces are not necessarily self-explanatory. AT said that the submission template testing showed that people spend between 3 hours and 3 days for filling in the excel sheet.

→ CL volunteered AVPreserve to participate in the CCEx testing.

AVPreserve introduction

AVPreserve has looked at curation costs in the context of their work with universities and the New York Public Library on AV preservation. They look at workflows, do cost and effort analyses and compare inhouse vs. outsourcing solutions. They tend to focus on digitization, but look at surrounding issues, too, of which digital curation is an important one.

Cost of Inaction (CoI) Tool

CL introduced the background/motivation of the Cost of Inaction (CoI) tool: AVPreserve typically works with operational staff. It was sometimes found difficult to defend necessary investments against the executive management level, where typically questions on Return on Investments (RoI), opportunities for monetization are asked. There are usually no classical RoI arguments, but the fact mustn't be neglected that institutions have usually heavily invested in building, organizing and maintaining their AV collections. If institutions do not invest in their curation, they risk this investment (not to mention their reputation).

The CoI calculator visualizes this point. It is focused on AV media and takes into account their foreseeable obsolescence and degradation. The trickiest, but also a critical, part of the tool is the "investment to date for media". AVPreserve has talked to a number of institutions what kind of costs need to be included here, and found that a) this differs a lot and b) it is difficult for the institutions to come up with this numbers.

→ Maybe the 4C work will be of help here.

CL admitted that some tricky assumptions are included, e.g. the decrease of storage cost, or the increase of digitization cost. However, the pre-filled values result from expert consultation, and customers can change them if they disagree.

The CoI tool is work in progress, but can already be found on http://coi.avpreserve.com.

Sabine Schrimpf said that the tool will be of particular interest to the 4C Enhancement Group that looks at cost related concepts like "value", "risk", and "benefit" of digital curation.

→ She will share minutes of the meeting with the 4C project group and encourage project members to provide feedback on the CoI tool.

Appendix C Bedern Group meeting Report

22nd May 2014

Attendees

- Alex Thirifays, Danish National Archives
- Robert Dickinson, English Heritage
- Sarah Middleton, DPC
- Gareth Edwards, Royal Commission on the Ancient and Historical Monuments of Wales (RCAHMW)
- Helen Shalders, English Heritage
- Katarina Haage, German National Library
- Katie Green, Archaeology Data Service (ADS)
- Luis Faria, KEEP Solutions
- Sabine Schrimpf, German National Library

Agenda

1. Introduction round (all)
2. Introduction to the Bedern Group, its work etc.
3. Brief introduction to the 4C project (Katarina Haage, German National Library)
4. Presentation of the CCEx mock ups (Luis Faria, KEEP Solutions)
5. Walk through the submission template (Alex Thirifays, Danish National Archives)
6. Open discussion
7. Bedern and 4C—next steps?
8. Bedern and digital preservation costings

1. Introduction round (all)

2. Introduction to the Bedern Group, its work etc.

The Bedern group, is principally a working party chaired/facilitated by the DPC. In 2011 the DPC was asked to convene a group of members: Archaeology Data Service, English Heritage, the RCAHMS and RCAHMW, which preserve data pertaining to the historic environment. In particular this is data for heritage management rather than historical data. They collect everything from specialist research data to routine reports, data types include biological, geophysical etc. and they interface with construction companies, mineral extractors, and environmental agencies.

The name is taken from the location of the first meeting in York in 2011.

Bedern members have a joint work plan and declaration. They share an interest in digital preservation, protecting the historic environment and a commitment to preserving intellectual heritage for future generations.

Previously their processes were entirely paper based, and that worked then. It doesn't work anymore, particularly in archaeology which is a very destructive process. Data is all that remains following an excavation.

The work plan encompasses three core themes/activities:

- To co-ordinate and harmonise collection policies
- To develop shared data deposition standards
- To establish the costs of preservation services

Some members provide services to others: some free of charge, some at cost, some free at point of use.

There was an item arising from the last Bedern Group meeting to develop the third activity—to establish the costs of preservation services.

The activities of the 4C Project seemed like a good place to start with taking that action forward.

3. Brief introduction to the 4C project (Katarina Haage, German National Library)

The project is a collaboration to clarify the costs of curation. It is a 2 year EU funded project with 13 partners from 7 different countries.

Project Summary: The Collaboration to Clarify the Costs of Curation (4C) project will help organisations across Europe (and beyond) 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.

4C is an open and social project which listens to the needs of its stakeholders so comments and discussions are welcome in order to further develop the project outcomes.

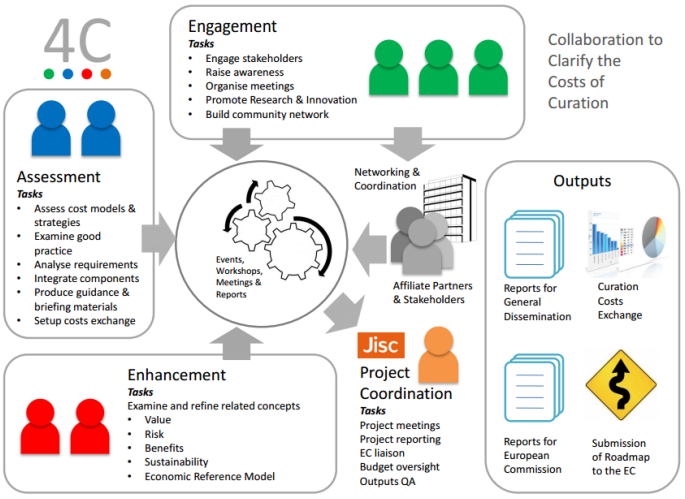


Figure 12—4C project work packages

The project is arranged into 5 work packages, managed by Jisc.

Understanding the cost of curation…

- helps to offer realistic and cost effective curation services to others
- can support strategic planning
- can support tactical decision-making
- can provide evidence of cost-effectiveness and value

Understanding economic drivers can help to strategically align an organisation. The project does not want to come up with another cost model—but rather provide the tools to help stakeholders better understand the cost of digital preservation.

Key resources and outputs:

- Indirect Economic Determinants (IED)—a taxonomy of terms which reflect the most important stakeholder motivators for preserving digital data
- Economic Sustainability Reference Model (ESRM)—a strategic reference tool to help in planning long term investments in terms of economic sustainability
- Needs and Gap Analysis—an evaluation of 10 cost models in terms of various criteria e.g. whether it is able to model cost variables, output types etc. Some of the gaps identified in the currently available models include a lack of reliability…
- Cost Concept Model (CCM)—currently in mind map form, this is a framework to support future research and explanins current challenges and themes in cost modelling for digital preservation.
- Curation Costs Exchange (CCEx)—a platform for users to access a range of information which is intended to inform the process of investment in digital preservation e.g. a cost comparison tool, a literature and articles library, information on cost models.

All outcomes are published to the project website for community review, comment and feedback, as well as publishes news and blogs: www.4cproject.eu

4. Presentation of the CCEx mock ups (Luis Faria, KEEP Solutions)

4C is developing web mockups of the Curation Costs Exchange (CCEx) at present, with a view to making this available online later this year.

The current iteration of the CCEx comprises the following features:

- Cost input—framework of comparable costs. Users map their costs to a set of predetermined categories to enable comparison—either global or peer-to-peer.
- Cost information—a collection of information to help users get started in using cost models, as well as a library of literature and articles for further reading
- Discuss and Share—platform for sharing experiences and connecting with other users
- News and Events—an aggregation of news from relevant sites as well as an calendar listing dates of useful and relevant events.

The completed CCEx is due for release in autumn 2014, and will be formally unveiled at the DPC/4C Conference, 'Investing in Opportunity: Policy Practice and Planning for a sustainable digital future' which will take place on 17th - 18th November.

This is not yet complete and feedback is welcomed at this stage to incorporate stakeholder ideas and preferences.

- 'Profile' tab—users are asked to define their organisations and their collections in order to enable comparison
- 'Cost input'—users define cost units, type and volume of their data and breakdown their costs by activity or capital procurement/labour
- 'Cost Analysis'—shows a summary of costs as entered, then option to compare with global average using data submitted by other users, or-peer-to-peer comparison based an average

of data submitted by organisations with a profile match. There is also the option to make contact with those similar organisations to share experiences

Bedern Group can trial the mock-ups at: http://4c.keep.pt/ccex/mockups/

5. Walk through the submission template (Alex Thirifays, Danish National Archives)

The submission template is an excel version of the mock-up Luis has presented—may not be great value in walking through this again as it will duplicate the previous presentation.

4C has a challenge to convince organisations to take the time to work through this process:

- approach financial departments
- extract data
- input to CCEx
- map and normalise to the CCEx categories—thus losing a degree of accuracy

Why should people make this effort?

6. Open discussion

Undertaking this exercise would create greater transparency of costs.

Would certainly be useful when making an internal case for digital preservation—you need to have significant data behind a case like that, CCEx can provide this.

Would the Bedern Group be prepared to share data and undertake the process?

- RCAHMW—working through the process with RCAHMS to try and work out what digital preservation costs in the first place, not sure they could submit anything at this stage.
- English Heritage—sharing is very important to create transparency. Perhaps the process could be iterative: fill in what we know, compare, validate, move on to the next stage?
- It is not impossible for English Heritage—just couldn't put everything in there at once.
- ADS—should be able to share data easily, just gone through a similar modelling process. ADS is a smaller organisation, more independent and more in control of what they do with their own costs.

Piecemeal costs would still be welcome. The profile section allows users to define the level of costing they are submitting. This could be made more granular still.

The idea is not to predict costs however, rather to compare, validate, benchmark, to be used as a self-assessment.

Bedern group happy for other organisations to contact them regarding their costs if they were of interest. Confidentiality not too much of an issue—confidential costs like salaries would be aggregated anyway and staff would not be named, so a problem is not envisaged with connecting organisation names to the data submitted.

Suggestion that the results should be printable or exportable for use in reports/business cases.

ESRM

This might also be used by the Bedern Group. The results could be made openly available, so that producers can decide where to deposit themselves?

7. Bedern and 4C—next steps?

Help progress agenda item for Bedern.

Set UK benchmark for costing.

Agree what the organisations are costing.

- For own planning purposes and presentation to budget holders
- Developer pays—what is a reasonable cost for this? Need to think about this. Want to cost the ideal, not what is taking place at the moment.

There is more to 4C to talk about than has been shown in this meeting.

- e.g. the 'Evaluation of Cost Models and Needs & Gap Analysis:' http://www.4cproject.eu/community-resources/outputs-and-deliverables/d3-1-evaluation-of-cost-models-and-needs-gaps-analysis
- May be a useful key to unlocking next steps.

Help progress CCEx for 4C

- Share link for mock-ups so Bedern Group can trial and provide feedback to 4C
- ADS to share data through submission template
- RCAHMW and English Heritage to identify relevant partial information which could be shared (at a later date)

8. Bedern and digital preservation costings

Action from last meeting on Collection Policies—Emily (RCAHMS) compiling by the end of April, WK to follow up.

Appendix D Nestor working group on costs meeting Report

Collaboration to Clarify the Cost of Curation

Minutes of 4C meeting with nestor on 25th March 2014 at DNB

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	31 March 2014	First draft	KH
1.00	9 January 2015	Final version	KH

Attendees

- Armin Straube (DNB), a.straube@dnb.de
- Karlheinz Schmitt (DNB), k.schmitt@dnb.de
- Katarina Haage (DNB), k.haage@dnb.de
- Sabine Schrimpf (DNB), s.schrimpf@dnb.de
- Martin Iordanidis (hbz), IORDANIDIS@hbz-nrw.de
- Michael Nelissen (hbz), NELISSEN@hbz-nrw.de
- Torsten Rathmann (DKRZ), rathmann@dkrz.de
- Yvonne Friese (ZBW), Y.Friese@zbw.eu
- Alex Thirifays (DNA), alt@sa.dk
- Neil Grindley (Jisc), n.grindley@jisc.ac.uk
- Paul Stokes (Jisc), p.stokes@jisc.ac.uk
- Sarah Norris (DPC), sarah@dpconline.org
- Ulla Bogvad Kejser (KBDK), ubk@kb.dk
- William Kilbride (DPC), william@dpconline.org

Agenda

- Welcome
- Housekeeping
- Introduction round
- Presentation 1: Introduction to nestor working group on costs
- Presentation 2: Introduction to 4C project
- Presentation 3: Economic Sustainability Reference Model
- Presentation 4: Curation Costs Exchange (Mockups and Submission Templates)

Minutes

The main purpose of this meeting was an exchange of experiences. The nestor working group has also tried to collect cost data and they have done some general cost model research and are finalizing a guideline / a set of recommendations on how to identify / calculate the cost of curation. In this way they look at typical preservation workflows and where the cost drivers may be.

After a short welcome and information about the housekeeping the attendees began with a short round of introductions.

Yvonne Friese (YF) then introduced the nestor working group on costs with a presentation and pointed out the structure of this working group, the purpose and approach to the task as well as the challenges. In the following discussion it became clear that nestor and 4C are dealing with the same sort of challenges when it comes to collecting external cost data or rather using external cost data to work with. Neil Grindley (NG) asked if 4C could use the cost data of ZBW YF had collected for the Curation Costs Exchange (CCEx). Unfortunately, YF had to refuse since the collected data of her institution is confidential. However, she also stated that it is important to encourage institutions to share costs in order to be able to compare these. Karlheinz Schmitt (KHS) then mentioned it would be sensible to talk about abstractions of costs, i.e. no longer calculate the processes but the categories in which the costs occur. Martin Iordanidis (MI) stated the importance of telling people or rather educating them on sharing cost data. Regarding the "Broker Service" which the nestor working group on costs would develop in the best case Alex Thirifays

(AT) proposed to meet in the middle with the CCEx and exchange experiences on developing a tool, service or platform.

NG then presented the 4C project to the nestor group on costs. In the subsequent discussion it was emphasized that not counting or calculating the costs of digital curation is the major issue but comparing these costs seems the most complicated part. The nestor members agreed on this statement. AT presented the CCEx with its mock-ups and the submission template and informed the group that 10 data sets of 4C members and 5 data sets of trusted members (Advisory Board) have been collected so far; publically available data sets from different public institutions are to be collected in the future. In order to encourage institutions to share their cost information via the CCEx, the 4C Engagement package will communicate the benefits of the CCEx and submitting cost data to it to the community via focus groups, webinars and expert interviews. Sabine Schrimpf (SS) asked the nestor members if they respectively their organisations would be willing to submit (anonymized) cost data to the CCEx submission template and they considered yes; Torsten Rathmann (TR) will submit the data sets to AT. TR also stated that he sees the issue of submitting costs two-fold: the staff costs on the one hand are easy to determine (in his organisation rather in working hours than salary), however, the storage costs on the other hand seem to be the nub of the matter. Another idea that came up is to link the CCEx to certification through nestor. In this context NG pointed to the Economic Sustainability Reference Model (ESRM) which he also presented to the nestor group. The overall consensus was that the ESRM is a useful tool to document the costs and investments in digital long term preservation for the particular organisation that it can get back to in case of a revision of costs and data etc. and to use it as a risk management tool. However, it was asked how the results of the ESRM Appendix can be analysed or evaluated. Thoughts on that are for example to define trends, important areas and gaps as well as finding key aspects of costs in digital curation. It could also be used as an “alert service” since it includes a timeline.

In conclusion the meeting was a great opportunity to compare notes between both initiatives and make sure to be aware of cross-overs and use these and other synergies to collaborate.

Appendix E Focus Group 1 Report

Collaboration to Clarify the Cost of Curation

4C Focus Group at iPRES 2013 in Lisbon, Portugal

Report

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	Jan 13 2014	First draft	SS
	Jan 14 2014	Edited	KH
	Mar 07 2014	Edited	KH

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	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 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

Minutes

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 top 6:

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 the IEDs separately.

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 does not have much choice in prioritizing or even selecting from the list of IEDs. The participants raised the question if it cannot 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 using 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 ESRSM exercise needs to be made clearer and needs to be pitched by the 4C project.

Appendix F Focus Group 2 Report

Collaboration to Clarify the Cost of Curation

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.

(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 approached 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.

Appendix G Focus Group 3 Report

Collaboration to Clarify the Cost of Curation

Minutes of "DPHEP Full Costs of Curation Workshop" on 13-14 January 2014 at CERN in Geneva

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)	

History

Version	Date	Changed pages / reason	Modified by
1.00	7 Feb 2014	First draft	Stephan Strodl
2.00	16 Feb 2014	Second draft	Anders Bo Nielsen

Minutes

In January 2014 the DPHEP Full Costs of Curation Workshop took place at CERN in Geneva. The 4C Project was invited to present the 4C project, and was represented by Neil Grindley, Stephan Strodl and Anders Bo Nielsen.

The aim of the workshop was to increase the understanding in detail of the full costs of curation for both past and current experiments from High Energy Physics (HEP). The international Collaboration for Data Preservation and Long Term Analysis in High Energy Physics (DPHEP) (http://www.dphep.org/) was established as a study group of the International Committee for Future Accelerators (ICFA).

On the first day of the two day workshop the existing experiences in the long term preservation of projects and experiments in the area of HEP were presented. The presentations illustrated the required effort and costs in curation of experiments after the data recording phase—including storage media migration and porting of computation software and experiment data to new systems.

In the area of HEP the scientific output of experiments does not stop with the end of data collection. For example 25% of the publications from HERA were published after 2007. HERA at DESY (Deutsches Elektronen-Synchrotron (German Electron Synchrotron)) include the H1 and the ZEUS detector was operated from 1994-2007. It is the world's only accelerator that is able to collide protons with either electrons or positrons. Already in 2008 the preservation activities for the HERA experiment started including work on the preservation and archival storage of the data (in total about 1050TB) and the preservation of the software and its environment. In collaboration with INSPIRE (the High Energy Physics information system made by CERN, DESY, Fermilab and SLAC) the documentation of the experiments should be maintained including digital as well as non-digital material.

The enhanced knowledge over time and improved methods such as Monte Carlo models make the preservation and re-analysis of old data and experiments in High Energy Physics scientifically mandatory. By the increase of precision and discovery new theoretical insights can be achieved as well as the improvement of current research work. During the workshop a series of curation activities were presented amongst others from the Large Electron–Positron Collider (LEP) at CERN. The particle accelerator was used from 1989 until 2000 with four detectors ALEPH, DELPHI, OPAL and L3. The ALEPH experiments were performed from 1989 to 1998 and about 15,000 runs were executed. In order to use new models and simulation for the data the preservation of the full software chain is required for new data production. A virtualisation approach was chosen to maintain the legacy software and data. The costs of the required resources including server and storage decreased over time significantly. A decade later the costs for the infrastructure are less than ten percent of the original cost. The optimization of resources was important during data taking, but becomes negligible after a decade. A challenge for the future is the preservation of the knowledge about the data and the software. Investments are required such as open data approaches and environments to share data analysis to support and share the understanding of experiments.

Another good example for curation effort was shown for the JADE experiment. The experiments were performed on the PETRA collider from 1979-1986. The curation activity provides a good example of the effort required for a system migration. The experiments were developed and performed on mainframes from 70s and 80s. In the late 90s the data and the software were migrated to IBM AIX on RS6000. During the software migration different challenges arose such as missing libraries, the use of pre-processors and nonstandard conform programming language usage. It took the work of a PhD and a Postdoc for a year to perform the software migration and required additional advice of JADE software experts. The re-analysis of the JADE data has led to significant improvements of the precision and provided measurements in a

unique energy domain. The costs for the JADE resurrection showed that personnel costs dominate over hardware expenses. Overall curation effort including physics analysis took so far more than 15 years and cost less than 1 million Euros.

Following each presentation many discussions took place. The following characteristics for curation for the area of HEP were made by the 4C participants based on these discussions.

Regarding the curation time horizon and curation cost the area of High Energy Physics is characterized by being able to reproduce experiments, in contrast to many other areas such as astronomy. Due to the theoretical and esp. technological development the time horizon is estimated by DPHEP to be somewhere around half a century. This is because DPHEP estimates that they will be able to get much more and more precise results reproducing an experiment after half a century instead of curating and analysing data from old experiments. On the other hand, analysing data and understanding experiments fully can take decades, and some of the best published results appear often many years after the experiments were made.

Regarding cost the need for curation depends on the cost of curating experiments compared to the cost of reproducing them. Due to the extreme cost of creating many of the experiment the curation cost for half a century is assumed to be a few percent. Nevertheless, funding for equipment and facilities are easier to get than for operations such as curation, and therefore curation cost should be shown as a part of the budget of establishing an experiment.

Neil Grindley from the 4C project presented the project and its goals at the DPHEP workshop. The collaboration aims at helping organisations to understand their costs and clarify the complex relationships between costs and other factors. An approach of the project is to gather data from different organisations and investigate the methods and structures that the organisations used for their costs of curation. The output of the project should help to understand and share costs for curation across organisations. The work of the 4C project also addresses related concepts such as risk, value, quality and sustainability.

The second day of the workshop focused on the planning of the future of curation. A cost planning tool for bit preservation was presented allowing calculating the approximate costs for data archives over 10, 20 and 30 years. Different scenarios showed the effects on storage costs for different collection growth rates. Virtualisation as enabling technology for preservation for HEP experiments and the advantages compared of porting software to new systems were discussed in a second session on day two. The reduced maintenance effort of virtualised systems was identified as a major advantage. Performance degradations should be solved by the increase of computation power of time and it can be used to distribute the workload across multiple VMs in parallel. Virtualisation can support long term preservation for the experiments as it allows the encapsulation of the data production system. Common services that are used by experiments such as data service or detector condition services need to be preserved by using virtualisation. The current CernVM was presented; it has been used for LHC data analysis since 2010. The virtual machines are integrated well with today's cloud infrastructures and can be used for local and Grid computation that can be flexibly contextualized. The Virtualisation File System is a network file system supporting versioning and snapshotting of file systems. The use of CernVM technology allows retrieving the historic data processing environment in the future.

In a break-out session the 'Curation Costs Exchange' (CCEx) of the 4C research project was introduced. It should allow contributors to exchange their curation cost information via an online platform. The platform allows gathering cost information from partner organisations and stakeholders using submission templates. The aim of the CCEx is to collect empirical knowledge about costs information and structure of

curation costs that can be used in the research project to increase the understanding and provide methods for sharing curation costs. The feedback for the 4C project and its cost exchange was very positive. It was remarked in the discussion that cost information is available in different organisations from different areas, but so far too little effort was made to collect and consolidate this information. The need and the positive effects of having cost comparable information received consent from all sides.

More information about the 4C project can be found at: http://www.4cproject.eu/

The activities and reports from the Collaboration for Data Preservation and Long Term Analysis in High Energy Physics (DPHEP) are published at: http://www.dphep.org

Appendix H Focus Group 4 Report

Collaboration to Clarify the Cost of Curation

Report of Focus Group Meeting on May 13th 2014 at ARCHIVING Berlin

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	19 May 2014	First draft	KH
1.00	23 May 2014	Ratified version	KH

Attendees

4C:

- Alex Thirifays, DNA
- Hervé L'Hours, UK Data Archive
- Katarina Haage, DNB
- Ulla Bogvad Kejser, KBDK

Participants:

- Dr. Harry Enke, aip Potsdam
- Elena Simukovic, HU Berlin
- Joost van der Nat, NCDD
- Karen Colbron, Presto4U
- Lindio Ligios, Presto4U
- Marcel Ras, NCDD
- Marcin Ostasz, Barcelona Supercomputing Centre
- Martin Iordanidis, hbz Köln
- Reiner Mauer, gesis

Agenda

Time	Topic	Responsible	Description
13.30-14.00	Introduction round Overview of the 4C Project	Katarina Haage, German National Library	
14.00-15.30	General introduction to CCM and GRS	Hervé L'Hours, UK Data Archive	A general introduction to the challenges of the Cost Concept Model/Gateway Requirements Specification and how we have chosen to structure our approach
	A 'specification' exercise	Hervé L'Hours	Exercise to consider your requirements from a cost modelling perspective
15.30-15.45	Break		
15.45-16.00	Excursus—Netherlands Coalition for Digital Preservation (NCDD)	Joost van der Nat and Marcel Ras, NCDD	Introduction to the NCDD and the present project on digital preservation and the relation to the 4C project
16.00-17.00	General introduction to the Curation Costs Exchange (CCEx)	Alex Thirifays, Danish National Archives	15 minutes introduction to the concept, the purpose, the hopes, the sustainability and the what's-in-it-for-you

	In-depth discussion about the submission of costs	Alex Thirifays	Discussions on the usefulness of sharing and comparing costs; exercises regarding the submission template; feedback and exchange of views and experiences
17.00-17.30	Round-up and possible future collaborations		

Minutes

After a welcome and brief overview of the 4C project by KH the participants introduced themselves and formulated their expectations of the meeting. Goals for the day were: Share knowledge across projects—learn and bring, work to find out what it costs and get a cost "template".

HLH presented the core concepts of theCost Concept Model (CCM) and Gateway Requirements Specification (GRS). What it is—and what it isn't.

Prior to the event attendees were provided with an outline of the deliverable structure and the workshop questions which would be presented

At the event the structure and content of the deliverable were described in more detail. 9 attendees and guest speakers were then split into two groups, mediated by members of the 4C project team. Though forms were provided for annotating responses to questions (see some extracts below) the general tone was informal as the project team took advantage of an opportunity to test concepts and definitions in the real world. Responses are anonymised.

Attendees were primarily those with a role similar to repository managers or data managers within a project scenario so potential implementers of curation cost models rather than theorists. This presented an opportunity to work on the themes of stakeholder context, organisational profile and attitude to benefits. Resources were addressed through a question relating to categorisation of labour forces and activities were addressed via questions about the approach to the structure and quality of activities and the structures of collections; these latter themes are represented under separate sections of the core cost concepts but have a clear relationship to the 'facts' about a system which form the organisational profile.

These themes together are critical to the structure of the framework as their clear communication between implementers and theorists is critical. Potential implementers of curation cost models must have a certain level of understanding and maturity around the issues to select the appropriate model and implement it with an appropriate degree of success. Theorists must understand these issues to correctly model the underlying curation systems relevant to their method. These subjects are essential for designing use cases or scenarios which are a common means of communication between the developers of a methodology and their uses.

Overall the structure of the deliverable was well received and provided an effective means of standardising communications around this complex area. As expected there are some artefacts from current cost models (e.g. common descriptions of collection profiles and a standard approach to grouping types of labour) which were not well established within the attendee's organisations. One area for improvement in the deliverable is a clearer distinction between stakeholder context (stakeholder identification and management, perhaps closest to Designated Community monitoring in OAIS terms) and Organisation Context which will be more clearly defined as relevant business intelligence derived from Stakeholder interaction in future versions of the deliverable.

Original question text below is quoted and in bold. Each bullet collects key parts of the responses from a participant.

1. Stakeholder Context

It became clear during discussions that the defining factor for the separation of stakeholder context and organisational profile levels of the systems was the artefacts they produce. Stakeholder engagement produces contact lists, questionnaires, minutes and other records to manage the communication process. The outcomes of that process are then operationalised into actionable parts of the organisation profile through the creation of targets, mission statements, service level agreements, reporting procedures, mandates, contracts, licences, risk registers etc.

"Do you identify relevant stakeholders for your organisation If so who are they…"

- Funders, legislators, standards developers, staff, users
- Some organisations are very complicated amalgams of projects and services, this increases the complexity of stakeholder identification and interaction.
- Funders, user communities, sites which harvest our content
- Customers, software providers, depositors, financial supporters (ministry)
- Creators, content providers, users, professionals, industry. Decision makers are differentiated from influencers.
- Creators, users, decision makers, influencers, depositors, professionals.
- Domain memory institutions, libraries and archives
- Customers, libraries, researchers, financial supporters
- Definition of use cases

"…and how do you manage and communicate with them?"

- Reporting
- We diagram them in terms of their roles and influence
- They support in developing mandates and mission statements
- 1 to 1 contact

2. Organisational Context

"What aspects of your organisation (your 'organisational context') do you consider critical to cost model issues?"

- Skills, budget, collection profiles, data standards, legislative environment
- Staff costs, software licencing, storage (network operations
- Sustainability ("forever"), zero cost to users, trustworthiness (TDR status), and openness of data (minimal access restrictions)
- Our Mission and Vision must be met by delivering our primary business processes
- Policies, processes, legal framework
- Business plans, use cases and business cases
- User needs (e.g. tiered storage and use on demand)

3. Incentives/Benefits

"Do you address Indirect Economic Determinants, incentives, risks and benefits explicitly within the organisation?

- Not at this stage, only risk is talked about
- Risk assessment
- Through development plans and strategies, preservation policy
- Efficiency is considers an operational issue, not a cost issue. Benefits are only addressed at a high level
- Delivering our tasks and responsibilities within budget and within our legal obligations is the focus. Risks are dealt with if we can't do that.
- Budgets are provided and we deal with this by making choices
- As a project we directly address risks and opportunities around all activities and review them annually
- Funder driven focus
- Long-term goals are damaging for science

"If so does this follow a formal approach (ISO standard, risk analysis etc.)?"

- Data Seal of Approval. Looking at the ISO16363 for trust and ISO27001 for information security.
- ISO16363, DRAMBORA, DIN31644
- Audit and Certification through the Data Seal of Approval is planned

4. Service/Activities

"How do you structure repository activities…"

- We're currently in a project phase so via a project management system
- Based on OAIS and organisational structure
- To be decided, we're still designing the system
- Organisational structure and divisions
- We sit at the middle and don't know what each site does in detail

"and manage their quality?"

- Validation of standardised XML but otherwise by trusting the opinion of expert teams who document their actions

5. Asset Adjustments

"How do you structure your collections? Are format, complexity and quantity the critical criteria?"

- A data collection is made up of data, metadata (structured documentation) and documentation (text files and descriptions which aren't that structured. Formats are approved because of value to users and preservation efficiency. No specific complexity consideration. Quantity is monitored at Archival Storage level but could be better monitored through quantity estimates pre-Ingest.
- This is done by the user, we just take what they give us

- PDF/A is the required format with quantity being an issue for both transmission and dissemination
- Structured by type and format, evaluated through significance 2.0 http://arts.gov.au/sites/default/files/resources-publications/significance-2.0/pdfs/significance-2.0.pdf

6. Labour

"Do you categorise your labour force? In terms of Qualifications, experience, training received, job description etc."

- No, the requirements are defined on a per role basis. It would not be useful to band staff costs by this method.
- Not really
- By speciality (developers vs maintenance), job description, hierarchy and salary scales

After a short coffee break Marcel Ras from the NCDD gave a guest presentation and introduced the participants to the structure of the organisation followed by an overview on NCDD's current project on cost modelling and the goals by his colleague Joost van der Nat.

After this the participants were introduced to the CCEx. Before the meeting information about the CCEx and the submission template were sent to the attendees to prepare the session. After a brief introduction to the CCEx by AT the group started a discussion and gave feedback on what they had received for preparation. In the following comments, questions and remarks on the CCEx are listed:

- It took some time in the beginning to understand how to start enter information in the submission template due to the fact that an example in the cost chart was missing
- A cost submission template is needed: If organisations were able to fill in the submission template effortlessly, they would not need it at all
- Terminology—be very clear about the definitions
- The products that are developed are not always only curation related and are used for many departments in an organisation, which means that activities costs cannot always be attributed to digital curation only (complicates the submission of costs)
- Create more filters/refine the 'Profile'tab
- Narrow the scope of stakeholders or develop various templates for types of stakeholders/institutions
- Focus on defining clearly the pre-ingest (Production) process and make clear on which assets you put the focus
- It's not possible to operate with FTE's only

The attendees were then asked to think about the benefits of the submission template or rather why they would submit data:

- Average calculation
- Benchmarking; look at the change from year 1 to year 2—have we done better than the year before? Have our plans, strategies, choices had the intended impact?
- Self-assessment
- For budgeting, planning, predictions
- Important information for funders
- Important information for sending bills

- Development of new services and to inform business models
- Communication
- Collect and compare figures and facts
- Raise awareness
- Submission of cost data requires clear outlining of issues of anonymity and confidentiality
- Doing the exercise spurs discussions within the organisation—the exercise itself is beneficial.
- Mentality change: Maximise the culture of sharing—it instigates faster development, enhances trust raises awareness and leads to increases in efficiency

In the end of this meeting further engagement between 4C and the NCDD and Presto4U was agreed upon.

AT mentioned that the Curation Costs Exchange Platform may encounter some sustainability issues beyond the lifetime of the project, but that the Digital Preservation Coalition and Nestor both have showed an interest in keeping it alive. NCDD stepped in at this point and gracefully offered to engage in sustainability negotiations with the 4C-project as well.

As an outlook the attendees mentioned the idea to calculate the benefits of curation and the benefits of collaboration. KH informed that all presentations will be online and circulated soon after the meeting.

Appendix I Focus Group 5 Report

Collaboration to Clarify the Cost of Curation

Report of Focus Group Meeting on Nov 3rd 2014

In London at

The Wesley, 81-103 Euston Street, London, NW1 2EZ, http://www.thewesley.co.uk/

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	8 Jan 2015	First draft	KH
0.02	16 Jan 2015	Edited	DS
1.00	22 Jan 2015	Final version	KH

Attendees

4C:

- Joy Davidson, DCC
- Magdalena Getler, DCC
- Sarah Middleton, DPC
- Diana Sisu, DCC

Participants:

- Catherine Hardman, University of York
- Grant Denkinson, University of Leicester
- Tim Clark, Harvard University
- Gareth Knight, London School of Hygiene and Tropical Medicine
- Stephen Grace, University of East London

Event schedule

- 13:00 Lunch
- 13:30 Welcome
- 13:40 Introduction to CCEx
- 14:00 Time for individual familiarisation with the tool
- 14:30 Focus group discussion on CCEx
- 15:15 Coffee break
- 15:30 Overview of Roadmap, http://4cproject.eu/roadmap
- 15:40 Focus group discussion on roadmap http://4cproject.eu/rmfeedbackand
- 16:00 END

Focus group analysis

1.0 Methodology

Focus group interviews are popular research method for exploring what individuals believe or feel as well as to understand why they behave the way they do.

The main aim was to understand, explain, and

- Find out what users think of CCEx
- To identify what users want / expect from the tool
- To find out what are the common problems when using it
- To collect recommendations for improvement
- Find out if they would have concerns about sharing their cost data and if so, why. What could be done to mitigate these concerns?

Data Analysis: Process

1. Familiarisation with the data (listening to the recording, reading the transcript—43 pages, reading the observational notes and summary notes after the interview)
2. Developing categories. Look for patterns: repeating ideas, larger themes. Also
    a. Consider the actual words used

   b. Consider the frequency and extensiveness of comments (how often a comment or view is made)
   c. Intensity of the comments
   d. Internal consistency
   e. Specificity of responses (greater attention is placed referring to personal experience)
3. Did we answer our research questions?
4. What theories develop?
5. Implications
   f. What does it mean?
   g. What major themes emerge?
   h. Is the knowledge acquired something we already know, or is it new?
   i. Does the knowledge confirm a hunch?
   j. How does the knowledge change our perspective?
   k. How are participants' environments or past experiences related to their behaviour and attitudes?
   l. **What else do we need to know?** (Additional topics for the next focus group)

2.0 Focus Group Interview

The focus group interview took place on 03.11.14, The Wesley Hotel, London.

FACILITATOR: Joy Davidson and Sarah Middleton

OBSERVERS/ NOTE TAKERS: Diana Sisu and Magdalena Getler

SIZE: 7 participants (respondents). Participants were selected on the criteria that they a) would have something to say on the topic and b) would be comfortable talking to the interviewer.

OVERALL IMPRESSION: The atmosphere was friendly and people felt comfortable in expressing their opinions.

LANGUAGE: Language used (selected quotes in yellow, below), special vocabulary—See Conclusions on p. 87.

NON-VERBAL INTERACTIONS: impact of the group dynamic. Three respondents dominated the discussion; more reserved participants could perhaps be probed more.

2.1 Questions asked at the focus group

- What are your genuine thoughts on the CCEx tool?
- Do you find it easy and intuitive to use?
- What do you like best/least about it?
- What are the barriers in using it?
- Are there redundant features?
- Is the language understandable or is there too much jargon?
- Would your organisation submit cost data? Who in your organization would be responsible for submitting the data?
- What kind of support are you currently using for costing digital curation?

2.2 Developing Categories

Feedback was extracted and categorised into Scope of the tool, Problems/uncertainties encountered by users, Suggestions for extra features/improvements, and Comments on the interface/workflow.

2.2.1 The Scope of the tool

- It was not immediately clear to participants what the tool does, how does it work, who it is aimed at:

  R6: I wasn't too sure what costs we were talking about. I'm used to doing planning ahead for services and working out the kind of business models that we're going to be talking about, how many staff we're going to need the kind of hardware, infrastructure kits and that kind of thing. It wasn't at all obvious to me that what we're talking about here really seems to be depositing data for long-term preservation…

  Because for me curation needs more than that and I think that needs to be clear up front what this is actually…what you're costing. I was confused at first because I didn't know what it was I was costing. I was trying to cost a service and none of it made any sense…

  just what the scope is because it's not necessarily obvious on the front page...

  R6: *I think in general it helps to be able to explore all this stuff to know what you're doing but you kind of need to know that before you actually start it which comes back to this idea of having to sign up and representing a university before you can even get an idea of what the processes actually involve. I think you need to be able to see what you're going to do and what kind of tool this is much more clearly before you have to sign up to anything. I didn't understand it at first, even though it looks lovely. I mean, it's a beautiful looking website, but I didn't understand it… at first when I came to this, I thought, oh, good, I've got a service I'm planning. I can do this and then I'll go to the CEO and say, can I have £2 million and I'll give you a curation service. Of course that's not the tool for this and it's not at all obvious that that's the case when you first look at the main website...*

2.2.2 Problems/uncertainties encountered by users

2.2.2.1 Conceptual:

- Participants agreed that they tend to add costs first and calculate the total later. It seemed *the other way round* to provide a cost upfront, as it is in the tool:

  R7: *I can see doing this two different ways round. The way it's done is we're going to spend this much money, 50 per cent of which is ingest. The other way, the way I'd probably get that is well, how much time are we going to spend time ingesting, how much hardware we need, how much software we need, what is the total? So it's sort of the other way round because I wouldn't know a percentage...*

  R7 continues: *But I can see that when I actually do this I'd probably do it the other way round. We'd say, well, for this much stuff we're thinking we're going to need about this much hardware and we're going to need this much people, the total being this much…*

  Respondents would prefer to asses resources first (look at purchases and staff mapping), then map to activities (decide how they would spend it per activity—ingest, storage, access) and get the total.

  R1: *When you're adding a cost unit it's confusing to be presented with the activity mapping first of all before you've actually worked out what resources you're allocating?*

- The issue of comparing costs with other organisations. Some participants expressed concerns about comparing their organisation's costs with other organisations. R4 could see the tool used to compare costs internally.

I didn't really know what I was putting my data in for. I suspect that what I'd probably end up finding most useful for this tool is for me putting in lots of data on different levels about our own costs for me to compare against itself rather than compare with someone else. I do think that we're so different actually in what we're trying to achieve that to try and…that comparison thing, it sounds awful because I know that's the whole point of your project...

R1 agrees it only make sense if comparisons are made between similar organisations:

But there must have been an impulse to look at that benchmarking because that's clearly part of what we're obviously setting out to do. I think there is a value in benchmarking as long as you're certain that you're comparing…

R4 adds: *That's right, like with like…*

R6 would like to see more questions about curation levels in order to make comparisons more meaningful:

I'm just to trying to think of this from the point of view of my librarians. They're going to be a bit terrified because in the costs of their curation during this kind of thing, senior management are going to look at it and say, that's bloody expensive, because what you've got on here, you've got the start of it and you can see how many copies people are making of it. There are other ways of measuring the kind of curation level in particular our librarians would be very nervous that there isn't a way of indicating that all the metadata surrounding the deposited data is actually reviewed by a librarian, so it has that personal review which obviously puts the price up sky high. They're never going to get away with it but they want to. It's a very different level of curation they're hoping to offer. It's going to make the prices look totally incomparable to the other averages you've got unless you can put it in a tick box or some slightly more complex option…

and continues

Yes, but there's three other parameters involved. Firstly whether all of the metadata has been reviewed on ingest. Secondly whether these copies you're capturing are spinning disk or tape, because that will make a difference to the costs. Thirdly whether there's geographical separation of where the different copies are being held because that will also increase, less of an issue than the other two factors, but they all allow people to make more accurate comparisons…

- Sharing costs: R3 didn't think the BBC could ever share cost information with anyone, even anonymously, because of the risk that tabloids could got hold of it and see how much public money had been spent on x, y or z then however well-intentioned it was it could be held against them.
- Reliability of cost data entered. This knowledge is shared between IT, Library, Research Support Office and reasearchers, not one single unit has the whole picture for costs:

  R2: How do you evaluate the reliability [inaudible 03:12] sadly they'll be less qualified to do that. Also I think you'll probably all admit it's incredibly time-consuming. Even working out the 75 per cent of your time [inaudible 03:30] when costs are [inaudible 03:33] funding may come from different sources as well. I'm not sure that, say nobody [03:40] in the BBC could actually sit down and do it and if so, it would be quite time-consuming. So people will input all sorts of numbers…

Related to the reliability issue is the issue of staff time: who has time to enter cost data. Most institutions do not employ a research data manager who has the time to chase up cost information so whose duty is it?

2.2.2.2 Technical:

- Terminology used in the tool:

  R6: *Is there a list of cost units, because I was a bit baffled at first as to what the cost unit really is…*

  R5: *I had the help page open to remind myself what the resources are and services…*

  R7: *We know that they're both going to be generic but getting a feel of the hierarchy between a data set and a unit would be helpful because a unit is a part of a thing in a thing in a data set. That makes sense when I look at it now, but because it's generic the data, it's not obvious what that means.*

- Asset types:

  R7: *Actually we probably won't split them all up individually for this sort of costing because even in one department I might have 50 different types of things. They'll just come under, kind of, instrument data. As I say, we'll start to think later human readable or not or that sort of thing but initially saying, here's some data we'll need someone else to describe what it is…*

- Data volume vs number of files. In some organisations costs are associated with number of files rather than their volume.

  R1: *Can I just make a very small comment on the data volume option? It would be useful to have a megabytes option because not all collections that I deal with are gigabyte size. I wasn't entirely sure whether that was the data volume of the SIP or whether it was the VIP or all of them all together...*

  R4: *I had a similar or the same sort of set up issues that Gareth's got is that I was unsure because we work on…we had a bit of a discussion of this on the table, we at ADS work on file types and numbers of files. It's not really about the gigabyteage because some of our more complex files might actually be quite small but we could spend more staff time on them and the gigabyte stuff is associated with our archival storage costs. The numbers of files and file complexity is associated with staff cost…*

- Graphs. Some respondents struggled to understand the chart:

  R4 comments: *you have your nice little pie chart that works all your percentages out and then your file types, your data types, we did have a specific question about whether graphics was the same as images, but it gives you a skewed assessment of the amount of work that's been going into an archive if…so most of our archives might be word process files or spreadsheets; in gigabyte terms they're quite small. We might have 100,000 images; in gigabyte terms they're massive but in terms of the percentage of time we spend on them, very small because we'll batch process them and they're easy. So that didn't reflect properly in my case how we would divvy up the work associated with it, all the costs associated with it...*

  R7 says: *I've noticed the graph doesn't show the percentage that we put in…*

And later adds: *I haven't tried to download just when I see activities, if I hover it gives me the amount, which is great, but the thing I put in was the percentage…*

2.2.3 Suggestions for extra features/improvements

- An option to upload an excel spreadsheet with data sets (or reintroduce a submission template)

  R 7: *Well, if I had a hundred data sets, having to type them all in to get a graph, I can just take it straight from Excel and do that. I couldn't do the comparison, but I could do a lot of the other things without the tool…*

  R1 suggests: *Does it sound like there needs to be some kind of standard Excel template that people can use to import stuff into and then we can do the comparisons?*

- An option to share costs internally before publishing

  R1 inquiries: *Can you share your in development stuff that you are putting into here with other people before it's published?*

- Inserting a 'department' field: participants were concerned about appearing to represent a whole institution.
- Being shown an example of what the tool does (also see Scope of the tool)

  R6 says: *I think in general it helps to be able to explore all this stuff to know what you're doing but you kind of need to know that before you actually start it which comes back to this idea of having to sign up and representing a university before you can even get an idea of what the processes actually involve. I think you need to be able to see what you're going to do and what kind of tool this is much more clearly before you have to sign up to anything. I didn't understand it at first, even though it looks lovely. I mean, it's a beautiful looking website, but I didn't understand it…*

  Facilitator1 adds the project is in the process of developing a wizard, a video tutorial.

  R6: *Yes, that would be really handy, just so you know what they're doing and can check that it is going to be useful, because at first when I came to this, I thought, oh, good, I've got a service I'm planning. I can do this and then I'll go to the CEO and say, can I have £2 million and I'll give you a curation service. Of course that's not the tool for this and it's not at all obvious that that's the case when you first look at the main website…*

- Be able to run a report to help with a business case

  R4 makes a suggestion: *There is a thing about who is this for and how you would perhaps use it best. I think a lot of it is about making the case to someone or something and that comes back to the import and export functions. It would be very, very nice to be able to run a report because what I want to be able to do is either send a PDF or slap two sides of A4 on a table and say, this is where we need investment. This is where we're overspending. This is why we're doing this. This is why we're doing that. That might be actually quite a nice way of doing things.*

- Being able to calculate and compare salaries (FTEs).

  Both R4 and R1 mentioned this and others approved. R1 mentioned a tool called Sirius which could perhaps be plugged into CCEx. The discussion on the salaries issue arose both during Sarah's presentation and later on.

2.2.4 Comments on the interface, workflow

- We think it is fair to say that users in the given time were not able to create a mental model of how the tool worked. For example, for returning users it was still not clear how to add more costs (having to go to 'Compare costs' in order to add data didn't make sense either).

  R7: *I think one of the things if I was using this, one of the things I quite often do is add more data. So from the page at the beginning I have to go to compare costs to do that, which is not what I'm doing. I'm putting in data. I'm going to compare costs in a minute. Then I have to go to manage cost data sets to add another one. There isn't an add button here…*

  Facilitator2 suggests: *So in terms of coming back to what Catherine was saying and the value of having your own organisational data, would you prefer to have adding your own data as a separate field altogether and then maybe once you've got that then go into compare costs? Would that be a more…?*

  [Participants agree]

  R7 adds: *having said that there's not many buttons and I'm used to computers and things, I would just press all the buttons until I find it so I'm not too worried. It's not for completely naïve users. I'd work it out very quickly...*

  R4: *because I had the same problems as X, I wasn't actually sure what I was adding when I was adding a cost unit and why I was adding it or how I added it. So I just basically stabbed around randomly.*

  [People laugh]

- Sliders vs manual entry: during the hands-on session, R1 stated the sliders were useful but they would have also liked to be able to enter figures manually.
- It was not clear also how you could delete data, where was the delete button:

  R7: *The other thing I'd want to be able to do is delete... It's not under management, which is where I'd expect to be able to delete…*

  R4 says: *Well, I haven't even found that. Oh, there's delete…*

  [Everyone laughs]

- Registration: R6 failed to receive registration confirmation. Fortunately, Facilitator1 was able to speed it up manually. This issue had arisen before during one of the previous usability testing sessions, when one user had to wait at least 9 minutes for registration and was tempted to create another account.

3.0 Themes

- The scope of the tool—unclear
  - What the tool does, how it works, who it is aimed at
- Uploading and comparing costs a major issue
  - Worries about reliability of cost data entered (by multiple staff)
  - Staff time—who has time to enter cost data?
  - Also doubts about meaningfulness of such comparisons (especially global comparisons)
  - Suggestion for more questions on curation levels to make comparisons more significant
- Terminology not clear

- Graphs not clear
- Suggestions for improvements:
    - Being able to upload excel spreadsheet with organisation's data
    - Option to share costs internally before publishing
    - Add 'department' field
    - Being shown an example of how the tool works
    - Being able to run a report to help make a business case
    - Being able to compare and calculate salaries

4.0 Conclusion

The language employed by users demonstrates that it is not immediately obvious how CCEx works, what costs are being added and why and how to employ CCEx to obtain accurate cost figures. R7 says *It's not for completely naïve users. I'd work it out very quickly...* while R4 says *I wasn't actually sure what I was adding when I was adding a cost unit and why I was adding it or how I added it. So I just basically stabbed around randomly*.

Others described it as time consuming and this fact may affect the quality of data users will enter. R3 said: Also I think you'll probably all admit it's incredibly time-consuming... I'm not sure that, say nobody in the [organisation] could actually sit down and do it and if so, it would be quite time-consuming. So people will input all sorts of numbers.

R4 and R7 asked to have examples, case studies of how CCEx can help make a difference and R4 even mentioned a prize. They seemed sceptical that people will use the tool; they think serious incentives are needed.

The Focus Group was a great method for assessing CCEx. Visualising people in action highlighted difficulties people encountered both technologically and conceptually.

The interface in itself seemed easy to use. The two major obstacles encountered were the jargon and the fact that people failed to see adding costs was a two-step process i.e. adding cost data sets then cost units.

The main problem however, is at conceptual level. People find it difficult to build a mental picture of the process they have to go through. First of all, users do not think of budgets in terms of data sets and various units associated with it. Secondly, the categories used to breakdown costs do not make sense or, better said, they are presented in the wrong order. It is not the way users normally calculate costs and, after trying out the tool, they did not appear inclined to follow the CCEx way either.

Please note, the group understood the importance of sharing and comparing costs so they are not short on motivation. There is a certain reluctance to share data due to worries about confidentiality and reliability of data, both of which can be easily alleviated with careful communication.

Appendix J Focus Group 6 Report

Collaboration to Clarify the Cost of Curation

4C Focus Group on CCEx

15th January 2015 in London, UK

Report

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	Jan 15 2015	First draft	KH
0.02	Jan 19 2015	Edited	KH
1.0	Jan 21 2015	Finalised	KH

Attendees 4C:

4C:

- Luis Faria, KEEPS
- Neil Grindley, Jisc
- Katarina Haage, DNB
- Hervé L'Hours, UK Data Archive
- Paul Stokes, Jisc

External Participants:

- Matthew Addis, Arkivum
- Robert Bley, ExLibris
- John Kaye, Jisc
- Natasa Milic-Frayling, Microsoft Research
- Mike Quinn, Preservica

Meeting Details:

Date: 15th January 2015

Time: 01:00 - 04:00 pm

Location: Jisc, Brettenham House, 5 Lancaster Place, London, WC2E 7EN

Main objectives:

- To find out what users think of CCEx
- To identify what users want / expect from the tool
- To find out what the common problems are when using it
- To collect recommendations for improvement
- To find out if there are concerns about sharing cost data and if so, why. What could be done to mitigate these concerns?

Agenda:

- 13:00 Lunch
- 13:30 Welcome
- 13:40 Introduction to CCEx
- 14:00 Time for individual familiarisation with the tool
- 14:30 Focus group discussion on CCEx based on prepared questions
- 15:15 Coffee break
- 15:30 Continued group discussion on the CCEx
- 16:00 END

Questions for CCEx discussion:

- What are your genuine thoughts on the CCEx tool?
- Do you find it easy and intuitive to use?
- What do you like best/least about it?
- What are the barriers in using it?

- Are there redundant features?
- Is the language understandable or is there too much jargon?
- Would your organisation submit cost data? Who in your organisation would be responsible for submitting the data?
- What kind of support are you currently using for costing digital curation?

The objectives and questions should be ideally considered from a vendor's and digital preservation solution provider's perspective and also, if possible, mirror their users' view and requirements.

Minutes

Note: Because the presentation slides are available on the 4C website[28], this report focuses on the discussion during the meeting.

After a welcome by project coordinator Neil Grindley (NG) and a brief introduction round, NG gave an overview to the 4C project and the project results to date were presented. Following this Luis Faria (LF) gave a live introduction to the Curation Costs Exchange (CCEx) website, focusing on the Cost Comparison Tool (CCT): http://www.curationexchange.org/

The participants have browsed the website beforehand the meeting and had the chance to ask questions they had already formulated. After this initial Q&A session the participants went into one-on-on sessions to discuss the website and the tool based on the first use experience. Since the participants had prepared the questions (see above) the following discussion was lively.

The following highlight remarks re CCEx and CCT were made and discussed in plenary discussion round:

The overall impression of the website is positive; it is well designed, looks good and is easy to use; it is also easy to create a profile, and to add cost data sets and costs. Also the Open Source aspect and integration into working practices evoked interest. There was some criticism concerning the lack of knowledge what happens when the submission of costs was completed (missing explanation) and also the question of what the comparison factors are. There is a need of more cost data sets to be able to compare costs, however, the comparison of costs vs. value seems rather difficult in the first place since defining value is complicated because it is too multi-faceted for and in each organisation; quick innovation makes it even more difficult to define value and benefits. Nevertheless, the demand for a tool like the CCT is existent and "The CCEx has an important role in connecting the community" (says Natasa from Microsoft Research). At the moment the main objectives of the CCEx/ the submission of cost data are to share it with other (peer or non-peer) organisations to start an information exchange on costs in curation. It was mentioned that vendors may be prepared to add "standard" off the shelf costs (Matthew from Arkivum opened up his 3 sets in the meeting). They would feel much more comfortable if those sets had some form of quality metric attached (Mike from Preservica suggested digital preservation capability maturity model might form a framework for quality).

Wish list:

- value added features
- put more data sets online and for this purpose communicate the benefits to the stakeholders

[28] http://4cproject.eu/community-resources/focus-groups

- Extension of the platform and add a function to make it a planning tool
- Import/export of the data submitted (possible future feature?)
- Mirror the numbers that have been submitted in the end results
- Have a maturity slider?
- Add a pop-up window after “save and close” the cost data sets to tell the user his/her State of the Art (“Your costs are average/high/low etc.”)

Future possibilities for the CCEx:

- Integration in Data Management Plans
- Separation of costs is getting increasingly difficult as services become more integrated
- Connect vendors and academics through the CCEx to support the discussion (and challenge) about funding universities and other public research institutions
- CCEx as a “market place”; it could be the catalyst for a community to show there is a market
- Get the research councils / funders to either use it as a (closed?) cost tracking tool?
- It could be the catalyst for a community to show there is a market; a vote/up vote down aspect could make it essentially self-regulating with little moderation needed

Conclusions from participants:

- a certain maturity is necessary to use it as a market/comparison tool
- concentrate on memory institutions since no commercial organisation/customer will most likely ever share their costs
- Total cost of ownership (TCO) is highly important and should be mirrored in the tool
- Communicate clearly the value and benefit of using the tool is necessary
- To support the dialogue and exchange it is a good tool for the community, however, value and benefits must be added as factors

Towards the end of the meeting other 4C outputs such as the DCSM (former ESRM) and the Roadmap in general and the messages and actions for vendors/solution providers in particular were presented to the participants and gained much interest. Reactions to the Roadmap were to shrink the aims by 2 or 3 years and to look outwards; also take into account accreditation and benchmarks (what is good and/or bad accreditation). Vendors believe it will be problem solved in 3 years (they will be out of business as digital preservation will be integrated and therefore be more “business as usual” and there will no longer be a need for a specialist preservation unit.

Conclusions from 4C:

- The Preservica view (in particular) supports the general view that 'benefits' are more important than costs. We need to design a 'benefits comparison tool' to complement the CCT. (There's 4C mark2 (the Horizon 2020 sequel) just designed)
- There is potentially scope to provide private instances of the CCT for more localised comparisons
- We can and should generate costs data with publicly available pricing information from vendors (we would need to look carefully at this and work out how plausible the data ends up being. The endgame might be to compel vendors to engage with us because they would want to be accurately represented in the CCEx!)

In terms of further engagement activities the participants and project members agreed on future information exchange re the CCEx, CCT and the Roadmap.

Appendix K Summary of Engagement Activities

Date	Location/ Country	Title	Audience	Partner
1 Communication and information exchange with EC-funded and other projects and organisations				
10th April 2013	UK	APARSEN Meeting	APARSEN Project	Jisc
26th June 2013	UK	Issue 5—SCAPE Newsletter	SCAPE project	DPC
28th June 2013	Europe	Volume 2 Issue—TIMBUS Times	TIMBUS Project	DPC
13th June 2013	Europe	Sustainability & Cost Models for Digital Preservation	APARSEN Project	Jisc
6th September 2013	Lisbon, Portugal	APARSEN Presentation, 4C Workshop, iPRES: http://ipres2013.ist.utl.pt/ws4-4C-iPRES%20Workshop%20Agenda.pdf	Stakeholder Groups	ALL
2nd - 6th 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=WyJwb3N0ZXIiXQ== Poster Session: http://vimeo.com/74101447 Minute Madness: http://vimeo.com/74097837	Stakeholder Groups	Jisc
7th - 8th October 2013	The Hague, Netherlands	Koninklijke Bibliotheek meeting (KB—National Library of the Netherlands)	Project Team/ KB	Jisc, DNA
22nd 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
28th - 30th 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
4th November 2013	Cranfield, UK	ENSURE Meeting	Project Team/ Ensure	Jisc, DPC
7th - 8th November 2013	Geneva, Switzerland	CERN Meeting	Project Team/ CERN	DNB
29th 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
5th - 6th December 2013	The Hague	UNESCO, IFLA and ICA Roadmap meeting	Public	Jisc
24th February 2014	Online	MiLoS—Presentation of CCEx Mockups to the MiLoS Project Meeting	MiLoS Project	DPC
18th March 2014	Webinar	Meeting with AVpreserve	AVpreserve	DNB, DPC, KEEPS
25th March 2014	DNB, Frankfurt	Nestor—Presentation of the 4C Project and outline descriptions of key project deliverables	nestor working group on costs	DNB

Date	Location/ Country	Title	Audience	Partner
2nd April 2014	The Hague, Netherlands	SCAPE—Presentation of the 4C Project and the 'Cost Quest' to the SCAPE and Open Planet Foundation Seminar, The Hague	SCAPE and OPF	DANS
22nd May 2014	Webinar	Meeting with Bedern Group	Bedern Group	DNB, DPC
10th June 2014	Webinar	Roadmap	All Stakeholders	Jisc, DPC
17th June 2014	Webinar	Roadmap	All Stakeholders	Jisc, DPC
4th July 2014	Split, Croatia	ADA Summer School	All Stakeholders	Jisc
11th July 2014	European	'Presto4U' Friday Rewind: Learning and Teaching Repository Newsletter on and circulated to its membership by email	AV preservation	DPC
July 2014	UK	'Inspire' e-newsletter	Libraries	None
2 Stakeholder focus groups and workshops				
17th May - 21st June 2013	Europe	4C Initial Consultation	Stakeholder Groups	All
12th July - 31st July 2013	Europe	Invitations to 4C Workshop and Focus Groups	Consultation respondees	All
6th September 2013	Lisbon, Portugal	4C Workshop#1 and Focus Group #1, iPRES 2013 Conference	Stakeholder Groups	ALL
12th December 2013	London, UK	4C Focus Group #2 (Industry)	Industry stakeholder group	DPC
13th/14th January 2014	Geneva, Switzerland	4C Focus Group #3 at DPHEP costs and cost model workshop at CERN	Big Data stakeholder group	Jisc, SBA, DNA
24th February 2014	San Francisco, USA	4C Workshop #2 at IDCC 2014 Conference	All	Jisc, DCC, KEEPS, DNB
13th May 2014	Berlin, Germany	4C Focus Group #4 at ARCHIVING 2014	All	DNA, DNB, KBDK, UK Data Archive
6th October 2014	Melbourne, Australia	4C Roadmap Workshop at iPres 2014 Conference	All	Jisc, KEEPS, KBDK, DNB
3rd November 2014	London, UK	4C Focus Group #5	Researcher	DCC, DPC
17th - 18th November 2014	London, UK	4C/DPC Conference	All	All

Date	Location/ Country	Title	Audience	Partner
15th January 2015	London, UK	4C Focus Group #6	DP Vendors, solution providers	Jisc, KEEPS, UEssex, DNB
3 Advisory Boa rd Meetings				
11th June 2013	London	4C Advisory Board Meeting (1)	Advisory Board	All
20th January 2014	The Hague	4C Advisory Board Meeting (2)	Advisory Board	All
30th June 2014	Edinburgh	4C Advisory Board Meeting (3)	Advisory Board	All
4 Project Website a nd Blog posts				
13th March 2013	International	Blog post—Lets Collaborate!	Public	All
18th March 2013	International	Blog post—There's room for everyone @4C	Public	All
25th March 2013	International	Press—Press Release: New EU collaboration to clarify the costs (and benefits) of curation	Public	DPC
5th April 2013	International	Blog post—"The Age of Exploration and the Curation Costs Exchange" by Alex Thirifays	Public	DNA
14th April 2013	International	Blog post—'Digital curation cost models for everybody' by Sabine Schrimpf	Public	DNB
29th April 2013	International	Blog post—'Cache in the Attic' by William Kilbride	Public	DPC
1st May 2013	International	Blog post—Be part of the action—Collaborate with 4C and help to Clarify the Costs of Curation	Public	KEEPS
7th May 2013	International	Blog post—'A very pragmatic European enterprise—reflections on cross border project involvement' by Paul Stokes	Public	Jisc
2nd June 2013	International	Guest Blog Post—Digital Lifecycles and the Costs of Curation by Paul Wheatley	Public	Guest
9th June 2013	International	Blog post—'Collaborating our way to success' by Kathrine Hougaard	Public	DNA
14th June 2013	International	Blog post—'Call for Curation Cost Models' by Ulla Bøgvad Kejser	Public	KBDK
20th June 2013	International	Blog post—'Communication is key…' by Sarah Norris	Public	DPC
2nd July 2013	International	Engagement—The 4C consultation	Public	DPC
2nd July 2013	International	Outputs and Deliverables—D2.5—Project Communication Plan	Public	DPC
2nd July 2013	International	Outputs and Deliverables—MS7—Functioning Information Dependency Profile	Public	DNA
10th July 2013	International	Community Resources—4C Project Glossary	Public	DPC
10th July 2013	International	Blog post—"Nothing is Static" by Katarina Haage	Public	DNB
10th July 2013	International	Community Resources—4C workshop and focus group #1: "What does it cost?—EU to Assess the Cost of Digital Curation"	Public	DPC
18th 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
31st July 2013	International	4C Project Website	Public	DPC
2nd Augu st 2013	International	Community Resources—D2.1—Baseline Study of Stakeholder & Stakeholder Initiatives	Public	KEEPS
2nd Augu st 2013	International	Outputs and Deliverables—D3.1—Evaluation of Cost Models and Needs & Gaps Analysis	Public	
19th Augu st 2013	International	4C Focus Group Game	Stakeholder	DPC

Date	Location/ Country	Title	Audience	Partner
			Groups	
19th Augu st 2013	International	Blog post—'4C's Cost Model Evaluation' by Joy Davidson	Public	DCC
5th September 2013	International	Community Resources—T4.2—Draft Economic Sustainability Reference Model (Summary)	Public	Jisc
6th September 2013	International	Community Resources—MS9—Draft Economic Sustainability Reference Model	Public	Jisc
6th September 2013	International	Community Resources—IPRES Workshop	Public	All
9th September 2013	International	Current News—4C Project Wins Best Poster Award at iPRES 2013	Public	JIsc
9th September 2013	International	Blog post—'How to cut costs and keep the quality of service?' by Raivo Ruusalepp	Public	NLE
11th September 2013	International	Blog post—'What the 4C Project Learnt in Lisbon' by Neil Grindley	Public	Jisc
19th September 2013	International	Blog post—'The Case of the Curious Machine' by Sarah Norris	Public	DPC
21st September 2013	International	Community Resources—Related Projects: http://www.4cproject.eu/community-resources/related-projects	Public	KEEPS
7th October 2013	International	Community Resources—D4.1—A prioritised assessment of the indirect economic determinants of digital curation	Public	NLE
8th October 2013	International	Current News—4C Project submits Deliverable 4.1 to the European Commission	Public	DPC
9th October 2013	International	Blog post—'How to cut costs and keep the quality of service?' by Raivo Ruusalepp	Public	NLE
21st October 2013	International	Blog post—'The Future of Curation Costs' by Heiko Tjalsma	Public	DANS
28th October 2013	International	Guest Blog post—'The Carrot and the Stick' by Matthew Addis	Public	Advisory Board
8th November 2013	International	Blog post—'No such thing as free digital preservation' by Jan Dalsten Sørensen	Public	DNA
26th November 2013	International	Community Resources—Focus Group #2—London/Frankfurt	Public	DPC
4th December 2013	International	Blog post—Friends and family (or "Is there anybody out there")	Public	Jisc
17th December 2013	International	Blog post—'Please help us draw a map!' by Alex Thirifays	Public	DNA
18th December 2013	International	Current News—4C at the Ninth International Digital Curation Conference (IDCC)	Public	DPC
3rd February 2014	International	Blog post—'Lessons from the Half Way Point' by Neil Grindley	Public	Jisc
3rd February 2014	International	Outputs and Deliverables—D2.6—Report on Communications Activities	Public	DPC
10th February 2014	International	Current News—Three milestones met for the 4C Project's Curation Costs Exchange	Public	DPC
10th February 2014	International	Current News—4C Project submits Deliverable 3.1 to the European Commission	Public	DPC
14th February 2014	International	Blog post—'Whistle for the start of the second half' by Katarina Haage	Public	DNB
27th February	International	Outputs and Deliverables—D1.1—Draft Sustainability & Benefits Realisation Plan	Public	Jisc
27th February 2014	International	Blog post—'Uncertainty: the final frontier' by José Borbinha	Public	INESC-ID
27th February 2014	International	Outputs and Deliverables—D4.3—Quality and trustworthiness as economic determinants in digital curation	Public	NLE
3rd March 2014	International	Current News—Sustaining the Benefits of the 4C Project	Public	DPC
3rd March 2014	International	Current News—Crunching the Numbers and Comparing the Costs at Jisc Digifest 2014	Public	DPC
6th March 2014	International	Community Resource—Workshop #2—IDCC	Public	DPC

Date	Location/ Country	Title	Audience	Partner
6th March 2014	International	Outputs and Deliverables—D2.8—Curation Costs Exchange	Public	DPC
7th March 2014	International	Blog Post—'ANADP II Action Session—4C case Studies and Quantitative Data Session' by Neil Grindley and Raivo Ruusalepp	Public	Jisc, NLE
13th March 2014	International	Community Resources—Digifest14—Expert Speaker Session	Public	Jisc
13th March 2014	International	Current News—Slides available from Digifest14 Expert Speaker Session	Public	DPC
17th March 2014	International	Guest Blog post—'Crunching Numbers and Comparing Costs' by Matthew Addis	Public	Arkivum
19th March 2014	International	Current News—Quality and Trustworthiness as Economic Determinants in Digital Curation	Public	DPC
31st March 2014	International	Current News—'Excellent progress' for 4C Project at first year review	Public	All
7th April 2014	International	Blog post—'Models and Methods and Tools, Oh My!' by Hervé L'Hours	Public	UK Data Archive
7th May 2014	International	Current News—Breaking down walls in Digital Preservation: The OPF and SCAPE Seminar	Public	DPC
12th May 2014	International	Community Resources —Focus Group #3—DPHEP CERN	Public	DPC
19th May 2014	International	Community Resources—Focus Group #4—Archiving 2014	Public	DNA, DNB, UEssex
22nd May 2014	International	Current News—Save the Date—DPC and 4C 'Investing in Opportunity: Policy Practice and Planning for a Sustainable Digital Future' Conference on 17th—18th November 2014	Public	DPC
29th May	International	Current News—Share digital curation costs with the 4C Project to gain greater efficiencies	Public	DPC
29th May 2014	International	Blog post—'Valuable feedback from Berlin' by Alex Thirifays, Katarina Haage and Hervé L'Hours	Public	DNA, DNB, UK Data Archive
26th June 2014	International	Current News—'Investing in Opportunity: Policy Practice and Planning for a Sustainable Digital Future'—Conference registration now open	Public	DPC
26th June 2014	International	Blog post—'Digital curation buys us options—invest in opportunity' by Sarah Middleton	Public	DPC
26th June 2014	International	Community Resource—'Investing in Opportunity Conference'	Public	All
27th June 2014	International	'Investing in Opportunity: Policy, Practice and Planning for a Sustainable Digital Future' section was established on the website	Public	DPC
18th July 2014	International	Blog post—'Ready for take-off' by Alex Thirifays and Sarah Middleton	Public	DNA, DPC
31st July 2014	International	Outputs and Deliverables—D5.1—Draft Roadmap	Public	All
8th August 2014	International	Outputs and Deliverables—D3.2—Cost Concept Model and Gateway Specification	Public	UEssex
11th August 2014	International	Current News—Investing in Curation: A Shared Path to Sustainability	Public	DPC
11th August 2014	International	Current News—Curation Costs Exchange beta release—supporting smarter investments by comparing digital curation costs	Public	DPC
18th September 2014	International	Current News—Defining a Roadmap for Economically Efficient Digital Curation	Public	DNB
1st October 2014	International	Blog post—'3 (more) reasons to head to iPRES 2014' by Sarah Middleton	Public	DPC
6th October 2014	International	Community Resources—iPRES 2014 Workshop and Presentations	Public	DPC

Date	Location/ Country	Title	Audience	Partner
27th October 2014	International	Guest Blog post—'Collaborating on sustainable services for curation' by Matthew Addis	Public	Arkivum
4th November 2014	International	Community Resources—Focus Group # 5—Research Data Management	Public	DPC, DCC
7th November 2014	International	Outputs and Deliverables—D3.3—Curation Costs Exchange Framework	Public	DNA
7th November 2014	International	Current News—Understanding and comparing digital curation costs to support smarter investments	Public	DPC
11th November 2014	International	Blog post—'Trust, certification, sustainability and framework agreements' by Matthew Addis	Public	Arkivum
21st November 2014	International	Outputs and Deliverables—D4.4—Report on Risk, Benefit, Impact and Value	Public	INESC-ID
21st November 2014	International	Current News—Assessing the Risk, Benefit, Impact and Value of Digital Curation	Public	DPC
28th November 2014	International	Outputs and Deliverables—D4.5—From Costs to Business Models	Public	SBA
5th December 2014	International	Blog post—'The Curation Costs Exchange unveiled and challenged' by Alex Thirifays	Public	DNA
17th January 2014	International	Blog post—Zettabyting off more than we can chew by Paul Stokes	Public	Jisc
23rd January 2014	International	Blog Post—How time flies—Final Blog post by project coordinator Neil Grindley	Public	Jisc
28th January 2015	International	Blog Post—Shaping the Curation Costs Exchange: sharing your feedback by Magdalena Getler	Public	DCC
30th January 2015	International	Guest Blog post—Why Cost Models are Risky by Sean Barker	Public	BAE Systems
5 Social Media				
1st February – 31st January 2015	International	Posts made using the @4c_Project may be found at: https://twitter.com/4c_project	Public	DPC
6 Conferences and Events				
2nd 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
6th May 2013	Florence, Italy	'Economics of Digital Curation Training and Education' at the DigCurV Conference	Stakeholder Groups	Jisc
13th June 2013	Web	'Sustainability & Cost Models for Digital Preservation' at the APARSEN Training Webinar	Stakeholder Groups	Jisc
17th – 23rd August 2013	Singapore	Flyers distributed at IFLA	Stakeholder Groups	NLE
2nd - 6th September 2013	Lisbon, Portugal	iPRES Conference Workshop and Focus Group, Poster Session, Minute Madness	Stakeholder Groups	DNB, Jisc
23rd – 26th September 2013	Valletta, Malta	Flyers distributed at TPDL 2013	Stakeholder Groups	NLE
9th – 13th October 2013	Frankfurt, Germany	Flyers distributed at Frankfurt Book Fair 2013	Stakeholder Groups	DNB
22nd October 2013	Web	'Implementing Sustainable Digital Preservation at the PASIG ASIS&T Webinar	Stakeholder Groups	Jisc

Date	Location/ Country	Title	Audience	Partner
29th October 2013	Rome, Italy	'Parallel Track III - Policy & Sustainability Issues at the EUDAT 2nd Conference	Stakeholder Groups	UEDIN-DCC
7th November 2013	Nottingham, UK	Flyers distributed at the Business Archives Conference Council 2013	Stakeholder Groups	DPC
18th - 20th November 2013	Barcelona, Spain	Panel Session: Chaired by Neil Grindley (Jisc), the Resource Alignment panel "How can we employ the resources when we have available to us most economically to achieve our digital preservation objectives?" Action Session: 4C Case Studies and Quantitative Data	Stakeholder Groups	Jisc
28th November 2013	London, UK	Panel session on 'Sustainability' at the Institute of Historical Research	Stakeholder Groups	Jisc
19th March 2014	Aarhus	Flyers distributed at Public Digitization 2014 (national conference)	Stakeholder Groups	DNA
13th and 14th January 2014	Geneva, Switzerland	CERN DPHEP Workshop and 4C Focus Group as Satellite Event	Big Data Science	Jisc, DNA
24th February 2014	San Francisco, USA	IDCC 2014	IDCC Conference	DCC, Jisc, DNB
11-12th March 2014	Birmingham, UK	Jisc Digifest 2014	Higher Education	Jisc
25th March 2014	Dublin, UK	Representing the 4C project in several WGs and IGs at 3rd RDA plenary meeting	Stakeholder Groups	SBA
2nd April 2014	The Hague	Guest Talk in SCAPE + OPF Seminar	Stakeholder Groups	DANS
3rd April 2014	Web	WebEx presentation about Cost Models and Sustainability Strategies State Electronic Records Initiative - Council of State Archivists (US)	Stakeholder Groups	Jisc
7th April 2014	Copenhagen	Presentation of 4C at industrial partners' meeting in Danish National Archives	Stakeholder Groups	DNA
13th May 2014	Rome	Representing 4C at EuroCRIS 2014	Stakeholder Groups	DCC
13th May 2014	Berlin	Presentation of WP3 T3.2: Evaluation of Cost Models and Needs & Gaps Analysis; Focus Group on CCEx and CCM at IS&T Archiving 2014	Stakeholder Groups	KBDK
21st May 2014	Vienna	Presentation at 3rd LIBER Workshop on Digital Curation - Satellite Event	Stakeholder Groups	SBA
3rd June 2014	Toronto, Canada	IASSIST Conference	Memory Institutions	DCC
3rd June 2014	Copenhagen	Presentation of 4C at Nordic Conference National Archives, "NorDig"	Stakeholder Groups	DNA

Date	Location/ Country	Title	Audience	Partner
3rd June 2014	Bremen	Presentation abou t Costs in digital presevation - Projects, activities and approaches in cost modelling at German Librarian Day 2014	Stakeholder Groups	DNB
4th July 2014	Split	Presentation/workshop at ADA Summer School	Stakeholder Groups	Jisc
16th August 2014	Lyon	Flyers distributed, general engagement at IFLA Conference 2014	Stakeholder Groups	DNA, DCC, NLE
26th August 2014	Roosta	Flyers and roadmap distributed at Digital Cultural Heritage: Strategies for Added Value through Digitisation (national workshop)	Stakeholder Groups	NLE
5th September 2014	Budapest	Presentation and brief CCEx hands-on at 3rd INNET conference - Best practices in digital archiving of language and music data	Stakeholder Groups	DANS
8th September 2014	London	Representation of 4C at JCDL/TPDL	Stakeholder Groups	INESC-ID
16th September 2014	Karlsruhe	Flyers distributed at PASIG	Stakeholder Groups	DNB
22nd September 2014	Amsterdam	Flyers distributed at RDA Plenary 4	Stakeholder Groups	DANS
24th September 2014	Amsterdam	Distribution of flyers and promoting the CCEx and Roadmap at APA/RDA workshop	Stakeholder Groups	DNB
24th September 2014	Amsterdam	Flyers distributed at Data Seal of Approval Conference	Stakeholder Groups	DANS
25th September 2014	Amsterdam	Presentation of A comparison of existing cost models; The Cost Comparison Tool, Distribution of 4C flyers and CCEx flyers at EUDAT	Stakeholder Groups	DNA
6th October 2014	Melbourne	Roadmap Workshop, CCEx Poster Demo and CCM Paper at iPres 2014	Stakeholder Groups	Jisc, KEEPS, KBDK, DNB
9th October 2014	Tallinn	Flyers distributed at Nordic IUG meeting	Stakeholder Groups	NLE
9th - 10th October 2014	Stockholm	Flyers distributed and panel discussion at Bibliotheca Baltica	Stakeholder Groups	NLE
13th - 14th October 2014	Tallinn	Workshop, focus group, flyers at Meeting of national libraries of Estonia, Latvia and Lithuania	Stakeholder Groups	NLE
15th October 2014	Tallinn	Flyers distributed at Europeana Newspapers Project Information Day	Stakeholder Groups	NLE
21st October 2014	Brussels	Distribution of flyers and promoting the CCEx and Roadmap at APA	Stakeholder Groups	DNB

Date	Location/ Country	Title	Audience	Partner
22nd October 2014	Tallinn	Flyers distributed at Research and specialized libraries workshop	Stakeholder Groups	NLE
20th October 2014	Santa Maria, Brazil	Keynote and workshop at Brazilian Nacional Conference for Archives	Stakeholder Groups	KEEPS
5th November 2014	Copenhagen	Flyers, 4C and CCEx presented at Statens Arkivers konference om Forvaltningsetik og dokumentation i praksis	Stakeholder Groups	DNA
9th November 2014	Copenhagen	Presentation of CCEx Website at National Archives of Singapore - European Tour	Stakeholder Groups	DNA
13th November 2014	University of Essex	Integration of cost/4C topics into seminar at Consortium of European Social Science Data Archives (CESSDA) Expert Group Meeting	Stakeholder Groups	UEssex
14th November 2014	Karlsruhe	Roadmap promotion at Digital Archiving—Mission and strategies	Stakeholder Groups	DNB
17th - 18th November 2014	London	4C/DPC conference: presentation of all project results and outputs; Emphasis on Sustainability, Economics, Benefits (Roadmap, CCEx)	Stakeholder Groups	All
28th November 2014	Cologne	4C slides in broader DM presentation at Second data management workshop	Stakeholder Groups	DANS
4th December 2014	Riga	Flyers distributed at Meeting of national libraries of Estonia, Latvia and Lithuania	Stakeholder Groups	NLE
7 Publications				
7th 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
23rd March 2013	Denmark	Press release on the EU project 4C http://digitalbevaring.dk/pressemeddelelse-4c/	Public	KBDK/ DNA
25th 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
25th 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
25th 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
25th March 2013	UK	Collaboration to Clarify the Costs of Curation (4C) http://www.dcc.ac.uk/projects/4c	Public	DCC
25th 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
25th 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
27th 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
1st 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
1st 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
18th April 2013	Germany	PROJECT STARTED—4C http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/Kurzmeldungen/4C.html	Public	DNB
18th April 2013	Germany	PROJECT STARTED—4C http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/Kurzmeldungen/4C.html	Public	DNB/ Nestor
1st 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
17th May 2013	UK	The 4C consultation is now open! http://www.dpconline.org/newsroom/latest-news/1014-4c-consultation-now-live	Public	DPC
20th 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
26th 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
26th 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
28th June 2013	Europe	European Partnerships in TIMBUS Times:	TIMBUS	DPC

Date	Location/ Country	Title	Audience	Partner
		http://timbusproject.net/about/publicity-material	Readership	
2nd 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
13th July 2013	US	4C - Collaboration to Clarify the Costs of Curation http://www.dlib.org/dlib/july13/07inbrief.html	D-Lib Readership	DPC
10th August 2013	Germany	"Nothing is static" in "Dialog mit Bibliotheken": http://www.dnb.de/DE/Service/Publikationen/dialog201302.html	Stakeholder Groups	DNB
10th November 2014	UK	Curation Costs Exchange: Supporting Smarter Investments in Digital Curation: http://www.educause.edu/ero/article/curation-costs-exchange-supporting-smarter-investments-digital-curation	Stakeholder Groups	DPC
24th November 2014	US	Curation Costs Exchange: Supporting Smarter Investments in Digital Curation as Blog Post on “Digital Preservation matters”: http://preservationmatters.blogspot.nl/2014/11/curation-costs-exchange-supporting.html	Stakeholder Groups	External

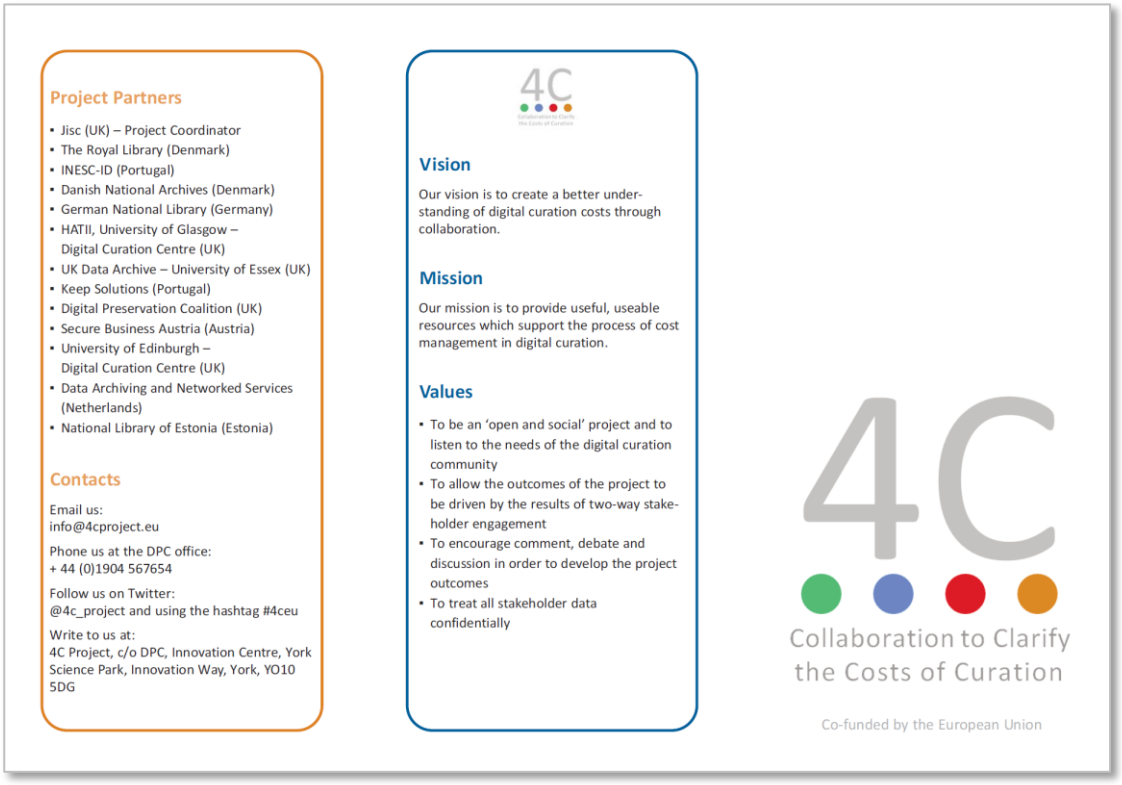
Table 6—Summary of Engagement Activities

Appendix L Printed Information and Materials

4C Flyer

English

http://www.4cproject.eu/component/docman/doc_download/95-4c-flyer-english



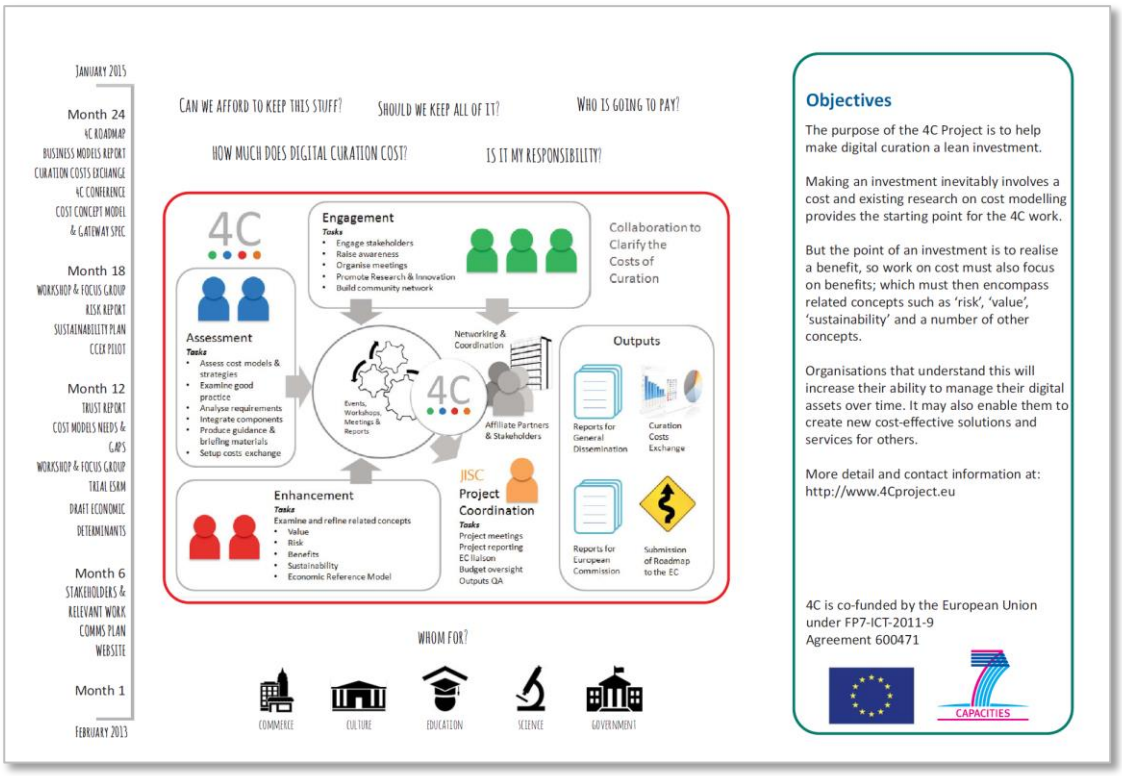


Figure 13—4C Flyer—English

German

http://www.4cproject.eu/component/docman/doc_download/96-4c-flyer-german

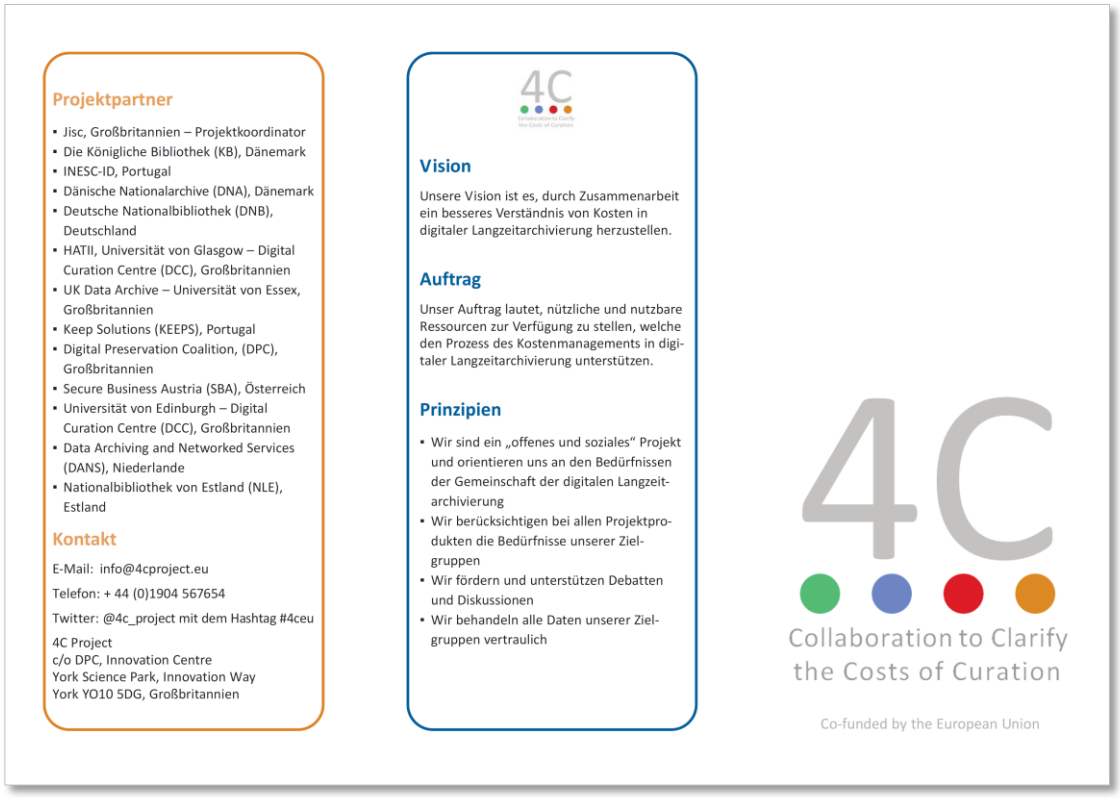


Figure 14—4C Flyer—German

CCEx Flyer

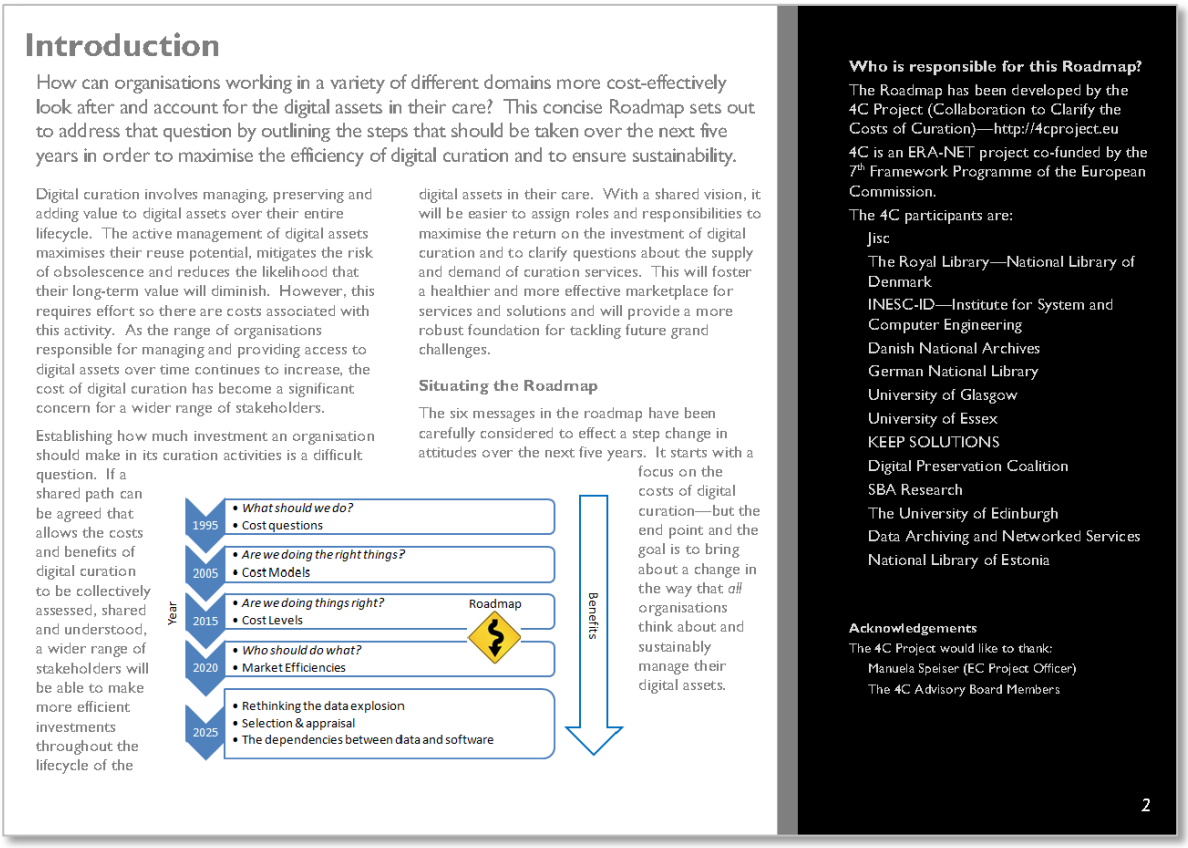
http://www.4cproject.eu/component/docman/doc_download/94-ccex-flyer

Figure 15—CCEx Flyer

Roadmap

Full roadmap

http://www.4cproject.eu/component/docman/doc_download/58-d5-1-draft-roadmap



Introduction

How can organisations working in a variety of different domains more cost-effectively look after and account for the digital assets in their care? This concise Roadmap sets out to address that question by outlining the steps that should be taken over the next five years in order to maximise the efficiency of digital curation and to ensure sustainability.

Digital curation involves managing, preserving and adding value to digital assets over their entire lifecycle. The active management of digital assets maximises their reuse potential, mitigates the risk of obsolescence and reduces the likelihood that their long-term value will diminish. However, this requires effort so there are costs associated with this activity. As the range of organisations responsible for managing and providing access to digital assets over time continues to increase, the cost of digital curation has become a significant concern for a wider range of stakeholders.

Establishing how much investment an organisation should make in its curation activities is a difficult question. If a shared path can be agreed that allows the costs and benefits of digital curation to be collectively assessed, shared and understood, a wider range of stakeholders will be able to make more efficient investments throughout the lifecycle of the digital assets in their care. With a shared vision, it will be easier to assign roles and responsibilities to maximise the return on the investment of digital curation and to clarify questions about the supply and demand of curation services. This will foster a healthier and more effective marketplace for services and solutions and will provide a more robust foundation for tackling future grand challenges.

Situating the Roadmap

The six messages in the roadmap have been carefully considered to effect a step change in attitudes over the next five years. It starts with a focus on the costs of digital curation—but the end point and the goal is to bring about a change in the way that *all* organisations think about and sustainably manage their digital assets.

Who is responsible for this Roadmap?

The Roadmap has been developed by the 4C Project (Collaboration to Clarify the Costs of Curation)—http://4cproject.eu

4C is an ERA-NET project co-funded by the 7th Framework Programme of the European Commission.

The 4C participants are:

- Jisc
- The Royal Library—National Library of Denmark
- INESC-ID—Institute for System and Computer Engineering
- Danish National Archives
- German National Library
- University of Glasgow
- University of Essex
- KEEP SOLUTIONS
- Digital Preservation Coalition
- SBA Research
- The University of Edinburgh
- Data Archiving and Networked Services
- National Library of Estonia

Acknowledgements

The 4C Project would like to thank:

- Manuela Speiser (EC Project Officer)
- The 4C Advisory Board Members

2

The Vision

In five years time (2020) it will be easier to design or procure more cost effective and efficient digital curation services because the costs, benefits and the business cases for doing so will be more widely understood across the curation lifecycle and by all relevant stakeholders. Cost modelling will be part of the planning and management activities of all digital repositories.

Who should be interested?

Curation Practitioners

Those with direct responsibility for managing digital assets and appropriate knowledge about digital curation processes and techniques.

For example: digital curators, digital preservation officers, digital archivists, records managers and digital repository/data/collections managers with enough technical expertise to assume responsibility for the long-term management of assets.

Curation Researchers

Those with the remit and the expertise (or the appropriate guidance) to tackle emerging digital curation challenges and to define new methods and processes for the long-term management of digital assets.

For example: university research teams, research teams in larger memory institution, funded research consortia, research arms of commercial entities (e.g. Microsoft, Google, IBM).

Data Users (and re-users)

Those with an interest in using and re-using the curated data. Also known as the 'designated community' when it comes to determining why and for whose benefit investment is being considered to curate the digital assets.

For example: data scientists, researchers, cultural heritage professionals, authors, analysts, media and broadcast organisations, and any data-consuming business.

Managers (and financial officers)

Those within organisations or groups that have little or no digital curation expertise themselves but are required to integrate, coordinate, facilitate or manage digital curation activity as an integral part of the business function of the organisation.

For example: heads of library and information systems, IT managers, finance managers, administrators,

Member Organisations

Those who represent the interests of subscribing member organisations and the wider community to promote and support best practice and policy-making in the domain of digital curation or in related areas.

For example: Alliance for Permanent Access, Archives & Records Association (UK), Digital Preservation Coalition, International Council on Archives, International Federation of Library Associations, LIBER, Nestor, Netherlands Coalition for Digital Preservation, Open Planets Foundation

Policy Makers (Resource Providers / Data Owners)

Those with responsibility for dictating the type and quality of digital curation activity that is required; those responsible for making the resources available to support that activity (funding); and those responsible for establishing the framework of ownership around data.

For example: research councils, funding agencies, government departments, charitable bodies, senior information risk owners, , publishers, and any senior management within data dependent corporations.

Solution Providers

Those with incentives (commercially or community-driven) to develop and disseminate products that will support digital curation activity at either the infrastructure (services) or systems (solutions) level.

For example: Archivematica, Arkivum, CERN, DuraSpace, Ex Libris, LOCKSS, OCLC, Portico, Tessella.

4

Identify the value of digital assets and make choices

1: Identify the value of digital assets and make choices

What the message means and who should act

Not all digital objects are digital assets. Only those which store value and will realise future benefit can be described as assets. Those which won't are liabilities. Trying to distinguish these is difficult but it is no harder than the many other business decisions that organisations are faced with on a regular basis. And although it might seem cheaper to preserve everything than to spend time doing this selection, such an approach is unlikely to be sustainable or result in assets that are findable, understandable and reusable.

> *"One in five of the UK's largest companies now measure the value of corporate data on their balance sheets. Businesses realise that finding better ways of analysing data is the key to unlocking [their] profitability"*
>
> Alwin Magimay,
> KPMG UK Head of Digital and Analytics

It has long been true, but appraisal and selection of valuable assets is of increasing relevance given the upward curve of data creation. Even for organisations that have explicit—limiting—policies on the types or quality of the assets that they manage, budgeting for the curation of rapidly increasing volumes is a challenge.

The resources available to ensure long-term availability of data are unlikely to grow at the same rate as data volumes. Secondly, despite the long-standing tradition of human appraisal of assets (i.e. deciding what to retain), for many organisations data has grown to such an extent that it is no longer feasible for this to be done by a person. Appraisal has to be (at least) semi-automated to be scalable and "value" is an essential concept that will need to be algorithmically defined.

Designing how human appraisal knowledge and skills can be combined with machine-based appraisal to result in semi-automated decision making process is a major topic for research.

However, some key aspects can be identified:

- Value is an indirect economic determinant on the cost of curating an asset. The perception of value will affect the methods chosen and how much investment is required. That perception is best established by the designated community for whom the asset is being curated.
- Content owners should have clear policies regarding the scope of their collections, the type of assets sought, the preferred file formats. They must also identify the designated community using the assets and monitor usage intentions over time. From this, decisions can be made about which properties or attributes of the asset should be prioritised for preservation.
- Establishing, formalising and codifying value criteria for assets requires active effort and should be a costed component of curation. This should be done in conjunction with an understanding that certain types of assets can be re-generated or re-captured relatively easily, thereby avoiding curation costs

Establishing 'value' is a challenging exercise. The myriad contexts in which organisations operate and the differing perceptions of stakeholders about the current and potential use cases for digital assets makes the concept difficult to quantify and difficult to compare. A mixed approach, however, in which automated appraisal leads to selection advice for the human expert would mean an important reduction of workload during appraisal and selection.

6

1: Identify the value of digital assets and make choices

Benefits and positive outcomes

- Focusing on the value of digital assets and their likely return on investment will foster a deeper sense of tactical and strategic alignment at all levels within an organisation.
- Questions will usefully arise about whether existing data and digital collections are being used, have potential users, are being adequately exposed or are sufficiently discoverable.
- The effort to automate the identification of value could be combined with improving the overall efficiency of curation systems (see message 2).
- Similarly, the information that must be explicit for automated appraisal will also be valuable when digital repositories seek to validate their procedures.
- Co-operation and collaboration between organisations will become more commonplace as organisations work together to effect 'handoffs' of data and agree long-term archiving arrangements.
- An investment now into research relating to automated selection and appraisal techniques will lay the groundwork for increasingly sophisticated and critical work beyond 2020 when global data volumes dwarf current levels.
- The articulation of demand for automated selection and appraisal products will drive solution provider activity and provoke action within the marketplace to supply that demand.

Actions

Who	What	When 2015	2016	2017	2018	2019
Curation Practitioners	Lobby management into proper resourcing of selection and appraisal practice and focus on cost-effective digital curation activity	●	●			
Curation Researchers	Conduct research into automatic appraisal and selection techniques based on codified value criteria	●	●	●		
Data Users	Content experts to work with technologists to establish value criteria and represent 'designated communities'	●	●			
Managers	Incorporate the concept of 'value' into strategic and tactical decision-making		●	●	●	
Member Organisations	Help establish relationships between organisations to facilitate the transfer or 'handoff' of digital assets	●	●	●	●	●
Policy Makers	Establish requirements for digital asset value assessment as part of data management and curation planning			●	●	●
Solution Providers	Build on existing tools (e.g. file format registries) to provide automated selection & appraisal tools		●	●	●	

7

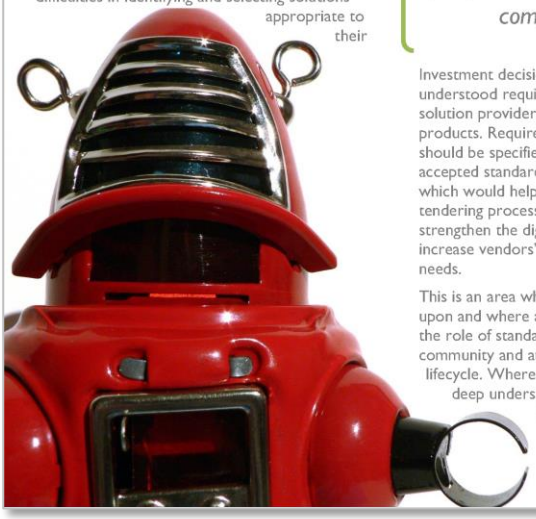
Demand and choose more efficient systems

2: Demand and choose more efficient systems

What the message means and who should act

The concept of supply and demand is a fundamental economic principle and should underpin decisions about service design, business modelling and sustainability. In a fully functional marketplace, a clearly articulated demand will be met by a competitive range of solutions, at least one of which should be able to meet or even surpass the specification and do so at an affordable price.

Digital curation remains an immature market for systems and some data managers report difficulties in identifying and selecting solutions appropriate to their organisational requirements. The question arises, is this a supply-side or a demand-side problem? If the supply is not adequately responding to demand, one practical response is to look closely at how that demand is being articulated and whether there are ways it can be simplified, amplified or just expressed more clearly.



A common understanding and clear specifications are prerequisites for a competitive market

Investment decisions should be based on well understood requirements which in turn will allow solution providers to supply new or enhanced products. Requirements for curation services should be specified according to a range of widely accepted standards or established best practices which would help to encourage competitive tendering processes. Standardisation would strengthen the digital curation market and increase vendors' responsiveness to curation needs.

This is an area where existing practice can be built upon and where a more uniform understanding of the role of standards is needed across the community and at all stages of the digital asset lifecycle. Where organisations already have a deep understanding of developing and implementing standards (and of procuring and implementing digital curation solutions) this expertise should be sought, synthesised and disseminated for the benefit of other types of stakeholders.

Information about existing well-established methods and concepts should be made as accessible as possible and might include plain-language guidance or simple implementation tools that address such topics as: risk management (ISO 31000), information security (ISO 27001), records management (ISO15489), digital preservation (ISO 14721), or digital repository trustworthiness (ISO 16363).

A common understanding and clear specifications are prerequisites for a competitive market and this can also be fostered by adopting good practice approaches as well as adherence to formal standards. Third party formal certification of services and systems may helpfully increase the comparability of products but well designed and widely endorsed self-assessment tools, or peer-reviewing will also help to improve knowledge across a variety of domains and allow a broader range of stakeholders to better understand the types of systems they should be seeking to procure and implement.

More knowledgeable customers demanding better specified and standardised functionality will mean that products can mature more quickly. It is this transaction that will over time create a virtuous circle of supply and demand and result in more effective and efficient systems.

9

2: Demand and choose more efficient systems

Benefits and positive outcomes

- Digital curation requires a significant investment of time in order to acquire expertise. The spread and adoption of standardised practices helps to lower the barriers to entry for new practitioners.
- Standardisation supports easier institutional decision making and will add efficiencies to operational environments.
- Standardisation may not be possible or applicable for institutions with unusual or unique digital holdings, but may still provide practical reference points for customisations and extensions.
- For institutions where curation is not their core business best effort approaches are often sufficient to address their needs. Being clear about where local practice deviates from standard practice and documenting the reasons in policies will be helpful in maintaining an effective operating environment and to align understanding (within the organisation) of current capability and the organisational mission.

Actions

Who	What	When 2015	2016	2017	2018	2019
Curation Practitioners	Establish a common understanding of curation. Share experiences and empirical evidence about tools & methods to provide institutions with baseline curation requirements	●	●			
Curation Researchers	Undertake research work to minimise subjectivity and clarify & standardise definitions of benefits. Develop tools that facilitate the implementation of standards	●	●			
Data Users	Demand better and more standardised interfaces to data and metadata making data more usable and thus demonstrating its value	●	●			
Managers	Setup agreements between organisations to share infrastructure for more efficient utilisation of available resources	●	●	●		
Member Organisations	Evangelise for the standardisation of practice across domains and produce advice & guidance that will help organisations to act upon this message. Work with solution providers & customers to translate and improve system specifications	●	●	●		
Policy Makers	Promote good practice and training so that integrated and standardised digital curation tools and services have a higher profile			●	●	●
Solution Providers	Work with customers and the community to develop, explain and simplify standard practices. Meet customers half-way in specifying solutions and by making pricing models and implementation options clear & understandable	●	●	●		

10

Develop scalable services and infrastructure

3: Develop scalable services and infrastructure

What the message means and who should act

While some organisations will need to be able to provide intensive curation services, others may only need to provide basic functions. To help organisations develop sustainable business models that fit their particular needs they need to understand what drives their investment and where it will have the most impact.

This will require decisions around appropriate infrastructure—not only in terms of hardware and software—but also in terms of the skills and resources that can be employed within the organisation. Optimising the impact of investments may be achieved through:

- Information and knowledge exchange, including cost data, to enable the identification of opportunities for improved efficiencies
- Sharing infrastructure, resources and effort among complementary institutions

It may also require a high level of commitment to collaboration and a realisation that retaining effective local control might mean letting go of some tasks and commissioning external parties to do things more efficiently on a contractual basis. This feeds into a wider issue around maturing strategy and practice right across the digital curation domain.

The switch to collaboration, sharing information and sharing resources to manage budgets for digital curation may be easily justified in financial terms. Nevertheless a programme of "education" and "culture change" is required to encourage this approach.

It may be possible to do this from the "top-down" and from the "bottom up":

- Mature national and international support networks, with endorsements from national sector leaders and funders, mentoring less mature or less well equipped organisations, facilitating lessons learned and identifying opportunities for further sharing and collaboration.
- Local or sectoral organisations actively seeking peers and establish platforms for information exchange and the sharing of resources.

It is realistic and prudent to assume that curation budgets are unlikely to be raised in line with the enormous growth in volumes of content, so investment needs to be strategically targeted to the right places to create economies of scale and scope. Where organisations have sufficient resources, capability and need to design their own infrastructure, additional budget must be found for ensuring that evaluation, advocacy and sustainability planning are built into the ongoing cost of maintaining the infrastructure.

> *"Collaborating & sharing infrastructure, resources & effort is a valuable approach for local institutions who want to improve their digital curation practices on limited budgets."*
>
> Matt Greenhall,
> Programmes Manager at The National Archives

Whether organisations are reliant on local or external curation infrastructures, they should all be aiming to work smarter every year and should be able to demonstrate the impact of their investments year on year.

This will remain the case all the way up the infrastructure stack towards national and international provision of infrastructure. The measures of effectiveness may change radically depending on the context but the need to demonstrate the efficiency and effectiveness of investment remains constant.

12

3: Develop scalable services and infrastructure

Benefits and positive outcomes

- Collaborating this way opens a forum for mentoring, knowledge exchange, application of standards and continuous development; reducing the "gap" between organisations seen as more "mature" in the field of digital preservation and those who are relatively new to the practice.
- An assessment of local capability versus outsourcing shines a light on skills gaps within the organisation and should highlight training and staff development opportunities.
- Organisations will be able to identify opportunities for the introduction of cost efficiencies by comparing their own activities with those of similar organisations.
- Shared infrastructure, resources and effort will also enable the realisation of further cost reductions by improving efficiency of the workflows necessary to undertake digital curation.

Actions

Who	What	When 2015	2016	2017	2018	2019
Curation Practitioners	Make realistic assessments of institutional capability to provide scalable services & infrastructure and compare this with the cost effectiveness & suitability of external service provision		●	●		
Curation Researchers	Optimise workflows and design procedures that will handle large volumes and complex digital objects	●	●	●		
Data Users	Demand delivery of assets and access to resources that suit the needs of users rather than fit within the constraints of current services and infrastructure	●	●	●	●	●
Managers	Setup agreements between organisations to share infrastructure for more efficient utilisation of available resources. Support practitioners to make realistic assessments of local capability	●	●	●		
Member Organisations	Identify and share lessons learnt relating to the economic benefits of using shared infrastructures and the value of planning for scalability over time. Provide a neutral environment to build trust for the negotiation of sharing agreements.	●	●			
Policy Makers	Provide domain-wide shared infrastructures to exploit economies of scale	●	●	●		
Solution Providers	Pay close attention to the need to build scalability into services. Offer solutions that are vigorously tested and provide transparent, benchmarked performance in response to more sophisticated specifications	●	●	●	●	

13

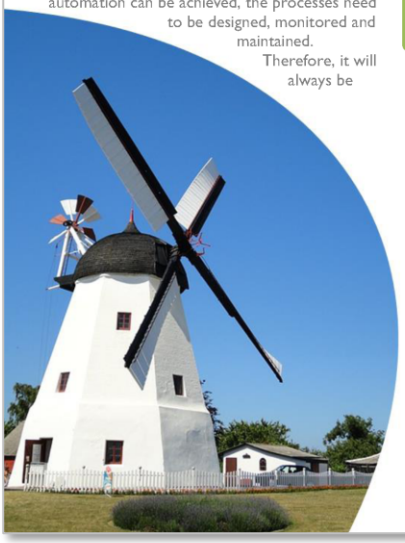
Design digital curation as a sustainable service

4: Design digital curation as a sustainable service

What the message means and who should act

In most cases, the ongoing management of digital assets will be an accumulative challenge as new data is added to existing managed information environments. Even where strict retention schedules dictate that obsolete data is deleted in a timely fashion, the trend of information growth will be upward.

Also, effective digital curation requires active management throughout the whole lifecycle of a digital object. 'Active' implies effort . Even where automation can be achieved, the processes need to be designed, monitored and maintained. Therefore, it will always be necessary to find resources to fund curation, and the level of resources required will need to be regularly reviewed.



Whilst the likelihood in most organisations is that the amount of digital assets that need curating will steadily (or even dramatically) increase over time, it must also be understood that solutions and processes can be employed more effectively and efficiently over time to keep pace with or even overtake resourcing requirements.

Curation should be undertaken with a stated purpose in mind

But this can only be achieved with a purposeful focus on planning for increased scale (see message 3) and by anticipating - on a regular basis - the need to enhance and mature the current curation environment.

This drive towards maturity is often characterised in practice by a shift from ad hoc or reactive activities towards a situation where curation is planned into the organisational culture and becomes a service-type activity.

As well as implying a planned and continuous provision of capacity and capability. The transactional nature of the work illustrates the supply-side and a demand-sides of service provision and consumption.

Curation should be undertaken with a stated purpose. Even in cases where there is no formal requirement for a business model to be declared, understanding who requires it to happen is fundamental to arguing the case for resources to support it.

The designation of curation as a service further embeds the activity into the normal business function of an organisation. As part of the infrastructure of an organisation, a curation service becomes as necessary and unremarkable as the human resources section or the estates department and relies on similar levels of mature alignment of practice across organisations and across sectors. It also implies that the mechanisms and systems used to curate digital assets should be interoperable, joined up and easily scalable.

Where the provision of a curation service within the organisation is not viable or practical, services must be easily procurable from outside the organisation. This requires structural services offering competitively priced and appropriate digital curation capability to be available.

15

4: Design digital curation as a sustainable service

Benefits and positive outcomes

- By considering curation activity in terms of service provision, organisations will be able to specify more clearly the costs of digital curation and better establish their incentives and the methods they should use to manage their digital assets over time.
- Rather than digital curation happening within the organisation as an ad-hoc activity or an accidental adjunct to other tasks, it will become a strategic business function, underpinned by appropriate cost/benefit analyses.
- The result of those analyses will provide a clearer view and a better understanding of the value of digital assets and will help to refine the mission and objectives of an organisation
- Designing digital curation as a service should help to make activity more comparable across all sectors and should help to align and standardise practice.
- This in turn should promote the market for the provision of solutions and services and should lead to a wider range of competitively priced offerings from a broader range of suppliers.

Actions

Who	What	When 2015	2016	2017	2018	2019
Curation Practitioners	Work with digital curation service consumers (users) to model the current costs and benefits of digital curation activity	●	●			
Curation Researchers	Continue research into sustainable business models and examine how to standardise divergent current practices	●	●			
Data Users	Methodically and empirically assert the value of digital assets and work with practitioners and managers to undertake cost/benefit analyses		●	●	●	
Managers	Seek proof that digital curation activity within the organisation is: optimally & sustainably resourced; works within a defined supply & demand framework; is providing an efficient & effective service		●	●		
Member Organisations	Provide practitioner advocacy material to promote activities within organisations. Help solution providers to publicise & promote their offerings to enhance the marketplace for services & solutions	●	●	●	●	●
Policy Makers	Provide domain-wide shared infrastructures to exploit economies of scale. Design funding constraints to ensure that sustainable digital curation is underpinned by proven cost-effectiveness	●	●	●		
Solution Providers	Participate in setting standards and focus on long-term interoperability of design in software & infrastructure. Focus on openness & collaboration and building a sustainable & inclusive market place	●	●	●	●	

16

Make funding dependent on costing digital assets across the whole lifecycle

5: Make funding dependent on costing digital assets across the whole lifecycle

What the message means and who should act

Digital curation activity requires a flow of resources and whether that means salaries, skills acquisition, building infrastructure or systems procurement, a resource provider must make a commitment to provide sufficient resources for that activity to proceed.

Many sectors call these resource providers 'funders' and the most straightforward implication of this message would be to recommend that funds are not awarded to initiatives (e.g. research projects, development projects) that aren't able to give a plausible estimate of how much it will cost to sustain and make available the data they will be funded to create.

For this message to have broad applicability the term 'funder' needs to be widely defined as does the timescale for funding. Some digital assets may need to be preserved in perpetuity but others will have a much more predictable and shorter life-span.

The overall message should, therefore, be understood as being very context sensitive and particularly aimed at situations where a demonstrably efficient use of funding is an important principle and a critical component of any case that is made for sustaining assets into the future.

> *Digital curation activity requires a flow of resources to support it*

All stakeholders involved at any point in the curation lifecycle will need to understand their fiscal responsibilities for managing and curating the asset until such time that the asset is transferred to another steward in the lifecycle chain.

Using the management of research data as an example:

- Universities and researchers need to be able to estimate the cost of curating research data during the active phase of the research project and be able to request all or some of these costs to be covered in new grant applications.
- Data centres need to be able to assess the costs associated with the long-term retention of data beyond the life of the project along with requirements relating to access and functionality (e.g., restricted access, specific software required to render, analyse and/or manipulate the data).
- Re-users of data may need to understand if there are any costs associated with access and reuse of in new data intensive activities.

In all domains organisations have to operate within funding constraints and the general principle of anticipating costs as much as possible in advance tends to appeal to budget holders and resource providers everywhere. What will also be necessary is for those resource providers to have a way of assessing whether the requested costs are reasonable and for it to be clear that the benchmarks and costing practices being used by those seeking funds are legitimate.

18

5: Make funding dependent on costing digital assets across the whole lifecycle

Benefits and positive outcomes

- With more clarity on the costs associated with each stage of the curation lifecycle, transfers of assets from one managed environment to another are likely to be handled more smoothly.
- The ability to make realistic estimates of future liabilities will integrate digital asset management more firmly into the ordinary planning activities of organisations.
- This in turn will raise awareness of the value and importance of digital assets and may prompt an increased desire to exploit that value creatively.
- Resource Providers will be better placed to identify areas where centralised support may realise greater curation efficiencies, potentially leading to more shared infrastructure becoming available.
- A focus on lifecycle costs may incentivise organisations and resource providers to avoid re-creating data that already exists, or to create data in such a way that the prospects for its sustainability and reusability are optimised from the outset.
- This, in turn, may positively affect the quality of data created allowing re-users to have greater confidence in the data they use and, subsequently, produce more useful results.

Actions

Who	What	When 2015	2016	2017	2018	2019
Curation Practitioners	Collaborate with peer organisations and engage with tools to establish the cost and benefits of digital curation. Be prepared to clarify whole lifecycle costs for managing digital assets	●	●	●	●	
Curation Researchers	Further develop resources that will simplify cost modelling & comparison for digital curation. Engage in additional pathfinder research to refine methods & decrease costs	●	●			
Data Users	Work with practitioners, researchers & policy makers to establish a better understanding of the variable asset value across the digital lifecycle & the impact of digital curation on that value		●	●	●	●
Managers	Establish clarity within organisations about roles & responsibilities for costing curation & resource it appropriately. Provide additional training for finance & accounting staff to understand digital asset management budgeting issues	●				
Member Organisations	Help establish relationships between organisations to facilitate the transfer or 'handoff' of digital assets. Promote tools & methods for whole lifecycle costing and disseminate good practice	●	●	●	●	●
Policy Makers	Identify where the maintenance of digital assets is a priority & design clauses in support agreements that require an estimation of the whole lifecycle costs of sustaining the assets for as long as they may be needed			●	●	●
Solution Providers	Work with practitioners and researchers to build accounting and budgeting modules into curation systems		●	●		

19

Be collaborative and transparent to drive down costs

6: Be collaborative and transparent to drive down costs

What the message means and who should act

Drivers for managing and curating digital assets vary greatly between stakeholders, but essentially each is looking to realise a return on their investment—either through mitigation of risk or through derived benefits. Comparing operational costs and effort with peers is essential for identifying where efficiencies and savings can be made and to improve an organisation's ability to make informed investment decisions. The only way organisations can compare costs is if they– and others–are prepared to be transparent about their costs.

> *"The Curation Costs Exchange (CCEx) will help funders realise the benefit of their investments. By being transparent about their costs and plugging them into this platform, projects can demonstrate that the taxpayer is getting value for money."*
>
> Ron Dekker, the Netherlands Organisation for Scientific Research (NWO)

Whilst transparency of cost data is urgently needed, it must in some cases be anonymised, and properly contextualised. This might include information about: the environment in which the costs were incurred; the assumptions of quality and trustworthiness of curation that have been made ; the complexity of the objects being managed; the scale of working; and a host of other issues that will allow proper interpretation of the overall value of the investments that have been made.

Up until now, there have been no mechanisms to help stakeholders find out what their peers are spending, to share their own cost data and to provide contextual information to better identify risks and benefits. The Curation Costs Exchange (CCEx) has been developed to address this problem.

The critical issue for the CCEx – and for the whole concept of being transparent about costs – is that collaboration is *key* and requires trust between the parties sharing their information.

Institutions where digital curation is a significant part of their core business, e.g. national memory institutions or large content-rich organisations, may not only already have some experience of trying to cost curation but may also have a publicly funded mandate to be transparent and accountable. Where this is the case, those types of organisations may be able to take a lead and start sharing existing data (anonymised if necessary).

In return, those organisations can expect information that will immediately help them to optimise their investments. In addition policy makers should promote and support a culture of sharing cost data, then it should be possible to build a critical mass of data relatively quickly that would be of benefit to all.

If those who provide digital curation services can be contextually descriptive about their products and transparent about their pricing structures, this will enhance possible comparisons, drive competitiveness and lead the market to maturity.

If a whole range of organisations creating and managing digital assets can share emerging cost data and contextual information, this will help everyone to identify points in the curation lifecycle where efficiencies and savings can be realised.

6: Be collaborative and transparent to drive down costs

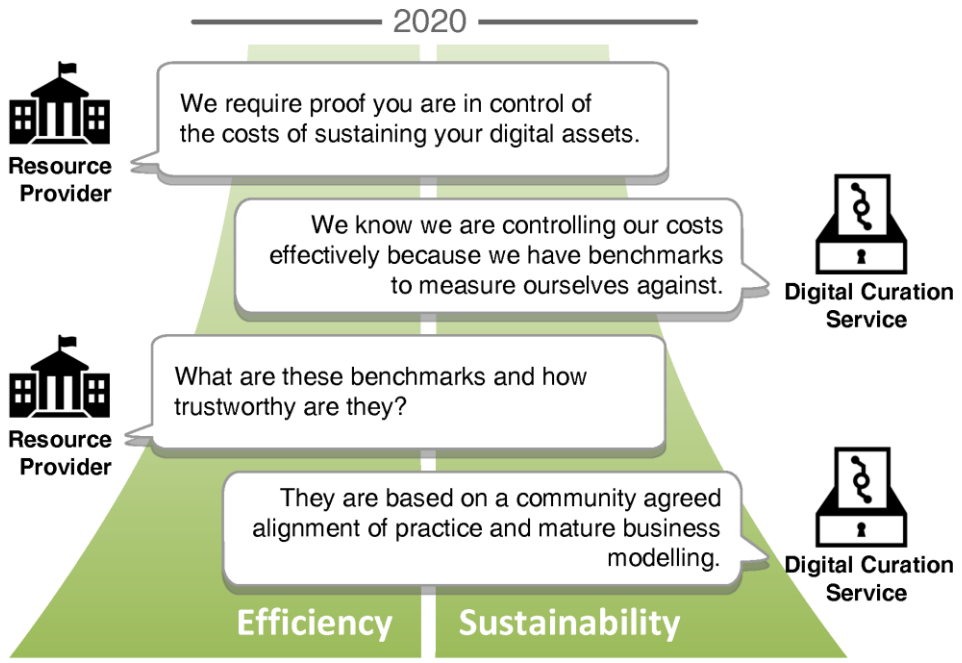
Benefits and positive outcomes

- Being transparent about costs is a short-cut for organisations and content holders to obtain reciprocal information from their peers.
- The analysis of this information should afford opportunities to optimise curation strategies and practices, identify efficiencies, create networks, and enhance communication with peers, designated communities and other stakeholders.
- Better business cases, scenario planning and calculation of different scenarios will be easier to perform for all parties that are involved with and active in digital curation.
- Better informed investments in digital curation will create value and trust.
- A demonstrable increase in organisational transparency could have important positive reputational implications and could be used as an instrument for changing public perceptions.
- Having accurate and comprehensive data on which to base decisions will benefit all types of organisation and should universally provide advantage. Businesses and other types of organisations where there are sensitivities around openly revealing the economic basis of their activities can still contribute with carefully contextualised and anonymised data.

Actions

Who	What	When 2015	2016	2017	2018	2019
Curation Practitioners	Devote resources to clarifying the costs & benefits of curation and then share the findings with the wider community. Ask for reciprocal information from others	●	●	●		
Curation Researchers	Examine, evaluate, assess and report on the impact of being collaborative and transparent about costs and benefits information		●	●		
Data Users	Understand the role and purpose of the 'designated community' for curation and ensure that managers & policy makers include users in consultation and steering groups for digital curation initiatives	●	●			
Managers	Ensure that curation activity within organisations is aligned with organisational objectives and that curation practitioners are correctly identifying & emphasising curation benefits when they are outlining curation costs	●	●			
Member Organisations	Synthesise & disseminate the data on costs & benefits and adopt a neutral & universal approach to help all organisations drive down the costs of curation. Foster a culture of trust among members	●	●	●	●	●
Policy Makers	Foster a culture of collaboration to understand the costs and benefits of digital curation	●	●	●		
Solution Providers	Come up with good descriptions of the benefits frameworks and the curation objectives that systems & solutions support to complement clear pricing & costs information	●	●			

The shape of things to come?



23

How can the 4C Project Help?

The outputs of the 4C Project are designed to help stakeholders manage and control the costs of digital curation and to assess those costs against critically related concepts such as benefits, value, risk and sustainability.

The main outputs of the project are:

The Curation Costs Exchange

A trustworthy and sustainable community resource for depositing and accessing curation costs data and related information. Its purpose is to make the sharing and comparison of data as easy as possible.

http://curationexchange.org

A Cost Concept Model and Gateway Specification

A framework that allows current and future cost models to be compared and benchmarked against a comprehensive set of cost concepts. The model and the associated gateway specification are designed to support future cost modelling activities

http://4cproject.eu/d3-2-ccm

An Evaluation of Costs Models and a Needs & Gap Analysis

An analysis of existing research related to the economics of digital curation and how well current cost and benefit models meet stakeholders' needs for calculating and comparing financial information.

http://4cproject.eu/d3-1

A Summary of Current Cost Models

A summary and description of 10 openly available cost models

http://4cproject.eu/summary-of-cost-models

An Economic Sustainability Reference Model

A strategic tool to facilitate discussion and to support planning of successful sustainability strategies for digital curation.

http://4cproject.eu/ms9-draft-esrm

A Report on the Indirect Economic Determinants of Digital Curation

A description of the indirect factors and concepts that organisations need to be aware of when clarifying the costs of curation.

http://4cproject.eu/d4-1-ied

A Report on Quality & Trustworthiness as an Indirect Economic Determinant

A case study report on the overhead, cost, intellectual input and the eventual benefits that may accrue of undergoing audit and certification procedures to become a 'trusted digital repository' or similar.

http://4cproject.eu/d4-3-quality-and-trustworthiness

A report on Risk as an Indirect Economic Determinant

A report on the role of risk and risk assessment in relation to digital curation and its impact on costs.

(Work in progress).

From Costs to Business Models for Digital Curation

An examination of potential business models, an analysis of the types of services needed, the ways that these can be provided, and options for fee structures.

(Work in progress).

Baseline Study of Stakeholders & Stakeholder Initiatives

A review of relevant work on the economics of digital curation and the results of a stakeholder survey on current practice and stakeholder needs.

http://4cproject.eu/d2-1-stakeholders

4C

24

So what do you think?

This version of the Roadmap (August 2014) is being published as a draft because we want your feedback. The messages are based on extensive research and engagement with the community but they are *opinions*.

We want to know what you think.

What have we got right?

What have we got wrong?

What have we missed out?

We have an online survey so when you've read the draft and thought a bit please go on-line and let rip!
http://4cproject.eu/rmfeedback

Don't hold back.
*Tell us what you **really** think!*

We want the final roadmap to speak to all stakeholders and we want it to be useful, practical and for people to act on the messages it contains. For that we need it to contain targets and goals that people can get behind. We know that not all of the messages will be applicable to all stakeholders, but we do want there to be at least something in the Roadmap for everyone. If you think that your community will see no benefit then tell us (and just as importantly tell us why).

Digital curation is important and this roadmap represents an opportunity to raise awareness of that fact across the board.

The questions

Before you leap away to answer the questions we thought we'd give you an opportunity to see them here.

Firstly, we'd like to know if you share the overall vision we outlined up front. Or do you not share it? Either way we'd like to know why.

Then for each of the 6 messages we'd like you to consider:

- Is the message meaningful to you?
- If the message applies to you, are you prepared to act on it?
- Do you agree with the message?
- Is this message aimed at the right audiences?

As with the vision, we'd also like to know why.

Pass it on

Please do. If you know of someone who you think might be interest then please do send them a copy. Alternatively you can point them at the web-site— http://4cproject.eu/d5-1-draft-roadmap —where they can download their own copy.

From all of us here at the 4C Project, thanks for participating.

Disclaimer and Licence

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.

"Investing in Curation—A Shared Path to Sustainability" by 4cproject.eu is licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

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.

Contact

The 4C project and partners can be contacted via our Web site — http://4cproject.eu . Other means of contact include:

Email info@4cproject.eu

Phone + 44 (0)1904 567654

Twitter @4c_project and using the hashtag #4ceu

Mail 4C Project, c/o DPC, Innovation Centre, York Science Park, Innovation Way, York, YO10 5DG

Acknowledgements

Pictures provided by:

"NYC bridge Central Park Manhattan" by Mary Schiller / PD CC0
"Euro Bills" by Tax Credits / CC BY
"Introducing C.A.N.E." by Lamont Cranston / CC BY-NC-SA
"Arbolote_2112" by Omar Bárcena / CC BY-NC
"Windmill Mill Bornholm Denmark " by Agropyron / PD CC0
"The Price Tag from My Brown Suede Jacket" by abbyladybug / CC BY-NC
"LDDVD Comparison" by Kevin586 / CC BY-SA
"Decaying Media" by Paul Stokes / CC BY-SA

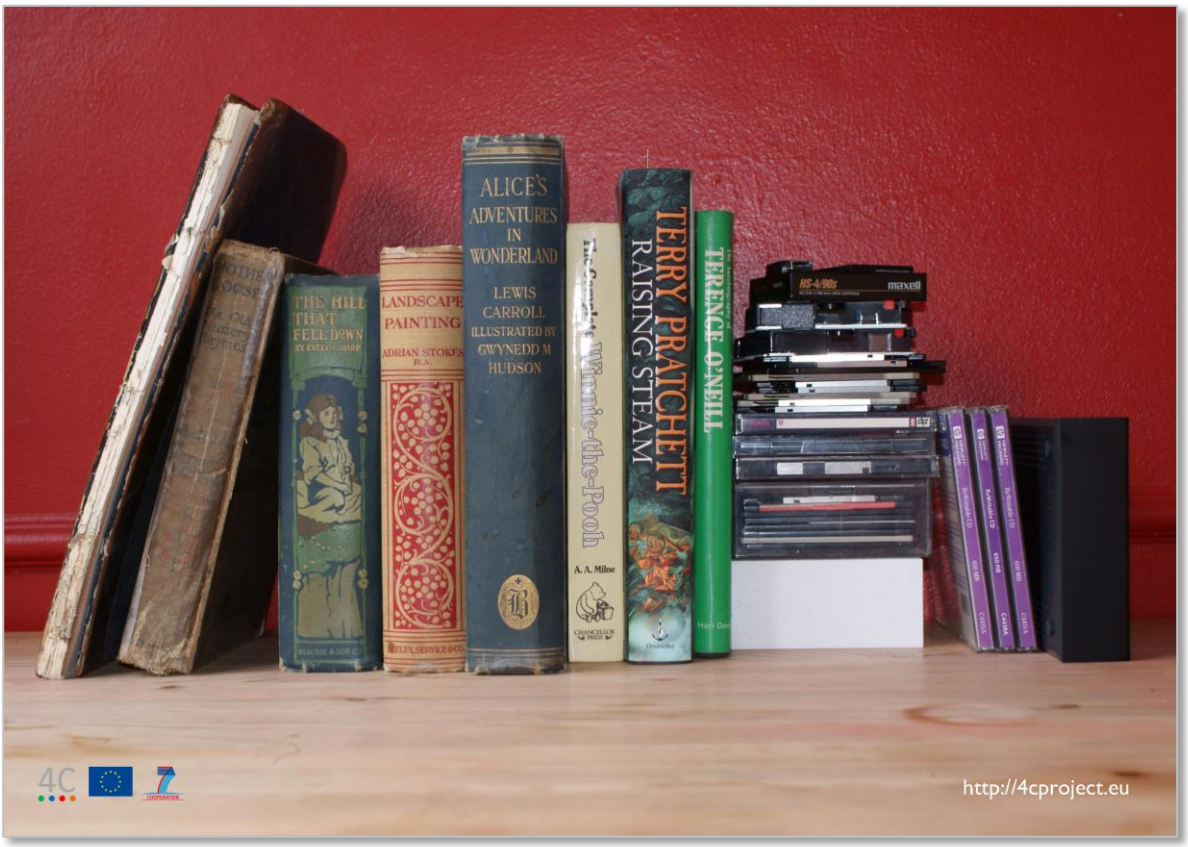
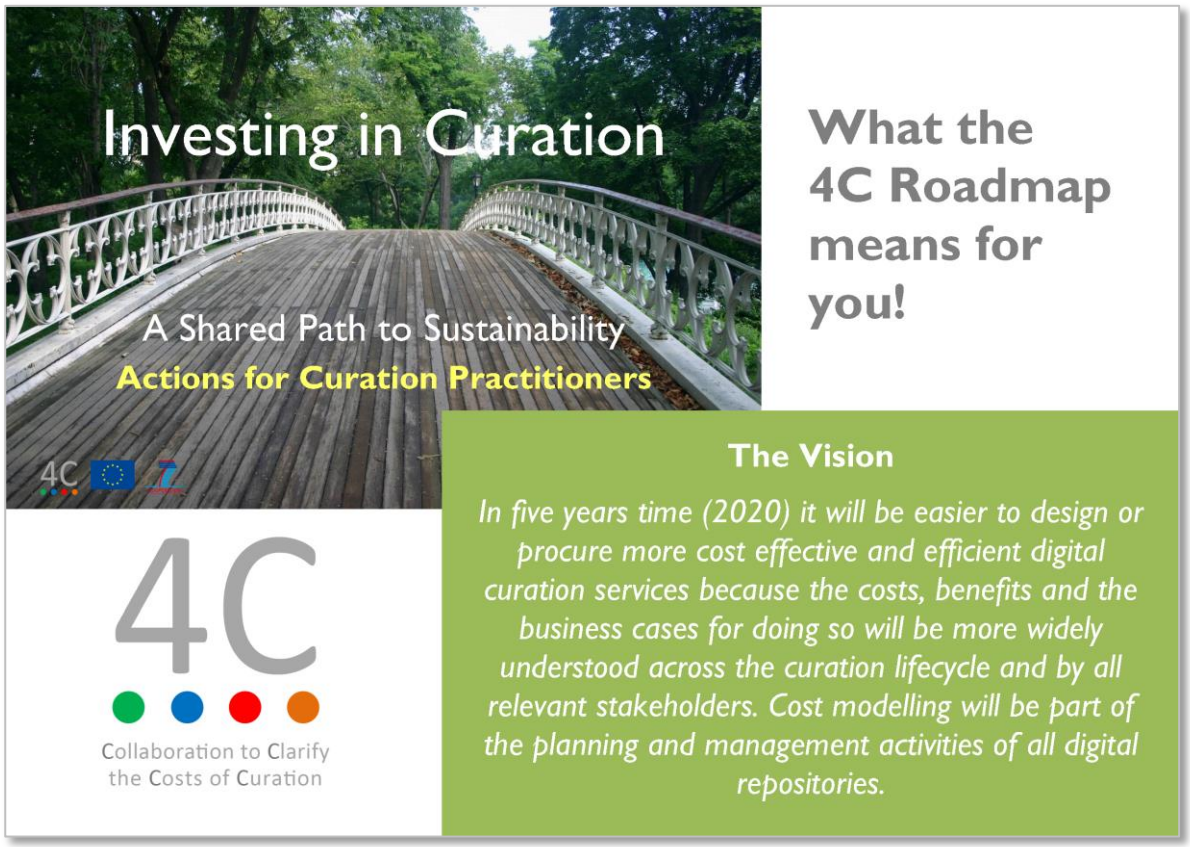


Figure 16—Roadmap booklet

Stakeholder actions postcards—English

http://www.4cproject.eu/roadmap-resources#English



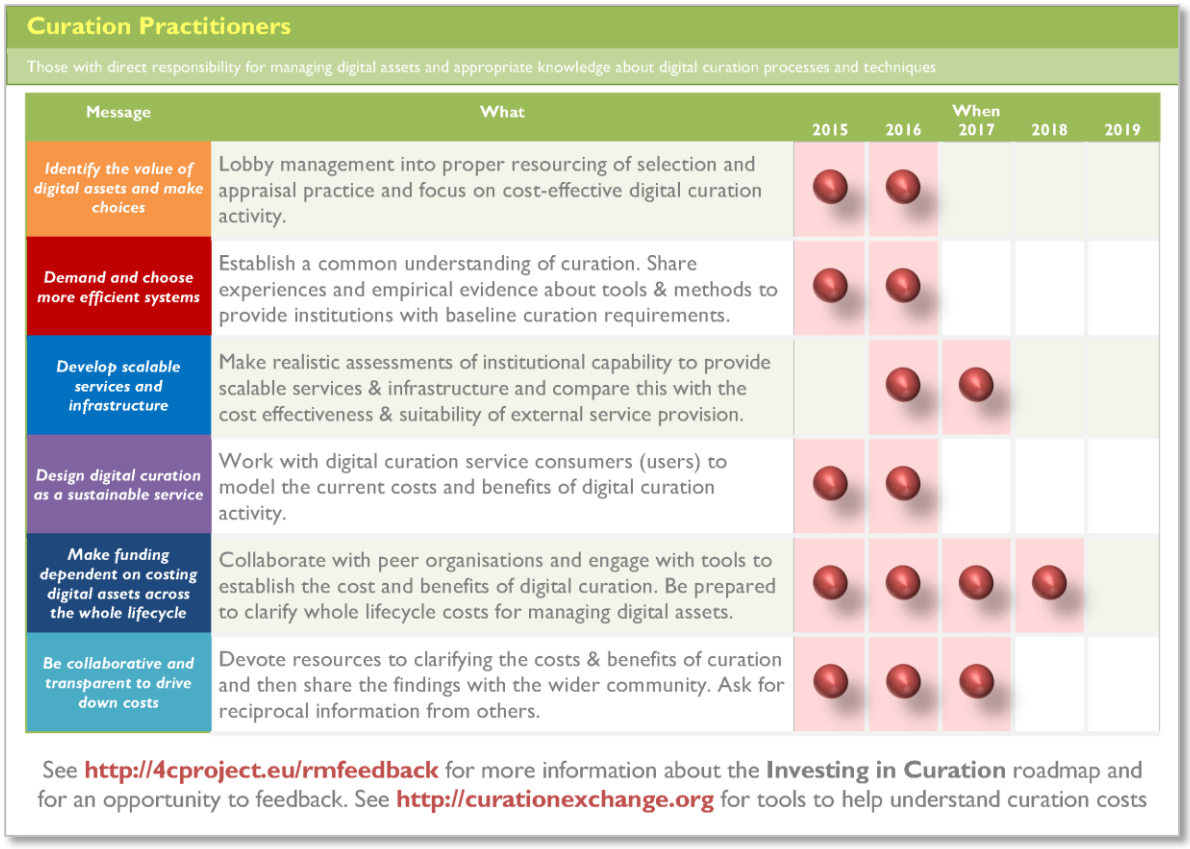


Figure 17—Actions for Curation Practitioners—English

Figure 18—Actions for Curation Researchers—English

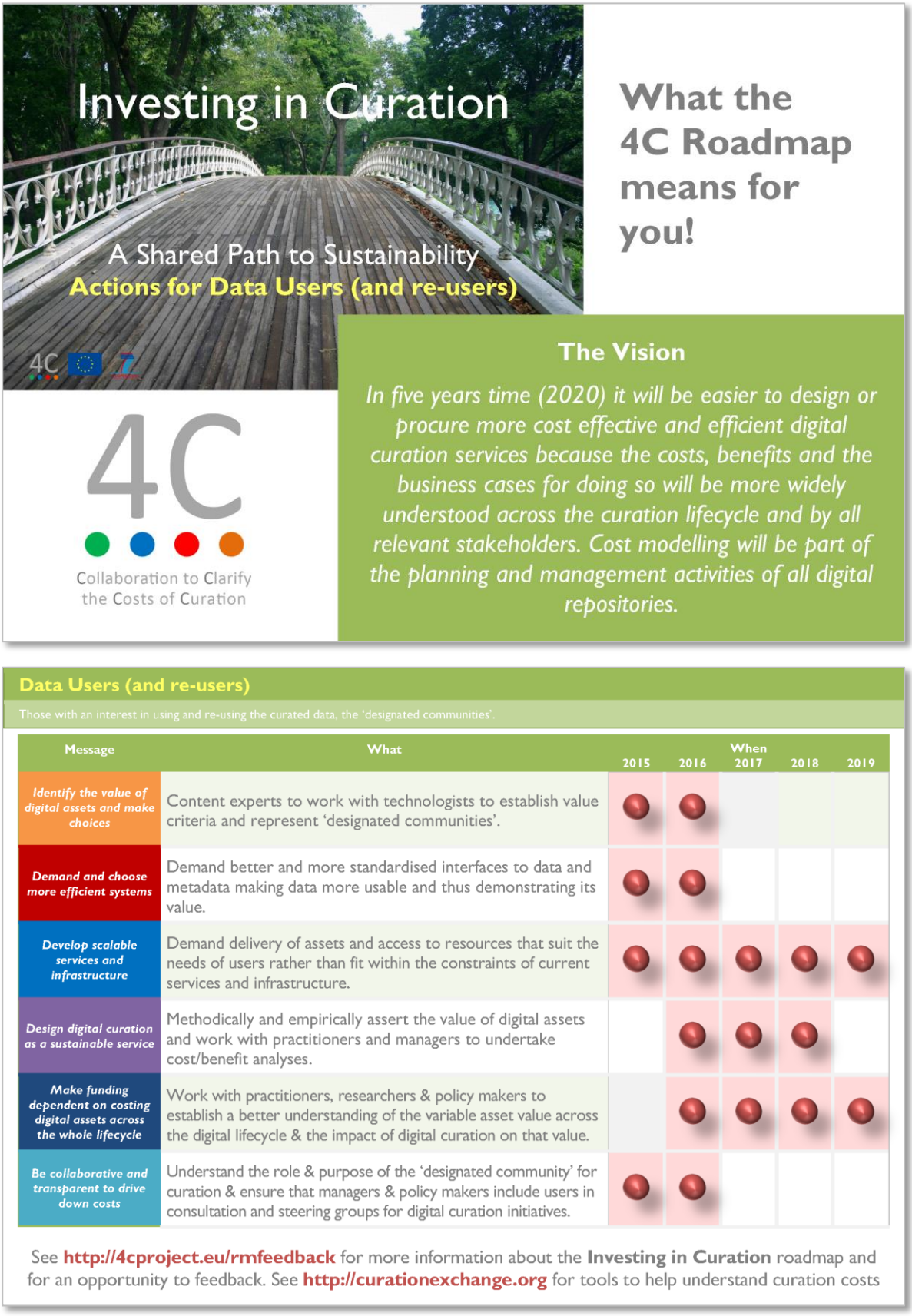


Figure 19—Actions for Data Users—English

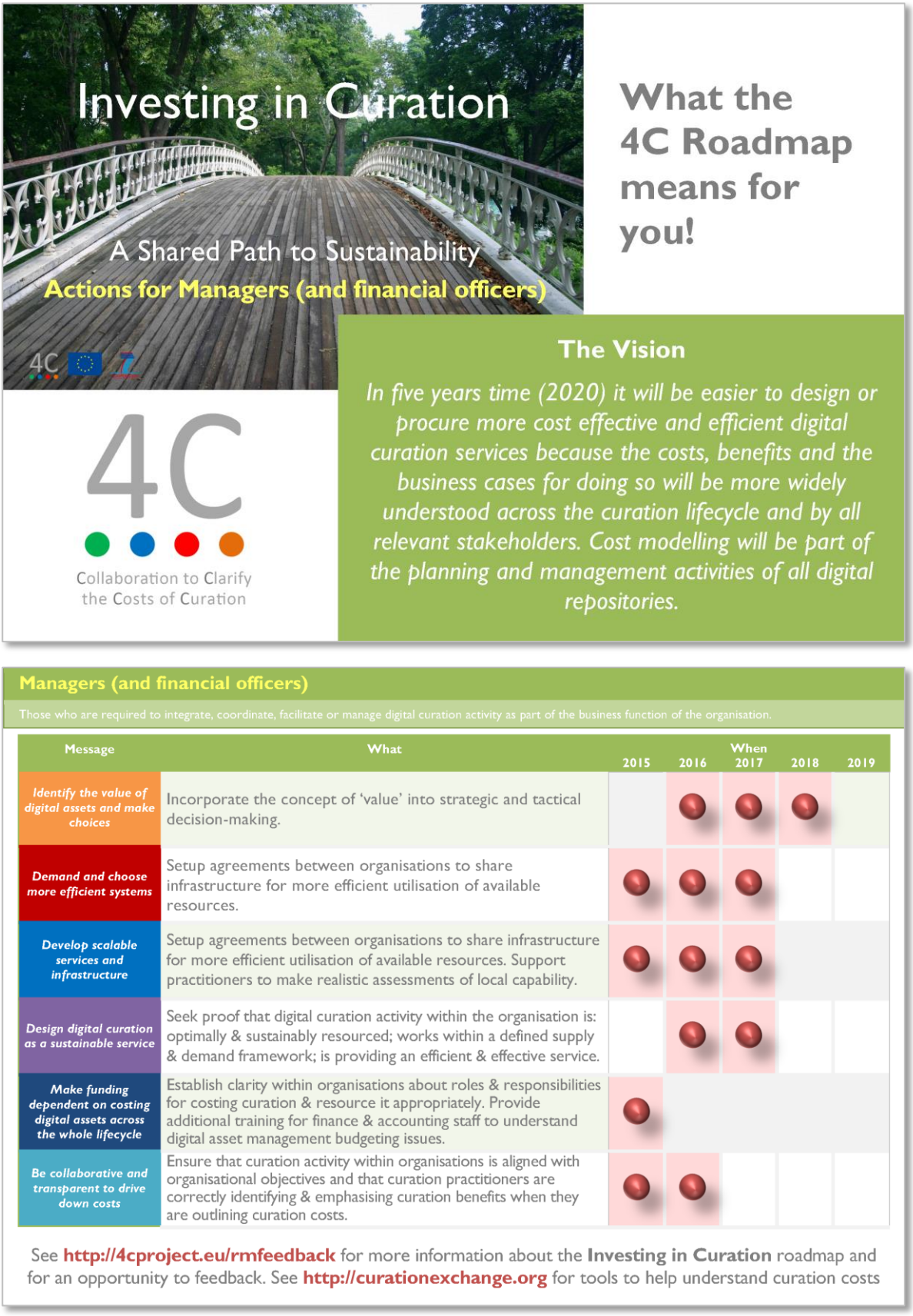


Figure 20—Actions for Managers—English

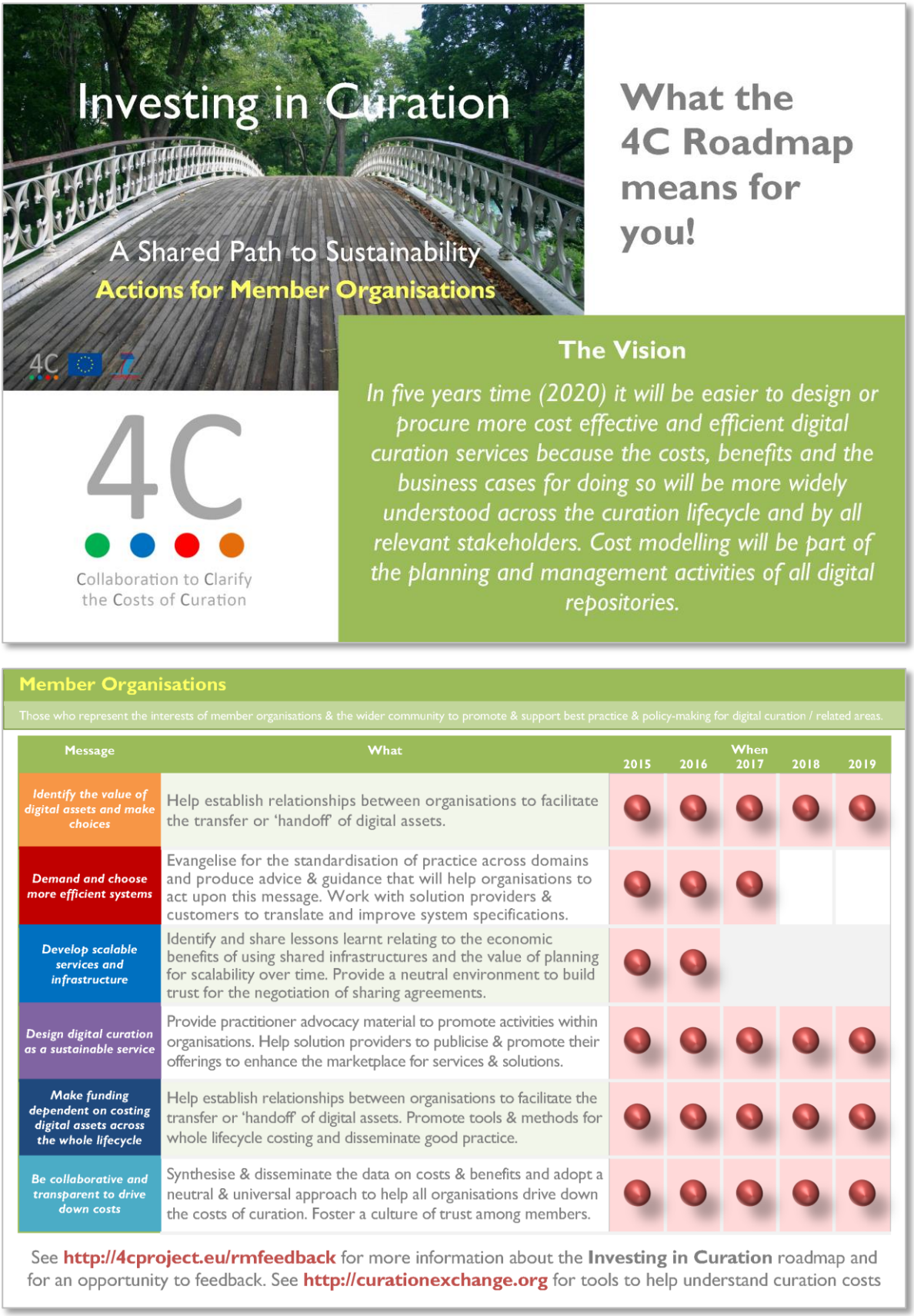


Figure 21—Actions for Member Organisations—English

Figure 22—Actions for Policy Makers—English

Figure 23—Actions for Solution Providers—English

Stakeholder actions postcards—German

http://www.4cproject.eu/roadmap-resources#German

Figure 24—Actions for Curation Practitioners—German

Figure 25—Actions for Curation Researchers—German

Was die 4C Roadmap für Sie konkret bedeutet!

4C

Collaboration to Clarify the Costs of Curation

Die Vision

In fünf Jahren (2020) wird es einfacher sein, einen digitalen Langzeitarchivierungsservice kosteneffektiver und effizienter zu betreiben. Relevante Stakeholder werden ein besseres Verständnis sowohl von Kosten und Nutzen, als auch von zugrundeliegenden Geschäftsmodellen rund um die langfristige Erhaltung digitaler Daten haben.

Datennutzer

Jene mit Interesse in Nutzung und Nachnutzung der langzeitarchivierten Daten, also die vorgesehenen Datennutzer ('designated communities')

What		Wann 2015	2016	2017	2018	2019
Bewerten Sie den Nutzen Ihrer digitalen Objekte und treffen Sie eine gezielte Auswahl	Experten arbeiten mit Technikern zusammen, um Bewertungskriterien zu etablieren und die vorgesehenen Datennutzer ('designated communities') zu repräsentieren.	●	●			
Fordern und nutzen Sie den Einsatz von effizienteren Systemen	Forderung nach besseren und mehr standardisierten Schnittstellen zu Daten und Metadaten, um diese schneller und besser verarbeiten zu können und so deren Nutzen zu zeigen.	●	●			
Entwickeln Sie skalierbare Services und Infrastrukturen	Fordern Sie Zugriff auf jene digitale Assets und Ressourcen, die den Bedürfnissen der Benutzer entsprechen und weniger jenen, die durch die Einschränkungen von derzeitigen Systemen geliefert werden.	●	●	●	●	●
Implementieren Sie digitale Langzeitarchivierung als nachhaltigen Service	Stellen Sie methodisch und empirisch den Wert und Nutzen von digitalen Artefakten fest. Arbeiten Sie mit Anwendern und Manager für Kosten-/Nutzenanalysen zusammen.		●	●	●	
Bewerten Sie die finanzielle Förderung am gesamten Lebenszyklus von digitalen Beständen	Arbeiten Sie mit Anwendern, Forscher und Entscheidungsträgern zusammen, um ein besseres Verständnis von sich ändernden Werten von digitalen Artefakten über deren Lebenszyklus hinaus zu erhalten.		●	●	●	●
Bewerten Sie den Nutzen Ihrer digitalen Objekte und treffen Sie eine gezielte Auswahl	Verstehen Sie die Rollen und die Absichten der vorgesehenen Anwender ('designated community') und stellen Sie sicher, dass Manager und Entscheidungsträger Anwender in ihren Überlegungen miteinbeziehen.	●	●			

Gehen Sie auf **http://4cproject.eu/rmfeedback** für mehr Informationen zur Roadmap **Investition in digitale Bewahrung** und geben Sie Feedback. **http://curationexchange.org** hilft Ihnen Kosten zu verstehen

Figure 26—Actions for Data Users—German

Figure 27—Actions for Managers—German

Mitgliedsorganisationen

Für jene die, die Interessen von Mitgliedsorganisationen vertreten, für die erweiterte Community und für Entscheidungsträger im Bereich Langzeitarchivierung und themenverwandter Gebieten

Was		Wann 2015	2016	2017	2018	2019
Bewerten Sie den Nutzen Ihrer digitalen Objekte und treffen Sie eine gezielte Auswahl	Unterstützen Sie die Zusammenarbeit zwischen Organisationen, um den Transfer oder die Übergabe von digitalen Artefakten zu unterstützen.	●	●	●	●	●
Fordern und nutzen Sie den Einsatz von effizienteren Systemen	Forcieren sie Standardisierung von Praktiken domainübergreifend, stellen Sie Ratschläge und Anleitungen für Organisationen als Hilfestellung zur Verfügung; arbeiten Sie mit Nutzern und Anbietern zusammen, um Anforderungen für Systeme zu verbessern.	●	●	●		
Entwickeln Sie skalierbare Services und Infrastrukturen	Teilen Sie Ihre Erfahrungen in Bezug auf ökonomischen Nutzen bei der Nutzung von gemeinsamer Infrastruktur und den Vorteil von Skalierbarkeit. Stellen Sie neutrale Umgebungen zur Verfügung, um gemeinsame Nutzungsverträge zu ermöglichen.	●	●			
Implementieren Sie digitale Langzeitarchivierung als nachhaltigen Service	Stellen Sie Anwendern Materialien zur Verfügung, um deren Aktivitäten innerhalb der Organisation zu fördern. Unterstützen Sie Anbietern ihre Angebote zu verbreiten und zu bewerben, um den Markt für Langzeitarchivierung zu stärken.	●	●	●	●	●
Bewerten Sie die finanzielle Förderung am gesamten Lebenszyklus von digitalen Beständen	Unterstützen Sie die Zusammenarbeit zwischen Organisationen, um den Transfer oder Übergabe von digitalen Artefakten zu unterstützen. Fördern Sie die Bewerbung von Tools und Methoden für Kostenberechnung des gesamten Lebenszyklus und verbreiten Sie ‚good practice' Beispiele.	●	●	●	●	●
Minimieren Sie Kosten durch Kollaborationen und ermöglichen Sie Transparenz	Erstellen und teilen Sie Daten zu den Kosten und Nutzen von Langzeitarchivierung. Entwickeln Sie neutrale und universelle Ansätze anwendbar für alle Organisationen, um Kosten zu reduzieren. Fördern Sie eine Kultur des Vertrauens zwischen Mitgliedern.	●	●	●	●	●

Gehen Sie auf **http://4cproject.eu/rmfeedback** für mehr Informationen zur Roadmap **Investition in digitale Bewahrung** und geben Sie Feedback. **http://curationexchange.org** hilft Ihnen Kosten zu verstehen

Figure 28—Actions for Member Organisations—German

Entscheidungsträger (Ressourcenanbieter/ Datenverwalter)

Für Personen mit Verantwortung für Anweisungen zu Art und Qualität der gefragten Bestandserhaltungsaktivität/Finanzierung/Rahmenschaffung für Dateneigentum

Inhalt	Was	Wann 2015	2016	2017	2018	2019
Bewerten Sie den Nutzen Ihrer digitalen Objekte und treffen Sie eine gezielte Auswahl	Spezifizieren Sie die Anforderungen an die Wertgutachten von digitalen Beständen hinsichtlich des Datenmanagements und der Bestandserhaltungsplanung.			●	●	●
Fordern und nutzen Sie den Einsatz von effizienteren Systemen	Setzen Sie sich für „Good Practice" und Training ein, so dass integrierte und standardisierte Tools und Dienstleistungen in digitaler Bestandserhaltung einen höheren Stellenwert haben.			●	●	●
Entwickeln Sie skalierbare Services und Infrastrukturen	Stellen Sie domänenübergreifend genutzte Infrastrukturen bereit, um Skaleneffekte auszuschöpfen.	●	●	●		
Implementieren Sie digitale Langzeitarchivierung als nachhaltigen Service	Stellen Sie domänenübergreifend genutzte Infrastrukturen bereit, um Skaleneffekte auszuschöpfen. Konzipieren Sie Kapitalbegrenzungen, um eine nachhaltige digitale Bestandserhaltung zu gewährleisten, die sich durch eine erprobte Wirtschaftlichkeit auszeichnet.	●	●	●		
Bewerten Sie die finanzielle Förderung am gesamten Lebenszyklus von digitalen Beständen	Ermitteln Sie, wo die Pflege digitaler Bestandserhaltung Priorität hat und entwickeln Sie Klauseln in Pflegeverträgen, die eine Kostenabschätzung voraussetzen, die den gesamten Lebenszyklus für den Erhalt digitaler Bestände berücksichtigt(, solange wie diese benötigt werden).			●	●	●
Minimieren Sie Kosten durch Kollaborationen und ermöglichen Sie Transparenz	Fördern Sie die Kultur der Zusammenarbeit, um die Kosten und Nutzen der digitalen Bestandserhaltung gemeinsam zu verstehen.	●	●	●		

Gehen Sie auf http://4cproject.eu/rmfeedback für mehr Informationen zur Roadmap **Investition in digitale Bestandserhaltung** und geben Sie Feedback. http://curationexchange.org hilft Ihnen Kosten zu verstehen

Figure 29—Actions for Policy Makers—German

Was die 4C Roadmap für Sie konkret bedeutet!

4C

Collaboration to Clarify the Costs of Curation

Die Vision

In fünf Jahren (2020) wird es einfacher sein, einen digitalen Langzeitarchivierungsservice kosteneffektiver und effizienter zu betreiben. Relevante Stakeholder werden ein besseres Verständnis sowohl von Kosten und Nutzen, als auch von zugrundeliegenden Geschäftsmodellen rund um die langfristige Erhaltung digitaler Daten haben.

Aufgaben für Anbieter

Für Personen mit (kommerziellem oder gemeinschaftlichem) Anreiz, Produkte zu entwickeln und zu verbreiten, die digitale Langzeitarchivierung unterstützen.

Inhalt	Was	Wann 2015	2016	2017	2018	2019
Bewerten Sie den Nutzen Ihrer digitalen Objekte und treffen Sie eine gezielte Auswahl	Nutzen Sie vorhandene Tools (z.B. Dateiformat Register), um eine automatisierte Auswahl und Beurteilung von Tools anbieten zu können.		●	●	●	
Fordern und nutzen Sie den Einsatz von effizienteren Systemen	Arbeiten Sie mit Nutzern zusammen, um Standardpraktiken zu entwickeln und zu vereinfachen. Treffen Sie Nutzer auf halbem Weg, indem Sie Lösungen spezifizieren sowie Preismodelle und Umsetzungsmöglichkeiten klar und verständlich machen.	●	●	●		
Entwickeln Sie skalierbare Services und Infrastrukturen	Achten Sie besonders auf die Notwendigkeit von Skalierbarkeit der Leistungen. Bieten Sie als Reaktion auf wachsende Ansprüche getestete Lösungen und transparente sowie geprüfte Leistungen an.	●	●	●	●	
Implementieren Sie digitale Langzeitarchivierung als nachhaltigen Service	Beteiligen Sie sich an der Bildung von Standards. Setzen Sie den Fokus auf langfristige Interoperabilität in Software und Infrastruktur. Konzentrieren Sie sich auf Offenheit, Zusammenarbeit und Aufbau eines nachhaltigen und integrativen Marktplatzes.	●	●	●	●	
Bewerten Sie die finanzielle Förderung am gesamten Lebenszyklus von digitalen Beständen	Arbeiten Sie mit Anwendern und Forschern zusammen, um Kosten und Budgetmodelle als feste Module in Langzeitarchivsysteme zu verankern.		●	●		
Minimieren Sie Kosten durch Kollaborationen und ermöglichen Sie Transparenz	Liefern Sie gute Beschreibungen des Nutzen und der Vorteile von digitalen Langzeitarchivsystemen, die klare Preisgestaltung und Kosteninformation unterstützen.	●	●			

Gehen Sie auf http://4cproject.eu/rmfeedback für mehr Informationen zur Roadmap **Investition in digitale Bestandserhaltung** und geben Sie Feedback. http://curationexchange.org hilft Ihnen Kosten zu verstehen

Figure 30—Actions for Solution Providers—German

Stakeholder actions postcards—French

http://www.4cproject.eu/roadmap-resources#French

Figure 31—Actions for Curation Practitioners—French

Chercheurs dans la préservation numérique

Pour ceux qui ont la capacité et la connaissance permettant de saisir les défis de la préservation numérique et de définir de nouvelles méthodes et procédés pour la gestion à long terme des biens numériques.

Message	Quoi	Quand 2015	2016	2017	2018	2019
Identifier la valeur des biens numériques et faire les choix nécessaires	Faire de la recherche concernant les techniques d'évaluation et de sélection automatisées employant des critères basés sur la valeur.	●	●	●		
Demander et choisir des systèmes plus efficaces	Entreprendre un travail de recherche pour minimiser la subjectivité ainsi que clarifier et standardiser les définitions des bénéfices. Développer des outils facilitant la mise en place de standards.	●	●			
Développer des services et une infrastructure évolutifs	Optimiser les méthodes de travail et concevoir des procédures permettant de gérer de grandes quantités de données numériques ainsi que des données numériques complexes.	●	●	●		
Concevoir la préservation numérique comme un service durable	Continuer la recherche pour obtenir un modèle d'affaires durable et examiner comment standardiser les pratiques courantes divergentes.	●	●			
Etablir une dépendance entre coûts et financements de la préservation numérique tout au long du cycle de vie	Poursuivre le développement de produits permettant de simplifier l'estimation des coûts ainsi que l'étude comparative de la préservation numérique. S'engager dans des recherches supplémentaires pour redéfinir les méthodes et réduire les coûts.	●	●			
Etre collaboratif et transparent afin de réduire les coûts	Examiner et évaluer l'impact sur les coûts et les bénéfices qu'apportent la collaboration et la transparence de données financières		●	●		

Vous pouvez vous reporter au site **http://4cproject.eu/rmfeedback** pour plus d'information sur la feuille de route **Investir dans la préservation numérique** et nous faire part de vos commentaires. Vous pouvez consulter **http://curationexchange.org** pour accéder aux outils permettant de comprendre les coûts de préservation.

Figure 32—Actions for Curation Researchers—French

Ce que la feuille de route 4C signifie pour vous!

4C

Collaboration to Clarify the Costs of Curation

La Vision

Dans cinq ans (2020) il sera plus facile de concevoir ou d'offrir des services de préservation numérique plus économiques et plus efficaces, car les coûts, les bénéfices et les analyses de rentabilisation pour le permettre seront plus généralement compréhensibles pour les acteurs concernés. L'estimation des coûts deviendra alors partie intégrante de la planification et des activités de gestion de tous les systèmes de préservation numérique.

Utilisateurs de biens numériques

Destiné à ceux qui ont un intérêt à utiliser et à réutiliser les données preservées. La "Communauté d'utilisateurs cible".

Message	Quoi	Quand 2015	2016	2017	2018	2019
Identifier la valeur des biens numériques et faire les choix nécessaires	Aux experts métiers de travailler avec les techniciens afin d'établir des critères de valeur et de représenter la "Communauté d'utilisateurs cible".	●	●			
Demander et choisir des systèmes plus efficaces	Demander des interfaces de meilleure qualité et standardisées pour les données et les metadonnées numériques, les rendant plus exploitables, démontrant ainsi leur valeur.	●	●			
Développer des services et une infrastructure évolutifs	Demander la consultation des données et l'accès aux outils nécessaires pour couvrir les besoins des utilisateurs afin de passer outre les contraintes des services et infrastructures existants.	●	●	●	●	●
Concevoir la préservation numérique comme un service durable	Faire valoir de façon méthodique et empirique la valeur des biens numériques et travailler avec les praticiens et les gestionnaires pour entamer une analyse coûts-bénéfices.		●	●	●	
Etablir une dépendance entre coûts et financements de la préservation numérique tout au long du cycle de vie	Travailler avec les praticiens, les chercheurs et les responsables pour établir une meilleure connaissance de la valeur des données numériques et une meilleure connaissance de la valeur ajoutée qu'apportent les activités de la préservation numérique.		●	●	●	●
Etre collaboratif et transparent afin de réduire les coûts	Comprendre le rôle et le but de la "Communauté d'utilisateurs cible" et assurer que les gestionnaires et les responsables impliquent les utilisateurs dans les initiatives de préservation numérique par le biais de la consultation ou de la participation.	●	●			

Vous pouvez vous reporter au site **http://4cproject.eu/rmfeedback** pour plus d'information sur la feuille de route **Investir dans la préservation numérique** et nous faire part de vos commentaires. Vous pouvez consulter **http://curationexchange.org** pour accéder aux outils permettant de comprendre les coûts de préservation.

Figure 33—Actions for Data Users—French

Les gestionnaires et responsables financiers

Destiné à ceux qui doivent intégrer, coordonner, faciliter ou gérer les activités de préservation numérique considerées comme faisant partie des fonctions commerciales de l'organisation

Message	Quoi	Quand 2015	2016	2017	2018	2019
Identifier la valeur des biens numériques et faire les choix nécessaires	Incorporer le concept de "valeur" dans la prise de décision stratégique et tactique.		●	●	●	
Demander et choisir des systèmes plus efficaces	Mettre en place des accords entre les organisations pour partager les infrastructures pour une meilleure utilisation des ressources disponibles.	●	●	●		
Développer des services et une infrastructure évolutifs	Mettre en place des accords entre les organisations pour partager les infrastructures pour une meilleure utilisation des ressources disponibles. Apporter un support aux praticiens pour faire des évaluations réalistes des capacités locales.	●	●	●		
Concevoir la préservation numérique comme un service durable	S'assurer que les activités de préservation numérique dans l'organisation sont dotées de ressources optimales et durables, encadrées dans une logique d'offre et de demande et offrent un service efficace.		●	●		
Etablir une dépendance entre coûts et financements de la préservation numérique tout au long du cycle de vie	Etablir de claires définitions des rôles et responsabilités au sein de l'organisation pour établir le coût de la préservation et assurer les ressources adéquates. Offrir une formation supplémentaire aux équipes financières et comptables afin de les sensibiliser aux problèmes liés à la gestion des biens numériques.	●				
Etre collaboratif et transparent afin de réduire les coûts	Assurer que les activités de préservation numérique dans l'organisation sont en accord avec les objectifs de cette organisation, que les praticiens soient correctement identifiés et qu'ils mettent suffisamment en avant les bénéfices lorsqu'ils présentent les coûts de la préservation numérique.	●	●			

Vous pouvez vous reporter au site **http://4cproject.eu/rmfeedback** pour plus d'information sur la feuille de route **Investir dans la préservation numérique** et nous faire part de vos commentaires. Vous pouvez consulter **http://curationexchange.org** pour accéder aux outils permettant de comprendre les coûts de préservation.

Figure 34—Actions for Managers—French

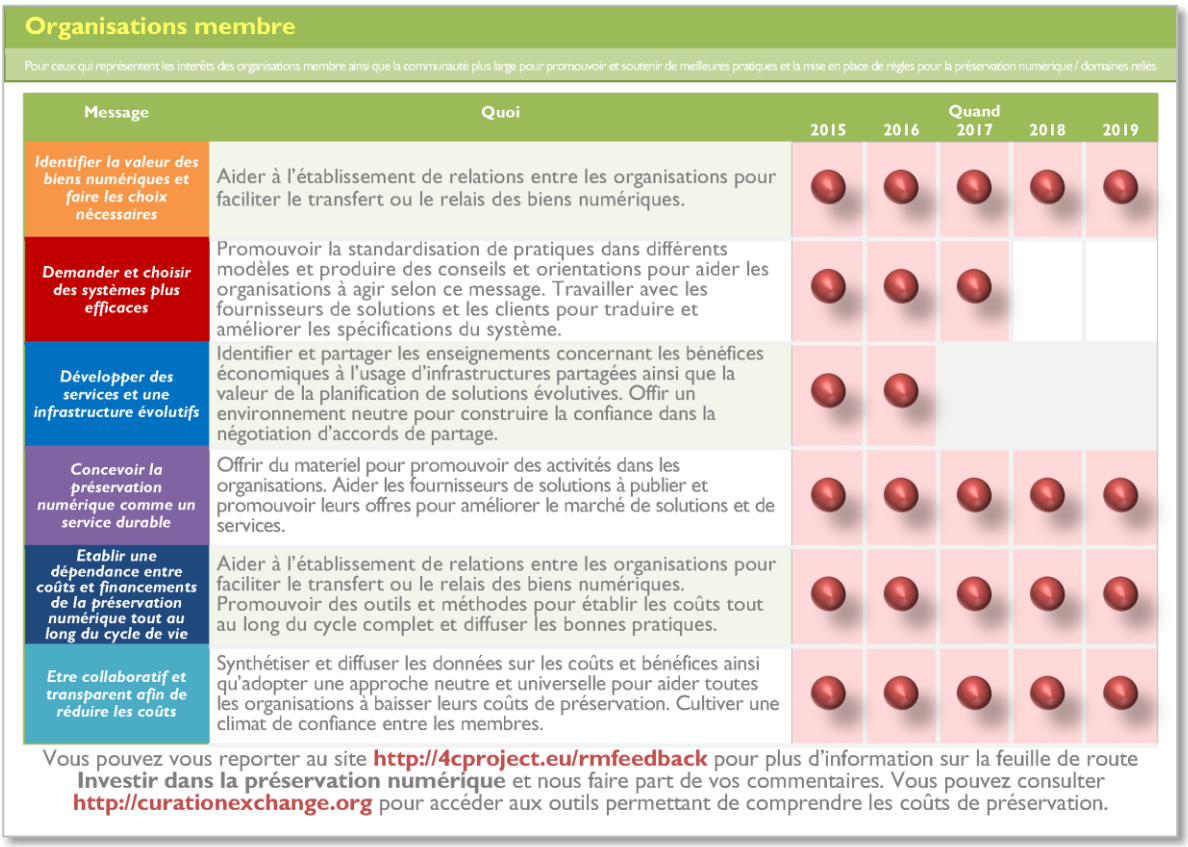


Figure 35—Actions for Member Organisations—French

Ce que la feuille de route 4C signifie pour vous!

4C

Collaboration to Clarify the Costs of Curation

La Vision

Dans cinq ans (2020) il sera plus facile de concevoir ou d'offrir des services de préservation numérique plus économiques et plus efficaces, car les coûts, les bénéfices et les analyses de rentabilisation pour le permettre seront plus généralement compréhensibles pour les acteurs concernés. L'estimation des coûts deviendra alors partie intégrante de la planification et des activités de gestion de tous les systèmes de préservation numérique.

Décideurs (Les bailleurs de fonds / Proprietaires de données)

Destiné aux responsables de la qualité des activités relatives à la préservation numérique, de son financement, et de la mise en place d'un réseau de propriétaires de données.

Message	Quoi	Quand 2015	2016	2017	2018	2019
Identifier la valeur des biens numériques et faire les choix nécessaires	Établir les exigences permettant d'évaluer la valeur du bien numérique comme étant un élément de la gestion des données et de la planification de la préservation numérique.			●	●	●
Demander et choisir des systèmes plus efficaces	Promouvoir les bonnes pratiques et des formations afin que les outils standard de préservation numérique acquièrent un meilleur profil.			●	●	●
Développer des services et une infrastructure évolutifs	Offrir des infrastructures partagées et transparentes pour exploiter les économies d'échelle	●	●	●		
Concevoir la préservation numérique comme un service durable	Offrir des infrastructures partagées et transparentes pour exploiter les économies d'échelle. Créer des contraintes budgétaires afin d'assurer qu'une préservation numérique durable soit étayée par une rentabilité avérée.	●	●	●		
Etablir une dépendance entre coûts et financements de la préservation numérique tout au long du cycle de vie	Identifier le moment où l'entretien de biens numériques est une priorité et concevoir des clauses dans les contrats de support permettant d'estimer les coûts du cycle de vie de la préservation des biens pour la durée requise.			●	●	●
Etre collaboratif et transparent afin de réduire les coûts	Entretenir une culture de collaboration pour permettre de sensibiliser aux coûts et bénéfices de la préservation numérique.	●	●	●		

Vous pouvez vous reporter au site **http://4cproject.eu/rmfeedback** pour plus d'information sur la feuille de route **Investir dans la préservation numérique** et nous faire part de vos commentaires. Vous pouvez consulter **http://curationexchange.org** pour accéder aux outils permettant de comprendre les coûts de préservation.

Figure 36—Actions for Policy Makers—French

Fournisseurs de solutions

Destiné à ceux dont les incitations sont à développer et diffuser des produits pour maintenir l'activité de préservation numérique au niveau des services et des solutions.

Message	Quoi	Quand 2015	2016	2017	2018	2019
Identifier la valeur des biens numériques et faire les choix nécessaires	Construire sur des outils existants (par exemple des registres de formats) pour offrir des outils de sélection et d'évaluation automatisés.		●	●	●	
Demander et choisir des systèmes plus efficaces	Travailler avec les clients et la communauté pour développer, expliquer et simplifier les pratiques courantes. Rencontrer les clients à mi-chemin en specifiant les solutions tout en faisant des modèles tarifaires et mettant en œuvre des options claires et compréhensibles.	●	●	●		
Développer des services et une infrastructure évolutifs	Accorder une attention particulière au besoin de construire des services évolutifs. Offrir des solutions qui sont rigoureusement testées ainsi qu'une performance transparente et de référence en réponse à des spécifications plus sophistiquées.	●	●	●	●	
Concevoir la préservation numérique comme un service durable	Participer à la mise en oeuvre de standards et mettre l'accent sur une interopérabilité durable dans la construction de logiciels et d'infrastructures. Mettre l'accent sur la transparence et la collaboration et construire un marché durable et inclusif.	●	●	●	●	
Etablir une dépendance entre coûts et financements de la préservation numérique tout au long du cycle de vie	Travailler avec des praticiens et des chercheurs pour incorporer des modules de comptabilité et de budget dans les systèmes de préservation numérique.		●	●		
Etre collaboratif et transparent afin de réduire les coûts	Présenter de bonnes descriptions des cadres des bénéfices ainsi que des objectifs de la préservation que les systèmes et les solutions épaulent afin d'obtenir un modèle tarifaire clair et des informations claires sur les coûts.	●	●			

Vous pouvez vous reporter au site **http://4cproject.eu/rmfeedback** pour plus d'information sur la feuille de route **Investir dans la préservation numérique** et nous faire part de vos commentaires. Vous pouvez consulter **http://curationexchange.org** pour accéder aux outils permettant de comprendre les coûts de préservation.

Figure 37—Actions for Solution Providers—English

Stakeholder actions postcards—Portuguese

http://www.4cproject.eu/roadmap-resources#Portuguese

Figure 38—Actions for Curation Practitioners—Portuguese

Figure 39—Actions for Curation Researchers—Portuguese

Figure 40—Actions for Data Users—Portuguese

Figure 41—Actions for Managers—Portuguese

Figure 42—Actions for Member Organisations—Portuguese

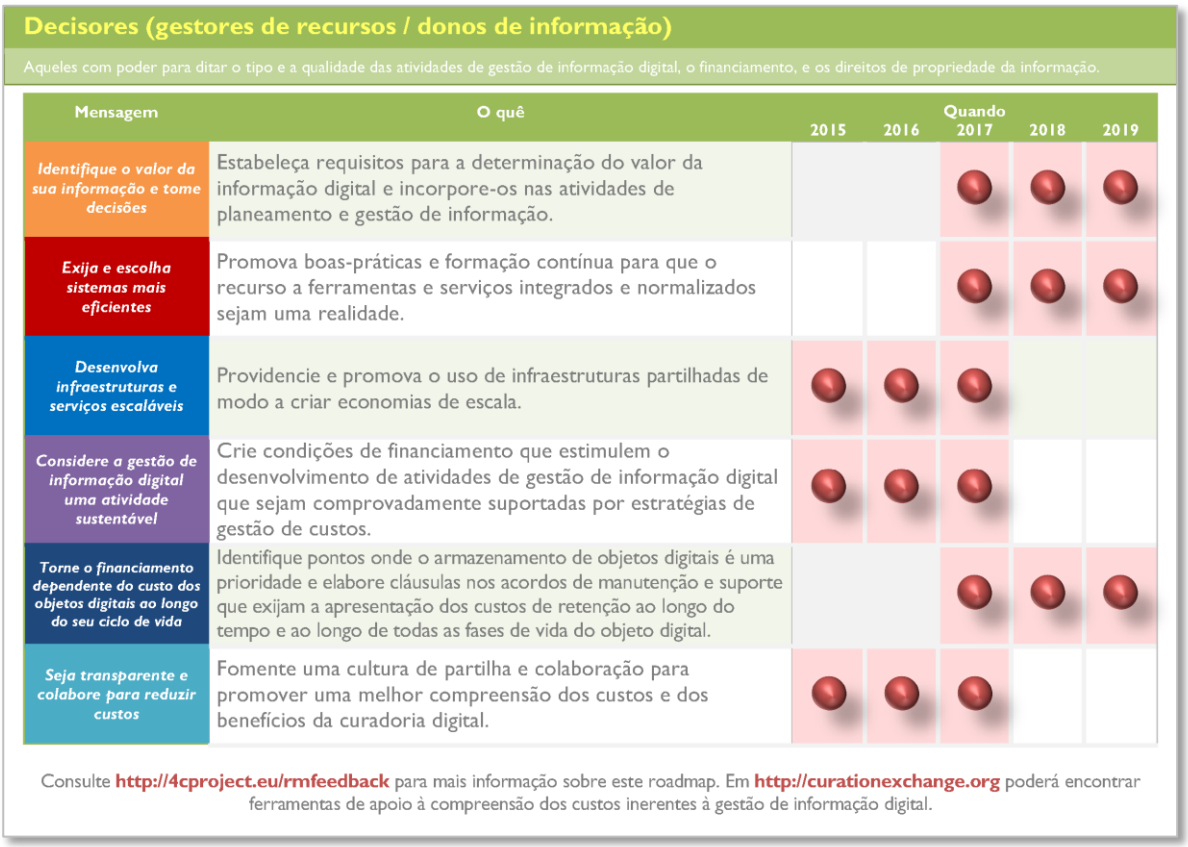


Figure 43—Actions for Policy Makers—Portuguese

Figure 44—Actions for Solution Providers—Portuguese

Stakeholder actions postcards—Dutch

http://www.4cproject.eu/roadmap-resources#Dutch

Figure 45—Actions for Curation Practitioners—Dutch

Figure 46—Actions for Curation Researchers—Dutch

Figure 47—Actions for Data Users—Dutch

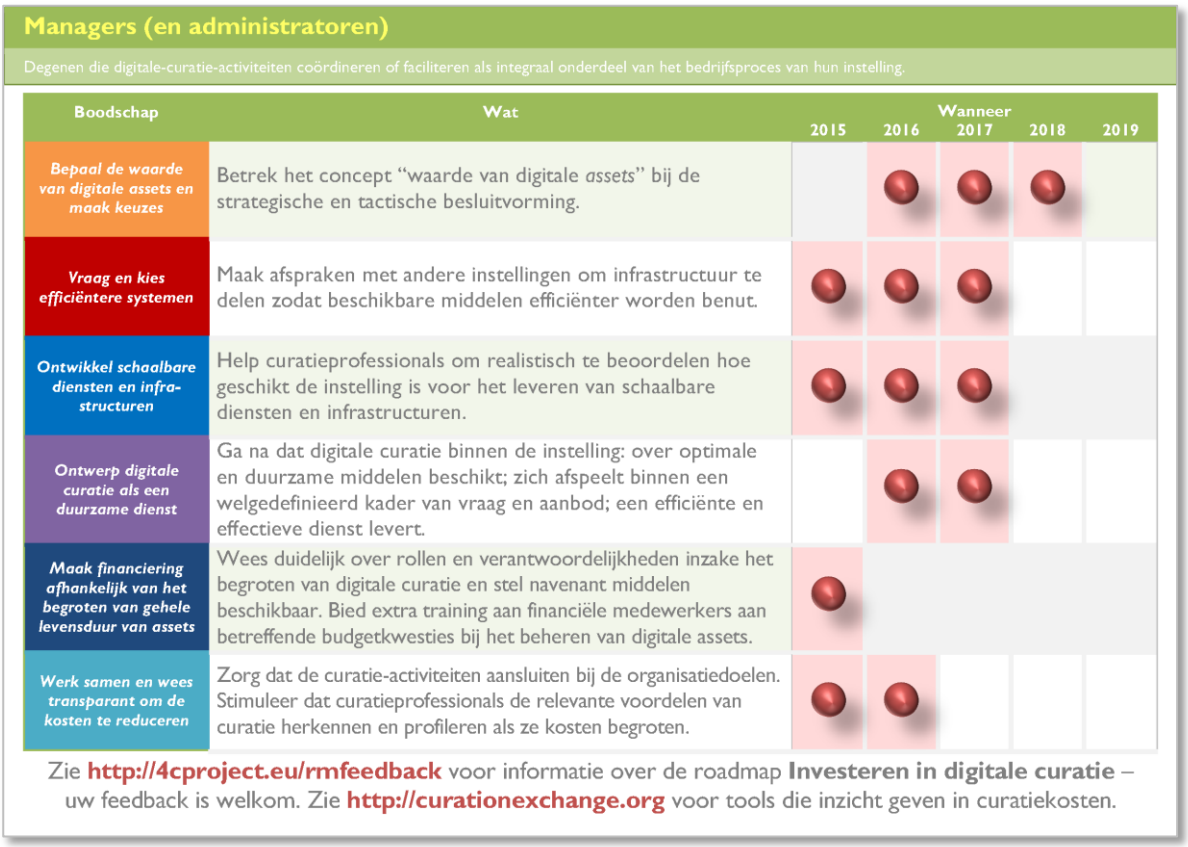


Figure 48—Actions for Managers—Dutch

Wat de 4C-roadmap voor u betekent!

4C

Collaboration to Clarify the Costs of Curation

De visie

Over vijf jaar (in 2020) is het eenvoudiger om kosteneffectieve en efficiënte diensten voor digitale curatie te ontwerpen of af te nemen, omdat de kosten, de baten en de business cases hiervoor breder begrepen zullen worden: door alle betrokkenen en in alle stadia van het curatieproces. Het modelleren van kosten zal deel uitmaken van de planning- en beheeractiviteiten van alle digitale repositories.

Leden- en koepelorganisaties

Degenen die de belangen van lidorganisaties en de bredere community vertegenwoordigen bij het bevorderen van *best practices* en beleid voor digitale curatie.

Boodschap	Wat	Wanneer 2015	2016	2017	2018	2019
Bepaal de waarde van digitale assets en maak keuzes	Bevorder dat instellingen relaties aangaan waarmee de overdracht van digitale *assets* wordt bevorderd.	●	●	●	●	●
Vraag en kies efficiëntere systemen	Stimuleer dat de digitale-curatiepraktijk convergeert naar een disciplineoverstijgende standaard; help organisaties om hiernaar te handelen. Werk samen met dienstverleners en gebruikers aan betere systeemspecificaties.	●	●	●		
Ontwikkel schaalbare diensten en infra-structuren	Benoem en deel inzichten in de financiële voordelen van gedeelde infrastructuur en in het belang om bij planning rekening te houden met schaalvergroting. Wees neutraal en betrouwbaar terrein voor onderhandelingen over samenwerking.	●	●			
Ontwerp digitale curatie als een duurzame dienst	Stimuleer activiteiten binnen organisaties met promotiemateriaal voor curatieprofessionals. Help dienstverleners om hun aanbod bekend te maken, zodat de dienstenmarkt beter wordt.	●	●	●	●	●
Maak financiering afhankelijk van het begroten van gehele levensduur van assets	Bevorder het gebruik van instrumenten en methoden voor het budgetteren van de hele levenscyclus van digitale assets. Geef bekendheid aan *good practices*.	●	●	●	●	●
Werk samen en wees transparant om de kosten te reduceren	Verzamel en verspreid gegevens over kosten en baten van digitale curatie. Kies een neutrale, generieke aanpak om alle organisaties te helpen hun curatiekosten te beperken. Bevorder een cultuur van vertrouwen bij de leden.	●	●	●	●	●

Zie **http://4cproject.eu/rmfeedback** voor informatie over de roadmap **Investeren in digitale curatie** – uw feedback is welkom. Zie **http://curationexchange.org** voor tools die inzicht geven in curatiekosten.

Figure 49—Actions for Member Organisations—Dutch

Figure 50—Actions for Policy Makers—Dutch

Figure 51—Actions for Solution Providers—Dutch